



Universiteit
Leiden
The Netherlands

Hierarchical systems

Hamers, A.S.

Citation

Hamers, A. S. (2016, June 21). *Hierarchical systems*. Retrieved from <https://hdl.handle.net/1887/41202>

Version: Not Applicable (or Unknown)

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/41202>

Note: To cite this publication please use the final published version (if applicable).

Cover Page



Universiteit Leiden



The handle <http://hdl.handle.net/1887/41202> holds various files of this Leiden University dissertation

Author: Hamers, Adrian Sven

Title: Hierarchical systems

Issue Date: 2016-06-21

Hierarchical Systems

Proefschrift

ter verkrijging van
de graad van Doctor aan de Universiteit Leiden,
op gezag van Rector Magnificus prof. mr. C.J.J.M. Stolker,
volgens besluit van het College voor Promoties
te verdedigen op dinsdag 21 juni 2016
klokke 13:45 uur

door

Adrian Sven Hamers
geboren te Utrecht
in 1988

Promotiecommissie

Promotor:	Prof. dr. Simon Portegies Zwart	(Universiteit Leiden)
Co-Promotor:	Prof. dr. Hagai Perets	(Israel Institute of Technology)
Overige leden:	Prof. dr. Pau Amaro-Seoane	(Max Planck Institute for Gravitational Physics)
	Prof. dr. Rosemary Mardling	(Monash University)
	Prof. dr. Alice Quillen	(University of Rochester)
	Prof. dr. Huub Röttgering	(Universiteit Leiden)
	Prof. dr. Scott Tremaine	(Institute for Advanced Study)

ISBN: 978-94-6233-304-8
© 2016 Adrian S. Hamers

Dit proefschrift werd financieel ondersteund door NWO, NOVA en het Leids Kerkhoven Bosscha fonds.

Continuing in the family line of physicists

The endeavor to understand is the first and only basis of virtue.
—Baruch Spinoza

Contents

1	Introduction	1
1.1	Gravity	1
1.2	History of the N -body problem	2
1.3	Hierarchies in nature	3
1.4	Secular evolution	6
1.4.1	Principles	6
1.4.2	Implications	8
1.5	This thesis	9
1.5.1	Chapter 2 – Relativistic dynamics around an SBH	9
1.5.2	Chapter 3 – Planetesimals in the GC	12
1.5.3	Chapter 4 – Secular evolution of hierarchical quadruple systems	12
1.5.4	Chapter 5 – Explaining the lack of circumbinary planets around short-period binaries	13
1.5.5	Chapter 6 – Secular evolution of hierarchical multiple systems	13
1.5.6	Chapter 7 – Hot Jupiters in multiplanet systems	13
2	Relativistic dynamics of stars near a supermassive black hole	15
2.1	Introduction	16
2.2	Time-scales	17
2.3	Method	20
2.4	Orbital evolution below the SB; small- N simulations	22
2.4.1	Initial conditions	22
2.4.2	Qualitative behaviour	23
2.4.3	Capture rates	24
2.4.4	Eccentricity oscillations below the SB	25
2.4.5	Diffusion in angular momentum above and below the SB	29
2.5	Dynamical evolution of the S-stars	36
2.5.1	Initial conditions	36
2.5.2	Orbital evolution	38
2.5.3	Tidally disrupted and ejected stars	42
2.5.4	Diffusion coefficients	42
2.5.5	A new criterion for the location of the barrier	50
2.6	Steady-state distribution	51

2.6.1	Analytic solutions	54
2.6.2	Numerical solutions	58
2.7	Discussion	62
2.7.1	Limits on the typical S-star age from the N -body simulations	62
2.7.2	S-star relaxation times for different field-star models	63
2.7.3	Generalizations to other galactic nuclei	66
2.7.4	Caveats of TPI	68
2.7.5	Location of the sign change of $\langle \Delta \ell \rangle$ at high ℓ	69
2.7.6	Comparison of steady-state solutions	69
2.8	Conclusions	73
2.A	Simple tests of TPI	75
2.A.1	Test stars orbiting the SBH	75
2.A.2	Test stars orbiting a field star that orbits the SBH	78
2.B	Dependence of the NRR diffusion coefficients on γ	78
2.C	Extrapolating the shape of the cumulative eccentricity distribution for the S-star simulations	81
2.D	Dependence of derived diffusion coefficients on time lag in the simulations	83
2.E	Steady-state solutions to the Fokker-Planck equation	83
2.E.1	Solution of the steady-state equation	83
2.E.2	Analytic expressions for the diffusion coefficients	85
2.E.3	Explicit analytic steady-state solutions	85
2.F	N -body simulations with $\gamma = 1$	86
2.G	Equivalence of two critical radii	88
3	Probing the formation of planetesimals in the Galactic Centre using Sgr A* flares	91
3.1	Introduction	92
3.2	Setting the stage	93
3.2.1	Models of the GC	93
3.2.2	Disc models	96
3.3	Stripping planetesimals from stars in the GC	97
3.3.1	Stripping by the SBH	97
3.3.2	Stripping by gravitational encounters with other stars	99
3.4	Dynamics of planetesimals orbiting the SBH	103
3.4.1	Fokker-Planck equation	103
3.4.2	Boundary and initial conditions	105
3.4.3	Results: distributions and disruption rates	105
3.5	Discussion	117
3.5.1	Comparison to observations: constraints on $N_{a/\star}$	117
3.5.2	Internal scattering of planetesimals by planets	117
3.5.3	Scaling of the disruption rate: tidal disruption of planets	117
3.5.4	Changes of the stellar orbit prior to stripping	119
3.5.5	Special case: a burst of flares?	120
3.6	Conclusions	124
3.A	Planetesimal stripping	126
3.A.1	Stripping by the SBH	126
3.A.2	Stripping by gravitational encounters	127

3.B	Terms appearing in the Fokker-Planck equation	128
3.B.1	Gravitational scattering flux	128
3.B.2	Flux into the loss cone	129
3.B.3	Collision flux	131
3.C	Approximate semi-analytic solutions to the time-dependent Fokker-Planck equation	133
3.D	Scaling of the disruption rate	134
4	Secular dynamics of hierarchical quadruple systems: the case of a triple system orbited by a fourth body	137
4.1	Introduction	138
4.2	Methods	139
4.2.1	Expansion of the Hamiltonian	139
4.2.2	Orbit averaging	142
4.2.3	Equations of motion and numerical algorithm	142
4.2.4	The importance of the octupole-order cross terms	144
4.2.5	Comparisons to direct N -body integrations	150
4.3	Global evolution of highly hierarchical systems	151
4.3.1	Examples: A and B initially coplanar	152
4.3.2	Examples: A and B initially highly inclined	155
4.3.3	Qualitative trends	156
4.3.4	Quantitative dependence on $\mathcal{R}_0$	157
4.3.5	Behaviour near $\mathcal{R}_0 = 1$	161
4.3.6	General relativistic effects	161
4.4	Discussion	163
4.4.1	Application: planetary systems	163
4.4.2	Application: observed stellar quadruples	168
4.5	Conclusions	171
4.A	The Hamiltonian for hierarchical quadruple systems	175
4.A.1	Hierarchical triple system orbited by a fourth body	175
4.A.2	Two binaries orbiting each other	182
4.B	Test of the SecularQuadruple algorithm for three-body systems	184
5	A triple origin for the lack of tight coplanar circumbinary planets around short-period binaries	187
5.1	Introduction	188
5.2	Methods and assumptions	190
5.3	The planet-shielding effect	191
5.3.1	<i>Kepler</i> transiting circumbinary systems	191
5.3.2	A triple with a shrinking inner binary orbit	198
5.4	Planets in triples with short-period inner binaries	201
5.4.1	Initial conditions	201
5.4.2	Results	209
5.5	Discussion	220
5.5.1	An approximate analytic condition for planet shielding – implications for other systems	220

5.5.2	Shielding of the planetary orbit by the inner binary	220
5.5.3	The fate of planets with unstable orbits	221
5.5.4	Implications for planets around blue straggler stars	222
5.5.5	Approximations in the integrations	222
5.5.6	Other dissipative effects	223
5.6	Conclusions	224
5.A	<i>Kepler</i> transiting circumbinary planets	226
5.B	Magnetic braking in triples	228
6	Secular dynamics of hierarchical multiple systems composed of nested binaries, with an arbitrary number of bodies and arbitrary hierarchical structure	231
6.1	Introduction	232
6.2	The generalized Hamiltonian for hierarchical multiple systems	234
6.2.1	Definition and description of the system	234
6.2.2	The expanded Hamiltonian	236
6.2.3	Orbit averaging	237
6.2.4	General implications	238
6.2.5	Numerical algorithm	242
6.3	Tests: S-type multiplanet systems in single and multiple stellar systems	243
6.3.1	Single-star systems	243
6.3.2	Binary-star systems	255
6.3.3	Triple-star systems	268
6.4	Secular constraints in observed systems	273
6.4.1	30 Arietis	273
6.4.2	Mizar and Alcor	285
6.5	Discussion	286
6.5.1	The validity of orbit averaging	286
6.5.2	Short-term instabilities	288
6.6	Conclusions	288
6.A	Derivation of the expanded and orbit-averaged hierarchical multiple Hamiltonian	291
6.A.1	Description of the system structure	291
6.A.2	Expansion of the potential energy	297
6.A.3	Derivation of the kinetic energy term	301
6.A.4	Rewriting summations in the Hamiltonian	304
6.A.5	Orbit averaging	334
6.A.6	Estimate of the importance of non-binary terms for nested planetary systems	344
6.A.7	Ad hoc expression for the 1PN Hamiltonian	348
7	Implications of secular evolution on high eccentricity migration giving rise to hot Jupiters in multiplanet systems	351
7.1	Introduction	352
7.2	Methods and assumptions	354
7.2.1	Notations and overview	354
7.2.2	Secular dynamics	354
7.2.3	Tidal evolution	354

7.3	Verification of the secular method with N -body integrations	356
7.3.1	Initial conditions	356
7.3.2	Results	357
7.3.3	Orbit crossings	360
7.4	Conditions for producing high eccentricities due to secular evolution	361
7.4.1	Dependence of the maximum eccentricities on the AMD with varying eccentricities and inclinations	361
7.4.2	Dependence of the maximum eccentricities on the semimajor axes ratios	365
7.4.3	Tidal disruption versus tidal dissipation	371
7.5	Population synthesis	380
7.5.1	Initial conditions	381
7.5.2	Stopping conditions	382
7.5.3	Results	382
7.6	Discussion	398
7.6.1	HJ contribution and comparisons to other variants of high- e migration .	398
7.7	Conclusions	399
	Samenvatting	403
	Summary	409
	Bibliography	414
	Curriculum Vitae	423
	List of publications	425
	Acknowledgements	427

1 Introduction

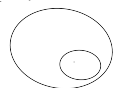
1.1 Gravity

Gravity is one of the four fundamental forces of nature. It plays a crucial role in the Universe at all scales, from the onset of the Big Bang, at the present day, and in the future. On Earth, gravity ensures that objects are not flung into space. In the Solar system, it binds the planets to the Sun, which itself was formed by the gravitational collapse of a giant gas cloud. Most stars are bound by gravity in galaxies like the Milky Way and smaller clusters like globular clusters. The largest known gravitationally bound and orbiting systems are clusters of galaxies, with scales on the order of Mpc, and crossing times as long as Gyr. On the other hand, at the time of writing, ESO's Rosetta spacecraft is bound to and orbiting comet 67P at a mere distance of 10 km.

Despite the crucial role of gravity on these disparate scales, its fundamental nature is not yet completely understood. On the scale of galaxies (kpc), measurements of the velocities of galaxies in galaxy clusters, and the rotation rate of stars suggest the presence of additional unseen, i.e. dark matter (Zwicky 1933; Rubin et al. 1980), or a modification of the laws of gravity on these scales, giving rise to Modified Newtonian Dynamics, or MOND (Milgrom 1983). On the other extreme end of the length scale spectrum of the Planck length (1.6×10^{-35} m), classical general relativity (GR) currently cannot be reconciled theoretically with quantum physics.

Between these extremes of length scales, gravity is generally well described in terms of the theory of GR. To extraordinary accuracy, pulse timing measurements of pulsars in binary systems have confirmed predictions by GR of the apsidal precession and the shrinkage and circularisation of the binary orbit due to the emission of gravitational waves (GWs) (Hulse & Taylor 1975; Burgay et al. 2003; Antoniadis et al. 2013). Until very recently, the existence of GWs was only proven indirectly through the orbital decay of these pulsar binaries. In a spectacular discovery, the first direct detection of GWs by the ground-based LIGO detector was announced on February 11 2016 (Abbott et al. 2016). The source, GW150914, was a stellar black hole binary (with masses of ≈ 36 and $\approx 29 M_{\odot}$) that merged to a single black hole of $\approx 62 M_{\odot}$, releasing $\approx 3 M_{\odot}$ of mass-energy in the form of GWs. The observed waveform was in extremely good agreement with the predictions of GR.

The GW150914 system clearly resides in the regime of *strong* gravitational fields. In the limit of *weak* gravitational fields, GR reduces to the Newtonian laws of motion and gravity, which describe well the dynamics of most planetary and stellar systems. In some cases, it is necessary to include in the Newtonian description corrections from GR, which are usually well described in terms of the post-Newtonian (PN) framework (Einstein et al. 1938). For example,



in the Solar system, GR corrections contribute to the apsidal precession of Mercury, and together with Newtonian gravitational perturbations from other planets, amount precisely to the measured apsidal precession rate.

1.2 History of the N -body problem

The Newtonian laws of motion and gravitation have been the subject of research by mathematicians and (astro)physicists for centuries. In the 17th century, Isaac Newton, in his *Principia* (Newton 1687), famously solved his laws of motion and gravity for two-body systems, providing a theoretical understanding of Kepler's laws of planetary motion (Kepler 1609). In modern notation, these laws are formulated mathematically as

$$m_i \ddot{\mathbf{R}}_i = \sum_{j \neq i}^N \frac{Gm_i m_j (\mathbf{R}_j - \mathbf{R}_i)}{||\mathbf{R}_j - \mathbf{R}_i||^3}, \quad (1.1)$$

where m_i and $\mathbf{R}_i$ are the mass and the position vector of body i , respectively. Although seemingly simple, Newton's law of gravity turn out to be notoriously hard to solve for systems with more than two bodies. In the 18th century, progress was made by Euler and Lagrange for specific three-body configurations (notably, the circular restricted three-body problem). In the context of the Solar system, Laplace and Lagrange developed approximate solutions for the long-term evolution of the planetary orbits. Although very useful, these solutions were approximate and apply only to planetary systems with a central massive body with nearly circular and coplanar orbits. Also, strong interactions between planets, which could lead to collisions or ejections, are not taken into account, as well as orbital resonances.

The difficulty of finding a general solution to the 3- or N -body problem led to the establishment in 1885 of a prize by the journal *Acta Mathematica* in honour of the 60th birthday of King Oscar II of Sweden and Norway. The problem was posed as follows.

Given a system of arbitrarily many mass points that attract each according to Newton's law, under the assumption that no two points ever collide, try to find a representation of the coordinates of each point as a series in a variable that is some known function of time and for all of whose values the series converges uniformly.

Although he did not solve this specific problem, Poincaré received the prize for his work on the three-body problem. In particular, he showed that no solutions exist in terms of known 'simple' analytic functions, and in the process, he laid the foundation for chaos theory.

The actual solution to King Oscar's prize as it was originally formulated, although restricted to $N = 3$, was found by Sundman (Sundman 1912). For the case of three bodies, he derived a solution in terms of a convergent power series of $t^{1/3}$. Although formally a solution, it did not give physical insight into the three-body problem. In addition, the series converges extremely slowly, making it not useful for practical applications — in order to be useful for astronomical observations, one would need at least $10^{8 \times 10^6}$ terms (Belorizsky 1930).

Nearly 80 years later, Sundman's solution for $N = 3$ was generalised to arbitrary N by Wang (1991). As before, the solution does not give much physical insight, and converges extremely slowly.

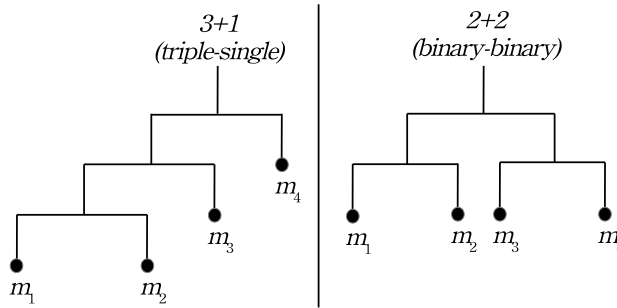


Figure 1.1: A schematic representation of two hierarchical quadruple systems in a mobile diagram. Left: a ‘3+1’ or ‘triple-single’ configuration. Right: a ‘2+2’ or ‘binary-binary’ system.

Although no physically meaningful solutions exist to the general N -body problem, analytic or semi-analytic results can be obtained for certain restricted systems. In particular, in this thesis, we will consider systems which are ordered in a hierarchy, i.e. hierarchical systems.

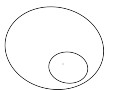
1.3 Hierarchies in nature

Planetary and stellar systems are often ordered in a hierarchical fashion. This can be explained intuitively by noting that on short time-scales, i.e. time-scales comparable to the orbital periods or crossing times, non-hierarchical systems are usually unstable, or stable but unstable to perturbations¹. For example, in the Solar system, all planets orbit a common centre, the Sun (more correctly, the centre of mass of the solar system). Their orbits are sufficiently separated to ensure stability, at least on time-scales that are short compared to the lifetime of the Sun. The Solar system also contains ‘deeper’ hierarchical layers: most planets are orbited by moons, which in turn can be orbited by lighter objects (e.g. the Apollo spacecraft around the Moon).

Other, more exotic, hierarchies have been observed outside of the Solar system. For example, planets have been observed to orbit stellar binaries (circumbinary planets). A very useful tool in the study of hierarchical systems is the *mobile diagram*, a tree diagram in which the hierarchy of the system is depicted schematically². An example is shown in the left part of Fig. 1.1, where a quadruple system of the ‘triple-single’ or ‘3+1’ configuration is depicted. This could represent a circumbinary planet (m_3) orbiting a stellar binary ($m_1 + m_2$); the stellar binary itself is part of a stellar triple system (third star m_4). Another possible hierarchical configuration with four bodies is the ‘binary-binary’ or ‘2+2’ configuration, depicted in the right part of Fig. 1.1. In the latter configuration, two binary systems are orbiting each other. One might expect that these configurations of four bodies are very rare and are ‘freaks of nature’, if existent at all. This is not the case, however; ≈ 1 per cent of all solar-type stellar systems within 67 pc of the Sun are hierarchical quadruple systems (Tokovinin 2014a,b). About 2/3 of those appear in the ‘2+2’ configuration, and 1/3 in the ‘3+1’ configuration.

¹An exception in the case of a three-body system (i.e. star-planet-asteroid) is the occurrence of two stable points for asteroids orbiting the star in the same orbit as the planet, but advanced or delayed in phase (i.e. the 4th and 5th Lagrange points). Notable examples are the Trojan and Greek asteroids that share their orbit with Jupiter.

²This term was first introduced by Evans (1968).



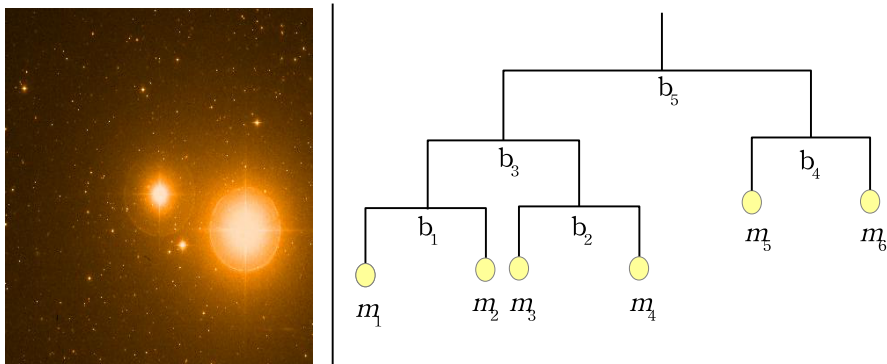


Figure 1.2: Left: image of Mizar and Alcor (image credit: ESO Online Digitized Sky Survey). Right: mobile diagram of the stellar sextuple system.

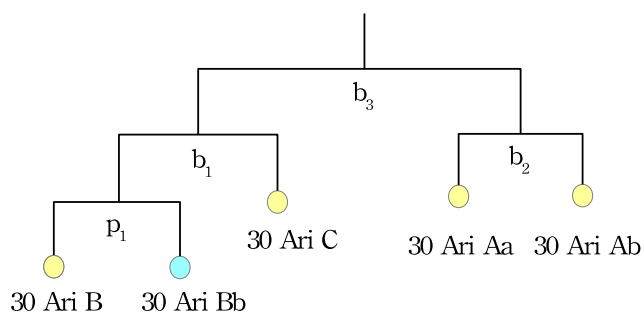


Figure 1.3: Mobile diagram of the stellar quadruple system 30 Ari with an additional planet, 30 Ari Bb.

Even more complicated stellar systems are known. For example, the famous visual binary Mizar and Alcor (Fig. 1.2, left) has been used as a vision test since antiquity (Allen 1899). Mizar itself is also a visual binary, known from as early as 1617, when Benedetto Castelli reported resolving it in a letter to Galileo Galilei (Ondra 2004; Siebert 2005). The components in Mizar, Mizar A and B, later turned out to be binaries themselves. In fact, Mizar A is the first spectroscopic binary known, found by Antonia Maury and reported by Pickering (1890). A few years later, the binary nature of Mizar B, also a spectroscopic binary, was discovered independently by Frost (1908) and Ludendorff (1908). Nearly a century later, Alcor was found to be a binary as well (Mamajek et al. 2010; Zimmerman et al. 2010). The mobile diagram of Mizar and Alcor, which is effectively a sextuple system, is shown in Fig. 1.2 (right). Another example is 30 Ari, which is a stellar quadruple system in the ‘2+2’ configuration in which a planet has been discovered around one of the stars (Fig. 1.3).

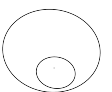
The hierarchies discussed so far are of the ‘simplex’ type composed of nested binaries, i.e. with each parent having only two children. Another type is the ‘multiplex’ type, in which each parent can have more than two children, and the children are not necessarily in hierarchical orbits with respect to each other. This latter means that crossing orbits and, therefore, close encounters, are possible.

Multiplex systems are observed in nature in addition to simplex systems. Most notably, in the Galactic Centre (GC), which is a distance of ≈ 8 kpc from the Sun, $\sim 10^6$ stars all orbit a supermassive black hole (SBH) of mass $\approx 4 \times 10^6 M_\odot$, at a distance ranging from only a few mpc (≈ 206 AU), to a few pc. These nuclear star clusters (NSCs) are known to exist in other galaxies as well; they may in fact be present in the centre of *every* galaxy.

Apart from dominating their motion on short time-scales, the SBH in an NSC can have dramatic effects on the lives of stars and other objects such as compact objects and planetesimals residing in such an environment. When passing within the tidal radius of the SBH, tidal forces can break up and disrupt a star, producing an energetic multiwavelength outburst known as a tidal disruption event, of which a handful have been observed so far.

Similarly, planetesimals with sufficiently large radius passing within the tidal radius of the SBH are separated into smaller fragments. These fragments, which are not bound by gravity but by molecular forces (and are therefore not broken up into even smaller pieces by the tidal force of the SBH), can subsequently be vaporised due to friction with the ambient gas around the SBH. This produces a potentially observable short-duration flare (Zubovas et al. 2012). The latter flares have been proposed to explain the near infrared and X-ray flares observed on a nearly daily basis from the SBH in the GC, Sgr A*.

More tightly bound compact objects orbiting the SBH such as white dwarfs, neutron stars and stellar-mass black holes are not tidally disrupted; however, the gradual emission of GWs can cause the compact object to spiral into the SBH. This process can take thousands of orbits to complete, producing a high signal-to-noise GW signal that is expected to be observable by future space-based GW detectors.



1.4 Secular evolution

1.4.1 Principles

As mentioned before, by necessity, hierarchical systems are stable on short times-scales, i.e. time-scales on the order of the orbital periods. On time-scales much longer than the orbital periods, the orbits can be interpreted as infinitesimally thin wires with a mass density that is proportional to the time spent at the corresponding point in the orbit. In the simplest case of a circular orbit (i.e. zero eccentricity), this implies a circle with uniform mass density. As the eccentricity of the orbit is increased, the time spent at apocentre increases relative to the time spent at pericentre, therefore the mass density is higher at apocentre.

These wires interact gravitationally and exert mutual torques on each other, thus transferring angular momentum between them. On the other hand, the orbital energies (i.e. the semimajor axes) remain constant. This implies that the eccentricities and the orientations of the orbits change. This type of long-term gravitational interaction is generally referred to as secular evolution. Depending on the type of system, secular evolution can have large implications. In the most extreme cases, eccentricities can be driven to nearly unity, e.g. leading to collisions or mergers of stars, planets or compact objects.

The simplest, and best-studied hierarchical system in which secular evolution can take place is the hierarchical three-body system (cf. the left part of Fig. 1.1, with m_4 removed). In this system, secular evolution is manifested in the form of Lidov-Kozai (LK) oscillations, named after its discoverers Lidov (Lidov 1962) and Kozai (Kozai 1962)³. If the system is highly hierarchical, i.e. if the orbital period of the ‘outer’ orbit, $P_{\text{orb,out}}$, is much ($\gtrsim 10^2$ times) longer than the ‘inner’ orbital period, $P_{\text{orb,in}}$, and if the outer orbit is not highly eccentric, then LK oscillations are highly regular. Furthermore, if the initial mutual inclination between the orbits, i_{tot} , lies within a certain range, $39.2^\circ \lesssim i_{\text{tot}} \lesssim 140.8^\circ$, then the eccentricity of the inner orbit oscillates, reaching a maximum value that depends mainly on i_{tot} ⁴. In this limit, the eccentricity of the outer orbit is unaffected. This can be understood intuitively by noting that the angular momentum of the outer orbit is much larger than that of the inner orbit; therefore, angular momentum exchanges have a much larger effect on the inner orbit. Assuming the inner orbit is initially circular, the maximum eccentricity reached in the inner orbit during the LK cycles is

$$e_{\text{max,in}} = \sqrt{1 - \frac{5}{3} \cos^2(i_{\text{tot}})}. \quad (1.2)$$

Equation (1.2) shows that $e_{\text{max,in}}$ increases as i_{tot} approaches 90° . In other words, the larger the initial mutual inclination, the larger the maximum inner orbit eccentricity.

The eccentricity oscillations, as well as the oscillations of the orbital orientations, occur on a time-scale of the order of P_{LK} , where

$$P_{\text{LK}} = \frac{P_{\text{orb,out}}^2}{P_{\text{orb,in}}} \frac{m_1 + m_2 + m_3}{m_3} (1 - e_{\text{out}}^2)^{3/2}. \quad (1.3)$$

³Until a few years ago, LK oscillations were almost invariably referred to in the literature as ‘Kozai’ cycles, without reference to Lidov. Lidov was, in fact, the first to publish his results – in Russian – in 1961; the English version appeared in 1962. In 1962, Lidov presented his results at a conference that Kozai also attended. Kozai, likely guided by Lidov’s work, published a paper on the secular evolution of asteroids around Jupiter later that year in 1962.

⁴Strictly, this is only for initially circulating orbits with zero initial eccentricity. In other cases, eccentricity oscillations are possible for lower inclinations.

Here, e_{out} is the eccentricity of the outer orbit, m_1 and m_2 are the masses of the bodies in the inner orbit, and m_3 is the mass of the tertiary body in the outer orbit. The precise period of the oscillations also depends on the initial orientation of the orbit; the latter dependence, which is much weaker compared to that of the other quantities, is not included in equation (1.3). Equation (1.3) shows that secular evolution takes place on a time-scale that is typically much longer than the outer orbital period. In addition to the orbital periods, the mass ratio and the outer orbit eccentricity are important factors in the period of the LK oscillations.

The above results apply only to highly hierarchical systems in which the tertiary body never closely approaches the inner binary (i.e. $P_{\text{orb,out}} \gg P_{\text{orb,in}}$ and $e_{\text{out}} \ll 1$). In this case, the system can be described by the lowest-order terms in the expansion of the Hamiltonian in terms of the small separation ratio $r_{\text{in}}/r_{\text{out}}$, known as the ‘quadrupole’ order⁵. At this order, and making another assumption relating to angular momentum conservation which is valid if $P_{\text{out}} \gg P_{\text{in}}$, the LK oscillations can be described analytically (Kinoshita & Nakai 1999, 2007). The properties equations (1.2) and (1.3) follow from these analytical solutions.

However, if the triple system is not extremely, but still moderately hierarchical ($10^1 \gtrsim P_{\text{orb,out}}/P_{\text{orb,in}} \gtrsim 10^2$), then LK oscillations are more complex, and they are typically no longer amenable to analytical solutions. In particular, the next-order terms in the expansion of the Hamiltonian, known as the ‘octupole-order’ terms, give rise to modulations of the quadrupole-order oscillations, on a time-scale that is given approximately by $P_{\text{LK}}/\epsilon_{\text{oct}}$, where the ‘octupole parameter’ $\epsilon_{\text{oct}} < 1$ is given by (Lithwick & Naoz 2011; Katz et al. 2011; Naoz et al. 2013a; Teyssandier et al. 2013; Li et al. 2014b)

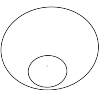
$$\epsilon_{\text{oct}} = \frac{|m_1 - m_2|}{m_1 + m_2} \left[\frac{m_1 + m_2}{m_1 + m_2 + m_3} \left(\frac{P_{\text{in}}}{P_{\text{out}}} \right)^3 \right]^{1/2} \frac{e_{\text{out}}}{1 - e_{\text{out}}^2}. \quad (1.4)$$

When ϵ_{oct} is large, typically $\epsilon_{\text{oct}} \gtrsim 10^{-3}$, the octupole-order terms are important, and this can give rise to ‘flips’ of the mutual orbital planes from prograde to retrograde, and vice versa. These orbital flips are associated with extremely high eccentricities in the inner orbit, in some cases as high as $1 - e_{\text{in}} \sim 1 - 10^{-7}$. In stellar triple systems with compact objects, these extreme eccentricities can lead to violent head-on collisions of carbon-oxygen (CO) white dwarfs (WDs) (Katz & Dong 2012). Such collisions are likely to result in a type Ia supernova (SNe Ia) explosion because of temperatures reaching $\sim 10^9$ K during the collision, high enough to trigger a thermonuclear runaway explosion. Although certainly spectacular, current estimates of the event rates are very low, i.e. $\sim 10^{-3}$ of the SNe Ia rate computed from binary population synthesis studies (assuming mergers of CO WDs in circular orbits after common-envelope evolution) (Hamers et al. 2013). The latter are about a factor 10 lower compared to observations.

Even when eccentricities are not high enough for direct collisions, in many cases they can be high enough to trigger strong orbital energy dissipation, notably due to tides raised on bodies, or due to the emission of GWs. Both processes are highly sensitive to the orbital eccentricity, and the dissipation rate peaks at the maximum eccentricities of the LK cycle. Either quickly or gradually, the orbit shrinks (i.e. $P_{\text{orb,in}}$ decreases) and is circularised, and the oscillations are quenched.

This quenching phenomenon can be understood by noting that as the inner orbital period decreases, the LK time-scale increases (cf. equation 1.3), whereas the precession rate of the

⁵Note: the term ‘quadrupole’ is easily confused with ‘quadruple’.



inner orbit due to additional forces increases. Notable additional forces arise from GR and tidal bulges. Their associated precession rates increase with decreasing $P_{\text{orb, in}}$. With an increasing LK time-scale and a decreasing precession time-scale due to additional forces, LK cycles are typically damped, or completely quenched.

1.4.2 Implications

This process of LK cycles combined with orbital energy dissipation has important astrophysical implications. In stellar triple systems, it is believed that LK cycles with tidal friction (LKCTF) drive the production of short-period binaries (orbital periods $\sim 3 - 6$ d) (Fabrycky & Tremaine 2007; Naoz & Fabrycky 2014). Such short-period binaries are believed not to have formed in isolation; in the case of LKCTF, the initial period can be as long as 10^4 d. This is consistent with observations of the period distribution of binaries, which shows a peak around $\sim 3 - 6$ d (Tokovinin et al. 2006; Tokovinin 2014a), and with the tertiary fraction of binaries, which is nearly unity for binaries with periods shorter than 3 d (Tokovinin et al. 2006).

Another important implication relates to the origin of hot Jupiter (HJ) planets, i.e. gas giant planets of order Jupiter mass that orbit their host star in extremely tight orbits, with periods between $\sim 3 - 10$ d. It is commonly thought that these planets cannot have formed in situ because of the low gas densities and high temperatures during the protoplanetary disk phase in this region so close to the star (however, e.g. recently Batygin et al. 2015 suggested that in situ formation due to core accretion is possible). Instead, they might have formed at more distant regions from the star, i.e. beyond the ‘snow line’ where volatiles condense into solid ice grains. The precise location of the snow line is uncertain and depends on various parameters (including the disk age), but it is typically a few AU from the host star (i.e. an orbital period of a few 100 days). This implies that the planet must have migrated by two orders of magnitude in orbital period.

Two main migration scenarios have been proposed: disk migration and high-eccentricity (or high- e) migration. In the case of disk migration, the planet migrates because of gas drag within the protoplanetary disk. There are a number of concerns with this type of migration. Firstly, theoretical models for disk migration do not predict a preferred orbital period at which migration stops. Observations, however, show a pile-up of HJs at periods near 5 d. Furthermore, the observed obliquities of HJs, i.e. the angles between the stellar spin and the orbit of the HJ, are found to be both consistent with zero, and ranging between 0 and 180 degrees, i.e. retrograde orbits are also possible. Such high-obliquity orbits are at odds with disk migration models, in which the obliquity remains close to, or exactly zero.

An alternative migration scenario for the formation of HJs is high- e migration. In this case, the eccentricity of the orbit of the planet is driven to high values by one or more external perturbers. The high eccentricities lead to strong tidal evolution and hence orbital energy dissipation and circularisation, eventually leading to a HJ. The eccentricity excitation could be due to planet-planet scattering interactions, or due to secular interactions induced by a stellar binary companion or other (inclined) planets. In the case of secular interactions, a pile-up is predicted, around 3 days, because of the stalling mechanism mentioned above. This is (approximately) consistent with observations.

However, there are problems with these secular high- e scenarios. Evidently, a third body (in addition to the star and the Jupiter-like planet) is required. If the third body is another Jupiter-like planet, then the former needs to be either in a very close, eccentric and/or inclined orbit

with respect to the inner planet. This is inconsistent with observations of the majority of HJs, which show that HJs do not have any close-in planetary companions. Furthermore, high mutual planetary inclinations seem unlikely if the star is isolated (this could be different if there is a more distant binary companion which could have torqued and hence warped the protoplanetary disk). If the third body is a star (i.e. binary companion), then it can be further away and a high initial mutual inclination with respect to the (inner) planet seems more plausible. However, such binary companions are observed for only ~ 0.5 of the observed HJs (Ngo et al. 2015).

In an alternative scenario proposed by Wu & Lithwick (2011), the planetary system consists of three or more planets in orbits that are mildly eccentric and inclined (eccentricities on the order of $0.2 - 0.4$ and inclinations on the order of $5 - 20^\circ$). The planets need not be very closely separated; their semimajor axes could e.g. be 1, 6 and 16 AU. In this case, chaotic motion is likely to occur. Very high eccentricities can be excited, particularly in the innermost orbit, during highly irregular secular oscillations. The time-scale for reaching high eccentricities can be as long as a few Gyr. This process of ‘secular chaos’ is likely very important for determining the long-term evolution of planetary systems. For example, in the Solar system, the orbit of Mercury is unstable due to secular chaos on a time-scale of ~ 5 Gyr relative to the current Solar system (Laskar 1994; Ito & Tanikawa 2002; Laskar 2008; Laskar & Gastineau 2009).

Secular chaos is an interesting scenario for producing HJs on time-scales of Gyr, without requiring close-in planets. This is currently consistent with observations, which cannot yet detect planets at these larger separations with good certainty. Although promising, this scenario of high- e migration merits further investigation, in particular with consideration of a larger range of parameter space compared to the work of Wu & Lithwick (2011), who considered a limited number of systems. In particular, it is unclear what range of HJs properties can be attained.

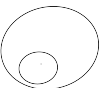
1.5 This thesis

This thesis deals with the dynamical evolution of hierarchical systems with various hierarchies, in various astrophysical contexts. We begin by considering systems of the ‘multiplex’ type, dominated by an SBH (Chapters 2 and 3). Subsequently, we focus on simplex-type systems, beginning with the general dynamics of quadruple systems (Chapter 4), and continuing with an application to circumbinary planets (Chapter 5). Finally, we generalise our methods to hierarchical simplex-type systems with an *arbitrary* hierarchical structure and an *arbitrary* number of bodies (Chapter 6), and apply these methods to the production of HJs through secular evolution in multiplanet systems (Chapter 7).

1.5.1 Chapter 2 – Relativistic dynamics around an SBH

In Chapter 2, we consider systems of the ‘multiplex’ type, and study the long-term evolution of orbits around an SBH in a NSC, with particular emphasis on Sgr A*, the SBH in the GC. About 20 B-type stars, known as the S-stars, have been observed to orbit around a single point in the GC. Two decades of observations have revealed their detailed orbital properties, and most importantly, the mass of the central object. Several studies have shown that this central object, beyond any reasonable doubt, should be an SBH with a mass of $\sim 4 \times 10^6 M_\odot$ (see Alexander 2005 and Genzel et al. 2010, and references therein).

As mentioned before, orbits around an SBH, when highly eccentric, can give rise to tidal



disruption events, or high signal-to-noise GW signals in the case of the inspiral of a compact object onto the SBH. Here, we refer to both events simply as ‘disruptions’. Objects on orbits that are initially highly eccentric can be brought close to the SBH and be disrupted already during the first orbit; the corresponding orbits are referred to the ‘loss cone’. Because the orbital periods around the SBH are at most a few Myr (in the case of the GC), objects on initially loss cone orbits are rapidly depleted, and do not contribute to the long-term disruption rate (i.e. on Gyr time-scales).

However, because of gravitational interactions between objects on nearby orbits, orbits near the SBH are not static, and change due to various effects. First of all, random encounters, or ‘fly-bys’ of two (or more) objects change all orbital properties (the energy, angular momentum and orientation). When taking into account many of such encounters, the effect is to stochastically alter the orbital properties. Because these encounters are essentially random, this effect is referred to as ‘non-resonant relaxation’ (NRR). However, in the case of a NSC, objects move on (nearly) Keplerian orbits, and some encounters are, therefore, correlated. This gives rise to secular interactions, which are similar to LK cycles in the case of the hierarchical three-body problem. However, in this case, for a given ‘subject’ orbit, there are many ‘inner’ and ‘outer’ orbits, and the secular effects are much more complicated and irregular. Nevertheless, orbits can still be driven to high eccentricities, potentially leading to disruption events. This type of evolution is known as ‘resonant relaxation’ (RR).

Close to the SBH, where orbital periods are short and close encounters are very rapid (i.e. the velocity dispersion is large), RR is very effective at changing the orbital properties and dominates compared to NRR. Taking only into account the Newtonian effects of NRR and RR, objects on these close orbits would rapidly evolve to loss cone orbits. However, close to the SBH where orbital speeds v can reach a few per cent of the speed of light c , GR effects are important, and these affect the dynamics. To lowest order in v/c , the effect of GR is apsidal precession of the orbit; a familiar example is the relativistic precession of the orbit of Mercury around the Sun. In the case of a NSC, however, relativistic precession rates can be much higher compared to Mercury. Consequently, the ‘efficiency’ of the Newtonian torques associated with RR can be reduced substantially, thereby dampening, or completely mitigating the effects of RR.

This effect was first shown by Merritt et al. (2011), who carried out detailed N -body simulations of a small cluster of 50 stellar-mass black holes, each with a mass of $50 M_{\odot}$, orbiting an SBH with a mass of $10^6 M_{\odot}$. In their simulations, Merritt et al. (2011) noticed that as an orbit becomes more eccentric (due to RR), the relativistic precession rate increases, until at a certain semimajor axis and eccentricity, the orbit ‘bounces’ back to lower eccentricity. Because the relativistic precession is also known as Schwarzschild precession, Merritt et al. (2011) coined this phenomenon the ‘Schwarzschild barrier’ (SB). Equating the time-scales associated with Schwarzschild precession and RR, Merritt et al. (2011) obtained an expression for the semimajor axis and eccentricity associated with the SB, i.e. the ‘location’ of the SB.

Direct N -body simulations scale with N^2 . Combined with the fact that the typical required time-step is very short (fractions of orbits which can be as short as a few yr, whereas the desired simulation time is on the order of a few Myr, or even Gyr to achieve a steady state), this means that further work was limited by the computational complexity of detailed simulations. In particular, realistic representations of the centres of NSCs require on the order of 10^3 stars, which is not feasible with currently available direct N -body codes.

In order to make progress, we developed a new special-purpose N -body code, `TEST PARTICLE`

INTEGRATOR (TPI). By making a number of simplifying assumptions about the motion of particles, this code allows efficient integration with a large number of particles around the SBH (order 10^3). It is based on a splitting of particles into field particles that move on (nearly) fixed Keplerian orbits around the SBH, and test particles that respond to the time-dependent potential of the field stars and the SBH. In this approach, interactions between field particles are ignored apart from effects associated with mass precession and relativistic precession. Furthermore, the mass of test particles is neglected, and, therefore, test particle-test particle interactions are ignored as well. Nevertheless, by comparing with slower, but more accurate N -body codes that do not make these assumptions, we show that TPI can be used to simulate the long-term angular momentum evolution of orbits around the SBH.

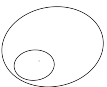
We apply TPI to the S-stars in the GC. We assume that the S-stars were formed through binary disruption. In this scenario, first proposed by Hills (Hills 1988), a stellar binary approaches the SBH from a wide but highly eccentric, or a hyperbolic orbit. Due to the tidal force of the SBH, the binary is unbound, and one of the stars is ejected and escapes from the SBH, whereas the other remains bound to the SBH in a tight and highly eccentric orbit. The escaped star is believed to give rise to hypervelocity stars (HVSs), a group of stars observed in the Milky Way with velocities on the order of 10^3 kms^{-1} . The majority of HVSs are consistent with originating from the GC, and their velocities (taking into account deceleration in the Galactic potential) are consistent with the escape speed of the SBH. Furthermore, their number is consistent with the number of S-stars (Perets et al. 2007).

The eccentricities of the orbits of the stars that remain bound to the SBH are expected to be very high, typically $e \sim 0.98$. However, this is inconsistent with the current observed eccentricity distribution, which shows that the typical eccentricity is high (~ 0.7), but not as high as ~ 0.98 . To explain the statistically-lower observed eccentricities, we assume that the orbits of the S-stars, in particular their eccentricities, are perturbed by gravitational interactions with a background cluster of stellar-mass black holes. In TPI, the latter cluster is represented by 4800 field particles, ≈ 100 times larger compared to the simulations of Merritt et al. (2011), and predicted by previous studies of stellar populations in the GC (Hopman & Alexander 2006b).

In our simulations, the eccentricity distribution of the S-stars indeed evolves from a highly eccentric distribution corresponding to the initial assumed formation through binary disruption, to a less eccentric distribution which is statistically consistent with the observed eccentricity distribution of the S-stars. This gives indirect evidence for both the existence of the cluster of stellar-mass black holes, and the formation mechanism of the S-stars. In particular, when assuming an alternative background distribution of lower-mass main-sequence stars, the time-scale for the eccentricity distribution of the S-stars to relax to the observed distribution is much longer, and is inconsistent with the age of the S-stars.

As a next step, we determine specific coefficients from the simulations that describe diffusion in angular momentum. These diffusion coefficients are very useful in Fokker-Planck simulations, which can be used to describe the statistical dynamical evolution on very long time-scales (exceeding Gyr), i.e. time-scales that are currently inaccessible with direct N -body integrations. Furthermore, the coefficients yield analytical insight into the long-term orbital evolution. As a function of the angular momentum variable $\ell = \sqrt{1 - e^2}$, we find that the SB can be associated with a rapid drop of the coefficients with decreasing ℓ . The value of ℓ for which this occurs, turns out to be approximately consistent with the previous expression for the SB derived by Merritt et al. (2011).

Finally, we give simple analytic expressions for the diffusion coefficients as a function of



ℓ , and we apply them to the Fokker-Planck equation to obtain the steady-state distribution.

1.5.2 Chapter 3 – Planetesimals in the GC

In the previous work, we considered the long-term dynamical evolution of orbits around an SBH, and applied our results to the S-stars. In this chapter, we focus on another type of bodies in the GC, i.e. planetesimals, and relate these to observations of flares observed from Sgr A* in the near infrared and X-ray (Baganoff et al. 2001, 2003; Genzel et al. 2003a; Dodds-Eden et al. 2011; Barrière et al. 2014). These flares occur approximately once per day, and are 3-100 times more luminous than the quiescent emission of the central radio source, Sgr A*.

One of the proposed explanations for the flares is that they result from the tidal disruption of planetesimals with radius $\gtrsim 10$ km (Zubovas et al. 2012). The latter authors showed that if such a planetesimal passes within ~ 1 AU of the SBH, it is broken up into smaller fragments by tidal forces; the fragments subsequently vapourize because of friction with the ambient gas. When the vapourized material is mixed with the accretion flow onto Sgr A*, enough energy could be released to produce an observable flare.

However, very little is known about the formation and evolution of planetesimals in galactic nuclei like the GC. One possibility is that they are formed in a large-scale spherical cloud orbiting the SBH. Another possibility is that they are born in debris discs around stars, and are stripped by the tidal force of the SBH or gravitational encounters with other stars. In this chapter, we investigate both scenarios by means of numerical integrations of the Fokker-Planck equation. We model the orbital energy evolution around the SBH, taking into account the effects of gravitational perturbations from various background perturbers.

We show that the predicted present-day disruption rates in the GC differ very little between the two scenarios and that this conclusion depends weakly on the details of the perturbers or other assumptions. In both scenarios, we find a disruption rate of $\sim 1 \text{ d}^{-1}$ assuming that the number of planetesimals per (late-type) star is $N_{a/\star} = 2 \times 10^7$. The number $N_{a/\star} = 2 \times 10^7$ is consistent with debris discs observed around stars in the Solar neighbourhood. In the first scenario, in which the planetesimals are formed in a large cloud, this implies that the number of bodies formed is strongly correlated with the number of stars, and this requires finetuning of the quantity $N_{a/\star}$. We favour the more natural explanation that planetesimals in galactic nuclei similar to the GC are formed in debris discs around stars, no differently than planetesimals around stars in the Solar neighbourhood.

1.5.3 Chapter 4 – Secular evolution of hierarchical quadruple systems

The first two chapters dealt with hierarchical systems of the multiplex type. In the remainder of this thesis, we shall focus on hierarchical systems of the simplex type, i.e. systems consisting of nested binary orbits. Each binary has two children, and the children can be either bodies or binaries themselves (see e.g. Fig 1.1). The simplest of these, the hierarchical triple, has been studied in detail before. However, this is not the case for hierarchical systems with more than three bodies.

Here, we begin with the simplest possible extension and consider four bodies in the ‘3+1’ configuration, i.e. a hierarchical triple orbited by a fourth body. We assume that the system is sufficiently hierarchical, i.e. the ratios x_i of the binary separations are small. In that case,

it is appropriate to expand the Hamiltonian in terms of the x_i . Subsequently, we orbit average the expanded Hamiltonian, and implement the equations of motion into a computer code, SECULARQUADRUPLE.

Subsequently, we study the secular evolution of highly hierarchical systems that are well described by the lowest order terms in the Hamiltonian, and characterise the evolution in terms of ratios of LK time-scales applied to different binary pairs.

1.5.4 Chapter 5 – Explaining the lack of circumbinary planets around short-period binaries

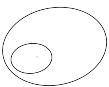
As mentioned before, planets have been observed around stellar binaries (i.e. circumbinary planets). The detection of transiting circumbinary planets is more tractable around short-period binaries, i.e. binaries with periods less than ~ 7 d. However, sofar, no such binaries have been found. As also mentioned before, short-period main sequence binaries have been suggested to form in triple systems, through a combination of LK cycles and tidal friction, i.e. LKCTF. Here, we apply the method developed in Chapter 4, and we show that coplanar circumbinary transiting planets are unlikely to exist around short-period binaries, due to secular evolution. This is consistent with the currently observed lack of transiting circumbinary planets around short-period binaries.

1.5.5 Chapter 6 – Secular evolution of hierarchical multiple systems

This chapter generalises the method and the algorithm developed in the earlier Chapter 4. Here, we derive the expanded Hamiltonian for systems consisting of nested binary (i.e. simplex) orbits, with an *arbitrary* number of bodies, and an *arbitrary* hierarchical configuration. This includes hierarchical triple and quadruple systems (in both the ‘2+2’ and ‘3+1’ configurations), but also much more complicated systems. For example, it applies to multistar systems with multiplanet subsystems around any of the stars. We develop an algorithm to implement the generised method, making long-term integrations of hierarchical systems with such complex hierarchies feasible, and present first applications. In particular, we apply our method to Mizar and Alcor and to 30 Ari.

1.5.6 Chapter 7 – Hot Jupiters in multiplanet systems

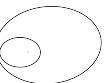
In the above, we mentioned a scenario for forming HJs in multiplanet systems through secular eccentricity excitation in multiplanet systems. Current studies of this variant of high- e migration have, until recently, been limited to a number of N -body integrations. Although accurate, such simulations are computationally costly, making large parameter space exploration unfeasible. Here, we apply the algorithm developed in Chapter 6 to study the formation of HJs in multiplanet systems with at least three planets. In particular, we investigate the dependence of the final orbital period distribution of the HJs, and compare this distribution to observations.



2 | Relativistic dynamics of stars near a supermassive black hole

General relativistic precession limits the ability of gravitational encounters to increase the eccentricity e of orbits near a supermassive black hole (SBH). This ‘Schwarzschild barrier’ (SB) has been shown to play an important role in the orbital evolution of stars like the Galactic Centre S-stars. However, the evolution of orbits below the SB, $e > e_{\text{SB}}$, is not well understood; the main current limitation is the computational complexity of detailed simulations. Here, we present an N -body algorithm that allows us to efficiently integrate orbits of test stars around an SBH including general relativistic corrections to the equations of motion and interactions with a large ($\gtrsim 10^3$) number of field stars. We apply our algorithm to the S-stars and extract diffusion coefficients describing the evolution in angular momentum L . We identify three angular-momentum regimes, in which the diffusion coefficients depend in functionally different ways on L . Regimes of lowest and highest L are well described in terms of non-resonant relaxation and resonant relaxation, respectively. In addition, we find a new regime of ‘anomalous relaxation’. We present analytic expressions, in terms of physical parameters, that describe the diffusion coefficients in all three regimes, and propose a new, empirical criterion for the location of the SB in terms of the L -dependence of the diffusion coefficients. Subsequently we apply our results to obtain the steady-state distribution of angular momentum for orbits near an SBH.

Hamers, A.S., Portegies Zwart, S.F., Merritt, D. MNRAS 443, 1, 355-387 (2014)



2.1 Introduction

Near a supermassive black hole (SBH), evolution of stellar orbits due to gravitational encounters is influenced by three factors. (1) Orbits are nearly Keplerian. (2) The number, $N_*(r)$, of stars contained within radius r is likely to be small. (3) Relativistic corrections to the equations of motion can be important. Considerations (1) and (2) are the basis of ‘resonant relaxation’ (RR; Rauch & Tremaine 1996), which identifies changes in orbital angular momenta with torques due to the nearly stationary mass rings corresponding to the Keplerian orbits. General relativity (GR) appears in this theory as one of several mechanisms capable of inducing orbital precession, hence setting the ‘coherence time’ over which the torques can act (Rauch & Tremaine 1996). But recent work reveals that GR can play a much more essential role, particularly in the case of orbits that are highly eccentric. Such orbits precess due to GR at a higher rate than most other orbits at the same radii. This rapid precession tends to quench the effects of the torques (Hopman & Alexander 2006b), but it also leads to a less obvious, and more striking, phenomenon: a ‘barrier’ in angular momentum that ‘reflects’ stars that strike it from above (i.e. from orbits of higher angular momentum; Merritt et al. 2011, hereafter MAMW11). Following MAMW11, we refer to the locus in (energy, angular-momentum) space where these phenomena occur as the ‘Schwarzschild barrier’ (SB), in recognition of the fact that the precession that underlies the phenomenon is due to the spinless, or Schwarzschild, part of the SBH metric. Compact objects are expected to dominate the stellar population at these small radii, and the existence of the SB is expected to mediate their capture by the SBH (MAMW11; Brem et al. 2014; however, for spinning SBHs highly eccentric orbits may not suffer a blockade, Amaro-Seoane et al. 2013). Capture events, or EMRIs (extreme-mass-ratio inspirals; Sigurdsson & Rees 1997), would otherwise be expected to be a potentially observable source of low-frequency gravitational waves (Amaro-Seoane 2012).

Many processes exist that can deposit stars on to highly eccentric orbits around an SBH. These processes include close encounters between stars (Goodman 1983), encounters between stars and massive perturbers (Perets et al. 2007) or a stellar disc (Chen & Amaro-Seoane 2014), and the tidal disruption of stellar binaries that approach the SBH on nearly radial orbits (Hills 1988). These ideas are relevant to models that attempt to explain the presence of young stars very near to the SBH in the Galactic Centre (GC). Some of these stars, the so-called S-stars, have orbits of high enough eccentricity that they must lie below the predicted location of the SB (Antonini & Merritt 2013). If the S-stars were deposited initially on to orbits with even higher eccentricities than observed today (which would be the case, for instance, in the binary disruption model), then the fraction of S-stars initially below the SB was even higher in the past. The evolution of such highly eccentric orbits over Myr time-scales is not well described by existing theory of resonant or non-resonant relaxation; it depends in critical ways on the barrier phenomena described above (Antonini & Merritt 2013).

Progress in understanding the relativistic dynamics of nuclear star clusters has been driven in large part by the recent development of extremely accurate and efficient computer codes for solving the (small-) N -body problem (Mikkola & Aarseth 1993, 2002; Mikkola & Tanikawa 1999; Mikkola & Merritt 2008). But the new results summarized above also imply that the number of stars in a real galaxy that are subject to GR phenomena is probably much larger than can be handled efficiently by these codes. For instance, in the Milky Way, the number of stars and stellar remnants inside $r = a_{\text{SB,max}}$, the largest semimajor axis for which the SB exists, is probably of the order of $10^3 - 10^4$. Efficient, Monte Carlo algorithms for evolving test-orbits

near the SB were developed in MAMW11 and applied to the S-star problem by Antonini & Merritt (2013), but these algorithms are based on an extremely simple model for the torquing potential and its time dependence.

A major goal of this paper is to develop an alternate algorithm that represents the field-star forces much more accurately than the Monte Carlo routines in MAMW11, but which nevertheless is efficient enough to be used for realistically large N -values. Our code, called `TEST PARTICLE INTEGRATOR (TPI)`, explicitly follows the motion of the field stars along their precessing, Keplerian orbits, but ignores interactions between them. The motion of the test stars is then followed by direct integration in the time-varying potential produced by the N field stars. Relativistic terms are included in the equations of motion of both test and field stars via the post-Newtonian (PN) approximation. This algorithm contains all of the dynamics which are believed to be important for the evolution of orbits due to RR in the presence of relativity, excluding only the changes in the field-star distribution that would be due to the RR torques themselves, or to perturbations from the test stars.

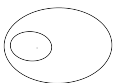
In Section 2.3, we describe `TPI` and perform a number of basic tests. In Section 2.4, the orbital evolution below the SB is studied using simulations similar to those performed by MAMW11. By restricting to a small number of particles we can compare our results to results obtained from N -body codes in which the simplifying assumptions adopted in `TPI` are relaxed. In Section 2.5, we apply our code to the S-star cluster; similar simulations with the other N -body codes used in Section 2.4 are currently not feasible. Assuming that the S-stars are formed in highly eccentric orbits, which is consistent with the binary disruption model, and adopting a cusp of stellar black holes, we study the orbital evolution of the S-stars after their formation.

The models explored here were designed to represent the GC, but it is useful to ask how our results would generalize to other nuclei. To this end, in Sections 2.4.5 and 2.5.4 we extract angular-momentum diffusion coefficients from the simulations and compare them with existing theory. We argue in Section 2.2 that diffusion in angular momentum should be well described by non-resonant relaxation (NRR) at very low L ($e \gg e_{\text{SB}}$), and by RR at high L ($e \ll e_{\text{SB}}$). But in the angular-momentum regime near and ‘below’ the SB (i.e. $e \gtrsim e_{\text{SB}}$), neither RR nor NRR is applicable (MAMW11). By computing angular-momentum diffusion coefficients from the simulations, we are able, for the first time, to demonstrate the existence of the three regimes and to quantify their L -dependence. This allows us, in Section 2.6, to estimate the steady-state angular-momentum distribution implied by the Fokker-Planck equation. In Section 2.7, we discuss the implications of our results and we conclude in Section 2.8.

2.2 Time-scales

The focus in this paper is on orbits near an SBH that are very eccentric compared with the typical eccentricity expected in, say, a ‘thermal’ distribution, $\langle e \rangle = 2/3$. The time-scale over which such eccentric orbits evolve due to gravitational encounters with other stars can depend strongly on e . We begin by summarizing what is known about that dependence. As we will see, in regimes near or below the SB, i.e. $e \gtrsim e_{\text{SB}}$, the eccentricity dependence is still poorly understood and that is one motivation for carrying out the simulations described below.

The top panel of Fig. 2.1 plots several curves in the (a, ℓ) (semimajor axis, normalized angular momentum) plane that are relevant to stars orbiting near an SBH. This figure adopts an SBH mass $M_{\bullet} = 4 \times 10^6 M_{\odot}$, the value in the Milky Way.



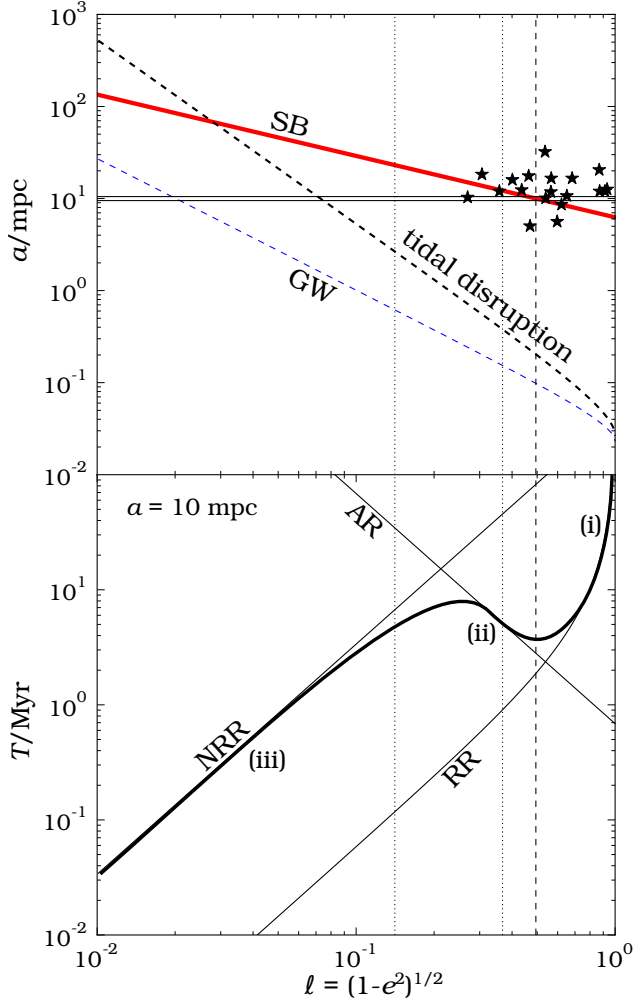


Figure 2.1: Top panel: characteristic curves in the (a, ℓ) (semimajor axis, normalized angular momentum) plane for stars orbiting near the MW SBH. Curves are defined in the text; the assumed stellar mass and radius are $10 M_{\odot}$ and $8 R_{\odot}$, respectively. Shown as stars are the (a, ℓ) values of the 19 S-stars with well-determined orbits (Gillessen et al. 2009). Bottom panel: approximate time-scales for ℓ to change by of order itself due to gravitational interactions with field stars, at a ‘slice’ in semimajor axis at 10 mpc. RR = resonant relaxation; AR = anomalous relaxation; NRR = non-resonant relaxation. Three angular-momentum regimes are indicated; see the text. The vertical dotted lines indicate the range in ℓ expected for the tidal disruption of a stellar binary by the SBH (Hills 1988; see also Section 2.5). The vertical dashed line marks the value of ℓ_{SB} at $a = 10$ mpc, the semimajor axis in the bottom panel.

The red curve labelled ‘SB’ is given by

$$(1 - e^2)^{1/2} \equiv \ell = \frac{r_g}{a} \frac{M_\bullet}{M_\star(a)} \sqrt{N_\star(a)}. \quad (2.1)$$

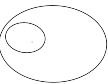
Here, $r_g \equiv GM_\bullet/c^2$ is the gravitational radius of the SBH, $M_\star(r)$ is the mass in stars within radius r , and $N_\star(r)$ is the number of stars within r ; $M_\star = m_\star N_\star$. The quantity ℓ is a normalized angular momentum: $\ell = L/L_c$ with $L_c(a) = \sqrt{GM_\bullet a}$ the angular momentum of a circular orbit of semimajor axis a . Equation (2.1) is the approximate locus in the (a, ℓ) plane where the change in the angular momentum L of a test star due to torques from the other stars, in one relativistic precession cycle (equation 2.2), is of the order of L . In their (small-) N -body simulations, MAMW11 found that equation (2.1) predicts very well the maximum eccentricity reached by orbits as they evolve due to gravitational encounters. In Fig. 2.1, we adopted for $M_\star(r)$ and $N_\star(r)$ the expressions given in Section 2.5.1.

The curve labelled ‘tidal disruption’ is the locus of orbits having periapses at the tidal disruption radius defined in Section 2.5.1. The curve labelled ‘GW’ is an estimate of where changes in orbital eccentricity due to gravitational-wave energy loss occur at the same rate as changes due to two-body relaxation (MAMW11, equation 62).

The lower panel of Fig. 2.1 shows how time-scales for changes in orbital angular momentum are believed to depend on ℓ . Here, ‘time-scale’ is defined as the time for L to change by of order itself (and not, for instance, for L to change by of order L_c). We identify three regimes. (i) Above the SB, RR is effective (Rauch & Tremaine 1996). On time intervals longer than the ‘coherence time’ – the time for an orbit of typical eccentricity to precess – orbital angular momenta are expected to undergo a random walk due to torques from the $\sqrt{N}$ asymmetry in the mass distribution. The associated time-scale is $\sim \ell^2$ times t_{RR} ; the latter is given by equation (2.14). (ii) At $\ell \lesssim \ell_{SB}$, GR precession strongly reduces the ability of the $\sqrt{N}$ torques to change ℓ . While no adequate theory yet exists for diffusive evolution in this regime, approximate arguments (MAMW11) suggest that the evolution time-scale should increase rapidly with decreasing ℓ . The curve labelled AR (‘anomalous relaxation’) shows, qualitatively, how the evolution time-scale might depend on ℓ in this regime. (iii) At sufficiently small ℓ , the diffusion time due to AR is expected to become very long due to the rapid GR precession. But NRR is not affected by the precession, and because the time-scale for NRR to change ℓ is proportional to ℓ^2 , at sufficiently small ℓ , this time must become shorter than the AR time-scale.

As we discuss below, the value of ℓ at the transition between regimes (ii) and (iii) depends on various quantities, including the number of stars (for a given $M_\star$, say). This result is likely to be important when calculating rates of ‘barrier penetration’, since the dominant mechanism driving diffusion past the SB will be different in different nuclear models. We will argue that in the simulations of MAMW11, the particle number was small enough that the NRR regime extended all the way to the SB for some values of the semimajor axis; while in more realistic nuclear models, diffusion below the SB would need to contend with AR.

The time-scales plotted in Fig. 2.1 do not tell the whole story. For instance, there is a separate time-scale associated with drift in angular momentum (due to the first-order diffusion coefficients) and that time-scale is relevant to the ‘bounce’ phenomenon that occurs near the SB, as described below. At sufficiently small L , orbits around a spinning SBH will also be affected by spin-orbit (Lense-Thirring) torques (Merritt & Vasiliev 2012), a phenomenon whose consequences for the angular-momentum evolution will not be explored here.



2.3 Method

In TPI, we exploit the property that well within the sphere of influence of an SBH the motion of the stars is dominated by the SBH, i.e. the stellar motion is well described in terms of perturbed Keplerian orbits. Torques acting on these stars give rise to exchange of angular momentum between stars. This process is known as resonant relaxation (RR) and affects the eccentricities of the orbits. Furthermore, two-body (non-resonant) interactions affect the orbital energies in addition to their angular momenta. When considering a large ensemble of stars, however, these processes should not strongly affect the *mean* angular momenta and energies provided that the system is dynamically relaxed. On the other hand, energy exchange and RR are important when considering individual stars. This consideration motivates a split between dynamically relaxed field stars and test stars that evolve dynamically in time as a consequence of both angular momentum and energy exchanges with field stars. We define a test star as a particle with zero mass, i.e. a particle that does not affect the field stars and other test stars.

The field stars are assumed to follow uniformly precessing Kepler orbits with constant semi-major axis a , eccentricity e , inclination i and longitude of the ascending node Ω . The argument of periastris ω is advanced linearly in time according to the rate prescribed by analytical formulae that include precession due to GR (Schwarzschild precession) and Newtonian precession due to the distributed mass in stars (mass precession). The advance per orbital period P due to the Schwarzschild precession, to first PN order, is given by (Weinberg 1972)

$$\Delta\omega_{\text{1PN},P} = 6\pi \frac{G(m_\star + M_\bullet)}{a(1-e^2)c^2}. \quad (2.2)$$

Here, $m_\star$ is the field star mass, $M_\bullet$ is the SBH mass, G is the gravitational constant and c is the speed of light. Periastris advance due to mass precession depends on the detailed distribution of the mass. In all the models considered here, we assume a spherical field-star distribution with density $\rho_\star(r) \propto r^{-2}$. In this case, the apsidal advance due to mass precession per orbital period is (Merritt 2013a, equation 4.87)

$$\Delta\omega_{\text{MP},P} = -2\pi \frac{M_\star(a)}{M_\bullet} \frac{\sqrt{1-e^2}}{1+\sqrt{1-e^2}}. \quad (2.3)$$

Here, $M_\star(a)$ is the total field star mass within radius $r = a$.

In TPI, the motion of the field stars is calculated with a Kepler solver that advances the positions and velocities for a given time interval assuming unperturbed Keplerian ellipses. The resulting positions and velocities $\mathbf{r}$ and $\mathbf{v}$ are subsequently rotated in the orbital plane to account for the in-plane precession resulting from both Schwarzschild and mass precession:

$$\mathbf{r} \rightarrow \cos(\Delta\omega) \mathbf{r} + \sin(\Delta\omega) \hat{\ell} \times \mathbf{r}, \quad (2.4a)$$

$$\mathbf{v} \rightarrow \cos(\Delta\omega) \mathbf{v} + \sin(\Delta\omega) \hat{\ell} \times \mathbf{v}. \quad (2.4b)$$

Here, $\Delta\omega = \Delta\omega_{\text{1PN},\Delta t} + \Delta\omega_{\text{MP},\Delta t}$ is the total precession angle in time interval Δt , and $\hat{\ell} = \mathbf{r} \times \mathbf{v} / \|\mathbf{r} \times \mathbf{v}\|$ is the unit specific angular-momentum vector. By treating the motion of the field stars in this way, the interactions between field stars are modelled in an approximate method that neglects two-body encounters and resonant torques. This makes it computationally feasible to include a large ($\gtrsim 10^3$) number of field stars.

The test stars are integrated using a direct-summation N -body code. It is advantageous to employ Kustaanheimo-Stiefel regularization (Kustaanheimo & Stiefel 1965) for their motion around the SBH. Tests have shown that in the absence of field stars this method reduces the required number of integration steps while at the same time it increases the accuracy. In TPI, each test star forms a regularized and independent two-body system with the SBH. The perturbing acceleration $\mathbf{a}$ is given by:

$$\mathbf{a} = \mathbf{a}_{\text{SBH, PN}} + \mathbf{a}_{\text{field, 0PN}}. \quad (2.5)$$

Here, $\mathbf{a}_{\text{SBH, PN}}$ is the PN acceleration from the SBH. We have implemented 1PN, 2PN and 2.5PN terms for a non-spinning SBH (Damour & Deruelle 1981), and 1.5PN and 2.0PN terms that arise from spin of the SBH (Kidder 1995). In most of the simulations presented here we restrict to including only the 1PN terms. The quantity $\mathbf{a}_{\text{field, 0PN}}$ is the Newtonian acceleration from the field stars. To integrate the regularized equations of motion, we use a standard fourth-order Hermite predict, evaluate and correct integration scheme (Makino 1991).

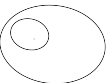
The implementation of the PN terms in our algorithm can be compared to other algorithms based on geodesic solvers (Brem et al. 2014). In the latter algorithms, the Schwarzschild metric is used to obtain relativistic corrections to the equations of motion, thereby assuming that the metric is determined solely by the SBH, i.e. that the mass m of the particle orbiting the SBH can be neglected compared to $M_\bullet$. This is similar to our algorithm, in which m in the PN terms is set to zero (i.e. the symmetric mass ratio $\nu = mM_\bullet/(m + M_\bullet)^2 = 0$). A major difference is that in our algorithm the corrections are included to finite order of v/c , whereas a geodesic solver is in principle accurate to arbitrary order, provided that m is sufficiently small and that therefore there is no dissipation due to gravitational waves.

Test stars have individual block time steps Δt_{block} that are determined dynamically using time symmetrization (Funato et al. 1996). At the end of each integration step, a new time-symmetric time step Δt is calculated from:

$$\begin{aligned} \Delta t &= \frac{1}{2} \|\mathbf{u}\|^2 \left[f(\mathbf{u}_b^{(2)}, \mathbf{u}_b^{(3)}, \mathbf{u}_b^{(4)}) + f(\mathbf{u}_e^{(2)}, \mathbf{u}_e^{(3)}, \mathbf{u}_e^{(4)}) \right]; \\ f(\mathbf{u}^{(2)}, \mathbf{u}^{(3)}, \mathbf{u}^{(4)}) &= \eta \times \min \left[\frac{\|\mathbf{u}^{(2)}\|}{\|\mathbf{u}^{(3)}\|}, \left(\frac{\|\mathbf{u}^{(2)}\|}{\|\mathbf{u}^{(4)}\|} \right)^{1/2} \right]. \end{aligned} \quad (2.6)$$

Here, $\mathbf{u}$ is the regularized position vector, (i) indicates the i th derivative with respect to the regularized time, η is a time step parameter and the indices b and e indicate the beginning and end of the current step, respectively. Subsequently, the block time step is computed from $\Delta t_{\text{block}} = 2^k \times \Delta t_{\text{min}}$, where k is the largest positive integer such that $2^k \times \Delta t_{\text{min}} < \Delta t$ and $\Delta t_{\text{min}} = 1 \times 10^{-14}$ yr is the minimum time step that we allow in the simulations. The time between iterations is given by the minimum of the test star block time steps. At each iteration the positions and velocities of the field stars are shifted in their Kepler orbits and rotated according to equation (2.4).

In the case of a large number of test and/or field stars (typically if either number is $\gtrsim 10^3$) the evaluation of $\mathbf{a}_{\text{field, 0PN}}$ for all the test stars is the most computationally expensive part of the integration of the equations of motion. For this reason, we have implemented parallel computation of $\mathbf{a}_{\text{field, 0PN}}$ on CPUs using OPENMP as well as GPU-accelerated evaluation of $\mathbf{a}_{\text{field, 0PN}}$ using the SAPPORO library (Gaburov et al. 2009).



In TPI , the detection of captures of test stars by the SBH is implemented. We assume that during the integration step the test star moves in a straight line $\mathbf{r}(s) = \mathbf{r}_b + s(\mathbf{r}_e - \mathbf{r}_b)$ where $\mathbf{r}_b$ and $\mathbf{r}_e$ are the (non-regularized) position vectors at the beginning and the end of the integration step, and $s \in [0, 1]$ is a parameter. We check if any of the points on this trajectory satisfies $\mathbf{r}(s)^2 = r_{\text{capt}}^2$ for $s \in [0, 1]$, where r_{capt} is the capture radius. If this is the case then the test star has either just grazed or penetrated the capture sphere and we register a capture event. After a test star has been captured it is recorded and removed from the simulation.

To validate TPI , we have performed several simple tests of interactions between test stars and the SBH and between test and field stars. These tests are described in Appendix 2.A. In this paper, the time step parameter is set to $\eta = 0.02$; this choice is motivated in the latter appendix. Tests of TPI in the regime below the SB, which is the main focus of this paper, are described in detail in Section 2.4, where we also compare our results with those from other, slower, N -body codes in which the simplifying assumptions adopted in TPI are relaxed.

Before describing the results of the code comparisons, we note that even the more accurate algorithms discussed below contain potentially important approximations. These codes include the Newtonian terms from the SBH and the N bodies, plus the 1PN terms from the SBH alone. The latter terms are proportional to $G^2 M_\bullet^2 / r^3 c^2$, or to $(GM_\bullet / r)(v^2 / c^2)$, with r the distance from the SBH. At 1PN order, one can potentially do better, since the full N -body Hamiltonian is known, the so-called EIH (Einstein-Infeld-Hoffmann) Hamiltonian (Einstein et al. 1938). The EIH equations of motion also include terms of the order of $M_\bullet m_\star$ and $m_\star^2$ (Will 2014b). Given the small values of $m_\star / M_\bullet$ considered here, only the former, or ‘cross’, terms are likely to matter. In the context of apsidal precession, one expects the cross terms to induce changes of order

$$(\Delta\omega)_{\text{cross}} \approx (\Delta\omega)_{\text{M}} \times (\Delta\omega)_{\text{PN}}, \quad (2.7)$$

that is, the product of the shifts due to mass precession and to the Schwarzschild precession considered individually. Over sufficiently long times, the effects of the cross terms will of course accumulate, and it is an open question whether this might significantly impact the evolution of orbits near the SB.

2.4 Orbital evolution below the SB; small- N simulations

2.4.1 Initial conditions

As mentioned in Section 2.1 there exist several processes that can deposit stars below the SB on time-scales of the order of the Kepler period $P(a)$. Here, we study the evolution of orbits after deposition below the SB using simulations with TPI . We also include simulations performed with two direct-summation N -body codes, MI6 (Nitadori & Makino 2008; Iwasawa et al. 2011) and ARCHAIN (Mikkola & Merritt 2008). MI6 uses a mixed fourth-order and sixth-order Hermite integration scheme. The SBH is kept fixed at the origin, simplifying the equations of motion. In particular, this allows for PN accelerations to be calculated for star-SBH interactions only, avoiding the calculation of PN accelerations for star-star interactions. The latter are assumed to be negligible compared to the former. In MI6 1PN and 2.5PN accelerations are included. The ARCHAIN code is an essentially exact N -body code owing to chain regularization and it includes 1PN, 2PN and 2.5PN terms.

The initial conditions of our simulations were similar to those of the N -body simulations performed by MAMW11 and Brem et al. (2014). We sampled field stars of mass $m_\star = 50 M_\odot$ in Kepler orbits around an SBH of $M_\bullet = 1.0 \times 10^6 M_\odot$ with the following orbital distributions: semimajor axes a were sampled randomly between $a_{\min} = 0.1$ mpc and $a_{\max} = 10$ mpc corresponding to a stellar density distribution $\rho_\star(r) \propto r^{-2}$; a thermal eccentricity distribution was assumed and orbital angles were sampled randomly. The total number of field stars was $N_{\max} \equiv N_\star(a_{\max}) = 50$. We also carried out simulations with larger $N_\star(a_{\max})$ with TPI; the latter are discussed in Section 2.4.5.

In the case of ARCHAIN and MI6 we placed five of the field stars below the SB in the (a, e) parameter space. We define above and below the SB as $\ell \equiv \sqrt{1 - e^2} > \ell_{\text{SB}}$ and $\ell < \ell_{\text{SB}}$, respectively, where $\ell_{\text{SB}} = \ell_{\text{SB}}(a)$ is defined in equation (2.1); for the N -body simulations $M_\star(a) = m_\star N_\star(a) \approx m_\star N_{\max}(a/a_{\max})$. We will refer to these five stars as test stars, but we note that in the case of ARCHAIN and MI6 these stars are not massless and have the same mass as the field stars. In the case of TPI we initiated five test stars below the SB at the same values of a and e as those of the five test stars in ARCHAIN and MI6. In each simulation the five test stars shared a common value of a and e but were initiated with different (random) orbital angles and phases. We carried out a series of simulations with the following combinations of the initial values of a and e :

$$\begin{aligned} a &= 2 \text{ mpc}; & \log_{10}(1 - e) &\in \{-3.0, -2.5, -2.0\}; \\ a &= 4 \text{ mpc}; & \log_{10}(1 - e) &\in \{-3.3, -2.9\}; \\ a &= 8 \text{ mpc}; & \log_{10}(1 - e) &= -3.8. \end{aligned} \tag{2.8}$$

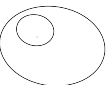
For each combination of a and e (i.e. each simulation with five test stars below the SB) we ran simulations with five different random realizations, obtaining 25 time series for each (a, e) pair.

The integration time per simulation was set to 1 Myr. The capture radius was $r_{\text{capt}} = 8 r_g \approx 3.8 \times 10^{-4}$ mpc, consistent with the capture radius of a compact object on to a non-spinning SBH (Will 2012). In all simulations we included 1PN terms; we also carried out integrations in which the 2.5PN terms were included (in the case of ARCHAIN, 2PN terms are included as well). However, because the 2.5PN terms cannot be included self-consistently in TPI we present in Section 2.4 only results in which the 2.5PN terms were excluded, with the exception of Section 2.4.3.

2.4.2 Qualitative behaviour

We show in Fig. 2.2 the eccentricity evolution for a test star with initially $a = 2$ mpc and $\log_{10}(1 - e) = -2.5$ as computed with each of the three codes, without the 2.5PN terms. We select two cases (corresponding to the two rows) in which the test star crosses the SB from below to above. Note that the initial conditions differ in each panel of Fig. 2.2, hence the panels should not be compared directly. Based on these and similar plots, we make the following qualitative observations.

1. Below the SB the eccentricity varies in an approximately periodic fashion, on a (short) time-scale consistent with the Schwarzschild precession time. There is also a component of its evolution that can be described as a random walk. (The latter was referred to as ‘anomalous relaxation’ in Section 2.2.)



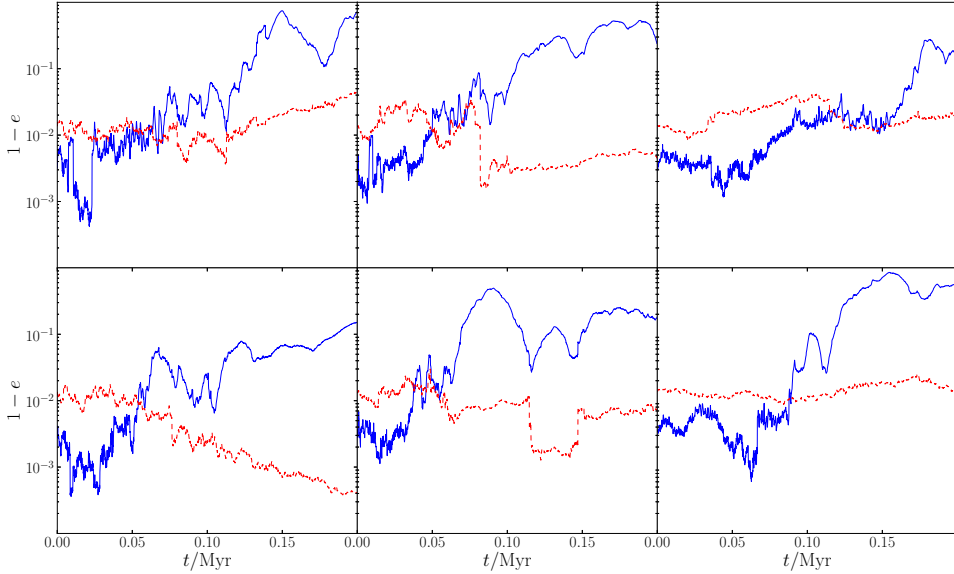


Figure 2.2: Eccentricity evolution (blue solid lines) and the SB (equation 2.1; red dashed lines) as function of time for test stars with initially $a = 2$ mpc and $\log_{10}(1 - e) = -2.5$ as computed with ARCHAIN (first column) MI6 (second column) and TPI (third column). We show cases in which the test star is not captured and moves to above the SB. The initial conditions differ in each panel, hence the panels should not be compared directly.

2. Above the SB the eccentricity variations are much larger, extending to $e \approx 0$, and have a longer associated time-scale. These features can be explained qualitatively in terms of RR, which is not quenched above the barrier.
3. Stars above the SB tend to remain there, since their trajectories ‘bounce’ on striking the barrier from above.
4. As a consequence of items (ii) and (iii), the SB acts as a diode or a one-way membrane: stars can only easily cross it in one direction, from below (high e) to above (low e).

In Sections 2.4.4 and 2.4.5, we explore some of these properties more quantitatively, and we also use them as a means of comparing the different codes.

2.4.3 Capture rates

Before turning to our observations from Section 2.4.2 we present in Table 2.1 the number of captured stars at the end of the simulation for the three codes ARCHAIN, MI6 and TPI. We include both simulations with (W) and without (WO) 2.5PN terms. Although the exact number of captured stars N_{capt} differs between the three codes, in all three cases there is a similar trend of increasing N_{capt} with both a and e . For example, for each of the codes without the 2.5PN terms

a/mpc	$\log_{10}(1-e)$	ARCHAIN		N_{capt} MI6		TPI	
		W	WO	W	WO	W	WO
2	-2.0	4	8	7	6	1	4
2	-2.5	9	14	11	9	13	8
2	-3.0	13	15	18	16	15	17
4	-2.9	11	13	9	7	13	11
4	-3.3	18	17	18	17	11	12
8	-3.8	18	13	22	15	21	21

Table 2.1: Number of captured test stars at $t = 1$ Myr for the three different codes. A distinction is made between simulations with (W) and without (WO) 2.5PN terms.

N_{capt} increases by a factor $\sim 3 - 4$ from $a = 2$ mpc and $\log_{10}(1-e) = -2.0$ to $a = 8$ mpc and $\log_{10}(1-e) = -3.8$.

2.4.4 Eccentricity oscillations below the SB

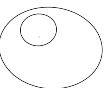
Frequency of oscillations

We obtained power spectra of the eccentricity and argument of periapsis from the simulations below the SB using the following method. For each time in the simulation t_0 we computed the time-scale $2t_{\text{GR}}$ for the Schwarzschild precession to change ω by 2π , $2t_{\text{GR}} = 2\pi P / \Delta\omega_{1\text{PN},P} = (1/3)(a/r_g)(1-e^2)P$ (cf. equation 2.2). Subsequently, we recalculated $2t_{\text{GR}}$ based on the mean values of a and e in the interval $t_0 < t < t_0 + \Delta t$, where $\Delta t = 8t_{\text{GR}}^1$. This procedure was repeated until convergence with respect to t_{GR} had occurred. The interval was rejected if for any of the points within it the star was above the SB, the number of points was less than 50 or the fractional changes in a and e satisfied $|\Delta a/a| > 0.04$ and $|\Delta e/e| > 0.04$, respectively. The latter criteria serve to minimize noise in the power spectra induced by sudden changes in a and e due to NRR. Power spectra of the eccentricity and argument of periapsis were subsequently computed for the accepted intervals. The starting search time for the subsequent interval was $t_0 + \Delta t$.

We show in Fig. 2.3 an example of power spectra obtained using the above method in a simulation with initially $a = 2$ mpc and $\log_{10}(1-e) = -2.5$, as computed with TPI. There is a peak in both power spectra at $f = f_{\text{GR}}$, where $f_{\text{GR}} \equiv 1/(2t_{\text{GR}})$. This is consistent with our observation in Section 2.4.2 that below the SB the eccentricity oscillations occur on the Schwarzschild precession time-scale. The peak in the power spectrum at $f \approx f_{\text{GR}}$ is higher for the argument of periapsis compared to the eccentricity because the Schwarzschild precession affects the argument of periapsis directly, whereas the effect on the eccentricity is indirect, i.e., through the $\sqrt{N}$ torques.

We applied the above method to all simulated (a, e) pairs of the test stars (cf. equation 2.8). For the obtained power spectra we determined the local maxima (shown for one example in the top panel of Fig. 2.3 with bullets) and we recorded the corresponding frequencies f_{max}

¹The factor 8 in $\Delta t = 8t_{\text{GR}}$ is a compromise between a long sampling interval (leading to much noise induced by two-body encounters) and a short sampling interval (leading to too few data points).



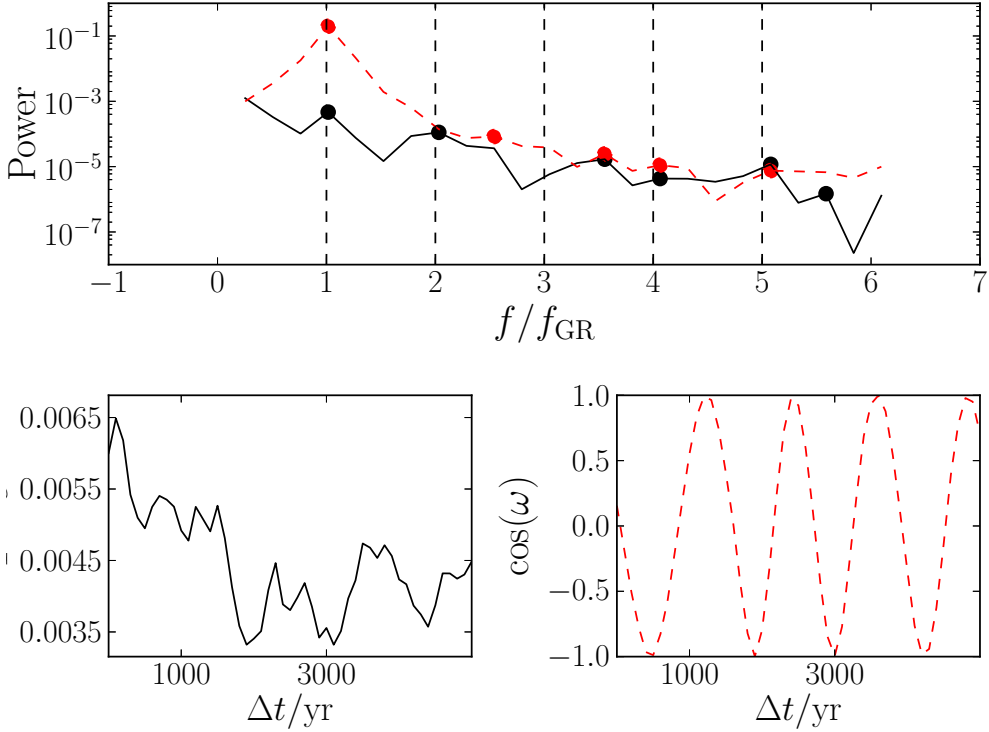


Figure 2.3: Top panel: power spectra of the eccentricity (black solid line) and the cosine of the argument of periapsis (red dashed line) for a simulation with initially $a = 2$ mpc and $\log_{10}(1-e) = -2.5$ as computed with `TPI`. Local maxima are indicated with bullets. The time series of the eccentricity and argument of periapsis used for the power spectra are shown in the bottom left and bottom right panels (time is expressed with respect to the beginning of the sampling interval). Refer to the text for the method used to select the sampling interval. Most of the power is contained at $f = f_{\text{GR}}$.

and amplitudes A , where A is the square root of the power. We show in the first column of Fig. 2.4 the resulting distributions of f_{max} for the three codes. There is a clear peak in the eccentricity spectra at $f_{\text{max}} \approx f_{\text{GR}}$. This peak can be interpreted as implying that the torquing potential (due to the $\mathcal{O}(N^{-1/2})$ asymmetry in the field star distribution) is basically lopsided, or $m = 1$, in character (MAMW11). Higher order terms in the multipole expansion of the field-star potential would give rise to eccentricity oscillations at higher integer frequencies of f_{GR} . The results shown in the first column of Fig. 2.4 indicate that these higher-order contributions are important, though typically not dominant.

We determined the amplitudes A_n of the peaks at higher integer frequencies $f_{\text{max}} \approx n f_{\text{GR}}$ and we normalized these to A_1 , the amplitude at $f_{\text{max}} \approx f_{\text{GR}}$. Frequencies for each n were selected from data satisfying $n - 0.3 < f_{\text{max}}/f_{\text{GR}} < n + 0.3$, where the limits are motivated by the distributions shown in first column of Fig. 2.4. We show the resulting distributions of A_2/A_1 for the three codes in the second column of Fig. 2.4. These distributions have peaks at roughly similar locations for each of the three codes, $A_2/A_1 \approx 0.4$. In the case of `ARCHAIN`

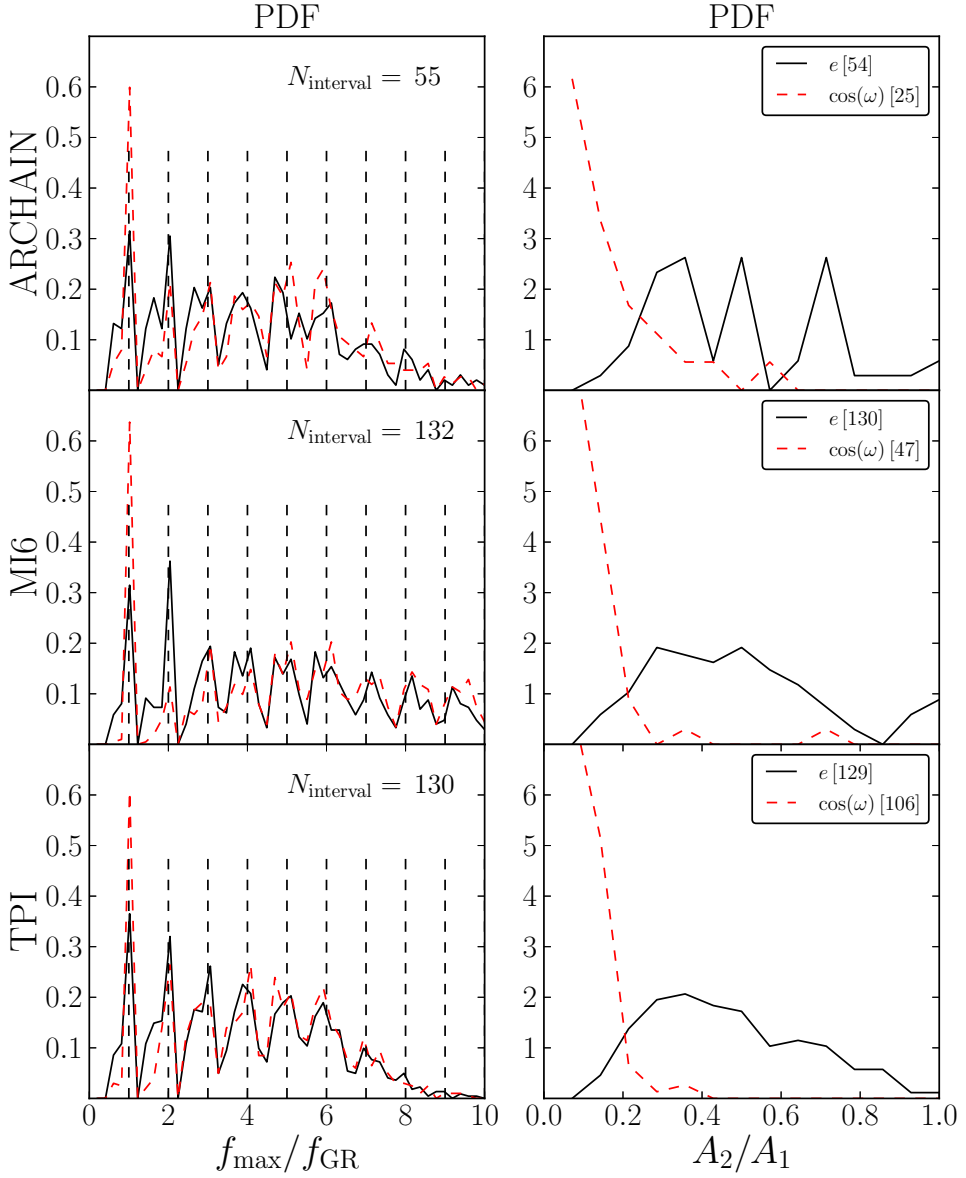
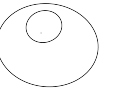


Figure 2.4: Left column: distributions of the local frequency maxima determined from the power spectra for the three codes. Black solid lines apply to the eccentricity and red dashed lines apply to the argument of periaapsis. Right column: distributions of the ratios of the amplitude at $f_{\max} \approx 2f_{\text{GR}}$ to the amplitude at $f_{\max} \approx f_{\text{GR}}$. The numbers in brackets are the number of data points in the frequency bin.



n	A_n/A_1		
	ARCHAIN	MI6	TPI
2	0.46 ± 0.19 [54]	0.55 ± 0.29 [130]	0.42 ± 0.16 [129]
3	0.22 ± 0.09 [53]	0.23 ± 0.09 [131]	0.25 ± 0.10 [126]
4	0.18 ± 0.06 [46]	0.14 ± 0.06 [113]	0.18 ± 0.06 [123]
5	0.13 ± 0.05 [43]	0.11 ± 0.06 [111]	0.13 ± 0.06 [113]
6	0.12 ± 0.06 [38]	0.12 ± 0.05 [106]	0.12 ± 0.05 [93]
7	0.08 ± 0.05 [21]	0.09 ± 0.05 [82]	0.07 ± 0.02 [46]
8	0.09 ± 0.05 [15]	0.08 ± 0.05 [75]	0.06 ± 0.03 [18]

Table 2.2: Amplitudes A_n for the eccentricity power spectra at frequencies $f_{\max} \approx n f_{\text{GR}}$ normalized to A_1 . The values are the median values and the median absolute deviations. The numbers in brackets are the number of data points in the frequency bin.

there appear to be several peaks, which may be due to a lack of data. We show in Table 2.2 the median and median absolute deviations of A_n/A_1 for $n \leq 8$. These values are consistent between the three codes.

It is conceivable that precession of the field stars affects the ratios A_n/A_1 in the above analysis. We have verified that for the duration of the sampling intervals the field stars do not precess by a large amount, i.e. in the case of MI6 on average only 3 (out of 45) field star orbits precess over an angle $\Delta\omega > \pi/2$ during Δt . In addition, we have carried out the analysis for TPI without field star precession (i.e. by setting $\Delta\omega = 0$ in equation 2.4) and we found no substantially different results, e.g. $A_2/A_1 \approx 0.52 \pm 0.21$, consistent with $A_2/A_1 \approx 0.42 \pm 0.16$ in the case where precession of the field stars is included (cf. Table 2.2).

Amplitude of oscillations

Here, a method is presented to obtain the amplitude of eccentricity oscillations below the SB and the results are compared to theoretical predictions. We expect the amplitude of the latter oscillations to depend on the (dimensionless) angular momentum $\ell = \sqrt{1 - e^2}$ and hence the distance in angular momentum to the SB. This is due to more rapid Schwarzschild precession for lower ℓ and therefore more efficient quenching of the effects of the $\sqrt{N}$ torques that would otherwise drive RR.

We adopt the Hamiltonian model presented in MAMW11 that includes the Schwarzschild precession, mass precession and the effects of a lopsided mass distribution, assumed to be oriented with respect to the orbit with an angle α . Let $\Delta\ell \equiv \ell_{\max} - \ell_{\min}$ be the amplitude of oscillations in ℓ , where $\ell_{\min}$ and $\ell_{\max}$ are the minimum and maximum angular momenta during one oscillation of duration $2t_{\text{GR}}$, respectively. Then if ℓ is sufficiently small, i.e. if the second and third terms of MAMW11 equation (41a) can be neglected with respect to the first term, $\Delta\ell$ depends on the average angular momentum $\langle\ell\rangle \equiv (1/2)(\ell_{\min} + \ell_{\max})$ via the relation (MAMW11 equation 46):

$$\Delta\ell \approx 2\langle\ell\rangle^2 A_D \sin(\alpha). \quad (2.9)$$

Here, A_D is a dimensionless parameter that specifies the strength of the lopsided component of the distributed mass in the Hamiltonian model. In terms of the model parameters A_D is

expressed by (MAMW11 equation 43b²):

$$A_D = \frac{C_{AD}}{2} \frac{S}{GM_\bullet/a^2} \frac{a}{r_g} = \frac{C_{AD}}{2\sqrt{N_\star(a)}} \frac{M_\star(a)}{M_\bullet} \frac{a}{r_g}. \quad (2.10)$$

Here, S is the amplitude of the lopsided distortion, and $M_\star(a)$ and $N_\star(a) = M_\star(a)/m_\star$ are the enclosed stellar mass and the number of enclosed stars, respectively. The parameter C_{AD} captures unspecified uncertainties in this model.

We obtained $\Delta\ell$ from the simulations with a method similar to that used to obtain power spectra. In order to minimize the effect of directed changes in ℓ over time-scales longer than t_{GR} the sampling interval was shortened to $\Delta t = 4 t_{GR}$ and the number of required points per sampling interval was reduced to 20. For the resulting sampling intervals we recorded the minimum and maximum values of ℓ and the mean value of a , $\langle a \rangle$. We binned the data into 100 bins of $\langle a \rangle$ with $1 < \langle a \rangle/\text{mpc} < 10$ and 10 bins of $\langle \ell \rangle$ with $0 < \langle \ell \rangle < 0.3$. For each bin we computed the mean values of $\langle \ell \rangle$ and $\Delta\ell$, which amounts to averaging these quantities over the angle α . We rejected bins if the bin size was less than or equal to 5.

In order to compare results from the simulations to the prediction of equation (2.9) we average this equation over the unit sphere and substitute A_D using equation (2.10) with $N_\star(a) \approx (a/a_{\max})N_{\max}$. Subsequently, we obtain

$$\langle a \rangle_\alpha^{-3/2} \times \Delta\ell_\alpha \approx C_{AD} \frac{\pi}{4} \frac{m_\star}{M_\bullet} \sqrt{\frac{N_{\max}}{a_{\max}}} r_g^{-1} \langle \ell \rangle_\alpha^2. \quad (2.11)$$

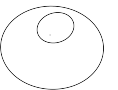
Here, the subscript α indicates the average over the unit sphere. We show in Fig. 2.5 the resulting amplitudes for the three codes and linear least-squares fit to the data according to equation (2.11), where we used the number of points in each bin as relative weights. The data from the simulations is consistent with the prediction $\Delta\ell \propto \langle \ell \rangle^2$, for each of the three codes. The resulting values of C_{AD} that we obtain from the fits are shown in the top left of Fig. 2.5.

2.4.5 Diffusion in angular momentum above and below the SB

In Section 2.4.4, we described the eccentricity oscillations that occur below the SB. However, in our simulations, not all orbits remain below the SB indefinitely. To illustrate this we show in Fig. 2.6 the cumulative distributions of ℓ/ℓ_{SB} , with ℓ_{SB} given by equation (2.1), at three time intervals $0 < t/\text{Myr} < 0.05$, $0.50 < t/\text{Myr} < 0.55$ and $0.95 < t/\text{Myr} < 1.0$. At the earliest time in the simulations the majority of orbits are below the SB ($\ell/\ell_{SB} < 1$). As time progresses the latter quantity gradually increases and by the end of the simulation nearly all orbits ($\gtrsim 90\%$) have diffused above the SB (i.e. $\ell \gg \ell_{SB}$). The deviations in the cumulative distributions of ℓ/ℓ_{SB} between the three codes appear to increase with time. This may be due to various reasons, including exponential divergence in the gravitational N -body problem and the increase in the amplitude of eccentricity oscillations above the SB (cf. Fig. 2.2), therefore reducing the number of data points for larger ℓ . Nevertheless, there does not appear to be a systematic difference between the distributions for the three codes.

In this section, we carry out a quantitative analysis of the angular-momentum diffusion. We obtained from the simulations the first-order ($n = 1$) and second-order ($n = 2$) diffusion

²The factor $(1/3)$ in MAMW11 equation (43b) should be replaced by $(1/2)$ (Merritt, private communication).



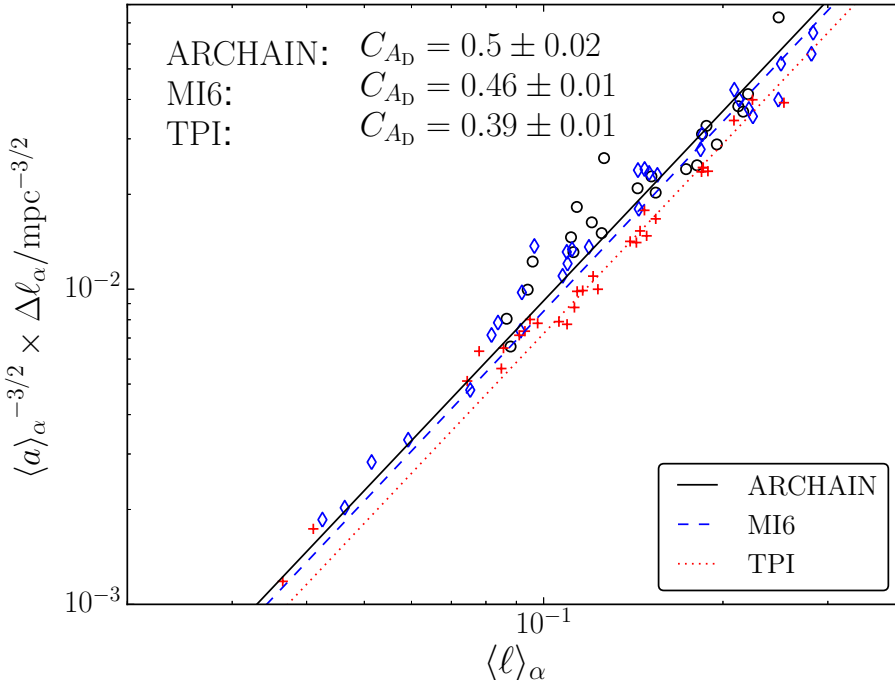


Figure 2.5: Amplitude of eccentricity oscillations binned in mean value of a and $\langle \ell \rangle$ as function of $\langle \ell \rangle$, averaged over all orientations α . Black circles: ARCHAIN; blue diamonds: MI6; red pluses: TPI. The solid, dashed and dotted lines show least-squares fit to the data for the three codes, respectively, weighted by the bin size corresponding to each data point. The fitted values of C_{A_D} are shown in the top left.

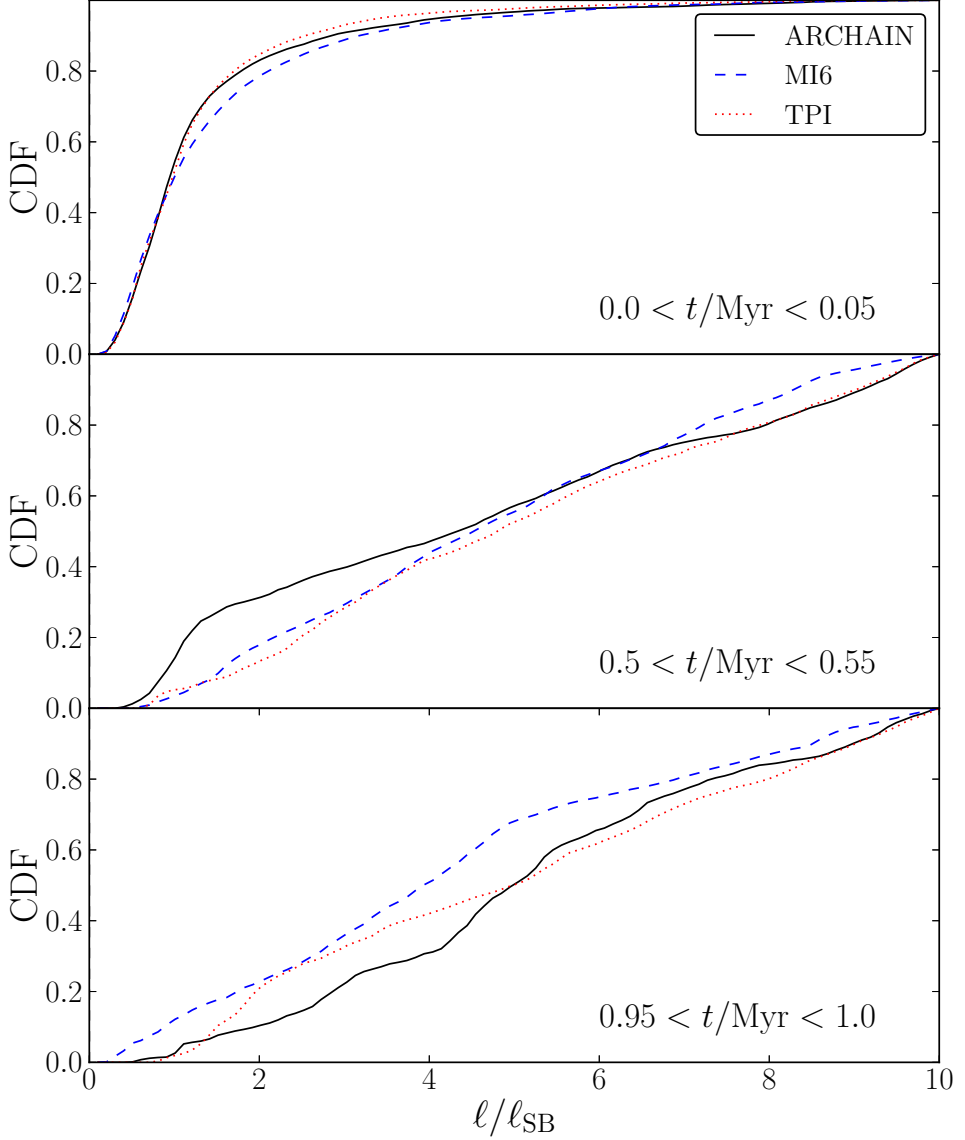
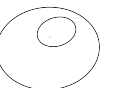


Figure 2.6: Cumulative distributions of the dimensionless angular momentum ℓ normalized to the angular momentum associated with the SB, ℓ_{SB} (cf. equation 2.1). Three time intervals are shown for the three codes. Black solid lines: ARCHAIN; blue dashed lines: MI6; red dotted lines: TPI.



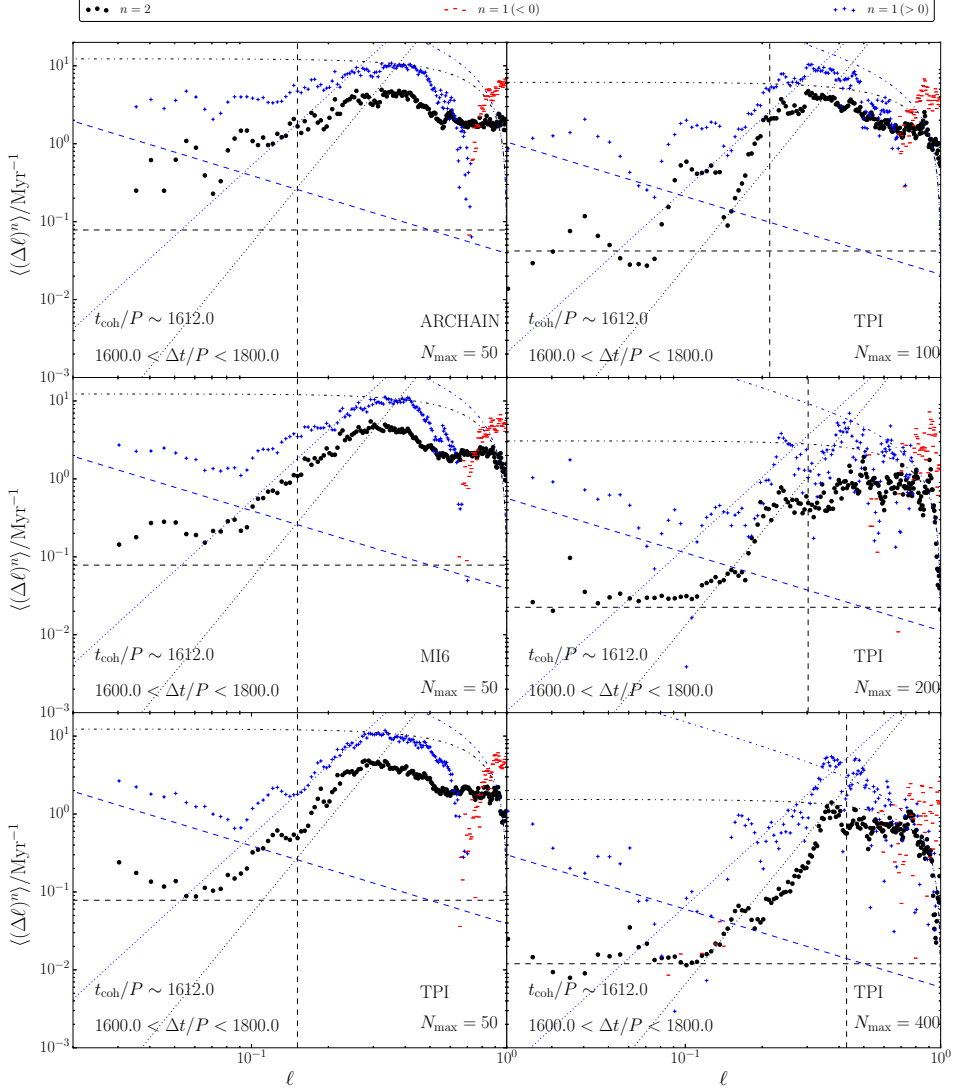


Figure 2.7: First-order and second-order diffusion coefficients as function of $\ell \equiv L/L_c$. Left-hand column: based on the $N_{\text{max}} = 50$ simulations, distinguishing between the three codes (initially $a = 2$ mpc and $\log_{10}(1 - e) = -2.0, -2.5$ or -3.0). Right-hand column: based on simulations with TPI with $N_{\text{max}} = 100, 200$ and 400 (initially $a = 2$ mpc and $\log_{10}(1 - e) = -2.5$). Positive (negative) first-order diffusion coefficients are shown in blue (red); second-order diffusion coefficients are shown in black. Minuses, plusses and bullets: quantities obtained from the simulations. Dashed lines: the predicted NRR diffusion coefficient in the limit $\ell \rightarrow 0$, equation (2.12). Black dot-dashed lines: the second-order incoherent RR prediction, equation (2.14), with $\beta_s = 1.6 \sqrt{1 - \ell^2}$ (Gürkan & Hopman 2007). The blue dot-dashed lines show an ad hoc relation for the first-order RR coefficient, equation (2.15). Dotted lines: predictions for $\ell \lesssim \ell_{\text{SB}}$ according to the model presented in Section 2.5.4. The vertical black dashed line shows the predicted value of ℓ at the SB, equation (2.1). In each panel the time lags shown are comparable to the coherence time (see the text).

coefficients, $\langle(\Delta\ell)^n\rangle$, describing changes in ℓ . Each diffusion coefficient was computed for a given initial value of ℓ and for a time interval, Δt , normalized to the orbital period P : $\tau = \Delta t/P$. These quantities were binned in linear bins of size 200 with $0 < \ell < 1$ and size 5 with $10^3 < \tau < 2 \times 10^3$, respectively. For each time t_i in the simulation we selected times $t_j > t_i$ with associated time lags $\tau_{ij} = (t_j - t_i)/P_i$ in the range $\tau < \tau_{ij} < \tau + \Delta\tau$ with $\Delta\tau = 10$. We rejected any time t_j if the absolute value of the change of the semimajor axis at time t_j relative to t_i , $|(a_i - a_j)/a_i|$, exceeded 0.2, or if the test star was captured or unbound at t_j . For the remaining t_j , we computed the corresponding change of orbital angular momentum (normalized to the angular momentum of a circular orbit L_c), $\Delta\ell_{ij} = \ell_j - \ell_i$. Subsequently, we computed the first-order diffusion coefficients from $\langle\Delta\ell\rangle = \text{mean}(\Delta\ell_{ij})/\text{mean}(\tau_{ij}P)$ and the second-order diffusion coefficients from $\langle(\Delta\ell)^2\rangle = \text{mean}[(\Delta\ell_{ij})^2]/\text{mean}(\tau_{ij}P)$, where the mean is taken over each bin of ℓ and τ .

In the method described above, the diffusion coefficients are functions of the time lag Δt . One expects that over some finite range in Δt , the results will not depend too strongly on Δt . According to van Kampen (1992), when evaluating diffusion coefficients in some quantity x , Δt must be ‘so small that x cannot change very much during Δt , but large enough for the Markov assumption to apply’. In our case, an additional condition applies: for $\ell < \ell_{\text{SB}}$, a lower limit on Δt is given by $t_{\text{GR}}(\ell) \propto \ell^2$, since we are interested in directed changes in the *mean* value of ℓ below the SB, averaged over the time-scale of the angular-momentum oscillations, which is $\sim t_{\text{GR}}$ (cf. Section 2.4.4).

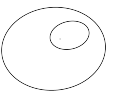
We argue in Section 2.5.4 that for $\ell = \ell_{\text{SB}}$, the characteristic time for ℓ to change by of order itself is the ‘coherence time’ $t_{\text{coh}}(a)$, defined as the time for a typical field-star orbit, of semimajor axis a , to change its orientation. We adopt $t_{\text{coh}}^{-1} = \langle t_{\text{MP}} \rangle^{-1} + \langle t_{\text{GR}} \rangle^{-1}$, where $\langle t_{\text{MP}} \rangle$ and $\langle t_{\text{GR}} \rangle$ are the field-star mass precession and the Schwarzschild precession time-scales averaged over a thermal distribution in eccentricity, respectively. As indicated in Fig. 2.1, the time-scale for changes in ℓ increases away from the SB, both towards higher and lower ℓ . Thus, in both the AR and RR angular-momentum regimes defined in that figure, we expect that setting $\Delta t \lesssim t_{\text{coh}}$ will ensure that ℓ ‘does not change very much’ during Δt .

The time-scale $t_{\text{GR}}(\ell)$ decreases rapidly as ℓ decreases from ℓ_{SB} ; clearly, for $0 < \ell < \ell_{\text{SB}}$, $t_{\text{GR}}(\ell)$ is maximal for $\ell = \ell_{\text{SB}}$. In our simulations $t_{\text{GR}}(\ell_{\text{SB}})$ is typically comparable to or smaller than t_{coh} . For example, for $a = 2$ mpc, $t_{\text{GR}}(\ell_{\text{SB}}) \approx 1.3 \times 10^3 P$, whereas $t_{\text{coh}} \approx 1.6 \times 10^3 P$. This demonstrates that, by choosing $\Delta t \sim t_{\text{coh}}$, we satisfy both lower and upper limits of Δt for $\ell < \ell_{\text{SB}}$.

These arguments aside, the validity of an assumed value of Δt can be checked by comparing the values of the diffusion coefficients derived for larger and smaller Δt . An example is given in Appendix 2.D.

Eilon et al. (2009) also carried out extensive N -body simulations to study the efficiency of RR in small- N systems. Their pioneering work differed from ours in two important respects: their integrations were Newtonian, i.e., the effects of relativistic precession were not included, and they did not investigate the L -dependence of the diffusion rate, choosing instead to focus on the time dependence of the mean angular momentum changes induced by the torques. We can, however, compare our results to theirs in the high- L regime where the effects of relativity are unimportant, as shown below.

We show in Fig. 2.7 our derived diffusion coefficients $\langle\Delta\ell\rangle$ (blue pluses and red minuses for positive and negative values, respectively) and $\langle(\Delta\ell)^2\rangle$ (black bullets) as function of ℓ . In the left-hand column results are shown for the three codes and the simulations with $N_{\text{max}} = 50$,



combining data from the test stars for the runs with initially $a = 2$ mpc and $\log_{10}(1 - e) = -2.0, -2.5$ and -3.0 . In each panel of Fig. 2.7, the value of ℓ associated with the SB, ℓ_{SB} (cf. equation 2.1), is indicated with the vertical black dashed line. The coherence times and the adopted time lag bins, expressed in units of orbital period, are indicated in the bottom left of each panel. There appear to be no systematic differences in the diffusion coefficients between the different codes shown in the first column of Fig. 2.7.

As mentioned in Section 2.1, the simulations with $N_{\text{max}} = 50$ field stars likely do not give a good description of the environment close to an SBH because the number of field stars within the initial volume of the simulation is too low (50, whereas $10^3 - 10^4$ would be more realistic). For this reason, we carried out additional simulations with TPI with larger numbers of field particles, i.e. $N_{\text{max}} = 100, 200$ and 400 (simulations with $N_{\text{max}} = 4800$ are discussed in Section 2.5). The field star mass m_* was adjusted to keep the enclosed stellar mass within any radius constant with respect to the $N_{\text{max}} = 50$ simulations. The adopted values are $m_* = 25, 12.5$ and $6.25 M_{\odot}$ for $N_{\text{max}} = 100, 200$ and 400 , respectively. The initial orbital elements of the test stars were $a = 2.0$ mpc and $\log_{10}(1 - e) = -2.5$. Other parameters were identical to those in the $N_{\text{max}} = 50$ simulations (cf. Section 2.4.1). The diffusion coefficients derived from these simulations with larger N_{max} are shown in the right-hand column of Fig. 2.7.

Some theoretical predictions exist for the dependence of the diffusion coefficients on ℓ , and we can compare these predictions with our results. We refer the reader to Fig. 2.1 which identifies the three regimes in angular momentum.

As discussed in Section 2.2, we expect that NRR will dominate diffusion in angular momentum in the limit $\ell \rightarrow 0$. Our argument was that – by definition – NRR is unaffected by coherence-time arguments, and hence that the rapid GR precession that occurs in this low- ℓ regime has no consequence for the rate of non-resonant diffusion in angular momentum.

The orbit-averaged, NRR diffusion coefficients in the limit $\ell \rightarrow 0$ for test masses near an SBH are (Cohn & Kulsrud 1978; Cohn 1979):

$$\langle \Delta \ell \rangle_{\text{NRR}} \rightarrow \frac{1}{4\ell} A(E); \quad (2.12a)$$

$$\langle (\Delta \ell)^2 \rangle_{\text{NRR}} \rightarrow \frac{1}{2} A(E). \quad (2.12b)$$

Here $A(E)$ is a function of orbital energy, or, equivalently, of semimajor axis. It is given by (e.g. appendix B of MAMW11)

$$A(E)^{-1} = \frac{C_{\text{NRR}}(\gamma)}{\log(\Lambda)} \left(\frac{M_{\bullet}}{M_*(a)} \right)^2 N_*(a) P(a). \quad (2.13)$$

Here, $C_{\text{NRR}}(\gamma)$ is a dimensionless quantity that depends on the field-star density slope γ . It can be evaluated using the procedure outlined in appendix B of MAMW11. Explicit expressions for $C_{\text{NRR}}(\gamma)$ as function of γ are included in Appendix 2.B. The value that applies to the simulations presented here is $C_{\text{NRR}}(2) = (9/7)\{1/[12 \log(2) - 1]\} \approx 0.18$. For the Coulomb logarithm Λ , we adopt $\Lambda = 2M_{\bullet}/m_*$ (MAMW11). The diffusion coefficients described by equation (2.12) are plotted in Fig. 2.7 with the dashed blue and black lines for the first-order and second-order coefficients, respectively.

For the simulations with $N_{\text{max}} = 50$, it can be seen in Fig. 2.7 that the first- and second-order diffusion coefficients gradually approach the NRR predictions for $\ell \ll \ell_{\text{SB}}$. As N_{max} is

increased, the correspondence between measured and predicted diffusion coefficients becomes quite good in this regime. This reinforces the hypothesis that NRR is indeed the mechanism that is primarily responsible for changes in ℓ as $\ell \rightarrow 0$.

The other limiting case is $\ell \gg \ell_{\text{SB}}$. In this high-angular-momentum regime, we expect that the dominant diffusion mechanism is (incoherent) RR (Merritt 2013a, p. 274). Only a limited set of predictions are available for the dependence of the RR diffusion coefficients on ℓ , and as far as we are aware, no attempt has ever been made to compute the first-order coefficient in the incoherent RR regime.

The second-order coefficient can be written in the form

$$\langle (\Delta\ell)^2 \rangle_{\text{RR}}^{-1} = \beta_s^{-2} \left[\frac{M_\bullet}{M_\star(a)} \right]^2 N_\star(a) \frac{P(a)^2}{t_{\text{coh}}}. \quad (2.14)$$

Here, t_{coh} is the ‘coherence time’ as introduced above, and β_s is a parameter describing the efficiency of RR in the coherent regime, i.e. for time intervals $\Delta t \lesssim t_{\text{coh}}$ during which ℓ increases approximately linearly with time. Our adopted coherence time is given by $t_{\text{coh}}^{-1} = \langle t_{\text{MP}} \rangle^{-1} + \langle t_{\text{GR}} \rangle^{-1}$, where $\langle t_{\text{MP}} \rangle$ and $\langle t_{\text{GR}} \rangle$ are the field-star mass precession and Schwarzschild precession time-scales for ω to change by π radians, averaged over a thermal distribution in eccentricity, respectively. These quantities are given explicitly by $\langle t_{\text{GR}} \rangle = (1/12)(a/r_g)P(a)$ and $\langle t_{\text{MP}} \rangle = (3/2)[M_\bullet/M_\star(a)]P(a)$ for $\gamma = 2$ (cf. equations 2.2 and 2.3).

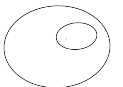
In Fig. 2.7, we show equation (2.14) with the black dot-dashed lines assuming $\beta_s = 1.6\sqrt{1-\ell^2}$, which is the Newtonian result obtained by Gürkan & Hopman (2007). For the $N_{\text{max}} = 50$ case, there are some systematic differences between the observed and predicted diffusion coefficients, which are similar in all three codes. The measured values are systematically smaller, and there is also a local minimum in $\langle (\Delta\ell)^2 \rangle$ at $\ell \approx 0.8$, which is not predicted. However, as N_{max} is increased, the local minimum gradually disappears, and the second-order coefficient is increasingly better described by the RR prediction. We can not claim to have a good explanation for the larger disagreement for smaller N_{max} , but one possibility may be the increased importance of NRR when N_{max} is small. We include in this category processes like strong encounters and multi-body effects, which, although not well described by equations like (2.12), are likewise unaffected by coherence-time arguments and which become increasingly important in stellar systems as N is decreased. We note that the simulations with even larger N_{max} (cf. Section 2.5.4) also show good agreement with theory in this regime.

We remark that the result $\beta_s = 1.6\sqrt{1-\ell^2}$ in the high- L regime can be compared to the work of Eilon et al. (2009). Eilon et al. (2009, cf. section 4.3 of the latter paper) determined a value of $\beta_{s,\text{EKA09}} = 1.05$ averaged over their simulations (cf. their table 1), in which a thermal distribution of eccentricities was assumed. Averaging $\beta_s = 1.6\sqrt{1-\ell^2} = 1.6e$ over a thermal eccentricity distribution one finds $\langle \beta_s \rangle \approx 1.07$, which is in excellent agreement with the result of Eilon et al. (2009).

As noted above, there does not appear to be any discussion in the literature about the expected form of the first-order RR diffusion coefficient. Fig. 2.7 plots the ad hoc expression

$$\langle \Delta\ell \rangle_{\text{RR}} \approx \ell^{-1} \langle (\Delta\ell)^2 \rangle_{\text{RR}}. \quad (2.15)$$

This expression must be very approximate; it is clear that it cannot be valid for $\ell \approx 1$ because the first-order diffusion coefficient is expected (and is observed) to become negative as $\ell \rightarrow 1$.



Nevertheless, as Fig. 2.7 suggests, it is a reasonable approximation for $\ell_{\text{SB}} \lesssim \ell \ll 1$ (in Fig. 2.7 equation 2.15 is plotted without modifying the normalization).

Finally, we consider the diffusion coefficients in the third of the three angular-momentum regimes defined in Fig. 2.1: $\ell \lesssim \ell_{\text{SB}}$, called ‘anomalous relaxation’ (AR) in that figure. Fig. 2.7 suggests that this regime becomes increasingly well defined in the simulations as N_{max} increases: a distinct ‘knee’ appears at $\ell \approx \ell_{\text{SB}}$, below which $\langle \Delta \ell \rangle$ and $\langle (\Delta \ell)^2 \rangle$ both drop rapidly towards smaller ℓ , before flattening off in the NRR regime. We interpret this behaviour as a manifestation of the rapid quenching of RR below the SB; indeed the location of this knee might be taken as an empirical definition of the location of the barrier. (The location of the knee is consistent with the value of ℓ_{SB} as predicted by equation (2.1), even though the nuclear model in Fig. 2.7 is rather different than the one considered in MAMW11.) As shown below, the ‘knee’ becomes even better defined in simulations with still larger values of N_{max} ; we will argue that this is due to a greater separation between the AR and NRR regimes.

A mechanism that would drive angular-momentum diffusion in the $\ell \lesssim \ell_{\text{SB}}$ region was discussed in MAMW11. Here we note – following the discussion in that paper – that diffusion in this regime is not expected to be well described either in terms of resonant or non-resonant relaxation. In Section 2.5.4, we return to the behaviour of diffusion in angular momentum for $\ell \lesssim \ell_{\text{SB}}$ and present a theoretical model for diffusion in this regime.

2.5 Dynamical evolution of the S-stars

2.5.1 Initial conditions

We have demonstrated the validity of the results of TPI in Section 2.4 using comparisons to more accurate, but slower, N -body codes. Here, we proceed with simulations of the S-star cluster in which the number of field particles is larger by a factor of $\sim 10^2$; such simulations are currently not feasible with the other N -body codes discussed in Section 2.4.

The S-star cluster consists of main-sequence (MS) B-type stars at projected distances $r_p \lesssim 0.8 \text{ arcsec} \approx 32 \text{ mpc}$ (Genzel et al. 2003b; Eisenhauer et al. 2005; Ghez et al. 2008; Gillessen et al. 2009) from the central SBH, Sgr A* (assuming a distance to the GC of 8.3 kpc; Gillessen et al. 2009). The strong tidal field of the SBH at these radii makes it unlikely that the S-stars formed in situ (Morris 1993), hence various formation scenarios have been proposed in which the S-stars formed elsewhere and migrated to their current locations (see Alexander 2005 and Genzel et al. 2010 for reviews). Antonini & Merritt (2013, hereafter AM13) used Monte Carlo simulations to show that binary disruption best matches the observed eccentricity distribution of the S-stars. In this process a stellar binary is tidally disrupted by the SBH, unbinding one of the stars from the SBH and leaving the other star in a tight and highly eccentric ($0.93 \lesssim e \lesssim 0.99$) orbit around the SBH (Hills 1988). For the observed semimajor axes of the S-stars the high initial eccentricities predicted by this process imply that some of the S-stars were deposited below the SB (cf. fig. 1 of AM13). The number of stars that were deposited below the SB in this case depends on the assumed distribution of the field stars near the SBH. The diffusion processes discussed in Sections 2.2 and 2.4 could therefore be important for the dynamical evolution of some of the S-stars in the first few Myr after being deposited in the GC.

In the Monte Carlo simulations of AM13, the orbital evolution of the S-stars was described using equations of motion derived from an orbit-averaged Hamiltonian, and two-body relax-

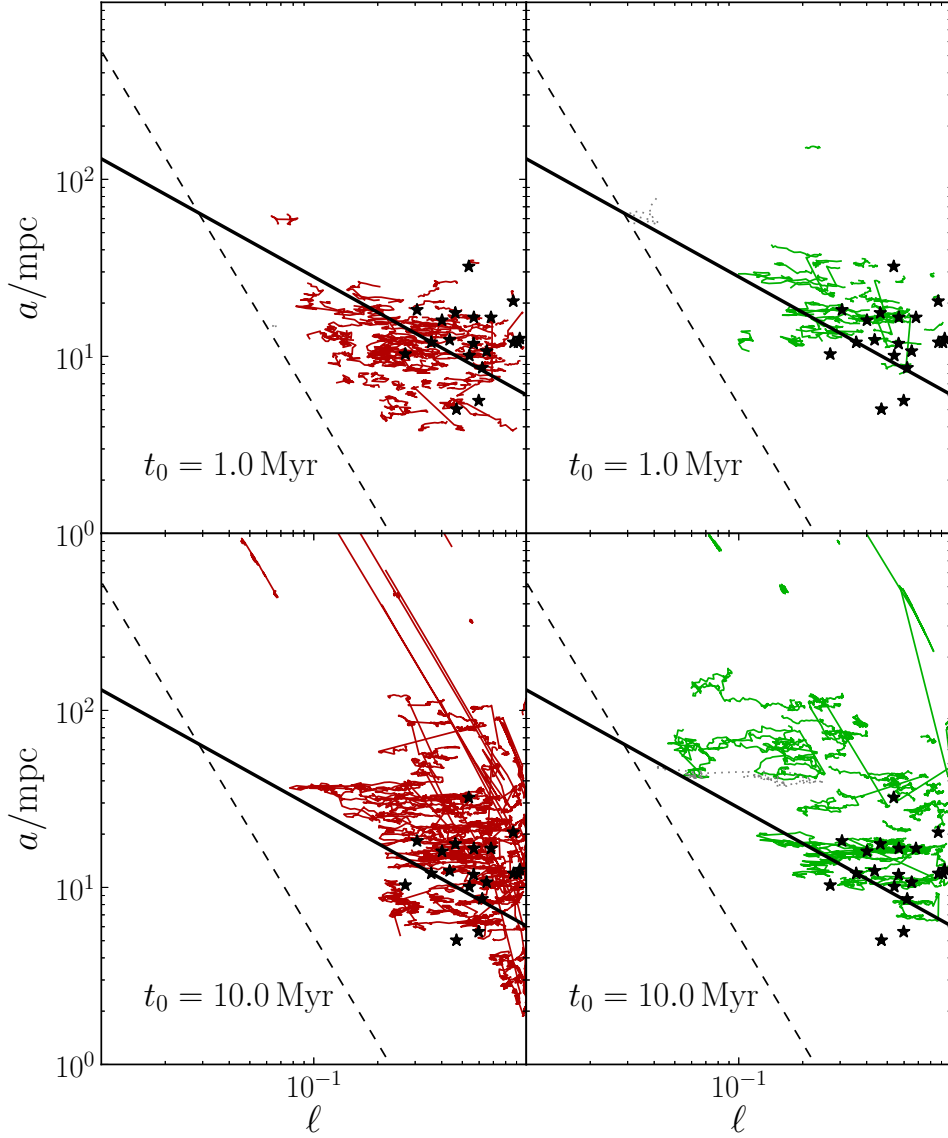
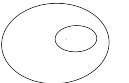


Figure 2.8: Planes of semimajor axis versus dimensionless angular momentum ℓ for all 10 realizations of the 19 S-stars in our simulations with TPI. Tracks are shown for two time ranges t_0 , with $0.8t_0 < t < t_0$. Red (green) tracks apply to orbits that are initially below (above) the SB; dotted grey tracks apply to stars that become captured or unbound during the interval shown. The predicted position of the SB (equation 2.16) is indicated with the black solid line. The black dashed line shows the pericentre distance corresponding to the tidal disruption radius. Black stars indicate the orbital elements of the observed S-stars with $a < 32.2$ mpc from Gillessen et al. (2009).



ation effects were not taken into account. Here, we do take into account diffusion driven by two-body relaxation using TPI. (As discussed in more detail below, we confirm that the neglect of NRR by those authors was a reasonable approximation, at least as far as the eccentricity distribution is concerned.) We adopted from AM13 a field star number distribution $N_\star(a) = N_{\max}(a/a_{\max})^{3-\gamma}$ with $N_{\max} = 4.8 \times 10^3$, $a_{\max} = 0.2 \text{ pc}$ and $\gamma = 2$; the field star mass was set to $m_\star = 10 M_\odot$. This distribution is consistent with steady-state models of the GC of a cusp of stellar remnants (Hopman & Alexander 2006b). We simulated 19 S-stars, adopting the semimajor axes with $a < 32.2 \text{ mpc}$ from the sample of S-stars for which orbital fits were obtained by Gillessen et al. (2009). In our simulations the S-stars were treated as test stars; for each S-star there were 10 random realizations, each with an initial eccentricity sampled from a thermal distribution with $0.93 < e < 0.99$ and a random orientation, consistent with binary disruption. The probability for $e > e_{\text{SB}}$ in this model is ≈ 0.72 . The capture radius was set to an approximation of the tidal disruption radius, $r_{\text{capt}} = 2R(M_\bullet/m)^{1/3}$, where $R = 8 R_\odot$ and $m = 10 M_\odot$ (Antonini et al. 2011). We included only 1PN terms in the simulations and therefore we assumed a non-spinning SBH. The integration time was constrained by computational limitations and was set to 20 Myr.

In these simulations, equation (2.1) predicts:

$$\begin{aligned} \ell_{\text{SB}}(a) &= \frac{r_g}{a} \frac{M_\bullet}{M_\star(a)} \sqrt{N_\star(a)} \\ &= \ell_{\text{SB},10} \left(\frac{a}{10 \text{ mpc}} \right)^{-3/2}, \\ \ell_{\text{SB},10} &\approx 0.49. \end{aligned} \tag{2.16}$$

2.5.2 Orbital evolution

We show in Fig. 2.8 the $(a, 1-e)$ -plane for all 10 realizations of the 19 S-stars in our simulations for $t < 10 \text{ Myr}$. Tracks are shown for two times t_0 in the simulations. Red (green) tracks apply to orbits that are initially below (above) the SB; dotted grey tracks apply to stars that become captured or unbound during the interval shown (cf. Section 2.5.3). A fraction ~ 0.72 of the stars start below the SB (red solid line in Fig. 2.8). The orbits of the majority of these rapidly diffuse to larger semimajor axis and/or smaller eccentricity: by 10 Myr, most of them have evolved to locations above the SB. The orbits that are initially above the SB, on the other hand, tend to remain in this region. Note that some penetration to regions below the SB does occur, however, and that some stars remain below the SB even after 10 Myr. In what follows, we discuss this evolution in more detail.

In Fig. 2.9 we show the eccentricity evolution for six realizations of S-stars in the simulations with initial eccentricity $e_0 > e_{\text{SB}}$. By 10 Myr these orbits have diffused to locations above the SB. At several instances the orbit, after having diffused to $e < e_{\text{SB}}$, becomes more eccentric again and reaches $e \approx e_{\text{SB}}$. The orbit is then ‘reflected’, however, to lower eccentricity. This behaviour is consistent with that seen in the N -body simulations of MAMW11.

It is significant that the relation proposed by MAMW11 for the location of the SB, and which is plotted as the red dashed line in Fig. 2.9, appears to predict remarkably well the value of the eccentricity at which RR ‘turns on’ in these simulations. This, in spite of the fact that the number of stars in the new simulations is a factor $\sim 10^2$ higher than in those of MAMW11.

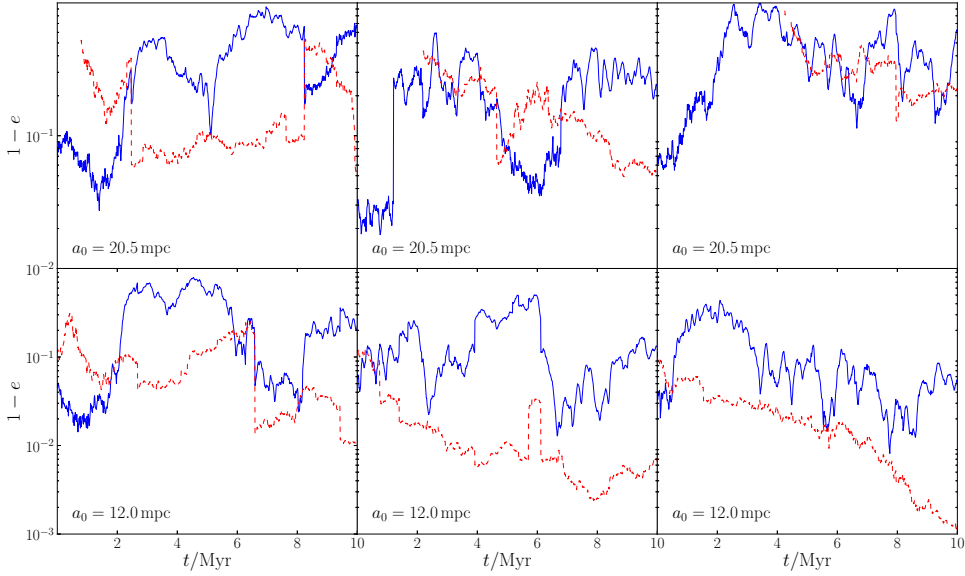
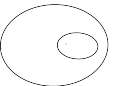


Figure 2.9: The eccentricity evolution (blue solid lines) for six realizations of the S-stars in the simulations with initial eccentricity $e_0 > e_{\text{SB}}$. Red dashed lines show the predicted value of e_{SB} (cf. equation 2.16). Tracks are selected for which no capture or unbinding event occurs within the time interval shown. The initial semimajor axis is indicated in the bottom left of each panel.

We interpret this success as confirming, to a greater degree than was possible in MAMW11, the general validity of the relation (2.1).

We show in Fig. 2.10 the evolution of the cumulative eccentricity distribution for all realizations of the S-stars in our simulations. This distribution evolves rapidly from a near δ -function at $e \sim 1$ that reflects the initial conditions, to a much more uniform distribution. The distribution does not appear to converge to a ‘thermal’ form, $N(e) = e^2$, but on average remains more eccentric. Interestingly, the distribution appears to converge to a form that is closer to the observed, ‘super-thermal’ distribution of the S-stars (Gillessen et al. 2009), shown in Fig. 2.10 with the blue solid line.

To investigate this apparent correspondence with observations more quantitatively, we fitted the cumulative eccentricity distribution in our simulations to a power law, $N(e) = e^p$, and we show the time evolution of p in Fig. 2.11 with solid and dashed black lines. There is an initial rapid decrease of p from ~ 22 to ~ 7 over the course of ~ 1 Myr. The subsequent evolution is slower, with p decreasing to ~ 3 after 4 Myr. The form of $p(t)$ for $1 \lesssim t/\text{Myr} \lesssim 7$ is well fitted by a decaying exponential function, $p_{\text{fit}}(t) = c_0 + c_1 \exp(-c_2 t)$; we find best-fitting values $c_0 \approx 2.11$, $c_1 \approx 6.27$ and $c_2 \approx 0.34 \text{ Myr}^{-1}$ (the fitted curve is shown with the blue dashed line in Fig. 2.11). After ~ 7 Myr, the detailed evolution of $p(t)$ deviates slightly from a decaying exponential function. The overall evolution is still consistent with a decaying exponential, however. Interestingly, in our simulations $p(t)$ appears to oscillate roughly between $p \approx 2$, consistent with a thermal distribution, and $p \approx 2.6$, consistent with observations ($p_{\text{obs}} = 2.6 \pm 0.9$).



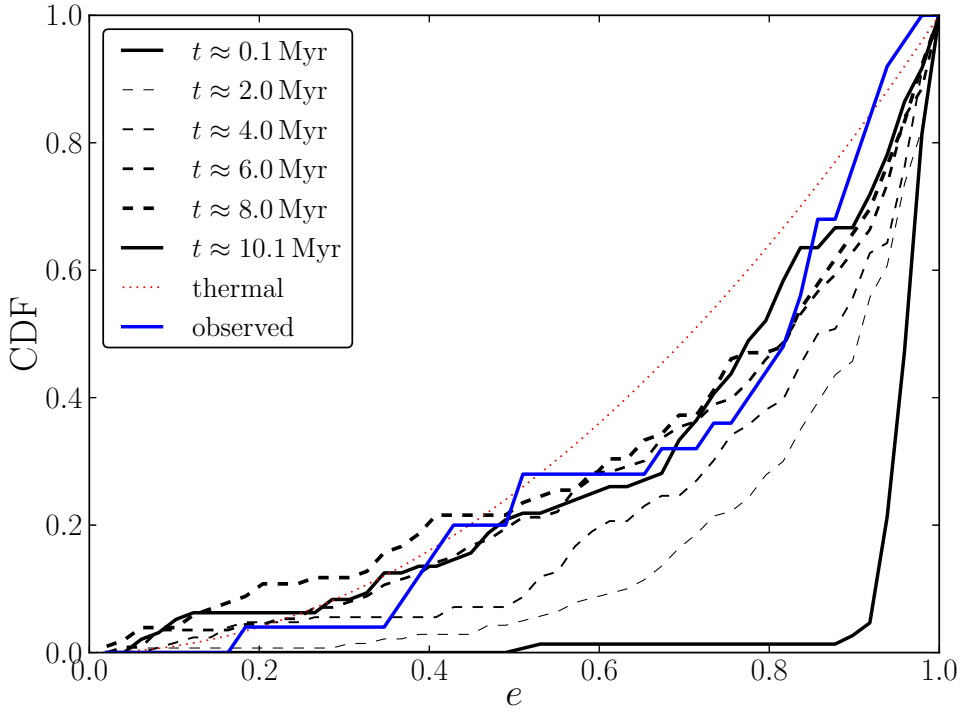


Figure 2.10: The cumulative eccentricity distribution of all realizations of the S-stars in our simulations at various times between $t \approx 0$ and 10 Myr, assuming burst formation. The initial and final distributions are shown with black solid lines. Intermediate times are shown with black dashed lines; the thickness increases with time. The blue solid line shows the observed distribution of the S-stars (Gillessen et al. 2009). The red dotted line shows a thermal distribution $N(e) = e^2$.

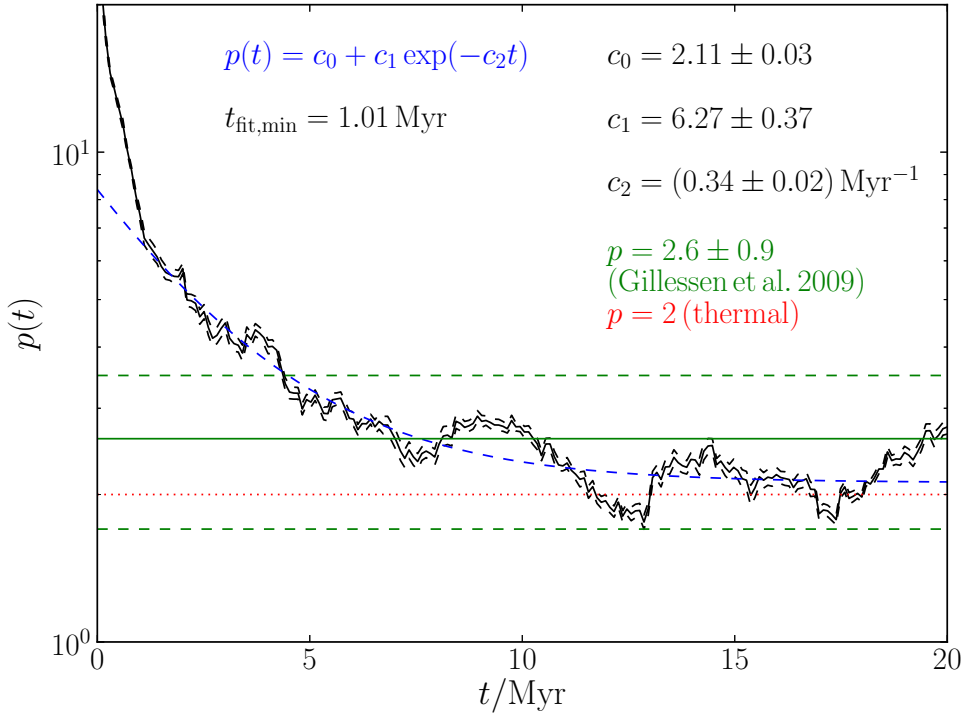
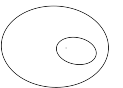


Figure 2.11: The evolution of the slope p with time for all realizations of the S-stars in our simulations (black solid line), in the case of formation in a burst. We exclude four S-stars, for which the computation had not advanced to 20 Myr by the time of writing. The uncertainty in p is indicated with black dashed lines. The blue dashed line shows a least-squares fit of the form $p_{\text{fit}}(t) = c_0 + c_1 \exp(-c_2 t)$; we exclude data for $t < 1.01 \text{ Myr}$. The green solid and dashed lines indicate the observed value $p_{\text{obs}} = 2.6 \pm 0.9$ (Gillessen et al. 2009). The red dashed line shows $p = 2$ (thermal distribution).



Gillessen et al. 2009).

In the results presented above it was assumed that all 19 S-stars are deposited in the GC in a single burst at $t = 0$. We used these results as a template to estimate the evolution of $p(t)$ in the case of continuous formation of S-stars. The details are presented in Appendix 2.C. The effect of continuous formation is to slow the evolution of p as function of time. We find that in the case of continuous formation the time for p to decrease to $p = 2.6$ is increased by a factor of ~ 3.6 from ~ 7 to ~ 25 Myr. We discuss implications of the evolution of the eccentricity distribution in Section 2.7.1.

2.5.3 Tidally disrupted and ejected stars

We show in the top panel of Fig. 2.12 the cumulative fraction of S-stars that are tidally disrupted, i.e. the stars that at some time in the simulation approach the SBH within the assumed tidal disruption radius $r_{\text{capt}} = 2R (M_{\bullet}/m)^{1/3}$ with $R = 8 R_{\odot}$ and $m = 10 M_{\odot}$. The majority of disruptions occurs at $t < 10$ Myr: initially the orbits are highly eccentric, making stars susceptible to disruption. As the eccentricity decreases and the orbits reach the SB the probability for capture decreases. This is borne out by the bottom panel of Fig. 2.12 in which we show orbital tracks in the (a, ℓ) -plane prior to disruption. Most of the orbits are close to the disruption boundary prior to disruption and most of the latter orbits are below the SB. We note that the eccentricity oscillations described in Section 2.4 potentially enhance disruptions because during the oscillations the eccentricity can reach a higher value than the mean eccentricity. Only few (2 out of 19) disruptions occur above the SB and with relatively high angular momentum ($\ell > 10^{-1}$), in which case a strong two-body encounter is required to produce the required small pericentre distance (i.e. an interaction typically associated with the full loss-cone). The cumulative fraction of disrupted stars is ~ 0.10 after 20 Myr, which is an order of magnitude larger than the fraction of $\lesssim 0.01$ found by AM13. This may suggest that NRR, which was not taken into account in the calculations of AM13, is important for determining the rate of tidal disruptions.

Furthermore we show in the top panel of Fig. 2.13 the cumulative fraction of S-stars that become unbound from the SBH (i.e. stars with orbital energy $E > 0$). Unlike the fraction of tidally disrupted stars, the fraction of unbound stars continues to increase steadily after $t \approx 10$ Myr. This likely reflects the property that strong two-body encounters leading to ejection can in principle occur at any eccentricity and semimajor axis, whereas two-body encounters leading to tidal disruption are more likely if the eccentricity is high, in which case a small perturbation to the orbit is required for disruption. After 20 Myr the cumulative fraction of unbound stars is ≈ 0.14 . The distribution of the escape velocity $v_{\text{esc}} = \sqrt{2E}$ from the SBH (not taking into account deceleration from the Galactic bulge) is plotted in the bottom panel of Fig. 2.13. This distribution is peaked near $v_{\text{esc}} \sim 10^3$ km/s, which is comparable to the escape velocities of hypervelocity stars (Hills 1988).

2.5.4 Diffusion coefficients

Diffusion coefficients in the S-star simulations were computed using the same technique as in Section 2.4.5. In addition to binning the data with respect to the initial value of ℓ and the time lag τ , here data were also binned with respect to semimajor axis. We show in Fig. 2.14 the resulting first-order and second-order diffusion coefficients for all realizations of the S-stars

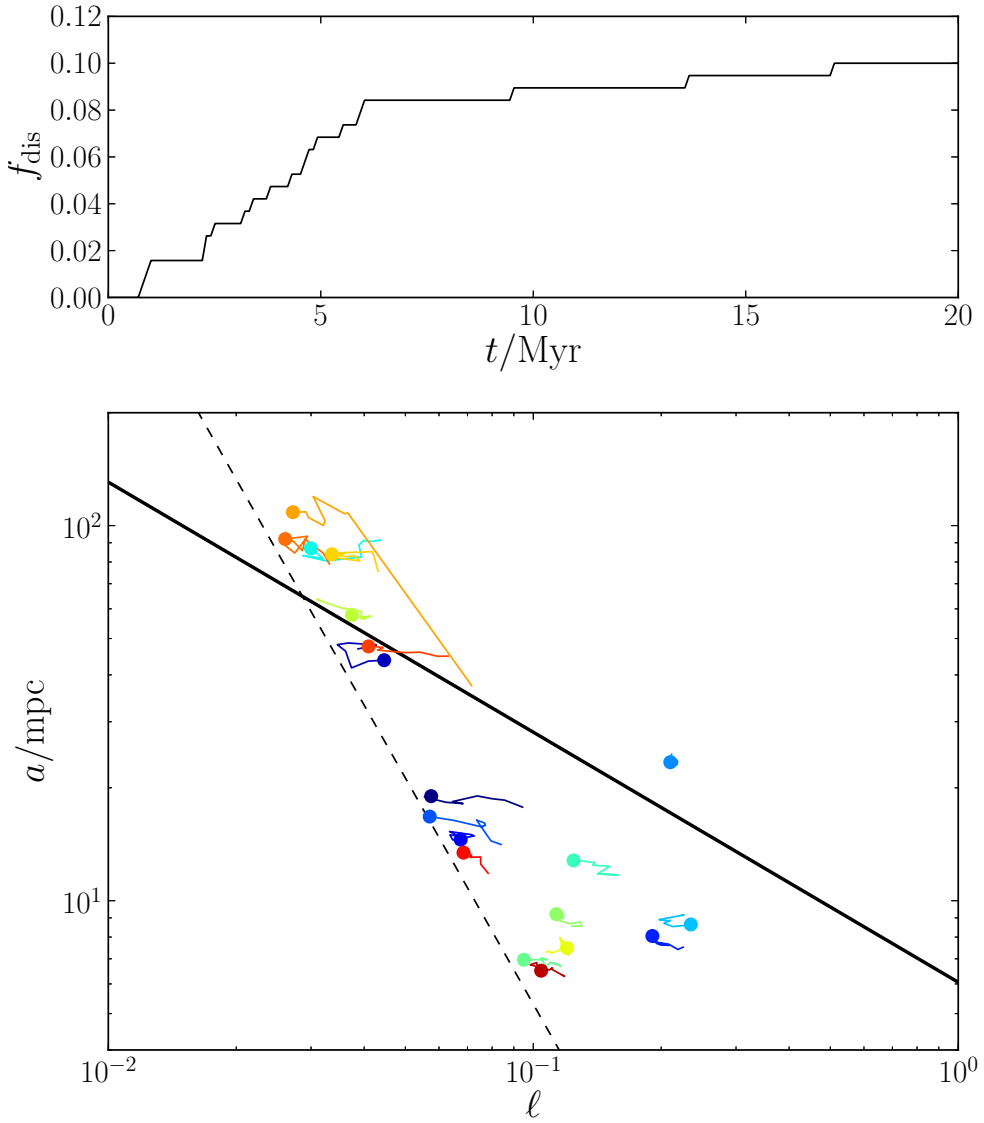
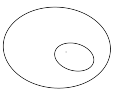


Figure 2.12: Top: the cumulative fraction of tidally disrupted S-stars (for all 10 realizations) as function of time. Bottom: orbital tracks prior to disruption. The orbital elements prior to the disruption event (determined at apocentre) are shown with bullets; in addition, tracks of 10 orbital periods prior to disruption are shown. The black solid line shows the SB according to equation (2.16).



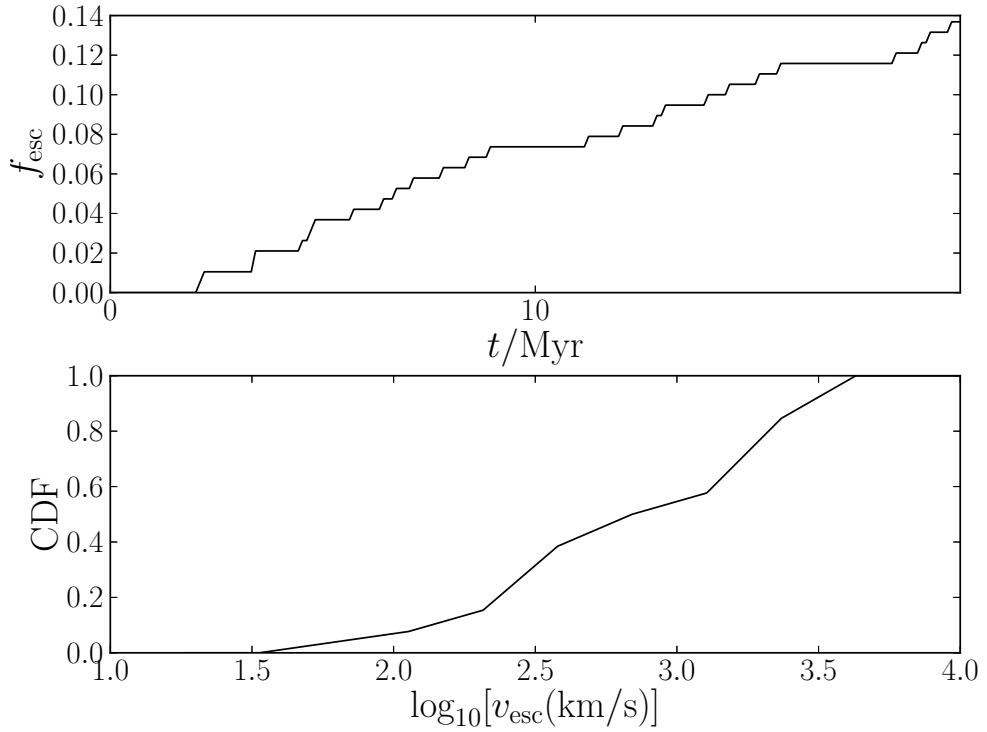


Figure 2.13: Top: the cumulative fraction of unbound S-stars (for all 10 realizations) as function of time. Bottom: the cumulative distribution of the escape velocity $v_{\text{esc}} = \sqrt{2E}$ from the SBH for the unbound stars, not taking into account the deceleration from the Galactic bulge.

in our simulations for six ranges of the semimajor axis. As in Fig. 2.7, the time lags shown in Fig. 2.14 are chosen such that the coherence time lies within the time lag bin. We note that by setting the time lag to values that are substantially longer, the diffusion coefficient plots tend to change in appearance. This is illustrated and explained in Appendix 2.D.

In the regime $\ell \gg \ell_{\text{SB}}$ the second-order diffusion coefficients from our simulations are consistent with the RR prediction, equation (2.14), with $\beta_s = 1.6 \sqrt{1 - \ell^2}$ (Gürkan & Hopman 2007). As noted in Section 2.4.5, this agreement is increasingly good with increasing N_{max} , a trend that continues here. Furthermore, the ‘knee’ feature of the diffusion coefficients near $\ell \approx \ell_{\text{SB}}$, which was observed in Section 2.4.5 as N_{max} was increased, is also clearly present in Fig. 2.14.

As noted above, the form of the diffusion coefficients in the $\ell \lesssim \ell_{\text{SB}}$ regime (‘anomalous relaxation’, AR) is not well understood theoretically. The rather abrupt decrease in the measured diffusion coefficients as ℓ decreases past $\sim \ell_{\text{SB}}$ is expected, at least qualitatively, since the SB is defined as the value of ℓ for which the rapid GR precession quenches the effects of the $\sqrt{N}$ torques. We find from our simulations that the dependence of the diffusion times,

$$T_n \equiv \left| \frac{\langle (\Delta \ell)^n \rangle}{\ell^n} \right|^{-1}, \quad n = \{1, 2\}, \quad (2.17)$$

on ℓ in this regime is often well fit by a relation of the form

$$T_{1,2}(a, \ell) = \text{constant}(a) \times \ell^{-2}, \quad \ell \lesssim \ell_{\text{SB}}. \quad (2.18)$$

This is the dependence that was assumed in making the lower panel of Fig. 2.1.

In fact, an ℓ^{-2} dependence below the barrier is predicted by the simple Hamiltonian model described in section VB of MAMW11, in which a random walk in ℓ results from assuming sudden, random changes in the direction of the $\sqrt{N}$ torquing potential each $\sim t_{\text{coh}}$. We briefly summarize here the results of an analytic calculation based on that model (Merriitt 2013, unpublished).

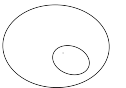
In the small- ℓ limit, the averaged Hamiltonian of MAMW11 predicts, for times $\Delta t \lesssim t_{\text{coh}}$,

$$\ell^{-1}(\omega) = \frac{1}{2\ell_1\ell_2} [(\ell_2 - \ell_1) \sin(\omega) + (\ell_1 + \ell_2)] \quad (2.19a)$$

$$= A_D [\sin(\omega) + h]. \quad (2.19b)$$

Here, $\{\ell_1, \ell_2\}$ are the extreme values of ℓ during a GR precession cycle, $\ell_{\text{av}} = (1/2)(\ell_1 + \ell_2)$, and $h = -H/A_D = (\ell_1^{-1} + \ell_2^{-1})/(2A_D) \approx 1/(A_D \ell_{\text{av}})$ is a normalized, averaged (secular) Hamiltonian H . (We have set $\sin i = \pi/2$ in equation 41 of MAMW11, i.e., the torquing potential is assumed to be aligned with the x axis.) Equation (2.19) describes changes in ℓ due to the $\sqrt{N}$ torques as the orbit precesses, at a (slightly) non-constant rate, due to GR. As noted above, the amplitude of the ℓ -oscillations in this regime scales as $\sim \ell_{\text{av}}^2$. These oscillations, by themselves, do not imply any directed evolution in ℓ_{av} . But if the direction of the torquing potential is suddenly changed, after a time $\sim t_{\text{coh}}$, the orbit will have been given a new value of h and correspondingly different values of ℓ_1 and ℓ_2 . Assuming that the changes in the direction of the torquing potential each t_{coh} are random, one finds for the first- and second-order diffusion coefficients of h in this model

$$\langle \Delta h \rangle \approx -\frac{1}{t_{\text{coh}}} \frac{1}{h}, \quad \langle (\Delta h)^2 \rangle \approx \frac{1}{t_{\text{coh}}}. \quad (2.20)$$



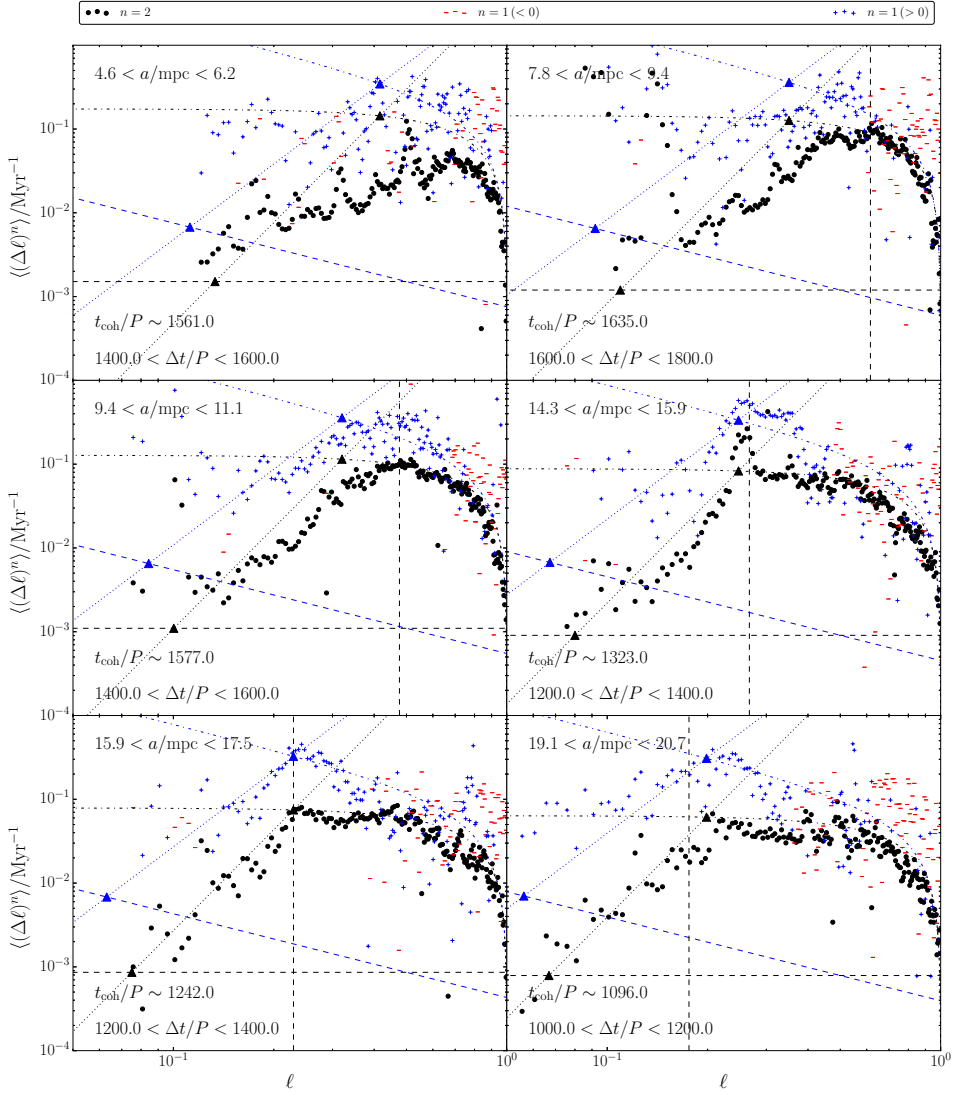


Figure 2.14: First-order and second-order diffusion coefficients as function of $\ell \equiv L/L_c$ obtained from the S-star simulations. Positive (negative) first-order diffusion coefficients are shown in blue (red); second-order diffusion coefficients are shown in black. Minuses, pluses and bullets: quantities obtained from the simulations. Dashed lines: the predicted NRR diffusion coefficient in the limit $\ell \rightarrow 0$, equation (2.12). Black dot-dashed lines: the second-order incoherent RR prediction, equation (2.14), with $\beta_s = 1.6 \sqrt{1 - \ell^2}$ (Gürkan & Hopman 2007). The blue dot-dashed lines show an ad hoc relation for the first-order RR coefficient, equation (2.15). Dotted lines: predictions according to the model presented in Section 2.5.4; we have set $C_{A_D} = 0.5$ and $C_1 = C_2 = 2.6$. The vertical black dashed line shows the predicted value of ℓ at the SB, equation (2.16). In each panel the time lags shown are comparable to the coherence time (see the text). The triangles indicate the quantities $\ell_{a,n}$ and $\ell_{b,n} > \ell_{a,n}$ (cf. equation 2.24); blue: $n = 1$; black: $n = 2$.

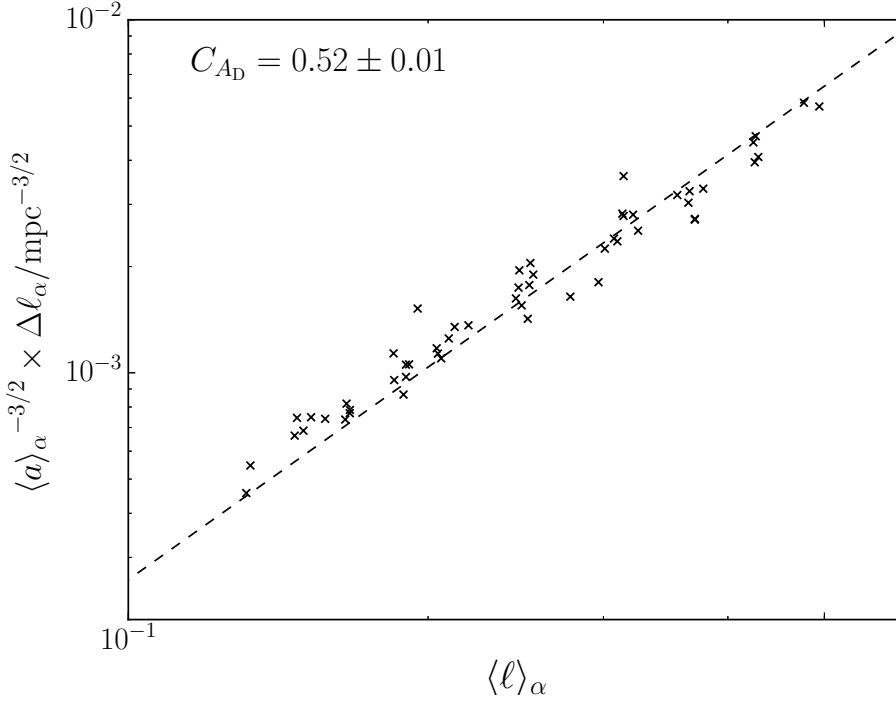


Figure 2.15: Amplitude of eccentricity oscillations for the S-star simulations, binned in the mean value of a and $\langle \ell \rangle$ as function of $\langle \ell \rangle$, and averaged over all orientations α . The best-fitting curve is shown with the black dashed line; the fitted value of C_{A_D} is shown in the top left.

The corresponding time scales are

$$\left| \frac{\langle \Delta h \rangle}{h} \right|^{-1} \approx \left| \frac{\langle (\Delta h)^2 \rangle}{h^2} \right|^{-1} \approx h^2 t_{\text{coh}} \approx \frac{t_{\text{coh}}}{A_D^2 \ell_{\text{av}}^2}, \quad (2.21)$$

consistent with the $\sim \ell^{-2}$ dependence observed in the simulations.

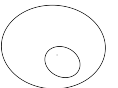
Accordingly, we suggest the following functional forms for the diffusion coefficients in the AR regime:

$$\langle \Delta \ell \rangle \approx \frac{C_1}{\tau} \ell^3; \quad (2.22a)$$

$$\langle (\Delta \ell)^2 \rangle \approx \frac{C_2}{\tau} \ell^4, \quad (2.22b)$$

with $\tau = t_{\text{coh}}/A_D^2$; furthermore, if the simple model presented above is valid, we expect $C_1 \approx C_2 = \mathcal{O}(10^0)$.

The quantities in equation (2.22) depend on the parameter A_D and the latter contains the fit parameter C_{A_D} (cf. equation 2.10). We used the same technique based on the amplitude



of the eccentricity oscillations as in Section 2.4.4 to determine this parameter for the S-star simulations; the results are shown in Fig. 2.15. Based on this result we adopt $C_{A_D} = 0.5$ and we plot the predicted diffusion coefficients, equation (2.22), in Fig. 2.14 with the dotted lines. We find best agreement with the data for $C_1 \approx C_2 \approx 2.6$. For reference we have also included these predictions for the simulations, that were discussed in Section 2.4, in Fig. 2.7.

Based on the results presented in Fig. 2.14, we can approximate the first and- second-order diffusion coefficients as piecewise-continuous functions of ℓ ,

$$\langle (\Delta\ell)^n \rangle (\ell) \approx \begin{cases} \langle (\Delta\ell)^n \rangle_{\text{NRR}} (\ell), & \ell < \ell_{a,n}; \\ \langle (\Delta\ell)^n \rangle_{\text{AR}} (\ell), & \ell_{a,n} \leq \ell \leq \ell_{b,n}; \\ \langle (\Delta\ell)^n \rangle_{\text{RR}} (\ell), & \ell > \ell_{b,n}. \end{cases} \quad (2.23)$$

Here, $\langle (\Delta\ell)^n \rangle_{\text{NRR}}$, $\langle (\Delta\ell)^n \rangle_{\text{AR}}$, $\langle \Delta\ell \rangle_{\text{RR}}$ and $\langle (\Delta\ell)^2 \rangle_{\text{RR}}$ are given explicitly by equations (2.12), (2.22), (2.15) and (2.14), respectively (in the latter equation we adopt $\beta_s = \alpha_s \sqrt{1 - \ell^2}$ with $\alpha_s = 1.6$). We emphasize that equation (2.15) is ad hoc and theoretically not well motivated, as discussed in Section 2.4.5. Moreover, it fails to describe the simulations for $\ell \gg \ell_{b,1}$. In Section 2.6, we present a modified (but still not theoretically motivated) analytic prescription for $\langle \Delta\ell \rangle_{\text{RR}}$ that better describes the data for $\ell \gg \ell_{b,1}$.

The quantities $\ell_{a,n}$ and $\ell_{b,n}$ are defined such that $\langle (\Delta\ell)^n \rangle (\ell)$ is a continuous function of ℓ . From equations (2.12), (2.22), (2.15) and (2.14) it follows that

$$\ell_{a,n} = \left[\frac{n \log(\Lambda)}{C_{\text{NRR}}(\gamma) C_n C_{A_D}^2} \left(\frac{r_g}{a} \right)^2 \frac{t_{\text{coh}}(a)}{P(a)} \right]^{1/4}; \quad (2.24a)$$

$$\ell_{b,n} = \frac{1}{\sqrt{2}} \left[-C_{b,n} + (C_{b,n}^2 + 4C_{b,n})^{1/2} \right]^{1/2}; \quad (2.24b)$$

$$C_{b,n} \equiv \frac{4\alpha_s^2}{C_n C_{A_D}^2} \left(\frac{r_g}{a} \right)^2 \left[\frac{t_{\text{coh}}(a)}{P(a)} \right]^2. \quad (2.24c)$$

We note that in the nuclear models considered here, $C_{b,n} \ll 1$, hence $\ell_{b,n} \approx C_{b,n}^{1/4}$. In Fig. 2.14 $\ell_{a,n}$ and $\ell_{b,n} > \ell_{a,n}$ are indicated with the two blue (black) triangles for $n = 1$ ($n = 2$). We also note that if $C_1 = C_2$, which we observe is approximately the case in our simulations, then $\ell_{a,2} = 2^{1/4} \ell_{a,1}$ and $\ell_{b,1} = \ell_{b,2}$.

Using equation (2.23), it is possible to estimate the time $\Delta t(\ell_1 \rightarrow \ell_2)$ to diffuse in angular momentum for any specified interval in ℓ . We are most interested here in the time $\Delta t(\ell_0 \rightarrow \ell_{\text{SB}})$ to diffuse from an initial value $\ell_0 < \ell_{\text{SB}}$ to ℓ_{SB} . Assuming – as is appropriate for these nuclear models – that the diffusion time is dominated by the AR regime, equation (2.22) implies

$$\frac{d\ell}{dt} \approx \langle \Delta\ell \rangle \approx \frac{C_1 \ell^3}{\tau} \Rightarrow \quad (2.25a)$$

$$\Delta t(\ell_0 \rightarrow \ell_{\text{SB}}) \approx \frac{\tau}{2C_1} (\ell_0^{-2} - \ell_{\text{SB}}^{-2}) \quad (2.25b)$$

$$\approx \frac{\tau}{2C_1 \ell_0^2}. \quad (2.25c)$$

The last step applies if $\ell_0 \ll \ell_{\text{SB}}$. We tested equation (2.25), and hence equation (2.22), by using the former to compute the time-scales to diffuse from the initial value of ℓ , ℓ_0 , to ℓ_{SB} , in

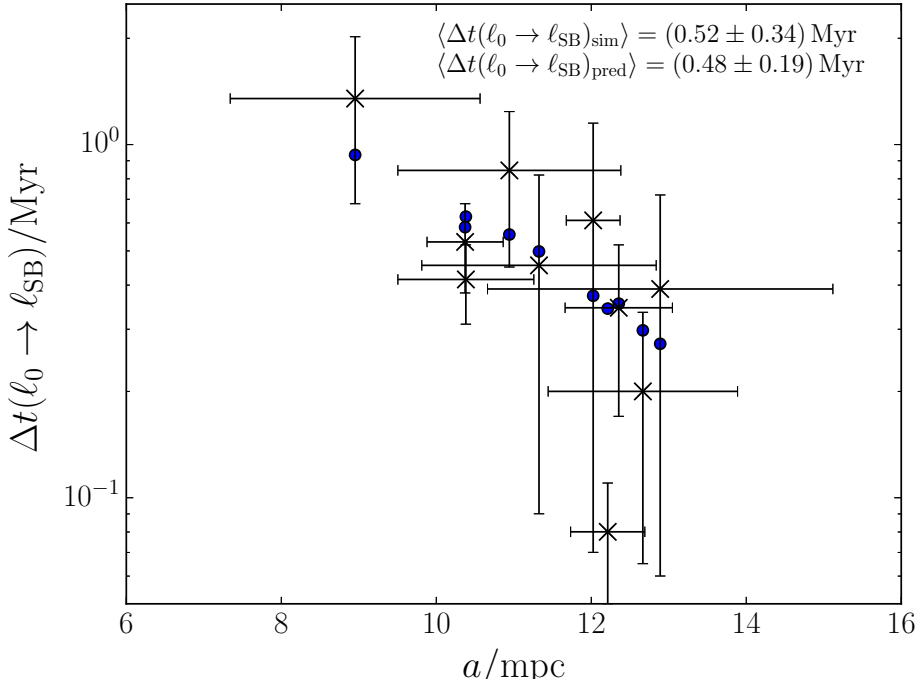
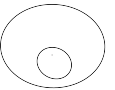


Figure 2.16: The times $\Delta t(\ell_0 \rightarrow \ell_{\text{SB}})$ required to diffuse from the initial angular momentum ℓ_0 to the SB if $\ell_0 < \ell_{\text{SB}}$. Crosses with error bars show diffusion times obtained from the S-star simulations. Horizontal error bars show the standard deviation of the values of a from $t = 0$ until reaching the SB for the first time for all realizations of the S-stars with $\ell_0 < \ell_{\text{SB}}$. Vertical error bars show the median absolute value of the diffusion times based on the latter realizations. Blue bullets show the predicted times computed from equation (2.25), where the mean semimajor axis from the simulations at $\ell = \ell_{\text{SB}}$ was used.



the context of our simulations of the S-stars. In the latter simulations, the initial values of ℓ are $0.14 \lesssim \ell_0 \lesssim 0.37$ with $dN/d\ell_0 = 2\ell_0$ (cf. Section 2.5.1). From the simulations, we selected the S-stars with $\ell_0 < \ell_{\text{SB}}$ and we recorded the time $\Delta t(\ell_0 \rightarrow \ell_{\text{SB}})_{\text{sim}}$ it takes for ℓ to increase from ℓ_0 to ℓ_{SB} . We also recorded the mean semimajor axis in this time interval. The latter value was used to compute the predicted time $\Delta t(\ell_0 \rightarrow \ell_{\text{SB}})_{\text{pred}}$ based on 1000 realizations of the initial value of ℓ (sampled similarly as in the simulations). For each realization we evaluated equation (2.25) if $\ell_0 < \ell_{\text{SB}}$, with $C_{A\text{p}} = 0.5$ and $C_1 = C_2 = 2.6$. From these realizations the mean was adopted as $\Delta t(\ell_0 \rightarrow \ell_{\text{SB}})_{\text{pred}}$.

The predicted times are shown as function of semimajor axis with blue bullets in Fig. 2.16; the times extracted from the simulations are shown as crosses with error bars. According to the prediction the time $\Delta t(\ell_0 \rightarrow \ell_{\text{SB}})$ decreases with increasing a which is borne out by the simulations, although there is considerable scatter. The mean values of $\Delta t(\ell_0 \rightarrow \ell_{\text{SB}})$, averaged over the semimajor axes, are $\langle \Delta t(\ell_0 \rightarrow \ell_{\text{SB}})_{\text{sim}} \rangle \approx (0.5 \pm 0.3)$ Myr and $\langle \Delta t(\ell_0 \rightarrow \ell_{\text{SB}})_{\text{pred}} \rangle \approx (0.5 \pm 0.2)$ Myr for the simulations and predictions, respectively, and are consistent with each other.

2.5.5 A new criterion for the location of the barrier

In MAMW11, the SB was first observed as a locus in the $\log a$ versus $\log(1 - e)$ plane where the N -body trajectories ‘bounced’ in the course of their RR-driven random walk in L . Equation (2.1), which was derived from a simple time-scale argument, was found to reproduce the ‘bounce’ location $e_{\text{SB}}(a)$ with acceptable accuracy in those simulations.

The location of the barrier in the MAMW11 simulations was determined by eye from the $\log a$ versus $\log(1 - e)$ plane. Figs 2.1, 2.7 and 2.14 from this paper suggest a new, more robust criterion for $e_{\text{SB}}(a)$ in terms of the diffusion time-scales or diffusion coefficients.

Under the influence of RR, the diffusion coefficient in ℓ , $\langle (\Delta\ell)^2 \rangle$, first increases towards smaller ℓ , then sharply drops when ℓ is small enough that GR precession suppresses the effects of the torques. A natural definition for the angular momentum associated with the barrier at radius a is the value $\ell = \ell_{\text{p}}(a)$ at which $\langle (\Delta\ell)^2 \rangle$ peaks.

We can implement this criterion in two ways: using our analytic expressions for $\langle (\Delta\ell)^2 \rangle$, or using the numerically computed diffusion coefficients. To the extent that the analytic expressions correctly predict the numerical results, the two approaches should yield similar answers.

The analytic expressions for $\langle (\Delta\ell)^2 \rangle$, equations (2.14) and (2.22), imply a maximum at $\ell = \ell_{\text{b},2}(a)$, the latter given by equation (2.24b). From that expression, the dependence of $\ell_{\text{b},2}$ on $C_{\text{b},2}$ in the limits of large and small a is easily shown to be

$$\ell_{\text{b},2} \rightarrow C_{\text{b},2}^{1/4}, \quad a \rightarrow \infty \quad (C_{\text{b},2} \ll 1) \quad (2.26a)$$

$$\rightarrow 1 - \frac{2}{C_{\text{b},2}}, \quad a \rightarrow 0 \quad (C_{\text{b},2} \gg 1). \quad (2.26b)$$

In the models considered here, $C_{\text{b},2} < 1$ at the radii of interest. Equation (2.26a) implies

$$\ell_{\text{b},2}^2(a) \approx \frac{2\alpha_s}{\sqrt{C_2}C_{A\text{p}}} \left(\frac{r_{\text{g}}}{a} \right) \left[\frac{t_{\text{coh}}}{P(a)} \right] \quad (2.27a)$$

$$\approx 4.0 \left(\frac{r_{\text{g}}}{a} \right) \left[\frac{t_{\text{coh}}}{P(a)} \right]. \quad (2.27b)$$

Unlike equation (2.1), the new expression (2.27) for the barrier location depends explicitly on the coherence time; in fact, $\ell_{b,2}$ is roughly the angular momentum for which the GR precession time equals t_{coh} .

We can also estimate $\ell_p(a)$ directly from the numerically-computed diffusion coefficients. Since the numerical data are noisy, we fit smoothing splines to the measured (X, Y) values in Fig. 2.14, where $X = \log \ell$ and $Y = \log \langle (\Delta \ell)^2 \rangle$. The optimal choice of smoothing parameter for each data set was determined via the standard technique of generalized cross validation (Wahba 1990). An estimate of the uncertainty associated with the location of the peak at each a was then made via the bootstrap, by resampling at random from the measured points and repeating the spline fits, recalculating the smoothing parameter with each new bootstrap sample.

Fig. 2.17 shows the results, for data having $7 \lesssim a/\text{mpc} \lesssim 20$. Values of ℓ_p derived from data both at large and small a are problematic: the former because the data are noisy, the latter because there tends not to be a well-defined maximum. Excluding the two data points at largest and smallest a in Fig. 2.17 results in a set of points that define a good power law; least-squares fit of a straight line to this subset of the data yields

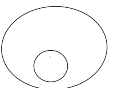
$$\begin{aligned} \ell_p(a) &= \ell_{p,10} \left(\frac{a}{10 \text{ mpc}} \right)^\beta, \\ \ell_{p,10} &= 0.51 \pm 0.016, \\ \beta &= -1.43 \pm 0.086. \end{aligned} \quad (2.28)$$

This relation is statistically indistinguishable from equation (2.16), the SB as defined in MAMW11. Interestingly, that relation is a *better* fit to the points than $\ell_{b,2}(a)$, which is also plotted in Fig. 2.17. The departure of the measured peak-values from the analytic prediction can be understood by referring to Fig. 2.14, which shows that for $a \lesssim 10$ mpc, the peak of the measured diffusion coefficients occurs increasingly at $\ell > \ell_{b,2}$.

The good agreement which we find between the barrier location as defined in MAMW11, and by our new criterion based on the diffusion coefficient, may be partly fortuitous. Nevertheless the agreement is encouraging, since it suggests that the ‘barrier’ that was identified in MAMW11, based on the short-term behaviour of orbits, can be recovered in a robust and quantitative way from simulations. It is also interesting to note that a single relation appears to define the barrier location both in these simulations and those of MAMW11, verifying that equation (2.1) holds true in systems with very different particle numbers and particle masses. At the same time, given the uncertainties in the numerical coefficients, we do not feel confident that we have necessarily ruled out our alternate expression (2.27) for the barrier location and we suggest that future work should compare both that expression and the one given in MAMW11 with the results of numerical simulations.

2.6 Steady-state distribution

In a nucleus where evolution in angular momentum was dominated by NRR, the steady-state phase-space density would be isotropic, $f = f(E)$, and the eccentricity distribution at any energy would be $dN/de = 2e$, a ‘thermal’ distribution. The steady-state eccentricity distribution under the influence of RR has not been well established. The semi-empirical model of Madi-gan et al. (2011, hereafter MHL11) predicts an eccentricity distribution that is bimodal with



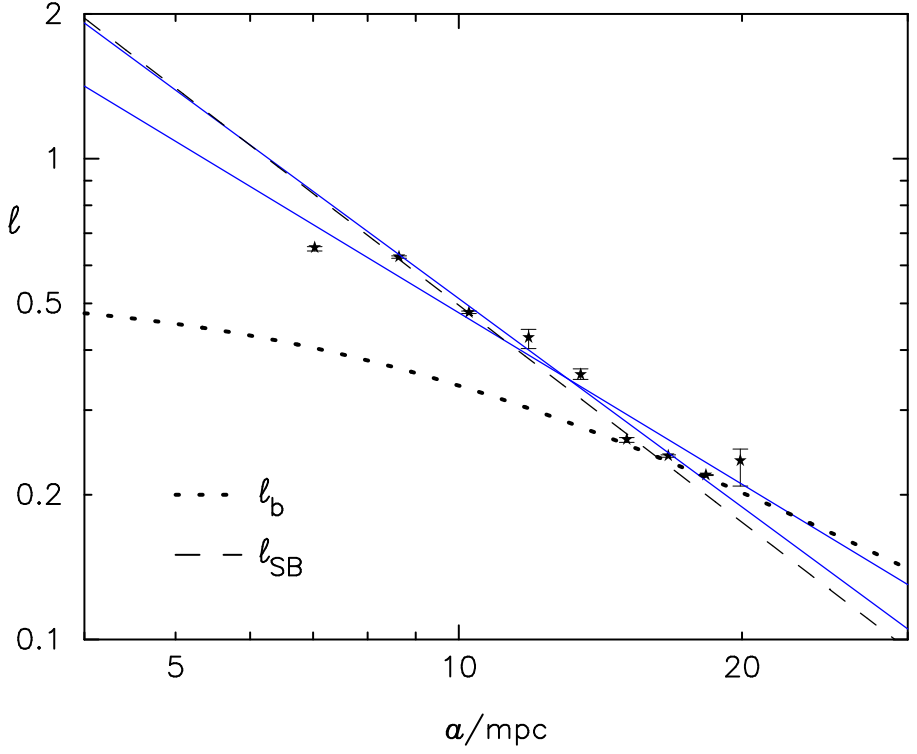


Figure 2.17: Points with error bars are the values of ℓ at which the numerically-computed diffusion coefficients $\langle(\Delta\ell)^2\rangle$ peak in the S-star simulations (cf. Fig. 2.14). Thick (blue) lines are best-fitting power laws: to all of the points (shallow slope) and to a subset that excludes the two data points at largest and smallest a (steep slope). The dashed (black) line is equation (2.1) or (2.16), and the dotted (black) curve is $\ell_{b,2}(a)$, equation (2.24b).

peaks at both low (~ 0.2) and high (~ 0.9) eccentricities at small semimajor axes. Our N -body simulations include the effects of both NRR and RR on the orbital angular momenta, and relativistic corrections to the equations of motion are also taken into account (the latter were not included by MHL11). Using the diffusion coefficients that we obtained in Section 2.5.4 it is therefore possible to investigate, for the first time, the expected steady-state distribution in angular momentum near an SBH under the joint influence of RR, NRR and GR.

Let $N(E, R, t) dR dE$ be the number of stars at time t in angular-momentum interval dR , where $R \equiv L^2/L_c^2(E) \approx 1 - e^2 \equiv \ell^2$, and energy interval dE . The orbit-averaged Fokker-Planck equation is (Merritt 2013a, 5.5.1)

$$\begin{aligned} & \frac{\partial N(E, R, t)}{\partial t} \\ &= -\frac{\partial}{\partial R} [N(E, R, t) \langle \Delta R \rangle] + \frac{1}{2} \frac{\partial^2}{\partial R^2} \left[N(E, R, t) \langle (\Delta R)^2 \rangle \right]. \end{aligned} \quad (2.29)$$

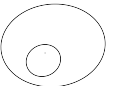
Here, $\langle \Delta R \rangle = \langle \Delta R \rangle(E, R)$ and $\langle (\Delta R)^2 \rangle = \langle (\Delta R)^2 \rangle(E, R)$ are the first- and second-order, orbit-averaged diffusion coefficients in R ; of course, the diffusion coefficients that we extract numerically from the N -body integrations are also orbit averaged. Our motivation for expressing the Fokker-Planck equation in terms of the variable R , rather than L or ℓ , is that the first-order NRR diffusion coefficient in the limit $\ell \rightarrow 0$ diverges as $1/\ell$ (cf. equation 2.12), whereas this divergence in the equivalent limit $R \rightarrow 0$ does not occur if expressed in terms of R . The diffusion coefficients in R can be related, without approximation, to diffusion coefficients in ℓ , i.e. $\langle \Delta \ell \rangle$ and $\langle (\Delta \ell)^2 \rangle$ (Merritt 2013a, eq. 5.167).

Before proceeding, we note the following caveats.

(1) We are finding the steady-state distribution of a set of test stars as they respond dynamically to a *specified* field-star distribution. In reality, the distribution of field stars would also evolve towards a steady state, both with respect to angular momentum (on the RR time-scale) and energy (on the longer NRR time-scale). It is often argued (e.g. Cohn & Kulsrud 1978) that calculating diffusion coefficients from a non-self-consistent L -distribution is an adequate approximation, and in fact this was done in almost all studies prior to ours, including that of MHL11.

(2) Orbit averaging is a way of removing the short time-scale (the radial orbital period) from the problem, by assuming that integrals like L are fixed over this time-scale. In the Newtonian problem, angular momentum is conserved (in a spherical cluster) in the absence of gravitational encounters. In the problem we are solving, there is a second short time-scale when $\ell \lesssim \ell_{\text{SB}}$: the time for GR precession. As noted above, L is not precisely conserved over a GR precessional cycle: it oscillates in response to the (nearly) fixed torques from the field stars. One way to deal with this additional short time-scale would be to express the Fokker-Planck equation in terms of a new quantity that is conserved during the precession; for instance, the ‘secular Hamiltonian’ mentioned in Section 2.5.4. Instead, when applying the Fokker-Planck equation to the $\ell \lesssim \ell_{\text{SB}}$ regime, we interpret ℓ as ℓ_{av} , its average value over a GR precessional cycle. This interpretation is fully consistent with the manner in which the diffusion coefficients were extracted from the simulations. Furthermore, as noted in Section 2.5.4, the ‘secular Hamiltonian’ is essentially ℓ_{av} .

(3) The Fokker-Planck equation assumes that the diffusion coefficients of third and higher order are negligible. In the case of diffusion driven by NRR, this approximation can be justified for intermediate and long time-scales as compared to the relaxation time-scale (e.g. Spitzer 1987); at short time-scales this is likely not the case (Bar-Or et al. 2013). We are not aware



of a justification of the neglect of higher order diffusion coefficients in the case of RR and AR. In fact, extraction of the angular-momentum transition probabilities from N -body simulations (Merritt, unpublished) reveals that the probability distributions are often extremely skewed near the SB, implying non-negligible third-order coefficients. The skewness is related to the ‘bounce’ phenomenon near the SB, and by neglecting it in what follows, our results for the steady-state solutions are likely to have systematic errors near the SB.

(4) We are assuming either zero or constant flux C of orbits in the L -direction (cf. equation 2.30). This assumption cannot be strictly correct because C must be zero at $L = L_c$, the angular momentum of a circular orbit, whereas it is nonzero near the loss cone L_{lc} . In reality, there must therefore also be a flux in the energy direction which supplies the loss of stars near the loss cone. In order to relax our assumption of constant flux, it would be necessary to solve the 2D Fokker-Planck equation for $f(E, L)$, which is beyond the scope of the current paper. We expect, however, that the functional dependence of the steady-state distribution on L is not strongly affected by assuming a constant flux in the L -direction.

(5) As inner boundary condition, we set $f = 0$ for orbits that satisfy the capture criterion that was defined in Section 2.5.1. In some contexts, a more appropriate condition would be to set $f = 0$ at the smaller L corresponding to orbits that intersect the sphere for capture of compact remnants; or at the larger L for which the angular-momentum diffusion time equals the time for gravitational-wave energy loss (cf. Fig. 2.1). Our inner boundary condition is only strictly correct for test stars that have zero mass and radius, and this choice will affect both the steady-state solutions and the implied flux.

With these caveats in mind, we return to equation (2.29) and set $\partial N / \partial t = 0$. The result is

$$-N(R)\langle\Delta R\rangle + \frac{1}{2}\frac{\partial}{\partial R}\left(N(R)\left\langle(\Delta R)^2\right\rangle\right) = C. \quad (2.30)$$

Here, C is an ‘angular momentum flux’. The dependence of both N and C on E (i.e. a) is understood. Equation (2.30) has two types of solutions: those with $C = 0$ (homogeneous; zero flux) and those with $C \neq 0$ (inhomogeneous; constant flux). Exact solutions exist for both cases and are derived in Appendix 2.E.1. These solutions require knowledge of the diffusion coefficients at arbitrary values of the angular momentum.

2.6.1 Analytic solutions

In equations (2.23) and (2.24), we presented approximate analytic expressions for the diffusion coefficients. The second-order coefficients from our N -body simulations are well described in terms of equation (2.23), as was demonstrated in Fig. 2.14. In the case of the first-order coefficients, the agreement of equation (2.23) with the data is good for $\ell \lesssim \ell_{b,1}$. For $\ell \gg \ell_{b,1}$, however, this agreement is poor: $\langle\Delta\ell\rangle$ is expected and observed to be negative at $\ell \approx 1$; the latter feature is not described by equation (2.15). In Fig. 2.14 the value of ℓ for which $\langle\Delta\ell\rangle$ becomes negative, ℓ_c , is weakly dependent on semimajor axis. In addition, the results in Fig. 2.7 suggest that ℓ_c also only weakly depends on $N_{\max}$. We therefore assume that ℓ_c is constant for our present purposes, and adopt the value $\ell_c \approx 0.7$. Furthermore, we adopt $C_1 = C_2 = 2.6$, hence $\ell_{a,2} = 2^{1/4}\ell_{a,1}$ and $\ell_{b,1} = \ell_{b,2}$ (cf. equation 2.24).

To take into account the sign change of $\langle\Delta\ell\rangle$ at $\ell \approx \ell_c$ observed in our N -body simulations, we make the following two changes to equation (2.15). (1) Instead of letting $\langle\Delta\ell\rangle \rightarrow 0$ as $\ell \rightarrow 1$, we let $\langle\Delta\ell\rangle \rightarrow 0$ as $\ell \rightarrow \ell_c$. As ℓ increases to $\ell > \ell_c$, then $\langle\Delta\ell\rangle < 0$. (2) We multiply

the resulting expression by a constant factor to ensure that $\langle \Delta \ell \rangle$ is continuous at $\ell = \ell_{b,1}$. The explicit form of $\langle \Delta \ell \rangle$ in the range $\ell_{b,1} < \ell < 1$ is included in Appendix 2.E.2. For completeness, we have also included there explicit expressions of the diffusion coefficients in ℓ in the other regimes, based on equation (2.23).

In the top panel of Fig. 2.18, we show the analytic functions for the diffusion coefficients described in Appendix 2.E.2 and we compare these to the coefficients obtained from the S-star simulations (cf. Section 2.5.4), for a single semimajor axis bin. The analytic relations capture the basic features of the coefficients obtained from the simulations. The boundary values given by equation (2.24) have been indicated in all of the panels of Fig. 2.18. The quantity R_{loss} is the value of R that corresponds to disruption of the star by the SBH; $R_{\text{loss}}(a) = r_{\text{capt}}/a (2 - r_{\text{capt}}/a)$. In the simulations $r_{\text{capt}} = 2R (M_{\bullet}/m)^{1/3} \approx 3 \times 10^{-2}$ mpc with $R = 8 R_{\odot}$ and $m = 10 M_{\odot}$, giving $R_{\text{loss}} = \mathcal{O}(10^{-2})$ for the semimajor axes of interest.

In the second panel of Fig. 2.18, we show the analytic coefficients transformed to R (cf. equation 2.49). We note that we defined R_c as the value of R for which $\langle \Delta \ell \rangle$ changes sign from positive to negative values. In general, the latter is different from the value of R for which $\langle \Delta R \rangle$ changes sign from positive to negative values, as illustrated in the first and second panels of Fig. 2.18.

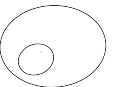
In the third panel of Fig. 2.18, we show the analytic solution of equation (2.30) assuming $C \neq 0$, computed from equation (2.46) (black dashed line). The latter solution is given explicitly in Appendix 2.E.3. For verification of the analytical results we also include results from numerical integrations using the analytic diffusion coefficients (black solid line). The eccentricity distribution $N(e)$ that follows from the solution $N(R)$ is shown in the fourth panel of Fig. 2.18. In this figure and those that follow in this section, all probability density functions are normalized to unit total number.

We note the following features in the analytic solutions based on our analytic approximations of the diffusion coefficients obtained from the N -body simulations:

1. For $R_{\text{loss}} < R < R_{a,1}$, $N(R)$ increases logarithmically with R , i.e. $N(R) \propto \log(R/R_{\text{loss}})$. This is the well-known NRR ‘empty loss cone’ result (Cohn & Kulsrud 1978) and reflects our assumed form of the diffusion coefficients in this regime. For $R_{a,1} < R < R_{a,2}$ the trend of increasing $N(R)$ continues, although the dependence on R is no longer strictly logarithmic.
2. For $R_{a,2} < R_{b,1}$, $N(R)$ decreases with R . Approximately, $N \propto R^{-2}$ for $R \approx R_{b,1}$, independent of C_1 or C_2 if $C_1 = C_2$ (cf. Appendix 2.E.3).
3. For $R_{b,1} < R_c$, $N(R)$ once again increases with R . As R increases to $R > R_c$, $N(R)$ drops. The latter reflects the rapid drop of $\langle (\Delta R)^2 \rangle$ as $R \rightarrow 1$, which can be interpreted as arising from the strongly reduced efficiency of RR as $R \rightarrow 1$.

The above features imply that there are *two* local maxima and *three* local minima in $N(R)$ (and, similarly, in the eccentricity distribution $N(e)$): two maxima at $R_{a,2}$ and near R_c , and three minima at R_{loss} , $R_{b,1}$ and $R = 1$. The local minimum at $R_{b,1}$ is near the ‘knee’ feature that was observed in the diffusion coefficients and, furthermore, $R_{b,1}$ is comparable to R_{SB} . This suggests that the SB can be associated with a *deficit* of orbits in the steady-state angular-momentum distribution.

An interesting feature of the steady-state solution $N(R)$ is a local maximum in N at $R < R_{\text{SB}}$. We suggest that this can be explained by the inefficiency of AR, which is the dominant



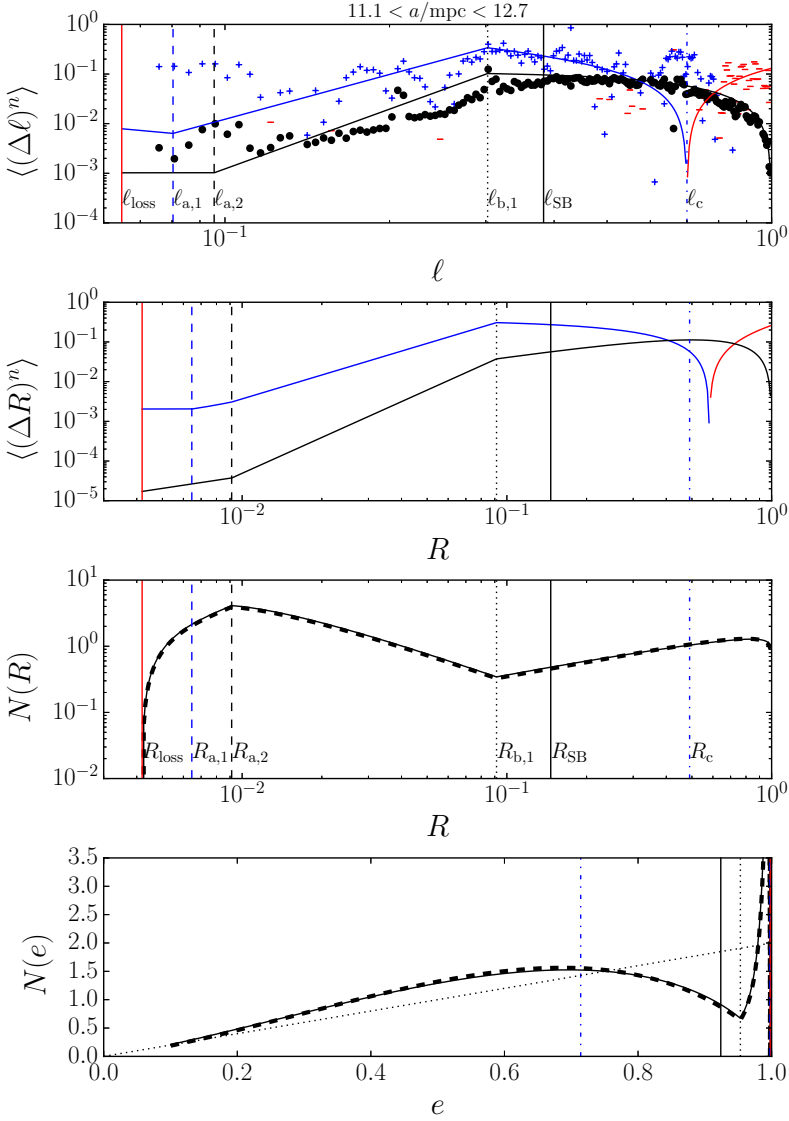


Figure 2.18: Analytic solutions of the steady-state distributions in R and e . Top panel: bullets, minuses and pluses show the diffusion coefficients in ℓ obtained from the S-star simulations (cf. Section 2.5.4). The continuous lines show our adopted analytic model, which is described explicitly in Appendix 2.E.2. Here we set a equal to the mean of the semimajor axis bin from the simulations. Second-order quantities are shown in black; positive (negative) first-order quantities are shown in blue (red). Various boundaries in terms of ℓ (and in terms of $R \equiv \ell^2$ and $e = \sqrt{1 - \ell^2}$ in the other panels) are indicated with vertical lines (cf. equation 2.24). Second panel: the analytic coefficients transformed to R using equation (2.49). Third panel: the solution to equation (2.30) in terms of R . Thick dashed lines: analytic expressions, given in Appendix 2.E.3; solid lines: numerical solutions. Fourth panel: the corresponding solutions in terms of e . The black dotted line shows a thermal distribution.

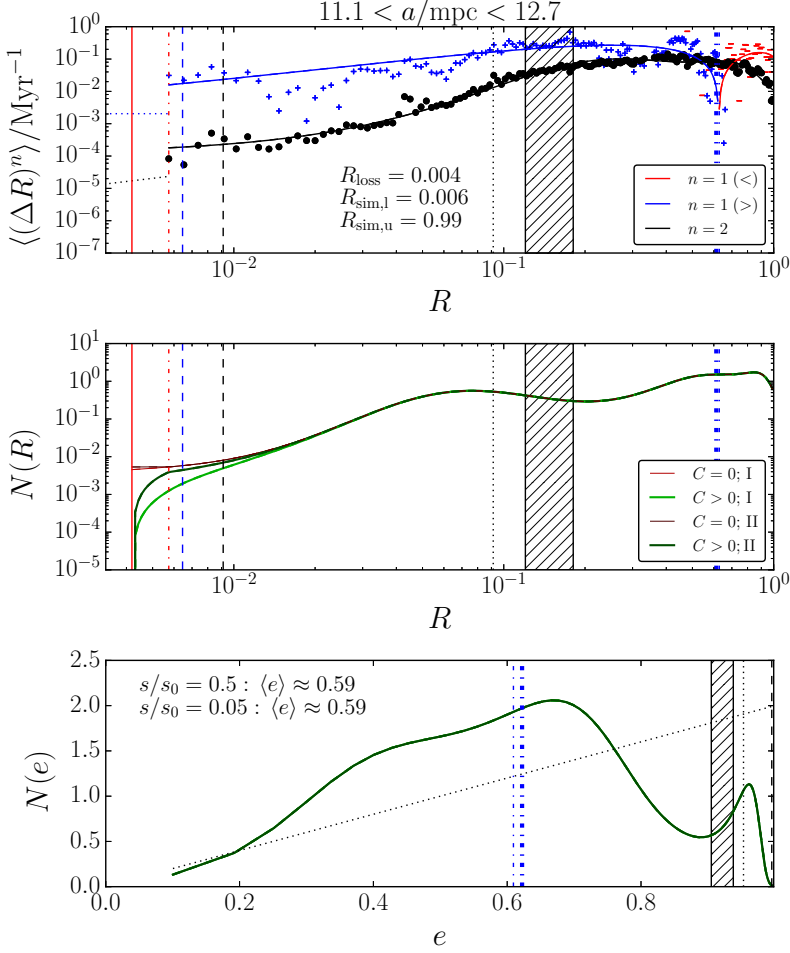
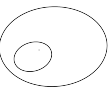


Figure 2.19: Numerical solutions of the steady-state distributions in R and e based on interpolations of the diffusion coefficients extracted from the N -body simulations, for one semimajor axis bin. Top panel: bullets, minuses and pluses show the transformed diffusion coefficients in R derived from the S-star simulations (cf. Section 2.5.4). The continuous lines show fifth-order spline interpolations to the data. Two different values of the smoothing parameter s are adopted: $s = 0.5 s_0$ (solid lines) and $s = 0.05 s_0$ (dashed lines); in this case, however, the results for both values of s are identical. Various boundaries in terms of ℓ and $R \equiv \ell^2$ and $e = \sqrt{1 - \ell^2}$ are indicated with vertical lines (cf. Fig. 2.18). The red vertical dot-dashed line indicates $R_{\text{sim,l}}$, the smallest value of R for which the diffusion coefficients were determined. The blue vertical thick dot-dashed line shows the value of angular momenta for which the interpolated $\langle\Delta\ell\rangle$ changes sign; the blue vertical thin dot-dashed line shows the value of angular momentum for which the interpolated $\langle\Delta R\rangle$ changes sign. The SB, which has a range in angular momentum because there is a range of semimajor axes, is indicated with the black hatched region. Second panel: the numerical solution to equation (2.30) in terms of R based on the interpolations. Two methods were used to extrapolate to the region $R_{\text{loss}} < R < R_{\text{sim,l}}$ that is missing in the data, cf. equation (2.31). Method I: light colour; method II: darker colour. We include solutions with $C = 0$ (red lines) and $C \neq 0$ (green lines). Third panel: the corresponding solution in terms of e . The black dotted line shows a thermal distribution.



form of relaxation in the angular-momentum regime $R_{a,2} < R < R_{b,1}$. If we imagine that the region below the SB was initially unpopulated, stars would diffuse to $R < R_{SB}$ at some rate determined by $\langle(\Delta R)^2\rangle_{SB}$. Once ‘below the barrier’, stars would experience diffusion at much lower rates, causing them to ‘pile up’ until reaching a high enough density that the fluxes in the AR and RR regions are equalized. Apparently, achieving this equality can result in higher values of N below the SB than above – a non-intuitive result given the difficulty of crossing the SB from above. At even lower R , N drops again because of losses to the SBH.

In the next section, we show that, although still clearly present, the increase in the value of N below the SB is probably less extreme than suggested by these analytic solutions.

2.6.2 Numerical solutions

In Section 2.6.1, we presented analytic functions that approximate the diffusion coefficients obtained from our N -body simulations, and we obtained analytic solutions for the steady-state angular-momentum distribution. This method facilitates insight into the steady-state solutions, but it turns out to be inaccurate in so far as the relative heights of the peaks in $N(R)$ are concerned. We also obtained numerical solutions by fitting splines to the diffusion coefficients obtained from the simulations. Although we find the same basic features in $N(R)$ discussed above, the analytic method fails to accurately describe the relative importance of the two local maxima in N . This is likely due to the sensitivity of the solution $N(R)$ to $\langle\Delta\ell\rangle$ in the regime $\ell_{b,1} < \ell < 1$. For example, by multiplying $\langle\Delta\ell\rangle$ by factors of a few in the analytic prescription (this does not make the fit to the data much worse), we find that the peak near R_c becomes much more dominant compared to the peak near $R_{a,2}$.

In this section, we present numerical steady-state solutions based on fifth-order spline fitting³ of the diffusion coefficients in R . The latter were derived from transformation of the measured coefficients in ℓ to R . For all the results shown in this section we adopted the same criterion for the time lags as in Section 2.5.4. We show an example of the spline fitting in the top panel of Fig. 2.19. To obtain better fit results for a large range in R we fitted the logarithm of $\langle(\Delta R)^2\rangle$; this is not the case for $\langle\Delta R\rangle$, which changes sign at $R \approx 0.6$. A parameter that affects the result of the interpolation is the smoothness s of the interpolated spline. For a data set (x_i, y_i) this parameter is defined via the condition that $\sum_i [y_i - h(x_i)]^2 \leq s$, where $h(x)$ is the interpolation function. In order to obtain a measure of uncertainty associated with the choice of s we adopted two values, $s = 0.5 s_0$ and $s = 0.05 s_0$, where $s_0 = N - \sqrt{2N}$ and N is the number of data points. Generally, the former value yields a smooth interpolation, whereas the latter yields a more detailed, but less smooth interpolation, which is more sensitive to scatter in the data.

The interpolated diffusion coefficients have a range $R_{sim,l} \leq R \leq R_{sim,u}$; the boundaries vary per semimajor axis bin. Typically, $R_{sim,l} \sim 10^{-2}$ and $R_{sim,u} \sim 1 - 10^{-2}$. The lower limit $R_{sim,l}$ is comparable to, but slightly larger than the value of R that corresponds to the assumed tidal disruption radius in the simulations, $R_{loss}(a) = r_{capt}/a (2 - r_{capt}/a) = \mathcal{O}(10^{-2})$. For the semimajor axis range shown in the top panel of Fig. 2.19, $R_{loss} \approx 0.004$, whereas $R_{sim,l} \approx 0.006$.

As mentioned above, in the solutions with $C \neq 0$, $N(R)$ is set to zero at R_{loss} . This constraint is physically desirable since close to the SBH the distribution function should be zero at

³We used the SPLPREP routine implemented in SCIPY, a PYTHON library.

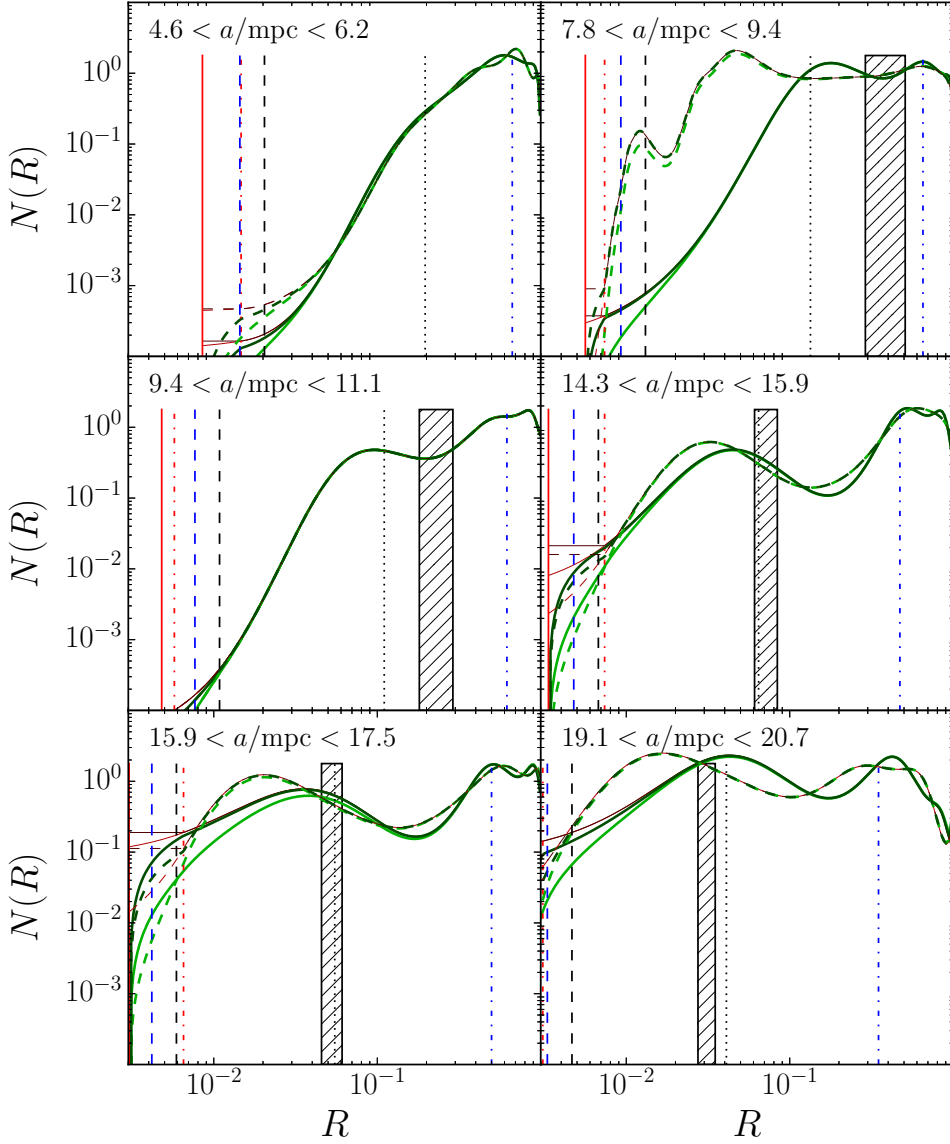
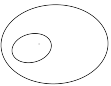


Figure 2.20: Numerical steady-state distributions in R based on interpolated diffusion coefficients extracted from the N -body simulations as in Fig. 2.19, for different semimajor axis bins. Line colours and styles are the same as in Fig. 2.19. The blue vertical dot-dashed line shows the value of R for which the interpolated $\langle \Delta \ell \rangle$ changes sign. The time lags for the diffusion coefficients are chosen according to the criterion discussed in Section 2.4.5.



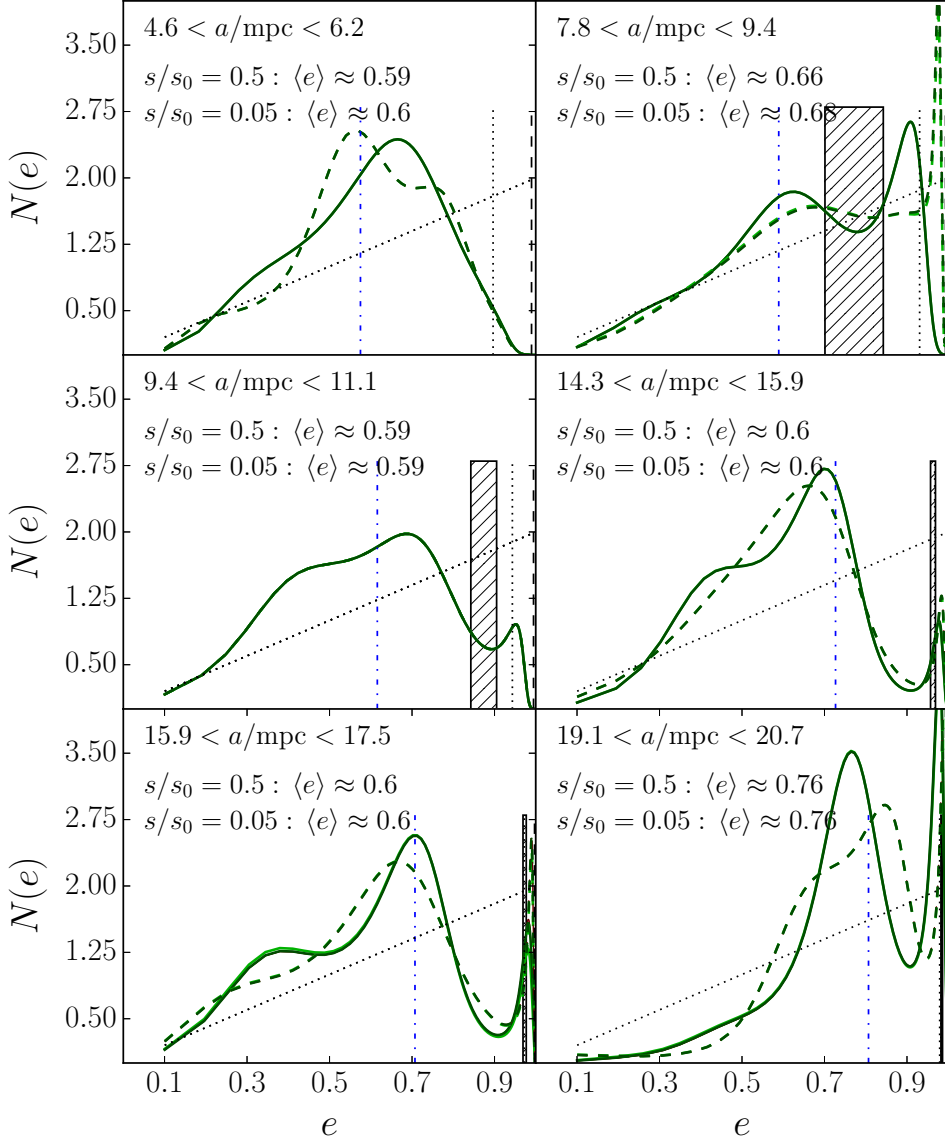


Figure 2.21: Steady-state eccentricity distributions based on the solutions shown in Fig. 2.20.

$R < R_{\text{loss}}$ ('empty loss cone'). Implementing $N(R_{\text{loss}}) = 0$ requires knowledge of the diffusion coefficients in the range $R_{\text{loss}} < R < R_{\text{sim},1}$, which is not available in our data. Therefore, we imposed two different extrapolations for the diffusion coefficients in this regime,

$$\left\{ \begin{array}{l} \langle (\Delta R)^n \rangle (R) = \langle (\Delta R)^n \rangle (R_{\text{sim},1}); \quad (\text{method I}) \\ \left\{ \begin{array}{l} \langle \Delta R \rangle (R) = A(E); \\ \langle (\Delta R)^2 \rangle (R) = 2RA(E). \end{array} \right. \quad (\text{method II}) \end{array} \right. \quad (2.31)$$

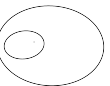
Here, $A(E)$ is given by equation (2.13). Method I amounts to imposing the constant values at $R_{\text{sim},1}$, whereas method II adopts the NRR diffusion coefficients in the limit $R \rightarrow 0$. The latter coefficients are shown with dotted lines in the first panel of Fig. 2.19 in the range $R_{\text{loss}} < R < R_{\text{sim},1}$.

In the middle panel of Fig. 2.19, we show the steady-state solution $N(R)$ for a single semimajor axis range, computed for the cases $C = 0$ (red lines) and $C \neq 0$ (green lines). In both cases, we employ the two different extrapolation methods in the regime $R_{\text{loss}} < R < R_{\text{sim},1}$ (light colour: method I; dark colour: method II). Dashed (solid) lines correspond to $s = 0.05 s_0$ ($s = 0.5 s_0$). The corresponding eccentricity distribution is shown in the bottom panel of Fig. 2.19. While qualitatively similar to the corresponding plots in Fig. 2.18, there are important differences. Most notably, the peak in $N(R)$ below the SB is much less dominant compared to the peak above the SB.

We show similar results for different semimajor axis bins in Figs 2.20 (in terms of R) and 2.21 (in terms of e). In these plots, at low R , $R_{\text{loss}} < R < R_{\text{sim},1}$, the solutions with $C = 0$ and $C \neq 0$ deviate from each other. However, for $R \gtrsim R_{\text{sim},1} \sim 10^{-2}$ the solutions are indistinguishable for the same smoothness s . Likewise, the choice of extrapolation in the regime $R_{\text{loss}} < R < R_{\text{sim},1}$ (i.e. method I or II) does not noticeably affect the solution for $R \gtrsim R_{\text{sim},1}$. The latter value of R corresponds to a very high eccentricity, $e \sim 0.995$. Consequently, the eccentricity distributions (cf. Fig. 2.21) are visually unaffected by the flux constraints nor by the choice of extrapolation. The only parameter that does noticeably affect the solutions is the interpolation smoothness parameter s (i.e. compare the solid and dashed lines). Nevertheless, the solutions are qualitatively similar for both values of s .

Except for $4.6 < a/\text{mpc} < 6.2$, the SB is present in the semimajor axis bins shown in Figs 2.20 and 2.21, i.e. $a > C_{\text{SB}}$. Here, $C_{\text{SB}} = [r_g a_{\text{max}}^{1/2} (M_{\bullet}/m_{\star})(1/N_{\text{max}}^{1/2})]^{3/2} \approx 6.25 \text{ mpc}$ is the smallest value of a for which the SB exists (MAMW11). Although the solutions for these larger a values do depend somewhat on the degree of smoothing, two local maxima and three local minima can always be observed in the distributions in R . These extrema have the following locations:

1. A minimum near $R \approx 1$.
2. A maximum near R_c , which we determine from the value of R for which the (interpolated) $\langle \Delta \ell \rangle$ changes sign (blue vertical dot-dashed line).
3. A minimum near or slightly above R_{SB} .
4. A maximum between $R_{a,2}$ (black dashed line) and R_{SB} .
5. A minimum near R_{loss} (red solid line).



These locations are generally consistent with those found using our analytic expressions for the diffusion coefficients (cf. Section 2.6.1). An exception is the maximum between $R_{a,2}$ and R_{SB} which, according to the analytic solutions, should occur near $R_{a,2}$. In the solutions based on the interpolations, this maximum occurs at a somewhat larger value of R .

For the smallest values of semimajor axis, shown in the top-left panel of Figs 2.20 and 2.21, the SB does not exist, i.e. $a < C_{SB}$. It is not surprising that the steady-state solutions at these small radii are systematically different compared to those farther out. In this regime, orbits at all R are strongly affected by GR precession and RR is not effective at any R . Our analytic prescription of the diffusion coefficients breaks down in this regime, as illustrated by the first panel of Fig. 2.14, where the measured diffusion coefficients are systematically lower than our predictions. A detailed description of diffusion in this regime is beyond the scope of this paper. Nevertheless, the numerical solutions indicate that the local minimum that was observed near R_{SB} for larger semimajor axes, disappears. This is not surprising, considering that the SB does not exist in this radial range. The maximum near R_c , on the other hand, becomes more pronounced.

2.7 Discussion

2.7.1 Limits on the typical S-star age from the N -body simulations

In our simulations of the S-stars, we assumed that their orbits about the SBH were initially very eccentric, $0.93 < e_0 < 0.99$. We considered two possibilities for the nature of the rate of supply of S-stars to the GC: formation in a burst or continuous formation. The former assumption is consistent with the infall of a young stellar cluster, possibly with a central intermediate mass black hole, into the GC that subsequently dissolves and leaves massive stars tightly bound to the SBH (Hansen & Milosavljević 2003; Berukoff & Hansen 2006; Fujii et al. 2010). There are numerous problems with this scenario, however (see e.g. Perets & Gualandris 2010). An alternative possibility is binary disruption, in which case the rate of supply of S-stars to the GC is expected to be continuous, if averaged over a sufficiently long time. Massive perturbers like giant molecular clouds are a promising candidate for strongly perturbing the orbits of stellar binaries outside the central parsec into loss-cone orbits at a rate that is high enough to account for the current number of S-stars and high-velocity stars (Perets et al. 2007; Perets & Gualandris 2010).

In the case of burst formation, we have found in our simulations that the cumulative eccentricity distribution rapidly evolves to a distribution that is consistent with observations in ~ 7 Myr (cf. Fig. 2.11). These results have also been extrapolated to include continuous formation and we have found that in this case the minimum time to evolve to the observed distribution is ~ 25 Myr (cf. Fig. 2.30). If our assumptions of the formation process (high initial eccentricities) and the field-star distribution (a cusp of stellar black holes) are correct, then the consistency of the eccentricity distribution with observations after a certain time implies a lower limit on the typical S-star lifetime and hence an upper limit on the typical S-star mass. We emphasize that only conclusions can be drawn for the typical age, because $p(t)$ applies to the S-stars as a whole population. Assuming solar metallicity, the lower limit of the typical age of ~ 7 Myr in the burst scenario corresponds to an upper limit of the typical mass of $\sim 24 M_\odot$. The lower limit of the typical age of ~ 25 Myr in the case of continuous S-star formation corresponds to

an upper limit of the typical mass of $\sim 10 M_\odot$. The observed spectral types of the S-stars range from B0 V to B9 V (Eisenhauer et al. 2005), or $3 \lesssim m/M_\odot \lesssim 20$. Furthermore, the initial mass function (IMF) of the S-stars is consistent with a Salpeter IMF, $dN/dm \propto m^{-2.15 \pm 0.3}$ (Bartko et al. 2010), which implies a mean mass of $\langle m \rangle = (6.4 \pm 0.5) M_\odot$. The latter mass is consistent with our upper limits of the typical mass for both burst and continuous formation.

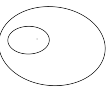
2.7.2 S-star relaxation times for different field-star models

In the N -body simulations of the S-stars, a cusp of stellar black holes was assumed. Although predicted by theory (Bahcall & Wolf 1976), so far no direct evidence for the presence of such a cusp in the GC has been found. Observations of late-type stars in the GC (Buchholz et al. 2009; Do et al. 2009; Bartko et al. 2010) indicate that there is a core of size ~ 0.5 pc in the distribution of these stars, which is well outside the radial extent of the S-star cluster. Such a core can be represented by a density slope $\gamma = 1/2$ (Merritt 2010), which is the lowest possible value consistent with an isotropic velocity distribution.

In order to estimate the effect of a core of late-type stars on the typical time-scale for the orbits of the S-stars to evolve to eccentricities consistent with observations (as opposed to a cusp of stellar black holes), we applied equation (2.25) using a similar method as in Section 2.5.4. Here, we adopted $\gamma = 1/2$, $N_{\max} = 8 \times 10^4$ and $m_\star = 1.0 M_\odot$ from the stellar core model that was assumed in AM13.

For a range of semimajor axes, an initial value of ℓ consistent with binary disruption ($0.93 < e_0 < 0.99$), ℓ_0 , was sampled in 10^4 Monte Carlo realizations. In each of these, a value $\ell_p > \ell_0$ was sampled from the cumulative distribution $\text{CDF}(\ell) = 1 - (1 - \ell^2)^{p/2}$ which corresponds to a cumulative eccentricity distribution $\text{CDF}(e) = e^p$, where $p = 2.6$ was adopted to be consistent with observations (Gillessen et al. 2009). The time-scale for ℓ to increase from ℓ_0 to ℓ_p was then computed as follows. For $\ell \leq \ell_{\text{SB}}$, equation (2.25) was applied assuming $C_{\text{Ad}} = 0.5$ and $C_1 = 2.6$; the mass precession time-scale $\langle t_{\text{MP}} \rangle$ was approximated by $\langle t_{\text{MP}} \rangle \approx (1/1.2)[M_\bullet/M_\star(a)]P(a)$ (Merritt 2013a, 4.4.1). For $\ell > \ell_{\text{SB}}$, the estimate $\Delta t_{\text{RR}} = t_{\text{RR},0} \Delta \ell^2$ was applied, where $t_{\text{RR},0} = [M_\bullet/M_\star(a)]N_\star(a)P(a)^2/t_{\text{coh}}(a)$. In the latter estimate, the dependence of the RR time-scale on ℓ was neglected for simplicity (cf. equation 2.14). Subsequently, by averaging over the Monte Carlo realizations we obtained $\langle \Delta t(\ell_0 \rightarrow \ell_p) \rangle$, the approximate time-scale for ℓ to increase from a value consistent with a highly eccentric orbit (e.g. as a result of binary disruption) to a value consistent with the ‘super-thermal’ eccentricities of the S-stars. In this method, the semimajor axes were assumed to be constant during the relaxation process.

The resulting time-scales are plotted as function of semimajor axis in Fig. 2.22 (solid lines). In that figure, we also included similar calculations for a cusp of stellar black holes as was assumed in Section 2.5 (dotted lines). For $a < C_{\text{SB}}$, where C_{SB} is the smallest value of a for which the SB exists (MAMW11), the time-scale decreases with increasing a . This can be understood from equation (2.25): neglecting ℓ_{SB} in that equation it can be shown that $\Delta t \propto a^{\gamma-5/2}$ and $\propto a^{2\gamma-13/2}$ assuming that orbital precession is dominated by relativity and mass precession, respectively. Therefore, the time-scales decrease with a for both $\gamma = 1/2$ and $\gamma = 2$. Conversely, for $a > C_{\text{SB}}$ the time-scale increases with increasing a , which can be understood from the scaling of the RR time scale with a (cf. equation 2.14): assuming mass precession the scaling is $\Delta t \propto a^{3/2}$, independent of γ . Note that in the stellar core model most of the orbits of the S-stars lie below the SB for any value of ℓ (cf. the left-hand panel of fig. 1 of AM13).



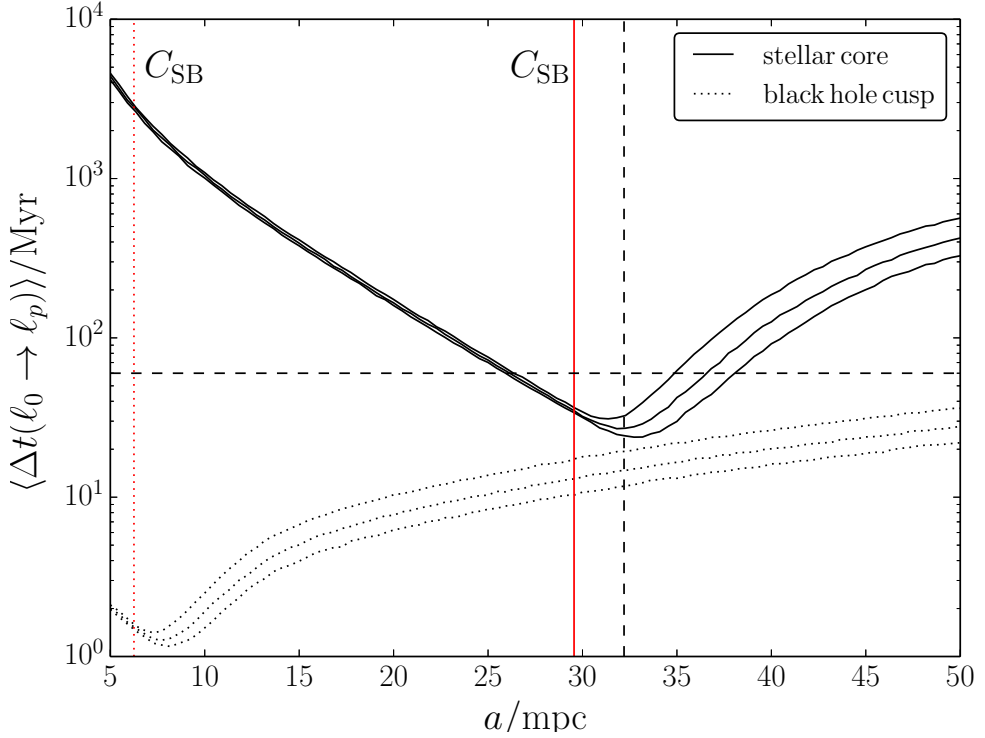


Figure 2.22: Evolution time-scales for the eccentricity of the S-stars to reach a value consistent with observations ($\ell = \ell_p$) starting from high eccentricities consistent with binary disruption, computed using Monte Carlo realizations of equations (2.14) and (2.25). Black solid lines: assuming a stellar core model; black dotted lines: assuming a black hole cusp model (see text). For each model, the time-scales were computed for three different values of p to take into account the uncertainty of the observed value of $p_{\text{obs}} = 2.6 \pm 0.9$ (Gillessen et al. 2009): $p = 2.6 - 0.9$, $p = 2.6$ and $p = 2.6 + 0.9$. The red solid and dotted lines indicate the value of C_{SB} for the stellar core and black hole cusp models, respectively, where C_{SB} is the minimum semimajor axis at which the SB is predicted to exist (cf. Section 2.6.2). The black vertical dashed line indicates 32.2 mpc, approximately the outer extent of the S-star cluster. The black horizontal dashed line indicates 60 Myr, an estimate of the mean age of the S-stars.

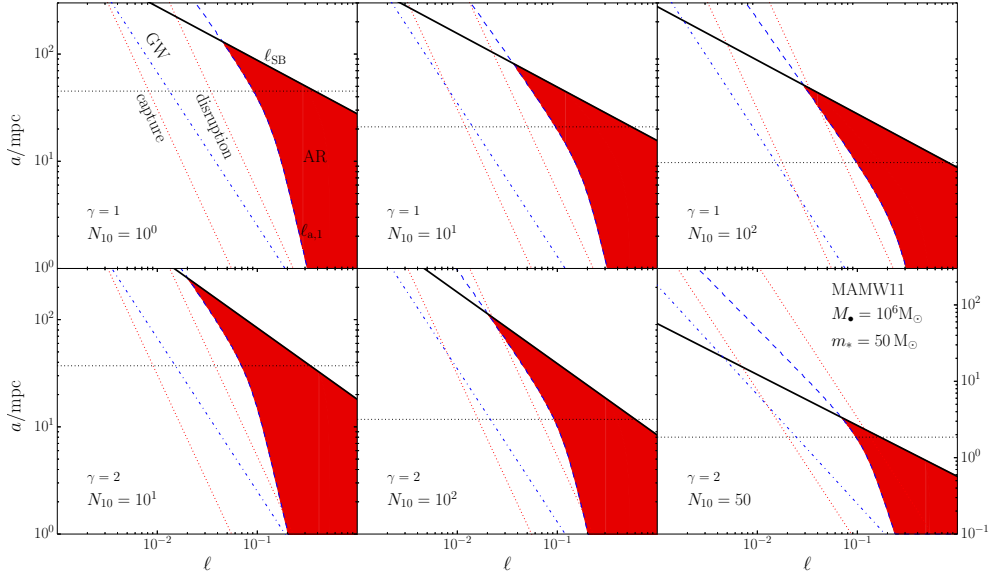
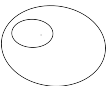


Figure 2.23: Several quantities of importance for AR shown in the (a, ℓ) -plane for nuclear models with different γ and N_{10} , the number of stars within 10 mpc. Black solid line: the SB (equation 2.1). Blue dashed lines: $\ell_{a,1}$ (equation 2.24). The black horizontal dotted line shows an estimate of the transition between GR and mass precession, a_{trans} (equation 2.35). Two red dotted lines: radii for capture of compact objects (left) and tidal disruption of stars (right) by the SBH. Blue dot-dashed line: an estimate of where changes in orbital eccentricity due to gravitational-wave energy loss occur at the same rate as changes due to two-body relaxation (MAMW11, equation 62). We have indicated with red shading the region in which we expect AR to dominate angular-momentum relaxation. In all panels $M_{\bullet} = 4 \times 10^6 M_{\odot}$ and $m_{\star} = 10 M_{\odot}$ is assumed, with the exception of the bottom right panel, where we adopted the model of MAMW11: $M_{\bullet} = 10^6 M_{\odot}$, $m_{\star} = 50 M_{\odot}$, $\gamma = 2$ and $N_{10} = 50$. Note that there is a different range of the vertical axis in the latter panel.

For the stellar core model, the evolution time-scale exceeds 60 Myr, an estimate of the mean S-star lifetime assuming a mean mass of $\sim 6 M_{\odot}$ (Eisenhauer et al. 2005), for a large range of semimajor axes. This is consistent with the result of AM13 that the eccentricity distribution of the S-stars cannot evolve to the observed distribution over the lifetime of the S-stars in the case of a stellar core (cf. the top-left panel of fig. 3 of AM13). The long evolution time-scales in the case of a stellar core would suggest that either (1) the assumption of the formation mechanism of the S-stars (i.e. high initial eccentricities) is incorrect, or that (2) a stellar core of is not the dominant cause of relaxation of the S-stars. Interestingly, the evolution time-scales are consistent with (i.e. shorter than) the ages of the S-stars when assuming a cusp of stellar black holes (cf. the dotted lines in Fig. 2.22).



2.7.3 Generalizations to other galactic nuclei

In Section 2.5.4, we presented analytic expressions for the diffusion coefficients which were calibrated using N -body simulations with assumed parameters $\rho_\star(r) \propto r^{-2}$, $m_\star = 10 M_\odot$ and $N_{\max} = 4800$. It is of interest to investigate whether these relations also apply to nuclear star clusters with different properties. We have also carried out a set of simulations with $\rho_\star(r) \propto r^{-1}$, $m_\star = 10 M_\odot$ and $N_{\max} = 2500$. The results of the latter simulations, i.e. with $\gamma = 1$, are presented in Appendix 2.F. The main conclusions that can be drawn from these additional simulations is that the features in the diffusion coefficients that are associated with AR and that were observed in the simulations with $\gamma = 2$, are also present in the simulations with $\gamma = 1$. In particular, the analytic approximation of the diffusion coefficients that was presented in equation (2.23) also describes the data well for $\gamma = 1$. There is an exception for larger semimajor axes, for which it appears that C_1 and C_2 (cf. equation 2.22) increase with semimajor axis.

These results indicate that it is justified to extrapolate the relations presented in Section 2.5.4 to nuclear star clusters with different properties. We adopt the reference values $M_\bullet = 4.0 \times 10^6 M_\odot$ and $m_\star = 10 M_\odot$, hence $\log(\Lambda) = \log[M_\bullet/(2m_\star)] \approx 12.2$. We consider two values of γ , $\gamma = 1$ and $\gamma = 2$, for which $C_{\text{NRR}}(1) \approx 0.07$ and $C_{\text{NRR}}(2) \approx 0.18$ (cf. equation 2.13 and Appendix 2.B). Furthermore, for both values of γ we adopt $C_{\text{Ad}} = 0.5$ (cf. Fig. 2.15) and $C_1 = C_2 = 2.6$ (cf. Fig. 2.14). For the coherence time t_{coh} we assume $t_{\text{coh}}^{-1} = \langle t_{\text{GR}} \rangle^{-1} + \langle t_{\text{MP}} \rangle^{-1}$ as before, with $\langle t_{\text{GR}} \rangle = (1/12)(a/r_g)P(a)$ and $\langle t_{\text{MP}} \rangle = C_{\text{MP}}(\gamma)[M_\bullet/M_\star(a)]P(a)$, where $C_{\text{MP}}(1) = 1$ and $C_{\text{MP}}(2) = 3/2$ (Merritt 2013a, 4.4.1).

We show the main relations in the (a, ℓ) -plane in Fig. 2.23. Models are included with the two values of γ and various values of N_{10} , the number of stars within 10 mpc. The SB (equation 2.1) is shown with the black solid line. We show two relevant periapsis distances that are associated with losses to the SBH: the stellar tidal disruption radius, $r_{\text{dis}} = 2R(M_\bullet/m)^{1/3} \approx 2.7 \times 10^{-3}$ mpc (assuming $R = 8 R_\odot$ and $m = 10 M_\odot$) (Antonini et al. 2011) and the radius for capture of compact remnants, $r_{\text{capt}} = 8 r_g \approx 1.5 \times 10^{-3}$ mpc (Will 2012).

In Section 2.5.4, an expression was presented for $\ell_{a,n}$, the lower boundary in ℓ for which we expect AR to dominate diffusion in angular momentum. Note that $\ell_{a,2} = 2^{1/4}\ell_{a,1}$ if $C_1 = C_2$, which we find is the case in our N -body simulations and which we adopt here. If precession of the field star orbits is dominated by GR precession, then $\ell_{a,n}$ can be written as:

$$\ell_{a,n} = n^{1/4} \tilde{C}_N(\gamma) \left(\frac{1}{12} \frac{r_g}{a} \right)^{1/4} \quad (2.32a)$$

$$\approx 0.11 n^{1/4} \left(\frac{M_\bullet}{4 \times 10^6 M_\odot} \right)^{1/4} \left(\frac{a}{10 \text{ mpc}} \right)^{-1/4}, \quad (2.32b)$$

the latter assuming $\gamma = 2$. Here, we defined

$$\tilde{C}_N = \tilde{C}_N(\gamma) = \left[\frac{\log(\Lambda)}{C_{\text{NRR}}(\gamma) C_n C_{\text{Ad}}^2} \right]^{1/4}. \quad (2.33)$$

On the other hand, if $t_{\text{MP}} \ll t_{\text{GR}}$, then precession is dominated by mass precession. In this

case,

$$\ell_{a,n} = n^{1/4} \tilde{C}_N(\gamma) C_{\text{MP}}^{1/4}(\gamma) \left(\frac{r_g}{a}\right)^{1/2} \left(\frac{M_\bullet}{m_\star}\right)^{1/4} N_\star(a)^{-1/4} \quad (2.34a)$$

$$\approx 0.12 n^{1/4} \left(\frac{M_\bullet}{4 \times 10^6 M_\odot}\right)^{3/4} \left(\frac{m_\star}{10 M_\odot}\right)^{-1/4} \times \left(\frac{N_\star(a)}{10^2}\right)^{-1/4} \left(\frac{a}{10 \text{ mpc}}\right)^{-1/2}, \quad (2.34b)$$

the latter assuming $\gamma = 2$.

The transition between the two regimes of field-star precession occurs near a_{trans} , which we define as the value of a for which $\langle t_{\text{GR}} \rangle = \langle t_{\text{MP}} \rangle$. With our assumptions, a_{trans} is given by

$$a_{\text{trans}} = \left[12 C_{\text{MP}}(\gamma) r_g a_{\text{max}}^{3-\gamma} N_{\text{max}}^{-1} \left(\frac{M_\bullet}{m_\star}\right) \right]^{1/(4-\gamma)} \quad (2.35a)$$

$$\approx 11.7 \text{ mpc} \left(\frac{M_\bullet}{4 \times 10^6 M_\odot}\right) \left(\frac{m_\star}{10 M_\odot}\right)^{-1/2} \left(\frac{N_{10}}{10^2}\right)^{-1/2}, \quad (2.35b)$$

the latter assuming $\gamma = 2$. In Fig. 2.23 a_{trans} is indicated with the horizontal black dotted line. Furthermore, we show in that figure $\ell_{a,1}$ (blue dashed line), with the coherence time computed from $t_{\text{coh}}^{-1} = \langle t_{\text{GR}} \rangle^{-1} + \langle t_{\text{MP}} \rangle^{-1}$.

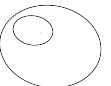
In Fig. 2.23, we have indicated with red shading the approximate region in which we expect that AR dominates evolution in angular momentum. In general, AR is expected to be important in the region $\ell_{a,1} \lesssim \ell \lesssim \ell_{\text{SB}}$. It can be seen in Fig. 2.23 that there is a critical value of a , $a_{\text{AR,max}}$, where $\ell_{a,1} = \ell_{\text{SB}}$. For $a > a_{\text{AR,max}}$, $\ell_{a,1} > \ell_{\text{SB}}$, and we expect AR not to be active at any ℓ . Instead, we expect that NRR dominates angular-momentum relaxation below the SB and that RR dominates above the SB. Hence, we expect that the AR regime disappears for $a > a_{\text{AR,max}}$.

As shown in Fig. 2.23, the value of $a_{\text{AR,max}}$ is large for the nuclear models considered here. Assuming that near $a_{\text{AR,max}}$ precession is dominated by mass precession, which is borne out by Fig. 2.23, and combining equations (2.1) and (2.24), we find

$$a_{\text{AR,max}} \approx \left[\tilde{C}_N^{-1}(\gamma) C_{\text{MP}}^{-1/4}(\gamma) r_g^{1/2} a_{\text{max}}^{3-\gamma} N_{\text{max}}^{-1/4} \left(\frac{M_\bullet}{m_\star}\right)^{3/4} \right]^{\frac{4}{5-\gamma}} \quad (2.36a)$$

$$\approx 114 \text{ mpc} \left(\frac{M_\bullet}{4 \times 10^6 M_\odot}\right)^{5/3} \left(\frac{m_\star}{10 M_\odot}\right)^{-1} \times \left(\frac{N_{10}}{10^2}\right)^{-1/3}, \quad (2.36b)$$

the latter assuming $\gamma = 2$. Unless N_{10} is very large, $N_{10} \gtrsim 10^3$, $a_{\text{AR,max}} \sim 10^2 \text{ mpc}$ is large compared to other values of a of interest in Fig. 2.23. This shows that for many models of galactic nuclei there is a large regime in the energy and angular-momentum space in which AR is important. We note that the quantity $a_{\text{AR,max}}$ derived above is the same as another critical semimajor axis that was defined in section VC of MAMW11. The latter quantity was argued



to be the minimum value of a for which NRR would allow orbits to ‘penetrate’ the SB. The equivalence of these two quantities is shown explicitly in Appendix 2.G.

For comparison purposes, we have also included in the bottom-right panel of Fig. 2.23 the N -body model that was adopted in MAMW11. In that model, $M_\bullet = 10^6 M_\odot$, $m_\star = 50 M_\odot$ and $N_{10} = 50$. Semimajor axes were sampled from a distribution consistent with $\gamma = 2$ with $0.1 < a/\text{mpc} < 10$. Equations (2.35) and (2.36) applied to this model give $a_{\text{trans}} \approx 1.9 \text{ mpc}$ and $a_{\text{AR,max}} \approx 3.1 \text{ mpc}$. Unlike the other models considered above, in the MAMW11 model a_{trans} and $a_{\text{AR,max}}$ are comparable, implying that in the AR regime field-star precession is driven mainly by relativistic precession. Moreover, in the latter model, $a_{\text{AR,max}} \approx 2 - 3 \text{ mpc}$, while the stellar orbits had $0.1 \text{ mpc} \leq a \leq 10 \text{ mpc}$. It follows that for the stars with the larger a -values in MAMW11, NRR was the dominant diffusion mechanism acting on stars after they had crossed the SB; only for $a \lesssim 3 \text{ mpc}$ was AR effective. Indeed it was shown in that paper that essentially all of the stars that were captured by the SBH had $a \gtrsim 2 \text{ mpc}$, and it was argued that ‘penetration’ of the SB was probably driven by NRR for these stars.

The trend seen in the panels in Fig. 2.7 with different N_{max} can similarly be explained by the scaling of $a_{\text{AR,max}}$ with N_{max} : as N_{max} increases, $a_{\text{AR,max}}$ increases, thereby increasing the importance of AR. More quantitatively, for fixed stellar mass $M_\star(a) = m_\star N_\star(a)$, as was assumed in Fig. 2.7, equation (2.36) implies $a_{\text{AR,max}} \propto N_\star(a)^{2/3}$. The values of $a_{\text{AR,max}}$ in the models shown in the different panels in Fig. 2.7 are $\approx 3.1, 4.8, 7.5$ and 11.7 mpc for $N_{\text{max}} = 50, 100, 200$ and 400 , respectively. In the latter model, $a_{\text{AR,max}} > a_{\text{max}} = 10 \text{ mpc}$.

2.7.4 Caveats of TPI

The code presented in Section 2.3 has the advantage of linear scaling with the number of field stars (for a fixed number of test stars), enabling simulations with much larger numbers of stars ($\gtrsim 10^3$) than are currently feasible using fully general N -body codes. The disadvantage is that the motion of the field stars is not reproduced precisely. By allowing the field-star orbits to precess, we do reproduce in an approximate way the dynamical effects of the smoothly distributed field-star mass (‘mass precession’) and of the 1PN relativistic corrections (‘Schwarzschild precession’). But the TPI algorithm does not reproduce either (i) interactions between field stars due to discreteness of the mass distribution, or (ii) the dynamical influence of the test stars on the field stars.

In the time- and spatial domains of interest here, discrete interactions between field stars can change both the magnitude and the direction of the field-star $\mathbf{L}$ -vectors (changes in energy occur on longer time-scales). Changes in the magnitude of $\mathbf{L}$, i.e. in orbital eccentricity, would cause $N(e)$ for the field stars to evolve with a characteristic time $\sim t_{\text{RR}}$ towards some steady-state distribution. In so far as the steady state $N(e)$ which we infer for the *test* stars is not hugely different from a ‘thermal’ distribution – the same distribution which we *assumed* for the field stars – we do not expect this evolution to be of much consequence for any of our results. Changes in the *direction* of the field-star $\mathbf{L}$ vectors on the other hand, constitute an additional form of precession and as such would play a role in determining the coherence time – which we recall is defined as the time for a typical (field) star orbit to precess and so is a function only of r or a . Changes in orbital orientation due to $\sqrt{N}$ torques occur on the coherent RR time-scale, $t_{\text{RR,coh}} \approx [M_\bullet/M_\star(a)]N_\star(a)^{1/2}P(a)$ (Merritt 2013a, p. 275); comparing this time-scale to the mass precession time-scale (cf. equation 2.3) one finds $t_{\text{MP}}/t_{\text{RR,coh}} \propto N_\star(a)^{-1/2}$. Therefore, for sufficiently large $N_\star(a)$, precession of orbital planes can be neglected compared to mass

precession. The consistency between the different codes in Section 2.4 suggests that $N_{\max} = 50$ is already sufficiently large for this to be the case.

With regard to (ii), i.e., neglect of test star – field-star perturbations, the consequences are less certain. Discussions in the literature of RR almost always ignore the dynamical influence of the test star on the stars producing the $\sqrt{N}$ torques. In the limit of small test star mass, that influence tends to zero, and so a test particle code like TPI is correct.

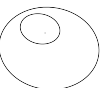
2.7.5 Location of the sign change of $\langle \Delta \ell \rangle$ at high ℓ

In the simulations presented in Sections 2.4 and 2.5, the first-order diffusion coefficient $\langle \Delta \ell \rangle$ was found to change sign from positive to negative values as ℓ increases towards unity. This is to be expected, considering that ℓ cannot exceed unity. The value of ℓ where this sign change occurs, ℓ_c , is well-defined in the simulations with $N_{\max} = 50$, but becomes less well-defined as $N_{\max}$ increases (cf. Fig. 2.7). In the S-star simulations, where $N_{\max}$ is greater still, ℓ_c is likewise not well-defined. A possible explanation for this trend with $N_{\max}$ is that for low $N_{\max}$ the number of data points in each bin at high ℓ is larger compared to this number at low ℓ , whereas for larger $N_{\max}$, the relative number of bins at high ℓ decreases. This is demonstrated in Fig. 2.24, where the number of data points in each bin of ℓ is plotted for the simulations of Section 2.4. The trend of the number of data points with $N_{\max}$ can be explained by the increase of the RR time-scale as $N_{\max}$ increases (cf. equation 2.14; note that here $M_\star$ is kept fixed): as $N_{\max}$ increases, the relative time spent at high ℓ in the simulations decreases, thereby decreasing the number of available data points.

The uncertainty of ℓ_c in simulations with large $N_{\max}$ is a caveat for our approximate analytic functions of the angular-momentum diffusion coefficients which depend on ℓ_c (cf. equation 2.48), and therefore for the analytic steady-state solutions (cf. Section 2.6.1). To explore the implications of this uncertainty, we show in Fig. 2.25 a figure similar to Fig. 2.18 for the steady-state solutions based on the analytic functions for the coefficients, but now also including a lower value of $\ell_{c,\text{II}} = 0.5$ and assuming the largest semimajor axis bin shown in Fig. 2.14. The latter bin is associated with large uncertainty in ℓ_c and we adopt $\ell_{c,\text{II}} = 0.5$ as an alternative value for ℓ_c for this semimajor axis. From Fig. 2.25 we conclude that the steady-state solutions for small ℓ ($\ell \ll \ell_c$) are not strongly affected by the uncertainty in ℓ_c . For larger ℓ the steady-state solution is sensitive to the value of ℓ_c , however. Nevertheless, our result that the maximum in the steady-state eccentricity distribution occurs near $e_c \equiv \sqrt{1 - \ell_c^2}$ is robust (cf. the third panel of Fig. 2.25).

2.7.6 Comparison of steady-state solutions

As mentioned in Section 2.6, MHL11 have previously investigated the effect of RR on the steady-state eccentricity distribution of stars near an SBH. MHL11 used a semi-empirical model and found a bimodal eccentricity distribution with two peaks at small semimajor axes (cf. fig. 18 of MHL11). Although we have also found a bimodal distribution with two peaks, the positions of these peaks are quite different in our work (cf. Fig. 2.21). In MHL11, the lower peak occurs at $e \sim 0.2$, whereas in our work the lower peak occurs at much higher eccentricity, $e \sim 0.7 - 0.8$. Furthermore, in MHL11 the upper peak occurs at $e \sim 0.9$, whereas in our work the upper peak occurs at even higher eccentricity, $e \sim 0.98$.



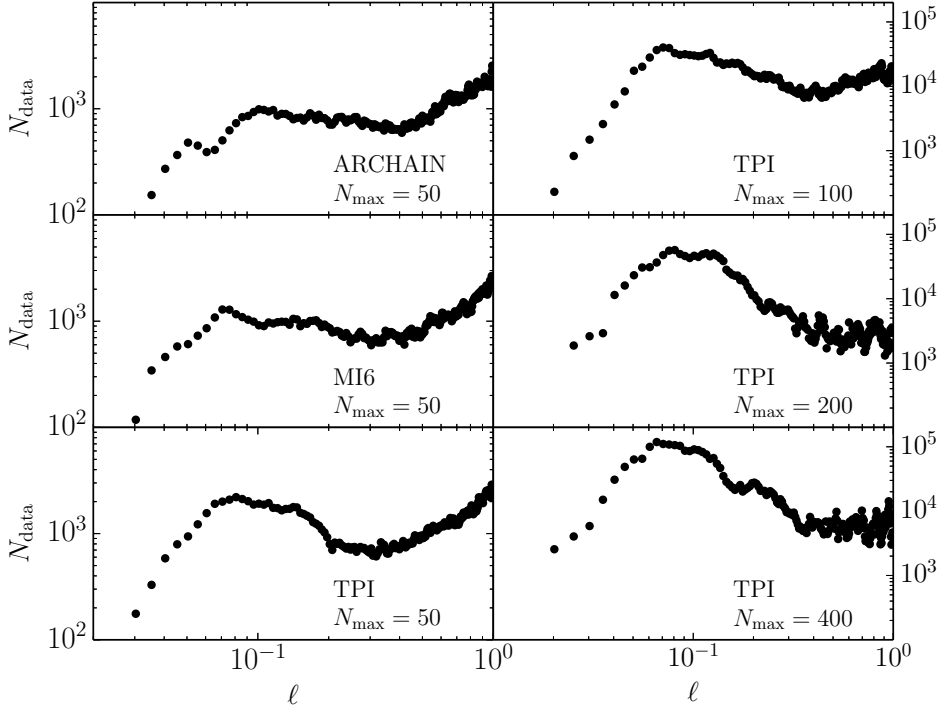


Figure 2.24: The number of data points N_{data} in each bin of ℓ and the time lag bin that was adopted in this paper, i.e. corresponding to the coherence time (cf. Fig. 2.7). Left-hand column: simulations with ARCHAIN, MI6 and TPI with $N_{\text{max}} = 50$. Right-hand column: simulations with TPI with $N_{\text{max}} = 100, 200$ and 400 .

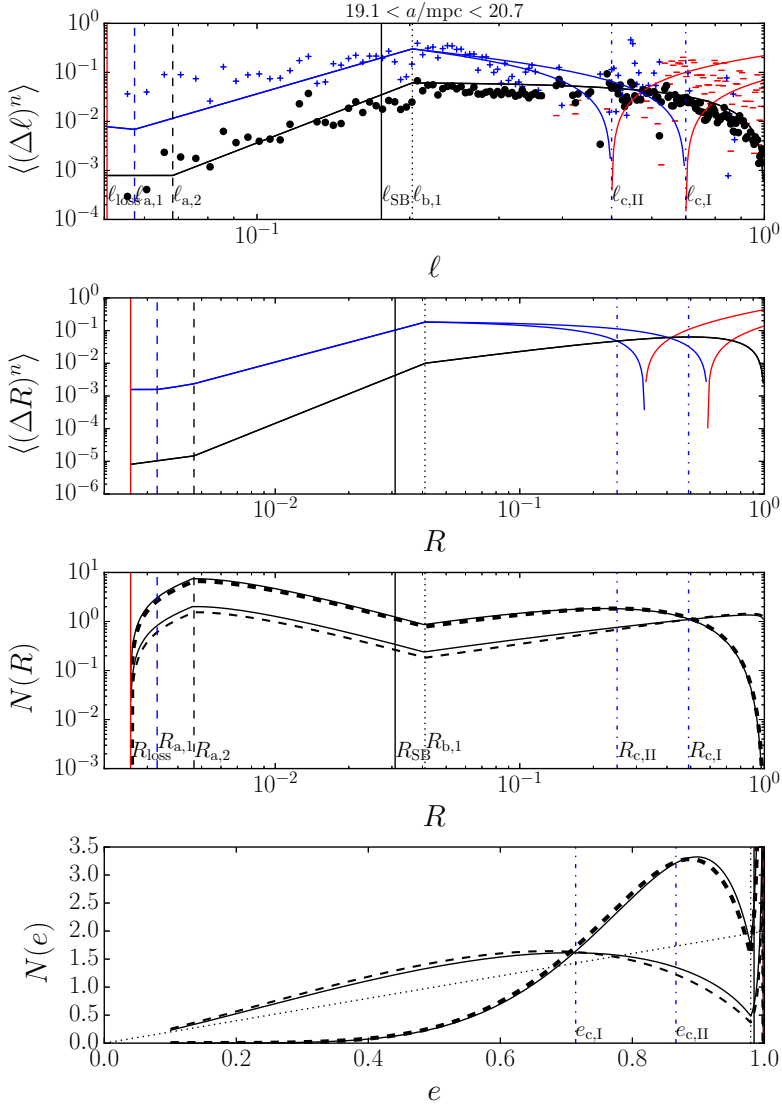
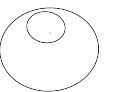


Figure 2.25: Steady-state solutions of the Fokker-Planck equation in angular-momentum space based on our analytic functions of the diffusion coefficients (cf. equation 2.48) similar to Fig. 2.18, but now assuming two values of ℓ_c : $\ell_{c,I} = 0.7$ (as before; thin black lines) and $\ell_{c,II} = 0.5$ (thick black lines), and for a larger semimajor axis bin.



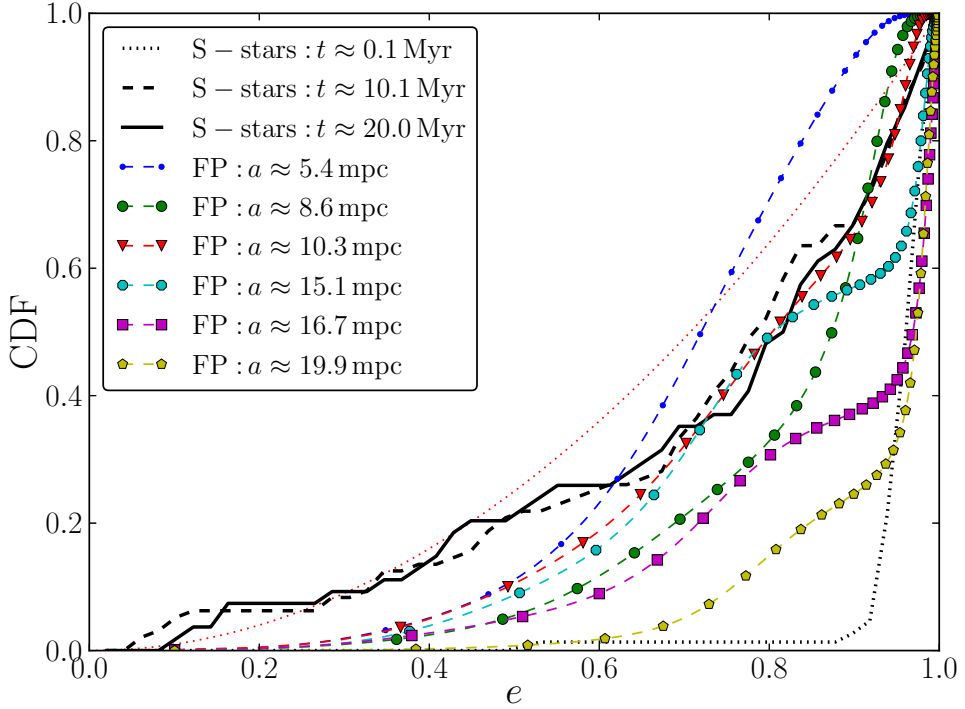


Figure 2.26: Cumulative eccentricity distributions obtained directly from the S-star simulations at three different times (black dotted, dashed and solid lines for $t \approx 0, 10$ and 20 Myr, respectively), and according to the steady-state solutions of the Fokker-Planck equation ($C > 0$; smoothness parameter $s = 0.5 s_0$ and method II, cf. Section 2.6.2). For the latter, the same semimajor axis bins are shown as in Fig. 2.21. The red dotted line shows a ‘thermal distribution’.

An important difference between our work and that of MHL11 is that in the latter general relativistic corrections in the equations of motion for the test stars were not taken into account, whereas these corrections were included here (cf. equation 2.5). This would suggest that these terms in the equations of motion tend to increase eccentricities in the steady-state distribution.

Finally, we briefly compare our distributions $N(e)$ obtained from solving the steady-state Fokker-Planck equation, equation (2.30), to the eccentricity distributions that we obtained directly from the N -body simulations of the S-stars in Section 2.5.2. One expects that the former apply in the limit $t \rightarrow \infty$. The latter are limited by the simulation time, in our case $t < 20$ Myr. We show both distributions in Fig. 2.26. In the case of the steady-state Fokker-Planck solutions, the semimajor axis bins are shown that were included in Figs 2.20 and 2.21; in the case of the distributions obtained directly from the N -body simulations, three times are shown. For semimajor axes that are comparable to the typical S-star semimajor axes, $a \sim 10$ mpc, the Fokker-Planck solutions are consistent with the direct S-star distributions for $t \gtrsim 10$ Myr at high eccentricities, $e \gtrsim 0.7$. For smaller eccentricities, the direct distributions are on average less eccentric than suggested by our steady-state solutions. The latter may be due to the follow-

ing reasons. (1) The simulated time of 20 Myr is too short for low-eccentricity orbits to reach a steady state. This would be consistent with the RR diffusion time-scale, which approaches $\sim 10^2$ Myr for $e \rightarrow 0$ in the assumed nuclear model (cf. Fig. 2.1). (2) At low eccentricities the diffusion coefficients that were obtained from the simulations suffer from large scatter. In particular, large scatter is present in the first-order diffusion coefficients shown in Fig. 2.14 at high ℓ , and this may produce bias in our results at low eccentricities.

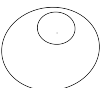
2.8 Conclusions

We have presented a new N -body algorithm, `TPI`, that allows us to efficiently integrate orbits of test stars around an SBH including PN corrections to the equations of motion and interactions with a large ($\gtrsim 10^3$) number of field stars. We compared results obtained with this code to results obtained with two more accurate, but slower, N -body codes `ARCHAIN` and `MI6`; we focused on the behaviour of orbits above and below the SB, the region in (energy, angular-momentum) space where relativistic precession vitiates torques from $\sqrt{N}$ asymmetries (RR). In addition, we have performed simulations of the GC to test models for the origin of the S-stars; these simulations used 4800 field stars close to the SBH, a number that is not currently feasible with other N -body algorithms. Our main conclusions are as follows.

1. The behaviour of test-particle orbits in `TPI` is consistent with what is found using the codes `ARCHAIN` and `MI6`, which do not make our simplifying assumptions.
2. We analysed several aspects of eccentricity oscillations below the SB ($e > e_{\text{SB}}$) that are associated with rapid GR precession in the presence of Newtonian torques from the field stars. Using power spectra of the eccentricity time series, we found evidence for enhanced power at higher integer frequencies than the relativistic frequency f_{GR} (Fig. 2.4). The peak at the latter frequency can be interpreted as implying that the torquing potential (due to the $\mathcal{O}(N^{-1/2})$ asymmetry in the field-star distribution) is basically lopsided, or $m = 1$, in character (MAMW11). Higher order terms in the multipole expansion of the field-star potential would give rise to eccentricity oscillations at higher integer frequencies of f_{GR} . Our results indicate that these higher order contributions are important, though typically not dominant.

In addition, we determined the amplitude of the eccentricity oscillations and we verified the expected dependence $\Delta\ell \propto \langle\ell\rangle^2$, where $\ell \equiv L/L_c \approx \sqrt{1-e^2}$ is the dimensionless angular momentum, $\Delta\ell$ is the amplitude of angular-momentum oscillations over a precessional cycle, and $\langle\ell\rangle$ is its average value. By fitting our data to the model of MAMW11, we also determined the fitting constant C_{A_D} that captures unspecified uncertainties in this model (Figs 2.5 and 2.15).

3. We applied the `TPI` algorithm to the evolution of the S-stars in the GC, assuming that they were deposited initially on to orbits of very high eccentricity. This is expected for the tidal disruption of a stellar binary. We adopted a distribution of field stars that is consistent with the steady-state distribution of stellar remnants at the GC. Assuming formation of S-stars in a burst, we found that their cumulative eccentricity distribution evolves to $N(e) \propto e^{2.6}$ on a time-scale of 7 ± 0.1 Myr. The latter distribution is consistent with observations. We also extrapolated our results to a continuous-formation model. Our results suggest a lower limit on the typical age of the S-stars of ~ 7 Myr in the case of burst formation and ~ 25 Myr in the case of continuous formation.



4. From our simulations, we extracted first- and second-order diffusion coefficients in the normalized angular-momentum variable ℓ . We identified three angular-momentum regimes, in which the diffusion coefficients depend in functionally different ways on ℓ . Regimes of lowest and highest ℓ are well described in terms of NRR and RR, respectively. Near and below the SB, a third regime exists, ‘anomalous relaxation’ (AR), which is not well described in terms of either NRR or RR. In this regime, the time-scale for angular-momentum diffusion increases rapidly with increasing eccentricity. We found that the features associated with the new AR regime are only clearly present in simulations with larger numbers of field stars than considered previously. We presented analytic expressions, in terms of physical parameters, that describe the diffusion coefficients in all three angular-momentum regimes.

5. We proposed a new, empirical criterion for the location of the barrier, based on the L -dependence of the diffusion coefficients. This criterion was found to predict essentially the same $\ell_{\text{SB}}(a)$ relation as equation (1) which was derived in MAMW11 from simple time-scale arguments. Our results also demonstrate the validity of that relation in systems that differ greatly in terms of particle number and mass.

6. We derived a simple expression for the typical time-scale of angular-momentum diffusion in the ‘anomalous’ (AR) regime (equation 2.25) and verified its correctness by applying it to the N -body simulations (cf. Fig. 2.16). We applied this relation assuming both a core of late-type stars and a cusp of stellar black holes in the GC, and confirmed the earlier result (AM13) that in the case of a core of late-type stars the time-scales for the S-stars to reach the observed ‘super-thermal’ distribution of eccentricities is much longer than the typical age of the S-stars (cf. Fig. 2.22).

7. Using our expressions for the angular-momentum diffusion coefficients, we derived the steady-state distribution of orbital angular momenta implied by the Fokker-Planck equation for stars near an SBH. This distribution differs significantly from the distribution predicted by NRR, $f(E, L) \propto f(E)$. There is a deficit of orbits near the SB and an excess just above it (i.e. $e < e_{\text{SB}}$). Furthermore, we found evidence for a local excess of orbits below the SB ($e > e_{\text{SB}}$) in a steady state, which can be attributed to the slow nature of diffusion in the AR regime, causing orbits to accumulate in this region.

8. Using our analytic expressions we derived an approximate relation for the maximum semi-major axis for which we expect AR to be important (cf. equation 2.36). This relation implies that AR is important in a large radial range for physically realistic nuclear star clusters.

Acknowledgements

We would like to thank M. Atakan Gürkan for making his Kepler solver freely available⁴, Jeroen Bédorf for invaluable help with implementing GPU acceleration in TPI using the SAPPORO library and Fabio Antonini for useful comments on the manuscript. We also thank the anonymous referee for providing comments that helped to improve the paper. We are grateful for the hospitality of the Institut Henri Poincaré where parts of this work were carried out. We also thank the organizers of the ‘Alájar Meeting 2013: Stellar dynamics and growth of massive black holes’ for a stimulating venue for discussions of issues related to this work. This

⁴The code can be downloaded from the web page http://home.strw.leidenuniv.nl/~gurkan/kepler/sol_kep/sol_kep.html.

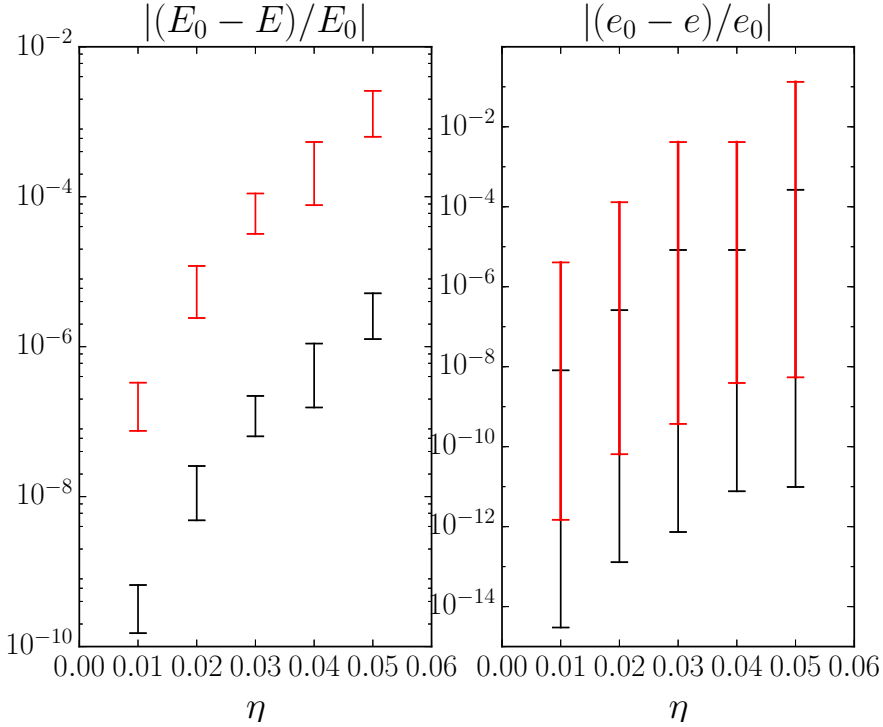


Figure 2.27: Accuracy of the TPI integrator for a single star orbiting an SBH; PN terms are excluded. For eight initial eccentricities $e_0 \in \{0.01, 0.1, 0.5, 0.9, 0.99, 0.999, 0.9999, 0.99999\}$, we show the range in the absolute values of the relative energy errors $(E_0 - E)/E_0$ (left-hand panel) and the relative eccentricity errors $(e_0 - e)/e_0$ (right-hand panel) as function of the time step parameter η . Black ranges are after 1000 orbits; red ranges are after 3.4×10^5 orbits.

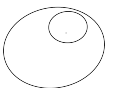
work was supported by the Netherlands Research Council NWO (grants #639.073.803 [VICI], #614.061.608 [AMUSE] and #612.071.305 [LGM]), the Netherlands Research School for Astronomy (NOVA), the National Science Foundation under grant no. AST 1211602 and the National Aeronautics and Space Administration under grant no. NNX13AG92G.

2.A Simple tests of TPI

2.A.1 Test stars orbiting the SBH

In the absence of PN and field-star perturbations, test stars should maintain fixed Kepler orbits about the SBH indefinitely. A sensitive test of the TPI algorithm is to check whether it conserves the Keplerian elements over many periods. For example, assuming constant a implies $t/P \approx 3.4 \times 10^5 (t/\text{Myr})(a/\text{mpc})^{-3/2}(M_\bullet/10^6 M_\odot)^{1/2}$ orbital revolutions after time t .

We initialized eight test stars in Kepler orbits around an SBH with $M_\bullet = 1.0 \times 10^6 M_\odot$



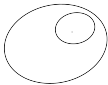
and initial semimajor axis $a_0 = 1$ mpc (this corresponds to an orbital period $P_0 \approx 2.96$ yr) and eight initial eccentricities $e_0 \in \{0.01, 0.1, 0.5, 0.9, 0.99, 0.999, 0.9999, 0.99999\}$. We integrated these stars with TPI for 3.4×10^5 orbital periods in the absence of PN terms and field stars. All orbital elements except the orbital phase should therefore remain constant. We show in Fig. 2.27 the relative energy errors $(E_0 - E)/E_0$ and the relative eccentricity errors $(e_0 - e)/e_0$ of the orbit around the SBH after 1000 orbits (black lines) and after 3.4×10^5 orbits (red lines). Energy and eccentricity errors are included for five values of the time step parameter η (cf. equation 2.6). The relative energy errors do not exceed 10^{-9} (10^{-5}) for $\eta = 0.01$ (0.05) after 1000 orbits and 10^{-6} (10^{-2}) for $\eta = 0.01$ (0.05) after 3.4×10^5 orbits. The relative eccentricity errors are larger than the relative energy errors but still do not exceed 10^{-3} for $\eta = 0.02$ after 3.4×10^5 orbits. We have made similar plots for the orbital angles i , ω and Ω and the relative errors of their cosines are $< 10^{-11}$ after 1000 orbits for $\eta \leq 0.05$ and $< 10^{-5}$ after 3.4×10^5 orbits for $\eta \leq 0.04$. The high precision in this test can be attributed to the use of regularization in the equations of motion (cf. Section 2.3).

Because we are interested in the regime where precession due to GR is important, we also tested the ability of the code to reproduce relativistic (Schwarzschild) precession of test stars. In the implementation of the 1PN terms, we assume that the SBH is fixed at the origin. To test whether this fixing of the SBH systematically affects the magnitude of secular precession expected from the theoretical expectation, equation (2.2), we integrated three different orbits with $a_0 = 0.1$ mpc and eccentricities $e_0 \in \{0.5, 0.9, 0.99\}$ with the 1PN terms included. The magnitude of pericentre shift during one orbit was computed by numerically determining the moments of two consecutive apocentre passages. By varying the numerically determined moments of apocentre by one output time, we obtained a measure of the error in the pericentre shift. We show the results in Table 2.3. The pericentre shift calculated with TPI is in very good agreement with the expected 1PN shift. For example, in the test with the smallest pericentre distance (corresponding to $r_p \approx 20.9 r_g$, where $r_g = GM_\bullet/c^2$ is the gravitational radius), the relative error is smaller than 0.005.

We also include in Table 2.3 results for the same test of the implementation of the 1PN terms carried out with the direct summation code M16 (Nitadori & Makino 2008; Iwasawa et al. 2011; see also Section 2.4.1). In the latter code, the SBH is assumed to be fixed as in TPI. The errors made in M16 should therefore be comparable to those of TPI. The close similarity of the errors made by the two codes suggests that this is indeed the case (cf. the last column of Table 2.3).

a/mpc	e	$\Delta\omega _{\text{sim}}/^{\circ}$		TPI	$\Delta\omega _{\text{pred}}/^{\circ}$		1PN error	
		MI6			MI6	TPI	MI6	TPI
0.1	0.5	0.69017 ± 0.00008		0.69010 ± 0.00010	0.68986	0.68979	-0.00046	-0.00045
0.1	0.9	2.75946 ± 0.00003		2.75913 ± 0.00002	2.75578	2.75548	-0.00135	-0.00133
0.1	0.99	30.141905 ± 0.000008		30.138717 ± 0.000005	30.28069	30.27620	0.00457	0.00454

Table 2.3: Pericenter shift per radial period for three different initial orbital configurations as computed by MI6 and TPI including 1PN terms. Shown are the values determined from the simulations ($\Delta\omega|_{\text{sim}}$), the predicted values computed from equation (2.2) based on a and e as determined in the simulations and the errors of the simulated values with respect to the predicted values, $(\Delta\omega|_{\text{pred}} - \Delta\omega|_{\text{sim}})/\Delta\omega|_{\text{pred}}$.



2.A.2 Test stars orbiting a field star that orbits the SBH

In TPI , we tested interactions between test and field stars by placing test stars in Kepler orbits around a single field star that orbits the SBH. The field star has a mass $m_\star = 1.0 M_\odot$ and its orbital parameters around the SBH ($M_\bullet = 1.0 \times 10^6 M_\odot$) are $a_\star = 10^6 \text{ AU} \approx 4.85 \text{ pc}$ and $e_\star = 0.01$. For the latter orbit, the star is situated far from the SBH at all times; hence, the orbital elements except the orbital phases of test stars orbiting the star should remain constant. The initial orbital elements of the test stars orbiting the field star are set to $a_0 = 1 \text{ AU}$ and $e_0 \in \{0.01, 0.1, 0.5, 0.9, 0.99\}$. The test stars are integrated for 100 orbital periods around the field star (i.e. 100 yr). We show in Fig. 2.28 the relative energy and eccentricity errors of the motion of the test stars around the field star. These errors are larger compared to those for the motion of test stars around the SBH, cf. Fig. 2.27. This is not surprising considering that in TPI the motion around the SBH is regularized, whereas the motion around the field stars is not. Nevertheless, after 100 orbital periods the relative energy error remains less than $\approx 10^{-3}$ even for rather eccentric ($e_0 = 0.99$) orbits.

We also tested interactions between test and field stars by initiating a test star and a field star in nearly circular and nearly intersecting orbits around the SBH. We set $\eta = 0.02$, $m_\star = 50 M_\odot$ and $M_\bullet = 1.0 \times 10^6 M_\odot$. The closest approach between the test and field star in the resulting interaction is $\approx 0.05 \text{ AU}$ and the orbit of the test star is strongly perturbed, i.e. $\Delta a/a \approx 0.85$ and $\Delta e/e \approx 8 \times 10^3$. We also computed the same encounter with the direct N -body code `HERMITE` implemented in `AMUSE` (Portegies Zwart et al. 2013). Here, the test star mass was set to $1.0 \times 10^{-30} M_\odot$. The discrepancies between the integrations with `HERMITE` and TPI are very small: the differences between the two codes in the test star semimajor axis and eccentricity after the strong encounter, are $(a_{\text{final}, \text{TPI}} - a_{\text{final}, \text{Hermite}})/a_{\text{final}, \text{TPI}} \approx 0.007$ and $(e_{\text{final}, \text{TPI}} - e_{\text{final}, \text{Hermite}})/e_{\text{final}, \text{TPI}} \approx 0.002$.

In TPI , decreasing η increases the number of integration steps and this decreases performance. Based on the above tests we have chosen the value $\eta = 0.02$ for the simulations presented in this paper, which we believe is a good compromise between accuracy and performance.

2.B Dependence of the NRR diffusion coefficients on γ

Here, we include an expression for the quantity $C_{\text{NRR}}(\gamma)$ for arbitrary γ . This expression appears in the NRR diffusion coefficients in which the limit $\ell \rightarrow 0$ was taken (cf. equations 2.12 and 2.13). It can be computed using the procedure described in appendix B of MAMW11 in which the potential of the stars was neglected. In that appendix, an explicit expression was derived for $\gamma = 2$,

$$C_{\text{NRR}}(2) = \frac{9}{7} \frac{1}{12 \log(2) - 1} \approx 0.175698. \quad (2.37)$$

We have derived an expression that is valid for arbitrary γ in the range $1/2 < \gamma < 3$,

$$C_{\text{NRR}}(\gamma) = \frac{3\pi}{64} \left[K_{1/2}(\gamma) - \frac{1}{5} K_{3/2}(\gamma) + \frac{5\pi}{8} \frac{1}{2\gamma - 1} \right]^{-1}. \quad (2.38)$$

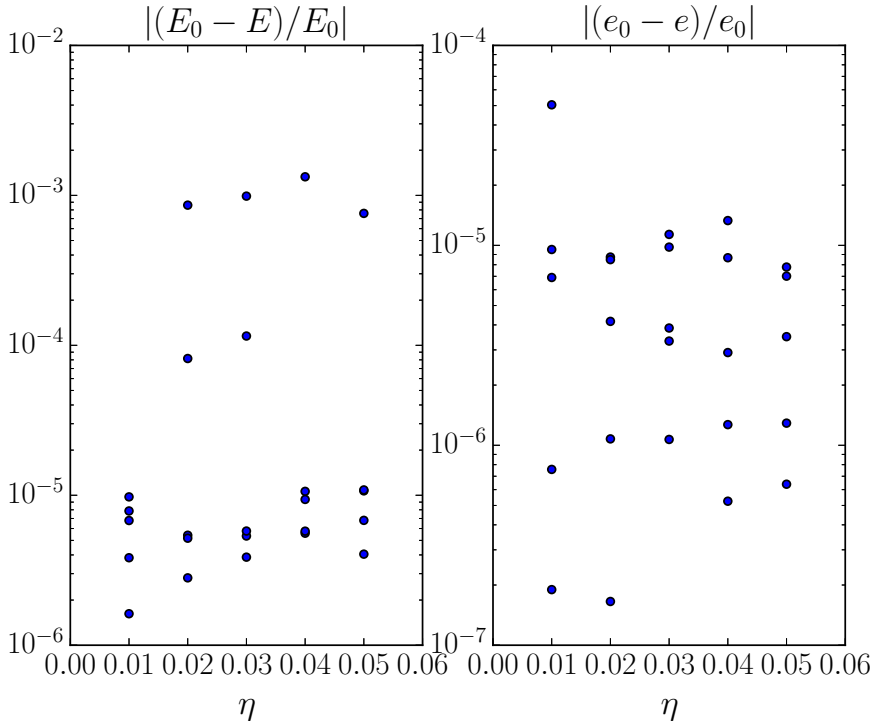
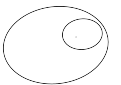


Figure 2.28: Similar to Fig. 2.27, now for test stars orbiting a single $1 M_{\odot}$ field star that orbits the SBH (without PN terms). The points correspond to the absolute values of the relative energy errors $(E_0 - E)/E_0$ (left-hand panel) and relative eccentricity errors $(e_0 - e)/e_0$ (right-hand panel) for five initial eccentricities $e_0 \in \{0.01, 0.1, 0.5, 0.9, 0.99\}$. Errors are shown after 100 orbits of the test stars around the field star.



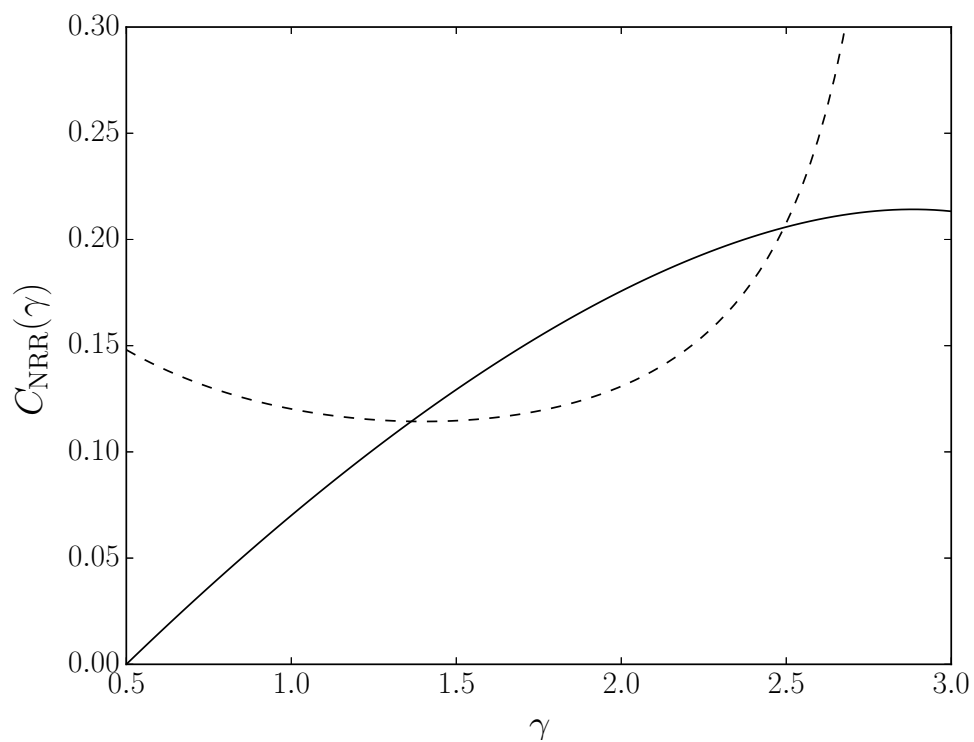


Figure 2.29: The quantity $C_{\text{NRR}}(\gamma)$ as a function of γ . Solid line: equation (2.38). Dashed line: a common approximation, $0.68/[(3 - \gamma)(1 + \gamma)^{3/2}]$.

Here, $K_{1/2}(\gamma)$ and $K_{3/2}(\gamma)$ are integral functions defined as

$$K_{1/2}(\gamma) = \int_0^1 x^{3-\gamma} \sqrt{1-x} {}_2F_1\left(\frac{3}{2}, \frac{3}{2} - \gamma, \frac{5}{2}, 1-x\right) dx; \quad (2.39a)$$

$$K_{3/2}(\gamma) = \int_0^1 x^{3-\gamma} \sqrt{1-x} {}_2F_1\left(\frac{5}{2}, \frac{3}{2} - \gamma, \frac{7}{2}, 1-x\right) dx, \quad (2.39b)$$

where ${}_2F_1(a, b, c; x)$ is the Gauss hypergeometric function. For $\gamma = 1$, equation (2.38) yields:

$$C_{\text{NRR}}(1) = \frac{1}{48 \log(2) - 19} \approx 0.0700719. \quad (2.40)$$

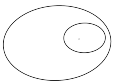
We show $C_{\text{NRR}}(\gamma)$ as function of γ in Fig. 2.29 (solid line). For reference we have also plotted in that figure with the dashed line a less accurate, but more common approximation, $0.68/[(3-\gamma)(1+\gamma)^{3/2}]$ (Merritt 2013a, p. 276) (both relations neglect the potential of the stars).

2.C Extrapolating the shape of the cumulative eccentricity distribution for the S-star simulations

Here, we present a method to extrapolate our results of $p(t)$ from the S-star simulations assuming burst formation, to the case of continuous formation (cf. Section 2.5.2). We assume that the probability density function (PDF) for a single S-star i is of the form $dN_i/de \propto e^{p_i(t-t_i)-1}$ for $t > t_i$ and $dN_i/de = 0$ for $t < t_i$. Here, t_i is the time at which star i is deposited. We assume that $t_i = xt_{\text{max}}$, where $x \in [0, 1]$ is a random number and t_{max} is a time-scale for which the upper limit is set by the MS lifetime of the S-star. Normalization of dN_i/de with $0 \leq e < 1$ yields $dN_i/de = p_i(t-t_i)e^{p_i(t-t_i)-1}$ for $t > t_i$. The PDF dN/de for the ensemble of $N_S = 19$ S-stars is composed of the PDFs for the individual S-stars and it is therefore given by the sum of the latter PDFs, i.e. $dN/de \propto \sum_{i=1}^{N_S} H(t-t_i) dN_i/de = \sum_{i=1}^{N_S} H(t-t_i) p_i(t-t_i) e^{p_i(t-t_i)-1}$, where H is the Heaviside step function. Normalization of the latter PDF with $0 \leq e < 1$ gives $dN/de = [\sum_{i=1}^{N_S} H(t-t_i)]^{-1} \sum_{i=1}^{N_S} [H(t-t_i) p_i(t-t_i) e^{p_i(t-t_i)-1}]$. This yields the following cumulative density function $N(e, t)$ for the ensemble of the S-stars at time $0 \leq t \leq t_{\text{max}}$:

$$N(e, t) = \left[\sum_{i=1}^{N_S} H(t-t_i) \right]^{-1} \sum_{i=1}^{N_S} \left[H(t-t_i) e^{p_i(t-t_i)} \right]. \quad (2.41)$$

It is assumed that $p_i(t-t_i) = p_{\text{fit}}(t-t_i)$, i.e. that the fitted curve $p_{\text{fit}}(t)$ in the burst scenario is representative for each individual S-star. For four different values of t_{max} , $t_{\text{max}}/\text{Myr} \in \{50, 100, 150, 200\}$, we take 100 different random realizations of t_i and for each realization we fit the CDF equation (2.41) to $N(e, t) = e^{p(t)}$. Subsequently, we average the curves $p(t)$ over the 100 different random realizations and we adopt the averaged curve as $p_{\text{cont}}(t)$, the value of p in the case of continuous formation. We show the results in Fig. 2.30 for the different values of t_{max} . As expected, $p(t)$ evolves more slowly for continuous formation of S-stars compared to formation in a burst (cf. Fig. 2.11); the speed of evolution is reduced by a factor of ~ 4 . The peaks are due to depositions of individual S-stars at the random times t_i . There is no strong dependence of $p_{\text{cont}}(t)$ on t_{max} ; after 50 Myr $p_{\text{cont}}(t)$ for the four values of t_{max} is nearly identical.



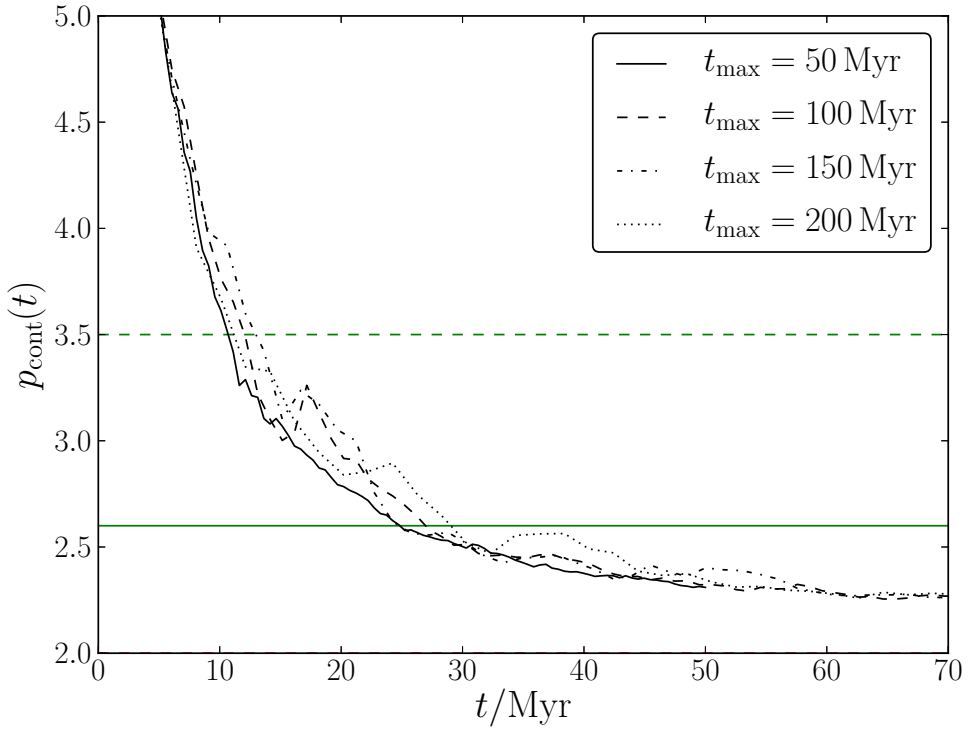


Figure 2.30: The evolution of the slope p with time extrapolated to the case of continuous formation of S-stars, determined using the fitted form $p_{\text{fit}}(t)$ shown in Fig. 2.11). Four different values are adopted for t_{max} , the maximum time of deposition. The green solid and dashed lines indicate the observed value $p = 2.6 \pm 0.9$ (Gillessen et al. 2009).

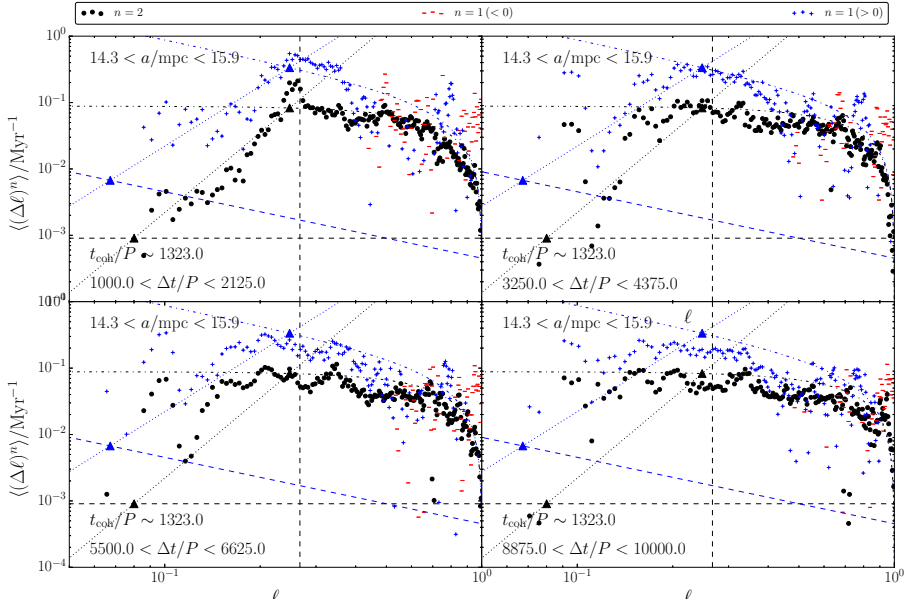


Figure 2.31: Diffusion coefficients obtained from the S-star simulations, similar to Fig. 2.14. Results are shown for a single semimajor axis bin and for different time lags.

2.D Dependence of derived diffusion coefficients on time lag in the simulations

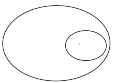
In Fig. 2.14, the time lags were chosen to match the coherence time t_{coh} . Here, we illustrate the importance of choosing the appropriate time lag, by showing in Fig. 2.31 an example of the dependence of the diffusion coefficients on time lag. We computed the coefficients for much longer time lags than in Section 2.5.4: $10^3 < \Delta t/P < 10^4$. For $\Delta t \sim t_{\text{coh}}$, the ‘knee’ feature (cf. Section 2.5.4) below the SB is clearly present; for $\Delta t \gg t_{\text{coh}}$ this feature gradually disappears. This can be understood from the argument that was presented in Section 2.4.5: below the SB, the time lag should not be much longer than t_{coh} , since for longer time lags the changes in ℓ become comparable to ℓ itself.

2.E Steady-state solutions to the Fokker-Planck equation

2.E.1 Solution of the steady-state equation

The aim is to solve equation (2.30), which we write as

$$-N(R)\mathcal{D}_1(R) + \frac{1}{2} \frac{\partial}{\partial R} [N(R)\mathcal{D}_2(R)] = C. \quad (2.42)$$



Here, we have used the notation $\langle \Delta R \rangle = \mathcal{D}_1$ and $\langle (\Delta R)^2 \rangle = \mathcal{D}_2$. We are interested in solutions $N(R)$ in the range $R_{\text{loss}} < R < R_{\text{up}}$. Here, $R_{\text{loss}} = r_{\text{capt}}/a (2 - r_{\text{capt}}/a) = \mathcal{O}(10^{-2})$ is the loss boundary and R_{up} is the largest value of R ; if the diffusion coefficients are completely known, then $R_{\text{up}} = 1$. If $C = 0$, then equation (2.42) can readily be integrated, with solution

$$\begin{aligned} N_{\text{H}}(R) &= N_{\text{H}}(R_{\text{loss}}) \exp \left[\int_{R_{\text{loss}}}^R \frac{2\mathcal{D}_1(R') - \mathcal{D}_2'(R')}{\mathcal{D}_2(R')} dR' \right] \\ &\equiv N_{\text{H}}(R_{\text{loss}}) g(R). \end{aligned} \quad (2.43)$$

Here, $\mathcal{D}'_n(R) \equiv (\partial/\partial R)\mathcal{D}_n(R)$ and we have implicitly defined the function $g(R)$; note that $g(R_{\text{loss}}) = 1$. The function $N_{\text{H}}(R)$ is a homogeneous solution to equation (2.42). To find the inhomogeneous solution, we apply the method of variation of constants and write $N(R) = N_{\text{H}}(R)N_{\text{I}}(R)$. Substituting the latter into equation (2.42), we find

$$C = -N_{\text{H}}N_{\text{I}}\mathcal{D}_1 + \frac{1}{2} \frac{\partial}{\partial R} (N_{\text{H}}N_{\text{I}}\mathcal{D}_2) \quad (2.44a)$$

$$= N_{\text{I}} \left[-N_{\text{H}}\mathcal{D}_1 + \frac{1}{2} \frac{\partial}{\partial R} (N_{\text{H}}\mathcal{D}_2) \right] + \frac{1}{2} N_{\text{H}}\mathcal{D}_2 \frac{\partial N_{\text{I}}}{\partial R} \quad (2.44b)$$

$$= \frac{1}{2} N_{\text{H}}\mathcal{D}_2 \frac{\partial N_{\text{I}}}{\partial R}. \quad (2.44c)$$

The last step is by virtue of equation (2.42) with $N = N_{\text{H}}$ and $C = 0$. Equation (2.44) is readily integrated,

$$N_{\text{I}}(R) = \int_{R_{\text{loss}}}^R \frac{2C}{N_{\text{H}}(R')\mathcal{D}_2(R')} dR' + C_1. \quad (2.45)$$

Here, C_1 is an integration constant. By imposing $N(R_{\text{loss}}) = 0$ and substituting the solutions equations (2.43) and (2.45), we find $C_1 = 0$. The general solution is therefore given by:

$$N(R) = 2Cg(R)I(R), \quad (2.46a)$$

$$g(R) = \exp \left[\int_{R_{\text{loss}}}^R \frac{2\mathcal{D}_1(R') - \mathcal{D}_2'(R')}{\mathcal{D}_2(R')} dR' \right], \quad (2.46b)$$

$$I(R) = \int_{R_{\text{loss}}}^R \frac{dR'}{\mathcal{D}_2(R')g(R')}. \quad (2.46c)$$

As expected for a second-order differential equation, the solution to equation (2.46) contains two parameters, R_{loss} and C (we do not consider R_{up} to be a free parameter). By imposing an additional constraint on the solution, the number of parameters is reduced by one. For example, requiring that $N(R)$ is normalized to unit total number, i.e. $\int N(R) dR = 1$, we find for the flux in terms of R_{loss} :

$$C = \frac{1}{2} \left[\int_{R_{\text{loss}}}^{R_{\text{up}}} g(R)I(R) dR \right]^{-1}. \quad (2.47)$$

2.E.2 Analytic expressions for the diffusion coefficients

For completeness, we give the explicit functional expressions for our approximation of the diffusion coefficients, equation (2.23),

$$\langle \Delta \ell \rangle = \begin{cases} 1/(4\ell t_{N1}), & \ell_{\text{loss}} < \ell < \ell_{a,1}; \\ C_1 \ell^3 / \tau, & \ell_{a,1} \leq \ell < \ell_{b,1}; \\ C_1 \ell_{b,1}^4 / (\tau \ell) \frac{\ell_c^2 - \ell^2}{\ell_c^2 - \ell_{b,1}^2}, & \ell_{b,1} \leq \ell \leq 1; \end{cases} \quad (2.48a)$$

$$\langle (\Delta \ell)^2 \rangle = \begin{cases} 1/(t_{N1}), & \ell_{\text{loss}} < \ell < \ell_{a,1}; \\ C_2 \ell^4 / \tau, & \ell_{a,1} \leq \ell < \ell_{b,1}; \\ (1 - \ell^2) \alpha_s^2 / t_{R1}, & \ell_{b,1} \leq \ell \leq 1. \end{cases} \quad (2.48b)$$

Here, $t_{N1} \equiv A(E)^{-1}$ (cf. equation 2.13), $\tau \equiv t_{\text{coh}}/A_D^2$ (cf. equation 2.22) and $t_{R1} \equiv [M_\bullet/M_\star(a)]^2 N_\star(a) P(a)^2 / t_{\text{coh}}$ (cf. equation 2.14). In equation (2.48), the first-order diffusion coefficient in the range $\ell_{b,1} < \ell < 1$ has been modified to account for negative $\langle \Delta \ell \rangle$ for $\ell > \ell_c$, as described in Section 2.6.1. A comparison of equation (2.48) to N -body data is given in Fig. 2.18.

2.E.3 Explicit analytic steady-state solutions

We derive explicit expressions for the steady-state distribution function $N(R)$ for the analytic functions of the diffusion coefficients presented in equation (2.48). First we transform $\langle (\Delta \ell)^n \rangle$ to $\langle (\Delta R)^n \rangle$ using the transformations (Cohn 1979; Merritt 2013a, eq. 5.167)

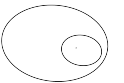
$$\begin{cases} \langle \Delta R \rangle = 2\ell \langle \Delta \ell \rangle + \langle (\Delta \ell)^2 \rangle; \\ \langle (\Delta R)^2 \rangle = 4\ell^2 \langle (\Delta \ell)^2 \rangle. \end{cases} \quad (2.49)$$

We subsequently substitute $\langle (\Delta R)^n \rangle$ into equation (2.46). Here, we assume that $\ell_{b,1} = \ell_{b,2}$, which is the case if $C_1 = C_2$. The result is

$$\bar{N}(R) = 2 C g(R) I(R); \quad (2.50a)$$

$$g(R) = \begin{cases} g_0, & R_{\text{loss}} < R < R_{a,1}; \\ g_1(R), & R_{a,1} \leq R < R_{a,2}; \\ g_2(R), & R_{a,2} \leq R < R_{b,1}; \\ g_3(R), & R_{b,1} \leq R < 1; \end{cases} \quad (2.50b)$$

$$I(R) = \begin{cases} I_0(R), & R_{\text{loss}} < R < R_{a,1}; \\ I_0(R_{a,1}) + I_1(R), & R_{a,1} \leq R < R_{a,2}; \\ I_0(R_{a,1}) + I_1(R_{a,2}) + I_2(R), & R_{a,2} \leq R < R_{b,1}; \\ I_0(R_{a,1}) + I_1(R_{a,2}) + I_2(R_{b,1}) \\ \quad + I_3(R), & R_{b,1} \leq R < 1. \end{cases} \quad (2.50c)$$



The auxiliary functions are given by

$$\begin{aligned}
 g_0 &= 1; \\
 g_1(R) &= g_0 \left(\frac{R}{R_{a,1}} \right)^{-\frac{1}{2}} \exp [c_N (R^2 - R_{a,1}^2)]; \\
 g_2(R) &= g_1(R_{a,2}) \left(\frac{R}{R_{a,2}} \right)^{\frac{2C_1 - 5C_2}{2C_2}}; \\
 g_3(R) &= g_2(R_{b,1}) \left(\frac{R}{R_{b,1}} \right)^{c_R c_B R_c - \frac{1}{2}} \left(\frac{1 - R_{b,1}}{1 - R} \right)^{c_R c_B (R_c - 1) + 1}
 \end{aligned} \tag{2.51}$$

and

$$\begin{aligned}
 I_0(R) &= \frac{t_{N1}}{2} \log \left(\frac{R}{R_{\text{loss}}} \right); \\
 I_1(R) &= \frac{t_{N1}}{4} R_{a,1}^{-\frac{1}{2}} c_N^{-\frac{1}{4}} \exp (c_N R_{a,1}^2) \\
 &\quad \times \left[\Gamma \left(\frac{1}{4}, c_N R_{a,1}^2 \right) - \Gamma \left(\frac{1}{4}, c_N R^2 \right) \right]; \\
 I_2(R) &= \frac{\tau}{2C_2} \left(\frac{R_{a,2}}{R_{a,1}} \right)^{\frac{1}{2}} \exp [c_N (R_{a,1}^2 - R_{a,2}^2)] \\
 &\quad \times \frac{C_2}{2C_1 - C_2} R_{a,2}^{-2} \left[1 - \left(\frac{R}{R_{a,2}} \right)^{\frac{C_2 - 2C_1}{2C_2}} \right]; \\
 I_3(R) &= \frac{t_{R1}}{2\alpha_s^2} \left(\frac{R_{a,2}}{R_{a,1}} \right)^{\frac{1}{2}} \exp [c_N (R_{a,1}^2 - R_{a,2}^2)] \left(\frac{R_{b,1}}{R_{a,2}} \right)^{\frac{5C_2 - 2C_1}{2C_2}} \\
 &\quad \times (1 - R_{b,1})^{-c_R c_B (R_c - 1) - 1} (2c_R c_B R_c - 1)^{-1} \\
 &\quad \times \left[\left(\frac{R}{R_{b,1}} \right)^{c_R c_B R_c - \frac{1}{2}} \times {}_2F_1 (d_1, d_2, d_3; R) \right. \\
 &\quad \left. - {}_2F_1 (d_1, d_2, d_3; R_{b,1}) \right].
 \end{aligned} \tag{2.52}$$

Here, $c_N \equiv (t_{N1}/\tau)C_1$, $c_R \equiv (t_{R1}/\tau)C_1$, $c_B \equiv (R_{b,1}^2/\alpha_s^2)[1/(R_c - R_{b,1})]$, $\Gamma(s, x) = \int_x^\infty t^{s-1} \exp(-t) dt$ is the upper incomplete Gauss function and ${}_2F_1(d_1, d_2, d_3; R)$ is the Gauss hypergeometric function, with $d_1 \equiv c_R c_B (1 - R_c)$, $d_2 \equiv c_R c_B R_c - \frac{1}{2}$ and $d_3 = c_R c_B R_c + \frac{1}{2}$.

2.F N -body simulations with $\gamma = 1$

In the simulations presented in Section 2.5, a field-star density profile $\rho_*(r) \propto r^{-2}$ was assumed. In order to establish whether the ‘knee’ feature in the diffusion coefficients that can be associated with AR is also present in simulations with different γ , we have carried out an

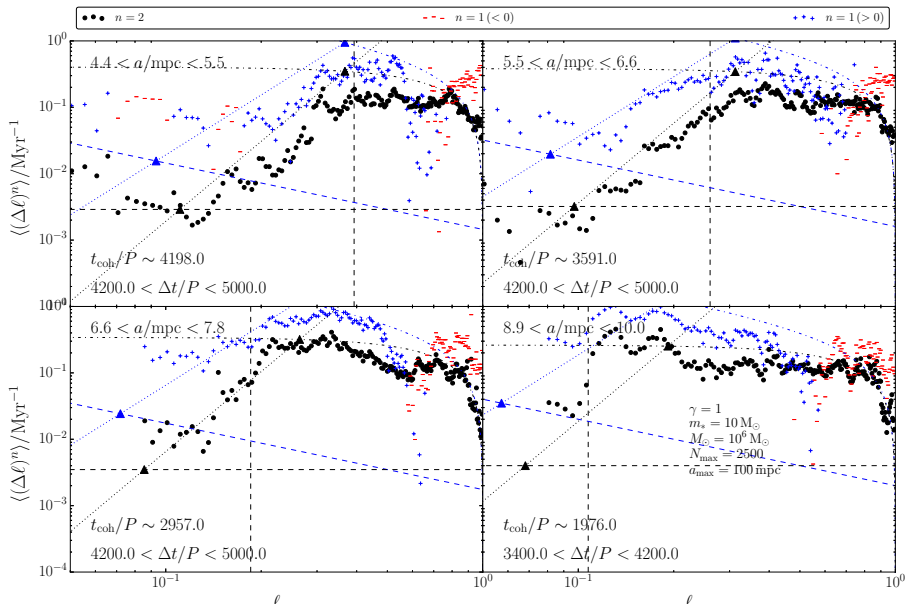


Figure 2.32: Diffusion coefficients obtained from an additional set of simulations with $\gamma = 1$; parameters are indicated in the bottom-right panel. The lines show the analytic model of equation (2.23) as in Fig. 2.14, but now evaluated for the corresponding nuclear model with $\gamma = 1$ (see Section 2.7.3 for details).

additional set of simulations with TPI with $\gamma = 1$. These additional simulations provide verification of some of our expectations for the regime in which AR is important, as discussed in Section 2.7.3.

The parameters of the additional set of simulations were as follows. The field-star mass was $m_\star = 10 M_\odot$ and the SBH mass was $M_\bullet = 10^6 M_\odot$. Field stars were distributed according to $N(a) = N_{\text{max}}(a/a_{\text{max}})^{3-\gamma}$, with $N_{\text{max}} = 2500$, $a_{\text{max}} = 100 \text{ mpc}$ and $\gamma = 1$, and their eccentricities were sampled from a ‘thermal’ distribution $dN/de = 2e$. In total 200 test particles were included, with initial semimajor axes sampled from $N(a) \propto a^{3-\gamma}$ with $3 \lesssim a/\text{mpc} \lesssim 14$ and $dN/de = 2e$. The orbits of the test and field stars were initially randomly oriented. The capture radius was $r_{\text{capt}} = 8 r_g$ and the integration time was 10 Myr. Only the 1PN terms were included.

The diffusion coefficients obtained from these simulations are shown for different semimajor axes in Fig. 2.32. In that figure, we have included the same analytical functions for the coefficients that were also included in Fig. 2.14 (cf. equation 2.23), but now evaluated for the model with $\gamma = 1$. The results for $\gamma = 1$ are consistent with those for $\gamma = 2$, which were presented in Fig. 2.14. In particular, the ‘knee’ feature is clearly present which, as we argued, can be associated with the rapid quenching of RR below the SB. The position of this ‘knee’ agrees well with the predicted position of the SB, equation (2.1), suggesting that this relation is also valid for nuclear models with $\gamma = 1$.

For small semimajor axes, our predictions for the AR diffusion coefficients with $C_1 = C_2 \approx 2.6$ (cf. equation 2.22) are in good agreement with the data obtained from the simulations with $\gamma = 1$. At larger semimajor axes, the slopes predicted by these relations are still consistent with the data, but the normalization is not: it appears that in order to remain consistent with the data, both C_1 and C_2 must *increase* with increasing semimajor axis. We note that this trend can also be observed in Fig. 2.14, although the dependence of C_1 and C_2 on semimajor axis appears to be weaker in the latter figure.

2.G Equivalence of two critical radii

Here, we show that the quantity $a_{\text{AR,max}}$ defined in Section 2.7.3 is the same as the critical semimajor axis that was defined in section VC of MAMW11. The latter quantity, which we here denote by a_{MAMW} , was argued to be the minimum value of a for which NRR would allow orbits to ‘penetrate’ the SB.

The criterion in MAMW11 was that – for orbits near the SB –

$$(\Delta\ell)_{\text{NRR}} \equiv \left(\frac{t_{\text{coh}}}{t_{\text{NRR}}} \right)^{1/2} \gtrsim \ell_{\text{max}} - \ell_{\text{min}} \approx 2\ell_{\text{av}}^2 A_{\text{D}} \quad (2.53)$$

(MAMW11, equations 66, 67). Thus, a_{MAMW} is the value of a for which

$$\frac{t_{\text{coh}}}{t_{\text{NRR}}} \approx 4A_{\text{D}}^2 \ell_{\text{SB}}^4. \quad (2.54)$$

The quantity defined as t_{NRR} in MAMW11 is essentially the inverse of the quantity $A(E)$ defined in this paper (cf. equation 2.13).

In Section 2.7.3 of this paper, $a_{\text{AR,max}}$ was defined as the value of a for which $\ell_{\text{a},1} = \ell_{\text{SB}}$. The quantity $\ell_{\text{a},1}$ was defined, in turn, as the angular momentum for which

$$\langle \Delta \ell \rangle_{\text{NRR}} \equiv \frac{1}{4\ell} A(E) = \langle \Delta \ell \rangle_{\text{AR}} \approx \frac{A_{\text{D}}^2 \ell^3}{t_{\text{coh}}} \quad (2.55)$$

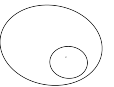
(equation 2.23). Thus,

$$\ell_{\text{a},1}^4 \approx \frac{t_{\text{coh}}}{A_{\text{D}}^2} \frac{A(E)}{4} \quad (2.56)$$

and setting $\ell_{\text{a},1} = \ell_{\text{SB}}$ then yields

$$4A_{\text{D}}^2 \ell_{\text{SB}}^4 \approx A(E)t_{\text{coh}}, \quad (2.57)$$

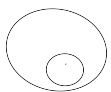
the same as equation (2.54).



3 | Probing the formation of planetesimals in the Galactic Centre using Sgr A* flares

Flares in X-ray and near-infrared are observed above the quiescent emission of the supermassive black hole (SBH) in the Galactic Centre (GC) at a rate of approximately once per day. One proposed energy source for these flares is the tidal disruption of planetesimals with radius $\gtrsim 10$ km passing within ~ 1 AU of the SBH. Very little is known about the formation and evolution of planetesimals in galactic nuclei such as the GC, making predictions for flaring event rates uncertain. We explore two scenarios for the formation of planetesimals in the GC: (1) in a large-scale cloud bound to the SBH, and (2) in debris discs around stars. We model their orbital evolution around the SBH using the Fokker-Planck equation and investigate the effect of gravitational interactions with various relevant perturbers. Our predicted flaring rate, $\approx 0.6 \text{ d}^{-1}$, is nearly independent of the distribution of perturbers. Moreover, it is insensitive to scenarios (1) or (2). The assumed number of planetesimals per star is consistent with debris discs around stars in the Solar neighbourhood. In scenario (1), this implies that the number of planetesimals formed in the large-scale cloud is strongly correlated with the number of stars, and this requires finetuning for our results to be consistent with the observed flaring rate. We favour the alternative explanation that planetesimals in the GC are formed in debris discs around stars, similar to the Solar neighbourhood.

*Hamers, A.S. & Portegies Zwart, S.F. MNRAS **446**, 1, 710-729 (2015)*



3.1 Introduction

The Galactic Centre (GC) contains a supermassive black hole (SBH) of mass $M_{\bullet} \approx 4 \times 10^6 M_{\odot}$ (Ghez et al. 2008; Gillessen et al. 2009). For the last decade, observations in the near-infrared and X-ray of the central region of the GC have revealed the existence of flares, occurring approximately once per day (Baganoff et al. 2001, 2003; Genzel et al. 2003a; Dodds-Eden et al. 2011; Barrière et al. 2014). These flares are 3-100 times more luminous than the quiescent emission of the central radio source Sgr A*, which is known to be very dim: its bolometric luminosity is only $\sim 10^{-8.5} L_{\text{Edd}}$ (Genzel et al. 2010). Observations indicate that the region from which the flares originate is very compact, extending no more than a few tens of gravitational radii $r_g = GM_{\bullet}/c^2$ from the SBH (Baganoff et al. 2001; Genzel et al. 2003a; Porquet et al. 2003; Shen et al. 2005).

Several explanations for the energy source and emission mechanism of the flares have been proposed. One of these is the tidal disruption of planetesimals with radius $\gtrsim 10$ km (Zubovas et al. 2012, hereafter ZNM12). ZNM12 showed that if the latter passes within ~ 1 AU of the SBH, it is broken up into smaller fragments by tidal forces; the fragments subsequently vapourize because of friction with the ambient gas. When the vapourized material is mixed with the accretion flow on to Sgr A*, enough energy could be released to produce an observable flare with X-ray luminosity $\sim 10^{34} - 10^{35}$ erg/s. In ZNM, it was assumed that the planetesimals were formed in debris discs around stars and were subsequently stripped by the tidal force of the SBH and by gravitational encounters with other stars. The resulting event rate of the flares was estimated using loss cone refilling arguments and a rate was found that is consistent with the observed rate of approximately once per day. The method employed by ZNM12 to calculate the flaring rate was very approximate; details about the distribution of stars in the GC and the spatial and temporal dependence of the stripping of planetesimals were not taken into account.

Very little is known about the formation and evolution of such planetesimals in galactic nuclei like the GC. One possibility is that they are formed in a large-scale spherical cloud orbiting the SBH. Another possibility is that they are born in debris discs around stars (see e.g. Wyatt 2008 for a review), and are stripped by the tidal force of the SBH or gravitational encounters with other stars (e.g. Nayakshin et al. 2012). In this paper, we investigate both scenarios by means of numerical integrations of the Fokker-Planck equation. We model the orbital energy evolution around the SBH, taking into account the effects of gravitational perturbations from late-type stars, a possible cusp of stellar black holes close to the SBH and giant molecular clouds further away, as well as the effects of physical collisions.

We will show that the predicted present-day disruption rates in the GC differ very little between the two scenarios and that this conclusion depends weakly on the details of the perturbers. In both scenarios, we find a disruption rate of $\sim 1 \text{ d}^{-1}$ assuming that the number of planetesimals per (late-type) star is $N_{a/\star} = 2 \times 10^7$. The number $N_{a/\star} = 2 \times 10^7$ is consistent with debris discs observed around stars in the Solar neighbourhood. In the first scenario, in which the planetesimals are formed in a large cloud, this implies that the number of bodies formed is strongly correlated with the number of stars, and this requires finetuning of the quantity $N_{a/\star}$. We favour the more natural explanation that planetesimals in galactic nuclei similar to the GC are formed in debris discs around stars, no differently than planetesimals around stars in the Solar neighbourhood.

The structure of this paper is as follows. In Section 3.2, we describe our models of the GC

and of debris discs around stars, in Section 3.3 we describe the stripping process and in Section 3.4 we present our integrations of the Fokker-Planck equation and the implied disruption rates. We extrapolate our results to different galactic nuclei in Section 3.5 and briefly discuss a special case that would lead to a burst of flares. We conclude in Section 3.6.

3.2 Setting the stage

Before describing the effects of the stripping of planetesimals from stars (Section 3.3) and presenting our main results of their orbital evolution around the SBH (Section 3.4), we discuss our models of the GC (Section 3.2.1) and of the adopted initial debris discs around stars (Section 3.2.2).

3.2.1 Models of the GC

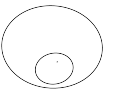
Here, we describe and motivate the models of the GC on which the calculations in Section 3.4 are based. For further details on the implementation of these models in the integrations of the Fokker-Planck equation (Section 3.4) we refer to Appendix 3.B.

Late-type stars

The majority (96%) of the observed stars in the GC are old (> 10 Gyr) late-type stars (Genzel et al. 2010). Number counts indicate that the 3D number density $n(r)$ of these stars approximately scales with the distance to the SBH as $n(r) \propto r^{-1.8}$ for $r \gtrsim 0.3$ pc, whereas there is a much flatter, possibly even declining density profile inside ~ 0.3 pc (Eckart et al. 1993; Genzel et al. 1996; Schödel et al. 2007; Trippe et al. 2008; Schödel et al. 2009; Oh et al. 2009; Do et al. 2013). There is evidence that at large radii the density drops more steeply than $n(r) \propto r^{-1.8}$, i.e. $n(r) \propto r^{-3}$ for $r \gtrsim 5$ pc (Schödel et al. 2014).

Most of the late-type stars likely formed ~ 10 Gyr ago, possibly coincident with the Galactic bulge (Blum et al. 2003; Maness et al. 2007; Pfuhl et al. 2011). From theoretical arguments these stars are expected to be distributed in a cusp, with number density $n(r) \propto r^{-7/4}$, after approximately a relaxation time-scale (Bahcall & Wolf 1976). This is consistent with the observed distribution of late-type stars at radii $0.3 \text{ pc} \lesssim r \lesssim 5 \text{ pc}$. At smaller radii, however, a core is observed. The nature of this core is not fully understood, but a possible explanation is that the relaxation time-scale in the GC at the radius of influence is likely > 10 Gyr, therefore the core could reflect the primordial population of the late-type stars (Merritt 2010). Merritt (2010) showed that if there initially was a core of late-type stars of size ~ 1 pc, this would have evolved to the observed size of ~ 0.3 pc today.

It is conceivable that the observed core of late-type stars arises from observational bias: within 0.3 pc the stellar light is dominated by bright early-type stars (see below), which complicates observations of late-type stars. To accommodate both possibilities (i.e. a core versus a cusp), we consider the following two (spherically symmetric) models of the present-day number



density $n_{\text{LT}}(r)$ of late-type stars in the GC (‘LT1’ and ‘LT2’),

$$n_{\text{LT}}(r) = \begin{cases} n_b \left(\frac{r}{r_b}\right)^{-\gamma_i} \left[1 + \left(\frac{r}{r_b}\right)^\alpha\right]^{\frac{\gamma_i - \gamma}{\alpha}} \left(1 + \frac{r}{r_o}\right)^{-\gamma_o}; & \text{(LT1)} \\ n_b \left(\frac{r}{r_b}\right)^{-\gamma} \left(1 + \frac{r}{r_o}\right)^{-\gamma_o}. & \text{(LT2)} \end{cases} \quad (3.1)$$

Here, the fixed parameters are $\gamma_i = 0.7$, $\gamma = 1.8$, $\gamma_o = 1.2$, $\alpha = 4$, $r_b = 0.3 \text{ pc}$ and $r_o = 5 \text{ pc}$. The values of the parameters γ_i , γ , α are adopted from Merritt (2010), who fitted the surface density as a function of the projected radius to the number counts of late-type stars in the sample of Buchholz et al. (2009, we adopt $\gamma_i = 0.7$ rather than $\gamma_i = 0.5$ for computational reasons). For LT1, and at small radii the number density has a slope of $-\gamma_i$, which turns over to a slope of $-\gamma$ at intermediate radii; the parameter α determines the smoothness of this transition. For both LT1 and LT2 and at larger radii $r \gtrsim r_o$ the slope $\approx -(\gamma + \gamma_o) = -3$, which is consistent with the observed slope at these larger radii (Schödel et al. 2014).

As shown by Merritt (2010) the break radius r_b may have been larger in the past. In particular, the present-day core can be explained by assuming an initial core of size $r_b \sim 1 \text{ pc}$. We will show in Section 3.4 that our results of the disruption rates at 10 Gyr vary little between models LT1 (with a core) and LT2 (without a core), indicating that the details of the core are not important for the purposes of this work.

We model the late-type stars as a single-mass stellar population with $m_{\text{LT}} = 1 M_\odot$. The normalization n_b in equation (3.1) is determined by equating the enclosed late-type stellar mass implied by this equation at a reference radius of $r_0 \equiv 1 \text{ pc}$, to the inferred distributed mass within r_0 of $M_0 \approx 1.5 \times 10^6 M_\odot$ as determined by Schödel et al. (2009). Here, we neglect the contribution of early-type stars and stellar black holes (discussed below). The latter populations have a total mass of $\sim 10^4 M_\odot$, which can safely be neglected compared to the total mass of the $\sim 10^6$ late-type stars within r_0 . We subsequently find $n_b \approx 1.7 \times 10^6 \text{ pc}^{-3}$ and $n_b \approx 1.4 \times 10^6 \text{ pc}^{-3}$ for LT1 and LT2, respectively.

The (negative of the) gravitational potential $\psi(r)$ is required for the calculations presented below. It is computed from equation (3.1) using the inverted Poisson equation (e.g. Cohn 1979; Merritt 2013b, 3.2)

$$\psi(r) = \frac{GM_\bullet}{r} + \frac{4\pi G m_{\text{LT}}}{r} \int_0^r dr' r'^2 n_{\text{LT}}(r') - 4\pi G m_{\text{LT}} \int_0^r dr' r' n_{\text{LT}}(r'). \quad (3.2)$$

Here, the contributions of early-type stars, stellar black holes and massive perturbers (see below) to the potential are neglected. In addition we define the arbitrary constant in the potential such that $\psi(r) \rightarrow GM_\bullet/r$ as $r \rightarrow 0$ assuming a power-law dependence of $n_{\text{LT}}(r)$ with $\gamma < 2$ for small r , which is consistent with equation (3.1).

Early-type stars

In addition to the late-type stars, $\sim 10^2$ young massive stars are observed within the central parsec of the GC, of which a fraction of ~ 0.2 is distributed within at least one disc structure

with an inner edge at projected radius ≈ 0.8 arcsec or ≈ 32 mpc (Paumard et al. 2006; Bartko et al. 2009; Lu et al. 2009). The age of these stars is $\sim 2 - 6$ Myr, and evidence exists for multiple formation events in the past, with an interval of ~ 100 Myr (Blum et al. 2003). Within the central arcsecond ($\lesssim 0.04$ pc), ~ 20 less massive ($3 \lesssim m/M_\odot \lesssim 15$) stars have been observed (Genzel et al. 2003a; Eisenhauer et al. 2005; Ghez et al. 2008; Gillessen et al. 2009). These stars, the S-stars, are likely older than $2 - 6$ Myr, and their formation process so close to the SBH is unclear, although disruption of a stellar binary through the Hills mechanism (Hills 1988) is currently favoured (see Alexander 2005 and Genzel et al. 2010 for reviews).

Their short lifetime, relatively low total mass ($\sim 10^4 M_\odot$) and limited radial extent make it unlikely that early-type stars are dynamically important for most planetesimals in the GC. Nevertheless, debris discs around early-type stars could be an important source of planetesimals on much longer time-scales ($> \text{Gyr}$) considering that planetesimals, once stripped by the tidal force of the SBH (cf. Section 3.3.1), would accumulate over time while early-type stars are formed episodically.

We estimate the latter contribution by the following source term (cf. equation 3.9)

$$F_{\text{strip,ET}}(\mathcal{E}) = N_{\text{a}/\star} N_{\text{ET}}(\mathcal{E}) / \tau_{\text{SF}}. \quad (3.3)$$

Here, $N_{\text{ET}}(\mathcal{E})$ is the number of early-type stars with orbital energies between $\mathcal{E}$ and $\mathcal{E} + d\mathcal{E}$, where the orbital energy $\mathcal{E} = -v^2 + \psi(r)$ is defined with respect to the SBH (note that the early-type stellar mass is not included in $\psi(r)$). The quantity $\tau_{\text{SF}} = 100$ Myr is an estimate of the time-scale at which the early-type stars are formed. In equation (3.3) it is assumed that all planetesimals are stripped within the lifetime of the early-type stars, which is justified considering the narrow radial extent of the early-type stars. Because likely not all planetesimals can be stripped within the lifetime of the early-type stars, equation (3.3) provides an upper limit for the stripping flux. Furthermore, it is assumed in equation (3.3) that over longer time-scales $\gg \tau_{\text{SF}}$, the source term can be considered constant.

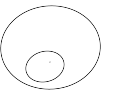
To compute $N_{\text{ET}}(\mathcal{E})$, we assume that the surface density of the early-type stars is $\Sigma_{\text{ET}}(r) \propto r^{-n}$ with $n = 2$ (e.g. Bartko et al. 2009), which implies $N_{\text{ET}} \propto \mathcal{E}^{n-3}$. Subsequently, $N_{\text{ET}}(\mathcal{E})$ is normalized by setting the total number of early-type stars formed in each formation event equal to 100 (roughly the current observed number) and assuming a distribution in energy bounded by $\mathcal{E}_1 < \mathcal{E} < \mathcal{E}_2$, where $\mathcal{E}_i = GM_\bullet / (2a_i)$, and $a_1 = 0.4$ pc and $a_2 = 0.04$ pc.

Stellar black holes

A population of stellar black holes distributed in a cusp close to the SBH could be dynamically important for planetesimals in the GC. Although so far not directly supported by observations, there are various theoretical motivations for the existence of such a population.

(1) Theoretical models of the GC based on the Fokker-Planck equation predict a dense population of stellar black holes close to the SBH arising from mass segregation (Hopman & Alexander 2006b).

(2) In the case of formation of the S-stars through the Hills mechanism, the predicted eccentricities are on average higher than the observed eccentricities of the S-stars. Dynamical evolution by field stars through resonant relaxation (RR) can modify the high eccentricities to a distribution that is consistent with observations. When assuming a core of late-type stars the time-scale of this relaxation process is much longer than the age of the S-stars, however. This problem can be remedied when assuming a cusp of stellar black holes; in this case the relaxation time-scale



is consistent with the age of the S-stars (Perets et al. 2009), in particular if relativistic effects are taken into account (Antonini & Merritt 2013; Antonini 2014; Hamers et al. 2014).

(3) If $\sim 10-100$ early-type stars in the GC are produced every ~ 100 Myr and each produce a stellar black hole $\sim 10 M_\odot$, then a black hole population with a total mass of $\sim 10^4-10^5 M_\odot$ could be produced over a time span of 10 Gyr (e.g. Merritt 2010).

We model a cusp of stellar black holes by assuming a total of 4800 stellar black holes, each $10 M_\odot$, distributed with an approximate number density $n(r) \propto r^{-7/4}$ within $r_d \equiv 0.2$ pc of the SBH, similar to what was assumed in Antonini & Merritt (2013). We assume that the stellar black hole distribution function vanishes for $\mathcal{E} < \mathcal{E}_d$, where $\mathcal{E}_d \equiv GM_\bullet/(2r_d)$, and neglect the effect of the black holes on the potential $\psi(r)$. We also include the possibility of a time dependence of the black hole density as implied by argument (3) above by assuming, for simplicity, that the black hole density increases linearly with time from 0 at $t = 0$ to the maximum density corresponding to $N_{\text{BH}} = 4800$ at $t = 10$ Gyr.

Massive perturbers

Massive perturbers such as giant molecular clouds can strongly enhance the rate of relaxation in galactic nuclei Perets et al. (2007, hereafter PHA07). They may therefore be relevant for the supply of planetesimals to the loss cone of the SBH. To evaluate the potential effect on the relaxation rate, we adopt the model GMC1 from PHA07 in which the effect on the relaxation rate is the largest, therefore giving an upper limit for this effect. We take into account the effect of the massive perturbers by multiplying the inverse relaxation time-scale by μ_2 , the ratio of the second moments of the mass distributions of massive perturbers and stars, at the two radial bins included in table 2 of PHA07. Here, we neglect the smaller Coulomb factor in the case of massive perturbers.

3.2.2 Disc models

In our second formation scenario for planetesimals in galactic nuclei, it is assumed that they are born in debris discs around stars (cf. Section 3.1). We assume a thin disc with surface density $\Sigma(d) \propto d^{-\beta}$, where d is the distance from the star, and adopt $\beta = 1.5$ (Hayashi 1981). The initial radial extent of the disc is $d_1 < d < d_2$, where two radial distributions are adopted for the disc as in Lestrade et al. (2011): (1) $d_1 = 40$ AU and $d_2 = 100$ AU (disc model 1; ‘DM1’) and (2) the more tightly bound distribution $d_1 = 10$ AU and $d_2 = 40$ AU (disc model 2; ‘DM2’). These two choices reflect uncertainties in the dependence of the radial extent of the disc on the mass of the parent star.

Based on the bulk energy and assuming a spherical shape, ZNM12 estimated that a planetesimal of radius $\gtrsim R_a \approx 10$ km could produce an observable flare ($L_X \sim 10^{34}$ erg/s) when tidally disrupted by the SBH. Assuming a distribution of radii $dN_a/dR_a \propto R_a^q$ with $q \approx -3.5$ (Wyatt 2008) and density $\rho_a = 1 \text{ g cm}^{-3}$, the number of planetesimals per star with radius $\geq R_a$ is $N_{a/\star} \sim 2 \times 10^7 [m_{a,\text{tot}}/(10^{-5} M_\odot)]$, where $m_{a,\text{tot}}$ is the total mass in planetesimals per star (ZNM12). Note that $m_{a,\text{tot}}$ is the most uncertain parameter in $N_{a/\star}$. Below we use the term ‘planetesimals’ to refer to planetesimals with radius $\geq R_a$.

We define the stripping radius d_{strip} as the maximum distance from a star for which a planetesimal, treated as a massless particle, is still bound to this star. The fraction f_{strip} of planetesimals that is stripped from a single star for a given d_{strip} and disc model (i.e. d_1 , d_2 and β) is

given by

$$f_{\text{strip}}(d_{\text{strip}}) = \begin{cases} 1, & d_{\text{strip}} < d_1; \\ \frac{d_2^{2-\beta} - d_{\text{strip}}^{2-\beta}}{d_2^{2-\beta} - d_1^{2-\beta}}, & d_1 \leq d_{\text{strip}} \leq d_2; \\ 0, & d_{\text{strip}} > d_2, \end{cases} \quad (3.4)$$

and the number of stripped planetesimals per star is $f_{\text{strip}} N_{a/\star}$. In Section 3.3, we consider two causes for stripping: the tidal force of the SBH (Section 3.3.1) and gravitational encounters with other stars (Section 3.3.2).

3.3 Stripping planetesimals from stars in the GC

3.3.1 Stripping by the SBH

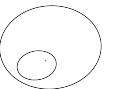
We determine d_{strip} (cf. Section 3.2.2) for planetesimals that are initially bound to a star with mass $m_\star$, that in turn is bound to the SBH in an orbit with pericentre distance r_p . A simple estimate of d_{strip} by the tidal force of the SBH is given by the radius of the Hill sphere (e.g. Hamilton & Burns 1992);

$$d_{\text{strip}} \approx r_p \left(\frac{m_\star}{3M_\bullet} \right)^{1/3}. \quad (3.5)$$

In the top panel of Fig. 3.1, the stripping radii according to equation (3.5) are plotted as a function of r_p . Two masses are adopted for the parent star: $1 M_\odot$ (thin line) and $50 M_\odot$ (thick line). The radial extents of the disc models DM1 and DM2 are indicated with the blue and red regions, respectively; these extents are used to compute the stripping fractions $f_{\text{strip}}(r_p)$ (cf. equation 3.4), and are shown as a function of r_p in the middle panel of Fig. 3.1. For both stellar masses and disc models, the SBH is ineffective at stripping planetesimals from stars with pericentre distances $\gtrsim 0.1$ pc. There is a strong dependence of $f_{\text{strip}}(r_p)$ on both the disc models and the mass of the parent star. For instance, the pericentre distance for which $f_{\text{strip}} = 0$ is ≈ 2.5 times larger for DM1 compared to DM2 (assuming $m_\star = 1 M_\odot$). The qualitative behaviour is intuitively easy to understand: in DM2 the planetesimals are more tightly bound to the parent star, therefore the parent star needs to be closer to the SBH for the planetesimals to be stripped. Similarly, more massive parent stars also need to be closer to the SBH for effective stripping of planetesimals by the tidal force of the SBH.

As a verification of equation (3.5), we carried out a series of N -body simulations of a star with mass $1 M_\odot$ orbiting the SBH. The star is orbited by a debris disc consisting of 100 particles of mass $2.0 \times 10^{-15} M_\odot$ in circular, coplanar orbits and with semimajor axes ranging between 10 and 1000 AU. The pericentre distance of the stellar orbit r_p is varied between 0.01 and 1 pc. The eccentricity e of the latter orbit is assumed to be either 0.01 or 0.9 and the semimajor axis is computed from $a = r_p/(1 - e)$. The system is integrated for the duration of an orbit of the star around the SBH with the `HERMITE0` code (Hut et al. 1995) in the `AMUSE` framework (Pelupessy et al. 2013; Portegies Zwart et al. 2013).

From each N -body simulation, we determined the orbital elements of the debris disc particles with respect to the star, and recorded which particles become unbound from the star at some



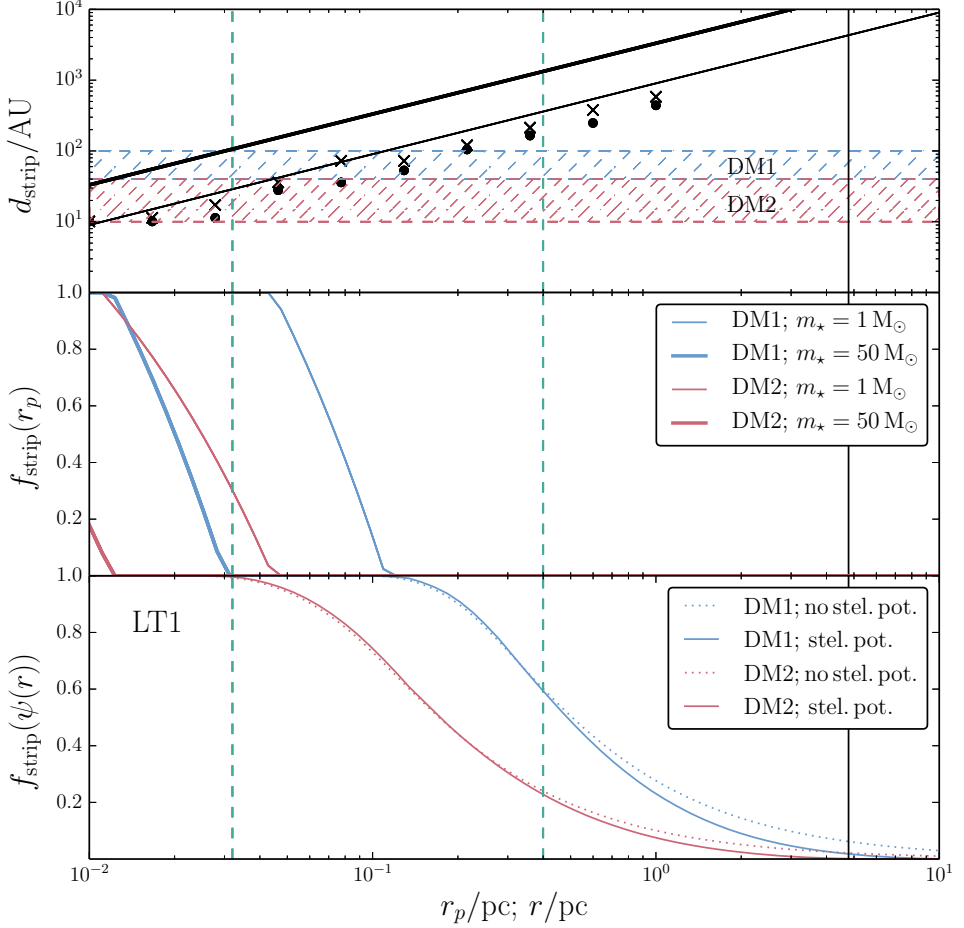


Figure 3.1: **Top panel:** the distance d_{strip} for which planetesimals are stripped from their parent star by the tidal force of the SBH as a function of r_p , the pericentre distance of the orbit of the star around the SBH, according to equation (3.5). Thin black line: assuming $m_\star = 1 M_\odot$ (i.e. late-type stars); thick black line: assuming $m_\star = 50 M_\odot$ (i.e. early-type stars). The horizontal regions indicate the two radial extents of planetesimals around stars considered in this paper (cf. Section 3.2.2): DM1 (blue) and DM2 (red). The green dashed vertical lines indicate the approximate region of the young massive stars in the GC. Black bullets (crosses): stripping radii determined from N -body simulations assuming $e = 0.01$ ($e = 0.9$), see the text. **Middle panel:** the stripping fraction as a function of r_p , computed from d_{strip} using equation (3.4). Blue (red) lines apply to DM1 (DM2). Thin lines: assuming $m_\star = 1 M_\odot$; thick lines: assuming $m_\star = 50 M_\odot$. **Bottom panel:** the angular-momentum-averaged stripping fraction (assuming an isotropic velocity distribution) as a function of $r = \psi^{-1}(\mathcal{E})$, assuming LT1. Solid lines: including the stellar potential; dashed lines: excluding the stellar potential (cf. Appendix 3.A.1). In all panels the black solid vertical line indicates the radius of influence r_h assuming LT1, defined as $M_{\text{LT}}(r_h) = 2M_\bullet$ (cf. equation 3.7).

point in the integration. For the latter particles, the minimum of the initial pericentre distances with respect to the star was adopted as the stripping radius d_{strip} .

In the top panel of Fig. 3.1, the stripping radii inferred from the N -body simulations are shown with the black bullets (crosses) assuming the stellar orbit has $e = 0.01$ ($e = 0.9$). The stripping radii determined from the N -body simulations are slightly lower compared to those implied by equation (3.5); the discrepancy is largest for the nearly circular orbit (black bullets). Nevertheless, the discrepancy, averaging the stripping radii from the N -body simulations for the two cases $e = 0.01$ and $e = 0.9$, is no larger than a factor of ≈ 1.9 . Therefore, we believe equation (3.5) is adequate for the purposes of this paper.

In equation (3.5), the stripping fraction f_{strip} is expressed in terms of the pericentre distance of the star. For the purposes of Section 3.4, we also compute f_{strip} in terms of the orbital energy $\mathcal{E}$ assuming an isotropic velocity distribution, and taking into account the stellar potential. The details are included in Appendix 3.A.1.

In the bottom panel of Fig. 3.1, the quantity $f_{\text{strip}}(\mathcal{E})$ is plotted as a function of radius by setting $\mathcal{E} = \psi(r)$, i.e. $r = \psi^{-1}(\mathcal{E})$, where $\psi^{-1}(\mathcal{E})$ is the inverse function of $\psi(r)$, and assuming late-type model LT1. Here, we include two cases: with (solid lines) and without (dotted lines) the stellar potential; in the latter case, analytic expressions can be derived for $f_{\text{strip}}(\mathcal{E})$ (cf. Appendix 3.A.1). As expected, at small r it is a good approximation to neglect the stellar potential; at larger r , however, neglect of the stellar potential causes the stripping fraction to be slightly overestimated.

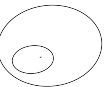
Compared to the case, if f_{strip} is expressed in terms of the pericentre distance r_p (cf. the middle panel of Fig. 3.1), if expressed in terms of orbital energy (cf. the third panel of Fig. 3.1), it is non-zero for a much larger radial range. This is because for a given r_p there exist many (eccentric) orbits with typical radius $r \gg r_p$.

3.3.2 Stripping by gravitational encounters with other stars

A planetesimal bound to a star orbiting the SBH can be treated as a binary system orbiting the SBH. Because of gravitational encounters with other stars, the intrinsic binary parameters, in particular the semimajor axis a_{bin} , change over time. A quantity that describes whether a_{bin} on average increases or decreases is the hardness parameter h , the ratio of the (negative) specific binding energy of the binary, $\mathcal{E}_{\text{bin}} = Gm_{\text{bin}}/(2a_{\text{bin}})$, to the squared stellar velocity dispersion, $\sigma^2(r)$ (Heggie 1975). Here, $m_{\text{bin}} \equiv m_* + m_a \approx m_*$. If $h \ll 1$, the binary is soft and gravitational encounters on average cause such binaries to become softer (i.e. to decrease h), until they dissolve as $h \rightarrow 0$ (e.g. Heggie 1975; Hut 1983; Hut & Bahcall 1983). On the other hand, if $h \gg 1$, the binary is hard and encounters tend to make it even harder (e.g. Heggie 1975; Hut 1993; Heggie et al. 1996).

To investigate which of these regimes applies to planetesimals in the GC, we compute $\sigma(r)$ from the isotropic Jeans equation, and assume that the number density and mass are dominated by late-type stars,

$$n_{\text{LT}}(r)\sigma^2(r) = \int_r^\infty dr' \frac{GM(r')n_{\text{LT}}(r')}{r'^2}, \quad (3.6)$$



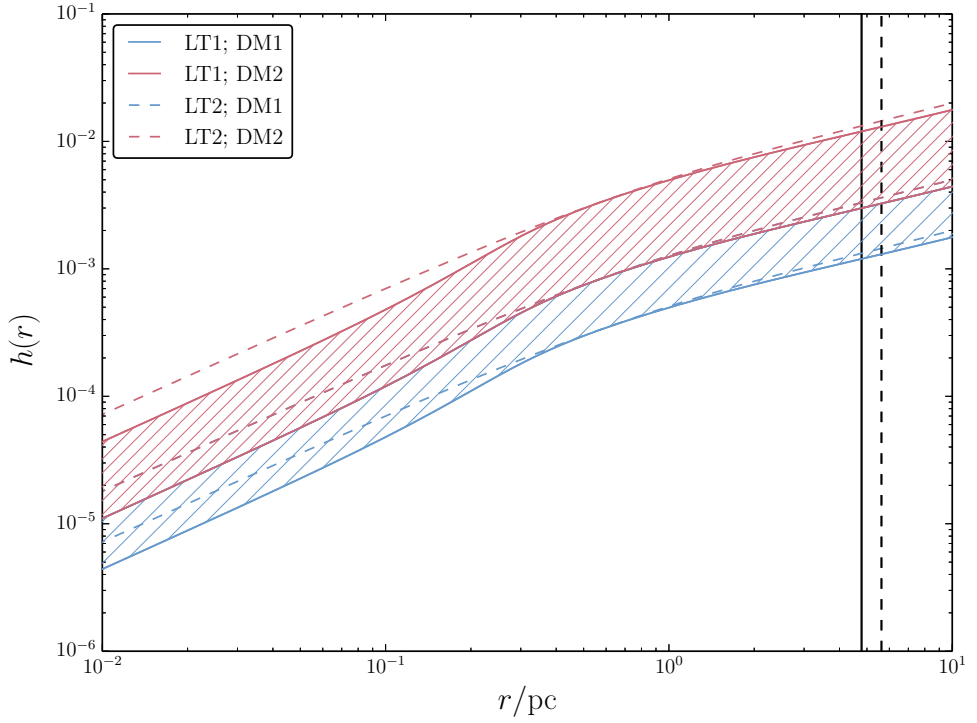


Figure 3.2: The hardness parameter h as a function of the distance r to the SBH for planetesimals distributed around the star according to DM1 (blue regions) and DM2 (red regions). The velocity dispersion is given by equation (3.6) and $m_{\text{bin}} = 1 M_{\odot}$ is assumed. Solid (dashed) lines apply to LT1 (LT2). Vertical lines: radius of influence in the models LT1 (solid) and LT2 (dashed).

where $M(r) = M_\bullet + M_{\text{LT}}(r)$ and

$$M_{\text{LT}}(r) = 4\pi m_{\text{LT}} \int_0^r n_{\text{LT}}(r') r'^2 dr'. \quad (3.7)$$

Adopting the semimajor axes from the disc models DM1 and DM2 and assuming $m_{\text{bin}} = m_{\text{LT}} = 1 M_\odot$, the hardness parameters that follow from equation (3.6) are plotted in Fig. 3.2 for LT1 (solid lines) and LT2 (dashed lines). For all combinations of disc models and distributions of late-type stars, $h < 10^{-2}$ for $r < r_h$, which shows that in our models the planetesimal+star binary is initially always soft. Therefore, gravitational encounters on average soften these binaries even further, to the point that eventually the planetesimals are no longer bound to their parent star.

For soft binaries, the stripping time-scale, i.e. the time-scale for the orbital energy of the binary to change by order itself because of gravitational encounters, can be estimated by (Binney & Tremaine 2008; Perets 2009; Alexander & Pfuhl 2014)

$$t_{\text{strip}} \equiv \frac{|\mathcal{E}_{\text{bin}}|}{\langle \mathcal{D}(\Delta \mathcal{E}_{\text{bin}}) \rangle} = \frac{1}{8} \sqrt{\frac{1+q_\sigma}{2\pi q_\sigma}} \frac{m_{\text{bin}} \sigma(r)}{G n(r) \langle m_\star^2 \rangle a_{\text{bin}} \log(\Lambda_{\text{bin}})}. \quad (3.8)$$

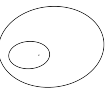
Here, $\langle \mathcal{D}(\Delta \mathcal{E}_{\text{bin}}) \rangle$ is a diffusion coefficient for the binary binding energy, $q_\sigma \equiv m_{\text{bin}}/m_\star \approx 1$, $\langle m_\star^2 \rangle$ is the second moment of the stellar mass function, and Λ_{bin} is the Coulomb factor for evaporation in the soft limit, which can be estimated by $\Lambda_{\text{bin}} \sim 3(1+1/q_\sigma)/(1+2/q_\sigma)[\sigma^2(r)/v_{\text{bin}}^2]$, where $v_{\text{bin}}^2 \equiv Gm_{\text{bin}}/a_{\text{bin}}$ (Alexander & Pfuhl 2014).

We assume that the parent star has mass $m_{\text{bin}} \approx m_{\text{LT}} = 1 M_\odot$ and we consider encounters with late-type stars only, such that $\langle m_\star^2 \rangle = m_{\text{LT}}^2$, $n(r) = n_{\text{LT}}(r)$ and σ is given by equation (3.6). Evidently, encounters with early-type stars also lead to evaporation. However, at the radial positions of the majority of the early-type stars in the GC, the SBH is already effective at stripping most planetesimals (cf. Fig. 3.1). Furthermore, because massive perturbers are extended objects they do not contribute to stripping (PHA07). Therefore, we expect, in the case of the GC, that stripping by encounters at larger radii is dominated by encounters with late-type stars.

Using equation (3.8), we estimate the stripping semimajor axis d_{strip} because of gravitational encounters with late-type stars as follows (here, we assume that the orbits of the planetesimals around the star are circular). For a given distance r to the SBH and age t of the star+planetesimal binary system, the stripping time-scale t_{strip} is equated to t . The resulting equation is solved numerically for $a_{\text{bin}} = d_{\text{strip}}$ (note that the logarithmic term $\log(\Lambda_{\text{bin}})$ depends on a_{bin}).

In the top panel of Fig. 3.3, we show the resulting d_{strip} as a function of the distance r of the centre of mass of the binary system, for the models LT1 (black solid lines) and LT2 (black dashed lines) and assuming different ages t . At small radii, the stripping radius is larger for LT1 than for LT2. This can be attributed to the lower number density at small radii in the case of LT1, implying less efficient stripping.

We show in the middle panel of Fig. 3.3 the stripping fraction as a function of r at various ages, indicated with numbers, for LT1 and assuming DM1 (blue lines) and DM2 (red lines). There is a strong dependence of $f_{\text{strip}}(r)$ on time. For instance, after 1 Myr and for DM1 the fraction at 0.4 pc, ≈ 0.5 , is comparable to the fraction in the case of stripping by the tidal force of the SBH alone (cf. the bottom panel of Fig. 3.1), whereas after ~ 100 Myr stripping has crept up all the way to the radius of influence.



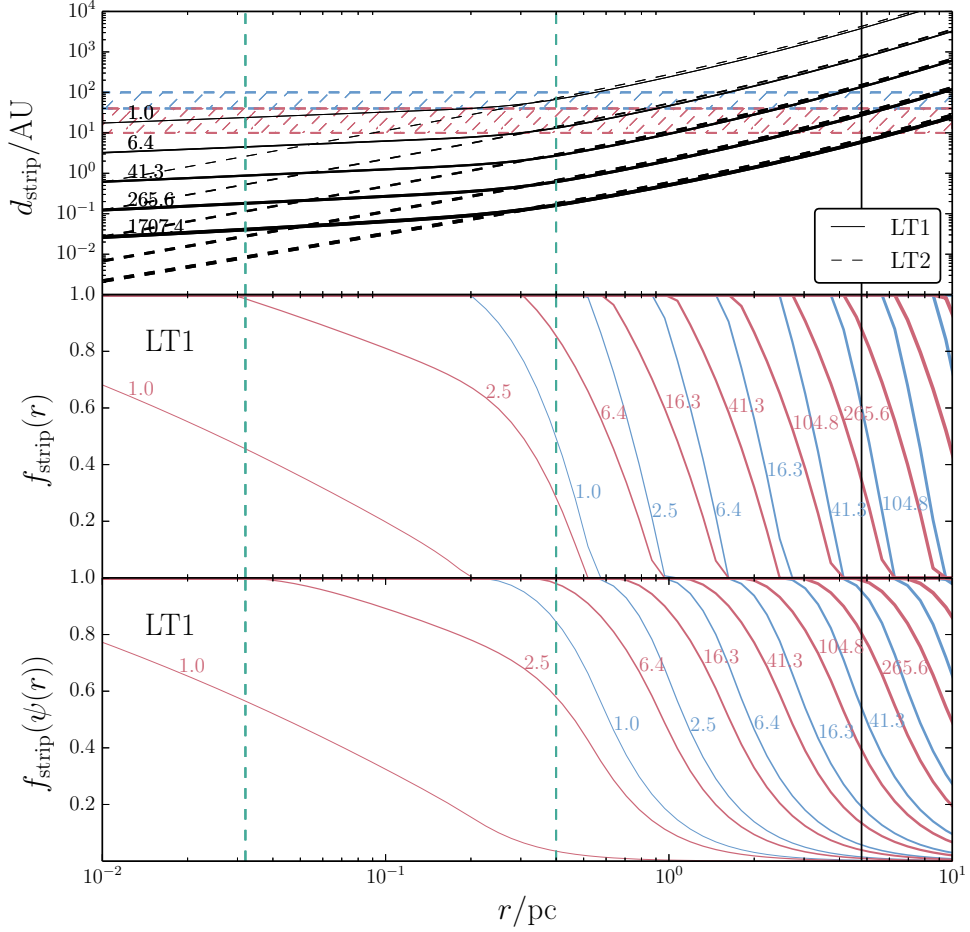


Figure 3.3: **Top panel:** the distance d_{strip} for which planetesimals are stripped from their parent star by gravitational encounters with other stars as a function of the distance r to the SBH, computed from equation (3.8) for different ages. Age increases with line thickness and is indicated (in Myr) with numbers. Black solid line: assuming LT1; black dashed line: assuming LT2; in both cases $m_{\text{bin}} = 1 M_{\odot}$. **Middle panel:** the stripping fraction as a function of r at various ages indicated in Myr with numbers, computed from d_{strip} using equation (3.4) and assuming LT1. Blue (red) lines apply to DM1 (DM2). **Bottom panel:** the angular-momentum- and orbit-averaged stripping fraction as a function of $r = \psi^{-1}(\mathcal{E})$ at various ages, assuming LT1. In all panels, the black solid vertical line indicates the radius of influence r_h for LT1.

In the above, the stripping fraction f_{strip} was computed as a function of r . For a typical orbit around the SBH r is not constant, however: the star will pass through environments with different densities and velocity dispersions. To take this into account, we performed an orbital average of $f_{\text{strip}}(r)$ weighted according to the time spent at each point in the orbit, and assuming an isotropic velocity distribution. See Appendix 3.A.2 for details.

In the bottom panel of Fig. 3.3, the orbit-averaged stripping fraction, $f_{\text{strip}}(\mathcal{E})$, is plotted as a function of $r = \psi^{-1}(\mathcal{E})$ for various ages and assuming LT1. Compared to the non-averaged $f_{\text{strip}}(r)$ (cf. the second panel of Fig. 3.3), the orbit-averaged stripping fractions are somewhat ‘smeared out’ to larger radii. In other words, allowing the orbital radius to vary during the orbit results in stripping by encounters to larger distances from the SBH.

3.4 Dynamics of planetesimals orbiting the SBH

3.4.1 Fokker-Planck equation

Planetesimals orbiting the SBH are susceptible to gravitational encounters with stars and other massive objects. The end result of this process is a steady-state in orbital angular momentum and energy. Previous studies (e.g. Bahcall & Wolf 1977; Merritt 2004; Merritt et al. 2007; Merritt 2010; Merritt 2013b, 7.1.2.2) have shown that in the case of scattering of a light population (i.e. planetesimals) by a massive population, the steady-state corresponds to a number density $n(r) \propto r^{-3/2}$. This is achieved on a time-scale $\sim t_{\mathcal{E}}$, the time-scale for non-resonant relaxation (NRR) to change the orbital energy by order itself.

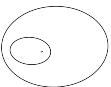
Here, we explore specifically the case of scattering of planetesimals by various perturbers in the GC. In addition, we take into account other potentially important effects such as the time- and energy-dependent influx of planetesimals from stars in our second scenario of planetesimal formation (cf. Section 3.1), as well as collisions and resonant relaxation (RR). Our main focus is the disruption rate of planetesimals by the SBH as a function of time, and in particular at $t = 10$ Gyr, which is approximately the age of the majority of the late-type stars in the GC (cf. Section 3.2.1).

To model the orbital evolution of planetesimals bound to the SBH, we solve the time-dependent Fokker-Planck equation in energy space for gravitational scattering of (nearly) massless particles (i.e. planetesimals) by massive particles with a central SBH (e.g. Merritt 2013b, 7.1.2.2), including several additional terms,

$$4\pi^2 p(\mathcal{E}) \frac{\partial f_a(\mathcal{E}, t)}{\partial t} = \frac{\partial \mathcal{F}_{\mathcal{E}}(\mathcal{E}, t)}{\partial \mathcal{E}} + F_{\text{strip}}(\mathcal{E}, t) - F_{\text{lc}}(\mathcal{E}, t) - F_{\text{col}}(\mathcal{E}, t) - F_{\text{RR}}(\mathcal{E}, t). \quad (3.9)$$

Here, $\mathcal{E} = -\frac{1}{2}v^2 + \psi(r)$ is the binding energy per unit mass of the planetesimals with respect to the SBH, $f_a(\mathcal{E}, t)$ is the planetesimal distribution function and $p(\mathcal{E})$ is the phase space volume per unit energy (cf. equation 3.28). The number of planetesimals with energies between $\mathcal{E}$ and $\mathcal{E} + d\mathcal{E}$ at time t is $N_a(\mathcal{E}, t) d\mathcal{E}$, where N_a is related to the distribution function via $N_a(\mathcal{E}, t) = 4\pi^2 p(\mathcal{E}) f_a(\mathcal{E}, t)$.

Equation (3.9) is derived under the assumption that the distribution function is (nearly) isotropic in velocity space and therefore independent of angular momentum (near the loss cone a logarithmic dependence on angular momentum is implicitly assumed, cf. Appendix 3.B.2). On



time-scales $\ll t_{\mathcal{E}}$ the evolution is expected to be dominated by changes in angular momentum, resulting in a steady-state with respect to that quantity (Frank & Rees 1976). The latter steady-state is subsequently expected to be maintained while the distribution in energies changes, because relaxation in energy occurs on the longer time-scale of $\sim t_{\mathcal{E}}$. In particular, for the GC it was shown by Merritt (2010) that an initially anisotropic distribution of late-type stars with a core in the GC evolves to near isotropy on a time-scale of 1 Gyr, which is much shorter than the relaxation time-scale at the radius of influence derived by Merritt (2010), $t_{\mathcal{E}} \sim 20 - 30$ Gyr. Here, we are interested in the long-term (> 1 Gyr) evolution of the planetesimals. For this purpose, it is therefore justified to assume isotropy in equation (3.9).

The term $\partial \mathcal{F}_{\mathcal{E}} / \partial \mathcal{E}$ in equation (3.9) describes the effect of gravitational scattering of planetesimals by more massive objects. In all our models, we consider scattering by late-type stars (cf. Section 3.2.1). In a number of models, we also consider scattering by a (hypothesized) cusp of stellar black holes (cf. Section 3.2.1) and massive perturbers (cf. Section 3.2.1). In each case, we assume that the distribution of scatterers is constant with time.

The terms $F_i(\mathcal{E}, t)$ in equation (3.9), which are all positive, represent sinks (negative sign preceding F_i) and sources (positive sign preceding F_i) of planetesimals. We consider the following effects. The source term F_{strip} describes the stripping of planetesimals from their parent stars. For late-type stars it is computed from

$$F_{\text{strip,LT}}(\mathcal{E}, t) = N_{a/\star} N_{\text{LT}}(\mathcal{E}) \frac{\partial f_{\text{strip,LT}}(\mathcal{E}, t)}{\partial t}. \quad (3.10)$$

Here, $N_{a/\star}$ is the number of planetesimals per star, $N_{\text{LT}}(\mathcal{E}) d\mathcal{E}$ is the number of late-type stars with energies between $\mathcal{E}$ and $\mathcal{E} + d\mathcal{E}$, and $f_{\text{strip,LT}}(\mathcal{E}, t)$ was computed in Section 3.3.2. In the case of stripping from early-type stars equation (3.3) is adopted. We note that our approach of including a source term in the Fokker-Planck equation in energy space to model inflow of planetesimals bound to the SBH is similar to the approach of Aharon & Perets (2015), who include a source term to model the formation of stars in nuclear star clusters. Furthermore, the sink term F_{lc} describes the effects of gravitational encounters into the loss cone of the SBH in angular-momentum space. In addition, physical collisions are described by the sink term F_{col} ; we consider collisions of planetesimals with late-type stars and with other planetesimals. Lastly, the sink term F_{RR} describes the effects of RR.

More details on the terms that appear on the right-hand side of equation (3.9) are included in Appendix 3.B. To investigate the relative importance of the effects described by these terms we solve equation (3.9) for each effect separately. Our models are described in Table 3.2.

In each integration of equation (3.9), we evaluate the (energy-integrated) disruption rate of planetesimals by the SBH, $F_{\text{dis}}(t)$, defined by

$$F_{\text{dis}}(t) = \int d\mathcal{E} [F_{\text{lc}}(\mathcal{E}, t) + F_{\text{RR}}(\mathcal{E}, t)] + \mathcal{F}_{\mathcal{E}}(\mathcal{E}_{\text{lc}}). \quad (3.11)$$

Here, $\mathcal{E}_{\text{lc}} = GM_{\bullet}/(2r_{\text{lc}})$ is the energy of an orbit with semimajor axis r_{lc} . The term $\mathcal{F}_{\mathcal{E}}(\mathcal{E}_{\text{lc}})$ represents the loss cone flux arising from scattering in energy space. Although included in the calculations for completeness, we find that $\mathcal{F}_{\mathcal{E}}(\mathcal{E}_{\text{lc}})$ is invariably negligible compared to the flux in angular-momentum space, which is represented by the first two terms on the right-hand side of equation (3.11). The largest fractional contribution of the energy flux to the total rate for the rates given in Table 3.1 $\approx 2 \times 10^{-5}$.

3.4.2 Boundary and initial conditions

Equation (3.9) is integrated in time using the `PYTHON FIPY` library (Guyer et al. 2009). For each model integration, 200 steps are taken in which the term $F_{\text{strip}}(\mathcal{E}, t)$, which explicitly depends on time, is updated. We have verified that increasing this number of steps does not change the results. The orbital energies are discretized on to a grid that is constructed using the relation $\mathcal{E} = \psi(r)$, where r is sampled from a logarithmic grid in the range $2r_{\text{lc}} \leq r \leq 100 \text{ pc}$, and $r_{\text{lc}} = 1 \text{ AU}$. The largest value of $\mathcal{E}$, $\mathcal{E}_{\text{lc}} = \psi(2r_{\text{lc}})$, corresponds to a semimajor axis of $a = r_{\text{lc}}$ under the assumption that $\psi(r_{\text{lc}}) = GM_{\bullet}/r_{\text{lc}}$. At this energy planetesimals are considered disrupted by the SBH, and we impose the boundary condition $f_a(\mathcal{E}_{\text{lc}}, t) = 0$.

The initial distribution in the case of planetesimals initially in discs around stars is

$$N_a(\mathcal{E}, 0) = N_{a/\star} f_{\text{strip;SBH}}(\mathcal{E}) N_{\text{LT}}(\mathcal{E}). \quad (3.12)$$

Here, $f_{\text{strip;SBH}}(\mathcal{E})$ is the stripping fraction due to the tidal force of the SBH only (cf. Section 3.3.1), and $N_{\text{LT}}(\mathcal{E}) = 4\pi^2 p(\mathcal{E}) f_{\text{LT}}(\mathcal{E})$ is number of late-type stars with energies between $\mathcal{E}$ and $\mathcal{E} + d\mathcal{E}$. In equation (3.12), we assume that any stripping by the SBH acts instantaneously. We justify this by noting that most of the stripping by the SBH occurs at radii $\lesssim 1 \text{ pc}$ (cf. the bottom panel of Fig. 3.1), for which the radial orbital period $\lesssim 7 \times 10^4 \text{ yr}$. We expect stripping by the SBH to occur on the latter time-scale, therefore stripping acts essentially instantaneously compared to our integration time.

In the case of planetesimals formed in a large-scale cloud around the SBH the initial distribution is unclear. Here, we assume that the distribution is similar to the (late-type) stellar distribution, i.e. $N_a(\mathcal{E}, 0) \propto N_{\text{LT}}(\mathcal{E})$. The constant of proportionality is unconstrained, but here we assume it is given by the number of planetesimals per star, $N_{a/\star} = 2 \times 10^7$, i.e. we assume that the initial distribution is

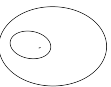
$$N_a(\mathcal{E}, 0) = N_{a/\star} N_{\text{LT}}(\mathcal{E}). \quad (3.13)$$

3.4.3 Results: distributions and disruption rates

The disruption rates at $t = 10 \text{ Gyr}$ are listed for various models in Table 3.1, assuming $N_{a/\star} = 2 \times 10^7$. In the first column the model number is listed; a succinct description of the models is given in Table 3.2. In model 1, the effect of scattering in energy and angular-momentum space by late-type stars is included; in other models late-type scattering is included, as well as other effects. For each of the models in Table 3.1, rates are quoted assuming the planetesimals were either formed in a large-scale cloud (‘cloud’) or in discs around stars (‘disc’). In the latter case, the stripping term, equation (3.10), is included assuming either DM1 or DM2 (cf. Section 3.2.2). Furthermore, rates are included for LT1 and LT2 (cf. Section 3.2.1). Below we discuss in more detail the results for the various models given in Table 3.1.

Scattering by late-type stars only

In the top four panels of Fig. 3.4, we show the differential number of planetesimals, $N_a(r, t)$, as a function of $r = \psi^{-1}(\mathcal{E})$ at various times. Here, $N_a(r, t)$ is defined as the number of planetesimals between radii r and $r + dr$, i.e. $N_a(\psi(r), t) d\psi(r) = N_a(r, t) dr$. The assumed model is model 1 (cf. Table 3.1). In the first (second) column results are shown for LT1 (LT2);



model	$F_{\text{dis}}(t = 10 \text{ Gyr})/\text{d}^{-1}$					
	LT1			LT2		
	cloud	disc		cloud	disc	
		DM1	DM2		DM1	DM2
1	0.55	0.54	0.53	0.61	0.59	0.57
2	0.55	0.54	0.53	0.61	0.59	0.57
3	0.38	0.37	0.36	0.51	0.49	0.47
4	0.38	0.37	0.36	0.51	0.49	0.47
5	0.55	0.54	0.53	0.59	0.57	0.54
6	0.54	0.53	0.52	0.59	0.57	0.55
7	0.55	0.54	0.53	0.61	0.59	0.56
8	0.58	0.57	0.56	0.63	0.61	0.58
9	0.55	0.54	0.53	0.61	0.59	0.57
10	0.71	0.70	0.69	0.72	0.69	0.67
11	0.58	0.57	0.56	0.63	0.60	0.58
12	0.71	0.70	0.69	0.72	0.69	0.67
13	0.58	0.57	0.56	0.63	0.60	0.58
14	0.24	0.18	0.14	0.42	0.31	0.23

Table 3.1: Disruption rates of planetesimals with radius ≥ 10 km passing the SBH in the GC at distances closer than 1 AU, possibly resulting in an observable NIR/X-ray flare (ZNM12), based on time-integrations of equation (3.9). The initial number of these planetesimals per star is assumed to be $N_{a/\star} = 2 \times 10^7$, consistent with observations of debris discs of stars in the Solar neighbourhood. Rates at the end of our integrations, $t = 10$ Gyr, are listed for various models, given in the first column, and for different combinations of the assumed distribution of late-type stars (LT1 or LT2, cf. Section 3.2.1). The planetesimals are assumed to be initially either distributed in a large-scale cloud around the SBH (‘cloud’) or in discs around stars (‘disc’; the two disc models DM1 and DM2 are defined in Section 3.2.2). Descriptions of the various models are included in Table 3.2.

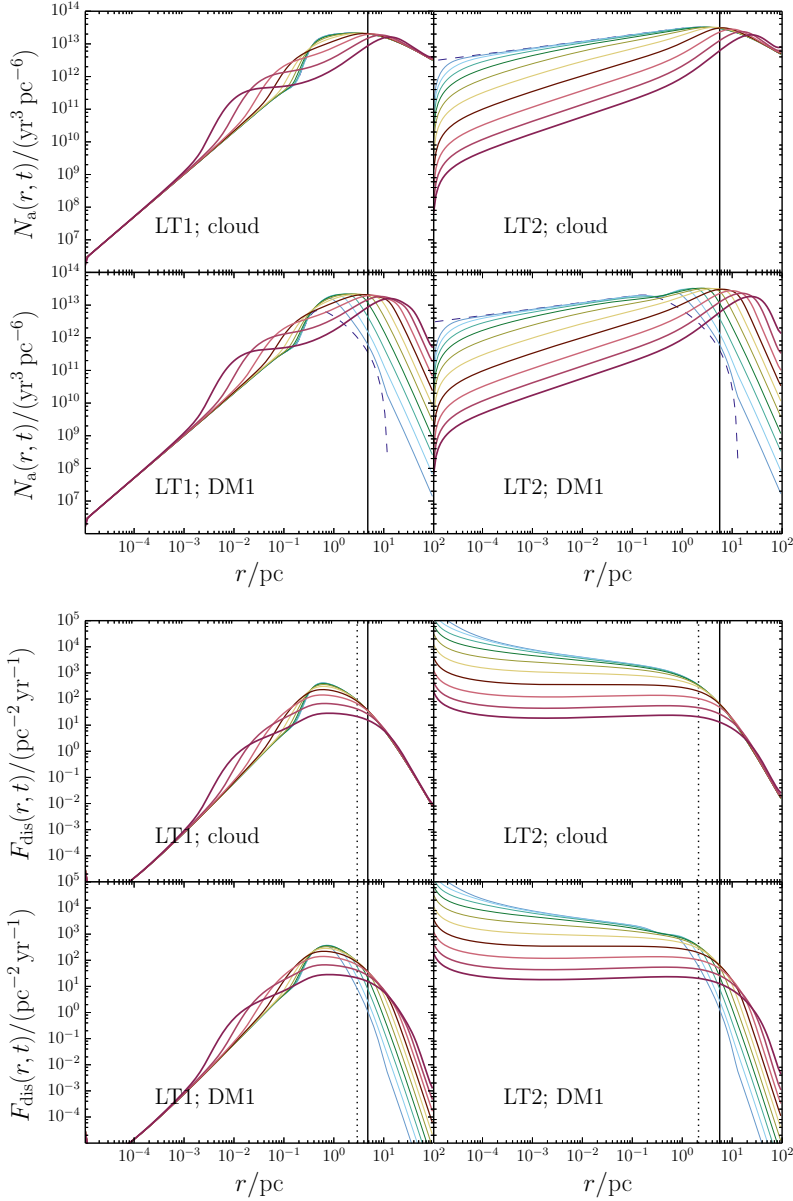
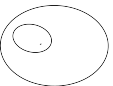


Figure 3.4: Top four panels: the quantity $N_a(r, t)$, defined such that $N_a(r, t) dr$ is the number of planetesimals between radii r and $r + dr$, as a function of r at various times, according to numerical integrations of equation (3.9) and assuming model 1 (cf. Table 3.2). The line thickness increases with time; the times shown are 0 (blue dashed line), 2.5, 6.4, 16.3, 41.3, 104.8, 265.6, 673.4, 1707.4, 4328.8 and 10000.0 Myr. The first (second) column applies to LT1 (LT2); the first (second) row applies to formation in a cloud (disc; DM1). The radii of influence r_h for LT1 and LT2 are indicated with the black vertical solid lines. **Bottom four panels:** the disruption flux of planetesimals by the SBH, $F_{\text{dis}}(r, t)$ (cf. Section 3.4.3), as a function of r at various times (the meaning of the line thicknesses and colours is the same as in the top four panels). The black vertical dotted line shows the boundary between the empty and full loss cone regimes, here defined as $\tilde{q} = 1$ (cf. Section 3.B.2).



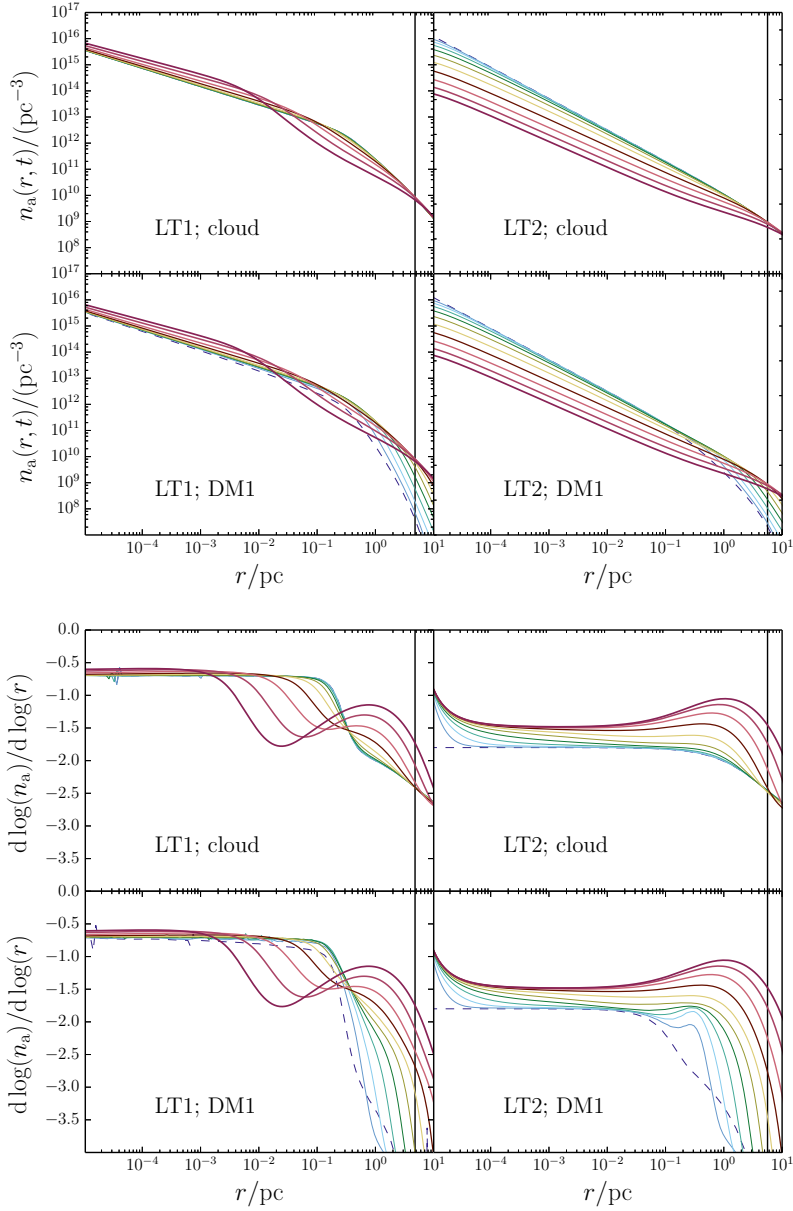


Figure 3.5: Similar to Fig. 3.4, here showing the planetesimal number density $n_a(r, t)$ as a function of r at the various times in the top four panels and the logarithmic number density derivative in the bottom four panels.

model	description
1	LT scattering
2	model 1 + source term: planetesimals from ET stars
3	model 1 + BH cusp (no RR)
4	model 1 + BH cusp (no RR); t -dependence
5	model 1 + LT-planetesimal collisions ($m_{\text{LT,init}} = 1 M_{\odot}$)
6	model 1 + LT-planetesimal collisions ($m_{\text{LT,init}} = 2 M_{\odot}$)
7	model 1 + planetesimal-planetesimal collisions
8	model 1 + RR; LT ($\chi_{\text{RR}} = 1$)
9	model 1 + RR; LT ($\chi_{\text{RR}} = 0.1$)
10	model 1 + RR; BH cusp ($\chi_{\text{RR}} = 1$)
11	model 1 + RR; BH cusp ($\chi_{\text{RR}} = 0.1$)
12	model 1 + RR; BH cusp ($\chi_{\text{RR}} = 1$; t -dependence)
13	model 1 + RR; BH cusp ($\chi_{\text{RR}} = 0.1$; t -dependence)
14	model 1 + massive perturbers (PHA07: GMC1)

Table 3.2: Brief descriptions of the various models included in Table 3.1. Refer to the text in Section 3.4.3 and Appendix 3.B for more details.

the first row applies to formation in a cloud, whereas the second row applies to formation in discs (DM1). The radii of influence, defined via $M_{\text{LT}}(r_h) = 2M_{\bullet}$, where $M_{\text{LT}}(r)$ is the distributed mass in late-type stars (cf. equation 3.7), are indicated with the vertical black solid lines.

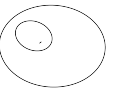
In the top four panels of Fig. 3.5, the planetesimal number density $n_a(r, t)$ is plotted at various times. It is computed from $f_a(\mathcal{E}, t)$ using (e.g. Merritt 2013b, 3.2.1)

$$n_a(r, t) = \int d^3v f_a(v, t) = 4\pi\sqrt{2} \int_{-\infty}^{\psi(r)} d\mathcal{E} f_a(\mathcal{E}, t) \sqrt{\psi(r) - \mathcal{E}}. \quad (3.14)$$

In Fig. 3.5, we also show in the bottom four panels the logarithmic derivative $d \log(n_a)/d \log(r)$.

We show the initial distributions in Figs 3.4 and 3.5 with blue dashed lines. In the case of formation in a cloud, the planetesimal number density n_a is proportional to the assumed number density of late-type stars at all radii. In the case of formation in discs this is the case at small radii because stripping is assumed to have occurred due to the tidal force of the SBH, whereas at larger radii planetesimals have not yet been stripped (cf. Section 3.3.1). Therefore, the number density drops rapidly for radii $r \gtrsim 0.3$ pc (cf. the middle panel of Fig. 3.1).

We show in Fig. 3.6 the stripping flux $F_{\text{strip}}(r, t)$ in the case of formation in discs at various times, for LT1 (solid lines) and LT2 (dashed lines). Here, we define $F_{\text{strip}}(r, t) dr$ as the number of stripped planetesimals per unit time between radii r and $r + dr$. Initially the stripping flux peaks at $r \approx 1$ pc. As time progresses, planetesimals are stripped at progressively increasing radii. After 100 Myr, the peak has shifted to outside the radius of influence, and the magnitude of the peak has decreased as well. There are no substantial differences in $F_{\text{strip}}(r, t)$ between the distributions of late-type stars. This can be explained by the lack of influx of stripping by encounters at small radii $\lesssim r_b = 0.3$ pc, at which the SBH is responsible for stripping, therefore the core in the case of LT1 does not affect the stripping process.



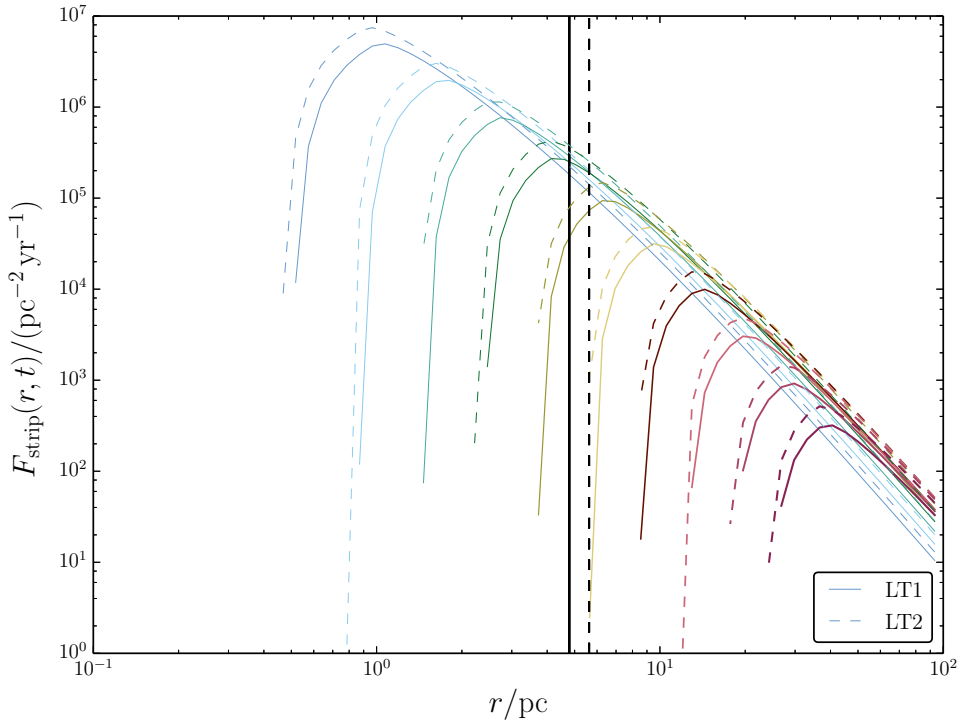


Figure 3.6: The stripping flux $F_{\text{strip}}(r, t)$ (cf. Section 3.4.1) at the same times as in Fig. 3.4 assuming model 1 (cf. 3.2) and DM1. Solid lines: LT1; dashed lines: LT2.

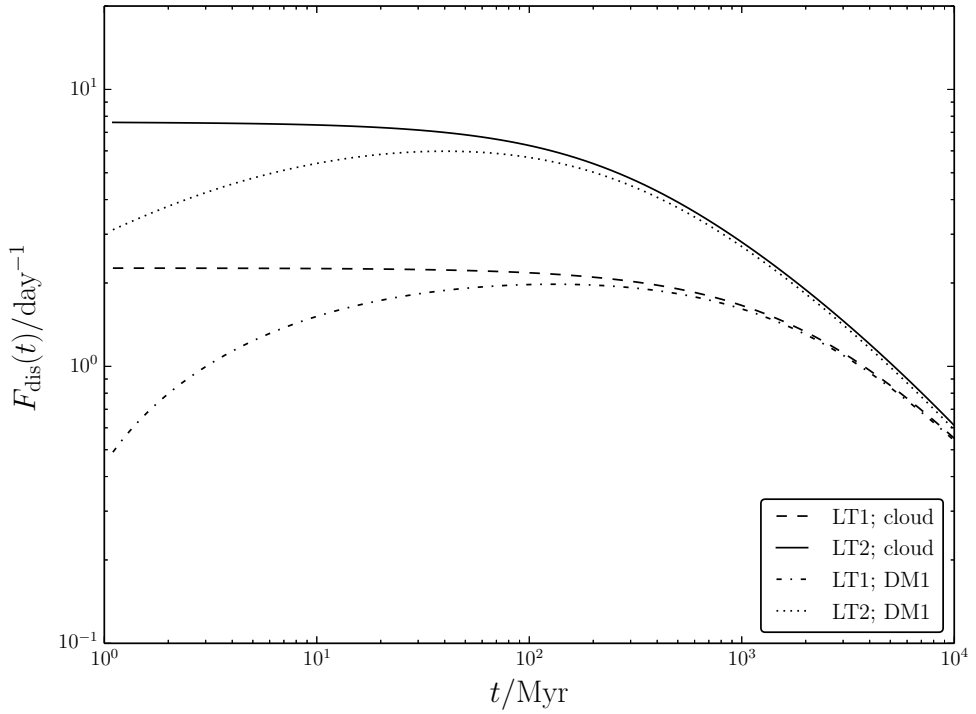
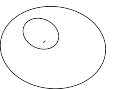


Figure 3.7: The energy-integrated disruption rate $F_{\text{dis}}(t)$ (cf. equation 3.11) as a function of time assuming model 1 (cf. Table 3.2). Results are included for LT1 and LT2 and formation in a cloud and disc (DM1).



In the bottom panels of Fig. 3.4, the disruption flux $F_{\text{dis}}(r, t)$ is shown, where $F_{\text{dis}}(r, t)$ is the number of disrupted planetesimals per unit time between radii r and $r + dr$. For model LT1, this flux peaks at radii somewhat smaller than r_h , whereas for LT2 it is much less peaked. This difference can be attributed to the core in LT1. The flux at radii $> r_h$ is initially more than two orders of magnitude larger for formation in a cloud compared to formation in discs. However, as time progresses and planetesimals are stripped at increasing radii, this difference decreases. Despite this influx at large radii, the peak in the disruption flux does not shift to larger radii, but rather it becomes wider.

In Fig. 3.7, we show the energy-integrated disruption rates $F_{\text{dis}}(t)$ (cf. equation 3.11) for LT1 and LT2, and for both formation scenarios. Initially, the rates for model LT2 are nearly an order of magnitude larger compared to model LT1. This can be attributed to the initially higher planetesimal density at small radii $r \lesssim 0.3$ pc in the case of LT2 because the late-type stars are assumed to be distributed in a cusp as opposed to a core (cf. the top four panels of Fig. 3.4). While planetesimals are depleted at these radii the density drops, and as a consequence the loss cone flux decreases as well (cf. the bottom four panels of Fig. 3.4).

The stripping from stars causes the disruption rate to increase initially, whereas for formation in a cloud the rate only decreases with time. The differences are small, however: initially the rates in the former case are a factor of ~ 2 higher, and as time progresses the rates in the two cases gradually approach each other. After ~ 100 Myr, the rates become effectively indistinguishable, because stripping has progressed to beyond the radius of influence (cf. Fig. 3.6), where the disruption flux peaks (cf. the bottom four panels of Fig. 3.4). By $t = 10$ Gyr, the rates in all cases are nearly the same at $F_{\text{dis}}(t = 10 \text{ Gyr}) \approx 0.6 \text{ d}^{-1}$ (cf. Table 3.1).

As mentioned in Section 3.4.1, the planetesimal steady-state distribution is expected to correspond to a number density $n(r) \propto r^{-3/2}$. The logarithmic slopes, shown in the bottom four panels of Fig. 3.5, can be used to assess this. In the case of LT1 and by $t = 10$ Gyr, the slopes are only consistent with $-3/2$ at a few radii $\ll r_h$, independent on the initial conditions. In particular, the slope is still close to the initial value of -0.7 (cf. Section 3.2.1) at radii $\lesssim 10^{-3}$ pc. Evidently, the core of late-type stars in the case of LT1 is unable to smoothen out the distribution at these radii, whereas this is the case in LT2 (note that the flattening in the later case for $r \lesssim 10^{-4}$ pc arises from the boundary condition $f_a(\mathcal{E}_{\text{lc}}, t) = 0$). For both LT1 and LT2 there is a flattening in the number density near radii of ~ 1 pc, causing a ‘bump’ in the logarithmic slope. This can be attributed to the disruption flux, which peaks near this radius (cf. Fig. 3.4).

Stripping from early-type stars

As mentioned in Section 3.2.1, planetesimals in the GC could also form in debris discs around early-type stars. This source of planetesimals has been included in model 2 (cf. Table 3.1) by adding the term $F_{\text{strip;ET}}(\mathcal{E})$ (cf. equation 3.3) to the right-hand side of equation (3.9). We find that the inclusion of this term has virtually no effect on the planetesimal distribution at all times. This is also reflected by the disruption rates, which change by less than 1 per cent (cf. Table 3.1). Considering that equation (3.3) is likely an upper limit for the supply of planetesimals from early-type stars, we conclude that stripping from early-type stars can be safely neglected compared to stripping from the much more numerous late-type stars.

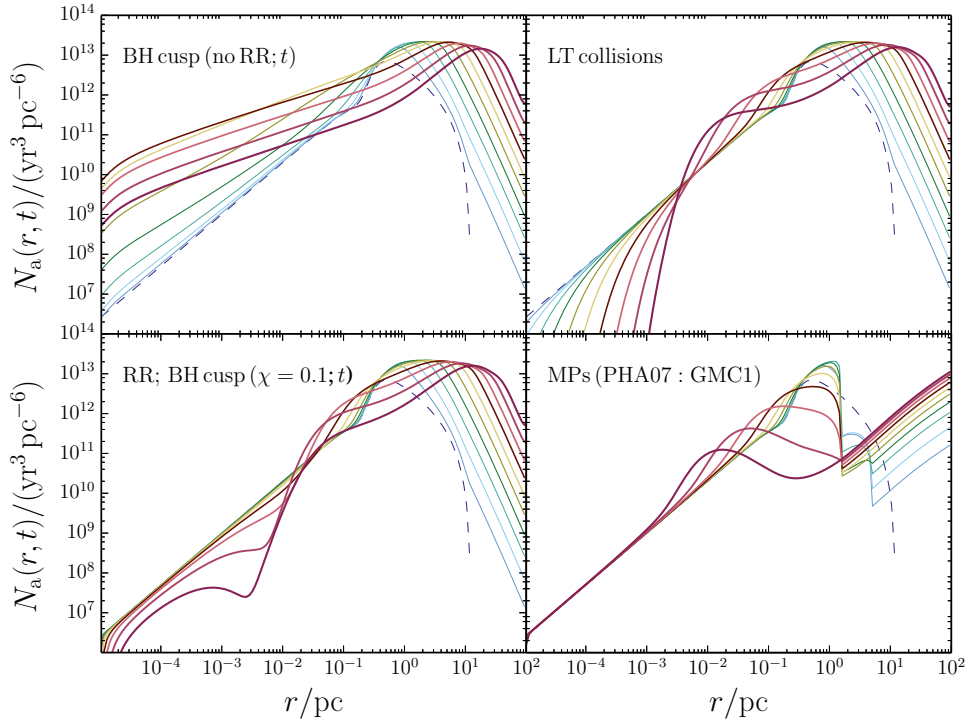


Figure 3.8: The quantity $N_a(r, t)$ for four models (models 4, 5, 14 and 15, cf. Table 3.2) as a function of r at various times. In all panels LT1 and formation in a disc (DM1) are assumed.

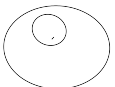
Scattering by a cusp of stellar black holes

In models 3 and 4, scattering in energy and angular momentum is taken into account assuming a (hypothesized) cusp of stellar black holes close to the SBH. In the case of model 4, the black hole number density is time-dependent and is assumed to increase linearly with time, reaching the number density that we assume at all times in model 3 (cf. Section 3.2.1), at $t = 10$ Gyr.

We show in the top-left panel of Fig. 3.8, the quantity $N_a(r, t)$ for the case of a time-dependent black hole cusp and for LT1 and DM1. The effect of a cusp of stellar black holes is initially to increase the density of planetesimals at small radii, $r \lesssim 0.1$ pc. At later times this density decreases again. This is reflected in the disruption rate, which is shown as a function of time in Fig. 3.9 (black dashed line). Initially the rate is slightly higher compared to the model 1 (blue solid line), but after ≈ 200 Myr it drops slightly below the latter rate.

Collisions

We consider three different models in which the effects of physical collisions are taken into account. In model 5, collisions between late-type stars and planetesimals are considered; the initial stellar mass is assumed to be $1 M_\odot$, and the mass and radius at subsequent times are



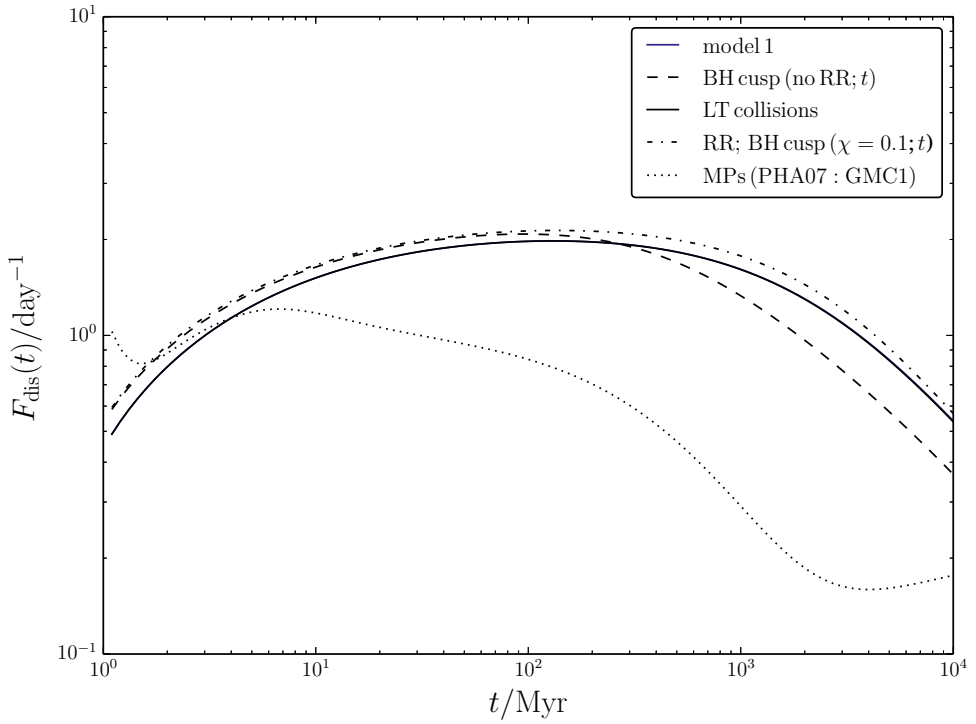


Figure 3.9: The total disruption rate $F_{\text{dis}}(t)$ as a function of time (cf. equation 3.11) for model 1 and the models included in Fig. 3.8. In all cases LT1 and DM1 are assumed.

adopted from a stellar evolution model with metallicity $Z = 0.02$ (the *sse* code; Hurley et al. 2000, is used as implemented in *AMUSE*, Pelupessy et al. 2013; Portegies Zwart et al. 2013). In the latter model, the radius changes only little during 10 Gyr; we have also included a model, model 6, in which the stellar mass is assumed to be $2M_{\odot}$ (resulting in faster evolution). In the third model, model 7, planetesimal-planetesimal collisions are taken into account with a number of simplifying assumptions (cf. Section 3.B.3).

The quantity $N_a(r, t)$ for model 5 is plotted in the top-right panel of Fig. 3.8 (results for models 6 and 7 are similar). It is reduced substantially compared to model 1 only at radii $\lesssim 10^{-2}$ pc (cf. the top-right panel of Fig. 3.8). The disruption flux peaks at much larger radii (cf. Fig. 3.4). Therefore, the effect of collisions on the disruption rates is very small. This is illustrated in Fig. 3.9, in which the rates according to model 5 (black solid line) are indistinguishable, from an observational perspective, from the rates in the model 1 (blue solid line).

RR

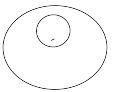
In six models, models 8-13, the effects of RR are included (cf. Section 3.B.3). In models 8 and 9, RR is assumed to arise from the late-type stars; in models 10-13 RR is assumed to arise from a cusp of stellar black holes, with and without a time dependence (cf. Section 3.2.1). Two values of the efficiency of RR, χ_{RR} (cf. Section 3.B.3), are assumed: $\chi_{RR} = 1$ and $\chi_{RR} = 0.1$.

In the case of RR arising from late-type stars, $N_a(r, t)$ is not affected noticeably compared to the case without RR. The disruption rates are enhanced only a few per cent, even if $\chi_{RR} = 1$ (cf. Table 3.1). A cusp of stellar black holes affects $N_a(r, t)$ more strongly, but this is sensitive to the RR efficiency χ_{RR} and whether or not the cusp is assumed to be time-dependent.

In the bottom left panel of Fig. 3.8, $N_a(r, t)$ is shown as a function of r for various times assuming a time-dependent cusp of stellar black holes with $\chi_{RR} = 0.1$ (model 13). This model can be considered as the most realistic model of RR with a black hole cusp. Only at late times, $t \sim 10$ Gyr, and at small radii, $r \lesssim 10^{-2}$ pc, is RR effective at reducing $N_a(r, t)$. Nevertheless, the energy-integrated disruption rates at $t = 10$ Gyr are increased by at most 3 per cent compared to model 1 (cf. Table 3.1).

Massive perturbers

In model 14, the effect of massive perturbers is taken into account (cf. Section 3.2.1). Contrary to what might be expected, the disruption rates at $t = 10$ Gyr are a factor of ~ 2 lower compared to those of model 1. The massive perturbers considered here are located at radii > 1.5 pc (cf. table 2 of PHA07), which corresponds to the full loss cone regime in the case of relaxation by late-type stars only (cf. the black vertical dotted lines in the bottom four panels of Fig. 3.4). The full loss cone rate corresponds to the maximum loss rate, and this explains why the disruption rates are initially only modestly larger, by a factor of ~ 2 , as shown in Fig. 3.9 (dotted lines). Relaxation in energy is also assumed to be much faster, however, therefore at the radial range of the massive perturbers, $1.5 \text{ pc} < r < 100 \text{ pc}$, planetesimals are efficiently transported from radii $r \sim 1 \text{ pc}$ to $r \gg 1 \text{ pc}$, as illustrated in the bottom-right panel of Fig. 3.8. Because the loss cone flux still peaks near 1 pc (cf. Fig. 3.4), this implies that the disruption rate must drop rapidly, which is reflected in Fig. 3.9. Nevertheless, after 10 Gyr the disruption rate is still $\approx 0.2 \text{ d}^{-1}$, which is only a factor of ≈ 3 lower compared to model 1.



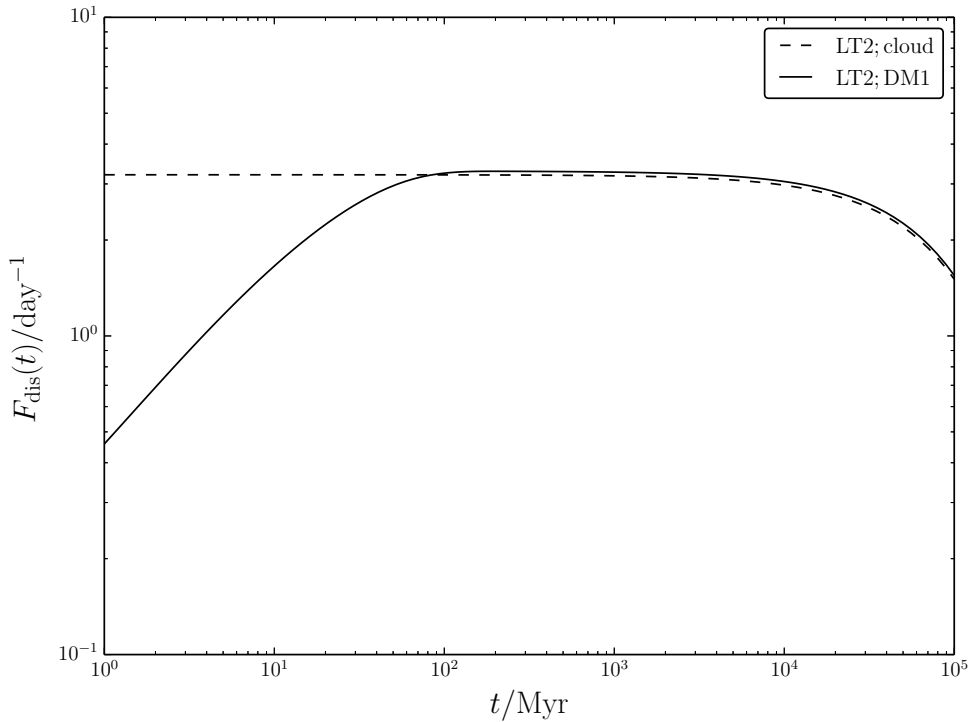


Figure 3.10: The energy-integrated disruption rate $F_{\text{dis}}(t)$ as a function of time as in Fig. 3.7, here according to approximate semi-analytic solutions to equation (3.9), for which it is assumed that a steady-state in energy applies at all times (cf. Section 3.4.3). Rates are shown for LT2 and the two formation scenarios.

Semi-analytic solutions

Equation (3.9) was solved numerically. We also obtained approximate semi-analytic solutions to this equation for model 1. In order to solve the equation, we assume that a steady-state in energy applies at all times. With this assumption, and setting the terms F_{col} and F_{RR} to zero in accordance with model 1, we find semi-analytical solutions of the disruption flux as a function of time (refer to Appendix 3.C for details).

These solutions are shown in Fig. 3.10 for LT2 and the two formation scenarios. The time-scale of the decrease of the disruption rate after $\sim 100 \text{ Myr}$, $\gg 10^4 \text{ Myr}$, is too long compared to the numerical solutions (cf. Fig. 3.7). Nevertheless, the semi-analytic solutions yield disruption rates of the same order of magnitude, therefore they should be adequate for the purposes of scaling the rates to different galactic nuclei (cf. Section 3.5.3).

3.5 Discussion

3.5.1 Comparison to observations: constraints on $N_{a/\star}$

The disruption rates at $t = 10$ Gyr as shown in Table 3.1 are generally $\approx 0.6 \text{ d}^{-1}$ for models 1-13; in the case of model 14 (massive perturbers) the rate is slightly lower, $\approx 0.2 \text{ d}^{-1}$. These rates are robust in the sense that there is no strong dependence on the assumed distribution of late-type stars (LT1 versus LT2), nor on the formation scenario (cloud versus disc). In particular, we note that our rates vary only a few per cent between the two disc models DM1 and DM2. This is because stripping by gravitational encounters occurs early in the evolution, and it is very effective at stripping all planetesimals at radii $\lesssim r_h$ for both DM1 and DM2, even though in DM2 planetesimals are more tightly bound to the star (cf. the bottom panel of Fig. 3.3).

Based on 3 Ms *Chandra* observations in 2012, the observed X-ray flaring rate is $1.1^{+0.2}_{-0.1} \text{ d}^{-1}$ (Nielsen et al. 2013). Strictly speaking, this would seem inconsistent with our rate of $\approx 0.6 \text{ d}^{-1}$ (5σ deviation).

We emphasize, however, that in Section 3.4 we assumed $N_{a/\star} = 2 \times 10^7$ and that this quantity is poorly constrained (the main uncertainty in $N_{a/\star}$ is the total mass of planetesimals per star). Our result can therefore be used to constrain $N_{a/\star}$ within the framework of our other assumptions. Using that the disruption rate is linearly proportional to $N_{a/\star}$ (cf. Section 3.5.3), it directly follows that according to models 1-13, $N_{a/\star} \approx 3.7 \times 10^7$, and according to model 14, $N_{a/\star} \approx 1.1 \times 10^8$.

3.5.2 Internal scattering of planetesimals by planets

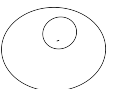
In Section 3.4, internal processes in the debris discs were not considered. Likely the most important of these processes is scattering of planetesimals by planets bound to the star, which could lead to the ejection of a significant fraction of planetesimals from the star, analogously to the Nice model (Gomes et al. 2005). The precise fraction of ejected planetesimals likely depends strongly on the distribution of planets around stars in the GC, which is currently completely unconstrained.

The scenarios considered in Section 3.4 can be interpreted as two extreme cases of internal scattering. In our cloud scenario, the number of planetesimals was chosen to be consistent with the number of stars and the number of planetesimals per star, and the planetesimals were assumed to be formed with an orbital distribution around the SBH similar to that of the stars. Modulo a likely delay between formation and ejection, this is consistent with an extreme case of internal scattering where *all* planetesimals are ejected from the debris disc. In contrast, in our disc scenario ejection was only assumed to occur due to encounters with other stars, i.e. no internal processes are taken into account.

The above implies that our cloud and disc scenarios can be used to evaluate the effect of internal processes on the flaring rate. The weak dependence of the latter rate on the assumed scenario, in particular at late times (cf. Fig. 3.7), suggests that this effect is very weak.

3.5.3 Scaling of the disruption rate: tidal disruption of planets

In Section 3.4, the number of planetesimals per star was assumed to be $N_{a/\star} = 2 \times 10^7$, consistent with observations of debris discs around stars in the Solar neighbourhood (ZNM12;



Wyatt 2008). Other important parameters that were assumed are the masses $M_\bullet = 4 \times 10^6 M_\odot$ and $m_{\text{LT}} = 1 M_\odot$, and the distributed mass in late-type stars within $r_0 = 1 \text{ pc}$, $M_0 = 1.5 \times 10^6 M_\odot$ (cf. Section 3.2.1), or, equivalently, the number of late-type stars within r_0 , $N_0 = M_0/m_{\text{LT}} = 1.5 \times 10^6$.

The SBHs with masses $\sim 10^6 M_\odot$ in several nearby spiral galaxies have quiescent X-ray luminosities of $\sim 10^{37} - 10^{39} \text{ erg/s}$ (Baganoff et al. 2003 and references therein). These luminosities are several orders of magnitude larger than the typical X-ray flare luminosity associated with the tidal disruption of planetesimals, $\sim 10^{34} - 10^{35} \text{ erg/s}$, making it unlikely that such flares could be observed in galactic nuclei other than the GC.

This conclusion could be different for the tidal disruption of planets by the SBH, which would produce a much more luminous flare. For example, ZNM12 estimated that the tidal disruption of a Jupiter-mass gas giant would produce a flare with an X-ray luminosity of $\sim 2 \times 10^{41} \text{ erg/s}$, and this would therefore dominate the quiescent luminosity. Such a flare could be observed with e.g. the *INTEGRAL* space telescope up to a distance of $\sim 50 \text{ Mpc}$; the number of galaxies within this distance is ~ 5000 (Nikołajuk & Walter 2013).

For this purpose, we obtain an approximate scaling of the disruption rate based on the semi-analytic solutions to equation (3.9) that were discussed in Section 3.4.3, with the inclusion of the stripping term (i.e. assuming formation in discs) and the angular-momentum loss cone term arising from scattering with late-type stars. As in Section 3.4.3, we assume that at all times a steady-state in the planetesimal distribution function has been reached with respect to the orbital energies. Furthermore, in order to find an analytic scaling we approximate the effect of stripping by gravitational encounters as an instantaneous process, we assume a single power-law number density of the perturbing stars, $n_\star \propto r^{-\gamma}$, and we neglect the stellar potential.

With these assumptions and approximations we find the following order-of-magnitude estimate of the disruption rate (refer to Appendix 3.D for more details)

$$F_{\text{dis}}(t) \sim \frac{N_{\text{a}/\star} N_0}{\tau} \exp(-t/\tau) [1 + \exp(t_c/\tau)], \quad (3.15)$$

where τ is given by

$$\begin{aligned} \tau^{-1} &\equiv \left[\int d\mathcal{E} S_{\text{lc}}(\mathcal{E}) \right] \left[4\pi^2 \int d\mathcal{E} p(\mathcal{E}) \right]^{-1} \\ &\sim N_0 (Gm_\star)^2 (GM_\bullet r_0)^{-3/2} \log(\Lambda) \left(\frac{2M_\bullet}{m_\star} N_0^{-1} \right)^{\frac{\gamma-3/2}{\gamma-3}}. \end{aligned} \quad (3.16)$$

Here, we neglected factors of order unity. The quantity t_c is the time-scale for which most planetesimals are assumed to have been stripped by gravitational encounters; $t_c \sim 100 \text{ Myr}$ for the GC (cf. Section 3.3.2). Equations (3.15) and (3.16) give for the GC parameters and $\gamma = 2$

at $t = 10 \text{ Gyr}$,

$$\begin{aligned}
 F_{\text{lc}}(t = 10 \text{ Gyr}) &\approx 6 \text{ d}^{-1} \left(\frac{N_{\text{a}/\star}}{2 \times 10^7} \right) \left(\frac{m_\star}{M_\odot} \right)^{5/2} \\
 &\times \left(\frac{N_0}{1.5 \times 10^6} \right)^{5/2} \left(\frac{r_0}{1.0 \text{ pc}} \right)^{-3/2} \left(\frac{M_\bullet}{4 \times 10^6 M_\odot} \right)^{-2} \\
 &\times \log \left(\frac{M_\bullet}{4 \times 10^6 M_\odot} \frac{1 M_\odot}{2 m_\star} \right). \quad (3.17)
 \end{aligned}$$

This estimate is correct within an order of magnitude (cf. Table 3.1). The scaling is linear with $N_{\text{a}/\star}$, which is intuitively easy to understand. The scaling with N_0 is $F_{\text{lc}} \propto N_0^{5/2}$. The latter can be understood as follows: increasing the number of stars increases the total supply of planetesimals but also accelerates the rate of scattering into the loss cone, hence the dependence is stronger than linear.

We use the above scaling relation to estimate the frequency at which *INTEGRAL* could detect planet disruptions. In equation (3.16), we set $r_0 = r_h$ and $N_0 = 2M_\bullet/m_\star$ with $m_\star = 1 M_\odot$, i.e. corresponding to the sphere of influence (with this choice there is no dependence of τ on γ). The radius of influence r_h is subsequently calculated from $r_h = GM_\bullet/\sigma^2$, where the velocity dispersion σ is computed from the $M_\bullet - \sigma$ -relation (Ferrarese & Merritt 2000). With these assumptions, the resulting expression for the disruption rate is a function of SBH mass, $F_{\text{dis}} = F_{\text{dis}}(M_\bullet)$. Here, we set $F_{\text{dis}}(M_\bullet) = 0$ for $M_\bullet \gtrsim 10^7 M_\odot$ because for the latter masses the relaxation time-scale is longer than a Hubble time, and steady-state in energy can no longer be assumed (e.g. Merritt 2013b, Eq. 3.6). (In addition, for $M_\bullet \gtrsim 10^8 M_\odot$ the tidal disruption radius of the planet, similar to that of a $1 M_\odot$ star, is smaller than the Schwarzschild radius (e.g. fig. 6.1 of Merritt 2013b), implying that the planet would not be disrupted and eventually produce a potentially observable flare, but captured whole.)

Subsequently, we average F_{dis} over the observed mass function of SBHs in the centres of local galaxies,

$$\overline{F}_{\text{dis}} = \frac{\int_{M_{\bullet,\text{low}}}^{M_{\bullet,\text{up}}} dM_\bullet F_{\text{dis}}(M_\bullet) dN/dM_\bullet}{\int_{M_{\bullet,\text{low}}}^{M_{\bullet,\text{up}}} dM_\bullet dN/dM_\bullet}, \quad (3.18)$$

where $dN/dM_\bullet$ is given by equation (3) and Table 2 (first row) of Vika et al. (2009) and $M_{\bullet,\text{low}} = 10^6 M_\odot$ and $M_{\bullet,\text{up}} = 10^{10} M_\odot$.

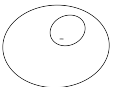
Multiplying the average rate implied by equation (3.18) by the number of galaxies within 50 Mpc, ~ 5000 , we find the following estimate for the rate of planet disruptions that are observable by *INTEGRAL*,

$$F_{\text{dis,planet}} \sim 0.05 N_{\text{p}/\star} \text{ yr}^{-1}, \quad (3.19)$$

where $N_{\text{p}/\star}$ is the number of planets per star. We therefore expect an observable planet disruption roughly every decade.

3.5.4 Changes of the stellar orbit prior to stripping

The debris disc of a star that was formed with pericentre distance $r_p \gg r_H \equiv d_2(3M_\bullet/m_\star)^{1/3}$ from the SBH is not tidally stripped by the SBH, where d_2 is the outer radius of the debris disc



(cf. Section 3.3.1). Gravitational encounters will gradually strip the debris disc over time at a time-scale of the order of $t_{\text{strip}} \sim 100$ Myr in the GC (cf. Section 3.3.2). In some cases, however, the time-scale for gravitational encounters with other stars to change the orbital properties of the *parent star* can be shorter than t_{strip} .

The time-scale for the orbital energy to change by order itself due to non-resonant relaxation (NRR), $t_{\mathcal{E}}(\mathcal{E}) \equiv [\langle (\Delta\mathcal{E})^2 \rangle / \mathcal{E}^2]^{-1}$, where $\langle (\Delta\mathcal{E})^2 \rangle$ is the second-order energy diffusion coefficient, is typically longer than the stripping time-scale. To illustrate this, we show the ratio $t_{\text{strip}}(\mathcal{E})/t_{\mathcal{E}}(\mathcal{E})$ as a function of $r = \psi^{-1}(\mathcal{E})$ in the top panel of Fig. 3.12. Here, t_{strip} is the orbit-averaged stripping time-scale where $a_{\text{bin}} = d_2$ (cf. Section 3.3.2). Furthermore, $t_{\mathcal{E}}(\mathcal{E})$ is computed from the late-type distribution function using the expression for $\langle (\Delta\mathcal{E})^2 \rangle$ in equation (21) of Cohn & Kulsrud (1978), without neglect of the stellar potential. At most radii $t_{\text{strip}}(\mathcal{E})/t_{\mathcal{E}}(\mathcal{E}) \ll 1$, demonstrating that stripping occurs faster than energy relaxation. However, the time-scale $t_L \sim (L/L_c)^2 t_{\mathcal{E}}$ for encounters to change the *orbital angular momentum* by order itself, where L_c is the angular momentum of a circular orbit, can be much shorter if the orbit is highly eccentric.

Nevertheless, even if $t_L \ll t_{\text{strip}}$ we do not expect our results in Section 3.4 to be much affected because NRR does not affect the *statistical* properties of the orbits of the parent stars in angular momentum. This is because the steady-state angular-momentum distribution due to NRR is an isotropic distribution (i.e. $f(\mathcal{E}, L) = f(\mathcal{E})$), which is the same distribution that was assumed in equation (3.9) (an exception is that $f(\mathcal{E}, L) \propto \log(L)$ for orbits very close to the SBH, Cohn & Kulsrud 1978).

3.5.5 Special case: a burst of flares?

In Sections 3.3 and 3.4, it was assumed that the planetesimals resulting in flares are bound to the SBH prior to their disruption, either because they were formed in a large-scale cloud bound to the SBH, or they were born in discs and stripped from their parent star. Here, we consider a special case in which the planetesimals are bound to a star prior to being disrupted by the SBH, possibly resulting in a burst of flares.

This special case could arise if the orbit of the star is highly eccentric with pericentre distance close to but slightly larger than the Hill radius r_H (cf. Section 3.5.4), and relaxation into an even more eccentric orbit subsequently drives the pericentre distance to $r'_p < r_H$. The tidal disruption radius of the star, $r_{\text{TD},*} \approx R_*(M_\bullet/m_*)^{1/3} \approx 0.7$ AU (assuming $m_* = 1 M_\odot$, $R_* = 1 R_\odot$ and $M_\bullet = 4 \times 10^6 M_\odot$), is approximately equal to that of the planetesimals. Therefore, if the new pericentre distance is $r'_p \approx r_{\text{lc}} = 1$ AU then we expect the star to be tidally disrupted by the SBH; the resulting flare would outshine the flares resulting from tidally disrupted planetesimals by many orders of magnitude.

However, if r'_p is slightly larger than the tidal disruption radius of the star then, owing to the extended size of the debris disc, some or all of the planetesimals might be stripped from the star, pass within 1 AU of the SBH and ultimately produce a flare, whereas the star is not tidally disrupted. To investigate this case, we performed a series of N -body simulations of a star with mass $1 M_\odot$ with a debris disc approaching the SBH ($M_\bullet = 4 \times 10^6 M_\odot$) on a highly eccentric orbit. These simulations, similar to those of Section 3.3.1, were carried out with the `HERMITE0` code (Hut et al. 1995) in the `AMUSE` framework (Pelupessy et al. 2013; Portegies Zwart et al. 2013). The orbit of the star has a semimajor of 1 pc, and the pericentre distance r'_p is varied between $0.1 r_{\text{lc}}$ and $2 r_H$, where $r_{\text{lc}} = 1$ AU and $r_H \equiv d_2(3M_\bullet/m_*)^{1/3} \approx 9.2 \times 10^3$ AU. The

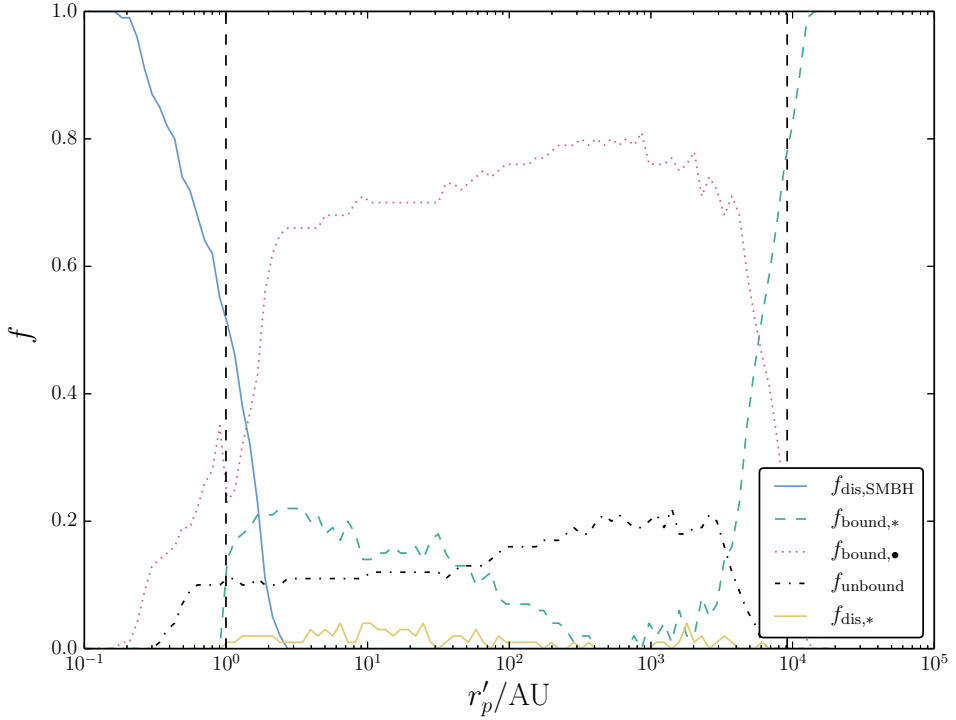
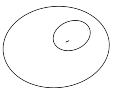


Figure 3.11: Fractions of the various outcomes as a function of the pericentre distance r'_p in the N -body simulations of a star with a disc approaching the SBH on a highly eccentric orbit, discussed in Section 3.5.5. Blue solid line: planetesimals that are disrupted by the SBH (i.e. pass the SBH within 1 AU); green dashed line: planetesimals that remain bound to the star; red dotted line: planetesimals that become bound to the SBH; black dot-dashed line: planetesimals that formally become unbound from the SBH (with low escape velocities, cf. Section 3.5.5); solid yellow line: planetesimals that collide with the star (the stellar radius is assumed to be $1 R_\odot$). The left (right) black vertical dashed lines indicate r_{lc} (r_{H}).



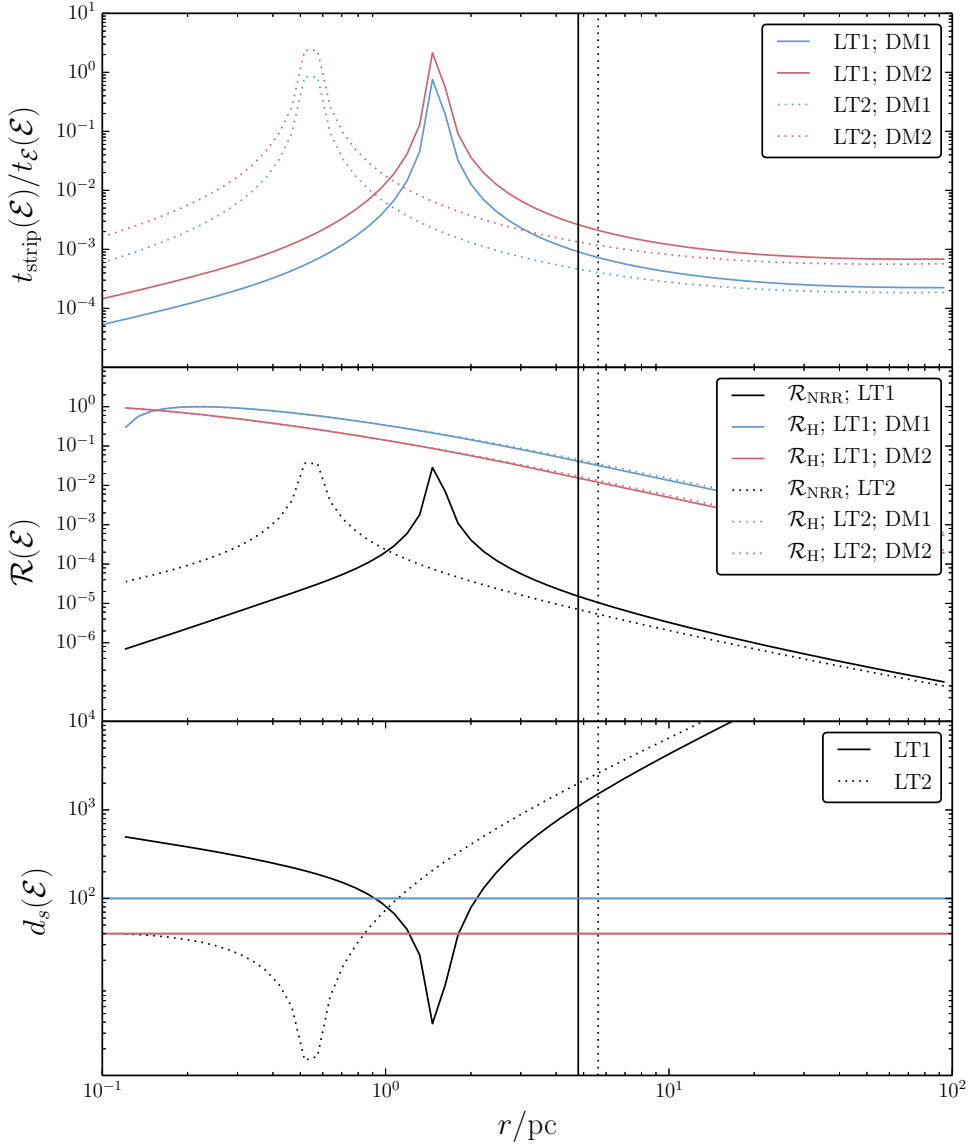


Figure 3.12: Various quantities as a function of $r = \psi^{-1}(\mathcal{E})$ discussed in Section 3.5.5. Two distributions of late-type stars are assumed: LT1 and LT2 (cf. Section 3.2.1). Blue (red) lines apply to DM1 (DM2) (cf. Section 3.2.2). **Top panel:** the ratio of the orbit-averaged stripping time-scale to the energy time-scale. **Middle panel:** the value of R , R_{NRR} , required for a gravitational encounter to sufficiently decrease the pericentre distance to produce a burst of flares, compared to R_{H} , the value of R for which the SBH strips the disc and a burst would not occur. **Bottom panel:** the stripping distance d_s at which the strong encounter would strip part or all of the disc.

debris disc is sampled consistently with DM1 (cf. Section 3.2.2) and the initial distance of the star to the SBH is $4 r_H$.

In Fig. 3.11, the fractions of the various outcomes in these simulations are shown as a function of r'_p . As expected, for $r'_p \gtrsim r_H$ (indicated with the right black vertical dashed line) all planetesimals remain bound to the star. For $r_{lc} < r'_p < r_H$ most (~ 0.8) planetesimals are stripped and become bound to the SBH. A smaller fraction (~ 0.2) becomes formally unbound from the SBH. We note that the escape speeds of these unbound planetesimals are low; the typical speed is 50 km/s. Therefore, it is likely that subsequent gravitational perturbations from stars could cause these planetesimals to be again bound to the SBH.

For $r'_p \lesssim r_{lc}$ (indicated with the left black vertical dashed line) most planetesimals are disrupted by the SBH. Disruption of planetesimals by the SBH also occurs for $r_p > r_{lc}$, but only for $r'_p \lesssim 2$ AU. Because the latter pericentre distance is close to $r_{TD,*}$, we expect that in this case the star would be either partially or fully disrupted, and therefore the resulting flare would be completely dominated by the star.

A burst of flares originating from multiple planetesimals can also be excluded based on an examination of the likelihood that a star with a debris disc would receive the perturbation necessary to change the pericentre from $r_p \geq r_H$ to $r'_p \ll r_H$. The following conditions need to be satisfied¹:

1. The initial pericentre distance $r_p = \theta r_H$ with $\theta \geq 1$, whereas the final pericentre distance $r'_p = \epsilon r_H$ with $r_{lc}/r_H < \epsilon \ll 1$. If $r_p < r_H$, then the star would have been stripped by the SBH at an earlier time. If ϵ is smaller than unity but not small enough, then most planetesimals will be stripped by the SBH but not pass within $r_{lc} = 1$ AU of the SBH (cf. Fig. 3.11). This also implies that the gravitational perturbation causing the pericentre change must act over a (radial) orbital time-scale $P_r(\mathcal{E})$.
2. The gravitational perturbations must be strong enough to produce the required decrease of the pericentre distance, but may not be too strong to disrupt the debris disc.

We assume that the pericentre changing encounters occur at apocentre and instantaneously change the apocentre velocity, whereas the apocentre distance and orbital energy remain unaffected. The change of the pericentre distance is therefore given by

$$\Delta r_p \approx \frac{\partial r_p(\mathcal{E}, \mathcal{R})}{\partial \mathcal{R}} \Delta \mathcal{R}, \quad (3.20)$$

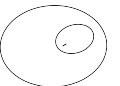
where $\mathcal{R} \equiv [L/L_c(\mathcal{E})]^2$ is a normalized angular-momentum variable. Furthermore, during $P_r(\mathcal{E})$ the change of angular momentum due to NRR is assumed to occur through a random walk,

$$\Delta L = L_c(\mathcal{E}) \left(\frac{P_r(\mathcal{E})}{t_{\mathcal{E}}(\mathcal{E})} \right)^{1/2}. \quad (3.21)$$

Combining equations (3.20) and (3.21) and writing the pericentre change as $\Delta r_p/r_p = \epsilon/\theta - 1 \approx -1$, we find the following relation for the angular momentum satisfying the requirement of the pericentre change,

$$\mathcal{R}_{\text{NRR}}(\mathcal{E}) \approx \frac{4P_r(\mathcal{E})}{t_{\mathcal{E}}(\mathcal{E})} \left(\frac{\epsilon}{\theta} - 1 \right)^{-2} \left(\frac{\partial \log r_p(\mathcal{E}, \mathcal{R})}{\partial \log \mathcal{R}} \right)^2. \quad (3.22)$$

¹The basic problem discussed here is similar to that of Murray-Clay & Loeb (2012).



The logarithmic derivative factor in equation (3.22) is a function of $\mathcal{R}$, implying that equation (3.22) cannot in general be solved analytically for $\mathcal{R}_{\text{NRR}}$. However, this factor is a weak function of $\mathcal{R}$ for $\mathcal{R} \ll 1$ and close to unity in this limit (in the case when the stellar potential is neglected, it is exactly equal to unity); for the computations below we set it equal to unity.

The quantity $\mathcal{R}_{\text{NRR}}(\mathcal{E})$ is shown as a function of $r = \psi^{-1}(\mathcal{E})$ with black solid (dotted) lines in the middle panel of Fig. 3.12 for late-type distribution LT1 (LT2), assuming $\epsilon/\theta = 10^{-3}$. There is a peak where the energy relaxation time-scale $t_{\mathcal{E}}(\mathcal{E})$ is shortest. In addition we show in the second panel of Fig. 3.12 with the blue (red) lines the quantity $\mathcal{R}_{\text{H}}$, defined as the value of R corresponding to $r_{\text{p}} = r_{\text{H}}$ assuming DM1 (DM2) and for both LT1 and LT2.

According to requirement (i) above $\mathcal{R} = \mathcal{R}_{\text{NRR}} > \mathcal{R}_{\text{H}}$ must be the case; however, this is not satisfied at any radii. In other words, in order for the encounter to be sufficiently strong the orbit must be highly eccentric, but in this case the tidal force of the SBH would have stripped the star at an earlier epoch. For completeness, we discuss the other requirement below.

The debris disc can survive the encounters if $d_2 < d_s$, where d_s is the stripping distance, which can be estimated from requiring that the relative velocity v_{rel} between the star and planetesimal exceeds the escape velocity v_{esc} from the star. The former can be estimated as $v_{\text{rel}} \sim (Gm_{\text{p}}/b^2)(2d/b)(2b/v_{\text{enc}})$ (Murray-Clay & Loeb 2012), where b is the impact parameter, m_{p} is the perturber mass, assumed to be $m_{\text{p}} = m_{\star} = 1 M_{\odot}$, and v_{enc} is the encounter velocity, which we approximate with the stellar velocity dispersion σ (cf. equation 3.6); the latter is given by $v_{\text{esc}} = (2Gm_{\star}/d)^{1/2}$. In terms of the change of the stellar speed, the impact parameter can be estimated by $b \sim 2Gm_{\text{p}}/(v_{\text{enc}}\Delta v_{\text{a}})$ where Δv_{a} is the change of the apocentre speed (assuming that the encounters occur at apocentre). Setting $v_{\text{rel}} = v_{\text{esc}}$ subsequently gives the following relation for the stripping distance,

$$d_s^3 \sim \frac{2(Gm_{\star})^3}{v_{\text{enc}}^2 \Delta v_{\text{a}}^4}. \quad (3.23)$$

From the assumptions $\Delta r_{\text{a}} = 0$ and $\Delta \mathcal{E} = 0$, the definition of $\mathcal{R}$ and the relation $L = r_{\text{a}}v_{\text{a}}$, it follows that

$$\Delta v_{\text{a}}(\mathcal{E}, \mathcal{R}) = \frac{L_{\text{c}}(\mathcal{E})}{2r_{\text{a}}(\mathcal{E}, \mathcal{R})} \frac{\Delta \mathcal{R}}{\sqrt{\mathcal{R}}}. \quad (3.24)$$

Substituting $\mathcal{R} = \mathcal{R}_{\text{NRR}}$ (cf. equation 3.22) and using equation (3.20) we find

$$\Delta v_{\text{a}}(\mathcal{E}) = \frac{L_{\text{c}}(\mathcal{E})}{r_{\text{a}}(\mathcal{E}, \mathcal{R}_{\text{NRR}})} \left(\frac{P(\mathcal{E})}{t_{\mathcal{E}}(\mathcal{E})} \right)^{1/2}. \quad (3.25)$$

Combining equations (3.23) and (3.25) then gives the stripping distance d_s as a function of $\mathcal{E}$, which is plotted in the bottom panel of Fig. 3.12 for LT1 and LT2. The outer disc radii in DM1 and DM2 are shown with horizontal lines. Destructive stripping, i.e. $d_s < d_2$, only occurs at a narrow radial range near $r \sim 1$ pc, therefore this does not pose a major problem in the scenario.

3.6 Conclusions

We have studied the dynamics of planetesimals in galactic nuclei focusing on the GC, and in the context of the tidal disruption of planetesimals by the SBH, possibly producing observable

near-infrared/X-ray flares. We assumed that the planetesimals were either formed in a large-scale cloud bound to the SBH, or in debris discs around stars. Our main conclusions are as follows.

(1) Assuming that planetesimals in the GC were initially formed in debris discs around stars, the tidal force of the SBH is effective at stripping the planetesimals from their parent stars at distances $r \lesssim 0.5$ pc from the SBH. Stripping by gravitational encounters with other stars is effective at stripping nearly all planetesimals within the radius of influence, ≈ 4 pc, after ~ 100 Myr.

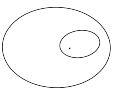
(2) We studied the orbital evolution of planetesimals bound to the SBH in response to gravitational scattering from late-type stars, a hypothesized cusp of stellar black holes close to the SBH, and massive perturbers. We also included other effects such as physical collisions and RR. We found that the disruption rate of planetesimals by the SBH at $t = 10$ Gyr is $\approx 0.6 \text{ d}^{-1}$, which is roughly consistent with the observed rate of the flares of once per day and the previous, less detailed, estimate by ZNM12, $\sim 1 \text{ d}^{-1}$. Moreover, the rate is insensitive to model assumptions, in particular the initial distribution of planetesimals, i.e. whether the planetesimals were formed in a large-scale cloud or in debris discs around stars, and details of the gravitational perturbers. By comparing our rates in the case of formation in discs to the observed rates of Neilsen et al. (2013), we have constrained the number of planetesimals per star to $N_{a/\star} \approx 3.7 \times 10^7$; taking into account the effect of massive perturbers, we find $N_{a/\star} \approx 1.1 \times 10^8$.

(3) The result that both formation in a large-scale cloud and formation in debris discs around stars are consistent with the observed flaring rate suggests that it is not possible to distinguish between these two cases on the basis of the latter observation alone. However, the assumed number of planetesimals per star, $N_{a/\star} = 2 \times 10^7$, is inferred from observations of debris discs around stars in the Solar neighbourhood. In the case of formation in a large-scale cloud, this implies that the number of planetesimals formed is strongly correlated with the number of stars, and this requires finetuning of the quantity $N_{a/\star}$. We favour the more natural explanation that planetesimals in galactic nuclei similar to the GC are formed no differently than planetesimals around stars in the Solar neighbourhood.

(4) We have extrapolated our results to different galactic nuclei and we have estimated the event rate of the tidal disruption of planets by SBHs. Assuming one planet per star, we expect an observable planet disruption originating within the local Universe ($D < 50$ Mpc) roughly every decade.

Acknowledgements

We would like to thank S. Markoff and S. Nayakshin for enlightening discussions on the disruption of planetesimals in the Galactic Centre, and the referee, Hagai Perets, for providing very useful comments that helped to improve the paper. This work was supported by the Netherlands Research Council NWO (grants #639.073.803 [VICI], #614.061.608 [AMUSE] and #612.071.305 [LGM]) and the Netherlands Research School for Astronomy (NOVA).



3.A Planetesimal stripping

3.A.1 Stripping by the SBH

In the middle panel of Fig. (3.1), the stripping fraction f_{strip} is expressed in terms of the pericentre distance r_p of the orbit of the star around the SBH. Here, we discuss our method to express f_{strip} in terms of the orbital energy, assuming an isotropic velocity distribution.

First, we express r_p in terms of energy $\mathcal{E}$ and $\mathcal{R}$, where angular momentum L is expressed in terms of $\mathcal{R} \equiv (L/L_c)^2 \in [0, 1]$; here, $L_c(\mathcal{E})$ is the angular momentum of a circular orbit with energy $\mathcal{E}$. The latter is given in terms of the potential by $L_c^2 = -r_c^3 d\psi/dr|_{r=r_c}$, with $r_c = r_c(\mathcal{E})$ the radius of a circular orbit with energy $\mathcal{E}$, which is the solution of the equation $2[\psi(r_c) - \mathcal{E}] + r_c d\psi/dr|_{r=r_c} = 0$ (e.g. Cohn 1979; Merritt 2013b, 5.5.1). We determine the pericentre distance r_p by finding the smallest solution $r = r(\mathcal{E}, \mathcal{R})$ for which the radial velocity $v_r = v_r(r, \mathcal{E}, \mathcal{R}) = \{2[\psi(r) - \mathcal{E}] - \mathcal{R} L_c^2(\mathcal{E})/r^2\}^{1/2}$ vanishes. From $r_p(\mathcal{E}, \mathcal{R})$, we compute the stripping fraction $f_{\text{strip}}(\mathcal{E}, \mathcal{R})$. Subsequently, we average this quantity over $\mathcal{R}$ assuming an isotropic velocity distribution, $f_{\text{strip}}(\mathcal{E}) \equiv \int_{\mathcal{R}_{\text{lc}}(\mathcal{E})}^1 d\mathcal{R} f_{\text{strip}}(\mathcal{E}, \mathcal{R})$. Here, $\mathcal{R}_{\text{lc}}(\mathcal{E})$ is the value of $\mathcal{R}$ that corresponds to the loss cone r_{lc} of the star; we adopt $r_{\text{lc}} = 1 \text{ AU}$, approximately the tidal disruption radius of a solar-type star. An orbit with energy $\mathcal{E}$ and $\mathcal{R} = \mathcal{R}_{\text{lc}}$ just grazing the loss cone r_{lc} at pericentre has $v_r(r_{\text{lc}}, \mathcal{E}, \mathcal{R}_{\text{lc}}) = v_r(r_p, \mathcal{E}, R_{\text{lc}}) = 0$, which implies $\mathcal{R}_{\text{lc}} = [2r_{\text{lc}}^2/L_c^2(\mathcal{E})][\psi(r_{\text{lc}}) - \mathcal{E}]$.

The resulting stripping fractions $f_{\text{strip}}(\mathcal{E})$ are plotted in the bottom panel of Fig. 3.1 with solid lines. Analytic expressions for $f_{\text{strip}}(\mathcal{E})$ can be derived if the stellar potential is neglected. In that case, the pericentre distance $r_p(\mathcal{E}, \mathcal{R}) = [GM_\bullet/(2\mathcal{E})][1 - \sqrt{1 - \mathcal{R}}] = a(1 - e)$ and $d_{\text{strip}} = af_H(1 - e)$, where $f_H \equiv [m_\star/(3M_\bullet)]^{1/3}$ (cf. equation 3.5). Using equation 3.4 and assuming that $\mathcal{R}_{\text{lc}} = 0$, we find for the angular-momentum-averaged stripping fraction $f_{\text{strip}}(a) \approx \int_0^1 d\mathcal{R} f_{\text{strip}}(\mathcal{E}, \mathcal{R}) = \int_0^1 da \, 2ef_{\text{strip}}(a, e)$, for $af_H > d_2$,

$$f_{\text{strip}}(a) = \frac{1}{d_2^{2-\beta} - d_1^{2-\beta}} \left\{ d_2^{2-\beta} \left[\left(1 - \frac{d_1}{af_H}\right)^2 - \left(1 - \frac{d_2}{af_H}\right)^2 \right] - 2 \left[(af_H)^{2+\beta} (\beta - 4)(\beta - 3) \right]^{-1} \right. \\ \left. \times \left[d_1^3 \left(\frac{d_1}{af_H} \right)^{-\beta} (af_H(\beta - 4) - d_1(\beta - 3)) - d_2^3 \left(\frac{d_2}{af_H} \right)^{-\beta} (af_H(\beta - 4) - d_2(\beta - 3)) \right] \right\} + 1 - \left(1 - \frac{d_1}{af_H}\right)^2 ;$$

for $d_1 \leq a f_H \leq d_2$,

$$f_{\text{strip}}(a) = \frac{1}{d_2^{2-\beta} - d_1^{2-\beta}} \left\{ d_2^{2-\beta} \left(1 - \frac{d_1}{a f_H} \right)^2 - 2(a f_H)^{2-\beta} \left[\frac{d_1^3 \left(\frac{d_1}{a f_H} \right)^{-\beta}}{(a f_H)^4 (\beta - 4)(\beta - 3)} + \frac{1}{12 - 7\beta + \beta^2} \right] \right\} + 1 - \left(1 - \frac{d_1}{a f_H} \right)^2 \quad (3.26)$$

and $f_{\text{strip}}(a) = 1$ for $a f_H < d_1$. The stripping fractions according to this approximation are shown with dashed lines in the bottom panel of Fig. 3.1.

3.A.2 Stripping by gravitational encounters

We compute the orbital average of the local stripping fraction arising from gravitational encounters using the relation (e.g. Merritt 2013b, 5.5.2)

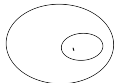
$$f_{\text{strip}}(\mathcal{E}) = \frac{4}{p(\mathcal{E})} \int_0^{r^{\psi^{-1}(\mathcal{E})}} dr r^2 v(r, \mathcal{E}) f_{\text{strip}}(r), \quad (3.27)$$

where $p(\mathcal{E})$ is the phase space volume per unit energy and is given by

$$p(\mathcal{E}) = 4 \int_0^{r^{\psi^{-1}(\mathcal{E})}} dr r^2 v(r, \mathcal{E}), \quad (3.28)$$

with $\psi^{-1}(\mathcal{E})$ the inverse function of $\psi(r)$ and $v(r, \mathcal{E}) = \sqrt{2[\psi(r) - \mathcal{E}]}$ the orbital speed at radius r for an orbit with energy $\mathcal{E}$. In equation (3.27), an isotropic velocity distribution is assumed; it was derived from the more general expression for the orbital average of a local function $f(r)$ (e.g. Cohn & Kulsrud 1978),

$$f(\mathcal{E}, \mathcal{R}) = \frac{2}{P(\mathcal{E}, \mathcal{R})} \int_{r_-(\mathcal{E}, \mathcal{R})}^{r_+(\mathcal{E}, \mathcal{R})} dr \frac{v_t(r, \mathcal{E}, \mathcal{R})}{v_t(r, \mathcal{E}, \mathcal{R})} f(r), \quad (3.29)$$



where $P(\mathcal{E}, \mathcal{R}) = 2 \int_{r_-}^{r_+} dr / v_r$ is the orbital period and $r_{\pm}(\mathcal{E}, \mathcal{R})$ are the turning points of the orbit, i.e. the solutions to $v_r(r_{\pm}, \mathcal{E}, R) = 0$. In this derivation, several approximations have been made. To verify the validity of the simplified orbit average, equation (3.27), we have computed $f_{\text{strip}}(\mathcal{E})$ numerically using both equations (3.27) and (3.29), in the latter case averaging over R after calculating $f(\mathcal{E}, R)$. We find that there is no discernable difference in the resulting averaged function $f_{\text{strip}}(\mathcal{E})$ between the two cases.

3.B Terms appearing in the Fokker-Planck equation

3.B.1 Gravitational scattering flux

The quantity $\mathcal{F}_{\mathcal{E}}$ represents a flux in $\mathcal{E}$ -space because of gravitational scattering by massive scatterers and is given by (e.g. Merritt 2013b, 7.1.2.2)

$$\mathcal{F}_{\mathcal{E}}(\mathcal{E}, t) = D_{\mathcal{E}\mathcal{E}}(\mathcal{E}) \frac{\partial f_a(\mathcal{E}, t)}{\partial \mathcal{E}}, \quad (3.30)$$

where $D_{\mathcal{E}\mathcal{E}}(\mathcal{E})$ is an energy diffusion coefficient that depends on the distribution function of the scatterers. Assuming a discrete mass spectrum of scatterers, $D_{\mathcal{E}\mathcal{E}}$ is given by

$$D_{\mathcal{E}\mathcal{E}}(\mathcal{E}) = \sum_j m_j^2 \ln(\Lambda_j) D_{\mathcal{E}\mathcal{E};j} = 64\pi^5 G^2 \sum_j m_j^2 \ln(\Lambda_j) \times \left[q(\mathcal{E}) \int_{-\infty}^{\mathcal{E}} d\mathcal{E}' f_j(\mathcal{E}') + \int_{\mathcal{E}}^{\infty} d\mathcal{E}' f_j(\mathcal{E}') q(\mathcal{E}') \right], \quad (3.31)$$

where $\ln(\Lambda_j)$ is the Coulomb logarithm, $f_j(\mathcal{E})$ is the distribution function of massive scatterer j with mass m_j and

$$q(\mathcal{E}) = \frac{4}{3} \int_0^{\psi^{-1}(\mathcal{E})} dr r^2 v^3(r, \mathcal{E}). \quad (3.32)$$

Here, we adopt $\ln(\Lambda_j) = \ln[M_{\bullet}/(2m_j)]$. The distribution functions $f_j(\mathcal{E})$ are computed from the number density $n_j(r)$ and potential $\psi(r)$ using Eddington's formula (Eddington 1916),

$$f_j(\mathcal{E}) = \frac{\sqrt{2}}{4\pi^2} \frac{\partial}{\partial \mathcal{E}} \int_{-\infty}^{\psi^{-1}(\mathcal{E})} dr \frac{dn_j(r)}{\sqrt{\mathcal{E} - \psi(r)}}. \quad (3.33)$$

Black hole cusp

In the case of a cusp of stellar black holes (cf. Section 3.2.1) we derive and adopt semi-analytic expressions for $D_{\mathcal{E}\mathcal{E};\text{BH}}(\mathcal{E})$. We assume that the distribution function $f_{\text{BH}}(\mathcal{E})$ has the form of a truncated power law,

$$f_{\text{BH}}(\mathcal{E}) = \begin{cases} f_0 \mathcal{E}^{\gamma-3/2}, & \mathcal{E} > \mathcal{E}_d; \\ 0, & \mathcal{E} \leq \mathcal{E}_d. \end{cases} \quad (3.34)$$

Here, $\mathcal{E}_d \equiv GM_\bullet/(2r_d)$ and $r_d = 0.2$ pc is assumed; furthermore we adopt $\gamma = 1.8$. The quantity f_0 can be expressed in terms of the number density of stellar black holes at $r = r_d$, $n_d = N_{\text{BH}}(3 - \gamma)/(4\pi r_d^3)$, where $N_{\text{BH}} = 4800$ is the total number of black holes within r_d , via

$$f_0 = \frac{\sqrt{2}}{4} \pi^{-3/2} \frac{\Gamma(\gamma + 1)}{\Gamma(\gamma - \frac{1}{2})} \left(\frac{r_d}{GM_\bullet} \right)^\gamma n_d, \quad (3.35)$$

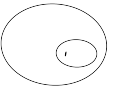
where $\Gamma(x)$ is the Gamma function. For $r \leq r_d$, it is well justified to neglect the stellar potential in $\psi(r)$. With this approximation equation (3.31) yields

$$D_{\mathcal{E}\mathcal{E};\text{BH}}(\mathcal{E}) = \frac{32}{3} \sqrt{2} \pi^6 G^2 (GM_\bullet)^3 f_0 \frac{1}{(2 - \gamma)(2\gamma - 1)} \times \begin{cases} 3\mathcal{E}^{\gamma-2} - 2(2 - \gamma)\mathcal{E}_d^{\gamma-1/2} \mathcal{E}^{-3/2}, & \mathcal{E} > \mathcal{E}_d; \\ (2\gamma - 1)\mathcal{E}_d^{\gamma-2}, & \mathcal{E} \leq \mathcal{E}_d. \end{cases} \quad (3.36)$$

3.B.2 Flux into the loss cone

As discussed in Section 3.4.1, relaxation through gravitational encounters typically occurs much faster in angular-momentum space than in energy space. Scattering into the loss cone of the SBH is therefore dominated by diffusion in angular momentum rather than diffusion in energy. The former is described in equation (3.9) by the term $F_{\text{lc}}(\mathcal{E}, t)$, the flux in angular-momentum space into the loss cone. We adopt the formalism of the Cohn-Kulsrud boundary layer, which is based on the Fokker-Planck equation in angular-momentum space (Cohn & Kulsrud 1978; Merritt 2013b, 6.1.2). By solving this equation both in local and orbit-averaged form and matching the two solutions, the following expression can be derived for the flux into the loss cone,

$$F_{\text{lc}}(\mathcal{E}, t) \approx 4\pi^2 P_{\text{t}}(\mathcal{E}) L_{\text{c}}^2(\mathcal{E}) \bar{\mu}(\mathcal{E}) \left\{ \ln [\mathcal{R}_0(\mathcal{E})^{-1}] \right\}^{-1} f_{\text{a}}(\mathcal{E}, t). \quad (3.37)$$



Here, $P_r(\mathcal{E}) = \int_0^{\psi^{-1}(\mathcal{E})} dr / v_r(r, \mathcal{E}, 0)$ is the radial orbital period and $\bar{\mu}(\mathcal{E})$ is an orbit-averaged angular-momentum diffusion coefficient in the limit $\mathcal{R} \rightarrow 0$, defined by

$$\bar{\mu}(\mathcal{E}) = \frac{2}{P_r(\mathcal{E})} \int_0^{\psi^{-1}(\mathcal{E})} dr \frac{\langle (\Delta \mathcal{R})^2 \rangle}{v_r(r, \mathcal{E}, 0)} \lim_{\mathcal{R} \rightarrow 0} \frac{\langle (\Delta \mathcal{R})^2 \rangle}{(2\mathcal{R})}, \quad (3.38)$$

where $\langle (\Delta \mathcal{R})^2 \rangle$ is the second-order diffusion coefficient in $\mathcal{R}$ arising from NRR. Using standard expressions for $\langle (\Delta \mathcal{R})^2 \rangle$ (e.g. Cohn & Kulsrud 1978), $\bar{\mu}(\mathcal{E})$ can be expressed in terms of f_j as

$$\bar{\mu}(\mathcal{E}) = \frac{3^2 \pi^2 G^2}{3 P_r(\mathcal{E}) L_c^2(\mathcal{E})} \sum_j m_j^2 \ln(\Lambda_j) [3 \bar{I}_{1/2;j}(\mathcal{E}) + 2 \bar{I}_{0;j}(\mathcal{E}) - \bar{I}_{3/2;j}(\mathcal{E})], \quad (3.39)$$

where

$$\begin{aligned} \bar{I}_{0;j}(\mathcal{E}) &= \int_0^{\psi^{-1}(\mathcal{E})} dr r^2 \frac{d\mathcal{E}'}{\sqrt{2[\psi(r) - \mathcal{E}']}} \int_{-\infty}^{\mathcal{E}} d\mathcal{E}' f_j(\mathcal{E}'); \\ \bar{I}_{n/2;j}(\mathcal{E}) &= \int_0^{\psi^{-1}(\mathcal{E})} dr r^2 \frac{d\mathcal{E}'}{\sqrt{2[\psi(r) - \mathcal{E}']}} \int_{\mathcal{E}}^{\psi(r)} d\mathcal{E}' \left(\frac{\psi(r) - \mathcal{E}'}{\psi(r) - \mathcal{E}} \right)^{n/2} f_j(\mathcal{E}'). \end{aligned} \quad (3.40)$$

In equation (3.37), it is assumed that the distribution function f_a decreases logarithmically with $\mathcal{R}$ as $\mathcal{R} \rightarrow 0$. The value of $\mathcal{R}$ for which f_a vanishes is not the capture boundary $\mathcal{R}_{lc}(\mathcal{E})$: far away from the SBH, where the orbital period is relatively long compared to the angular-momentum relaxation time-scale (the ‘full loss cone’ regime), a planetesimal can be scattered into and out of the loss cone without being immediately disrupted, i.e. it may be that $f_a(\mathcal{E}, \mathcal{R}_{lc}) > 0$. The value of $\mathcal{R}$ for which the distribution function does vanish is given by the function $\mathcal{R}_0(\mathcal{E}) = \mathcal{R}_{lc}(\mathcal{E}) \exp[-\bar{q}/\xi(\bar{q})] < \mathcal{R}_{lc}$, where $\bar{q} = \bar{q}(\mathcal{E}) = P_r(\mathcal{E}) \bar{\mu}(\mathcal{E}) / \mathcal{R}_{lc}(\mathcal{E})$ and $\xi(\bar{q})$ is given by

$$\xi(\bar{q}) \equiv 1 - \sum_{m=1}^{\infty} \frac{\exp(-\alpha_m^2 \bar{q}/4)}{\alpha_m^2},$$

where α_m is the m^{th} zero of the Bessel function $J_0(\alpha)$ of the first kind. We compute the quantity $\mathcal{R}_{lc}$ from $\mathcal{R}_{lc} = 2[r_{lc}^2 / L_c^2(\mathcal{E})][\psi(r_{lc}) - \mathcal{E}]$ with $r_{lc} = 1 \text{ AU}$.

Black hole cusp

Similarly as in Section 3.B.1, we derive and adopt semi-analytic expressions for $\bar{\mu}(\mathcal{E})$ for the case of a cusp of stellar black holes. With the same assumptions as in Section 3.B.1, the loss cone integrals $\bar{I}_{0;\text{BH}}$ and $\bar{I}_{n/2;\text{BH}}$ (cf. equation 3.40) can be expressed as

$$\bar{I}_{0;\text{BH}}(\mathcal{E}) = \frac{5\sqrt{2}\pi}{16} \frac{f_0}{2\gamma - 1} (GM_\bullet)^3 \times \begin{cases} \mathcal{E}^{\gamma-4} - \mathcal{E}_d^{\gamma-1/2} \mathcal{E}^{-7/2}, & \mathcal{E} > \mathcal{E}_d; \\ 0, & \mathcal{E} \leq \mathcal{E}_d; \end{cases} \quad (3.41)$$

$$\bar{I}_{n/2;\text{BH}}(\mathcal{E}) = \frac{\sqrt{2}}{2n+1} f_0 (GM_\bullet)^3 \times \begin{cases} g_n(\mathcal{E}, \mathcal{E}, \gamma), & \mathcal{E} > \mathcal{E}_d; \\ g_n(\mathcal{E}, \mathcal{E}_d, \gamma), & \mathcal{E} \leq \mathcal{E}_d, \end{cases} \quad (3.42)$$

where

$$g_1(\mathcal{E}, \mathcal{E}', \gamma) = \mathcal{E}'^{\gamma-4} \int_0^1 dx x^{3-\gamma} \left(1 - \frac{\mathcal{E}}{\mathcal{E}'x}\right)^{-1} (1-x)^{3/2} {}_2F_1\left(\frac{3}{2}, \frac{3}{2} - \gamma, \frac{5}{2}; 1-x\right);$$

$$g_3(\mathcal{E}, \mathcal{E}', \gamma) = \mathcal{E}'^{\gamma-4} \int_0^1 dx x^{3-\gamma} \left(1 - \frac{\mathcal{E}}{\mathcal{E}'x}\right)^{-2} (1-x)^{5/2} {}_2F_1\left(\frac{5}{2}, \frac{3}{2} - \gamma, \frac{7}{2}; 1-x\right).$$

Here, ${}_2F_1(a, b, c; x)$ is the Gauss hypergeometric function.

3.B.3 Collision flux

The term F_{col} in equation (3.9) describes losses of planetesimals because of physical collisions with other objects. We consider collisions of planetesimals with late-type stars and with other planetesimals of the same size. The collision cross-section including gravitational focusing is $\sigma_{\text{col};j} = \pi R_j^2 (1 + v_{\text{esc};j}^2/v_{\text{rel}}^2)$, where $R_j \approx R_{\text{L},T}$ in the case of collisions with late-type stars and $R_j = 2R_a$ in the case of collisions with other planetesimals; we adopt $R_a = 10 \text{ km}$ (cf. Section 3.2.2). The escape speed $v_{\text{esc};j} = (2Gm_j/R_j)^{1/2}$, where we assume $m_j = m_{\text{L},T}$ and $m_j = 2 \times 10^{-15} M_\odot$ for collisions with late-type stars and planetesimals, respectively, and $v_{\text{rel}} = |\mathbf{v} - \mathbf{v}'|$ is the relative speed at infinity. The collision rate $\nu_{\text{col};j}$ is subsequently found by replacing n in $n_j \sigma_{\text{col};j} v_{\text{rel}}$ with the impactor distribution function $f_j(v')$ and integrating

over all impactor velocities v' (e.g. Cohn & Kulsrud 1978),

$$\begin{aligned}\nu_{\text{col};j}(v) &= \int d^3v' f_j(v') \sigma_{\text{col};j}(|\mathbf{v} - \mathbf{v}'|) |\mathbf{v} - \mathbf{v}'| \\ &= 2\pi R_j^2 g_j(v) + 2\pi G m_j R_j h_j(v).\end{aligned}\quad (3.43)$$

Here, $g_j(v)$ and $h_j(v)$ are the ‘Rosenbluth potentials’ $h_j(v) = \int d^3v' f_j(v') |\mathbf{v} - \mathbf{v}'|$ and $h_j(v) = 2 \int d^3v' f_j(v') |\mathbf{v} - \mathbf{v}'|^{-1}$ (Rosenbluth et al. 1957). We subsequently average $\nu_{\text{col};j}(v)$ over a planetesimal orbit using an equation similar to equation (3.27) with $v = \sqrt{2[\psi(r) - \mathcal{E}]}$ and obtain $\nu_{\text{col};t;j}(\mathcal{E}) = 2\pi R_j^2 g_{t;j}(\mathcal{E}) + 2\pi G m_j R_j h_{t;j}(\mathcal{E})$, where

$$\begin{aligned}g_{t;j}(\mathcal{E}) &= \frac{4\pi}{3} \frac{4}{p(\mathcal{E})} \int_0^{\psi^{-1}(\mathcal{E})} dr r^{-2} v(r, \mathcal{E}) \int_{\mathcal{E}}^{\psi(r)} d\mathcal{E}' f_j(\mathcal{E}') \left[\left(\frac{\psi(r) - \mathcal{E}'}{\psi(r) - \mathcal{E}} \right)^{1/2} [8\psi(r) - 6\mathcal{E} - 2\mathcal{E}'] + 8\psi(r) - 2\mathcal{E} - 6\mathcal{E}' \right]; \\ h_{t;j}(\mathcal{E}) &= 8\pi \frac{4}{p(\mathcal{E})} \int_0^{\psi^{-1}(\mathcal{E})} dr r^{-2} v(r, \mathcal{E}) \int_{\mathcal{E}}^{\psi(r)} d\mathcal{E}' f_j(\mathcal{E}') \left(\frac{\psi(r) - \mathcal{E}'}{\psi(r) - \mathcal{E}} \right)^{1/2} + 8\pi \int_{-\infty}^{\mathcal{E}} d\mathcal{E}' f_j(\mathcal{E}').\end{aligned}$$

(note that in the orbit-averaged expressions for the collision rate in equation (43) of Cohn & Kulsrud 1978 the stellar potential was neglected).

The collision rate $\nu_{\text{col};t;j}(\mathcal{E})$ gives the rate of collisions with impactor j for a single planetesimal at energy $\mathcal{E}$; the flux for all planetesimals at energy $\mathcal{E}$, $F_{\text{col};j}(\mathcal{E}, t) = \nu_{\text{col};t;j}(\mathcal{E}) N_a(\mathcal{E}, t)$. The combined collision flux for both types of collisions is $F_{\text{col}}(\mathcal{E}, t) = \sum_j F_{\text{col};j}(\mathcal{E}, t)$.

In the case of planetesimal-planetesimal collisions, there is a non-linear dependence of $F_{\text{col};a}(\mathcal{E}, t)$ on the planetesimal distribution function f_a . To simplify the computations, we use the time-independent (and therefore strictly incorrect) distribution function $f_a(\mathcal{E}) = N_{a/*} \times f_{\text{LT}}(\mathcal{E})$ to compute $\nu_{\text{col};t;a}$ instead of the actual $f_a(\mathcal{E}, t)$, the function for which equation (3.9) is to be solved. The former distribution function is appropriate for complete stripping of all planetesimals from the late-type stars, and without further change of the distribution function. This gives an upper limit for the importance of planetesimal-planetesimal collisions. It turns out that planetesimal-planetesimal collisions, even assuming this upper limit, are negligible compared to planetesimal-late-type star collisions; e.g., the rates for planetesimal-planetesimal collisions are unaffected compared to the main model, model 1, whereas this is not the case for late-type star-planetesimal collisions, cf. Table 3.1. This justifies our simplified treatment of planetesimal-planetesimal collisions.

Flux into the loss cone from resonant relaxation

Close to the SBH, the stellar potential is not perfectly spherically symmetric because the number of stars is finite. This asymmetry gives rise to resonant relaxation (RR) and results in periodic changes of the orbital angular momenta on time-scales that can be much shorter than NRR

time-scale if sufficiently close to the SBH (Rauch & Tremaine 1996). The process of RR can potentially increase the rate of captures by the SBH arising from NRR in angular-momentum space alone (cf. Section 3.B.2). An approximate method to model (incoherent) RR arising from a given distribution of stars in equation (3.9) is to include the sink term $F_{\text{RR}}(\mathcal{E}, t)$, which is essentially the number of planetesimals with energies between $\mathcal{E}$ and $\mathcal{E} + d\mathcal{E}$, divided by the time-scale for RR to decrease the angular momentum to the value corresponding to disruption by the SBH (Hopman & Alexander 2006a; Merritt et al. 2011),

$$F_{\text{RR}}(\mathcal{E}, t) = \chi_{\text{RR}} \frac{N_a(\mathcal{E}, t)}{\ln [\mathcal{R}_{\text{lc}}(\mathcal{E})^{1/2}]} \frac{1}{T_{\text{RR}}};$$

$$T_{\text{RR}} = \beta_s^{-2} \left(\frac{M_\bullet}{m_\star} \right)^2 N(\mathcal{E})^{-1} \frac{P(\mathcal{E})^2}{t_{\text{coh}}(\mathcal{E})}. \quad (3.44)$$

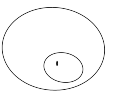
Here, β_s describes the efficiency of RR; numerical investigations indicate that $\beta_s \approx 1.6$ in the Newtonian regime (Gürkan & Hopman 2007; Eilon et al. 2009; Hamers et al. 2014). Averaging this quantity over a thermal distribution yields $\beta_s \approx 1.1$; here we adopt $\beta_s = 1$. The factor $|\ln[\mathcal{R}_{\text{lc}}^{1/2}(\mathcal{E})]|$ is approximately the typical number of relaxation time-scales for the angular momentum to decrease from the maximal value corresponding to a circular orbit, to the value corresponding to capture by the SBH.

The quantity $t_{\text{coh}}(\mathcal{E})$ in equation (3.44) is the coherence time-scale, i.e. the typical time-scale for field stars to change their orientation with respect to a test orbit with energy $\mathcal{E}$. Here, we adopt $t_{\text{coh}}^{-1}(\mathcal{E}) = t_{\text{MP}}^{-1}(\mathcal{E}) + t_{\text{GR}}^{-1}(\mathcal{E})$, where $t_{\text{MP}}(\mathcal{E}) = [M_\bullet/M_\star(\mathcal{E})]P(\mathcal{E})$ is an estimate of the (angular-momentum-averaged) Newtonian mass precession time-scale and $t_{\text{GR}}(\mathcal{E}) = (1/24)(c^2/\mathcal{E})P(\mathcal{E})$ is the (angular-momentum-averaged) relativistic precession time-scale, assuming a thermal eccentricity distribution.

At low angular momenta, the efficiency of RR is strongly reduced because in-plane relativistic precession tends to reduce the efficiency of the torques arising from the $\sqrt{N}$ -asymmetry. This effect, known as the Schwarzschild barrier, can reduce the flux implied by equation (3.44) by at least an order of magnitude (Merritt et al. 2011). It remains unclear, however, how the details of equation (3.44) are affected by the SB. To take into account this uncertainty we include in equation (3.44) the dimensionless (and poorly constrained) ad hoc factor χ , and adopt two values of χ_{RR} : $\chi_{\text{RR}} = 0.1$ and $\chi_{\text{RR}} = 1$.

3.C Approximate semi-analytic solutions to the time-dependent Fokker-Planck equation

We solve equation (3.9) with $F_{\text{col}} = F_{\text{RR}} = 0$ and assume that a steady-state is present in energy at all times, i.e. $f_a(\mathcal{E}, t) = f_0 g(t)$, where f_0 is a constant and $g(t)$ a time-dependent function with $g(0) = 1$ (cf. Section 3.4.3). In this case $\partial \mathcal{F}_\mathcal{E} / \partial \mathcal{E} = 0$; integrating both sides of



equation (3.9) with respect to energy, the latter equation can be written as

$$C_1 \frac{\partial g}{\partial t} = N_{a/*} \frac{\partial h}{\partial t} - C_2 g(t), \quad (3.45)$$

with

$$C_1 \equiv f_0 4\pi^2 \int d\mathcal{E} p(\mathcal{E}); \quad C_2 \equiv C_1 \frac{\int d\mathcal{E} S_{lc}}{4\pi^2 \int d\mathcal{E} p(\mathcal{E})}; \quad h(t) \equiv \int d\mathcal{E} N_{LT}(\mathcal{E}) f_{\text{strip}}(\mathcal{E}, t),$$

and $S_{lc}(\mathcal{E})$ is defined via $F_{lc}(\mathcal{E}, t) = S_{lc}(\mathcal{E}) f_a(\mathcal{E}, t)$ (cf. equation 3.37). In the case of formation in a cloud C_1 , and thereby f_0 , can be estimated from $C_1 \sim N_{a/*} \int d\mathcal{E} N_{LT}(\mathcal{E})$; in the case of formation in discs, $C_1 \sim N_{a/*} \int d\mathcal{E} N_{LT}(\mathcal{E}) f_{\text{strip;SBH}}(\mathcal{E})$. Equation (3.45) can be solved for $g(t)$ by separation of variables; writing $g(t) = g_H(t) g_I(t)$, the homogenous solution is

$$g_H(t) = g_H(0) \exp\left(\frac{-t}{\tau}\right), \quad (3.46)$$

where $\tau \equiv C_1/C_2$. Substituting the latter solution into equation (3.45) and solving for $g_I(t)$, we find

$$g_I(t) = \frac{N_{a/*}}{C_1 g_H(0)} \int_0^t dt' \frac{\partial h}{\partial t'} \exp\left(\frac{t'}{\tau}\right) + C, \quad (3.47)$$

where C is an integration constant. Setting $g(0) = 1$, the complete solution is

$$g(t) = \exp\left(\frac{-t}{\tau}\right) \left[1 + \frac{N_{a/*}}{C_1} \int_0^t dt' \frac{\partial h}{\partial t'} \exp\left(\frac{t'}{\tau}\right) \right]. \quad (3.48)$$

The implied disruption flux is $F_{\text{dis}}(t) \sim f_0 g(t) \int d\mathcal{E} S_{lc}(\mathcal{E})$ (cf. equation 3.37) and is shown in Fig. 3.10, where in case formation in a cloud we set $h(t) = 0$.

3.D Scaling of the disruption rate

To obtain the scaling of the disruption rate with $N_{a/*}$, m_* , $M_\bullet$ and N_0 (cf. Section 3.5.3) we adopt the approximate semi-analytic solution to equation (3.9) as described in Appendix 3.C. We approximate the effect of stripping by gravitational encounters as an instantaneous

process at all energies, i.e.

$$h(t) \sim \begin{cases} 0, & t < t_c; \\ \int d\mathcal{E} N_*(\mathcal{E}), & t \geq t_c, \end{cases} \quad (3.49)$$

where $N_*(\mathcal{E})$ is the number of stars with energies between $\mathcal{E}$ and $\mathcal{E} + d\mathcal{E}$ and t_c is the stripping time-scale; in the GC, $t_c \sim 100$ Myr (cf. Section 3.3.2). Integrating by parts, we subsequently obtain

$$\int_0^t dt' \frac{\partial h}{\partial t'} \exp\left(\frac{t'}{\tau}\right) \sim \begin{cases} 0, & t < t_c; \\ \exp(t_c/\tau) \int d\mathcal{E} N_*(\mathcal{E}), & t \geq t_c. \end{cases} \quad (3.50)$$

Neglecting the initial stripping by the SBH, $C_1 \sim N_{a/*} \int d\mathcal{E} N_*(\mathcal{E})$ (cf. Appendix 3.C), and equation (3.48) yields, for $t \geq t_c$,

$$g(t) \sim \exp\left(\frac{-t}{\tau}\right) \left[1 + \exp\left(\frac{t_c}{\tau}\right) \right]. \quad (3.51)$$

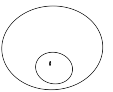
Note that $\tau \equiv C_1/C_2 \equiv 4\pi^2 \int d\mathcal{E} p(\mathcal{E}) / \int d\mathcal{E} S_{lc}(\mathcal{E})$ is independent of $N_{a/*}$, therefore $g(t)$ in equation (3.51) is also independent of $N_{a/*}$. From $f_a(\mathcal{E}, t) = f_0 g(t)$ and $f_0 \propto N_{a/*}$, it subsequently follows that the disruption flux (cf. equation 3.37) scales linearly with $N_{a/*}$, as might be intuitively expected.

We assume a power-law stellar number density distribution, $n_* = n_0(r/r_0)^{-\gamma}$, where $n_0 = N_0(3 - \gamma)/(4\pi r_0^3)$. The distribution function is then given by $f_*(\mathcal{E}) = f_0 \mathcal{E}^{\gamma-3/2}$, where f_0 is related to n_0 via an equation similar to equation (3.35). Neglecting, for simplicity, the integrals $\bar{I}_{1/2}(\mathcal{E})$ and $\bar{I}_{3/2}(\mathcal{E})$ in equation (3.39), and setting the integration limits in the energy integral to $\mathcal{E}_h < \mathcal{E} < \infty$, where $\mathcal{E}_h = [GM_*/r_0][(2M_*/m_*)(1/N_0)]^{1/(\gamma-3)}$ is the energy at the radius of influence neglecting the stellar potential, the energy integral of S_{lc} (cf. equation 3.37) is approximately given by

$$\int_{\mathcal{E}_h}^{\infty} d\mathcal{E} S_{lc}(\mathcal{E}) \sim 4\pi^2 \frac{32}{3} \pi^2 \frac{5\pi}{16} \frac{1}{\sqrt{2}} \frac{1}{3 - \gamma} (Gm_*)^2 (GM_*)^3 \ln(\Lambda) \mathcal{E}_h^{\gamma-3} f_0, \quad (3.52)$$

where we neglected the factor $\log(R_0^{-1})^{-1}$, the integrals $\bar{I}_{n/2}(\mathcal{E})$ and assumed $\gamma \leq 2$. With neglect of the stellar potential the function $p(\mathcal{E})$ is given by $p(\mathcal{E}) = (\sqrt{2}\pi/4)(GM_*)^3 \mathcal{E}^{-5/2}$ (e.g. Merritt 2013b, equation 5.182a). The time-scale τ is therefore approximately given by

$$\tau^{-1} \sim \frac{5}{4\sqrt{2}\pi} \frac{\Gamma(\gamma+1)}{\Gamma(\gamma-\frac{1}{2})} \log(\Lambda) (Gm_*)^2 (GM_* r_0)^{-3/2} \left(\frac{2M_*}{m_*} \frac{1}{N_0} \right)^{\frac{\gamma-3/2}{\gamma-3}} N_0. \quad (3.53)$$



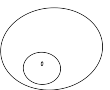
Setting $\int_{\mathcal{E}_0}^{\infty} d\mathcal{E} N_{\star}(\mathcal{E}) \sim N_0$, we find for the disruption flux

$$F_{\text{dis}}(t) \sim f_0 g(t) \int d\mathcal{E} S_{\text{lc}}(\mathcal{E}) \sim \frac{C_1 g(t)}{\tau} \sim \frac{N_{\text{a}/\star} N_0}{\tau} \exp(-t/\tau) [1 + \exp(t_c/\tau)]. \quad (3.54)$$

4 | Secular dynamics of hierarchical quadruple systems: the case of a triple system orbited by a fourth body

We study the secular gravitational dynamics of quadruple systems consisting of a hierarchical triple system orbited by a fourth body. These systems can be decomposed into three binary systems with increasing semimajor axes, binaries A, B and C. The Hamiltonian of the system is expanded in ratios of the three binary separations, and orbit averaged. Subsequently, we numerically solve the equations of motion. We study highly hierarchical systems that are well described by the lowest order terms in the Hamiltonian. We find that the qualitative behaviour is determined by the ratio $\mathcal{R}_0$ of the initial Kozai-Lidov (KL) time-scales of the binary pairs AB and BC. If $\mathcal{R}_0 \ll 1$, binaries AB remain coplanar if this is initially the case, and KL eccentricity oscillations in binary B are efficiently quenched. If $\mathcal{R}_0 \gg 1$, binaries AB become inclined, even if initially coplanar. However, there are no induced KL eccentricity oscillations in binary A. Lastly, if $\mathcal{R}_0 \sim 1$, complex KL eccentricity oscillations can occur in binary A that are coupled with the KL eccentricity oscillations in B. Even if binaries A and B are initially coplanar, the induced inclination can result in very high eccentricity oscillations in binary A. These extreme eccentricities could have significant implications for strong interactions such as tidal interactions, gravitational wave dissipation, and collisions and mergers of stars and compact objects. As an example, we apply our results to a planet+moon system orbiting a central star, which in turn is orbited by a distant and inclined stellar companion or planet, and to observed stellar quadruples.

*Hamers, A.S., Perets, H.B., Antonini, F. & Portegies Zwart, S.F. MNRAS **449**, 1, 4221-4245 (2015)*



4.1 Introduction

Hierarchical triple systems are known to be common among stellar systems. For example, a fraction of 0.076 of FG dwarfs systems in the catalogue of Tokovinin (2014a,b) are triple systems (in the fractions cited here from Tokovinin 2014b, completeness arguments have been taken into account; the observed number of triple systems in the sample of Tokovinin 2014a is 290, with a total number of 4847 systems). The triple fraction is likely higher for more massive stars. In such hierarchical systems, the torque of the outer binary can induce high-amplitude oscillations in the inner binary over time-scales that can vary from suborbital time-scales, to time-scales exceeding Gyr. These oscillations, known as Kozai-Lidov (KL) cycles (Lidov 1962; Kozai 1962), have important implications for a large range of astrophysical systems, in particular when the effects of tidal friction are also considered. The implications include the production of short-period binaries and hot Jupiters (Eggleton & Kiseleva-Eggleton 2001; Wu & Murray 2003; Eggleton & Kiseleva-Eggleton 2006; Fabrycky & Tremaine 2007; Wu et al. 2007; Correia et al. 2011; Naoz et al. 2011, 2012; Petrovich 2015a), accelerating the merging of compact objects (Blaes et al. 2002; Thompson 2011; Antonini & Perets 2012; Antonini et al. 2014), explaining some of the blue stragglers stars (Perets & Fabrycky 2009; Naoz & Fabrycky 2014), affecting the formation of binary minor planets (Perets & Naoz 2009), possibly producing a special type of type Ia supernovae through collisions of white dwarfs (Katz & Dong 2012; Hamers et al. 2013; Prodan et al. 2013), and modifying the evolution of stellar binaries that would not interact in the absence of a third star (Hamers et al. 2013).

Nature does not stop at $N = 3$, however. Although in the catalogue of Tokovinin (2014a,b) triple systems, with a fraction of 0.58 (observed: 290 of 350), are most common among systems with hierarchies ($N \geq 3$), quadruple systems also constitute a considerable fraction of hierarchical systems, i.e. a fraction of 0.32 (observed: 55 of 350). Unlike hierarchical triple systems, for which only one dynamically stable configuration is known to exist in nature, there are two different hierarchical configurations for which quadruples are known to be dynamically stable. One of these consists of two binary systems that orbit each other's barycentre, and this type of system constitutes a fraction of 0.74 (observed: 37 of 55) of the quadruple systems in the catalogue of Tokovinin (2014a,b). The long-term dynamical evolution of this configuration has been studied by Pejcha et al. (2013), who showed, by means of direct N -body simulations, that eccentricity oscillations, in particular orbital flips, can be enhanced in these systems relative to triples.

The other configuration consists of a hierarchical triple system that is orbited by a fourth body (referred to as a $3+1$ quadruple system in Tokovinin 2014b), and is the focus of this paper. In this case, three binary systems can be identified, and we will assume that they are each sufficiently separated from each other such that the quadruple system is dynamically stable. A stability analysis of these systems is beyond the scope of this work. Here, we shall always assume stability, although stability of some systems is borne out by our direct N -body integrations. We will refer to the binaries with the smallest, intermediate, and largest semimajor axes, as 'binary A', 'binary B' and 'binary C', respectively. A schematic depiction of our configuration is shown in Fig. 4.1.

Our hierarchical configuration not only applies to stellar quadruples, but also arises in other astrophysical systems. These include, but are not limited to, multiplanet, planet-moon and binary asteroid systems in single and binary star systems. Here, we study the case of a planet+moon system (binary A) that orbits a star (binary B), which in turn is orbited by a more

distant and inclined object (binary C), e.g. another planet or star. We assume that the orbit of the planet+moon system is initially coplanar with respect to that of the primary star. Therefore, in the absence of a distant body, no excitation of the eccentricity of the orbit of the planet+moon system is expected. However, we will show that, in the presence of an inclined fourth body, high-amplitude eccentricity oscillations can be induced in the planet+moon system through an intricate coupling of KL cycles.

The structure of this paper is as follows. In Section 4.2, we describe our methods. We expand the four-body Hamiltonian in terms of the separation ratios r_A/r_B , r_B/r_C and r_A/r_C . In order for our method to be suitable for the study of the long-term evolution of a large number of systems, we adopt the secular approximation, i.e. we average the Hamiltonian over the three binary orbits assuming unperturbed and bound orbits for time-scales shorter than the orbital periods. Subsequently, we numerically solve the equations of motion derived from the orbit-averaged Hamiltonian. We test our method by comparing to direct N -body integrations. In Section 4.3, we consider the general dynamics of highly hierarchical systems, i.e. systems that are well-described by the lowest-order terms in the Hamiltonian. We discuss our results in Section 4.4 and apply them to planetary and stellar systems. We give our conclusions in Section 4.5.

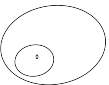
4.2 Methods

4.2.1 Expansion of the Hamiltonian

Our method to study the long-term evolution of quadruple systems is a natural extension to the orbit-averaged techniques that have been used extensively in the past to study the evolution of hierarchical triple systems, where an expansion was made in terms of the semimajor axis ratio $a_{\text{in}}/a_{\text{out}}$, with a_{in} and a_{out} the semimajor axes of the inner and outer orbit, respectively (Lidov 1962; Kozai 1962; Harrington 1968, 1969; Ford et al. 2000; Eggleton & Kiseleva-Eggleton 2001; Laskar & Boué 2010; Naoz et al. 2013a). We note that hierarchical systems with more complex configurations have also been studied using secular methods by Touma et al. (2009) and Boué & Fabrycky (2014). We expand the Hamiltonian in terms of the separation ratios r_A/r_B , r_B/r_C and r_A/r_C , where the separation vectors $\mathbf{r}_A$, $\mathbf{r}_B$ and $\mathbf{r}_C$ are defined in terms of the position vectors of the four bodies in equation (4.25). By assumption, $r_C \gg r_B \gg r_A$; therefore, these ratios are small and such an expansion is appropriate. The expansion is carried out to up and including fourth order in the separation ratios, i.e. including terms proportional to $(r_A/r_B)^i (r_B/r_C)^j (r_A/r_C)^k$, where $0 \leq i + j + k \leq 4$. The details are given in Appendix 4.A.1. For completeness, in addition to the configuration of a triple system orbited by a fourth body that is the focus of this paper, we have included results for the configuration of two binaries orbiting each other's barycentre in Appendix 4.A.2.

As derived in Appendix 4.A.1, at the lowest order, $i + j + k = 1$, the Hamiltonian consists of three terms that reduce to the binary binding energies of the three binaries A, B and C, assuming Kepler orbits. These terms therefore do not lead to secular orbital changes. At the next order, the ‘quadrupole’ order ($i + j + k = 2$)¹, we find three terms, each of which is mathematically equivalent to the quadrupole-order Hamiltonian in the three-body problem. These three terms can also be obtained from the three-body quadrupole order Hamiltonian by appropriate substitutions of the masses and separation vectors.

¹The term ‘quadrupole’ is not to be confused with ‘quadruple’.



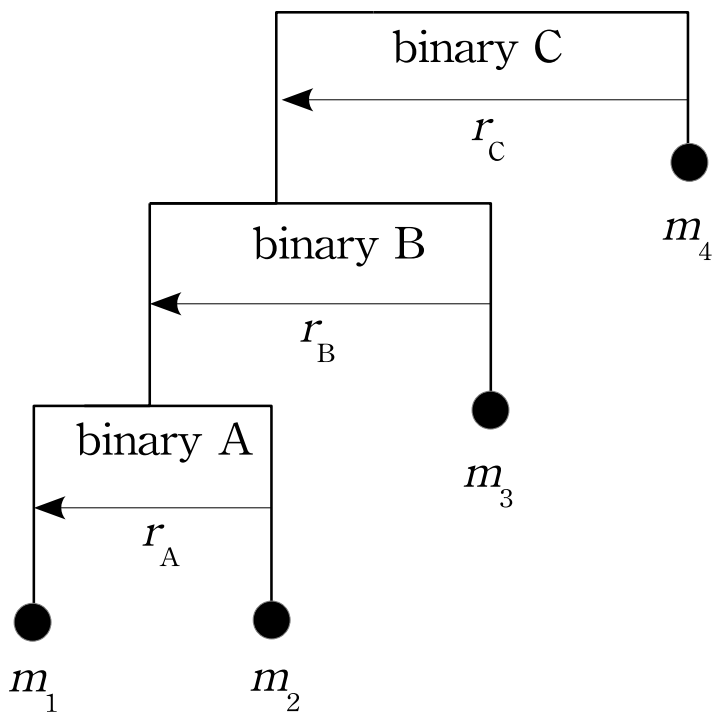


Figure 4.1: A schematic depiction of the hierarchical configuration of the quadruple systems considered in this paper.

More specifically, the (non-averaged) *three*-body Hamiltonian at the quadrupole order is given by

$$H_{\text{quad}} = -\frac{Gm_1m_2m_3}{m_1+m_2} \frac{1}{r_{\text{out}}} \left(\frac{r_{\text{in}}}{r_{\text{out}}} \right)^2 \frac{1}{2} \left[3 (\hat{\mathbf{r}}_{\text{in}} \cdot \hat{\mathbf{r}}_{\text{out}})^2 - 1 \right], \quad (4.1)$$

where $\mathbf{r}_{\text{in}}$ and $\mathbf{r}_{\text{out}}$ are the separation vectors of the inner and outer binary, respectively. In our *four*-body system, the Hamiltonian, to the corresponding level of approximation, is given by three terms. These are each obtained from equation (4.1) by the following substitutions of separation vectors,

1. $\mathbf{r}_{\text{in}} \rightarrow \mathbf{r}_A$ and $\mathbf{r}_{\text{out}} \rightarrow \mathbf{r}_B$ (AB);
2. $\mathbf{r}_{\text{in}} \rightarrow \mathbf{r}_B$ and $\mathbf{r}_{\text{out}} \rightarrow \mathbf{r}_C$ (BC);
3. $\mathbf{r}_{\text{in}} \rightarrow \mathbf{r}_A$ and $\mathbf{r}_{\text{out}} \rightarrow \mathbf{r}_C$ (AC),

and masses

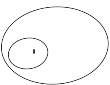
1. (no substitutions) (AB);
2. $m_1 \rightarrow m_1 + m_2$, $m_2 \rightarrow m_3$ and $m_3 \rightarrow m_4$ (BC);
3. $m_3 \rightarrow m_4$ (AC).

In the quadrupole-order approximation, there are no terms appearing in the Hamiltonian that depend on *all* three position vectors $\mathbf{r}_A$, $\mathbf{r}_B$ and $\mathbf{r}_C$. This is no longer the case for the next order, the ‘octupole’ order ($i + j + k = 3$). For the latter order, we find three terms that correspond to the octupole order terms in the three-body problem, and that can be obtained directly from the substitutions given above. In addition, we find a term that is a function of $\mathbf{r}_A$, $\mathbf{r}_B$ and $\mathbf{r}_C$. We will refer to such terms as ‘cross terms’. The cross term at octupole order is given by

$$\begin{aligned} H_{\text{oct, cross}} &= \frac{Gm_1m_2m_3m_4}{(m_1+m_2)(m_1+m_2+m_3)} \frac{1}{r_C} \left(\frac{r_A}{r_C} \right)^2 \left(\frac{r_B}{r_C} \right) \\ &\times \frac{1}{2} \left[15 (\hat{\mathbf{r}}_B \cdot \hat{\mathbf{r}}_C) (\hat{\mathbf{r}}_A \cdot \hat{\mathbf{r}}_C)^2 - 3 (\hat{\mathbf{r}}_B \cdot \hat{\mathbf{r}}_C) \right. \\ &\quad \left. - 6 (\hat{\mathbf{r}}_A \cdot \hat{\mathbf{r}}_C) (\hat{\mathbf{r}}_A \cdot \hat{\mathbf{r}}_B) \right]. \end{aligned} \quad (4.2)$$

In the systems of interest here, the three terms in the Hamiltonian that can be obtained by the substitutions discussed above from the corresponding terms in the three-body problem, are generally dominated by the terms that apply to the binary combinations AB and BC. This is because, by assumption, $r_A/r_B \gg r_A/r_C$ and $r_B/r_C \gg r_A/r_C$. For the same reason, the octupole-order cross term, which is proportional to $(r_A/r_C)^2(r_B/r_C)$, is also typically small. However, in the three-body problem, the octupole-order term vanishes for equal masses in the inner binary (cf. equation 4.30c). This implies that the octupole-order terms associated with the binary combinations AB and BC vanish if $m_1 = m_2$ and $m_1 + m_2 = m_3$, and suggests that the octupole-order cross term could be important in that case.

To investigate this further, we have also derived the terms of the next higher order, $i + j + k = 4$ (henceforth ‘hexadecupole’ order). Analogously to the lower orders, we find three terms



that depend only on quantities of two of the binaries and that satisfy the substitutions given above. Their general form is given by

$$H_{\text{hd}} = -\frac{Gmm'm''(m^2 - mm' + m'^2)}{(m + m')^3} \frac{1}{r_{\text{out}}} \left(\frac{r_{\text{in}}}{r_{\text{out}}}\right)^4 \times \frac{1}{8} \left[35 (\hat{\mathbf{r}}_{\text{in}} \cdot \hat{\mathbf{r}}_{\text{out}})^4 - 30 (\hat{\mathbf{r}}_{\text{in}} \cdot \hat{\mathbf{r}}_{\text{out}})^2 + 3 \right]. \quad (4.3)$$

These terms do not cancel if the masses in the inner binary are equal; in fact, they do not cancel for any non-trivial combination of masses m and m' . In addition to these terms, we find two terms that depend on quantities pertaining to all three binaries, i.e. two cross terms. Expressions for the latter terms are given in equations (4.30g) and (4.30h). Although in this work, we do not include the hexadecupole-order terms in numerical integrations, we use our results of the hexadecupole-order Hamiltonian to evaluate the relative importance of the octupole-order cross term in Section 4.2.4.

4.2.2 Orbit averaging

We carried out an orbital averaging of the Hamiltonian expanded to up and including the hexadecupole order. For the cross terms, this entails averaging over three orbits. We assumed unperturbed Kepler orbits.

A major advantage of the orbit-averaged approach compared to direct N -body integration, is the strongly reduced computational cost, in particular if the integration time is long compared to the orbital periods, and if a large number of systems is to be integrated. Furthermore, the orbit-averaged approach is a key instrument for the (semi)analytic understanding of the long-term behaviour (i.e. much longer than the orbital periods), as demonstrated e.g. below in Section 4.3.4.

The main disadvantage is that the dynamics on suborbital time-scales are averaged over, therefore potentially missing important effects (Antonini & Perets 2012; Antonini et al. 2014; Antognini et al. 2014). These effects can particularly be important in systems that are close to the limit of dynamical stability. However, for highly hierarchical systems, we do not expect these effects to be important, and these systems are the main focus of this work. In our numerical integrations, we check for the condition when the orbit-averaged approach likely breaks down (cf. Section 4.2.3).

In the orbit-averaging procedure, we express the angular momenta and orientations of each of the three binaries in terms of the triad of perpendicular orbital state vectors ($\mathbf{j}_k, \mathbf{e}_k, \mathbf{q}_k$), where $\mathbf{q}_k \equiv \mathbf{j}_k \times \mathbf{e}_k$ and $k \in \{A, B, C\}$. Here, $\mathbf{j}_k$ is a vector aligned with the angular momentum vector of the orbit and which has magnitude $j_k = \sqrt{1 - e_k^2}$, $\mathbf{e}_k$ is the eccentricity, or Laplace-Runge-Lenz vector, that is aligned with the major axis and which has magnitude e_k , the orbital eccentricity.

The orbit-averaged Hamiltonian is given in equation (4.33). For further details, we refer to Appendix 4.A.1.

4.2.3 Equations of motion and numerical algorithm

The equations of motion for the orbital vectors $\mathbf{j}_k$ and $\mathbf{e}_k$ of the three binary orbits are obtained by taking gradients of the orbit-averaged Hamiltonian $\overline{H}$ (Milankovitch 1939, see also e.g.

Musen 1961; Allan & Ward 1963; Allan & Cook 1964; Breiter & Ratajczak 2005; Tremaine et al. 2009; see Rosengren & Scheeres 2014 for a recent overview),

$$\frac{d\mathbf{j}_k}{dt} = -\frac{1}{\Lambda_k} [\mathbf{j}_k \times \nabla_{\mathbf{j}_k} \bar{H} + \mathbf{e}_k \times \nabla_{\mathbf{e}_k} \bar{H}] ; \quad (4.4a)$$

$$\frac{d\mathbf{e}_k}{dt} = -\frac{1}{\Lambda_k} [\mathbf{e}_k \times \nabla_{\mathbf{j}_k} \bar{H} + \mathbf{j}_k \times \nabla_{\mathbf{e}_k} \bar{H}] . \quad (4.4b)$$

Here, $\Lambda_k = mm' \sqrt{Ga_k/(m+m')}$, with $(m, m') = (m_1, m_2)$ for $k = A$, $(m, m') = (m_1 + m_2, m_3)$ for $k = B$ and $(m, m') = (m_1 + m_2 + m_3, m_4)$ for $k = C$.

To solve the equations of motion, we have developed a code written in C++, SECULARQUADRUPLE, that numerically solves the system of ordinary differential equations (ODEs) equations (4.4), to up and including octupole order. Because the ODEs are generally highly stiff, we used CVODE (Cohen et al. 1996), a library specifically designed to solve stiff ODEs. Our code is interfaced within the AMUSE framework (Pelupessy et al. 2013; Portegies Zwart et al. 2013). This allows for convenient comparison with direct N -body integration, i.e. without using the secular approximation, using any of the many N -body codes available in AMUSE. In addition, this facilitates the inclusion of effects modelled by other codes such as stellar and binary evolution. A test of the code for a hierarchical triple system is given in Appendix 4.B.

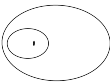
In the integrations with SECULARQUADRUPLE below, we included terms up and including octupole order, but without the octupole order cross terms. Here, we consider highly hierarchical systems, and it is shown in Section 4.2.4 that for these systems the octupole cross term does not dominate. Furthermore, neglect of this term is justified by the agreement with the N -body simulations, as shown in Section 4.2.5.

As mentioned above, situations can arise in which the orbit-averaged approximation breaks down. In particular, this can occur when the time-scale for changes of the angular momentum $\mathbf{j}_k$ is smaller than the orbital time-scale (Antonini & Perets 2012; Antonini et al. 2014). In SECULARQUADRUPLE, it is checked whether, at any time in the integration, any of the three binaries A, B or C satisfy this condition. This is implemented by means of a root-finding procedure: the integration is stopped whenever $t_{j,k} \leq P_{\text{orb},k}$, where $P_{\text{orb},k}$ is the orbital period of binary k and $t_{j,k}$ is the time-scale for the angular momentum of binary k to change by order itself, i.e.

$$t_{j,k} = \left| \frac{1}{j_k} \frac{dj_k}{dt} \right|^{-1} = \left| \frac{e_k}{1 - e_k^2} \frac{de_k}{dt} \right|^{-1} . \quad (4.5)$$

Although in SECULARQUADRUPLE the equations of motion are solved in terms of orbital vectors for numerical reasons, below we present our results in terms of the (generally easier to interpret) orbital elements $(e_k, i_k, \omega_k, \Omega_k)$, where i_k is the orbital inclination, ω_k is the argument of pericentre and Ω_k is the longitude of the ascending node. The latter quantities are defined with respect to a fixed reference frame (x, y, z) , and related to the orbital vectors $(\hat{\mathbf{e}}_k, \hat{\mathbf{j}}_k)$ according to

$$\begin{aligned} \hat{\mathbf{e}}_k &= [\cos(\Omega_k) \cos(\omega_k) - \sin(\Omega_k) \sin(\omega_k) \cos(i_k)] \hat{\mathbf{x}} \\ &\quad + [\sin(\Omega_k) \cos(\omega_k) + \cos(\Omega_k) \sin(\omega_k) \cos(i_k)] \hat{\mathbf{y}} \\ &\quad + \sin(\omega_k) \sin(i_k) \hat{\mathbf{z}}; \\ \hat{\mathbf{j}}_k &= \sin(\Omega_k) \sin(i_k) \hat{\mathbf{x}} - \cos(\Omega_k) \sin(i_k) \hat{\mathbf{y}} + \cos(i_k) \hat{\mathbf{z}}. \end{aligned} \quad (4.6)$$



description	m_1	m_2	m_3	m_4	e_A	e_B	e_C
reference	2	1	1	1	0.1	0.1	0.1
high e	2	1	1	1	0.99	0.99	0.99
no oct	1	1	2	1	0.1	0.1	0.1
high m_4	2	1	1	10^6	0.1	0.1	0.1

Table 4.1: Different combinations of the masses and eccentricities included in Fig. 4.2. Note that r depends on the masses only through their ratios, hence the mass unit is arbitrary.

In particular, i_k is defined as the angle between $\hat{\mathbf{j}}_k$ and the z -axis of the fixed reference frame. It is often useful to consider *mutual* inclinations i_{kl} between two orbits, rather than the *individual* inclinations i_k and i_l . They are related according to

$$\begin{aligned}\cos(i_{kl}) &= \hat{\mathbf{j}}_k \cdot \hat{\mathbf{j}}_l \\ &= \cos(i_k) \cos(i_l) + \sin(i_k) \sin(i_l) \cos(\Omega_k - \Omega_l).\end{aligned}\quad (4.7)$$

We note that in the hierarchical three-body problem, it is customary to define the orbital elements with respect to the *invariable plane*, i.e. a plane containing the total angular momentum vector (e.g. Naoz et al. 2013a). This implies $\Omega_k - \Omega_l = \pi$, and therefore the simple relation $i_{kl} = i_k + i_l$ can be applied. This is not the case here, where the z -axis of our frame of reference is not parallel to the total angular momentum. Therefore, one must resort to the more general equation (4.7).

Relativistic effects are also implemented in our algorithm. An important effect is relativistic precession of the argument of pericentre, associated with the Schwarzschild metric (Schwarzschild 1916). The associated time-scale for precession by 2π in binary k to the lowest post-Newtonian (PN) order is given by

$$t_{\text{1PN},k} = \frac{1}{3} P_{\text{orb},k} (1 - e_k^2) \frac{a_k}{r_{\text{g},k}}, \quad (4.8)$$

where $r_{\text{g},k} \equiv Gm_{\text{tot},k}$, with $m_{\text{tot},A} = m_1 + m_2$, $m_{\text{tot},B} = m_1 + m_2 + m_3$ and $m_{\text{tot},C} = m_1 + m_2 + m_3 + m_4$, is the gravitational radius. To take into account relativistic precession, the terms

$$\left. \frac{d\mathbf{e}_k}{dt} \right|_{\text{1PN}} = e_k \frac{2\pi}{t_{\text{1PN},k}} \hat{\mathbf{q}}_k \quad (4.9)$$

are added to the right-hand sides in equation (4.4b). Here, we neglect any possible additional ‘interaction terms’ between different binaries in the PN expansion that have been derived previously in the hierarchical three-body problem (Naoz et al. 2013b; Will 2014b,a), and that could also apply, in some form, to the configuration considered here.

4.2.4 The importance of the octupole-order cross terms

In Section 4.2.1 we derived a cross term in the Hamiltonian at octupole order. Here, we investigate further the importance of this term with respect to other terms at the octupole and the next

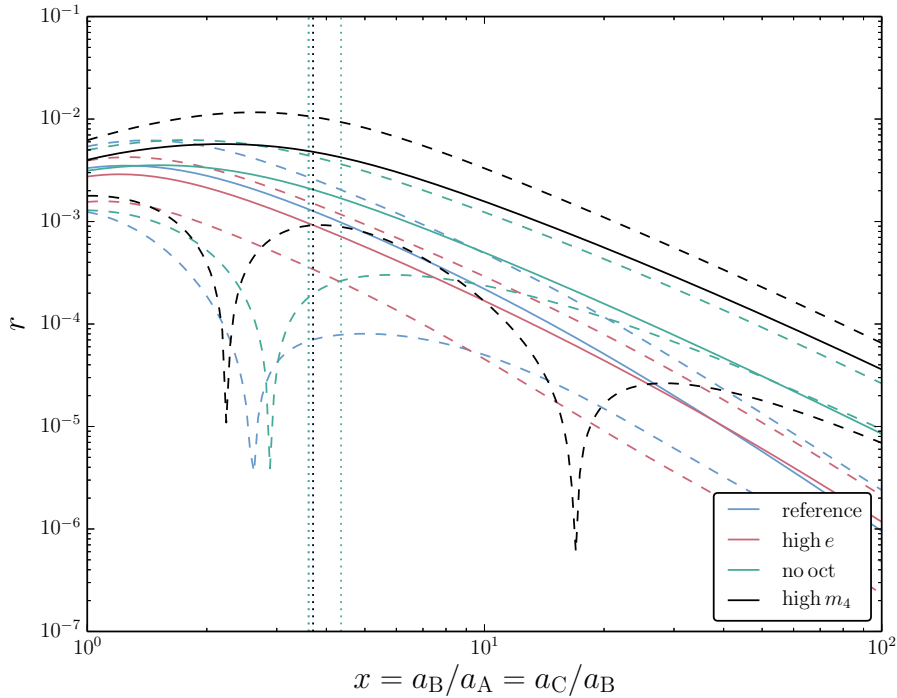
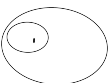


Figure 4.2: The ratio r of the absolute value of the orbit-averaged octupole-order cross term in the Hamiltonian to the absolute value of all other orbit-averaged terms at octupole and hexadecupole order (cf. equation 4.10), plotted as a function of $x = a_B/a_A = a_C/a_B$. An averaging over the orientations of the three binaries has been carried out (assuming random orbital orientations), and four different combinations of masses and eccentricities are assumed, which are given in Table 4.1. Solid lines: mean values of r for the different realizations of orbital orientations; (non-vertical) dashed lines: the same mean values, offset by the standard deviations (here, the absolute values are taken). Estimates for the minimum value of x for dynamical stability (based on the Mardling & Aarseth 2001 criterion applied to the AB and BC binaries) for each combination of parameters are indicated with vertical dashed lines. Note that for some of the ‘high m_4 ’ and ‘high e ’ combinations, these values are $> 10^2$, and are therefore beyond the range of the figure.



higher order, the hexacupole order. Long-term effects of the cross term can only be investigated by carrying out numerical integrations in time. However, a proxy for the short-term importance of the cross term is the ratio r of the absolute value of the orbit-averaged cross term, to the absolute value of all other orbit-averaged terms at octupole and hexadecupole order, i.e.

$$r \equiv \text{abs} \left[\overline{H}_{\text{oct, cross}} \left(\overline{H}_{\text{oct, AB}} + \overline{H}_{\text{oct, BC}} + \overline{H}_{\text{oct, AC}} + \overline{H}_{\text{oct, cross}} + \overline{H}_{\text{hd, AB}} + \overline{H}_{\text{hd, BC}} + \overline{H}_{\text{hd, AC}} + \overline{H}_{\text{hd, cross, 1}} + \overline{H}_{\text{hd, cross, 2}} \right)^{-1} \right]. \quad (4.10)$$

Here, $\overline{H}_{\text{oct, } kl}$ and $\overline{H}_{\text{hd, } kl}$ denote the orbit-averaged octupole-order and hexadecupole order terms corresponding to pair kl , respectively. They can be obtained directly from the general expressions equations (4.33c) and (4.33e), and using the substitutions discussed above in Section 4.2.1.

In principle, r can be maximized with respect to the parameters defining the properties and state of the quadruple system, i.e. with respect to the four m_i and the three a_k , e_k , $\hat{e}_k$ and $\hat{j}_k$ (with the orthogonality constraint $\hat{e}_k \cdot \hat{j}_k = 0$). This would yield the largest possible contribution of the cross term. However, the dimensionality (25) of this problem is very large, and this makes it computationally very difficult to find the *absolute* maximum. Here, we simplify the problem by restricting the parameter space.

In particular, we set $x \equiv a_B/a_A = a_C/a_B$, thereby reducing the dependence of the three semimajor axes to a single quantity. For given masses and eccentricities, we randomly sample the six unit vectors $\hat{e}_k$ and $\hat{j}_k$ with the orthogonality constraint $\hat{e}_k \cdot \hat{j}_k = 0$. We compute r for 20 of such realizations and each x , and subsequently, we compute the mean and standard deviations.

In Fig. 4.2, we show the resulting mean values (solid lines) and mean values offset by the standard deviations (dashed lines) of r as a function of x . We include four different combinations of masses and eccentricities, which are enumerated in Table 4.1. The minimum value of x for dynamical stability of the system is estimated by computing the critical semimajor axis ratio for stability of the AB and BC systems separately using the criterion of Mardling & Aarseth (2001). The latter two ratios are indicated for each combination of parameters in Fig. 4.2 with vertical dashed lines.

Regardless of our choice of parameters, r is typically small, in the sense that for values of x large enough for dynamical stability, $r \lesssim 10^{-2}$. For highly hierarchical systems, i.e. $x \gtrsim 100$, $r \lesssim 10^{-4}$. This indicates that typically the cross terms do not dominate the dynamics, at least for the short-term evolution. We note, however, that the octupole-order cross terms could give rise to important dynamical effects on long time-scales in less hierarchical systems.

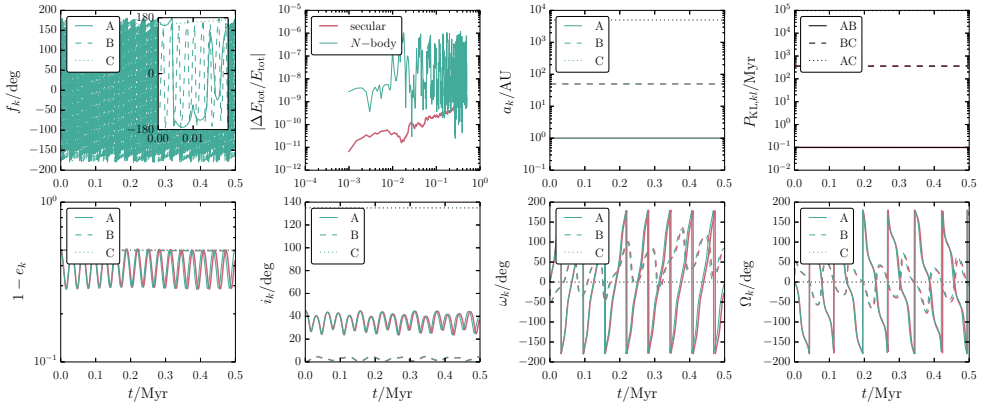


Figure 4.3: Comparison between the evolution of a quadruple system as computed with the orbit-averaged code `SECULARQUADRUPLE` developed in this work (red lines) and the direct N -body code `MIKKOLA` (Mikkola & Merritt 2008; green lines). The assumed initial parameters were semimajor axes $a_A = 1$ AU, $a_B = 5 \times 10^2$ AU and $a_C = 5 \times 10^3$ AU, masses $m_1 = m_3 = m_4 = 1 M_\odot$ and $m_2 = 0.5 M_\odot$, eccentricities $e_A = e_B = e_C = 0.5$, inclinations $i_A = 45^\circ$, $i_B = 0^\circ$, $i_C = 135^\circ$, arguments of pericentre $\omega_A = \omega_B = \omega_C = 0^\circ$ and longitudes of the ascending nodes $\Omega_A = \Omega_B = \Omega_C = 0^\circ$. When applicable to a single binary, solid, dashed and dotted curves correspond binaries A, B and C, respectively. When applicable to a binary pair, solid, dashed and dotted curves correspond to the binary pairs AB, BC and AC, respectively. The quantity $|\Delta E_{\text{tot}}/E_{\text{tot}}|$ is the absolute value of relative error in the total energy (the orbit-averaged Hamiltonian in the case of `SECULARQUADRUPLE`), and f_k is the true anomaly (applicable only to the N -body simulations). The inset in the top-left panel shows a magnification between $t = 0$ and 0.02 Myr. Note that the orbital period of binary A, $P_A \approx 0.8$ yr, is too short compared to the output resolution (≈ 500 yr) for f_A to be resolved. Also note that in the orbit-averaged code, the semimajor axes are constant by assumption, whereas the KL time-scales $P_{\text{KL},kl}$ in principle depend on time through the time-dependence of e_l (cf. equation 4.11). However, in this case, the dependence is extremely weak and not visible in the top-right panel.

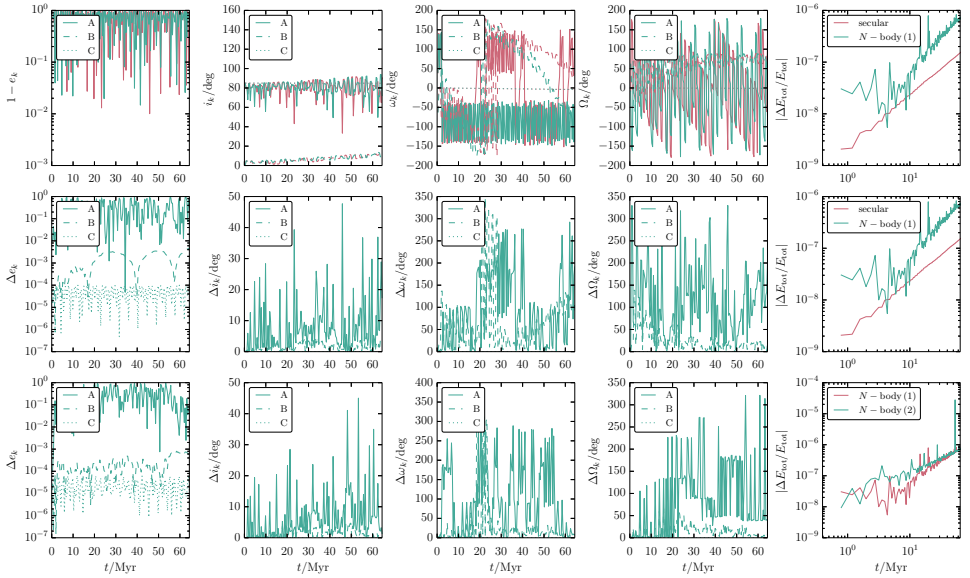
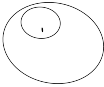


Figure 4.4: Top row: comparison between the evolution as computed with the orbit-averaged code SECULARQUADRUPLE (red lines) and the direct N -body code MIKKOLA (Mikkola & Merritt 2008; green lines). The system parameters are $a_A = 1\text{AU}$, $a_B = 10^2\text{AU}$ and $a_C = 5 \times 10^3\text{AU}$, $m_1 = m_3 = m_4 = 1\text{M}_\odot$ and $m_2 = 0.5\text{M}_\odot$, $e_A = e_B = e_C = 0.01$, $i_A = i_C = 85^\circ$ and $i_B = 0^\circ$, $\omega_A = \omega_B = \omega_C = 0^\circ$ and $\Omega_A = \Omega_B = \Omega_C = 0^\circ$ (the same as in panels 1-6 of Fig. 4.6). Middle row: the differences in e_k , i_k , ω_k and Ω_k between the secular code and one realization of the N -body code, as a function of time. Bottom row: the differences in e_k , i_k , ω_k and Ω_k between two realizations with the N -body code with different initial orbital phases f_k . Relative energy errors are shown in the last column.

K-S pair	e_k						ω_k						Ω_k					
	A		B		C		A		B		C		A		B		C	
	D	p	D	p	D	p	D	p	D	p	D	p	D	p	D	p	D	p
SN1	0.1	0.38	0.16	0.03	0.94	0.0	0.38	0.0	0.19	0.01	0.07	0.0	0.04	1.0	0.06	0.9	0.02	1.0
$\langle$ SN $\rangle$	0.08	0.65	0.18	0.02	0.73	0.0	0.34	0.11	0.17	0.02	0.09	0.58	0.05	0.95	0.07	0.85	0.02	1.0
NN1	0.09	0.46	0.11	0.31	0.1	0.38	0.45	0.0	0.04	1.0	0.08	0.38	0.04	1.0	0.06	0.96	0.01	1.0
$\langle$ NN $\rangle$	0.09	0.55	0.12	0.43	0.35	0.14	0.49	0.08	0.05	0.95	0.12	0.36	0.05	0.97	0.06	0.88	0.02	1.0

Table 4.2: Results of two-sided K-S tests (statistic D and the p -value) for time series in e_k , ω_k and Ω_k for the integrations shown in Fig. 4.4. In the first row, the secular code is compared to one realization of the N -body code, MIKKOLA (Mikkola & Merritt 2008). In the second row, the secular code is compared to five realizations of the N -body code (i.e. with different initial mean anomalies), and given are the resulting values of D and p averaged over individual K-S tests. In the third row, two realizations of the N -body code are compared, and in the fourth row, K-S tests are carried out for all combinations of the five realizations of the N -body codes, and the quoted values of D and p are averaged over these combinations.



4.2.5 Comparisons to direct N -body integrations

As a first demonstration of our algorithm, we show in Fig. 4.3 a comparison of a short-term integration with `SECULARQUADRUPLE` (red lines) and `MIKKOLA` (Mikkola & Merritt 2008), a highly accurate direct N -body code that uses chain regularization (green lines)². The assumed initial parameters were semimajor axes $a_A = 1$ AU, $a_B = 5 \times 10^2$ AU, $a_C = 5 \times 10^3$ AU, masses $m_1 = m_3 = m_4 = 1 M_\odot$ and $m_2 = 0.5 M_\odot$, eccentricities $e_A = e_B = e_C = 0.5$, inclinations $i_A = 45^\circ$, $i_B = 0^\circ$ and $i_C = 135^\circ$, arguments of pericentre $\omega_A = \omega_B = \omega_C = 0^\circ$ and longitudes of the ascending nodes $\Omega_A = \Omega_B = \Omega_C = 0^\circ$. Initially, i.e. during the first few KL oscillations in the AB pair, the two methods show very good agreement. However, as time progresses, noticeable deviations develop.

This poses a problem when comparing the two methods in longer integrations, i.e. for time-scales $\gg P_{\text{KL,AB}}$, where $P_{\text{KL,AB}}$ is the KL time-scale for the AB binary pair (cf. equation 4.11 below). To illustrate this, we show in the top row in Fig. 4.4 another example, where the integration time is $\sim 60 P_{\text{KL,AB}}$. In this case, we set $a_A = 1$ AU, $a_B = 10^2$ AU and $a_C = 5 \times 10^3$ AU, $m_1 = m_3 = m_4 = 1 M_\odot$ and $m_2 = 0.5 M_\odot$, $e_A = e_B = e_C = 0.01$, $i_A = i_C = 85^\circ$ and $i_B = 0^\circ$, $\omega_A = \omega_B = \omega_C = 0^\circ$ and $\Omega_A = \Omega_B = \Omega_C = 0^\circ$ (note that this system is the same as in panels 1-6 of Fig. 4.6). In the top row of Fig. 4.4, the quantities e_k , i_k , ω_k , Ω_k and the relative energy errors, pertaining to integrations with `SECULARQUADRUPLE` (`MIKKOLA`), are shown with red (green) lines. The differences in e_k , i_k , ω_k and Ω_k between the integrations with these codes are shown as a function of time in the middle row in Fig. 4.4. In this case, there is clearly no longer a one-to-one agreement between the two methods.

When comparing these results on long time-scales (i.e. long compared to $P_{\text{KL,AB}}$), it is important to take into account that for this system, the phase of the KL cycle in binary A becomes inherently chaotic on a time-scale that is shorter than $P_{\text{KL,AB}}$. To establish this, we determined the Lyapunov time-scale by carrying out pairs of integrations where in one realization, the initial value of $e_A = 0.01$, was increased by $\Delta e_A(0) = 10^{-4}$. We found that the difference $\Delta e_A(t)$ between the two integrations initially shows an exponential behaviour as a function of time. Subsequently, we fitted $\log[\Delta e_A(t)/\Delta e_A(0)]$ with time assuming a linear relation, i.e. $\log[\Delta e_A(t)/\Delta e_A(0)] = C + \lambda t$ where C is a constant, and we determined the Lyapunov time-scale t_{Ly} from the inverse of the resulting slope, i.e. $t_{\text{Ly}} = \lambda^{-1}$.

We find a Lyapunov time-scale of $t_{\text{Ly}} \approx 0.11$ Myr for this system, and this value is the same for either the secular and direct codes. Reducing the accuracy in the secular integrations does not affect the result, unless the accuracy is reduced such that the relative energy errors increase to > 0.1 . We also verified that $t_{\text{L}} \approx 0.11$ Myr for integrations with another N -body code, `SAKURA` (Gonçalves Ferrari et al. 2014).

This value of 0.11 Myr is shorter than $P_{\text{KL,AB}} \approx 1$ Myr, which suggests that the system is chaotic on a short time-scale. However, we find that this chaoticity arises from a slightly different phase of the KL cycle between the integrations with $\Delta e_A(0) = 10^{-4}$, whereas the shape of the $e_A(t)$ remains essentially the same. This result suggests that for long time-scales, it is not meaningful to compare the secular and N -body integrations on a one-to-one basis. However, given that the chaotic behaviour discussed above is associated with the phase of the KL cycle, it should still be appropriate to compare the two methods statistically.

We also note that when comparing the two methods, it is important to take into account that

²We remark that for this type of systems, it is essential to use a highly accurate N -body code because a large number of orbits, in particular in binary A, needs to be integrated very accurately.

in the N -body integrations, there is an additional dependence on the three initial orbital phases. We have also carried out N -body integrations with different initial orbital phases, where the initial mean anomaly was sampled randomly. We show the differences between two different N -body realizations as a function of time in the bottom row in Fig. 4.4. These differences are typically at least as large as the differences between the secular code, and a single realization with the N -body code. Furthermore, we have determined the Lyapunov time-scale as described above, where now $\Delta e_A(t)$ was determined from two short-term N -body integrations with different random mean anomalies. Again, we find that Δe_A increases exponentially with time, with a Lyapunov time-scale of $t_{Ly} \approx 0.1$ Myr. Therefore, the differences between the direct N -body integrations with different initial orbital phases can be ascribed to the chaotic nature of the phase of the KL cycles.

In Table 4.2, we show results of two-sided Kolmogorov-Smirnov (K-S) tests (Kolmogorov 1933; Smirnov 1948) between time series in e_k , ω_k and Ω_k obtained from the integration carried out with the secular code, and the integration of five different realizations with the N -body code (i.e. with different initial mean anomalies). For K-S tests between the secular and N -body integrations, and K-S tests between N -body integrations with different realizations, the D -values are generally low and the p -values are typically high. This shows that the integrations between the secular and N -body integrations are statistically consistent, and that the same applies to the N -body integrations with different realizations.

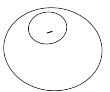
We conclude that, for the highly hierarchical systems considered here, the secular code gives results that are statistically consistent with the direct N -body code. The much greater speed makes the former highly suited for the long-term study of a large number of systems. For example, the integration with `SECULARQUADRUPLE` for one of the systems in Fig. 4.4 is $\sim 10^4$ times faster compared to `MIKKOLA`.

4.3 Global evolution of highly hierarchical systems

In principle, the `SECULARQUADRUPLE` algorithm can be used to perform a systematic parameter space study. Instead, here we choose to focus in detail on particular configurations to get insight into the typically complex dynamics that can arise. We consider the following two cases: (1) binaries A and B are initially coplanar ($i_{AB,0} = 0^\circ$) and highly inclined with respect to binary C ($i_{BC,0} = 85^\circ$), and (2) binaries A and B are initially highly inclined ($i_{AB,0} = 85^\circ$) while binary B is also highly inclined with respect to binary C ($i_{BC,0} = 85^\circ$). In both cases, we assume that the quadruple system is highly hierarchical at all times, i.e. $r_{p,A} \ll r_{p,B} \ll r_{p,C}$, where $r_{p,k}$ is the pericentre distance in binary k .

For both cases (1) and (2), we performed a sequence of integrations in which a_A was varied between 10^{-3} and 1 AU, and all other initial parameters were kept fixed. The latter were assumed to be semimajor axes $a_B = 10^2$ AU and $a_C = 5 \times 10^3$ AU, masses $m_1 = m_3 = m_4 = 1 M_\odot$ and $m_2 = 0.5 M_\odot$, eccentricities $e_A = e_B = e_C = 0.01$, arguments of pericentre $\omega_A = \omega_B = \omega_C = 0^\circ$ and longitudes of the ascending nodes $\Omega_A = \Omega_B = \Omega_C = 0^\circ$. The integration time for each system was set to $20 P_{KL,BC,0}$, where $P_{KL,BC,0}$ is the initial KL time-scale applied to binaries B and C, which we approximate by (Innanen et al. 1997)

$$P_{KL,kl} = \frac{P_l^2}{P_k} \frac{m_{k,p} + m_{k,s} + m_{l,s}}{m_{l,s}} (1 - e_l^2)^{3/2}, \quad (4.11)$$



where $m_{k,p} = m_1$, $m_{k,s} = m_2$ and $m_{l,s} = m_3$ in the case of $P_{KL,AB}$, and $m_{k,p} = m_1 + m_2$, $m_{k,s} = m_3$ and $m_{l,s} = m_4$ in the case of $P_{KL,BC}$ (cf. Section 4.2.1). Note that, contrary to triple systems and at the quadrupole-order approximation, the ‘outer’ orbit eccentricity e_l in equation (4.11) can change in time if this equation is applied to binaries A and B. This is addressed in more detail below.

For hierarchical triple systems, the octupole parameter

$$\epsilon_{\text{oct}} \equiv \frac{m_1 - m_2}{m_1 + m_2} \frac{a_{\text{in}}}{a_{\text{out}}} \frac{e_{\text{out}}}{1 - e_{\text{out}}^2} \quad (4.12)$$

is a useful proxy for the importance of octupole-order effects, in particular, orbital flips. The latter can occur if $\epsilon_{\text{oct}} \gtrsim 10^{-3}$, and are typically associated with very high eccentricities (Lithwick & Naoz 2011; Katz et al. 2011; Teyssandier et al. 2013; Li et al. 2014b). In the systems considered here, the initial octupole parameters ϵ_{oct} range between $\approx 3.3 \times 10^{-8}$ and $\approx 3.3 \times 10^{-5}$ for binary pair AB; for binary pair BC, $\epsilon_{\text{oct}} \approx 4.0 \times 10^{-5}$. This indicates that octupole-order terms are not important. Furthermore, the initial ratio r_0 of the orbit-averaged octupole-order cross term to all other orbit-averaged terms at octupole and hexadecupole order (cf. Section 4.2.4) ranges between $\approx 2 \times 10^{-12}$ and $\approx 3 \times 10^{-7}$, indicating that the orbit-averaged octupole-order cross term can similarly be neglected. The results presented below therefore demonstrate the dynamics that are manifested at the lowest possible, i.e. quadrupole, order.

4.3.1 Examples: A and B initially coplanar

In our first case, $i_{AB,0} = 0^\circ$ and $i_{BC,0} = 85^\circ$, which is achieved by setting the initial $i_A = i_B = 0^\circ$ and $i_C = 85^\circ$ (note that, initially, $\Omega_A = \Omega_B = \Omega_C$). In the absence of the fourth body, there would not be any excitation of the eccentricity in binaries A and B because they are not mutually inclined and only the quadrupole-order terms are important. We note that if the initial $e_B = 0.01$ were much larger (and therefore ϵ_{oct} would be much higher, cf. equation 4.12), owing to the greater importance of the octupole-order terms, orbital flips and very high eccentricity oscillations in binary A would be possible in certain conditions, even if $i_{AB,0}$ is close to zero (Li et al. 2014a). We show in Fig. 4.5 three examples of numerically integrated systems, in which a_A is either 1 (panels 1-6), 0.001 (panels 7-12) or 0.023 AU (panels 13-18).

For $a_A = 1$ AU, the *mutual* inclination between binaries A and B, i_{AB} , remains zero (cf. the solid line in panel 3 of Fig. 4.5). However, the *individual* inclinations of binaries A and B, i_A and i_B , which are initially zero, do change (cf. the solid and dashed lines in panel 4 of Fig. 4.5; note that these curves overlap). This can be understood from the large torque of binary B on binary A, compared to the torque of binary C on binary B. More quantitatively, the KL time-scales can be interpreted as proxies for the importance of these torques, and the initial KL time-scale for binaries A and B, $P_{KL,AB,0} \approx 1.2$ Myr, is much shorter (i.e. corresponding to a larger torque) than the initial KL time-scale for binaries B and C, $P_{KL,BC,0} \approx 2 \times 10^2$ Myr. The large torque of binary B on binary A enforces that zero *mutual* inclination between these binaries is maintained, despite the torque from binary C on binary B. The latter torque changes the *individual* inclination of binary B on the time-scale of $P_{KL,BC} \gg P_{KL,AB}$. Note that the mutual inclination is determined by the individual inclinations i_k and longitudes of the ascending nodes Ω_k (cf. equation 4.7). Therefore, both these angles for binaries A and B follow each other very closely (cf. panels 4 and 6 of Fig. 4.5).

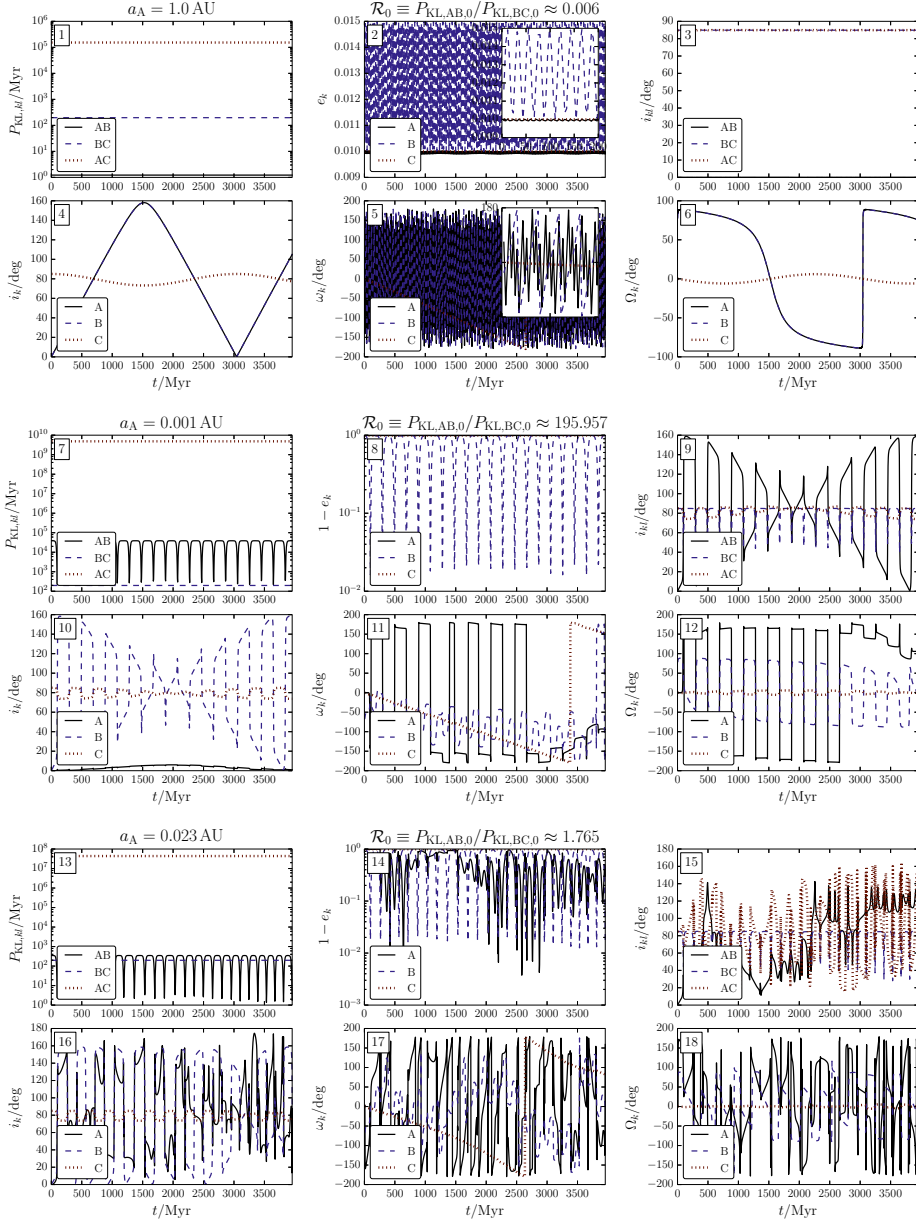
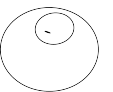


Figure 4.5: Evolution of three quadruple systems as discussed in Section 4.3.1, computed with SECULAR-QUADRUPLE. Binaries A and B are initially coplanar (as opposed to highly inclined in Fig. 4.6). Reference numbers are shown in the top left of each panel. Panels 1-6, 7-12 and 13-18 correspond to semimajor axes of binary A of 1, 0.001 and 0.023 AU, respectively. The other initial parameters are the same for these groups of panels, and are given by semimajor axes $a_B = 10^2$ AU and $a_C = 5 \times 10^3$ AU, masses $m_1 = m_3 = m_4 = 1 M_\odot$ and $m_2 = 0.5 M_\odot$, eccentricities $e_A = e_B = e_C = 0.01$, inclinations $i_A = i_B = 0^\circ$, $i_C = 85^\circ$, arguments of pericentre $\omega_A = \omega_B = \omega_C = 0^\circ$ and longitudes of the ascending nodes $\Omega_A = \Omega_B = \Omega_C = 0^\circ$. In panels 2 and 5, the abscissae in the inset range between $t = 0$ and 200 Myr.



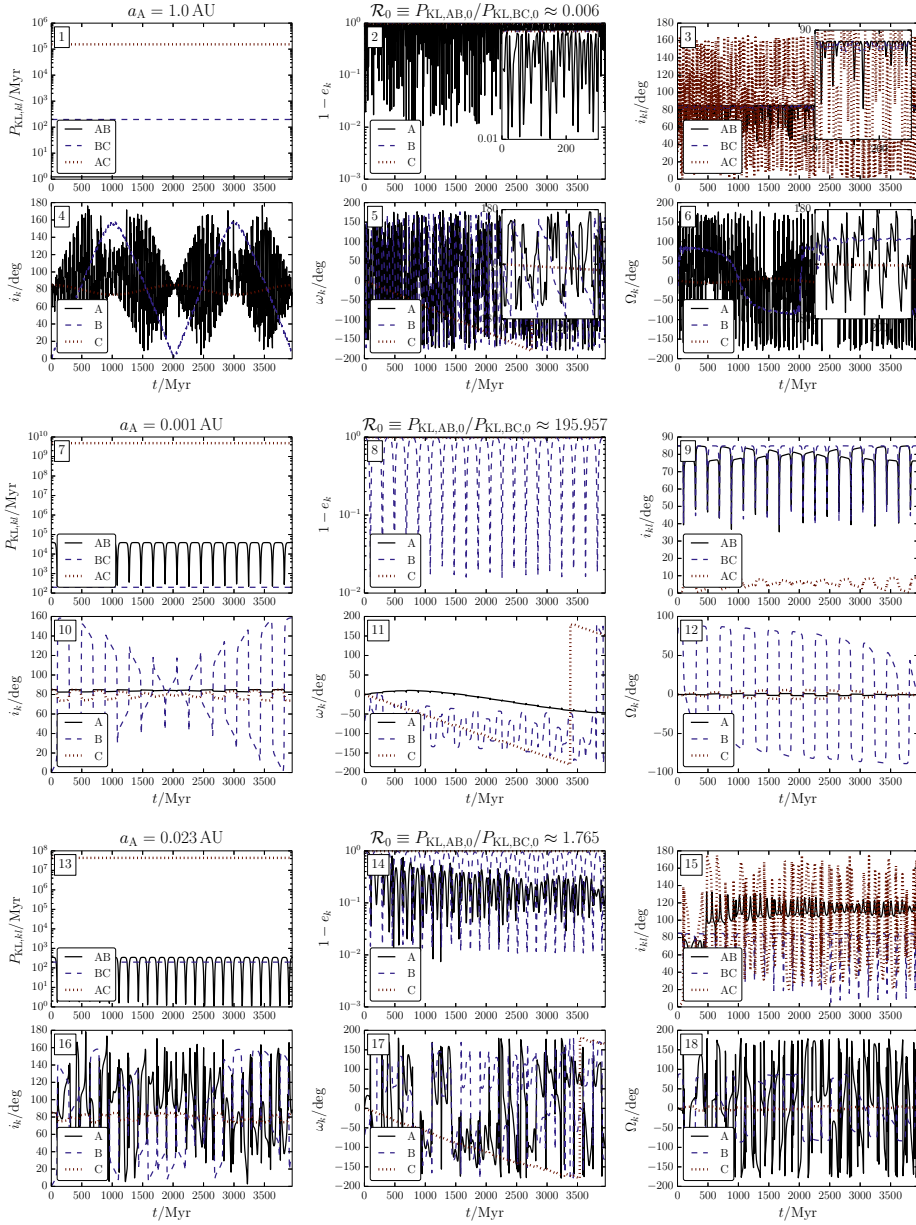


Figure 4.6: Evolution of three quadruple systems as discussed in Section 4.3.2, computed with SECULARQUADRUPLE. Binaries A and B are initially inclined by 85° (as opposed to 0° in Fig. 4.5). Panels 1-6, 7-12 and 13-18 correspond to semimajor axes of binary A of 1, 0.001 and 0.023 AU, respectively. The other parameters are the same for these groups of panels, and are given by semimajor axes $a_B = 10^2$ AU and $a_C = 5 \times 10^3$ AU, masses $m_1 = m_3 = m_4 = 1 M_\odot$ and $m_2 = 0.5 M_\odot$, eccentricities $e_A = e_B = e_C = 0.01$, inclinations $i_A = i_C = 85^\circ$, $i_B = 0^\circ$, arguments of pericentre $\omega_A = \omega_B = \omega_C = 0^\circ$ and longitudes of the ascending nodes $\Omega_A = \Omega_B = \Omega_C = 0^\circ$. In panels 2, 3, 5 and 6, the abscissae in the insets range between $t = 0$ and 300 Myr.

If binary A were replaced by a point mass, the eccentricity in binary B would oscillate as a result of the torque from binary C, with maxima of $1 - e_{B,\max} \approx 10^{-2}$. However, in the case of a quadruple system, the short KL time-scale in binary A with binary B causes rapid precession in both binaries A and B, on roughly the same time-scale (cf. the black solid and blue dashed lines in panel 5 of Fig. 4.5). Consequently, the rapid precession in binary B quenches any KL oscillations induced by the torque of binary C. This effect is analogous to the quenching of KL oscillations in triple systems due to additional sources of periaapse precession. Here, the additional precession is due to the extended nature of one of the components in the inner binary, rather than due to e.g. relativistic precession or tidal bulges. This quenching effect is discussed more quantitatively below, in Section 4.3.4.

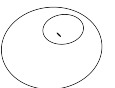
In panels 7-12 of Fig. 4.5, we show the evolution of an example system with $a_A = 10^{-3}$ AU. The initial KL time-scale for binaries A and B is $P_{\text{KL},AB,0} \approx 39 \text{ Gyr} \gg P_{\text{KL},BC,0} \approx 2 \times 10^2 \text{ Myr}$. Therefore, there is no induced precession of binary A on binary B, and KL eccentricity oscillations occur in binary B with maxima of $1 - e_{B,\max} \approx 10^{-2}$ (cf. the dashed lines in panel 8 of Fig. 4.5). Furthermore, the torque of binary C on binary B dominates compared to the torque of binary B on binary A. Consequently, the inclination of binary B changes rapidly, whereas the inclination of binary A hardly changes (cf. the solid and dashed lines in panel 10 of Fig. 4.5). However, this also changes the *mutual* inclination i_{AB} between binaries A and B. The latter increases very rapidly (cf. the solid line in panel 9 of Fig. 4.5). Nevertheless, binaries A and B are only highly mutually inclined (i_{AB} close to 90°) for short periods of time, and therefore, no significant eccentricity oscillations occur in binary A. In other words, the latter oscillations are impeded by rapid changes of the mutual inclination between binaries A and B, because of KL oscillations induced by binary C.

Finally, in panels 13-18 of Fig. 4.5, $a_A \approx 0.023 \text{ AU}$. The initial KL time-scales for the binary pairs AB and BC are comparable, i.e. $P_{\text{KL},AB,0} \approx 3 \times 10^2 \text{ Myr} \sim P_{\text{KL},BC,0} \approx 2 \times 10^2 \text{ Myr}$, and therefore the torques of binary B on binary A and of binary C on binary B are also comparable. Binaries A and B become mutually inclined, and the KL time-scale for the AB pair is short enough for large excitation of the eccentricity of binary A. The result is a complex evolution in which the oscillations in e_A are highly non-regular and strongly coupled with the oscillations of e_B . Interestingly, although binaries A and B started out with a mutual inclination of $i_{AB,0} = 85^\circ < 90^\circ$, the orientation between binaries A and B at $t \sim 400 \text{ Myr}$ changes from prograde to retrograde. Such orbital flips also occur at later times, and are associated with high eccentricities in binary A. The evolution of the eccentricity of binary B is also affected, although the effect is much smaller and the oscillations can still be considered as regular. In Section 4.3.4, we study the effect of the eccentricity of binary B in more detail.

4.3.2 Examples: A and B initially highly inclined

In our second case, we assume that both binaries A and B and binaries B and C are initially highly inclined, i.e. $i_{AB,0} = 85^\circ$ and $i_{BC,0} = 85^\circ$, which is achieved by setting $i_A = i_C = 85^\circ$ and $i_B = 0^\circ$. The evolution of three example systems, with other parameters identical to those in Section 4.3.1, is shown in Fig. 4.6. In the absence of the fourth body, high-eccentricity KL oscillations would be induced in binary A.

For $a_A = 1 \text{ AU}$ (panels 1-6 of Fig. 4.6), $P_{\text{KL},AB} \ll P_{\text{KL},BC}$, and for time-scales comparable to $P_{\text{KL},AB}$, KL eccentricity oscillations in binaries A and B are hardly affected by the torque of binary C. On much longer time-scales comparable to $P_{\text{KL},BC}$, i_B changes because of the torque



of binary C (cf. the blue dashed line in panel 4 of Fig. 4.6). However, the KL eccentricity oscillations between binaries A and B are not noticeably affected (note that in panels 2-6 of Fig. 4.6, the KL oscillations associated with binaries A and B are undersampled). Consequently, i_A , i_B , Ω_A and Ω_B are modulated on the $P_{\text{KL},\text{BC}}$ time-scale. We note that, as a consequence of KL oscillations in the AB pair, there is still short-time-scale precession induced on binary B, preventing any eccentricity excitation in binary B. This is similar to the previous case when binaries A and B are initially coplanar.

For $a_A = 0.001$ AU (panels 7-12 of Fig. 4.6), the evolution is qualitatively very similar to the case when $i_{\text{AB},0} = 0^\circ$. This may be surprising, given the high initial mutual inclination between binaries A and B. However, the latter changes strongly on the much shorter time-scale of $P_{\text{KL},\text{BC}}$, and this prevents any eccentricity excitation in binary A. Note that in this case, the quenching of KL eccentricity oscillations in binary A is not due to induced precession. As can be seen in panel 11 of Fig. 4.6, ω_A is not much affected on the $P_{\text{KL},\text{BC}}$ time-scale, although there is also a trend on a much longer time-scale of $\sim 4 \times 10^3$ Myr. The KL time-scale for the AB pair changes periodically as e_B oscillates (cf. panel 7 of Fig. 4.6). Therefore the time-scale of $\sim 4 \times 10^3$ Myr can, in this case, be interpreted as an effective KL time-scale for the AB pair.

When the KL time-scales for the AB and BC pairs are similar (cf. panels 13-18 of Fig. 4.6), the evolution of e_A is complex and high eccentricities are attained, similarly to the case when $i_{\text{AB},0} = 0^\circ$. Again, an orbital flip occurs around $t \sim 400$ Myr. Interestingly, subsequently there are no orbital flips, and the amplitude of the oscillations in e_A and i_{AB} gradually decreases.

4.3.3 Qualitative trends

The above examples suggest that the ratio of the (initial) KL time-scales for the AB and BC pairs,

$$\begin{aligned} \mathcal{R}_0 &\equiv \frac{P_{\text{KL},\text{AB},0}}{P_{\text{KL},\text{BC},0}} \\ &= \left(\frac{a_B^3}{a_A a_C^2} \right)^{3/2} \left(\frac{m_1 + m_2}{m_1 + m_2 + m_3} \right)^{1/2} \frac{m_4}{m_3} \left(\frac{1 - e_{\text{B},0}^2}{1 - e_{\text{C},0}^2} \right)^{3/2}, \end{aligned} \quad (4.13)$$

is an indication of the global trend of the inclination and eccentricity oscillations. We identify the following three regimes.

1. $\mathcal{R}_0 \ll 1$: binaries A and B remain coplanar if this was initially the case. If they are initially inclined, KL eccentricity oscillations in binary A are not much affected by the presence of the fourth body. In either case, KL eccentricity oscillations in binary B are quenched.
2. $\mathcal{R}_0 \gg 1$: binaries A and B become inclined if they are initially coplanar. However, there are no eccentricity oscillations in binary A, even if binaries A and B are initially highly inclined. This is because the mutual inclination between binaries A and B is large only for a small fraction of the KL time-scale for the AB pair, i.e. for a time of $< P_{\text{KL},\text{BC},0} = P_{\text{KL},\text{AB},0}/\mathcal{R}_0 \ll P_{\text{KL},\text{AB},0}$. Furthermore, KL eccentricity oscillations are not quenched in binary B.

3. $\mathcal{R}_0 \sim 1$: binaries A and B become inclined if they are initially coplanar; complex KL eccentricity oscillations arise in binary A that are coupled with the – much less affected – KL eccentricity oscillations in binary B.

These three regimes correspond to panels 1-6, 7-12 and 13-18 in Figs 4.5 and 4.6.

A complication in the above, is that $P_{\text{KL},\text{AB}}$ can change periodically with time because of KL eccentricity oscillations in binary B (cf. equation 4.11). Periodically higher values of e_{B} reduce $P_{\text{KL},\text{AB}}$ at the same times, therefore potentially increasing the range of $\mathcal{R}$ for which the eccentricity in binary A can be excited. Furthermore, for large enough values of e_{B} , higher order terms in the Hamiltonian become more important, and in extreme cases, the orbit-averaged approach could break down.

In principle, the time-dependence of e_{B} could be taken into account by e.g. averaging $P_{\text{KL},\text{AB}}$ over a KL cycle in binary B. However, except for a few simple cases, there are no analytic solutions for $e_{\text{B}}(t)$. Therefore, this would require numerical integration and hence not be of much practical use for predicting the behaviour *without* resorting to such integration. Nevertheless, because of the very peaked nature of $e_{\text{B}}(t)$ and the small width (in time) of the peaks, we expect the averaged value of $P_{\text{KL},\text{AB}}$ typically not to be very different from the value computed from $e_{\text{B},0}$, at least in systems in which the lowest order (quadrupole-order) terms dominate.

4.3.4 Quantitative dependence on $\mathcal{R}_0$

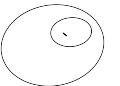
Results from numerical integrations

Here, we describe the dynamics outlined in Section 4.3.3 more quantitatively, focusing in particular on the effect of the quenching of KL eccentricity oscillations in binary B by the induced precession of binary A, and on the excitation of the eccentricity in binary A in the regime $\mathcal{R}_0 \sim 1$.

In Fig. 4.7, we show with black dots the maximum eccentricities in binaries A and B, the maximum inclination between binaries A and B, and the minimum inclination between binaries B and C as a function of $\mathcal{R}_0$, as determined from numerical integrations with `SECULARQUADRUPLE`. Here, $\mathcal{R}_0$ is varied by changing a_{A} (cf. equation 4.13) in the sequence of integrations described in the beginning of Section 4.3. In the left (right) panels, results are shown assuming that binaries A and B are initially coplanar (highly inclined).

For $i_{\text{AB},0} = 0^\circ$, $i_{\text{AB},\text{max}}$ is zero for $\mathcal{R}_0 \lesssim 1$ and rapidly increases for $\mathcal{R}_0 \gtrsim 1$; $e_{\text{A},\text{max}}$ is equal to the initial value for $\mathcal{R}_0 \lesssim 1$ and for $\mathcal{R}_0 \gtrsim 20$. This is consistent with the trend that was outlined in Section 4.3.3. Furthermore, if $\mathcal{R}_0 \lesssim 10^{-2}$, $e_{\text{B},\text{max}} \approx 0$, demonstrating that the induced precession of system A on B in this regime can completely quench any KL oscillations in binary B. Consequently, the minimum inclination between binaries B and C is constant and $\approx 85^\circ$, the initial value (note that for the regular KL oscillations in binary B, a maximum eccentricity corresponds to minimum inclination with respect to binary C). If $1 \lesssim \mathcal{R}_0 \lesssim 20$, $i_{\text{AB},\text{max}}$ is non-zero; $e_{\text{A},\text{max}}$ is also non-zero and reaches high values of up to $\approx 1 - 10^{-4}$. Although the behaviour of these two quantities as a function of $\mathcal{R}_0$ is non-regular, there is a general trend in which $i_{\text{AB},\text{max}}$ asymptotes to $\approx 160^\circ$. A general trend is also apparent in $e_{\text{A},\text{max}}$.

If binaries A and B are initially inclined by 85° (cf. the right-hand panels in Fig. 4.7), the dependence of $e_{\text{A},\text{max}}$ as a function of $\mathcal{R}_0$ is more complicated. For a large range in $\mathcal{R}_0$, $3 \times 10^{-2} \lesssim \mathcal{R}_0 \lesssim 50$, $e_{\text{A},\text{max}}$ fluctuates strongly with $\mathcal{R}_0$, reaching high values of $1 - e_{\text{A},\text{max}} \sim 10^{-4}$ for $\mathcal{R}_0$ already as low as $\mathcal{R}_0 \approx 3 \times 10^{-2}$. For $\mathcal{R}_0 \gtrsim 50$, $e_{\text{A},\text{max}}$ approaches $e_{\text{A},0}$, as was



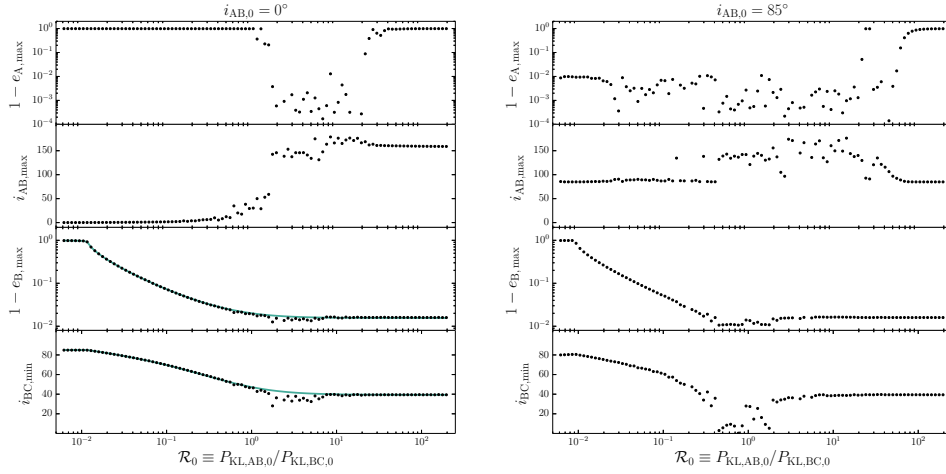


Figure 4.7: The maximum eccentricities in binaries A and B (first and third panels from the top), the maximum inclination between binaries A and B (second panels from the top), and the minimum inclination between binaries B and C (fourth panels from the top), as a function of the ratio $\mathcal{R}_0$ of the KL time-scales for the AB and BC pairs (cf. equation 4.13). Here, $\mathcal{R}_0$ is varied by changing the initial a_A and keeping the other initial semimajor axes, masses and eccentricities fixed. In both left- and right-hand panels, the initial conditions are the same as in Section 4.3.1 and Section 4.3.2, i.e. semimajor axes $a_B = 10^2$ AU and $a_C = 5 \times 10^3$ AU, masses $m_1 = m_3 = m_4 = 1 M_\odot$ and $m_2 = 0.5 M_\odot$, eccentricities $e_A = e_B = e_C = 0.01$, arguments of pericentre $\omega_A = \omega_B = \omega_C = 0^\circ$ and longitudes of the ascending nodes $\Omega_A = \Omega_B = \Omega_C = 0^\circ$. In the left- (right-hand) panels, binaries A and B are assumed to be initially coplanar (inclined by 85°), i.e. in the left-hand panel, $i_A = i_B = 0^\circ$ and $i_C = 85^\circ$, whereas in the right-hand panel, $i_A = i_C = 85^\circ$ and $i_B = 0^\circ$. Black dots: computed with `SECULARQUADRUPLE` (the integration time was $20 P_{\text{KL,BC}}$); solid lines: computed using the semianalytic method discussed in Section 4.3.4.

observed previously in Section 4.3.2. Furthermore, binary B is more affected compared to the coplanar case, in the sense that $i_{\text{BC},\text{min}}$ decreases more strongly in the regime $1 \lesssim \mathcal{R}_0 \lesssim 20$. The maximum eccentricity in binary B is similar to the coplanar case, however.

Semianalytic description

The maximum eccentricity (and hence minimum inclination) reached in binary B can be computed approximately using a semianalytic method based on conservation of the total energy (i.e. the Hamiltonian) and the total angular momentum. This method is similar to that used by Miller & Hamilton (2002); Blaes et al. (2002); Fabrycky & Tremaine (2007); Naoz et al. (2013a). We neglect any changes in binary A between the initial and final states, where the final state corresponds to a maximum eccentricity in binary B. To our knowledge, it is not possible to predict (i.e. without resorting to ‘brute-force’ numerical integrations as in Section 4.3.4) these changes in system A, and this is likely related to the generally chaotic nature of the evolution of binary A, in particular in the regime $1 \lesssim \mathcal{R}_0 \lesssim 20$ (cf. Section 4.3.5). Stated more mathematically, conservation of total energy and angular momentum and the condition that e_B is stationary, do not generally provide enough constraints to solve both for $e_{B,\text{max}}$ and the corresponding e_A .

In the Hamiltonian to quadrupole order and for the hierarchy considered here, the term corresponding to binaries A and C in the Hamiltonian can safely be neglected. This can readily be seen from equation (4.33b): the three terms at quadrupole order scale with the semimajor axes according to

$$\begin{aligned}\bar{H}_{\text{quad},\text{AB}} &\propto \frac{1}{a_B} \left(\frac{a_A}{a_B} \right)^2; & \bar{H}_{\text{quad},\text{BC}} &\propto \frac{1}{a_C} \left(\frac{a_B}{a_C} \right)^2; \\ \bar{H}_{\text{quad},\text{AC}} &\propto \frac{1}{a_C} \left(\frac{a_A}{a_C} \right)^2.\end{aligned}\quad (4.14)$$

Because, by assumption, $a_C \gg a_B \gg a_A$, this implies that $\bar{H}_{\text{quad},\text{AC}}$ can be neglected compared to $\bar{H}_{\text{quad},\text{AB}}$ and $\bar{H}_{\text{quad},\text{BC}}$. The Hamiltonian to quadrupole order is therefore well approximated by (cf. equation 4.33b)

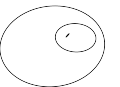
$$\begin{aligned}\bar{H}_0 &= C_{\text{AB}} j_B^{-5} \left[(1 - 6e_A^2) j_B^2 + 15 (\mathbf{e}_A \cdot \mathbf{j}_B)^2 - 3 (\mathbf{j}_A \cdot \mathbf{j}_B)^2 \right] \\ &+ C_{\text{BC}} j_C^{-5} \left[(1 - 6e_B^2) j_C^2 + 15 (\mathbf{e}_B \cdot \mathbf{j}_C)^2 - 3 (\mathbf{j}_B \cdot \mathbf{j}_C)^2 \right],\end{aligned}\quad (4.15)$$

where

$$\begin{aligned}C_{\text{AB}} &= \frac{1}{8} \frac{G m_1 m_2 m_3}{m_1 + m_2} \frac{1}{a_B} \left(\frac{a_A}{a_B} \right)^2; \\ C_{\text{BC}} &= \frac{1}{8} \frac{G (m_1 + m_2) m_3 m_4}{m_1 + m_2 + m_3} \frac{1}{a_C} \left(\frac{a_B}{a_C} \right)^2.\end{aligned}\quad (4.16)$$

The equation of motion for e_B that follows from equation (4.15) is given by (cf. equation 4.4)

$$\begin{aligned}\frac{de_B}{dt} &= \frac{6}{\Lambda_B} \\ &\times \left[C_{\text{AB}} j_B^{-5} \{ (\mathbf{j}_A \cdot \mathbf{j}_B) (\mathbf{e}_B \times \mathbf{j}_A) + 5 (\mathbf{e}_A \cdot \mathbf{j}_B) (\mathbf{e}_A \times \mathbf{e}_B) \} \right. \\ &\quad \left. + C_{\text{BC}} j_C^{-5} \{ (\mathbf{j}_B \cdot \mathbf{j}_C) (\mathbf{e}_B \times \mathbf{j}_C) - 5 (\mathbf{e}_B \cdot \mathbf{j}_C) (\mathbf{j}_B \times \mathbf{j}_C) \} \right].\end{aligned}$$



A stationary value of e_B corresponds to

$$0 = \frac{de_B}{dt} = \hat{e}_B \cdot \frac{d\mathbf{e}_B}{dt}. \quad (4.17)$$

Neglecting the terms proportional to C_{AB} in equation (4.17), this condition implies $e_B \cdot \dot{\mathbf{j}}_C = 0$, and/or $(\dot{\mathbf{j}}_B \times \dot{\mathbf{j}}_C) \cdot \mathbf{e}_B = 0$. The former cannot be generally true in the case of a maximum eccentricity, therefore the second condition must apply. The latter can be rewritten using the vector identity equation (4.34) as

$$\left(\hat{\mathbf{e}}_B \cdot \hat{\mathbf{j}}_C\right)^2 = 1 - \left(\hat{\mathbf{j}}_B \cdot \hat{\mathbf{j}}_C\right)^2. \quad (4.18)$$

The mutual inclination between binaries B and C can be related to e_B using conservation of the total angular momentum vector,

$$\mathbf{L}_{\text{tot}} = \Lambda_A \mathbf{j}_A + \Lambda_B \mathbf{j}_B + \Lambda_C \mathbf{j}_C. \quad (4.19)$$

At this level of approximation,

$$\nabla_{\mathbf{e}_C} \bar{H}_0 = \mathbf{0}, \quad (4.20)$$

therefore

$$\frac{de_C}{dt} = \hat{\mathbf{e}}_C \cdot \frac{d\mathbf{e}_C}{dt} = -\frac{1}{\Lambda_C} [\hat{\mathbf{e}}_C \cdot (\mathbf{e}_C \times \nabla_{\mathbf{j}_C} \bar{H}_0)] = 0, \quad (4.21)$$

and e_C is constant. Neglecting the term corresponding to binary A and writing $e_C = e_{C,0}$, equation (4.19) gives

$$\begin{aligned} \hat{\mathbf{j}}_B \cdot \hat{\mathbf{j}}_C &= \frac{1}{2\sqrt{1-e_B^2}\sqrt{1-e_{C,0}^2}} \left[2\sqrt{1-e_{B,0}^2}\sqrt{1-e_{C,0}^2} \right. \\ &\quad \left. \times \left(\hat{\mathbf{j}}_B \cdot \hat{\mathbf{j}}_C\right)_0 + \frac{\Lambda_B}{\Lambda_C} (e_B^2 - e_{B,0}^2) \right]. \end{aligned} \quad (4.22)$$

Furthermore, if any changes in binary A between the initial and final state are neglected, then the remaining unknown terms in equation (4.15) are simply given by $e_A = e_{A,0}$, $\dot{\mathbf{j}}_A \cdot \dot{\mathbf{j}}_B = (\dot{\mathbf{j}}_A \cdot \dot{\mathbf{j}}_B)_0$ and $\mathbf{e}_A \cdot \dot{\mathbf{j}}_B = (e_A \cdot \dot{\mathbf{j}}_B)_0$.

With these simplifications, equation (4.15) only contains the single unknown quantity e_B corresponding to stationary points. In general, this equation cannot be solved analytically. A notable exception is when the term proportional to C_{AB} in equation (4.15) is neglected (i.e. neglecting the contribution from binary A), as is the term proportional to Λ_B/Λ_C in equation (4.22) (i.e. assuming a highly hierarchical system). In that case, the solution corresponding to the maximum eccentricity is

$$e_{B,\text{max}} = \sqrt{1 - \frac{5}{3} \left(\hat{\mathbf{j}}_B \cdot \hat{\mathbf{j}}_C\right)_0^2}, \quad (4.23)$$

which is a well-known result for hierarchical triple systems applied to binaries B and C, and where binary A is essentially replaced by a point mass (note that $\hat{\mathbf{j}}_B \cdot \hat{\mathbf{j}}_C = \cos[i_{BC}]$). More general numerical solutions are shown in the bottom two panels of Fig. 4.7 with the solid lines, where $i_{BC,\text{min}}$ is computed using equation (4.22). Although the semianalytic curves do not capture the detailed behaviour of $e_{B,\text{max}}$ and $i_{BC,\text{min}}$ in the regime $1 \lesssim \mathcal{R}_0 \lesssim 20$, for other $\mathcal{R}_0$ they agree well with the results obtained from the numerical integrations with SECULARQUADRUPLE.

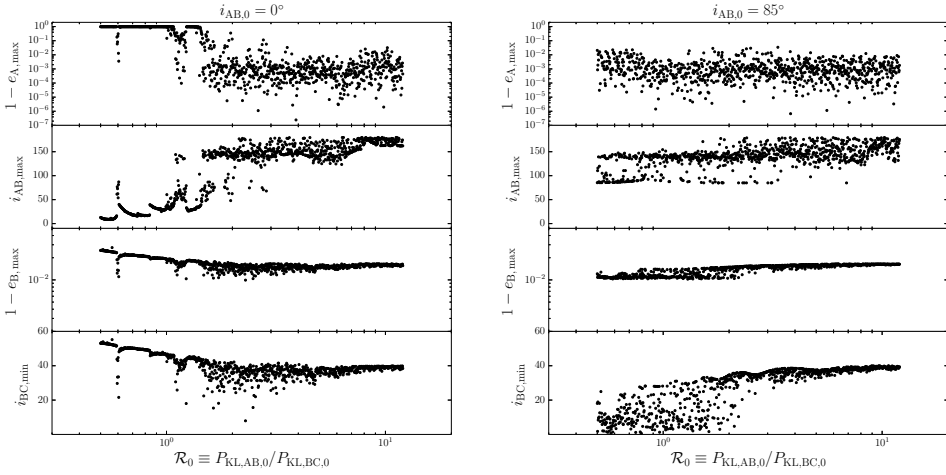


Figure 4.8: Similar to Fig. 4.7, showing greater detail near $\mathcal{R}_0 = 1$.

4.3.5 Behaviour near $\mathcal{R}_0 = 1$

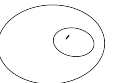
It is apparent from Fig. 4.7 that near $\mathcal{R}_0 = 1$, the behaviour of the maximum eccentricities of binaries A and B as a function of $\mathcal{R}_0$ is complex and non-regular. Here, we briefly discuss in more detail the behaviour in this regime based on numerical integrations with `SECULARQUADRUPLE`.

In Fig. 4.8, we show the same quantities as in Fig. 4.7, now based on 1000 numerical integrations within a smaller interval of $\mathcal{R}_0$ near $\mathcal{R}_0 = 1$. In the coplanar case and for $\mathcal{R}_0 \lesssim 1$, there are distinct peaks corresponding to enhanced eccentricities in both binaries A and B at specific values of $\mathcal{R}_0$. For $\mathcal{R}_0 \gtrsim 1$, individual peaks are harder to distinguish. We speculate that the peaked behaviour is due to resonances in the arguments of pericentre of binaries A and B that occur at specific integer ratios of the KL time-scales for the AB and BC pairs. In addition, for $\mathcal{R}_0 \gtrsim 1$ there may be an overlap of many resonances, thereby producing a chaotic behaviour as a function of $\mathcal{R}_0$ (Chirikov 1979). Interestingly, the peaks for $\mathcal{R}_0 \lesssim 1$ are much less pronounced, if not completely absent, in the highly inclined case.

These phenomena merit further study, but are beyond the scope of this work.

4.3.6 General relativistic effects

In the results presented above, all four bodies were assumed to be point masses and general relativistic effects were not included. In Fig. 4.9, we show the results of integrations with `SECULARQUADRUPLE` similar to those presented in Fig. 4.7, but now including 1PN precession in the equations of motion for all three binaries (cf. equation 4.9). We note that in the sequence of integrations shown in the left- and right-hand panels of Fig. 4.9, only a_A is varied; consequently, both $t_{1\text{PN},A}$ (cf. equation 4.8) and $P_{\text{KL},AB}$ (cf. equation 4.11) are affected. For the smallest value of $\mathcal{R}_0$ in Fig. 4.9 (largest value of $a_A = 1$ AU, cf. panels 1-6 in Figs 4.5 and 4.6), the initial $t_{1\text{PN},A} \approx 18.4$ Myr and $P_{\text{KL},AB} \approx 1.2$ Myr. For the largest value of $\mathcal{R}_0$ in Fig. 4.9 (smallest value



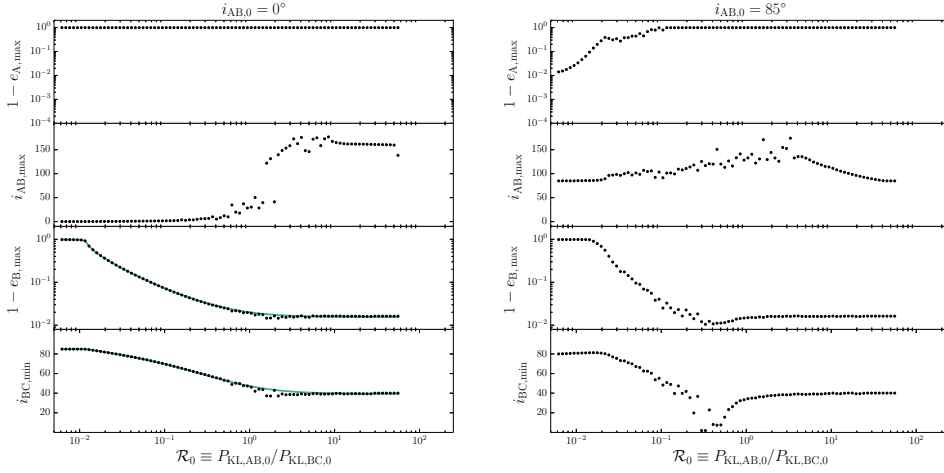


Figure 4.9: Similar to Fig. 4.7, but here with the inclusion of relativistic precession to 1PN order in all three binary systems.

of $a_A = 0.001$ AU, cf. panels 7-12 in Figs 4.5 and 4.6), the initial $t_{1\text{PN},A} \approx 5.8 \times 10^{-7}$ Myr and $P_{\text{KL},AB} \approx 3.9 \times 10^4$ Myr.

In the coplanar case, eccentricity oscillations in binary A are quenched due to relativistic precession, even if $\mathcal{R}_0 \sim 1$. We note, however, that the purely Newtonian results can be rescaled to other systems (in particular, with larger a_A), in which case relativistic precession in binary A becomes unimportant, whereas the purely Newtonian secular dynamics remain unaffected modulo a rescaling of the KL time-scales.

In the inclined case, the behaviour of the maximum eccentricity in binary A is more complicated (cf. the right-hand panel of Fig. 4.9). For the lowest $\mathcal{R}_0$, $P_{\text{KL},AB} < t_{1\text{PN},A}$ as mentioned above. As $\mathcal{R}_0$ is increased, the quantity $e_{A,\text{max}}$ decreases with increasing $\mathcal{R}_0$, which is due to the increasing relative importance of 1PN precession compared to the torque of binary B. However, the decrease of $e_{A,\text{max}}$ flattens around $\mathcal{R}_0 \approx 2 \times 10^{-2}$. The latter value of $\mathcal{R}_0$ corresponds to a significant increase of $e_{B,\text{max}}$. The flattening of $e_{A,\text{max}}$ as a function of $\mathcal{R}_0$ can be explained by considering that as $e_{B,\text{max}}$ increases, the KL time-scale for the AB pair decreases (cf. equation 4.11). Consequently, the latter KL time-scale can become comparable to the 1PN precession time-scale. Here, this is the case for $2 \times 10^{-2} \lesssim \mathcal{R}_0 \lesssim 10^{-1}$.

We show an example of this phenomenon in Fig. 4.10, where $a_A \approx 0.3$ AU and $\mathcal{R}_0 \approx 0.04$ (full parameters are given in the caption). At the maxima of e_B , the KL time-scale for the AB binary pair (black solid line in the top-left panel) decreases and becomes comparable to the 1PN precession time-scale in binary A (red solid line in the same panel). This gives rise to increased eccentricities in binary A, to much higher values if e_B were constant (cf. the top-middle panel). This is a mechanism for – at least partially – overcoming the well-known quenching of KL eccentricity cycles induced by 1PN precession. Note, however, that in this case, there is only a narrow region in $\mathcal{R}_0$ for which it is effective: as $\mathcal{R}_0$ increases, a_A decreases, therefore further decreasing $t_{1\text{PN},A}$.

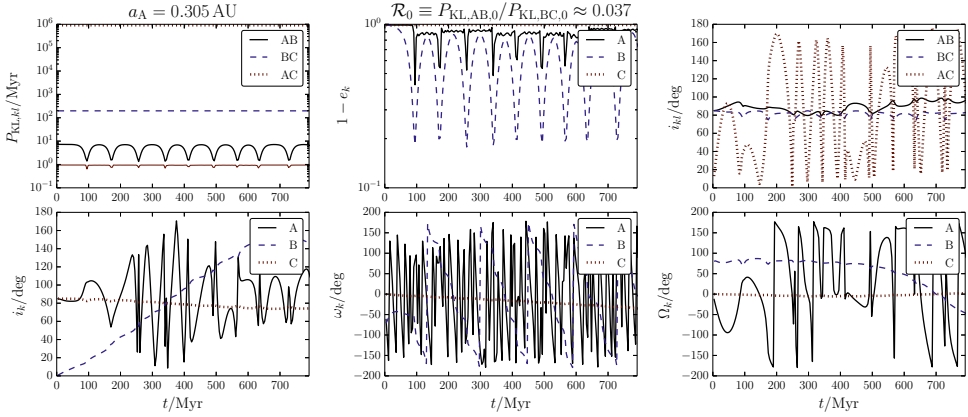


Figure 4.10: Evolution for a system taken from Fig. 4.9, demonstrating the effect of ‘overcoming’ 1PN precession by the periodically enhanced eccentricity of system B. The assumed initial parameters are semimajor axes $a_B = 0.305$ AU, $a_B = 10^2$ AU and $a_C = 5 \times 10^3$ AU, masses $m_1 = m_3 = m_4 = 1 M_\odot$ and $m_2 = 0.5 M_\odot$, eccentricities $e_A = e_B = e_C = 0.01$, inclinations $i_A = i_C = 85^\circ$ and $i_B = 0^\circ$, arguments of pericentre $\omega_A = \omega_B = \omega_C = 0^\circ$ and longitudes of the ascending nodes $\Omega_A = \Omega_B = \Omega_C = 0^\circ$. In the top-left panel, the solid red line shows the 1PN precession time-scale in binary A (cf. equation 4.8).

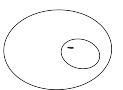
We note that this phenomenon is general, in the sense that it would also apply if precession in binary A is due to another effect, e.g. tidal effects or mass transfer in stellar systems.

4.4 Discussion

4.4.1 Application: planetary systems

As mentioned in Section 4.1, the hierarchical configuration considered in this work can be applied to planetary systems consisting of a planet+moon system (binary A) orbiting a central star (in binary B) that is orbited by a more distant and inclined planetary or stellar companion (in binary C). Both binaries A and B are assumed to be initially coplanar and circular. A pertinent question is whether the torque exerted by the fourth body causes the planet+moon system to become inclined with respect to the orbit of the central star, or whether coplanarity is maintained. We note that this is different from the question that has been addressed in the past in which case a different hierarchy was assumed, i.e. all bodies within the stellar binary were assumed to orbit the central star (Innanen et al. 1997; Takeda & Rasio 2005; Takeda et al. 2008).

Based on the qualitative results presented in Section 4.3.3, we expect that coplanarity between binaries A and B is maintained if $P_{KL,AB,0} \ll P_{KL,BC,0}$, i.e. if the binary companion is distant from the planetary orbit. In addition, we expect KL eccentricity oscillations in the orbit of the planet+moon system with respect to the central star due to the torque of the binary companion to be quenched. This effect could prevent the latter orbit from becoming highly eccentric, i.e. the presence of the moon could ‘shield’ the planet from disruption by the star as a



consequence of KL oscillations induced by the binary companion.

On the other hand, if $P_{\text{KL,AB},0} \gg P_{\text{KL,BC},0}$, the binary companion is close to the planetary orbit, and the planet+moon system can become inclined with respect to the orbit of the central star. However, in the latter case, the KL time-scale for the AB pair is long compared to that of the BC pair, such that there is no eccentricity excitation in the planet+moon system. In the intermediate regime where $P_{\text{KL,AB},0} \sim P_{\text{KL,BC},0}$, we expect significant eccentricity oscillations in the planet+moon system. These oscillations could lead to efficient tidal dissipation in cases where this would otherwise not have been important, and, in extreme cases, even to planet+moon collisions.

We explore in Section 4.4.1 some of the parameter space where significant KL eccentricity oscillations in the planet+moon system are expected, and give a number of examples in Section 4.4.1. A comprehensive population synthesis study is beyond the scope of this paper.

Expectations based on time-scale arguments

We assume a Jupiter-mass planet, $m_1 = M_J$, a moon with mass $m_2 = 10^{-4} m_1$ (the order of magnitude of the mass of Jupiter's heaviest moons), a central star with mass $m_3 = 1 M_\odot$, and a binary companion with mass $m_4 = 0.5 M_\odot$. The radii (of interest when considering collisions) are assumed to be $R_1 = R_J$, $R_2 = 10^{-2} R_1$ and $R_3 = 1 R_\odot$.

The semimajor axis of the planet+moon system is assumed to be either $a_A = 10^{-3}$ AU or $a_A = 10^{-2}$ AU; the semimajor axis a_B of the latter system with respect to the central star is either 1, 4 or 10 AU. The eccentricities of binaries A and B are assumed to be $e_A = e_B = 0.001$; the eccentricity of the orbit of the binary companion is either $e_C = 0.05$ or $e_C = 0.67$.

In Fig. 4.11, we show various time-scales of importance as a function of a_C , where in each panel different values are assumed for a_A and e_C . Quantities pertaining to the three values of a_B are indicated with blue, red and green lines for values of a_B of 1, 4 and 10 AU, respectively. The critical values of a_C corresponding to dynamical stability, computed using the three-body criterion of Mardling & Aarseth (2001) and where binary A is treated as a point mass, are indicated with vertical dashed lines for each value of a_B . Systems to the left of these lines are expected to be dynamically unstable.

Extrapolating our results from Section 4.3, we expect the region in parameter space in which e_A can be excited (in the absence of relativistic effects and other additional sources of apsidal motion), to be approximately $1 \lesssim \mathcal{R}_0 \lesssim 20$. The limiting values of a_C , for each value of a_B , are indicated with the vertical solid lines, and between these vertical lines the coloured horizontal (sloped) solid lines indicate the KL time-scales for pair AB (BC). We have indicated with hatched regions the ranges in a_C satisfying $1 < \mathcal{R}_0 < 20$, and the stability constraint.

In principle, the mechanism for producing high-amplitude oscillations in e_A in the regime $\mathcal{R}_0 \sim 1$ can be suppressed if KL oscillations in system B are quenched by relativistic precession in binary B. In all cases in Fig. 4.11, these time-scales are longer than 10 Myr, and therefore, precession in binary B is not important. Relativistic precession in binary A is of greater importance given the small values of a_A ; the associated time-scales are indicated in Fig. 4.11 with black dotted horizontal lines.

Based on Fig. 4.11, we expect eccentricity excitation in the planet+moon system for specific ranges in a_C . These ranges strongly depend on a_A , a_B and e_C . For small semimajor axes of the planet+moon system, i.e. $a_A = 10^{-2}$ AU, the criterion of dynamical stability of the orbit of the binary companion does not strongly reduce the parameter space. General relativistic

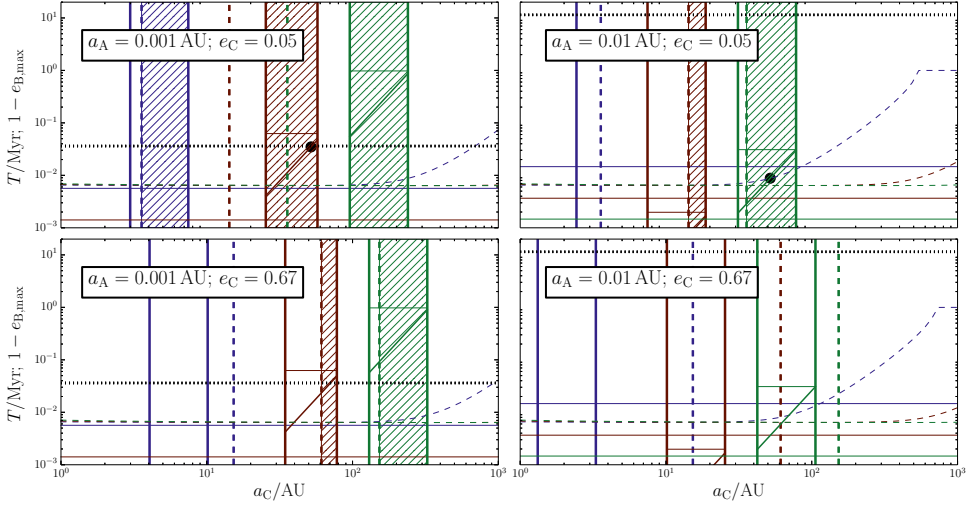
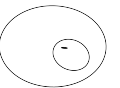


Figure 4.11: Exploration of the parameter space where eccentricity oscillations could be induced in a planet+moon system orbiting a central star that is orbited by an inclined binary companion (cf. Section 4.4.1). Blue, red and green lines correspond to semimajor axes a_B of the planet+moon system with respect to the central star of 1, 4 and 10 AU, respectively. The ranges of a_C for which $1 < \mathcal{R}_0 < 20$ are indicated with vertical thick coloured solid lines. Values of a_C corresponding to dynamical stability of the BC pair (according to the criterion of Mardling & Aarseth 2001) are indicated with vertical thick coloured dashed lines. The regions where we expect that the eccentricity of the planet+moon system is excited, are indicated with hatches. Near the hatches regions, the horizontal solid lines show $P_{KL,AB,0}$, whereas the sloped solid lines show $P_{KL,BC,0}$. The black horizontal dotted lines indicate the time-scale for relativistic precession in binary A. In addition to these time-scales, we show with non-vertical coloured dashed lines the maximum eccentricity in binary B, computed using the method of Section 4.3.4, and assuming $i_{BC,0} = 85^\circ$. The horizontal coloured solid lines show the value of $1 - e_B$ for which the planet+moon system is expected to be disrupted by the central star. In the top two panels, the black bullets correspond to the two example systems discussed in Section 4.4.1.



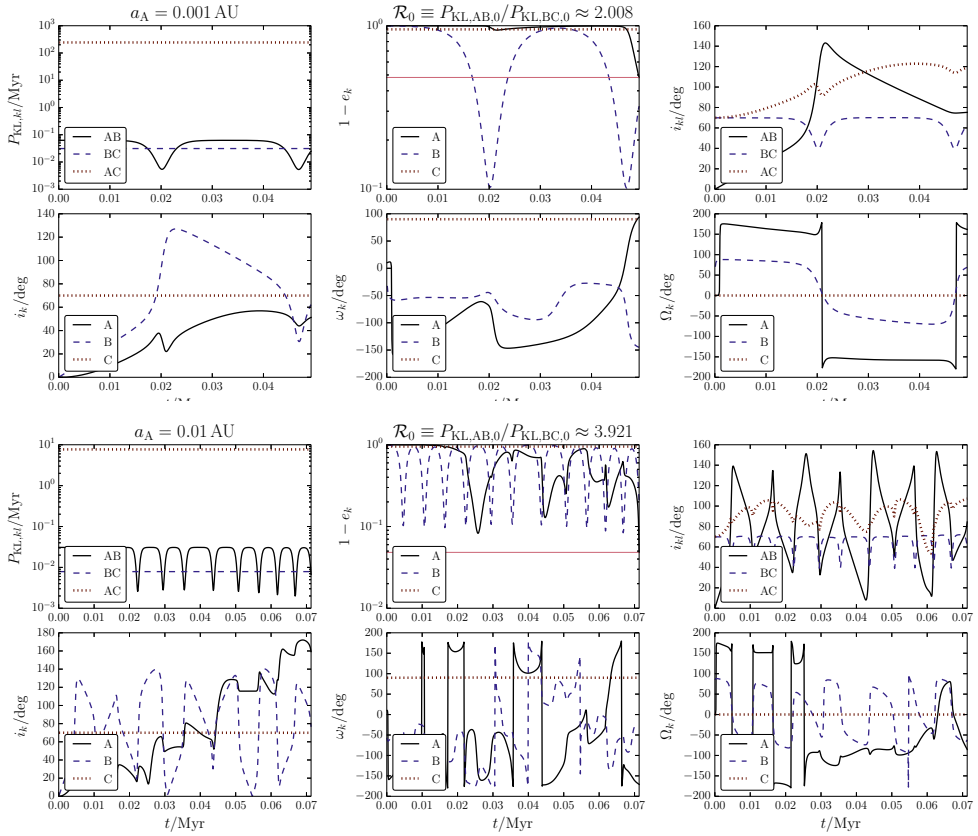


Figure 4.12: Evolution of two quadruple systems in the context of planetary systems as discussed in Section 4.4.1, computed with SECULARQUADRUPLE. First two rows: $a_A = 10^{-3}$ AU, $a_B = 4$ AU and $a_C = 50$ AU; second two rows: $a_A = 10^{-2}$ AU, $a_B = 10$ AU and $a_C = 50$ AU. Binaries A and B are initially coplanar, whereas $i_{BC,0} = 70^\circ$. In both examples, the other initial parameters were $m_1 = 1 M_J$, $m_2 = 10^{-3} M_J$, $m_3 = 1 M_\odot$ and $m_4 = 0.5 M_\odot$, $e_A = e_B = 0.001$ and $e_C = 0.05$, $i_A = i_B = 0^\circ$ and $i_C = 70^\circ$, $\omega_A = 10^\circ$, $\omega_B = 40^\circ$ and $\omega_C = 90^\circ$ and $\Omega_A = \Omega_B = \Omega_C = 0^\circ$. In the panels showing e_k , the horizontal solid red line shows the value of e_A for which the moon collides with its planet. Here, we assumed a planet radius $R_1 = 1 R_J$ and lunar radius $R_2 = 10^{-2} R_J$. The integrations were stopped when e_A reached this value.

precession is, however, also more important for smaller a_A . Nevertheless, for values of a_B of 4 and 10 AU, the relativistic precession time-scale in binary A is not much shorter than the KL time-scale for the AB pair. In those cases, there could still be high-eccentricity oscillations in binary A because of the reduction of the KL time-scale for the AB pair as a consequence of the eccentricity oscillations in binary B (cf. Section 4.3.6). This is demonstrated below in the first example in Section 4.4.1.

A larger eccentricity of the binary companion tends to reduce the parameter space of interest. The reason for this decrease is the larger range in a_C for which the system is not dynamically stable.

As discussed in Section 4.3, for $\mathcal{R}_0 \ll 1$, KL eccentricity oscillations in binary B are quenched because of the induced precession from binary A. We have plotted the maximum eccentricity in binary B as a function of a_C in Fig. 4.11 with dashed lines, computed using the semianalytic method described in Section 4.3.4. Here, we assumed $i_{BC,0} = 85^\circ$ to get a rough upper limit of the maximum eccentricity. The quenching effect is very effective for $a_B = 1$ AU and a_C larger than a few 100 AU. For large enough a_C , eccentricity oscillations in binary B are completely quenched.

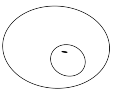
To illustrate the implications of this, we have indicated in Fig. 4.11 with horizontal coloured lines the values of $1 - e_B$ that satisfy $1 - e_B = (a_A + R_3)/a_B$, i.e. the eccentricity for which the pericentre distance of the orbit of binary B is equal to $a_A + R_3$. In the latter case, we expect the planet, the moon, or both, to be disrupted by the central star. For $a_A = 10^{-2}$ AU and $a_B = 1$ AU, the maximum eccentricity reached in binary B exceeds this value for $a_C \lesssim 100$ AU. However, for $a_C \gtrsim 100$ AU, a potentially catastrophic encounter of the planet+moon system with the central star is avoided because of quenching of the KL eccentricity oscillations in binary B. This shows more quantitatively the ‘shielding’ effect mentioned above.

To conclude, we expect that there exist regions in parameter space in which the eccentricity of the planet+moon system is excited, despite initial coplanarity. The region in parameter space is limited, however: the planet should be sufficiently far away from the central star, yet the orbit of the binary companion should also be dynamically stable. In addition, the latter orbit needs to be sufficiently inclined. In contrast, if the orbit of the binary companion is wide, the presence of the moon can prevent the orbit of the planet+moon system around the star from becoming highly eccentric.

Examples

To further illustrate the planetary system discussed here, we show in Fig. 4.12 two examples of integrations with SECULARQUADRUPE. In the first two rows, $a_A = 10^{-3}$ AU, $a_B = 4$ AU and $a_C = 50$ AU (cf. the black bullet in the top-left panel in Fig. 4.11); in the second two rows, $a_A = 10^{-2}$ AU, $a_B = 10$ AU and $a_C = 50$ AU (cf. the black bullet in the top-right panel in Fig. 4.11). In both cases, we assume $i_{BC,0} = 70^\circ$ and $e_{C,0} = 0.05$. For the other parameters, we refer to Section 4.4.1. In both examples, $R_0 \sim 1$, and high-eccentricity oscillations are expected in the planet+moon system.

The values of $1 - e_A$ corresponding to a collision between the planet and its moon are indicated with horizontal red lines in the corresponding panels in Fig. 4.12. Such collisions occur in both examples at ≈ 0.05 and 0.07 Myr, respectively, and the integrations were subsequently stopped. Note that the eccentricity of binary B does not become high enough for disruption of the planet+moon system by the central star. Particularly in the second example, e_A shows a



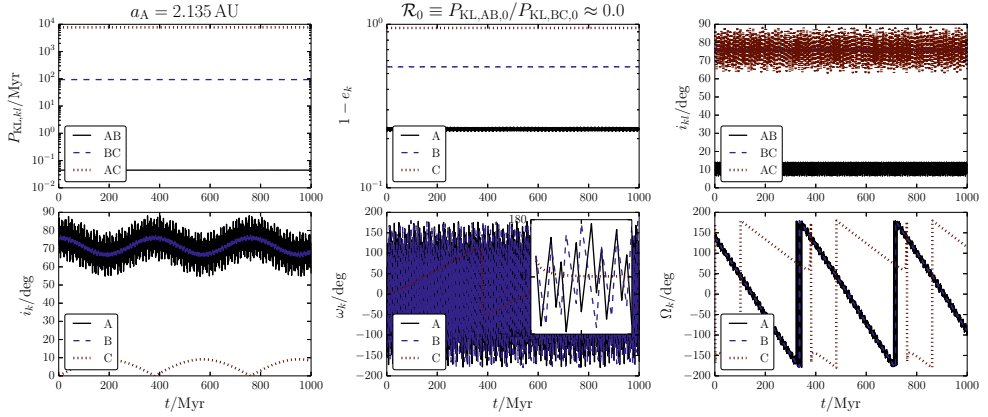


Figure 4.13: Example evolution for the quadruple system ADS 1652 as discussed in Section 4.4.2, computed with `SECULARQUADRUPLE`. The parameters are given in Table 4.3; the currently unconstrained parameters pertaining to the outermost orbit, binary C, are $e_C = 0.05$, $i_C = 0^\circ$, $\omega_C = 90.0^\circ$ and $\Omega_C = 130.0^\circ$. In the bottom middle panel, the inset shows a magnification for $t = 0$ to 20 Myr; note that both ω_A and ω_B are undersampled.

complicated behaviour as a function of time, changing rapidly each time i_{AB} passes 90° .

We remark that tidal dissipation was not included in these examples. This effect is likely important for the small pericentre distances reached during the evolution, therefore possibly not resulting in a collision, but a shrinking of the planet+moon orbit.

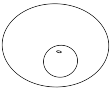
4.4.2 Application: observed stellar quadruples

ADS 1652

The quadruple system ADS 1652 (Tokovinin et al. 2014, and references therein) is composed of four main-sequence stars in the ‘3+1’ configuration. The system is likely old (age > 10 Gyr) considering the spectral types of its stellar components; the stars in binary A are of spectral type G9V, the star in binary B is of type K5V and the star in binary C is of type G8V. To date, ADS 1652 is one of few quadruple systems for which orbital fits have been obtained for multiple orbits.

m_1	m_2	m_3	m_4	a_A	a_B	a_C	e_A	e_B	e_C	i_A	i_B	i_C	ω_A	ω_B	ω_C	Ω_A	Ω_B	Ω_C
0.74	0.72	0.57	0.78	2.135	45.2	2500	0.769	0.45	–	75.9	76.0	–	287.9	175.0	–	127.2	140.0	–

Table 4.3: Parameters of the quadruple system ADS 1652 discussed in Section 4.4.2, adopted from Tokovinin et al. (2014), who fitted radial velocity and speckle measurements to the orbits of binaries A and B. The masses are obtained from these orbital fits (cf. the bottom row of table 7 of Tokovinin et al. 2014), the semimajor axes of the A and B binaries are computed from the fitted orbital periods (cf. table 4 of Tokovinin et al. 2014) using Kepler’s law. The eccentricity and orbital orientation of binary C are unknown (indicated with dashes), and are sampled with a Monte Carlo method in Section 4.4.2. Masses are expressed in solar masses, semimajor axes in AU, and angles in degrees.



Here, we apply the `SECULARQUADRUPLE` algorithm to ADS 1652 to explore its long-term secular dynamical evolution. We adopt the parameters that were obtained by Tokovinin et al. (2014), who fitted radial velocity and speckle measurements to the orbits of binaries A and B, and which are given in Table 4.3. Here, we adopted the component masses obtained from the orbital fits (cf. the bottom row of table 7 of Tokovinin et al. 2014), and computed the semimajor axes of the A and B binaries from the orbital periods (cf. table 4 of Tokovinin et al. 2014) using Kepler’s law. For the semimajor axis of the C binary, we adopt the observed projected distance of 2500 AU from binary A. Owing to its long orbital period of $\sim 10^5$ yr, the eccentricities and orbital angles of binary C are not known. Here, we proceed by sampling these quantities for 500 realizations of the system, where e_C is sampled from a thermal distribution, i_C from a distribution uniform in $\cos(i_C)$, and ω_C and Ω_C from a uniform distribution. In our integrations, we included terms up and including octupole order (excluding the cross term), and the 1PN relativistic precession terms in the three binaries. The integration time is $20 P_{\text{KL,BC}}$, which is typically a few Gyr (depending on e_C).

We show in Fig. 4.13 the evolution of an example system, where $e_C = 0.05$, $i_C = 0^\circ$, $\omega_C = 90.0^\circ$ and $\Omega_C = 130.0^\circ$. For this value of e_C , $\mathcal{R}_0 \approx 6.7 \times 10^{-4} \ll 1$, therefore the system is in the regime in which the torque of binary B on binary A dominates compared to the torque of binary C on binary B. Indeed, binaries A and B, which are initially nearly coplanar, remain nearly coplanar during the evolution (cf. the top-right panel in Fig. 4.13). Consequently, the KL eccentricity oscillations in binary A are of a very low amplitude, i.e. $e_{\text{A,max}} \approx 0.779$, whereas $e_{\text{A,0}} = 0.769$. Furthermore, KL eccentricity oscillations in binary B, which is initially inclined with respect to binary C with $i_{\text{BC,0}} \approx 70^\circ$, are completely quenched. This can be attributed to the rapid precession induced in B binary by binary A, on the time-scale of $P_{\text{KL,AB}} \approx 4 \times 10^{-2}$ Myr $\ll P_{\text{KL,BC}} \approx 10^2$ Myr (cf. the bottom middle panel of Fig. 4.13).

In Fig. 4.14, $e_{\text{C,0}} = 0.05$ was assumed to be low. The quantity $\mathcal{R}_0$ increases with increasing $e_{\text{C,0}}$ (cf. equation 4.13). Therefore, for larger $e_{\text{C,0}}$, the system could be in a very different regime in $\mathcal{R}_0$ in which the evolution is very different. This is not the case in our Monte Carlo realizations, however, for which the mean and standard deviations of $\mathcal{R}_0$ are $\approx 1.4 \times 10^{-3}$ and $\approx 1.0 \times 10^{-3}$, respectively. In Fig. 4.14, we show for the 500 integrations the maximum eccentricities in the A and B binaries, and the minimum and maximum inclinations between binaries A and B. There is very small spread in all of these quantities, showing that their dependence on e_C , as well as i_C , ω_C and Ω_C , is very weak.

We conclude that, based on the observed state of ADS 1652, the eccentricities of its orbits will remain very nearly constant for, at least, the remainder of the main-sequence time-scale of its constituents. This conclusion is independent of the currently unknown eccentricity and orientation of the outermost orbit. In particular, even if the latter orbit is highly inclined with respect to the intermediate orbit, any potential KL eccentricity oscillations in the intermediate orbit are efficiently quenched.

The Tokovinin sample of nearby FG dwarfs

As mentioned in Section 4.1, 55 of the 4847 observed systems of FG dwarfs in the catalogue of Tokovinin (2014a,b) are quadruple systems. From these, 18 are in the ‘3+1’ configuration, and for 13 of the latter, orbital periods and component masses are known for all three binaries. Here, we briefly explore in which dynamical regimes we expect these systems to be, by computing the associated value of $\mathcal{R}_0$ (cf. Section 4.3.3).

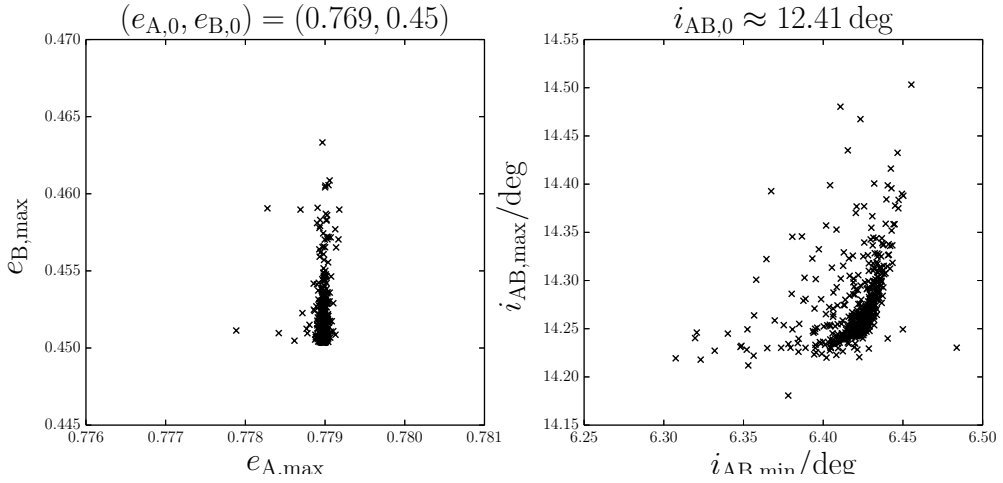


Figure 4.14: The maximum eccentricities in the A and B binaries (left-hand panel), and the minimum and maximum inclinations between binaries A and B (right-hand panel), computed from numerical integrations with `SECULARQUADRUPL` of 500 realizations of ADS 1652, where the parameters of the outermost orbit were varied (except a_C). The dependence on the latter parameters is very weak.

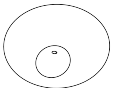
For the 13 systems mentioned above, the orbital elements, apart from the semimajor axes, are unknown. In order to compute $\mathcal{R}_0$, the eccentricities e_B and e_C are required (cf. equation 4.13). Therefore, for each of the 13 systems, we sample, in 1000 realizations, e_B and e_C from a thermal distribution. Here, we reject sampled eccentricities if either of the AB and BC pair would be unstable according to the dynamical stability criterion of Mardling & Aarseth (2001).

The distribution of the values of $\mathcal{R}_0$ obtained in this approach is shown in Fig. 4.15. The ratio $\mathcal{R}_0$ is typically small; ≈ 0.9 of the sampled systems have $\mathcal{R}_0 < 10^{-5}$. This is the regime in which the AB pair is effectively an isolated triple, and where induced precession of binary A on binary B quenches KL eccentricity oscillations in binary B, as a consequence of the torque of binary C.

We note that one might expect currently observed quadruples not to be in the regime $\mathcal{R}_0 \sim 1$. If $\mathcal{R}_0 \sim 1$, then the large eccentricities in the innermost binary would likely already have strongly affected the system, and possibly have resulted in a merger. Evidently, in this case, the system would not have been observed as a quadruple system, but as a triple system. Conversely, some of the observed quadruple systems may have been quintuple systems in the past, and triggered by secular dynamical evolution, evolved into quadruple systems through the merging of the stars in (likely) the shortest-period binary.

4.5 Conclusions

We have explored the global gravitational dynamics of hierarchical quadruple systems consisting of a hierarchical triple system orbited by a fourth body. Our main conclusions are as



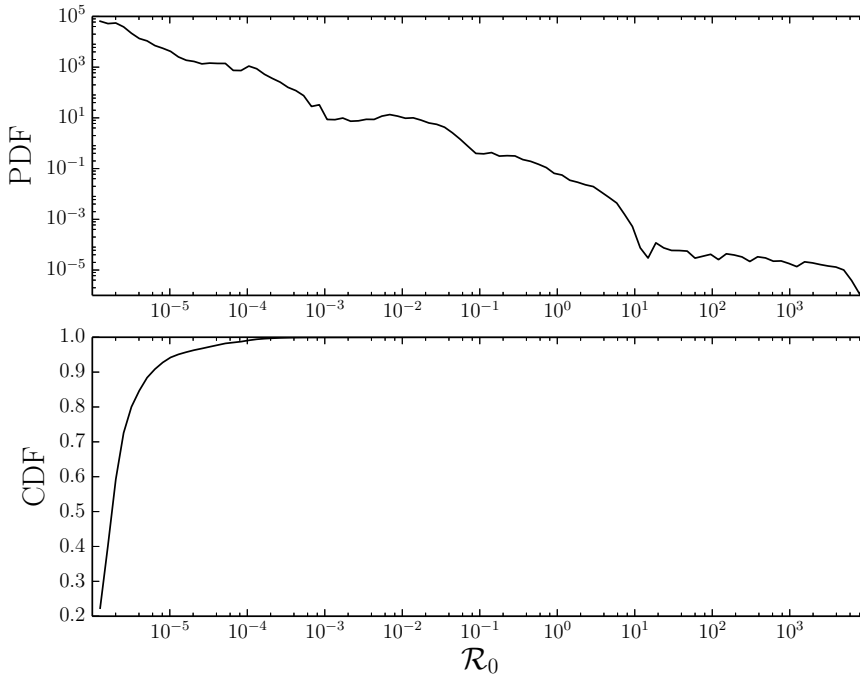


Figure 4.15: The distribution of the values of $\mathcal{R}_0$ for the ‘3+1’ quadruple systems in the catalogue of Tokovinin (2014a,b), obtained by sampling, in 1000 realizations, e_B and e_C from a thermal distribution. The probability (cumulative) density function is shown in the top (bottom) panel. For the majority of systems (≈ 0.9), $\mathcal{R}_0 < 10^{-5}$ is small.

follows.

1. The Hamiltonian for the system has been derived and expanded to up and including fourth order in the ratios of the binary separations r_A/r_B , r_B/r_C and r_A/r_C (cf. Fig. 4.1). At each order, we have found three terms that are each mathematically equivalent to the corresponding terms that appear in the hierarchical three-body problem, and that depend on the properties of only two binaries. In addition to these terms, for octupole and higher orders, we have found ‘cross terms’ that depend on properties of all three binaries. Subsequently, we have derived expressions for the orbit-averaged Hamiltonian. A preliminary analysis indicates that the cross terms are typically not important in highly hierarchical systems on short time-scales, i.e. not exceeding time-scales of order $P_{KL,BC}$, where $P_{KL,BC}$ is the KL time-scale of the BC pair. We have also derived the Hamiltonian for the configuration of two binaries orbiting each other’s barycentre (Appendix 4.A.2).

2. For highly hierarchical systems, i.e. in which the three binaries are widely separated, the global dynamics can be qualitatively described in terms of the (initial) ratio of the KL time-scales of the AB to the BC pairs, $\mathcal{R}_0 \equiv P_{KL,AB,0}/P_{KL,BC,0}$.

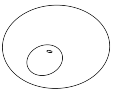
If $\mathcal{R}_0 \ll 1$, the torque of binary B on A dominates compared to the torque of binary C on binary B, and therefore binaries A and B remain coplanar if this was initially the case. If binaries A and B are initially inclined, KL eccentricity oscillations in binary A are not much affected by the presence of the fourth body. Eccentricity oscillations in binary B are efficiently quenched due to short time-scale precession induced on binary B by binary A.

If $\mathcal{R}_0 \gg 1$, the torque of binary C on binary B dominates compared to the torque of binary B on binary A. Initially, the inclination of binary B changes, whereas this is not the case for binary A. This induces a mutual inclination between binaries A and B, even if they are initially not inclined. However, rapid precession of binary B compared to the KL time-scale for the AB pair prevents any significant eccentricity oscillations in binary A, and even quenches KL oscillations if binaries A and B are initially inclined.

Lastly, if $\mathcal{R}_0 \sim 1$, complex KL eccentricity oscillations occur in binary A that are strongly coupled with the KL eccentricity oscillations in binary B. The latter are also affected compared to the situation in which binary A were replaced by a point mass, although this is typically a much smaller effect. Even if binaries A and B are initially coplanar, the induced inclination can result in very high eccentricity oscillations in binary A. These extreme eccentricities could have significant implications for strong interactions such as tidal interactions, gravitational wave dissipation, and collisions and mergers of stars and compact objects.

3. We also included the effects of general relativity, in particular relativistic precession. We have found that the range in the parameter space of the semimajor axis ratios a_B/a_A for which KL oscillations are important in binary A can be extended compared to hierarchical triple systems. This is due to a decrease of the KL time-scale of the AB pair when the eccentricity of binary B is at a maximum.

4. We have applied our results to a planetary configuration consisting of a planet+moon system orbiting a central star that is orbited by a more distant and inclined binary companion. We have found that there are regions in parameter space where a planet+moon system that is initially coplanar with respect to the central star, can become inclined and the eccentricity in the planet+moon system can be excited. This could result in significant tidal dissipation and even a collision of the planet with its moon. Furthermore, when the orbit of the binary companion



is wide, KL eccentricity oscillations in the orbit of the planet+moon system around the central star can be quenched because of induced precession from the planet+moon system. This effectively shields the planet from high-eccentricity KL oscillations from a binary companion, and, therefore, potential disruption by the central star could be avoided.

5. Lastly, we applied our results to stellar quadruple systems. In the case of ADS 1652, $\mathcal{R}_0 \sim 10^{-3}$ assuming a thermal distribution of the unknown e_C , and we find almost negligible KL eccentricity oscillations in both the innermost and intermediate orbits, binaries A and B. Even if the outer orbit, binary C, were highly inclined with respect to binary B, any potential KL eccentricity oscillations in binary B would be efficiently quenched.

For the ‘3+1’ FG stellar quadruples in the catalogue of Tokovinin (2014a,b), we estimate ≈ 0.9 of the systems to have $\mathcal{R}_0 < 10^{-5}$. Therefore, we expect that in the majority of these systems, KL eccentricity oscillations in the BC pair are quenched, and, from a secular dynamical point of view, the innermost AB pair can be considered as an isolated triple.

Acknowledgements

We thank the referee, Smadar Naoz, for providing very helpful comments that lead to improvement of the paper. This work was initiated during the International Summer-Institute for Modeling in Astrophysics (ISIMA) in 2014, hosted at CITA at the University of Toronto. It was supported by the Netherlands Research Council NWO (grants #639.073.803 [VICI], #614.061.608 [AMUSE] and #612.071.305 [LGM]) and the Netherlands Research School for Astronomy (NOVA). HBP acknowledges support from the ISF I-CORE programme 1829, The European FP-7 CIG programme ‘GRAND’ (333644), the BSF grant number 2012384 and the Asher foundation.

4.A The Hamiltonian for hierarchical quadruple systems

The general Newtonian four-body Hamiltonian is given by

$$H = T + U = \frac{1}{2} \sum_{i=1}^4 m_i \mathbf{V}_i^2 - \frac{1}{2} \sum_{i,j;i \neq j}^4 \frac{G m_i m_j}{\|\mathbf{R}_i - \mathbf{R}_j\|}, \quad (4.24)$$

where m_i , $\mathbf{R}_i$ and $\mathbf{V}_i$ denote the mass and position and velocity vectors of body i . Here, we consider dynamically stable hierarchical configurations. For four bodies, these consist of (1) a hierarchical triple system orbited by a fourth body in an orbit around the triple (discussed in detail in the main text), and (2) two binary systems orbiting each other's centre of mass. Below we discuss both configurations separately.

4.A.1 Hierarchical triple system orbited by a fourth body

In this configuration, we assume that bodies 1 and 2 are bound in binary A, body 3 is bound to the barycentre of bodies 1 and 2 in binary B, and body 4 is bound to the barycentre of bodies 1, 2 and 3 in binary C. It is convenient to define the following separation vectors,

$$\mathbf{r}_A \equiv \mathbf{R}_1 - \mathbf{R}_2; \quad (4.25a)$$

$$\mathbf{r}_B \equiv \frac{m_1 \mathbf{R}_1 + m_2 \mathbf{R}_2}{m_1 + m_2} - \mathbf{R}_3; \quad (4.25b)$$

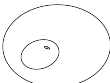
$$\mathbf{r}_C \equiv \frac{m_1 \mathbf{R}_1 + m_2 \mathbf{R}_2 + m_3 \mathbf{R}_3}{m_1 + m_2 + m_3} - \mathbf{R}_4. \quad (4.25c)$$

In addition, we define the centre of mass position of the four-body system,

$$\mathbf{r}_{\text{CM}} \equiv \left[\sum_{i=1}^4 m_i \right]^{-1} \sum_{i=1}^4 m_i \mathbf{R}_i, \quad (4.26)$$

which satisfies $\dot{\mathbf{r}}_{\text{CM}} = \mathbf{0}$. Equations (4.25) and (4.26) are easily inverted to give $\mathbf{R}_i$ in terms of $\mathbf{r}_A$, $\mathbf{r}_B$, $\mathbf{r}_C$ and $\mathbf{r}_{\text{CM}}$. Differentiating the resulting relations with respect to time, and assuming that the masses are constant, we find for the kinetic energy

$$T = \frac{1}{2} \frac{m_1 m_2}{m_1 + m_2} \dot{\mathbf{r}}_A^2 + \frac{1}{2} \frac{(m_1 + m_2) m_3}{m_1 + m_2 + m_3} \dot{\mathbf{r}}_B^2 + \frac{1}{2} \frac{(m_1 + m_2 + m_3) m_4}{m_1 + m_2 + m_3 + m_4} \dot{\mathbf{r}}_C^2. \quad (4.27)$$



To find the potential, we similarly invert equations (4.25) to give the difference vectors $\|\mathbf{R}_i - \mathbf{R}_j\|$ in terms of $\mathbf{r}_A$, $\mathbf{r}_B$ and $\mathbf{r}_C$. Substituting the resulting relations into the potential U yields six terms each of the form

$$\|\mathbf{r}_C + \alpha\mathbf{r}_B + \beta\mathbf{r}_A\|^{-1}, \quad \|\mathbf{r}_C + \alpha\mathbf{r}_B\|^{-1} \quad \text{and} \quad \|\mathbf{r}_B + \beta\mathbf{r}_A\|^{-1},$$

where α and β , which can be negative, are various mass ratios. We will assume that $r_C \gg |\alpha|r_B \gg |\beta|r_A$ (where $r \equiv \|\mathbf{r}\|$). In other words, we will assume that the system is sufficiently hierarchical in the sense that it is appropriate to expand the potential in terms of the relative distance ratios r_A/r_B , r_B/r_C and r_A/r_C , all of which are assumed to be small, and that the mass ratios are not too extreme.

We expand the potential in terms of the relative distance ratios using the general expansion

$$\begin{aligned} \|\mathbf{r} + \alpha\mathbf{r}' + \beta\mathbf{r}''\|^{-1} &= \frac{1}{r} \left[1 - \alpha \left(\frac{r'}{r} \right) (\hat{\mathbf{r}} \cdot \hat{\mathbf{r}}') - \beta \left(\frac{r''}{r} \right) (\hat{\mathbf{r}} \cdot \hat{\mathbf{r}}'') + \frac{1}{2} \alpha^2 \left(\frac{r'}{r} \right)^2 \left\{ 3 (\hat{\mathbf{r}} \cdot \hat{\mathbf{r}}')^2 - 1 \right\} + \frac{1}{2} \beta^2 \left(\frac{r''}{r} \right)^2 \left\{ 3 (\hat{\mathbf{r}} \cdot \hat{\mathbf{r}}'')^2 - 1 \right\} \right. \\ &\quad + \alpha\beta \left(\frac{r'}{r} \right) \left(\frac{r''}{r} \right) \left\{ 3 (\hat{\mathbf{r}} \cdot \hat{\mathbf{r}}') (\hat{\mathbf{r}} \cdot \hat{\mathbf{r}}'') - (\hat{\mathbf{r}}' \cdot \hat{\mathbf{r}}'') \right\} - \frac{1}{2} \alpha^3 \left(\frac{r'}{r} \right)^3 \left\{ 5 (\hat{\mathbf{r}} \cdot \hat{\mathbf{r}}')^3 - 3 (\hat{\mathbf{r}} \cdot \hat{\mathbf{r}}') \right\} - \frac{1}{2} \beta^3 \left(\frac{r''}{r} \right)^3 \left\{ 5 (\hat{\mathbf{r}} \cdot \hat{\mathbf{r}}'')^3 - 3 (\hat{\mathbf{r}} \cdot \hat{\mathbf{r}}'') \right\} \\ &\quad - \frac{1}{2} \alpha^2 \beta \left(\frac{r'}{r} \right)^2 \left(\frac{r''}{r} \right) \left\{ 15 (\hat{\mathbf{r}} \cdot \hat{\mathbf{r}}')^2 (\hat{\mathbf{r}} \cdot \hat{\mathbf{r}}'') - 3 (\hat{\mathbf{r}} \cdot \hat{\mathbf{r}}'') - 6 (\hat{\mathbf{r}} \cdot \hat{\mathbf{r}}') (\hat{\mathbf{r}}' \cdot \hat{\mathbf{r}}'') \right\} \\ &\quad - \frac{1}{2} \alpha \beta^2 \left(\frac{r'}{r} \right) \left(\frac{r''}{r} \right)^2 \left\{ 15 (\hat{\mathbf{r}} \cdot \hat{\mathbf{r}}'')^2 (\hat{\mathbf{r}} \cdot \hat{\mathbf{r}}') - 3 (\hat{\mathbf{r}} \cdot \hat{\mathbf{r}}') - 6 (\hat{\mathbf{r}} \cdot \hat{\mathbf{r}}'') (\hat{\mathbf{r}}' \cdot \hat{\mathbf{r}}'') \right\} \\ &\quad + \frac{1}{8} \alpha^4 \left(\frac{r'}{r} \right)^4 \left\{ 3 - 30 (\hat{\mathbf{r}} \cdot \hat{\mathbf{r}}')^2 + 35 (\hat{\mathbf{r}} \cdot \hat{\mathbf{r}}')^4 \right\} + \frac{1}{8} \beta^4 \left(\frac{r''}{r} \right)^4 \left\{ 3 - 30 (\hat{\mathbf{r}} \cdot \hat{\mathbf{r}}'')^2 + 35 (\hat{\mathbf{r}} \cdot \hat{\mathbf{r}}'')^4 \right\} \Bigg] \\ &\quad + \frac{1}{2} \alpha^3 \beta \left(\frac{r'}{r} \right)^3 \left(\frac{r''}{r} \right) \left\{ 35 (\hat{\mathbf{r}} \cdot \hat{\mathbf{r}}')^3 (\hat{\mathbf{r}} \cdot \hat{\mathbf{r}}'') - 15 (\hat{\mathbf{r}} \cdot \hat{\mathbf{r}}')^2 (\hat{\mathbf{r}}' \cdot \hat{\mathbf{r}}'') - 15 (\hat{\mathbf{r}} \cdot \hat{\mathbf{r}}') (\hat{\mathbf{r}} \cdot \hat{\mathbf{r}}'') + 3 (\hat{\mathbf{r}}' \cdot \hat{\mathbf{r}}'') \right\} \\ &\quad + \frac{1}{2} \alpha \beta^3 \left(\frac{r'}{r} \right) \left(\frac{r''}{r} \right)^3 \left\{ 35 (\hat{\mathbf{r}} \cdot \hat{\mathbf{r}}'')^3 (\hat{\mathbf{r}} \cdot \hat{\mathbf{r}}') - 15 (\hat{\mathbf{r}} \cdot \hat{\mathbf{r}}'')^2 (\hat{\mathbf{r}}' \cdot \hat{\mathbf{r}}') - 15 (\hat{\mathbf{r}} \cdot \hat{\mathbf{r}}') (\hat{\mathbf{r}} \cdot \hat{\mathbf{r}}'') + 3 (\hat{\mathbf{r}}' \cdot \hat{\mathbf{r}}'') \right\} \\ &\quad + \frac{1}{4} \alpha^2 \beta^2 \left(\frac{r'}{r} \right)^2 \left(\frac{r''}{r} \right)^2 \left\{ 105 (\hat{\mathbf{r}} \cdot \hat{\mathbf{r}}')^2 (\hat{\mathbf{r}} \cdot \hat{\mathbf{r}}'')^2 - 15 (\hat{\mathbf{r}} \cdot \hat{\mathbf{r}}')^2 - 15 (\hat{\mathbf{r}} \cdot \hat{\mathbf{r}}'')^2 - 15 (\hat{\mathbf{r}} \cdot \hat{\mathbf{r}}') (\hat{\mathbf{r}}' \cdot \hat{\mathbf{r}}'') + 3 (\hat{\mathbf{r}}' \cdot \hat{\mathbf{r}}'')^2 + 3 \right\} \\ &\quad + \mathcal{O} \left\{ \left(\frac{r'}{r} \right)^i \left(\frac{r''}{r} \right)^j \right\}. \end{aligned} \tag{4.28}$$

Here, $i + j \geq 5$. Substituting this expansion into equation (4.24), we find

$$\begin{aligned}
 H_{\text{ts}} = & H_{\text{bin}}(m_1, m_2, \mathbf{r}_A, \dot{\mathbf{r}}_A) + H_{\text{bin}}(m_1 + m_2, m_3, \mathbf{r}_B, \dot{\mathbf{r}}_B) + H_{\text{bin}}(m_1 + m_2 + m_3, m_4, \mathbf{r}_C, \dot{\mathbf{r}}_C) \\
 & + H_{\text{quad}}(m_1, m_2, m_3, \mathbf{r}_A, \mathbf{r}_B) + H_{\text{quad}}(m_1 + m_2, m_3, m_4, \mathbf{r}_B, \mathbf{r}_C) + H_{\text{quad}}(m_1, m_2, m_4, \mathbf{r}_A, \mathbf{r}_C) \\
 & + H_{\text{oct}}(m_1, m_2, m_3, \mathbf{r}_A, \mathbf{r}_B) + H_{\text{oct}}(m_1 + m_2, m_3, m_4, \mathbf{r}_B, \mathbf{r}_C) + H_{\text{oct}}(m_1, m_2, m_4, \mathbf{r}_A, \mathbf{r}_C) \\
 & + H_{\text{oct, cross}}(m_1, m_2, m_3, m_4, \mathbf{r}_A, \mathbf{r}_B, \mathbf{r}_C) \\
 & + H_{\text{hd}}(m_1, m_2, m_3, \mathbf{r}_A, \mathbf{r}_B) + H_{\text{hd}}(m_1 + m_2, m_3, m_4, \mathbf{r}_B, \mathbf{r}_C) + H_{\text{hd}}(m_1, m_2, m_4, \mathbf{r}_A, \mathbf{r}_C) \\
 & + H_{\text{hd, cross},1}(m_1, m_2, m_3, m_4, \mathbf{r}_A, \mathbf{r}_B, \mathbf{r}_C) + H_{\text{hd, cross},2}(m_1, m_2, m_3, m_4, \mathbf{r}_A, \mathbf{r}_B, \mathbf{r}_C) \\
 & + \mathcal{O} \left[\frac{1}{r_C} \left(\frac{r_A}{r_B} \right)^i \left(\frac{r_B}{r_C} \right)^j \left(\frac{r_A}{r_C} \right)^k \right],
 \end{aligned} \tag{4.29}$$

where ‘ts’ stands for ‘triple-star’, and $i + j + k \geq 5$. Here, the various functions are given by

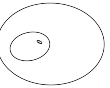
$$H_{\text{bin}}(m, m', \mathbf{r}, \dot{\mathbf{r}}) = \frac{1}{2} \frac{mm'}{m + m'} \dot{\mathbf{r}}^2 - \frac{Gmm'}{r}; \tag{4.30a}$$

$$H_{\text{quad}}(m, m', m'', \mathbf{r}, \mathbf{r}') = -\frac{Gmm'm''}{m + m'} \frac{1}{r'} \left(\frac{r}{r'} \right)^2 \frac{1}{2} \left[3(\hat{\mathbf{r}} \cdot \hat{\mathbf{r}}')^2 - 1 \right]; \tag{4.30b}$$

$$H_{\text{oct}}(m, m', m'', \mathbf{r}, \mathbf{r}') = -\frac{Gmm'm''(m - m')}{(m + m')^2} \frac{1}{r'} \left(\frac{r}{r'} \right)^3 \frac{1}{2} \left[5(\hat{\mathbf{r}} \cdot \hat{\mathbf{r}}')^3 - 3(\hat{\mathbf{r}} \cdot \hat{\mathbf{r}}') \right]; \tag{4.30c}$$

$$\begin{aligned}
 H_{\text{oct, cross}}(m_1, m_2, m_3, m_4, \mathbf{r}_A, \mathbf{r}_B, \mathbf{r}_C) = & \frac{Gm_1m_2m_3m_4}{(m_1 + m_2)(m_1 + m_2 + m_3)} \frac{1}{r_C} \left(\frac{r_A}{r_C} \right)^2 \left(\frac{r_B}{r_C} \right) \\
 & \times \frac{1}{2} \left[15(\hat{\mathbf{r}}_B \cdot \hat{\mathbf{r}}_C)(\hat{\mathbf{r}}_A \cdot \hat{\mathbf{r}}_C)^2 - 3(\hat{\mathbf{r}}_B \cdot \hat{\mathbf{r}}_C) - 6(\hat{\mathbf{r}}_A \cdot \hat{\mathbf{r}}_C)(\hat{\mathbf{r}}_A \cdot \hat{\mathbf{r}}_B) \right];
 \end{aligned} \tag{4.30d}$$

$$\tag{4.30e}$$



$$H_{\text{hd}}(m, m', m'', \mathbf{r}, \mathbf{r}') = -\frac{Gmm'm''(m^2 - mm' + m'^2)}{(m + m')^3} \frac{1}{r'} \left(\frac{r}{r'}\right)^4 \frac{1}{8} \left[35 (\hat{\mathbf{r}} \cdot \hat{\mathbf{r}}')^4 - 30 (\hat{\mathbf{r}} \cdot \hat{\mathbf{r}}')^2 + 3 \right]; \quad (4.30f)$$

$$H_{\text{hd, cross},1}(m_1, m_2, m_3, m_4, \mathbf{r}_A, \mathbf{r}_B, \mathbf{r}_C) = -\frac{Gm_1m_2(m_1 - m_2)m_3m_4}{(m_1 + m_2)^2(m_1 + m_2 + m_3)} \frac{1}{r_C} \left(\frac{r_A}{r_C}\right)^3 \left(\frac{r_B}{r_C}\right) \quad (4.30g)$$

$$\times \frac{1}{2} \left[3(\hat{\mathbf{r}}_A \cdot \hat{\mathbf{r}}_B) \left\{ 5(\hat{\mathbf{r}}_A \cdot \hat{\mathbf{r}}_C)^2 - 1 \right\} + 5(\hat{\mathbf{r}}_A \cdot \hat{\mathbf{r}}_C)(\hat{\mathbf{r}}_B \cdot \hat{\mathbf{r}}_C) \left\{ 3 - 7(\hat{\mathbf{r}}_A \cdot \hat{\mathbf{r}}_C)^2 \right\} \right];$$

$$H_{\text{hd, cross},2}(m_1, m_2, m_3, m_4, \mathbf{r}_A, \mathbf{r}_B, \mathbf{r}_C) = -\frac{Gm_1m_2m_3^2m_4}{(m_1 + m_2)(m_1 + m_2 + m_3)^2} \frac{1}{r_C} \left(\frac{r_A}{r_C}\right)^2 \left(\frac{r_B}{r_C}\right)^2 \quad (4.30h)$$

$$\times \frac{3}{4} \left[1 + 2(\hat{\mathbf{r}}_A \cdot \hat{\mathbf{r}}_B)^2 - 20(\hat{\mathbf{r}}_A \cdot \hat{\mathbf{r}}_B)(\hat{\mathbf{r}}_A \cdot \hat{\mathbf{r}}_C)(\hat{\mathbf{r}}_B \cdot \hat{\mathbf{r}}_C) - 5(\hat{\mathbf{r}}_B \cdot \hat{\mathbf{r}}_C)^2 + 5(\hat{\mathbf{r}}_A \cdot \hat{\mathbf{r}}_C)^2 \left\{ 7(\hat{\mathbf{r}}_B \cdot \hat{\mathbf{r}}_C)^2 - 1 \right\} \right].$$

The function $H_{\text{bin}}(m, m', \mathbf{r}, \dot{\mathbf{r}})$ is the Hamiltonian for an isolated two-body system with reduced mass $\mu = mm'/(m + m')$. It appears in equation (4.29) for the three binaries A, B and C, and reduces to the binding energy of each these binaries if Kepler orbits are assumed. Therefore, it does not lead to orbital changes.

The other functions defined in equations (4.30) do lead to orbital changes. We associate a term with the ‘quadrupole’ order if the combined power of (r_A/r_B) , (r_B/r_C) and (r_A/r_C) is equal to two, to ‘octupole’ order if the combined power is equal to three and ‘hexadecupole’ order if the combined power is equal to four. The functions $H_{\text{quad}}(m, m', m'', \mathbf{r}, \mathbf{r}')$, $H_{\text{oct}}(m, m', m'', \mathbf{r}, \mathbf{r}')$ and $H_{\text{hd}}(m, m', m'', \mathbf{r}, \mathbf{r}')$ are precisely the same functions that appear in the three-body problem. In this case, they each appear in the Hamiltonian three times by replacing the ‘inner’ and ‘outer’ binaries by the combinations AB, BC and AC.

At the quadrupole level, the three combinations of $H_{\text{quad}}(m, m', m'', \mathbf{r}, \mathbf{r}')$ are the only terms that appear; the ‘cross terms’ that are present in the expansion in equation (4.28) cancel. Such a cancellation does not occur at higher orders. At octupole order, we find the ‘cross term’ $H_{\text{oct, cross}}(m_1, m_2, m_3, m_4, \mathbf{r}_A, \mathbf{r}_B, \mathbf{r}_C)$ that is unique to the type of quadruple systems considered here. It depends on all four masses, all three semimajor axes and all three relative orientations between the binary separation vectors, and it is proportional to $(r_A/r_C)^2(r_B/r_C)$. At the hexadecupole order, we find two cross terms, $H_{\text{hd, cross},1} \propto (r_A/r_C)^3(r_B/r_C)$ and $H_{\text{hd, cross},2} \propto (r_A/r_C)^2(r_B/r_C)^2$. Equation (4.29) is exact if expanded to infinite order. Here, the expansion is truncated for $i + j + k \geq 5$.

We average the truncated Hamiltonian over the three binary orbits A, B and C, assuming that the Kepler orbit is unperturbed during this time-scale. For orbit k , the averaging is defined as

$$\langle H \rangle_k \equiv \frac{1}{2\pi} \int_0^{2\pi} H \, dl_k, \quad (4.31)$$

where l_k is the mean anomaly of orbit k . Depending on the sign of the power of r_k it is convenient to use either the true anomaly or eccentric anomaly in equation (4.31). If the power of r_k is negative, we use the true anomaly; if it is positive, we use the eccentric anomaly. Formally, averaging the Hamiltonian over the three orbits A, B and C according to equation (4.31) is not a canonical transformation. However, applying the Von Zeipel transformation technique to the unaveraged Hamiltonian (Brouwer 1959), a canonical transformation can be found that eliminates the short-period terms l_k from the Hamiltonian (cf. appendix A2 of Naoz et al. 2013a; the derivation presented there is straightforwardly extended to three, rather than two short-period variables). This transformation leads to a transformed Hamiltonian that is equivalent to the triply-averaged Hamiltonian $\langle\langle\langle H \rangle_{\text{C}} \rangle_{\text{B/A}} \rangle$ (cf. equation 4.31; note that the order of integration is arbitrary). Here, the transformed coordinates e_k^* and q_k^* differ from the original ones e_k and q_k . However, as noted by Naoz et al. 2013a, the differences between the untransformed and the transformed coordinates contribute to the Hamiltonian only at subleading order.

We express the angular momenta and orientations of each of the three binaries in terms of the triad of perpendicular orbital state vectors $(\hat{j}_k, \mathbf{e}_k, \mathbf{q}_k)$, where $\mathbf{q}_k \equiv \hat{j}_k \times \mathbf{e}_k$. Here, $\hat{j}_k$ is a vector aligned with the angular momentum vector of the orbit with magnitude $j_k = \sqrt{1 - e_k^2}$; $\mathbf{e}_k$ is the eccentricity or Laplace-Runge-Lenz vector that is aligned with the major axis and with magnitude the orbital eccentricity e_k (see e.g. Goldstein 1975, 1976; Goldstein et al. 2002 for historical overviews). In terms of these vectors and the true anomaly, the angle between two instantaneous separation vectors $\mathbf{r}_k$ and $\mathbf{r}_l$ can be expressed as

$$\begin{aligned} \hat{\mathbf{r}}_k \cdot \hat{\mathbf{r}}_l &= [\cos(f_k) \hat{\mathbf{e}}_k + \sin(f_k) \hat{\mathbf{q}}_k] \cdot [\cos(f_l) \hat{\mathbf{e}}_l + \sin(f_l) \hat{\mathbf{q}}_l] \\ &= \cos(f_k) \cos(f_l) (\hat{\mathbf{e}}_k \cdot \hat{\mathbf{e}}_l) + \cos(f_k) \sin(f_l) (\hat{\mathbf{e}}_k \cdot \hat{\mathbf{q}}_l) + \sin(f_k) \cos(f_l) (\hat{\mathbf{q}}_k \cdot \hat{\mathbf{e}}_l) + \sin(f_k) \sin(f_l) (\hat{\mathbf{q}}_k \cdot \hat{\mathbf{q}}_l). \end{aligned} \quad (4.32)$$

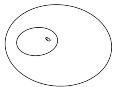
Our result of the orbit averaging $\overline{H}_{\text{ts}} \equiv \langle\langle\langle H_{\text{ts}} \rangle_{\text{C}} \rangle_{\text{B/A}} \rangle$ is³⁴

$$\overline{H}_{\text{bin}}(m, m', a) = -\frac{Gmm'}{2a}; \quad (4.33a)$$

$$\overline{H}_{\text{quad}}(m, m', m'', a, a', \mathbf{j}, \mathbf{e}, \mathbf{j}', \mathbf{e}') = \frac{Gmm'm''}{m+m'} \frac{1}{a'} \left(\frac{a}{a'}\right)^2 \frac{1}{8j^{1/3}} \left[1 - 6e^2 + 15(\mathbf{e} \cdot \hat{\mathbf{j}}')^2 - 3(1 - e^2)(\hat{\mathbf{j}} \cdot \hat{\mathbf{j}}')^2 \right]; \quad (4.33b)$$

³⁴The orbit averaging can be carried out in any order of A, B and C, e.g. C $\rightarrow$ B $\rightarrow$ A or C $\rightarrow$ A $\rightarrow$ B. Because the integration limits are constants, the result is not affected by the order of integration.

⁴We were unable to derive the simplification in $\overline{H}_{\text{ocd}}$ analytically, but we verified it by evaluating both sides numerically.



$$\begin{aligned}
\overline{H}_{\text{oct}}(m, m', m'', a, a', j, e, j', e') = & -\frac{Gmm'm''(m-m')}{(m+m')^2} \frac{1}{a'} \left(\frac{a}{a'}\right)^3 \frac{15}{64j^{5/5}} \left[(\mathbf{e} \cdot \mathbf{e}') \left\{ 2e^2 - 9 + 10(1-e^2)(\hat{\mathbf{e}} \cdot \hat{\mathbf{e}}')^2 \right. \right. \\
& + (10 + 25e^2)(\hat{\mathbf{e}} \cdot \hat{\mathbf{j}}')^2 + 10(1-e^2)(\hat{\mathbf{e}}' \cdot \hat{\mathbf{j}}')^2 \left. \left[1 - (\hat{\mathbf{e}} \cdot \hat{\mathbf{j}}')^2 \right] + 5(1-e^2)(\hat{\mathbf{j}} \cdot \hat{\mathbf{j}}')^2 \left[2(\hat{\mathbf{e}} \cdot \hat{\mathbf{e}}')^2 - 1 \right] \right\} \\
& - 10(1-e^2) \left[2(\hat{\mathbf{e}} \cdot \hat{\mathbf{e}}')^2 - 1 \right] (\mathbf{e} \cdot \hat{\mathbf{j}}') (\mathbf{e}' \cdot \hat{\mathbf{j}}) (\hat{\mathbf{j}} \cdot \hat{\mathbf{j}}') \Big] \\
= & -\frac{Gmm'm''(m-m')}{(m+m')^2} \frac{1}{a'} \left(\frac{a}{a'}\right)^3 \frac{15}{64j^{5/5}} \left[(\mathbf{e} \cdot \mathbf{e}') \left\{ 1 - 8e^2 + 35(\mathbf{e} \cdot \hat{\mathbf{j}}')^2 \right. \right. \\
& - 5(1-e^2)(\hat{\mathbf{j}} \cdot \hat{\mathbf{j}}')^2 \Big\} - 10(1-e^2)(\mathbf{e} \cdot \hat{\mathbf{j}}') (\mathbf{e}' \cdot \hat{\mathbf{j}}) (\hat{\mathbf{j}} \cdot \hat{\mathbf{j}}') \Big] ; \tag{4.33c}
\end{aligned}$$

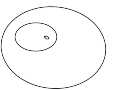
$$\begin{aligned}
\overline{H}_{\text{oct, cross}}(m_1, m_2, m_3, m_4, a_A, a_B, a_C, \mathbf{j}_A, \mathbf{e}_A, \mathbf{j}_B, \mathbf{e}_B, \mathbf{j}_C, \mathbf{e}_C) = & -\frac{Gm_1m_2m_3m_4}{(m_1+m_2)(m_1+m_2+m_3+m_4)} \frac{1}{a_C} \left(\frac{a_A}{a_C}\right)^2 \left(\frac{a_B}{a_C}\right) \frac{9}{32j_C^{5/5}} \\
& \times \left[2(1-e_A^2)(\mathbf{e}_B \cdot \hat{\mathbf{j}}_A)(\mathbf{e}_C \cdot \hat{\mathbf{j}}_A) \left\{ 4 - 5(\hat{\mathbf{e}}_A \cdot \hat{\mathbf{j}}_C)^2 \right\} - 10(\hat{\mathbf{e}}_A \cdot \mathbf{e}_C)(\hat{\mathbf{e}}_A \cdot \hat{\mathbf{j}}_C) \left\{ (1 + 4e_A^2)(\mathbf{e}_B \cdot \hat{\mathbf{j}}_C) \right. \right. \\
& - (1-e_A^2)(\mathbf{e}_B \cdot \hat{\mathbf{j}}_A)(\hat{\mathbf{j}}_A \cdot \hat{\mathbf{j}}_C) \Big\} + (\mathbf{e}_B \cdot \mathbf{e}_C) \left\{ - (1 - 6e_A^2) - 10(1-e_A^2)(\hat{\mathbf{e}}_C \cdot \hat{\mathbf{j}}_A)^2 - 5(\hat{\mathbf{e}}_A \cdot \hat{\mathbf{j}}_C)^2 \right. \\
& \times \left[5e_A^2 - 2(1-e_A^2)(\hat{\mathbf{e}}_C \cdot \hat{\mathbf{j}}_A)^2 \right] - 20(1-e_A^2)(\hat{\mathbf{e}}_A \cdot \hat{\mathbf{e}}_C)(\hat{\mathbf{e}}_A \cdot \hat{\mathbf{j}}_C)(\hat{\mathbf{e}}_C \cdot \hat{\mathbf{j}}_A) (\hat{\mathbf{j}}_A \cdot \hat{\mathbf{j}}_C) + 5(1-e_A^2)(\hat{\mathbf{j}}_A \cdot \hat{\mathbf{j}}_C)^2 \\
& - 10(1-e_A^2)(\hat{\mathbf{e}}_A \cdot \hat{\mathbf{e}}_C)^2 \left. \left[1 - (\hat{\mathbf{j}}_A \cdot \hat{\mathbf{j}}_C)^2 \right] \right\} + 10(\hat{\mathbf{e}}_A \cdot \mathbf{e}_B) \left\{ (1-e_A^2)(\hat{\mathbf{e}}_A \cdot \hat{\mathbf{j}}_C)(\hat{\mathbf{e}}_C \cdot \hat{\mathbf{j}}_A) (\hat{\mathbf{j}}_A \cdot \hat{\mathbf{j}}_C) \right. \\
& \left. \left. + (\hat{\mathbf{e}}_A \cdot \mathbf{e}_C) \left[1 - (1-e_A^2)(\hat{\mathbf{j}}_A \cdot \hat{\mathbf{j}}_C)^2 \right] \right\} \right] ; \tag{4.33d}
\end{aligned}$$

$$\begin{aligned}
\overline{H}_{\text{hd}}(m, m', m'', a, a', j, e, j', e') &= \frac{Gmm'm''(m^2 - mm' + m'^2)}{(m + m')^3} \frac{1}{a'} \left(\frac{a}{a'}\right)^4 \frac{3}{1024j^7} \\
&\times \left[262 + 423e'^2 - 40e^2 \{-38 - 3e'^2 + 6e^2(9 + 2e'^2)\} - 280 \left\{ (\hat{e} \cdot \hat{j}')^2 + (\hat{e}' \cdot \hat{j})^2 \right\} + 5 \left\{ -441e^4(2 + e'^2)(\hat{e} \cdot \hat{j}')^4 \right. \right. \\
&+ 8 \{ 7e^2(-5 + 6e^2) + 3(-4 + e^2 + 3e^4)e'^2 \} (\hat{e}' \cdot \hat{j})^2 + 56(-1 - 5e^2 + 6e^4)(2 + 3e'^2)(\hat{e} \cdot \hat{e}')(\hat{e} \cdot \hat{j}')(\hat{e}' \cdot \hat{j})(\hat{j} \cdot \hat{j}') \\
&- 112(-1 - 5e^2 + 6e^4)(2 + 3e'^2)(\hat{e} \cdot \hat{e}')^3(\hat{e} \cdot \hat{j}')(\hat{e}' \cdot \hat{j})(\hat{j} \cdot \hat{j}') - 2(1 - e^2) \{ 10 + 39e'^2 + 6e^2(38 + 11e'^2) \\
&- 42(1 - e^2)e'^2(\hat{e}' \cdot \hat{j})^2 \} (\hat{j} \cdot \hat{j}')^2 - 21(1 - e^2)^2(2 + e'^2)(\hat{j} \cdot \hat{j}')^4 - 56(-1 - 5e^2 + 6e^4)(2 + 3e'^2)(\hat{e} \cdot \hat{e}')^4 \\
&\times \left\{ 1 - (\hat{j} \cdot \hat{j}')^2 \right\} + 14(\hat{e} \cdot \hat{j}')^2 \{ -26e^2 + 84e^4 + 3(-2 - e^2 + 6e^4)e'^2 - (1 - e^2) \{ [-4 - 6e'^2 \\
&+ 6e^2(-4 + e'^2)\}(\hat{e}' \cdot \hat{j})^2 - 21e^2(2 + e'^2)(\hat{j} \cdot \hat{j}')^2 \} + 28(\hat{e} \cdot \hat{e}')^2 [-2(-1 - 5e^2 + 6e^4)(2 + 3e'^2)(\hat{e}' \cdot \hat{j})^2 \\
&+ (\hat{e} \cdot \hat{j}')^2 \{ 4 + 20e^2 - 24e^4 + 3(2 + 10e^2 + 9e^4)e'^2 + 2(-1 - 5e^2 + 6e^4)(2 + 3e'^2)(\hat{e}' \cdot \hat{j})^2 \\
&+ 3 \{ -2 - 3e'^2 + 2e^2[-5 + 6e^2 - 5(1 - e^2)e'^2] + (1 - e^2)[2 + 3e'^2 + e^2(12 + 11e'^2)](\hat{j} \cdot \hat{j}')^2 \} \} \} \Big]. \quad (4.33e)
\end{aligned}$$

In order to simplify the expressions in equation (4.33) we repeatedly used a vector identity for the dot product of two vector products, $(\mathbf{a} \times \mathbf{b}) \cdot (\mathbf{c} \times \mathbf{d}) = (\mathbf{a} \cdot \mathbf{c})(\mathbf{b} \cdot \mathbf{d}) - (\mathbf{b} \cdot \mathbf{c})(\mathbf{a} \cdot \mathbf{d})$, and the scalar product of two scalar triple products,

$$[(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c}][(\mathbf{d} \times \mathbf{e}) \cdot \mathbf{f}] = \det \begin{vmatrix} \mathbf{a} \cdot \mathbf{d} & \mathbf{a} \cdot \mathbf{e} & \mathbf{a} \cdot \mathbf{f} \\ \mathbf{b} \cdot \mathbf{d} & \mathbf{b} \cdot \mathbf{e} & \mathbf{b} \cdot \mathbf{f} \\ \mathbf{c} \cdot \mathbf{d} & \mathbf{c} \cdot \mathbf{e} & \mathbf{c} \cdot \mathbf{f} \end{vmatrix}. \quad (4.34)$$

Furthermore, we have omitted the explicit expressions for $\overline{H}_{\text{hd, cross, 1}}$ and $\overline{H}_{\text{hd, cross, 2}}$ because they are excessively long. The expressions for $\overline{H}_{\text{quad}}$ and $\overline{H}_{\text{oct}}$ are identical to those of Boué & Fabrycky (2014), who also adopted a description in terms of vectorial vectors.



4.A.2 Two binaries orbiting each other

In this configuration, we assume that bodies 1 and 2 are bound in binary A, bodies 3 and 4 are bound in binary B, and the barycentres of binaries A and B are bound in binary C. Although not explored in the main text, here, we present the formalism that can be used to study this hierarchy. The derivation closely parallels that of Section 4.A.1. We define the following instantaneous separations,

$$\mathbf{r}_A \equiv \mathbf{R}_1 - \mathbf{R}_2; \quad (4.35a)$$

$$\mathbf{r}_B \equiv \mathbf{R}_3 - \mathbf{R}_4; \quad (4.35b)$$

$$\mathbf{r}_C \equiv \frac{m_1 \mathbf{R}_1 + m_2 \mathbf{R}_2}{m_1 + m_2} - \frac{m_3 \mathbf{R}_3 + m_4 \mathbf{R}_4}{m_3 + m_4}. \quad (4.35c)$$

Again assuming that the centre of mass (cf. equation 4.26) is constant, the kinetic energy is given by

$$T = \frac{1}{2} \frac{m_1 m_2}{m_1 + m_2} \dot{\mathbf{r}}_A^2 + \frac{1}{2} \frac{m_3 m_4}{m_3 + m_4} \dot{\mathbf{r}}_B^2 + \frac{1}{2} \frac{(m_1 + m_2)(m_3 + m_4)}{m_1 + m_2 + m_3 + m_4} \dot{\mathbf{r}}_C^2. \quad (4.36)$$

In contrast to Section 4.A.1, we do not assume that $r_A \ll r_B \ll r_C$. Instead, we assume that $r_A \ll r_C$ and $r_B \ll r_C$, without making any (explicit) assumptions on the relation between r_A and r_B . Using the general expansion equation (4.28) we find for the Hamiltonian (cf. equation 4.24) in terms of the variables in equation (4.35)

$$\begin{aligned} H_{bb} = & H_{\text{bin}}(m_1, m_2, \mathbf{r}_A, \dot{\mathbf{r}}_A) + H_{\text{bin}}(m_3, m_4, \mathbf{r}_B, \dot{\mathbf{r}}_B) + H_{\text{bin}}(m_1 + m_2, m_3 + m_4, \mathbf{r}_C, \dot{\mathbf{r}}_C) \\ & + H_{\text{quad}}(m_1, m_2, m_3 + m_4, \mathbf{r}_A, \mathbf{r}_C) + H_{\text{quad}}(m_3, m_4, m_1 + m_2, \mathbf{r}_B, \mathbf{r}_C) \\ & + H_{\text{oct}}(m_1, m_2, m_3 + m_4, \mathbf{r}_A, \mathbf{r}_C) + H_{\text{oct}}(m_4, m_3, m_1 + m_2, \mathbf{r}_B, \mathbf{r}_C) \\ & + H_{\text{hd}}(m_1, m_2, m_3 + m_4, \mathbf{r}_A, \mathbf{r}_C) + H_{\text{hd}}(m_4, m_3, m_1 + m_2, \mathbf{r}_B, \mathbf{r}_C) + H_{\text{hd, cross,bb}}(m_1, m_2, m_3, m_4, \mathbf{r}_A, \mathbf{r}_B, \mathbf{r}_C) \\ & + \mathcal{O} \left[\frac{1}{r_C} \left(\frac{r_A}{r_B} \right)^i \left(\frac{r_B}{r_C} \right)^j \left(\frac{r_A}{r_C} \right)^k \right], \end{aligned} \quad (4.37)$$

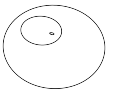
where ‘bb’ stands for ‘binary-binary’, and $i + j + k \geq 5$. Here, the various functions are the same as in equation (4.30). The cross term $H_{\text{hd, cross, bb}}$ is unique to the binary-binary configuration, and is given by

$$H_{\text{hd, cross, bb}}(m_1, m_2, m_3, m_4, \mathbf{r}_A, \mathbf{r}_B, \mathbf{r}_C) = -\frac{Gm_1m_2m_3m_4}{(m_1 + m_2)(m_3 + m_4)} \frac{1}{r_C} \left(\frac{r_A}{r_C}\right)^2 \left(\frac{r_B}{r_C}\right)^2 \\ \times \frac{3}{4} \left[1 - 5(\hat{\mathbf{r}}_A \cdot \hat{\mathbf{r}}_C)^2 - 5(\hat{\mathbf{r}}_B \cdot \hat{\mathbf{r}}_C)^2 + 35(\hat{\mathbf{r}}_A \cdot \hat{\mathbf{r}}_C)^2(\hat{\mathbf{r}}_B \cdot \hat{\mathbf{r}}_C)^2 + 2(\hat{\mathbf{r}}_A \cdot \hat{\mathbf{r}}_B)^2 - 20(\hat{\mathbf{r}}_A \cdot \hat{\mathbf{r}}_B)(\hat{\mathbf{r}}_A \cdot \hat{\mathbf{r}}_C)(\hat{\mathbf{r}}_B \cdot \hat{\mathbf{r}}_C) \right]. \quad (4.38)$$

As might be expected, equation (4.38) is invariant under interchange of the A and B binaries, i.e. $(m_1, m_2, \mathbf{r}_A) \leftrightarrow (m_3, m_4, \mathbf{r}_B)$.

It is interesting to compare the (non-averaged) Hamiltonian between the ‘triple-single’ and ‘binary-binary’ configurations. For the latter, the Hamiltonian is simpler, in the sense that at each order, there are only two, rather than three, terms that depend on the properties of two binaries (H_{quad} , H_{oct} and H_{hd} – for each of these, the combinations AC and BC occur). In addition, ‘cross terms’ appear at octupole order in the ‘triple-single’ configuration, whereas for the ‘binary-binary’ configuration, the lowest order at which these terms appear is the next higher, hexadecupole, order.

The orbit-averaged Hamiltonian is obtained directly by appropriate substitutions in equation (4.37), using the results of equations (4.33a), (4.33b), (4.33c) and (4.33e). The orbit-average of the cross term equation (4.38) is excessively long, and not included here.



4.B Test of the SecularQuadruple algorithm for three-body systems

We tested part of the `SECULARQUADRUPLE` algorithm by comparing to previously obtained integrations for three-body systems. We show an example in Fig. 4.16, where we assumed the same hierarchical three-body system as in fig. 3 of Naoz et al. (2013a). Here, we have both tested applying the triple parameters to the AB systems, choosing a very large value of a_C (essentially making AB an isolated triple), and to the BC systems, making a_A very small (essentially reducing binary A to a point mass). The results from `SECULARQUADRUPLE` agree very well with those of Naoz et al. (2013a), who performed both integrations based on the orbit-averaged equations, and direct- N body integrations. The maximum relative error of the ODE variables between time-steps was set to 10^{-15} . Consequently, the Hamiltonian is conserved to high accuracy; $|\Delta E/E| < 2 \times 10^{-9}$ for the integration shown in Fig. 4.16.

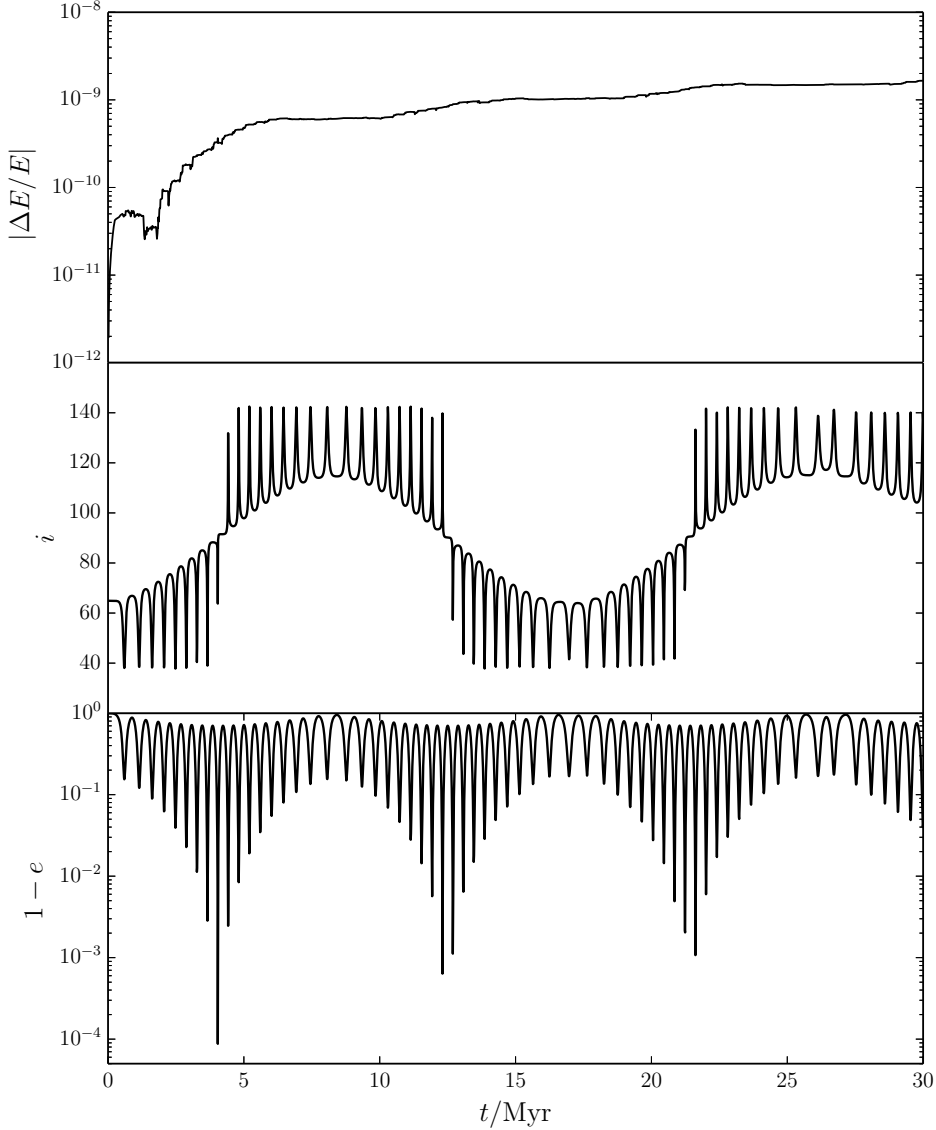
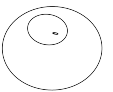


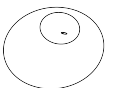
Figure 4.16: Test of the `SECULARQUADRUPLE` algorithm for a hierarchical three-body system; the parameters are set to mimic the system of fig. 3 of Naoz et al. (2013a). The inner binary consists of a star of mass $1 M_{\odot}$ and a planet of mass $1 M_J$ with semimajor axis $a_{\text{in}} = 6$ AU and initial eccentricity $e_{\text{in}} = 0.001$. The outer object is a brown dwarf with mass $40 M_J$ and the outer binary has semimajor axis $a_{\text{out}} = 100$ AU and initial eccentricity $e_{\text{out}} = 0.6$. The binary orbits are initially inclined by 65° .



5 | A triple origin for the lack of tight coplanar circumbinary planets around short-period binaries

Transiting circumbinary planets are more easily detected around short-period than long-period binaries, but none have yet been observed by *Kepler* orbiting binaries with periods shorter than seven days. In triple systems, secular Kozai-Lidov cycles and tidal friction (KLCTF) have been shown to reduce the inner orbital period from $\sim 10^4$ to a few days. Indeed, the majority of short-period binaries are observed to possess a third stellar companion. Using secular evolution analysis and population synthesis, we show that KLCTF makes it unlikely for circumbinary transiting planets to exist around short-period binaries. We find the following outcomes. (1) Sufficiently massive planets in tight and/or coplanar orbits around the inner binary can quench the KL evolution because they induce precession in the inner binary. The KLCTF process does not take place, preventing the formation of a short-period binary. (2) Secular evolution is not quenched and it drives the planetary orbit into a high eccentricity, giving rise to an unstable configuration, in which the planet is most likely ejected from the system. (3) Secular evolution is not quenched but the planet survives the KLCTF evolution. Its orbit is likely to be much wider than the currently observed inner binary orbit, and is likely to be eccentric and inclined with respect to the inner binary. These outcomes lead to two main conclusions: (1) it is unlikely to find a massive planet on a tight and coplanar orbit around a short-period binary, and (2) the properties of circumbinary planets in short-period binaries are constrained by secular evolution.

Hamers, A.S., Perets, H.B. & Portegies Zwart, S.F. MNRAS 445, 3, 3180-3200 (2016)



5.1 Introduction

An increasing number of transiting circumbinary planets around solar-type main-sequence (MS) binary stars are being discovered by the *Kepler* mission. Currently, 10 such planets are known (see Table 5.2; top rows). The orbital periods of the stellar binaries in the systems discovered so far have a mean value of 20.4 d; the shortest period is 7.45 d (Kepler 47; Orosz et al. 2012a). In contrast, the *Kepler* eclipsing binaries *without* circumbinary planets typically have shorter orbital periods, with a mean period of 2.8 d. There exists a bias for detecting *more* transiting circumbinary planets around shorter-period binaries because (1) the shorter period results in more transits in a given amount of time, and (2) the binary precession period is only a few years long, resulting in intersecting binary and planet orbits (as seen in the plane of the sky) during the *Kepler* mission (Martin & Triaud 2015). Because of this bias, many circumbinary planets are expected to have been observed around short-period *Kepler* eclipsing binaries. However, so far, none have been found.

If the apparent lack of (nearly) coplanar circumbinary planets around close binaries is intrinsic and not related to (yet unknown) observational bias(es), then this raises the question of its origin. The following two observations suggest that this origin is related to the (MS) evolution of solar-type triple star systems.

- The fraction of tertiary companions to spectroscopic binaries is a strong function of the inner period, increasing from 0.34 for inner periods $P_A > 12$ d to 0.96 for $P_A < 3$ d (Tokovinin et al. 2006). For the *Kepler* eclipsing binaries in particular, the tertiary fraction for inner binaries with periods $\lesssim 3$ d is $\sim 20\%$ (Rappaport et al. 2013; Conroy et al. 2014). The latter studies are limited (by observing time) to triples with outer periods $\lesssim 3$ yr. The complete tertiary fraction for these *Kepler* binaries, i.e. including any outer period, is likely much larger.
- In the survey of Tokovinin (2014a), a peak is found in the (inner) period distribution around 3 d. As shown by Fabrycky & Tremaine (2007); Naoz & Fabrycky (2014), this peak in the period distribution can be explained by the combination of the secular gravitational torque of the tertiary companion and tidal friction in the inner binary. The former, through Kozai-Lidov (KL) oscillations (Lidov 1962; Kozai 1962), can excite the inner binary orbit to high eccentricity. Consequently, tides are much more effective, leading to tidal dissipation and shrinkage of the orbit, and, eventually, to a tight and (nearly) circular orbit. In this scenario, the precursors of the binaries affected by KL cycles with tides have orbital periods ranging roughly between 10 and 10^4 d.

The lower limit of 10 d of precursor binaries is right in the ballpark of the orbital periods of the currently observed *Kepler* systems with transiting circumbinary planets. Also, there are a number of observed wider systems orbited by at least one circumbinary planet; some of them are listed in Table 5.2 (bottom rows). Therefore, a large fraction of these precursor binaries might have a circumbinary planet in a stable orbit in-between the ‘inner’ and ‘outer’ stellar orbits (henceforth, we refer to the inner and outer stellar binaries simply as the ‘inner’ and ‘outer’ binaries; see also Fig. 5.1).

Here, we show that in these stellar triples with a circumbinary planet around the inner binary, the planet can strongly affect the secular orbital evolution of the inner binary compared to the situation without a planet, provided that certain conditions are met (see below). This ‘shielding’

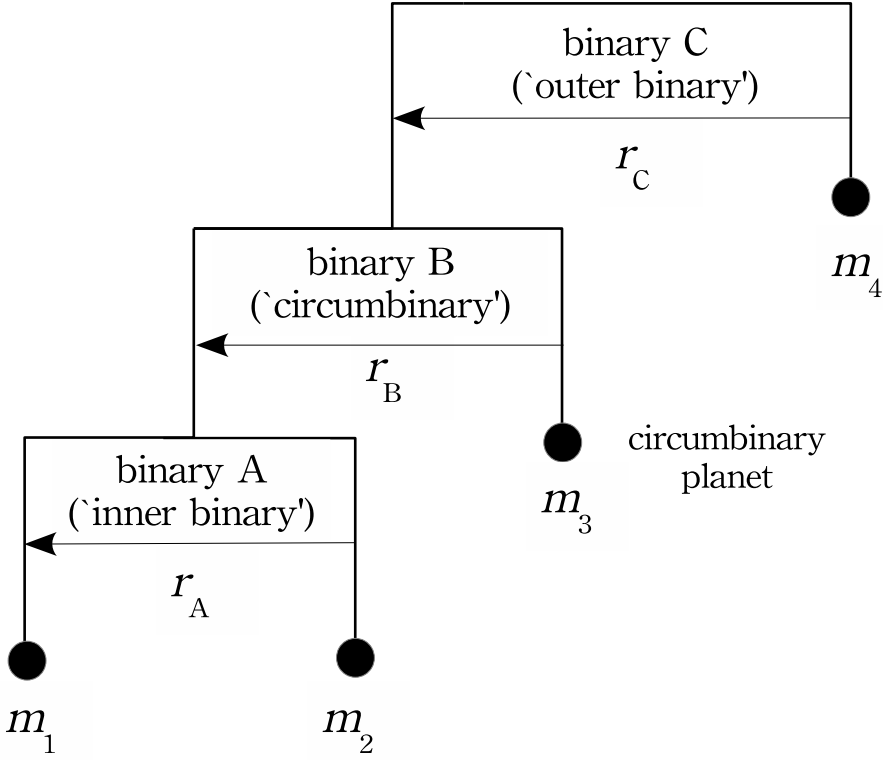
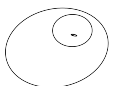


Figure 5.1: A schematic depiction (not to scale) of the hierarchical configuration of the systems considered in this paper. The circumbinary planet (m_3) orbits the inner binary (m_1 and m_2); the outermost star (m_4), in the outer binary, orbits the centre of mass of the inner binary+planet system.

effect arises from a quenching of the eccentricity oscillations in the inner binary induced by the torque of the outer binary, because of precession induced from the circumbinary planet. The shielding effect is similar to the quenching of KL oscillations in stellar triples because of additional sources of orbital precession (notably, precession associated with general relativity, tidal bulges and/or stellar rotation). In some cases, however, resonances can occur that *enhance* the eccentricity oscillations in the inner binary (such resonances can also occur in isolated triples in conjunction with general relativity; Naoz et al. 2013b).

Typically, shielding of KL oscillations by the presence of the planet is only effective for massive planets that are in a (nearly) coplanar and tight orbit with respect to the inner binary. This implies the following two scenarios for the secular evolution of the system.

- If the precursor binary is orbited by a massive planet in a tight and coplanar orbit, then planet shielding can prevent the binary from shrinking through the combined effects of KL cycles and tidal friction (KLCTF).



- On the other hand, if the planet is of low mass and in an inclined and wide orbit with respect to the inner binary, then shielding is typically ineffective, and the binary can shrink through KLCTF to become a short-period binary.

Assuming that short-period binaries are produced through KLCTF, then this implies an intrinsic lack of massive planets in tight and coplanar circumbinary orbits (i.e. strongly shielding planets), whereas far-away and inclined low-mass circumbinary planets (i.e. weakly shielding planets) could still be abundant. This is consistent with the current null-detections of transiting *Kepler* circumbinary planets around short-period binaries.

The goal of this paper is to quantify the above argument. The organization is as follows. In Section 5.2, we briefly describe our methods and assumptions. In Section 5.3, we study the secular gravitational dynamics of stellar triples with a circumbinary planet, and we quantify the conditions when the planet can affect the inner binary. In order to evaluate the effect of circumbinary planets in a population of triples in the field, we carry out a population synthesis study in Section 5.4. We discuss our results in Section 5.5, where we also describe, using direct N -body integrations, the fate of circumbinary planets that become unstable because of secular evolution. We conclude in Section 5.6.

We remark that nearing the completion of this paper, we became aware of two similar and independent studies on the lack of circumbinary planets around short-period binaries in stellar triples by Martin et al. (2015) and Muñoz & Lai (2015).

5.2 Methods and assumptions

We model the system of a circumbinary planet orbiting the inner binary in a stellar triple system as a hierarchical quadruple system in the ‘triple-single’ configuration, which we studied previously in Hamers et al. (2015) (hereafter HPAPZ15). In the latter work, the Hamiltonian of the system was derived and expanded in terms of the ratios of the three binary separations r_A , r_B and r_C , where, by assumption, $r_C \gg r_B \gg r_A$. In the current context, binary A corresponds to the inner binary, binary B to the orbit of the planet around the inner binary, and binary C to the orbit of the tertiary star around the centre of mass of the inner binary+planet system (evidently, the latter nearly coincides with the centre of mass of the inner binary). For consistency with HPAPZ15, we use indices 1 and 2 to denote quantities associated with the inner binary primary and secondary, respectively, index 3 for the planet, and index 4 for the tertiary star.

A schematic depiction of our configuration is shown in Fig. 5.1.

In HPAPZ15, the orbit-averaged Hamiltonian was derived, and a numerical algorithm was developed within the *AMUSE* framework (Portegies Zwart et al. 2013; Pelupessy et al. 2013) to solve the resulting equations of motion. Post-Newtonian (PN) dynamics at the 1PN order are taken into account as described by equation (9) of HPAPZ15.

Here, we extended this algorithm by also including the effects of tidal friction within the inner binary. Gravitational perturbations by the planet and the tertiary star with regards to the inner binary tidal evolution were ignored. We adopted the equilibrium tide model (Eggleton et al. 1998), in which it is assumed that each star quasi-hydrostatically adjusts its shape to the time-varying perturbing potential of its companion. Following Barker & Ogilvie (2009), we adopted a constant tidal quality factor Q related to the mean motion n and the tidal lag time τ via $Q = 1/(n\tau)$. Below, instead of Q , we use the directly related quantity $Q' \equiv 3Q/(2k)$ (Barker & Ogilvie 2009), where k is the second-order potential Love number. Typical values

of Q' for solar-type stars, as inferred from observations, are $Q' \sim 5.5 \times 10^5 - 2 \times 10^6$ (Meibom & Mathieu 2005; Ogilvie & Lin 2007).

A constant Q implies that $\tau = 1/(nQ) = P/(2\pi Q)$ scales with the orbital period, and, therefore, τ effectively decreases as the orbital period decreases due to tidal friction. This may not give an entirely accurate description; e.g. Socrates & Katz (2012) and Socrates et al. (2012) show that with a number of simplifying assumptions, τ is constant, and the results of Hansen (2010) suggest that Q' is not constant and dependent on the semimajor axis (cf. equation 13 of Hansen 2010). Nevertheless, as argued by Barker & Ogilvie (2009), given the current uncertainties in the underlying tidal dissipation mechanisms and, therefore, the efficiency of tidal dissipation, the assumption of a constant Q is useful for studying the general effects of tidal friction.

The equilibrium tide model is included in our algorithm by adding the relevant terms $d\mathbf{e}_A/dt|_{\text{TF}}$, $d\mathbf{h}_A/dt|_{\text{TF}}$, $d\mathbf{\Omega}_1/dt|_{\text{TF}}$ and $d\mathbf{\Omega}_2/dt|_{\text{TF}}$ in the equations of motion, as given by equations (A7)-(A15) of Barker & Ogilvie (2009). Here, $\mathbf{e}_k$ and $\mathbf{h}_k$ are the eccentricity vector and the specific angular momentum vector of orbit k , respectively, and $\mathbf{\Omega}_k$ is the spin angular momentum vector of star $k \in \{1, 2\}$. Throughout this paper, the initial spin vectors $\mathbf{\Omega}_k$ are assumed to be parallel with the inner orbital angular momentum vector $\mathbf{h}_A$.

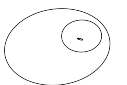
Our algorithm, that models the orbital evolution, is coupled within `AMUSE` with the stellar evolution code `SEBA` (Portegies Zwart & Verbunt 1996; Toonen et al. 2012), which is also interfaced within `AMUSE`. We use the latter code to compute the masses and radii during the integration. In this work, we focus on low-mass MS stars. Therefore, the time-dependence of the latter quantities is weak for time-scales less than a Hubble time. We remark, however, that this is no longer the case for post-MS evolution.

5.3 The planet-shielding effect

Before studying the effect of circumbinary planets in a population of triples (cf. Section 5.4), we first focus on a simplified case where the stellar triple system is kept fixed, and planets are ‘inserted’ at stable orbits at various loci in-between the inner and outer binaries. The aim is to gain quantitative insight into the effect of the planetary orbit on the eccentricity oscillations and the tidal evolution in the inner binary. For simplicity, we here focus on triples in which the octupole-order terms are unimportant. Note, however, that the (non-cross) octupole-order terms are always included in the integrations (cf. Section 5.5.5), and no restriction is made with respect to the importance on the octupole-order terms in Section 5.4.

5.3.1 *Kepler* transiting circumbinary systems

Here, we focus on systems similar to the *Kepler* transiting circumbinary systems. The inner semimajor axes in the latter are $a_A \sim 0.1 - 0.2$ AU (cf. the top rows of Table 5.2), and relativistic precession is therefore important in these systems. As is well known from previous studies (e.g., Holman et al. 1997; Blaes et al. 2002; Naoz et al. 2013b), general relativistic precession limits the range of semimajor axes and eccentricities of tertiary orbits for which the inner orbit eccentricity is excited. In general, an order-of-magnitude estimate of the KL time-scale for the



Section	5.3.1	5.3.2	5.5.3	5.4
$m_1/M_\odot$	1	1	0.83	0.5-1.2 ^a
$m_2/M_\odot$	0.5	0.5	0.50	0.1-1.2 ^b
m_3/M_J	0.01-1	0.001-3.16	0.01-1	0.01-1
$m_4/M_\odot$	1	1	0.73	0.1-1.2 ^c
a_A/AU	0.2	1	0.20	0.03-2.0 ^d
a_B/AU	1-5	10-50	1.09 - 1.22	0.14- 3.1×10^2 ^e
a_C/AU	20	160	7.73	1.5- 1.8×10^4 ^d
e_A	0.01	0.01	0.40	0.01-0.95 ^f
e_B	0.01	0.01	0.01	0.01
e_C	0.01	0.01	0.30	0.01-0.95 ^f
$i_A/^\circ$	0.01	0.01	0.01	0.01
$i_B/^\circ$	0.01-180	0.01-180	0.01-180	0.01-180
$i_C/^\circ$	75	85	73.7	0.01-180 ^g
$\omega_A/^\circ$	0.01	0.01	287.7	0.01-180 ^h
$\omega_B/^\circ$	0.01	0.01	281.0	0.01-180 ^h
$\omega_C/^\circ$	0.01	0.01	51.6	0.01-180 ^h
$\Omega_A/^\circ$	0.01	0.01	166.1	0.01-180 ^h
$\Omega_B/^\circ$	0.01	0.01	42.6	0.01-180 ^h
$\Omega_C/^\circ$	0.01	0.01	340.1	0.01-180 ^h
$P_{\text{spin},k}/\text{d}$	N/A	10	10	10
$Q'_k/10^6$	N/A	1	1.9	0.55-2 ^h
$r_{g,k}$	N/A	0.08	0.08	0.08

^a Salpeter distribution $dN/dm_1 \propto m_1^{-2.35}$.

^b Sampled from $m_2 = q_{\text{in}} m_1$, where $0 < q_{\text{in}} < 1$ is linearly distributed.

^c Sampled from $m_4 = q_{\text{out}}(m_1 + m_2)$, where $0 < q_{\text{out}} < 1$ is linearly distributed.

^d Lognormal distribution in the orbital period P_k with mean $\log_{10}(P_k/\text{d}) = 5.03$, standard deviation $\sigma_{\log_{10}(P_k/\text{d})} = 2.28$ and range $-2 < \log_{10}(P_k/\text{d}) < 10$. Triples are subject to stability constraints, and the inner binary should not merge in the absence of a planet.

^e Linearly sampled for each triple, with lower limit $a_{B,l} = 1.5 a_{B,\text{crit},AB}$, where $a_{B,\text{crit},AB}$ is the critical semimajor axis for dynamical stability of the planet in a coplanar orbit around the inner binary, and which is adopted from the fitting formula given by (Holman & Wiegert 1999). The upper limit is $a_{B,u} = 0.9 a_{B,\text{crit},BC}$, where $a_{B,\text{crit},BC}$ is the largest possible value of a_B for dynamical stability with respect to the orbit of the tertiary star, estimated by applying the Mardling & Aarseth 2001 criterion to the BC pair, with an ‘outer’ mass ratio of $q_{\text{out}} = m_4/(m_1 + m_2 + m_3)$.

^f Sampled from a Rayleigh distribution, $dN/de_k \propto e_k \exp(-\beta e_k^2)$, with rms $\langle e_k^2 \rangle^{1/2} = \beta^{-1/2} = 0.33$ (Raghavan et al. 2010).

^g Sampled from a linear distribution in $\cos(i_C)$, with $-1 < \cos(i_C) < 1$.

^h Sampled from a linear distribution.

Table 5.1: Initial conditions for the systems in Section 5.3.1 (first column, cf. Figs 5.2-5.4), for the systems in Section 5.3.2 (second column, cf. Figs 5.5 and 5.6), for the systems in Section 5.5.3 (third column, cf. Fig. 5.20), and for the population synthesis in Section 5.4 (fourth column). Note that for the third column, the systems are integrated with the secular code until the orbit of the planet intersects with the inner binary; subsequent evolution is carried out using direct N -body integration (cf. Section 5.5.3). The orbital angles (inclinations i_k , arguments of pericentre ω_k and longitudes of the ascending node Ω_k) are defined with respect to an arbitrary, but fixed reference frame; see equation (6) of HPAPZ15 for the relation between these angles and the orbital vectors. The spin periods $P_{\text{spin},k}$, tidal quality factors Q'_k and radii of gyration $r_{g,k}$ apply to the inner binary stars and are used for tidal evolution.

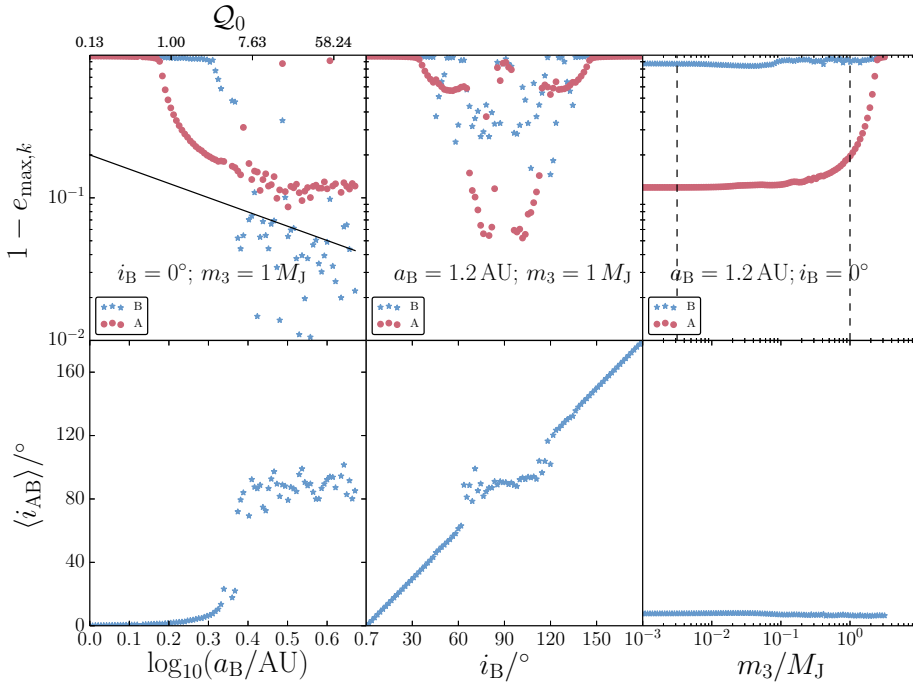
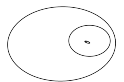


Figure 5.2: Top row: maximum eccentricities in the inner binary (red bullets) and circumbinary orbit (blue stars) for the integrations of Section 5.3.1 (cf. the left column of Table 5.1), as a function of the initial a_B (left column), i_B (middle column) and m_3 (right column). In the left column $i_B = 0^\circ$ and $m_3 = 1 M_J$; in the middle column $a_B = 1.2 \text{ AU}$ and $m_3 = 1 M_J$; and in the right column $a_B = 1.2 \text{ AU}$ and $i_B = 0^\circ$. The corresponding value of Q_0 (cf. equation 5.8) is indicated in the top left panel. In the bottom row, the average inclination between the inner and circumbinary orbits is shown. The two black vertical dashed lines in the top right panel indicate Earth and Jupiter masses, respectively.



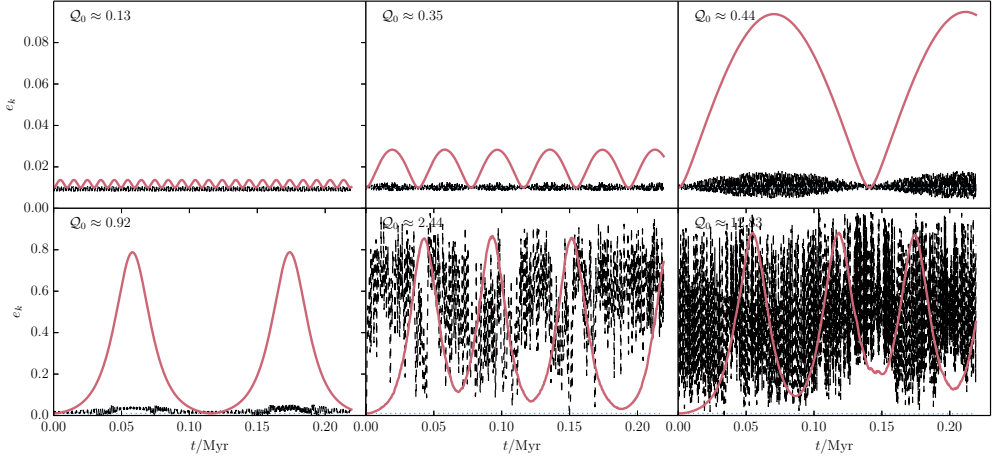


Figure 5.3: Illustration of the time evolution of the eccentricities of the inner orbit and the circumbinary, for six systems corresponding to the top-left panel of Fig. 5.2. In each panel, the value of Q_0 is indicated.

binary pair kl is given by

$$P_{\text{KL},kl} = \frac{P_l^2}{P_k} \frac{m_{k,p} + m_{k,s} + m_{l,s}}{m_{l,s}} (1 - e_l^2)^{3/2}, \quad (5.1)$$

where $m_{k,p} = m_1$, $m_{k,s} = m_2$ and $m_{l,s} = m_3$ in the case of $P_{\text{KL},\text{AB}}$, $m_{k,p} = m_1 + m_2$, $m_{k,s} = m_3$ and $m_{l,s} = m_4$ in the case of $P_{\text{KL},\text{BC}}$, and $m_{k,p} = m_1$, $m_{k,s} = m_2$ and $m_{l,s} = m_4$ in the case of $P_{\text{KL},\text{AC}}$ (Innanen et al. 1997, HPAPZ15, see also Antognini 2015). The (pairwise) 1PN time-scale for orbit k is given by

$$t_{\text{1PN},k} = \frac{1}{3} P_k (1 - e_k^2) \frac{a_k}{r_{g,k}}, \quad (5.2)$$

where $r_{g,k} \equiv Gm_{\text{tot},k}/c^2$, with $m_{\text{tot},\text{A}} = m_1 + m_2$, $m_{\text{tot},\text{B}} = m_1 + m_2 + m_3$ and $m_{\text{tot},\text{C}} = m_1 + m_2 + m_3 + m_4$, is the gravitational radius. Equating equation (5.1) as applied to the AC pair, to equation (5.2) as applied to binary A, we find that relativistic precession dominates in binary A if

$$a_C > (1 - e_C^2)^{-1/2} \left[\frac{a_A^4 c^2 (1 - e_A^2) m_4 (m_1 + m_2 + m_3 + m_4)}{3G(m_1 + m_2)^2 (m_1 + m_2 + m_4)} \right]^{1/3} \\ \approx 34 \text{ AU}, \quad (5.3)$$

where for the numerical estimate we assumed $e_A = e_C = 0.01$, $m_1 = 1 M_\odot$, $m_2 = 0.5 M_\odot$, $m_3 = 1 M_J$ and $m_4 = 1 M_\odot$. In the remainder of Section 5.3.1, we shall assume these values for the numerical estimates.

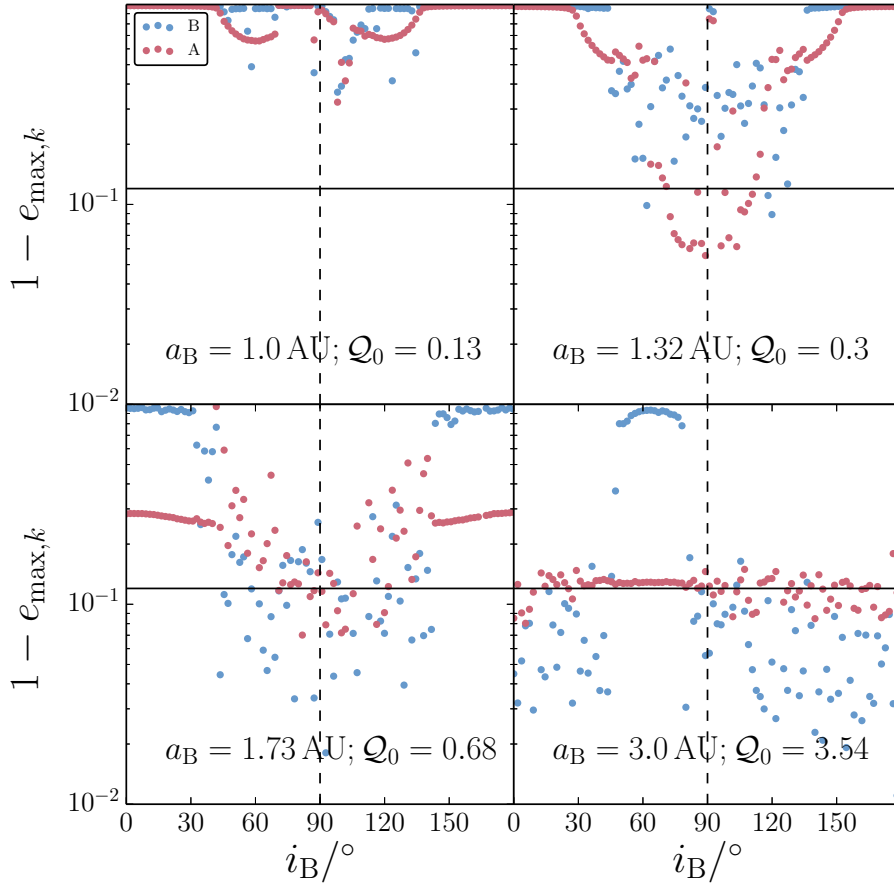
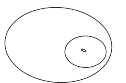


Figure 5.4: Illustration of the joint dependence of the maximum eccentricities in binaries A and B on a_A and i_B , as an extension of the top-middle panel of Fig. 5.2. In each panel, $1 - e_{\max,k}$ is plotted as a function of i_B . The value of a_B is indicated for each panel; for all panels, $m_3 = 1 M_J$. The black solid horizontal lines show the value of $1 - e_{\max,A}$ in the absence of the planet.



An estimate of the lower limit on a_C for dynamical stability with respect to the inner binary (not yet taking into account the circumbinary planet, and ignoring the dependence on mutual inclination) is given by (Mardling & Aarseth 2001)

$$a_C > \frac{2.8a_A}{1-e_C} \left[\left(1 + \frac{m_4}{m_1+m_2} \right) \frac{1+e_C}{\sqrt{1-e_C}} \right]^{2/5} \approx 0.67 \text{ AU}. \quad (5.4)$$

Here, we assume a fixed $a_C = 20 \text{ AU}$. With this choice, the octupole parameter $\epsilon_{\text{oct,AC}}$, defined as

$$\epsilon_{\text{oct,AC}} \equiv \frac{m_1 - m_2}{m_1 + m_2} \frac{a_A}{a_C} \frac{e_C}{1 - e_C^2}, \quad (5.5)$$

is $\epsilon_{\text{oct,AC}} \approx 1.7 \times 10^{-5}$, indicating that octupole-order terms in the Hamiltonian expansion for the AC pair are unimportant (Lithwick & Naoz 2011; Katz et al. 2011; Naoz et al. 2013a; Teyssandier et al. 2013; Li et al. 2014b). Assuming an initial mutual inclination of $i_{AC} = 75^\circ$ between binaries A and C, and with the inclusion of the 1PN terms, the maximum eccentricity attained in binary A, *without* a circumbinary planet, is $e_{A,\text{max}} \approx 0.88$.

As a next step, we include a circumbinary planet with mass m_3 in orbit around binary A with semimajor axis $a_B < a_C$, and with various mutual inclinations i_{AB} . For simplicity, the circumbinary orbit is assumed to be initially nearly circular, i.e. $e_B = 0.01$. Assuming $i_{AB} = 0^\circ$, for dynamical stability with respect to the inner binary, the semimajor axis a_B must satisfy (Holman & Wiegert 1999)

$$a_B > a_A f(e_A, \mu_A) \approx 0.48 \text{ AU}, \quad (5.6)$$

where $f(e_A, \mu_A)$ is a function of the inner binary eccentricity e_A and mass ratio $\mu_A \equiv m_2/(m_1 + m_2)$, given by equation (3) of Holman & Wiegert (1999). An estimate of the largest possible value of a_B for dynamical stability with respect to the outer orbit, binary C, ignoring dependencies on inclinations, is given by (Mardling & Aarseth 2001)

$$a_B < a_C \frac{1-e_C}{2.8} \left[\left(1 + \frac{m_4}{m_1+m_2+m_3} \right) \frac{1+e_C}{\sqrt{1-e_C}} \right]^{-2/5} \approx 5.8 \text{ AU}. \quad (5.7)$$

Note that, strictly speaking, the criterion of (Mardling & Aarseth 2001) only applies to three-body systems. However, by carrying out a number of N -body integrations with *MIKKOLA* (Mikkola & Merritt 2008) and *SAKURA* (Gonçalves Ferrari et al. 2014) within *AMUSE*, we find that circular circumbinary orbits with $a_B = 5.8 \text{ AU}$ in the system of our choice are indeed dynamically stable on time-scales of a few multiple periods of the outer binary.

Based on these estimates, we carried out integrations of the long-term secular dynamical evolution with our orbit-averaged algorithm (cf. Section 5.2), with a_B ranging between 1 and 5 AU, i_{AB} between 0 and 180° , and m_3 between 10^{-2} and $10^0 M_J$. To simplify the interpretation, tidal evolution is not included in Section 5.3.1. The integration time was set to $2P_{\text{KL,AC}}$, i.e. the time-scale for two KL oscillations in the AC pair, without taking into account the effect of the planet. A comprehensive list of initial parameters is given in the left column of Table 5.1.

In the left, middle and right panels of the first row of Fig. 5.2, we show the resulting maximum eccentricities in binaries A (red bullets) and B (blue stars) as a function of a_B (with

$i_B = 0^\circ$ and $m_3 = 1 M_J$), i_B (with $a_B = 1.2$ AU and $m_3 = 1 M_J$), and m_3 (with $a_B = 1.2$ AU and $i_B = 0^\circ$), respectively. In the bottom row of the same figure, we show the average inclinations between orbits AB as a function of a_B , i_B and m_3 . In the top left panel, we also show an equivalent dependence on the parameter Q_0 , defined as the initial ratio of the KL time-scales for the AB and AC pairs,

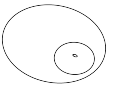
$$\begin{aligned}
 Q_0 &\equiv \frac{P_{\text{KL,AB},0}}{P_{\text{KL,AC},0}} \\
 &= \left(\frac{a_B}{a_C}\right)^3 \frac{m_4}{m_3} \frac{m_1 + m_2 + m_3 + m_4}{m_1 + m_2 + m_4} \left(\frac{1 - e_{B,0}^2}{1 - e_{C,0}^2}\right)^{3/2} \\
 &\approx \left(\frac{a_B}{a_C}\right)^3 \frac{m_4}{m_3} \left(\frac{1 - e_{B,0}^2}{1 - e_{C,0}^2}\right)^{3/2} \\
 &\approx 0.13 \left(\frac{a_B}{1 \text{ AU}}\right)^3.
 \end{aligned} \tag{5.8}$$

Here, in the third line, we neglected the planet mass compared to the sum of all stellar masses. We note that Q_0 is closely related to the quantity $\mathcal{R}_0$ defined by HPAPZ15, and is given by

$$\begin{aligned}
 \mathcal{R}_0 &\equiv \frac{P_{\text{KL,AB},0}}{P_{\text{KL,BC},0}} \\
 &= \left(\frac{a_B^3}{a_A a_C^2}\right)^{3/2} \left(\frac{m_1 + m_2}{m_1 + m_2 + m_3}\right)^{1/2} \frac{m_4}{m_3} \left(\frac{1 - e_{B,0}^2}{1 - e_{C,0}^2}\right)^{3/2} \\
 &\approx \left(\frac{a_B^3}{a_A a_C^2}\right)^{3/2} \frac{m_4}{m_3} \left(\frac{1 - e_{B,0}^2}{1 - e_{C,0}^2}\right)^{3/2} \\
 &\approx Q_0 \left(\frac{a_B}{a_A}\right)^{3/2}.
 \end{aligned} \tag{5.9}$$

If $i_B = 0^\circ$ and $Q_0 \gg 1$, $e_{\text{max,A}} \approx 0.88$ is the same as in the case without the planet (cf. the top-left panel of Fig. 5.2). In this limit, the torque of the outer orbit on the inner orbit (for which $P_{\text{KL,AC}}^{-1}$ is a proxy) is much larger than the torque of the circumbinary orbit on the inner orbit (for which $P_{\text{KL,AB}}^{-1} \ll P_{\text{KL,AC}}^{-1}$ is a proxy). The circumbinary orbit therefore only has a small perturbative effect on the inner orbit, and the maximum eccentricity is not affected. If $Q_0 \gg 1$, then also $\mathcal{R}_0 \gg 1$ (cf. equation 5.9). This implies that the torque of binary B is not strong enough to maintain coplanarity between binaries A and B, as is shown in the bottom-left panel of Fig. 5.2. Effectively, binary B decouples from binary A, and the torque of the outer orbit results in high-amplitude eccentricity oscillations in binary B.

The latter oscillations have large enough eccentricities for the circumbinary orbit to intersect with the inner binary, i.e. $r_{\text{p,B}} = a_B(1 - e_B) < r_{\text{a,A}} = a_A(1 + e_A)$. This is the case for all points below the black solid line in the top-left panel of Fig. 5.2. For the purposes of this section, no stopping conditions were imposed on the eccentricities during the integrations. However, in reality, the circumbinary orbit is dynamically unstable for points below the black solid line. This is borne out by integrating some of the systems below this line with direct N -body integration,



where the orbital parameters are set to correspond to the moment of maximum eccentricity of binary B. We investigate the possible outcomes for these cases of dynamical instability further in Section 5.5.3.

On the other hand, if $i_B = 0^\circ$ and $\mathcal{Q}_0 \ll 1$, $e_{\max,A} \approx e_{A,0}$. In this case, the net torque on the inner orbit is dominated by the circumbinary orbit. Consequently, the inner orbit precesses much more rapidly compared to the case without the planet, and this results in a quenching of the KL eccentricity oscillations otherwise induced by the outer orbit. Furthermore, the period of the latter oscillations is substantially reduced, as is illustrated in Fig. 5.3, where we show the time evolution of e_A and e_B for several values of $\mathcal{Q}_0$, assuming $i_B = 0^\circ$ and $m_3 = 1 M_J$. Note that the inclination between orbits A and B remains zero in this case (cf. the bottom-left panel of Fig. 5.2), and that octupole-order terms are not dominant, even for our smallest value $a_B = 1 \text{ AU}$ (for $a_B = 1 \text{ AU}$, $\epsilon_{\text{oct},AB} \approx 6.7 \times 10^{-4}$). Therefore, the circumbinary orbit, on its own, does not induce eccentricity oscillations in binary A.

Whenever the circumbinary orbit becomes unstable, the planet could collide with one of the stars, or be ejected from the system (see also Section 5.5.3). Evidently, in either case, the planet can no longer shield the inner binary from KL-eccentricity oscillations induced by the outer orbit. Typically, this occurs whenever the circumbinary orbit is relatively close to the outer orbit ($\mathcal{Q}_0 \gtrsim 1$), and is therefore strongly affected by the latter orbit's torque. In the same regime, however, shielding is ineffective. This implies that dynamical instability does not necessarily rule out the possibility for shielding by the planet. In the top-left panel of Fig. 5.2, this is indeed the case: the circumbinary orbits become unstable for $\mathcal{Q}_0 \gtrsim 1$, whereas shielding only occurs for $\mathcal{Q}_0 \lesssim 1$.

The dependence of the maximum eccentricities on i_B is shown in the top middle panel of Fig. 5.2 for $a_B = 1.2 \text{ AU}$. There is a strong dependence of the planet's shielding ability on i_B . For coplanar (either prograde or retrograde) orbits, $e_{A,\max}$ is close to the initial value $e_{A,0}$. As the initial orbit of the planet becomes more inclined with respect to the inner binary, shielding becomes less efficient. Interestingly, for i_B close to 90° , shielding is again *more* efficient.

The nature of the dependence on i_B also strongly depends on a_B . In Fig. 5.4, we show the dependence on i_B for other values of a_B , assuming $m_3 = 1 M_J$. Generally, $e_{\max,A}$ is weakly dependent on i_B for either small a_B or $\mathcal{Q}_0 \ll 1$ (in which case $e_{\max,A} \approx e_{A,0}$), or large a_B or $\mathcal{Q}_0 \gg 1$ (in which case $e_{\max,A}$ is approximately equal to the value in absence of the planet). In the intermediate regime, there is a complicated dependence on i_B . Generally, quenching of the KL eccentricity oscillations in binary A is most efficient for close to coplanar orbits (in the case of Fig. 5.4, $0^\circ \lesssim i_B \lesssim 30^\circ$ and $150^\circ \lesssim i_B \lesssim 180^\circ$). Furthermore, for some values of a_B , quenching can also be effective for $i_B \approx 90^\circ$ (cf. the top middle panel of Fig. 5.2 and the top left panel of Fig. 5.4).

The dependence on m_3 is illustrated in the top-right panel of Fig. 5.2 (for $a_B = 1.2 \text{ AU}$ and $i_B = 0^\circ$). The planet mass must be large enough for shielding to be effective, which is intuitively easy to understand. In our example, a Jupiter-mass planet can effectively shield the inner binary, assuming that it is coplanar with and close to the inner binary, whereas an Earth-mass planet cannot provide shielding, regardless of its orbit.

5.3.2 A triple with a shrinking inner binary orbit

Having considered in Section 5.3.1 the planet's ability to shield the inner binary from KL eccentricity oscillations induced by the outer orbit, we here extend the analysis by also including

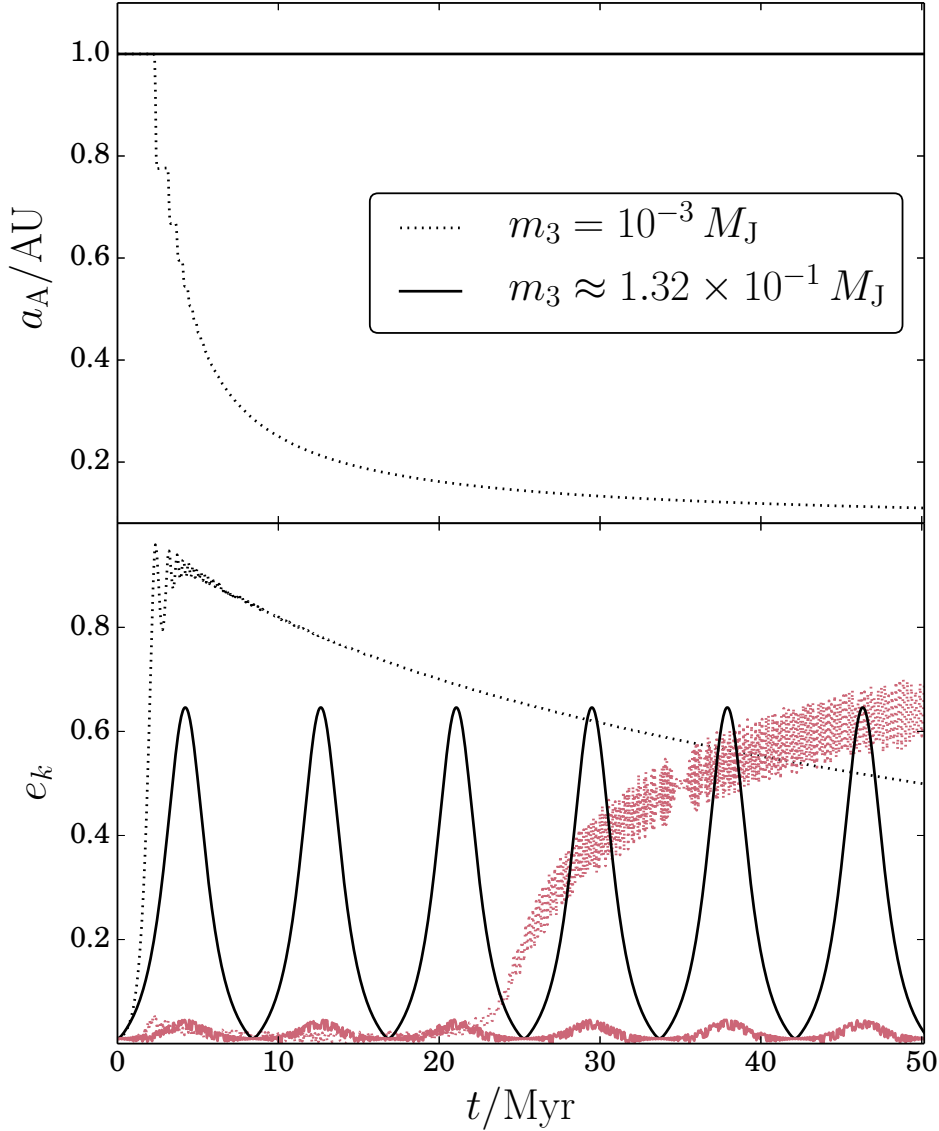
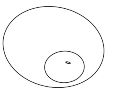


Figure 5.5: Example evolution of the inner binary semimajor axis (top panel), the inner binary eccentricity (bottom panel; black lines), and the circumbinary eccentricity (bottom panel; red lines), for the triple system discussed in Section 5.3.2. Two planet masses are assumed: $m_3 = 10^{-3} M_J$ (solid lines) and $m_3 \approx 1.32 \times 10^{-1} M_J$ (dashed lines). In either cases, $a_B = 40$ AU and $i_B = 0^\circ$ initially.



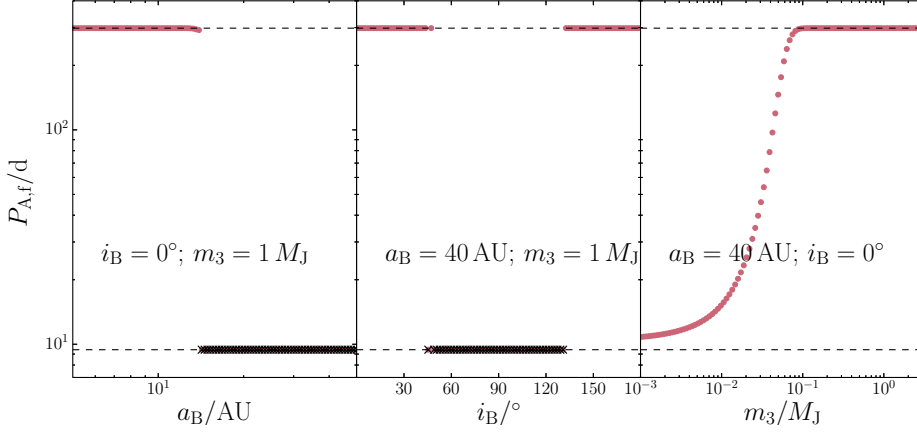


Figure 5.6: The final inner orbital period for the integrations of Section 5.3.2 (cf. the second column of Table 5.1) as a function of the initial a_B (left column), i_B (middle column) and m_3 (right column). In the left column $i_B = 0^\circ$ and $m_3 = 1 M_J$, in the middle column $a_B = 40 \text{ AU}$ and $m_3 = 1 M_J$, and in the right column $a_B = 40 \text{ AU}$ and $i_B = 0^\circ$. Cases where the planet became unstable are denoted with black crosses.

the effects of tidal friction in the inner binary. We choose a triple system with a relatively wide inner binary ($a_A = 1 \text{ AU}$ or $P_A \sim 300 \text{ d}$) and tight outer binary ($a_C = 160 \text{ AU}$), such that in the absence of a planet, the inner binary becomes highly eccentric because of KL eccentricity cycles induced by the tertiary, and shrinks to a tight binary with $P_A \approx 8 \text{ d}$.

As in Section 5.3.1, we carry out a set of integrations with a planet with various values of a_B , i_B and m_3 . A comprehensive list of the initial conditions is given in the second column of Table 5.1. For each system, the integration time is set to $10 P_{\text{KL},AC} \approx 50 \text{ Myr}$. In contrast to Section 5.3, here we stop the integration if the circumbinary orbit intersects with the inner binary, i.e. if $r_{p,B} = a_B(1 - e_B) \leq r_{a,A} = a_A(1 + e_A)$ (neglecting the radii of the stars and the planet). We assume that in the latter case, the planet either collides with one of the stars or is ejected from the system (see also Section 5.5.3), and, therefore, the subsequent evolution is equivalent to the situation without a planet, i.e. the inner binary shrinking to an inner final inner period $P_{A,f} \approx 8 \text{ d}$.

In Fig. 5.5, we show, as an example, the evolution of the inner binary semimajor axis a_A (top panel), the inner binary eccentricity e_A (bottom panel; black lines), and the circumbinary eccentricity e_B (bottom panel; red lines), where $a_B = 40 \text{ AU}$ and $i_B = 0^\circ$ initially. Two planet masses are assumed: $m_3 = 10^{-3} M_J$ (solid lines) and $m_3 \approx 1.32 \times 10^{-1} M_J$ (dashed lines). The values of Q_0 are ≈ 55 and ≈ 0.42 for these values of m_3 , respectively.

For the low planet mass ($Q_0 \approx 55$), shielding is ineffective, and the inner orbit eccentricity becomes highly excited, resulting in efficient shrinkage of the orbit. Interestingly, this also increases the relative strength of the torque of the outer orbit compared to the inner orbit on the circumbinary orbit. Consequently, the eccentricity of the circumbinary orbit becomes excited at $t \sim 25 \text{ Myr}$. Although not the case here, the latter eccentricity could become high enough for

the circumbinary orbit to intersect with the inner binary, and become unstable. In other words, KL cycles with tidal friction tend to destabilize the circumbinary planet, and this further reduces the chances of a circumbinary planet residing in a stable orbit around a binary that shrank due to KL cycles with tidal friction. This consequence of inner binary shrinkage on the planet orbital stability was also noted by Martin et al. (2015) and Muñoz & Lai (2015).

In contrast, for the high planet mass ($Q_0 \approx 0.42$), shielding is effective, and the inner binary eccentricity does not become high enough for efficient tidal friction. Nevertheless, e_A still oscillates with an amplitude of ≈ 0.6 . The circumbinary orbit eccentricity does not become excited in this case.

In the left, middle and right panels of Fig. 5.6, we show the final inner period $P_{A,f}$ as a function of a_B (with $i_B = 0^\circ$ and $m_3 = 1 M_J$), i_B (with $a_B = 40$ AU and $m_3 = 1 M_J$), and m_3 (with $a_B = 40$ AU and $i_B = 0^\circ$), respectively. Cases where the planet became unstable are denoted with black crosses. There is a strong dependence of the final inner period on the circumbinary parameters. If a_B is sufficiently small (the regime $Q_0 \ll 1$), shielding is effective, and the inner binary does not shrink. For larger a_B , the planet becomes unstable, and is therefore unable to shield the inner binary. Note that in this regime of $Q_0 \gtrsim 1$, shielding would have been ineffective even if the circumbinary orbit would be stable (cf. Fig. 5.2).

Similarly, there is a strong dependence on i_B . For coplanar (either prograde or retrograde) circumbinary orbits, shielding can be effective (if a_B is small enough) and the inner binary does not shrink. However, for highly inclined orbits, the planetary orbit is unstable and, subsequently, the inner binary shrinks. The boundaries between these regimes, $i_B \approx 45^\circ$ and $i_B \approx 135^\circ$, are very similar to the critical values for eccentricity excitation due to the KL mechanisms (assuming quadrupole order). We note that, for much more compact circumbinary planets with a small ratio a_B/a_A , instability can occur for even coplanar systems (Li et al. 2014a).

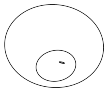
For the value of a_B and i_B in the right panel of Fig. 5.6, the planetary orbit is stable for all assumed masses m_3 . Effective shielding, and hence shrinkage of the inner orbit, can only occur if the planet is sufficiently massive, in this case, if $m_3 \gtrsim 0.1 M_J$.

5.4 Planets in triples with short-period inner binaries

5.4.1 Initial conditions

In Section 5.3, we demonstrated the planet's ability to shield the inner binary assuming a fixed triple system, and used specific worked-out examples to understand the evolution of circumbinary planets affected by a third companion. In the following, we study the effect of circumbinary planets in the more general case, and consider a large orbital phase-space through the study of a population of triples, for which the inner binary, in absence of a planet, would be affected by KL cycles with tidal friction.

For stellar triples, a number of surveys of solar-type MS stars (Duquennoy & Mayor 1991; Tokovinin et al. 2006; Raghavan et al. 2010; Tokovinin 2014a) have given insight into the initial distributions of the masses and orbital parameters. However, initial distributions for a circumbinary planet in orbit of the inner binary, for which observations are currently strongly biased, are very poorly constrained. Therefore, we generate initial conditions using a combination of Monte Carlo sampling for the stellar triple system, and fixed grids for the circumbinary planet. Our procedure consists of the following four steps (a succinct summary is given in the



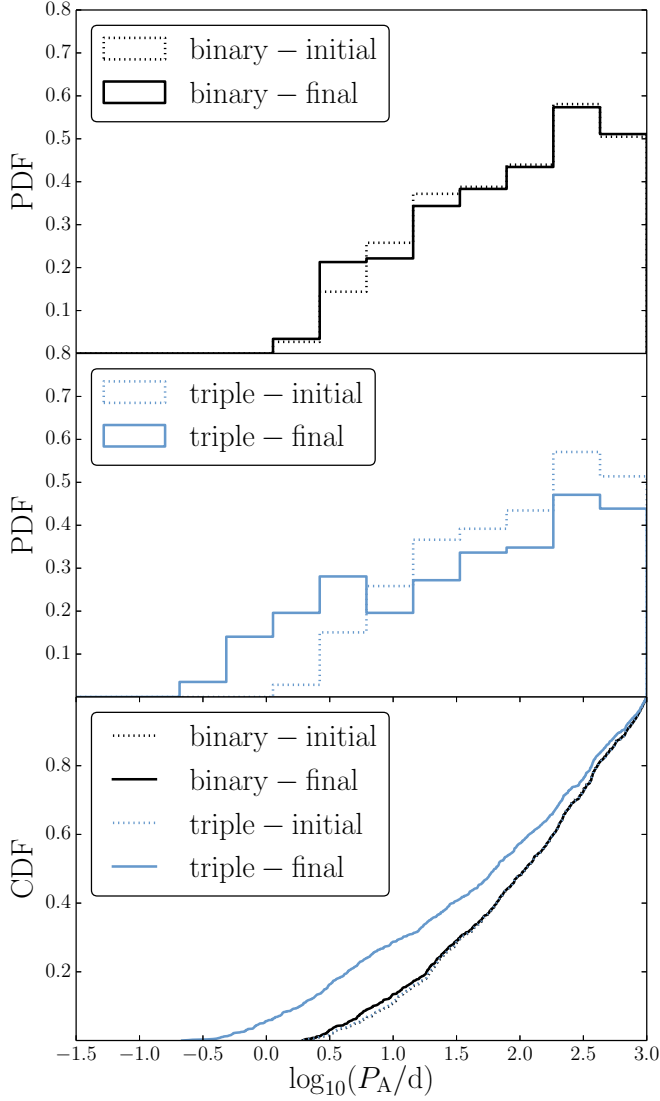


Figure 5.7: The initial (dotted lines) and final (solid lines) inner period distributions for the sampled triples as described in Section 5.4.1. Top row (black lines): isolated binary evolution, i.e. taking into account only tidal friction. Middle row (blue lines): isolated triple evolution, i.e. also taking into account the torque of the outer orbit. Cumulative distributions for both cases are shown in the bottom panel.

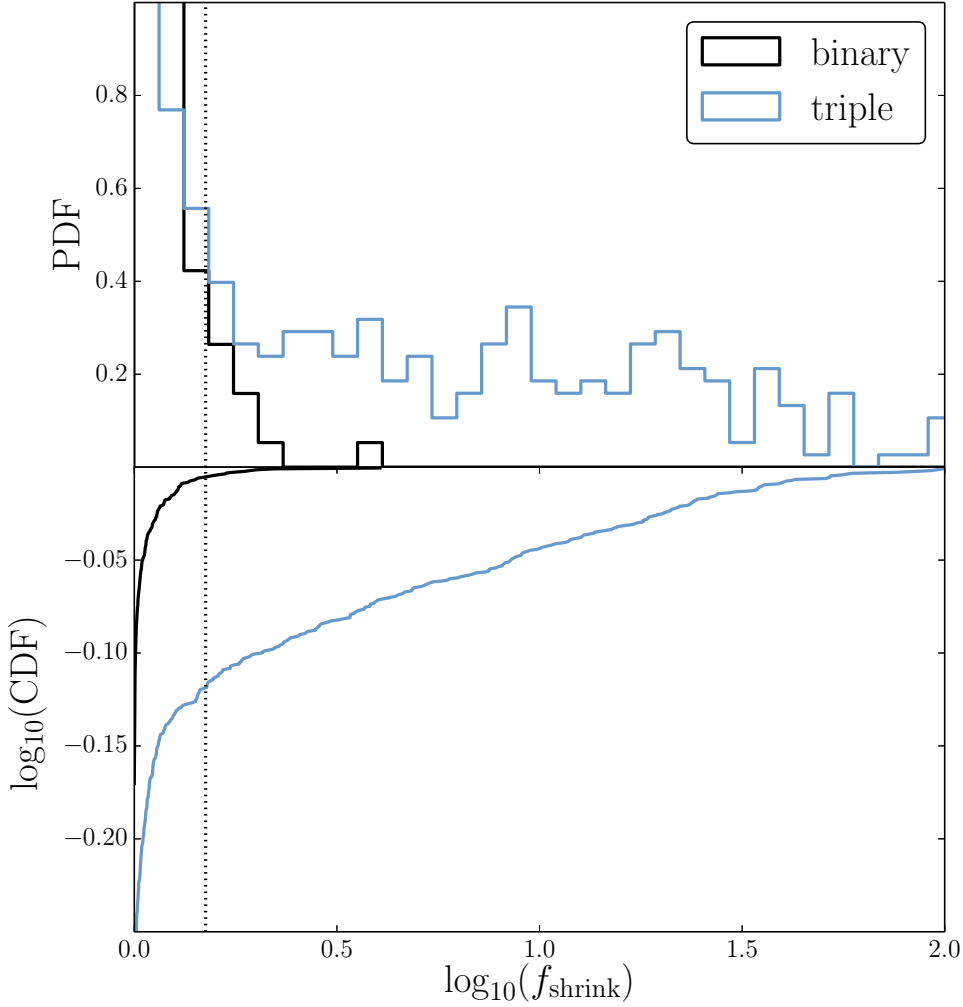
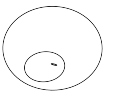


Figure 5.8: The factor $f_{\text{shrink}} \equiv a_{A,i}/a_{A,f}$ with which the inner binary semimajor axis shrinks for the sampled triples as described in Section 5.4.1. Black (blue) lines correspond to isolated binary (triple) evolution. The black vertical dotted line indicates the value of f_{shrink} chosen for the selection of systems with substantially shrinking inner binaries.



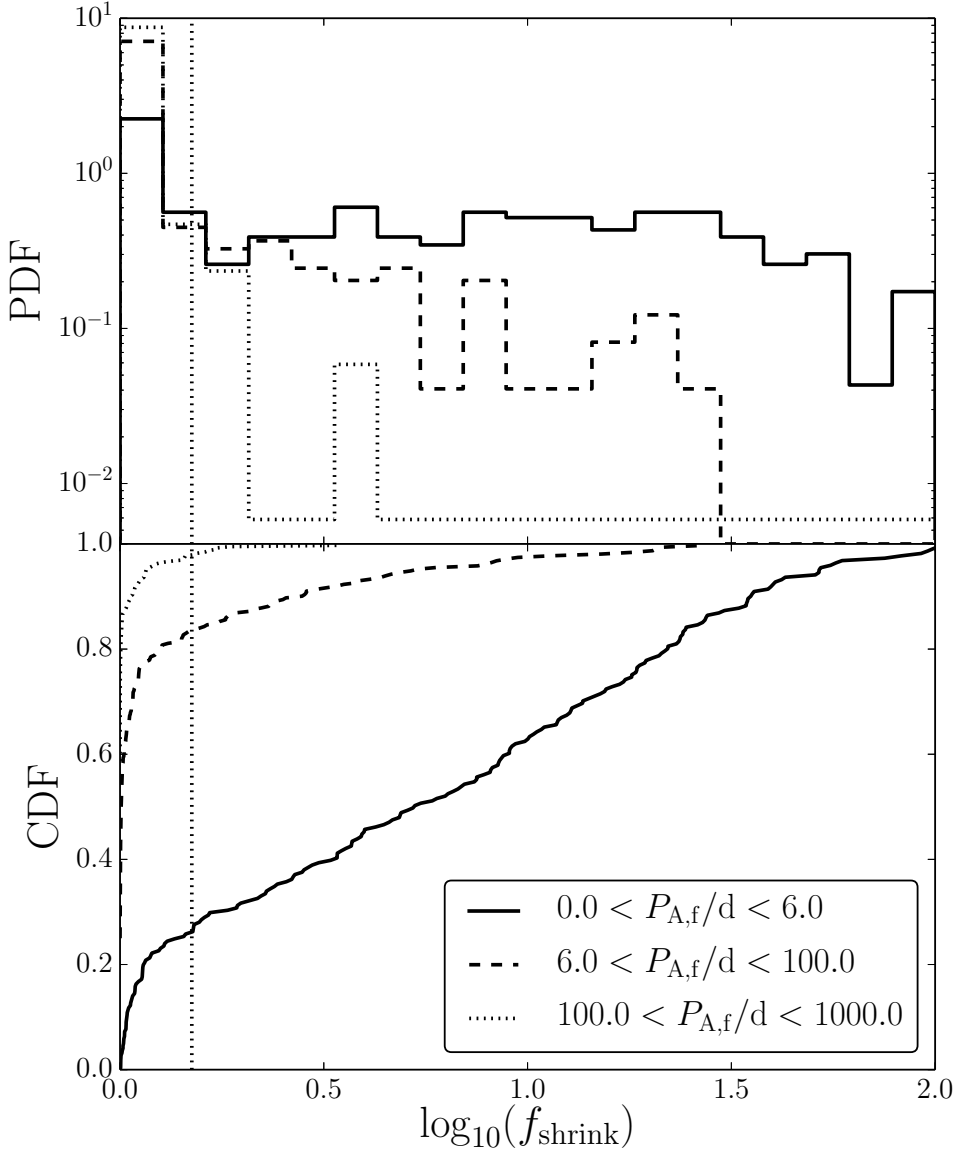


Figure 5.9: The factor $f_{\text{shrink}} \equiv a_{A,i}/a_{A,f}$ with which the inner binary semimajor axis shrinks for the sampled triples as described in Section 5.4.1, for isolated triple evolution as in Fig. 5.8, and here binned with respect to the final inner orbital period $P_{A,f}$. Three bins in $P_{A,f}$ are assumed, indicated in the legend in the bottom panel. The black vertical dotted line indicates the value of f_{shrink} chosen for the selection of systems with substantially shrinking inner binaries.

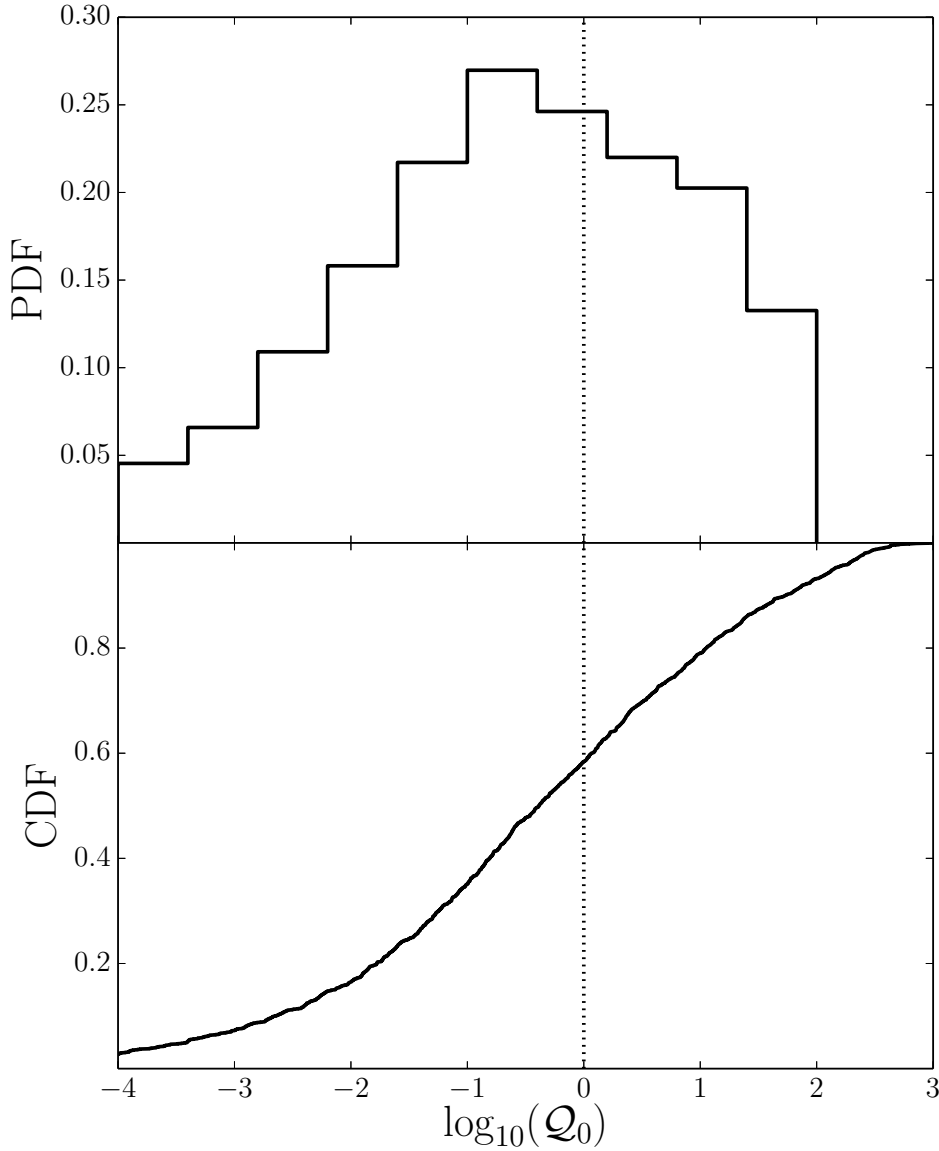
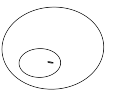


Figure 5.10: The distribution of Q_0 (cf. equation 5.8) for the systems that exceeded the imposed maximum wall time limit of 15 min. The black vertical dotted line shows the value $Q_0 = 1$ for which, for these systems, planets were assumed to be fully effective ($Q_0 < 1$) or ineffective ($Q_0 > 1$) at shielding.



fourth column of Table 5.1).

1. A stellar triple system is sampled, similarly to Fabrycky & Tremaine (2007); Naoz & Fabrycky (2014). For the inner binary, the primary mass m_1 is sampled between 0.5 and 1.2 $M_\odot$ assuming a Salpeter distribution, $dN/dm_1 \propto m^{-2.35}$ (Salpeter 1955). The secondary mass m_2 is sampled from $m_2 = q_{\text{in}} m_1$ with the constraint $0.1 < m_2/M_\odot < 1.2$, where q_{in} is sampled from a flat distribution with $0 < q_{\text{in}} < 1$. The mass of the tertiary star, m_4 , with $0.1 < m_4/M_\odot < 1.2$, is sampled from $m_4 = (m_1 + m_2)q_{\text{out}}$, where q_{out} is sampled from a flat distribution $0 < q_{\text{out}} < 1$. The stellar radii, R_1 , R_2 and R_4 , are computed initially, and at subsequent times, using the stellar evolution code *SEBA* as described in Section 5.2. The spins Ω_1 and Ω_2 of the stars in the inner binary are initially assumed to be parallel with the inner orbital angular momentum, and the spin periods are assumed to be $2\pi/||\Omega_k|| = 10$ d. The tidal quality factors Q'_k for the inner binary primary and secondary stars (cf. Section 5.2) are sampled linearly between 5.5×10^5 and 2×10^6 , adopted from Ogilvie & Lin 2007, who constrained Q'_k using observations of spectroscopic binaries by Meibom & Mathieu (2005). The gyration radii $r_{g,k}$ for the primary and secondary stars are assumed to be $r_{g,k} = 0.08$.

The inner and outer orbital periods, P_A and P_C , are both sampled from a lognormal distribution with mean $\log_{10}(P_k/\text{d}) = 5.03$, standard deviation $\sigma_{\log_{10}(P_k/\text{d})} = 2.28$ and range $-2 < \log_{10}(P_k/\text{d}) < 10$. The latter distribution is consistent with the orbital periods of binaries as found by Duquennoy & Mayor (1991); Raghavan et al. (2010); Tokovinin (2014a). We choose to impose the further restriction of $P_A < 10^3$ d, because most of the progenitor systems of short-period binaries produced through KL cycles with tidal friction initially have $P_A < 10^3$ d (Fabrycky & Tremaine 2007). The corresponding semimajor axes, a_A and a_C , are computed from the orbital periods using Kepler's law (neglecting the planet mass m_3).

The eccentricities e_A and e_C are sampled from a Rayleigh distribution, $dN/de_k \propto e_k \exp(-\beta e_k^2)$, with rms $\langle e_k^2 \rangle^{1/2} = \beta^{-1/2} = 0.33$, consistent with the results of Raghavan et al. 2010. In Raghavan et al. (2010), in contrast with Duquennoy & Mayor (1991), the eccentricity distribution is found not to be substantially different for periods $P_k < 10^3$ d and periods $P_k > 10^3$ d. Therefore, we choose not to sample eccentricities from a thermal distribution for binaries with $P_k > 10^3$ d as in Fabrycky & Tremaine (2007).

The inner and outer orbits are assumed to be randomly oriented, i.e. their arguments of pericentre ω_k and Ω_k are sampled from a random distribution, and their mutual inclination i_{AC} is sampled from a distribution that is linear in $\cos(i_{AC})$. Note that, without loss of generality, we fix the inner orbit to be aligned with the z axis of the coordinate system, i.e. we set $i_A = 0^\circ$. The *initial* mutual inclinations i_{AB} and i_{AC} are therefore equal to the initial individual inclinations i_B and i_C , respectively.

Sampled triple systems are rejected in the following cases at this stage.

- The stability criterion of Mardling & Aarseth 2001 is not satisfied.
- In isolation, the stars in the inner binary would experience strong tidal interaction, i.e. $a_A(1 - e_A) < 3(R_1 + R_2)$.
- In isolation, the stars in the inner binary would fill their Roche lobe, i.e. $R_k < R_{L,k}$ for $k \in \{1, 2\}$, where $R_{L,k}$ is the Roche lobe radius computed at pericentre according to the analytic fitting formulae given by Sepinsky et al. (2007).

2. For the sampled systems of step (1), additional selection is made based on the outcome of the integration of isolated systems. This is to ensure that in the absence of the planet, the system

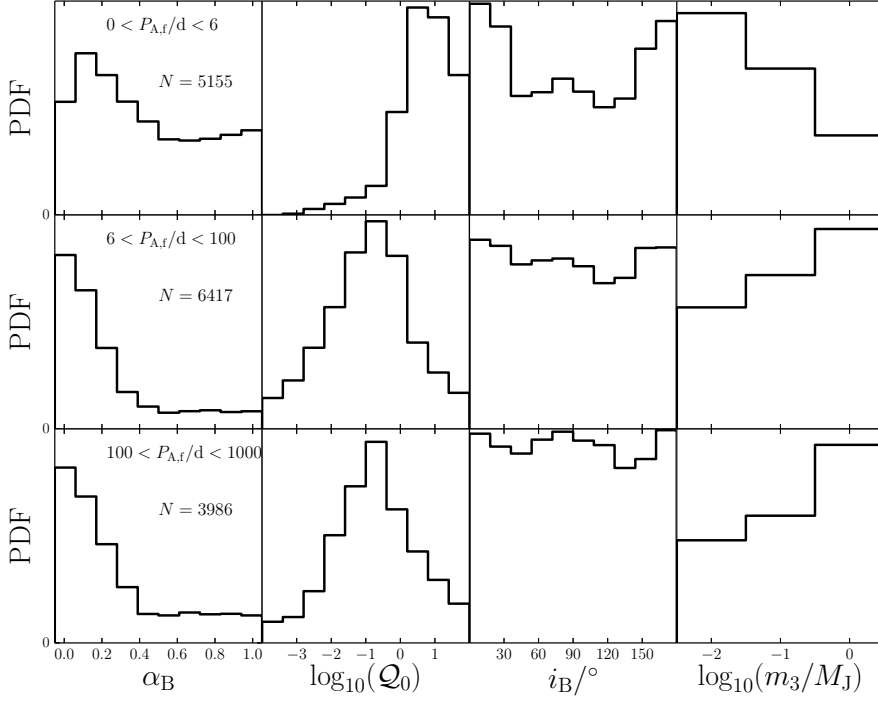


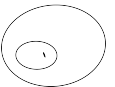
Figure 5.11: The distributions of the parameters α_B (first column), Q_0 (second column), i_B (third column) and m_3 (fourth column) for systems in which the orbit of the planet remains stable during the entire evolution, and no other stopping conditions occurred. Systems are binned with respect to the final inner orbital period $P_{A,f}$; each row corresponds to a different bin, indicated in the first column. Also indicated in the first column is the number of systems in the corresponding bin of $P_{A,f}$.

would remain dynamically stable, and that the inner binary is shrunk substantially because of KL cycles with tidal friction. The integration time for each system is sampled linearly between 1 and 10 Gyr. In addition, the following stopping conditions are always imposed.

- The stars in the inner binary collide or fill their Roche lobe, computed as above.
- One of the three stars evolves past the MS.

First, the sampled systems are integrated taking into account only the inner binary evolution, i.e. stellar and tidal evolution. From these systems, we reject those for which the stars in the inner binary collide, fill their Roche lobe, or evolve past the MS. The former two cases can occur because of tidal friction; note that our inner binaries are initially not circular, nor synchronized [see Section 5.5.6 and Appendix 5.B for discussion on magnetic braking (MB) effects].

Secondly, the sampled systems are integrated taking into account the evolution of the triple system without a circumbinary planet, i.e. taking into account KL cycles in the inner and outer



binaries, and tidal evolution in the inner binary. From these systems, we also reject those for which the inner binary stars collide or fill their Roche Lobe.

In Fig. 5.7, we show the initial (dotted lines) and final (solid lines) inner period distributions for the case of an isolated inner binary system (black lines), and an isolated triple system (blue lines). Consistent with previous studies on KLCTF (Mazeh & Shaham 1979; Eggleton & Kiseleva-Eggleton 2001; Fabrycky & Tremaine 2007; Naoz & Fabrycky 2014), the effect of the tertiary star is to produce an enhancement of systems with periods roughly between 1 and 6 d.

3. We select systems for which, in the case of an isolated triple, the inner binary shrinks substantially during the evolution. To quantify this criterion, we show in Fig. 5.8 the distribution of the factor $f_{\text{shrink}} \equiv a_{A,i}/a_{A,f}$ with which the inner binary semimajor axis decreases during the evolution for the case of isolated triples. Based on this distribution, we select systems for which $f_{\text{shrink}} \geq 1.5$. Effectively, this implies that the majority of the systems evolve to a system with a short-period inner period, $1 < P_{A,f} < 6$. This is demonstrated in Fig. 5.9, where we show the distribution of f_{shrink} in the triple case, for different bins of the final inner orbital period.

4. For each of the $N_{\text{tr}} = 192$ remaining sampled triple systems, we define a three-dimensional grid in (a_B, i_B, m_3) , with size $10 \times 10 \times 3 = 300$. The inclinations i_B range from 0 to 180° , and for the planet mass m_3 we assume three values, 10^{-2} , 10^{-1} and $10^0 M_J$. For a_B , we impose the lower and upper limits $a_{B,l}$ and $a_{B,u}$, respectively.

The lower limit is assumed to be $a_{B,l} = 1.5 a_{B,\text{crit},AB}$, where $a_{B,\text{crit},AB}$ is the critical semimajor axis for dynamical stability of the planet in a coplanar orbit around the inner binary, and which is adopted from the fitting formula given by (Holman & Wiegert 1999). The latter formula is a function of the inner binary mass ratio $\mu = m_2/m_1$ and eccentricity e_A . The factor of 1.5 in $a_{B,l} = 1.5 a_{B,\text{crit},AB}$ is a ‘safety factor’; close to the regime of dynamical stability, the orbits are non-Keplerian (e.g. Lee & Peale 2006; Leung & Lee 2013; Bromley & Kenyon 2015), and, therefore, the orbit-averaged approach no longer applies. The upper limit is assumed to be $a_{B,u} = 0.9 a_{B,\text{crit},BC}$, where $a_{B,\text{crit},BC}$ is the largest possible value of a_B for dynamical stability with respect to the orbit of the tertiary star. The latter value is estimated by applying the Mardling & Aarseth 2001 criterion to the BC pair, with an ‘outer’ mass ratio of $q_{\text{out}} = m_4/(m_1 + m_2 + m_3)$.

In the analysis below, we use the dimensionless quantity α_B which is closely related to a_B , and is defined as

$$\alpha_B \equiv \frac{a_B - a_{B,l}}{a_{B,u} - a_{B,l}}. \quad (5.10)$$

This quantity is motivated by the fact that the allowed range of initial a_B for dynamical stability varies per triple system. By construction, this is not the case for α_B , which, for any sampled triple, ranges between $\alpha_B = 0$ ($a_B = 1.5 a_{B,\text{crit},AB}$, i.e. the planet’s orbit is close to the dynamical stability limit with respect to the inner binary), and $\alpha_B = 1$ ($a_B = 0.9 a_{B,\text{crit},BC}$, i.e. the planet’s orbit is close to the dynamical stability limit with respect to the outer binary). In other words, by using α_B , we can evaluate the effect of the circumbinary planet for a population of triples with different parameters.

For the $N_{\text{quad}} = 300 \times 192 = 57600$ integrations with a planet, the following additional stopping conditions are set.

- The circumbinary orbit is unstable with respect to the inner and outer binaries according to the (Mardling & Aarseth 2001) criterion.

- The circumbinary orbit intersects with the inner binary, i.e. $a_B(1 - e_B) \leq a_A(1 + e_A)$, and is therefore unlikely to be stable.

For some quadruple systems, the integration proceeded slowly. We therefore imposed a maximum wall time of 15 min per system, which was reached for 19% of the integrated systems. In Fig. 5.10, we show the distribution of Q_0 (cf. equation 5.8) for the systems in which the maximum wall time was exceeded. For $\approx 60\%$ of these systems, $Q_0 \lesssim 1$. As shown previously in Section 5.3, for $Q_0 \lesssim 1$, shielding is typically effective, whereas it is not for $Q_0 \gtrsim 1$. Therefore, for the purposes of the analysis below, we will assume that in the systems that exceeded the integration wall time, if $Q_0 \lesssim 1$, shielding is completely effective and hence the inner orbital period does not change (i.e. $P_{A,f} = P_{A,i}$), and if $Q_0 \gtrsim 1$, shielding is completely ineffective and hence the inner orbital period shrinks as much as would be the case without a planet. Clearly, this is approximate because in reality, there is no sharp transition between the two regimes and there is also a dependence on the inclination i_B (cf. Fig. 5.2). However, we do not expect that the approximation strongly affects our conclusions.

5.4.2 Results

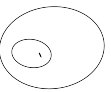
Initial parameters for surviving planets as a function of the final binary period

In Fig. 5.11, we show the distributions of the initial parameters α_B (cf. equation 5.10), Q_0 (cf. equation 5.8), i_B and m_3 in the first through fourth columns, respectively, for systems in which the orbit of the planet remains stable during the entire evolution according to the criteria discussed in Section 5.4.1. Each row in Fig. 5.11 corresponds to a different bin of the final inner orbital period, $P_{A,f}$.

For short periods, the planet tends to be further away from the inner binary (i.e. α_B and Q_0 are typically larger) compared to longer periods. This trend is compatible with the lack of tight coplanar planets around short-period binaries, although for our sampled population, there is certainly no *absolute* lack of planets in tight (i.e. $\alpha_B \lesssim 0.3$) orbits around short-period binaries.

In fact, for $1 \text{ d} < P_{A,f} < 6 \text{ d}$, the distribution of α_B is still peaked around $\alpha_B \approx 0.1$. There are two important factors that contribute to this.

1. In our initial triple population, there is a non-negligible number of systems with *initial* inner periods $1 \text{ d} < P_{A,i} < 6 \text{ d}$ (cf. Fig. 5.7). Evidently, for these systems, a possible outcome is a final period $1 \text{ d} < P_{A,f} < 6 \text{ d}$ when the inner orbit did not shrink due to effective shielding by the planet, likely corresponding to a massive planet in a tight and/or coplanar orbit. In other words, in principle, a short-period binary can simply be born with a tight and coplanar planet, and the planet would prevent the inner orbital period from becoming even shorter. Given the observational evidence for a third stellar companion around most short-period binaries (Tokovinin et al. 2006), one might consider the possibility that such binaries can not primordially form without the effects of KLCTF. If this is the case, then the primordially short-period binaries in our simulations should not be considered, and therefore, we would expect no tight coplanar circumbinary planets to exist at all around short-period binaries.
2. For $\alpha_B \gtrsim 0.3$ or $Q_0 \gtrsim 10^0$, the planetary orbit for a large fraction of systems becomes highly eccentric and intersects with the inner binary, and likely becomes unstable (cf.



Sections 5.4.2 and 5.5.3). These cases are not included in Fig. 5.11, and this produces a bias for an *absence* of stable systems with $\alpha_B \gtrsim 0.3$.

To take these complicating factors into account, we introduce the ‘shielding efficiency’ η_{shield} which quantifies the planet’s ability to shield the inner binary from KL eccentricity oscillations induced by the tertiary star, defined as

$$\eta_{\text{shield}} \equiv \frac{f_{\text{shrink,triple}} - f_{\text{shrink,quad}}}{f_{\text{shrink,triple}} - 1}. \quad (5.11)$$

Here, $f_{\text{shrink,triple}}$ is the factor with which the inner binary semimajor axis decreases in the absence of a planet, and $f_{\text{shrink,quad}}$ is the corresponding factor with a planet present. Note that we restrict to systems for which $f_{\text{shrink,triple}} \geq 1.5$. If the planet is able to fully shield the inner binary, $f_{\text{shrink,quad}} = 1$, and $\eta_{\text{shield}} = 1$. On the other hand, when shielding is completely ineffective, $f_{\text{shrink,quad}} = f_{\text{shrink,triple}}$, and $\eta_{\text{shield}} = 0$. For some systems, we find that the planet can *enhance* the inner binary eccentricity excitations; in this case, $\eta_{\text{shield}} < 0$ (cf. Section 5.4.2).

In Fig. 5.12, we show the distributions of α_B , Q_0 , i_B and m_3 , where the systems are now binned with respect to η_{shield} . When shielding is ineffective ($\eta_{\text{shield}} \approx 0$, cf. the second row in Fig. 5.12), Q_0 is invariably $\gtrsim 1$ and m_3 tends to be low (i.e., the large value of $m_3 = 1 M_J$ is disfavoured). In contrast, when shielding is highly effective ($\eta_{\text{shield}} \approx 1$, cf. the last row in Fig. 5.12), Q_0 is typically $\lesssim 1$ (with a peak near $Q_0 = 0.1$), and m_3 tends to be high (i.e., the large value of $m_3 = 1 M_J$ is favoured).

The distribution of the initial i_B in Fig. 5.12 is a strong function of α_B and m_3 . In Fig. 5.13, we show, for $0.9 < \eta_{\text{shield}} < 1$, the inclination distribution for different ranges of α_B and m_3 . For $\alpha_B < 0.3$, there is a tendency for coplanar orbits, for both low-mass and massive planets. For planets in wider orbits, $\alpha_B > 0.3$, the distribution of i_B is strongly peaked towards $i_B \approx 90^\circ$ for low planet masses, whereas for high planet masses, the distribution of i_B is much less peaked. These trends are qualitatively consistent with the smaller number of integrations that were carried out in Section 5.3 for fixed parameters of the stellar triple.

In Fig. 5.14 we show, for each of the final period bins of Fig. 5.11, the initial distributions of α_B , Q_0 , i_B and m_3 , where additional binning was made with respect to η_{shield} . Black (blue) lines correspond to $\eta_{\text{shield}} < 0.2$ ($\eta_{\text{shield}} > 0.2$). Cumulative distributions for the different final period bins and the two ranges of shielding efficiencies are shown in Fig. 5.15. As expected, the close planetary orbits from the top row in Fig. 5.11 ($1 \text{ d} < P_{A,f} < 6 \text{ d}$) typically correspond to a high shielding efficiency (cf. the blue lines in Fig. 5.14). This indicates that the inner binary did not shrink, but was simply formed with a planet in such an orbit. On the other hand, the wider orbits typically correspond to a low shielding efficiency (cf. the black lines in Fig. 5.14). In the latter case, the initial inner orbit was much wider, and shrank substantially.

Unstable planets

As mentioned in Section 5.4.2, for a large fraction of systems, the planetary orbit becomes highly eccentric because of the torque of the outer orbit, and intersects with the inner binary. This likely results in an unstable planetary orbit, with a high probability for the planet being ejected from the system (for further discussion on and calculations of the outcomes, see Section 5.5.3). Such orbit crossings typically occur early in the evolution. This is illustrated in Fig. 5.16, where we plot the distribution of the time of orbit crossing t_{cross} , when applicable. For

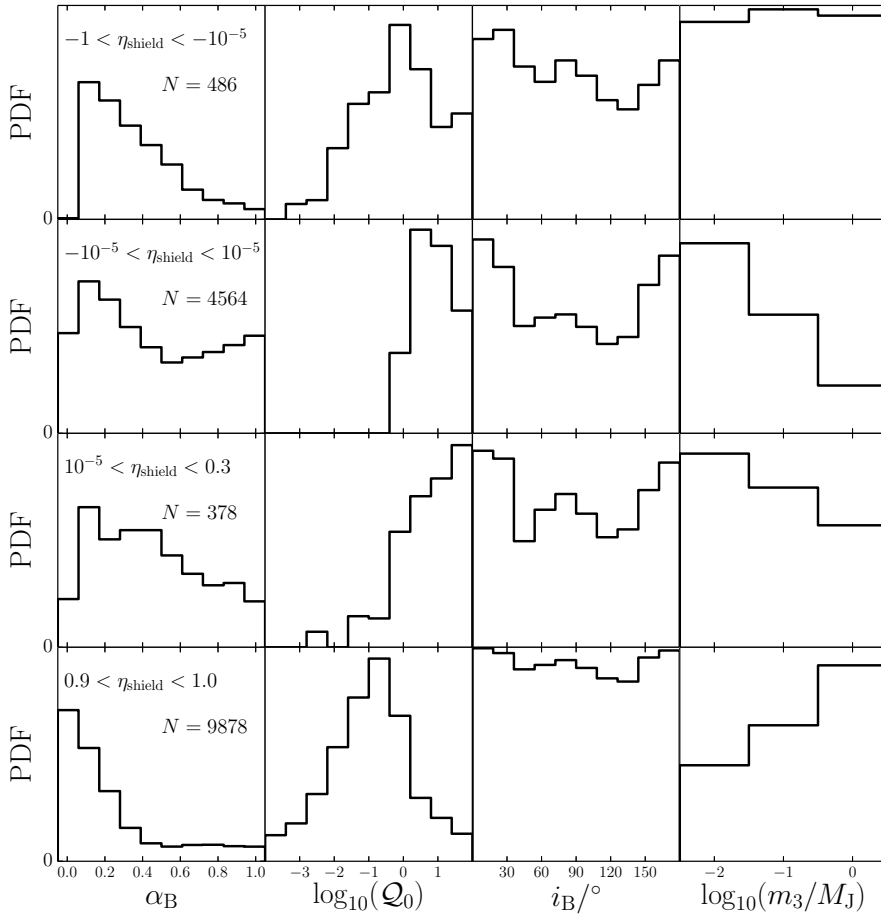
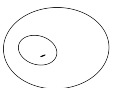


Figure 5.12: The distributions of the parameters α_B (first column), Q_0 (second column), i_B (third column) and m_3 (fourth column), where systems are binned with respect to the shielding efficiency η_{shield} (cf. equation 5.11); each row corresponds to a different bin, indicated in the first column. Also indicated in the first column is the number of systems in the corresponding bin of η_{shield} .



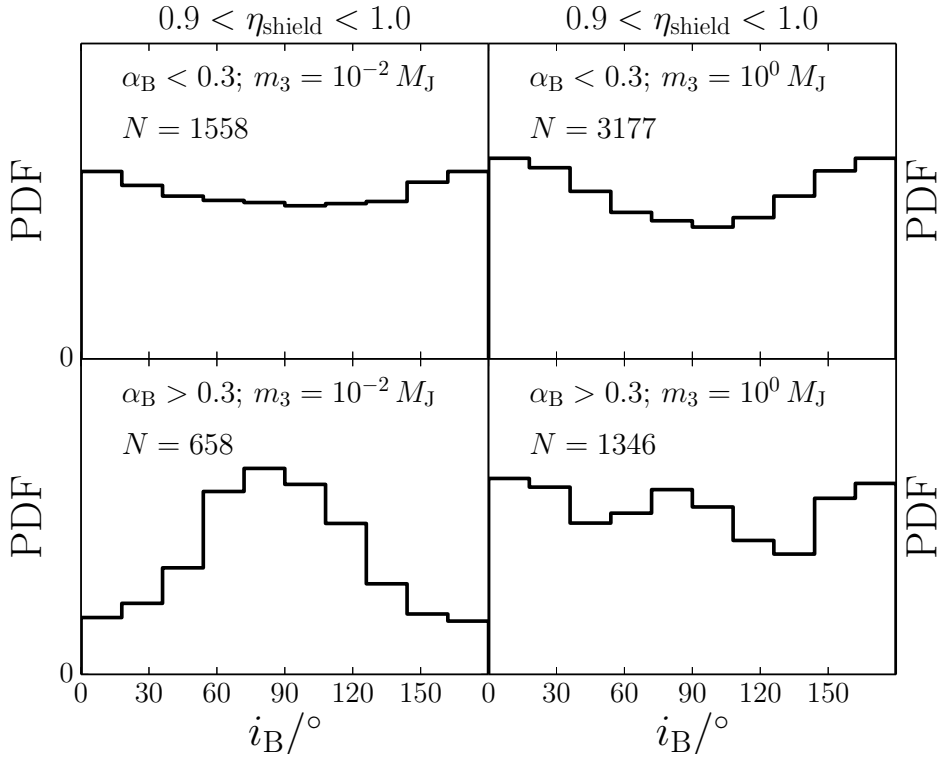


Figure 5.13: The distributions of the initial i_B for $0.9 < \eta_{\text{shield}} < 1$ (cf. Fig. 5.12), distinguishing between different ranges of α_B and m_3 . In the top (bottom) row, $\alpha_B < 0.3$ ($\alpha_B > 0.3$); in the left (right) column, $m_3 = 10^{-2} M_J$ ($m_3 = 10^0 M_J$). The number of systems in each case is indicated in each panel.

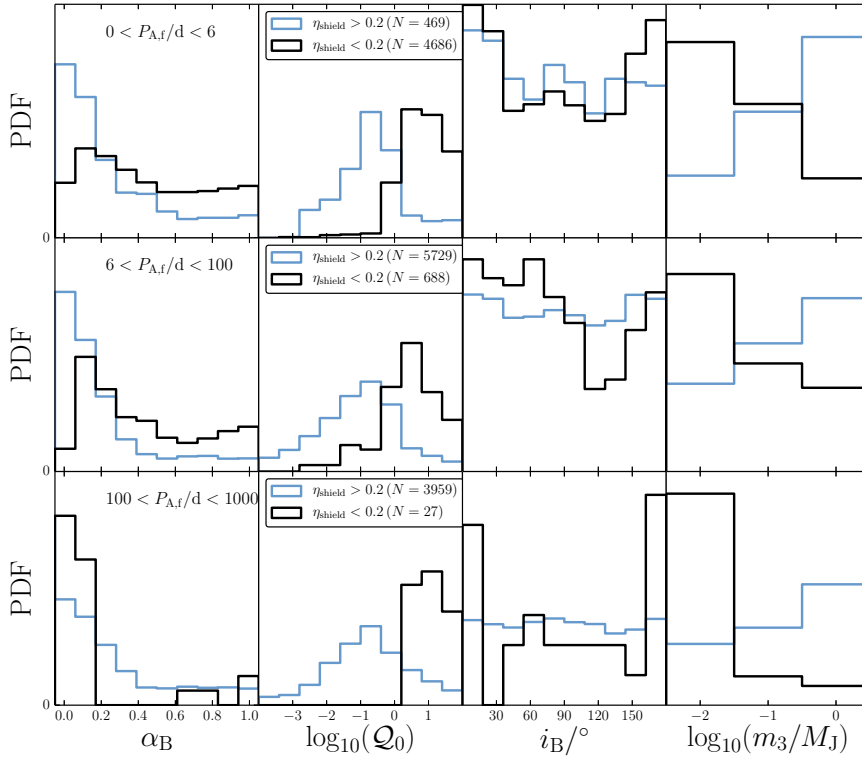
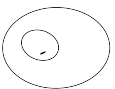


Figure 5.14: The distributions of the parameters α_B (first column), Q_0 (second column), i_B (third column) and m_3 (fourth column) for systems in which the orbit of the planet remains stable during the entire evolution, and no other stopping conditions occurred. Systems are binned with respect to the final inner orbital period $P_{A,f}$ as in Fig. 5.11; here, we also make a distinction between a low shielding efficiency ($\eta_{\text{shield}} < 0.2$; black lines) and high shielding efficiency ($\eta_{\text{shield}} > 0.2$; blue lines).



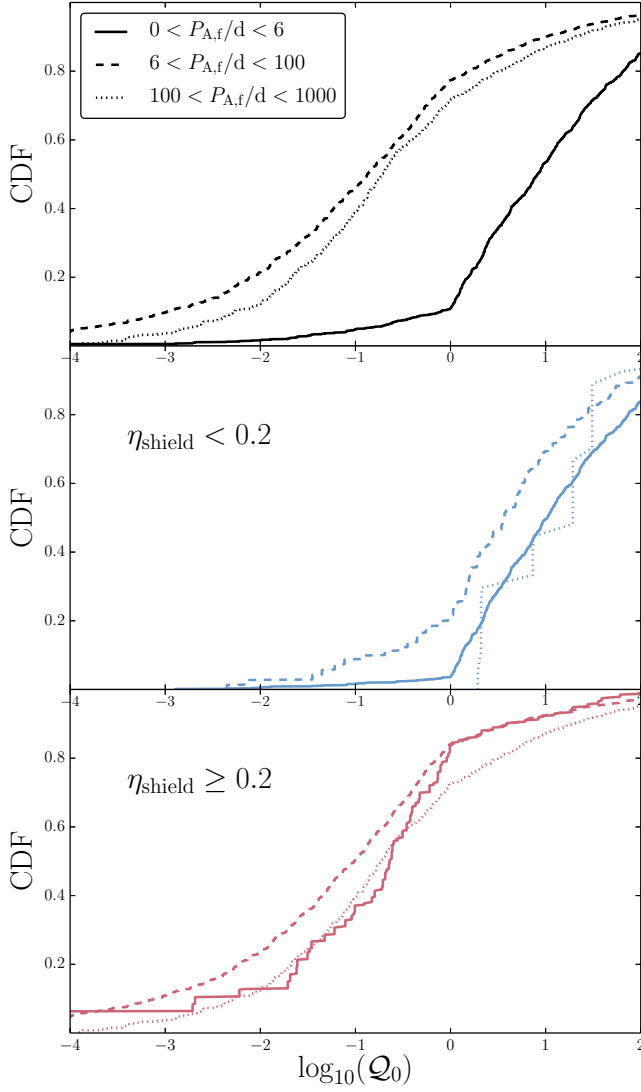


Figure 5.15: Cumulative distributions for Q_0 , binned with respect to the final inner orbital period $P_{A,f}$ (indicated in the legend in the top panel). In the middle (bottom) panels, a distinction is made with respect to the shielding efficiency η_{shield} (cf. equation 5.11): including the systems with $\eta_{\text{shield}} < 0.2$ (middle panel), and with $\eta_{\text{shield}} > 0.2$ (bottom panel). The meaning of the linestyles is the same between the different rows.

$\approx 80\%$ of the cases, $t_{\text{cross}} < 10$ Myr, which is very short compared to the MS time-scale of any of the stars (as described in Section 5.4.1, the integration time is sampled linearly between 1 and 10 Gyr). Because of this quick removal of the planet, it is unlikely that the planet could affect the inner binary evolution in these cases, and, therefore, we correspondingly assume $\eta_{\text{shield}} = 0$.

In Fig. 5.17, we show the fraction f_{cross} of systems with orbit crossing as a function of the initial parameters α_B , Q_0 , i_B and m_3 . This fraction increases strongly from $f_{\text{cross}} \approx 0.4$ for $\alpha_B \approx 0$ to $f_{\text{cross}} \approx 0.9$ for $\alpha_B \gtrsim 0.5$. This may seem counterintuitive: for larger α_B , the planet is further away from the inner binary, therefore orbit crossings are expected to be less likely. However, for larger α_B , the planet is also placed closer to the tertiary star, causing greater eccentricity excitation in the circumbinary orbit (e.g. the top-left panel in Fig. 5.2). Apparently, the latter effect is (much) stronger than the former, resulting in more orbit crossings for wider planetary orbits. As mentioned in Section 5.4.2, the high occurrence of orbit crossings produces a tendency for a lack of *stable* planets for $\alpha_B \gtrsim 0.3$.

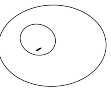
There is no strong dependence of f_{cross} on i_B , nor m_3 . This can be understood by noting that the distribution of the mutual inclination between the circumbinary and the outer orbits, i_{BC} , is the same as the distribution of the mutual inclination between the inner and outer orbits, i_{AC} (i.e. $\cos[i_{AC}]$ was sampled linearly between -1 and 1; cf. Section 5.4.1). Therefore, the distribution of i_{BC} is independent of i_B , and there is no preference with regard to f_{cross} for orbits with i_B close to 90° . Furthermore, because the planet mass m_3 is typically less than 10^{-3} of the stellar masses, $P_{\text{KL},BC} \propto (m_1 + m_2 + m_3 + m_4)/m_4$ is nearly independent of m_3 , explaining why f_{cross} is essentially independent of m_3 . Note that, in contrast, the shielding efficiency η_{shield} is a strong function of m_3 (cf. Fig. 5.12).

Final planet orbital eccentricities

In Figs 5.18 and 5.19, we show the distributions of the planet orbital eccentricity at the end of the integration, $e_{B,f}$, binned with respect to $P_{A,f}$ and η_{shield} , respectively. In the case of binning with respect to η_{shield} , orbit crossing cases are also included, in which case we set $\eta_{\text{shield}} = 0$ (cf. Section 5.4.2). For a crossing of the planetary orbit with the inner binary to occur, e_B needs to be high. This causes $e_{B,f}$ to be typically high if $\eta_{\text{shield}} \approx 0$; in $\approx 50\%$ of the cases, $e_{B,f} > 0.9$. On the other hand, if the planetary orbit remains stable, e_B tends to be much smaller. For cases of very effective shielding, i.e. $0.9 < \eta_{\text{shield}} < 1$, $\approx 90\%$ of the systems have $e_{B,f} < 0.2$. In Fig. 5.18, the planetary orbit remains stable in all cases, and $e_{B,f}$ tends to be low. There is little to no dependence of the distribution of $e_{B,f}$ on the final orbital period.

Reversed shielding

As shown in the top row of Fig. 5.12, in a relatively small number of cases, $\eta_{\text{shield}} < 0$, i.e. the inner binary shrinks *more* compared to the situation without a planet. The values of α_B for which this occurs are strongly peaked towards small values, and Q_0 is peaked near $Q_0 = 1$. For $Q_0 \approx 1$, the KL time-scales for the AB and AC pairs are approximately equal, and, as similarly mentioned with respect to the quantity $\mathcal{R}_0$ by HPAPZ15, this can give rise to chaotic dynamics and possible enhancement of the eccentricity in the inner binary, and hence a negative η_{shield} . Although interesting from a theoretical dynamical point of view, negative η_{shield} only occur for a small number of systems, i.e. a fraction of 0.032 of the stable systems, and 0.0084 of all systems.



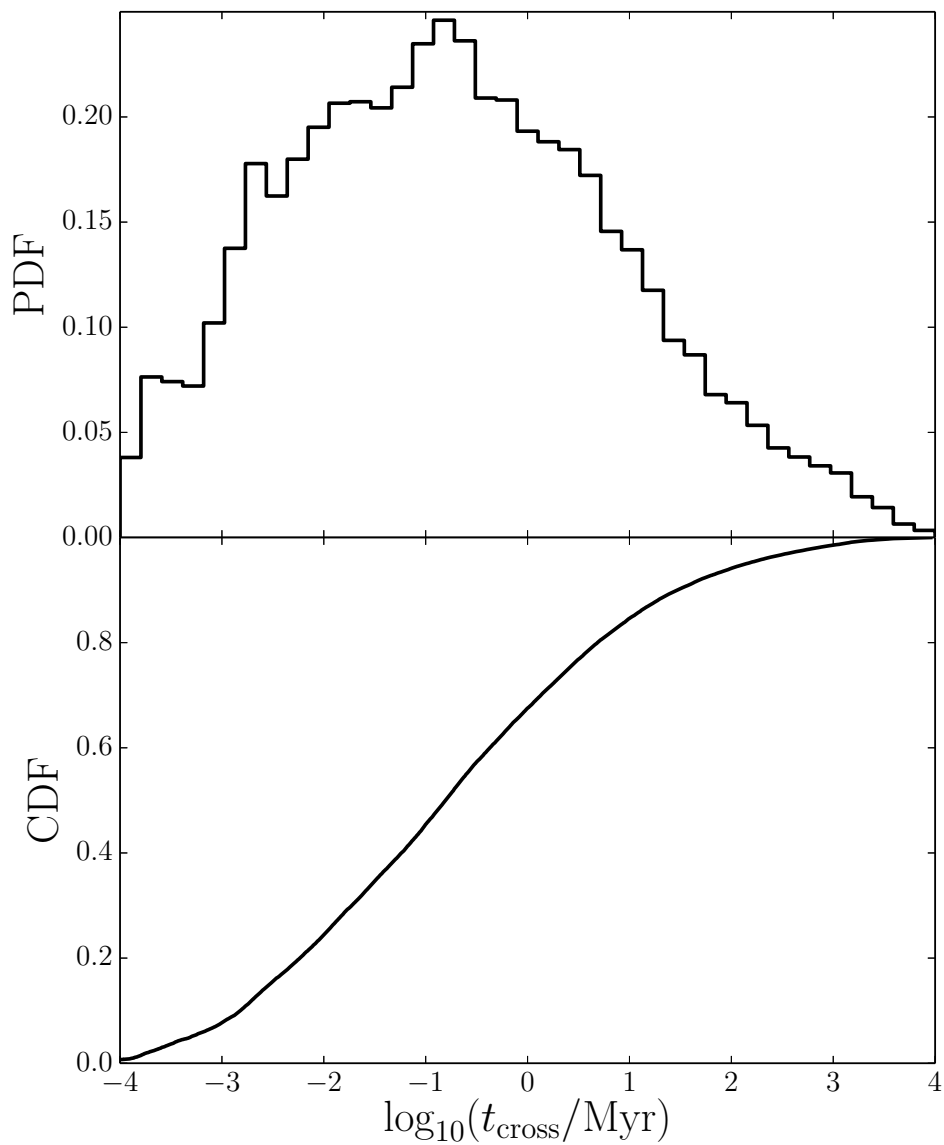


Figure 5.16: The distribution of the time of orbit crossing t_{cross} of the planet with the inner binary, for the systems in which this occurred.

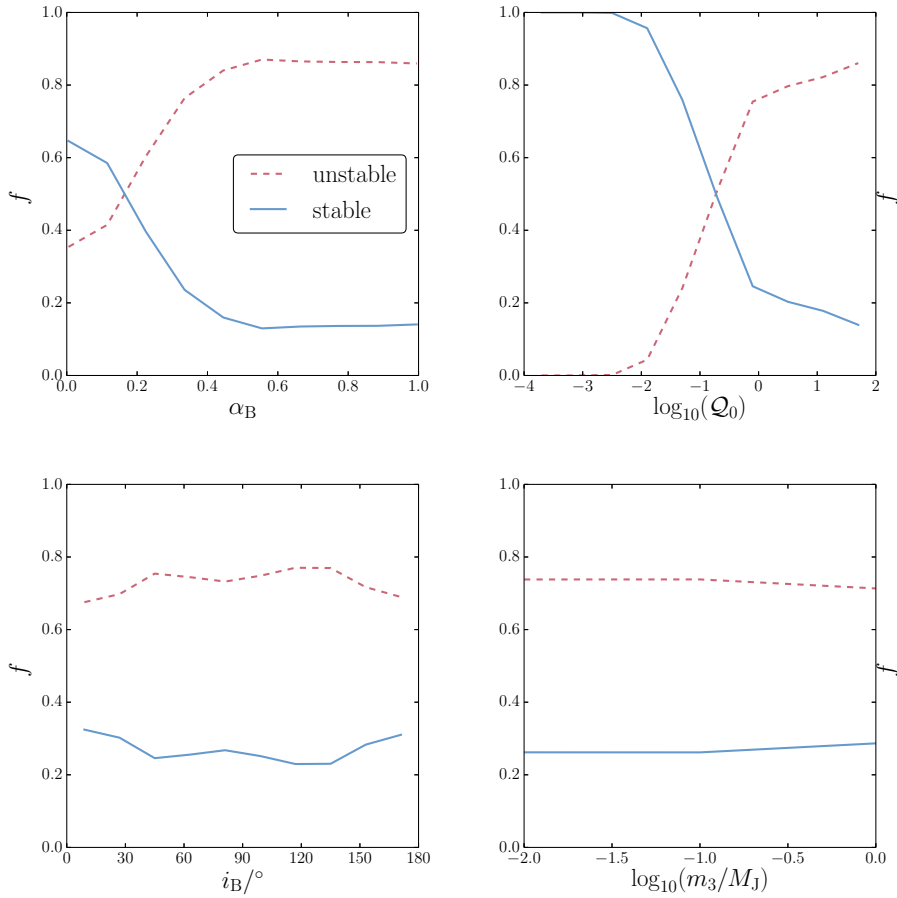
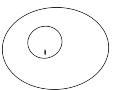


Figure 5.17: The fractions of systems for which the planet either remains stable (blue solid lines), or for which the orbit of the planet becomes highly eccentric and intersects with the inner binary, likely resulting in an unstable orbit (red dashed lines), as a function of α_B , Q_0 , i_B and m_3 .



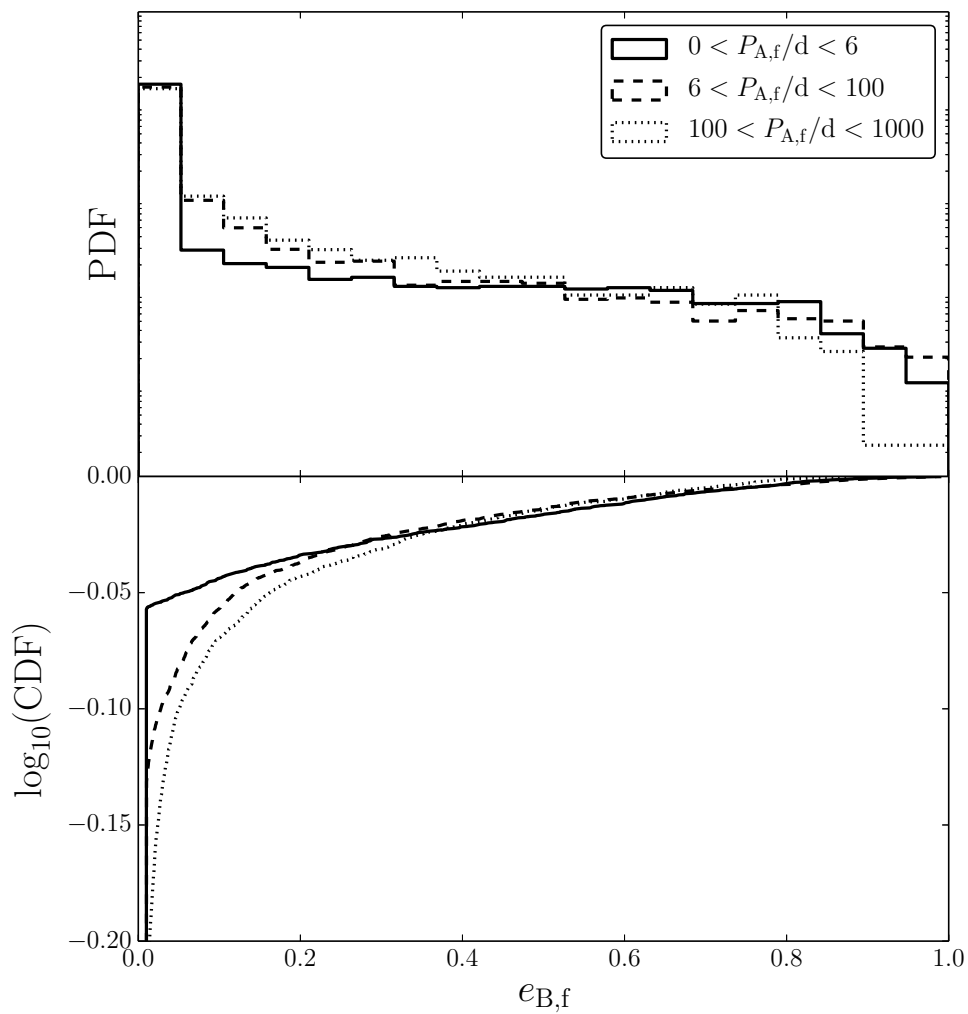


Figure 5.18: The distributions of the final eccentricity $e_{B,f}$ of the planetary orbit for stable systems and binned with respect to the final inner orbital period $P_{A,f}$.

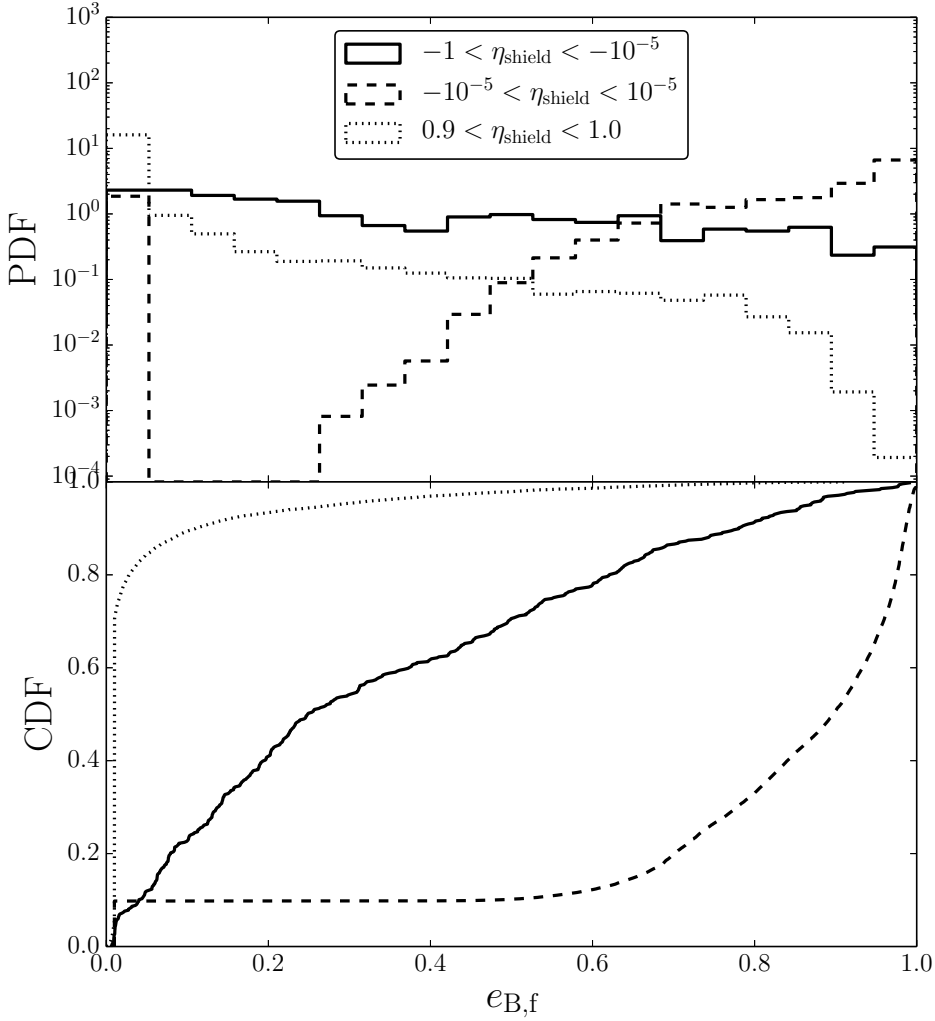
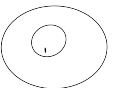


Figure 5.19: The distributions of the final eccentricity $e_{B,f}$ of the planetary orbit, binned with respect to the shielding efficiency η_{shield} . Here, cases are included when the planetary orbit intersects with the inner binary, in which case η_{shield} is assumed to be 0.



5.5 Discussion

5.5.1 An approximate analytic condition for planet shielding -- implications for other systems

In Section 5.3.1, we have shown that planet shielding is effective for $Q_0 \lesssim 1$ (cf. Fig. 5.2). We have also confirmed this for a much larger number of systems, and with the inclusion of tidal friction, in Section 5.4 (cf. Fig. 5.12). As demonstrated in e.g. Figs 5.2, 5.4 and 5.13, there is also a dependence of the planet shielding ability on the inclination i_B . Nevertheless, as a first approximation, the condition $Q_0 \lesssim 1$ can be used to evaluate the importance of planet-shielding in stellar triples. Using equation (5.8), the condition can be written as

$$\begin{aligned} a_B &\lesssim a_C \left(\frac{m_3}{m_4} \frac{m_1 + m_2 + m_4}{m_1 + m_2 + m_3 + m_4} \right)^{1/3} \left(\frac{1 - e_{C,0}^2}{1 - e_{B,0}^2} \right)^{1/2} \\ &\approx a_C \left(\frac{m_3}{m_4} \right)^{1/3} \left(\frac{1 - e_{C,0}^2}{1 - e_{B,0}^2} \right)^{1/2} \\ &= 0.1 a_C \left(\frac{m_3}{m_4} 10^3 \right)^{1/3}, \end{aligned} \quad (5.12)$$

where in the second line, we assumed that m_3 is negligible compared to the other masses, and in the last line, we assumed $e_{B,0} = e_{C,0} = 0$. For small planet masses, the scaling with m_3 is $a_B \propto m_3^{1/3}$. Equation 5.12 implies that for a planet of order Jupiter mass in a *stellar* triple (i.e. $m_k \sim 1 M_\odot$ for $k \in \{1, 2, 4\}$), the circumbinary semimajor axis should be less than approximately a tenth of the outer semimajor axis.

We remark that the fourth body orbiting the circumbinary planet system may also be less massive (by factors up to a thousand) than was assumed in Sections 5.3 and 5.4. In particular, it could be a brown dwarf or a massive planet orbiting a lower-mass circumbinary planet. As m_4 decreases, equation (5.12) implies that for fixed a_C , a_B can be larger for shielding to be effective. This can be explained intuitively by noting that for fixed m_3 and decreasing m_4 , the $P_{\text{KL,AB}}$ time-scale remains constant, whereas the $P_{\text{KL,AC}}$ time-scale increases, i.e. the outermost body becomes less dominant. This, combined with the lower efficiency of a tertiary with a lower mass to shrink the inner orbit, suggests that the lack of circumbinary planets around short-period binaries is even more severe for triples with low-mass tertiary companions. Similarly, shielding is expected to be more effective for lower inner binary masses m_1 and m_2 , suggesting a more severe lack of planets around low-mass short-period binaries.

5.5.2 Shielding of the planetary orbit by the inner binary

Although not discussed in detail here, shielding can also occur in the orbit of the planet: the inner binary can induce rapid precession in the planetary orbit, shielding the latter against high-amplitude KL eccentricity oscillations induced by the outer binary. This aspect was discussed in detail by HPAPZ15, where it was shown that the approximate condition for shielding of the

planetary orbit is $\mathcal{R}_0 \lesssim 1$, or (cf. equation 5.9)

$$\begin{aligned}
 a_B &\lesssim a_A^{1/3} a_C^{2/3} \left(\frac{m_3}{m_4} \right)^{2/9} \left(\frac{m_1 + m_2 + m_3}{m_1 + m_2} \right)^{1/9} \left(\frac{1 - e_{C,0}^2}{1 - e_{B,0}^2} \right)^{1/3} \\
 &\approx 0.22 a_A^{1/3} a_C^{2/3} \left(\frac{m_3}{m_4} 10^3 \right)^{2/9}, \\
 &\approx 0.92 \text{ AU} \left(\frac{a_A}{0.2 \text{ AU}} \right)^{1/3} \left(\frac{a_C}{20 \text{ AU}} \right)^{2/3} \left(\frac{m_3}{m_4} 10^3 \right)^{2/9}
 \end{aligned} \tag{5.13}$$

where in the second and last lines, we neglected m_3 compared to the other masses, and where we assumed circular orbits. In general, $\mathcal{R}_0 \gg \mathcal{Q}_0$ (cf. equation 5.9), indicating that typically, shielding of the inner binary by the planet ($\mathcal{Q}_0 \lesssim 1$) is more likely than shielding of the planet by the binary ($\mathcal{R}_0 \lesssim 1$).

We note that similar dynamics apply to satellites around the Pluto-Charon binary system in the Solar system. As shown by Michaely et al. (2015), if the orbit of a satellite is close enough to the Pluto-Charon binary, then precession induced by the Pluto-Charon binary on the satellite's orbit protects the latter from KL oscillations induced by secular perturbations of the Sun.

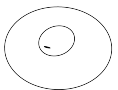
5.5.3 The fate of planets with unstable orbits

As mentioned in Sections 5.3 and 5.4, when the orbit of the planet is relatively close to the outer binary, the former can become highly eccentric because of KL eccentricity oscillations induced by the tertiary companion, and intersect with the inner binary. In fact, this is a likely scenario, as demonstrated by e.g. Fig. 5.17. Here, we investigate the possible outcomes of such orbit crossings and hence likely unstable orbits, by integrating the four-body system using the direct N -body code *HERMITE* (Hut et al. 1995) incorporated in *AMUSE* (Portegies Zwart et al. 2013; Pelupessy et al. 2013). Here, tidal effects are not included. The initial conditions are taken from the grid, of size 300, of the initial parameters a_B , i_B and m_3 for one the triple systems sampled in Section 5.4. Details of the parameters are given in the third column of Table 5.1.

First, the systems are integrated using our secular code until the circumbinary orbit intersects with the inner binary. Subsequently, we sample 100 different sets of random mean anomalies for the three orbits, and integrate the system for each set for the duration of 40 initial circumbinary orbital periods. In this manner, the most important outcomes are revealed for each combination of planet parameters. In total, $300 \times 100 = 3 \times 10^4$ direct N -body integrations were carried out.

We find the following outcomes, in order of decreasing likelihood f :

1. the planet becomes unbound from the system ($f \approx 0.657$);
2. the planet orbits the outer binary (i.e. a circumtriple planet; $f \approx 0.289$)
3. the planet collides with a star ($f \approx 0.029$);



4. the planet remains stable as a circumbinary planet, but with a different orbit ($f \approx 0.017$);
5. the planet becomes bound to a single star ($f \approx 0.009$).

The large probability of ejection of the planet is intuitively easy to understand from the low mass of the planet compared to that of the stars. Circumtriple planets are also common, but note that the integration time is limited and the orbits are typically eccentric, therefore not all systems may be stable indefinitely.

In the top row of Fig. 5.20, we show the distribution of the fractions of the outcomes. In the bottom row of the same figure, we show the distributions for collisions with stars (left column) and bindings to stars (right column), where a distinction is made between the different stars. Collisions are most common with the primary star. This is easily understood by noting that by definition, the primary star is the most massive, and, therefore, also has the largest radius (as described in Section 5.2, the radii are calculated using the *SEBA* stellar evolution code). With regards to the planet becoming bound to a single star, being bound to the tertiary is most likely. This may be the result of a lower orbital speed of the planet when it is close to the tertiary (roughly corresponding to the apocentre of the circumbinary orbit if it were still stable), as opposed to when it is close to the inner binary (roughly corresponding to the pericentre of the circumbinary orbit if it were still stable).

The outcomes found above – keeping in mind the caveat that they are based on integrations of only a single triple system and that the planetary parameters were taken from a linear grid – have interesting implications for stellar triples. Case (i), combined with the high fraction of orbit-crossing planets as found in Section 5.4.2, suggests that circumbinary planets in triples are likely to become unbound from their parent binary early in their evolution, and become free-floating planets. Case (ii) suggests that circumtriple planets should be fairly common. Case (iii) provides a scenario for polluting stars in the inner binary, the primary star in particular. Lastly, case (v) suggests that in stellar triples, circumbinary planets (i.e. P-type planets) can be transformed into circumstellar planets (i.e. S-type planets), albeit with a low probability.

5.5.4 Implications for planets around blue straggler stars

As first suggested by Perets & Fabrycky (2009), and later studied more quantitatively by Naoz & Fabrycky (2014), KLCTF may also lead to mergers or mass transfer in short-period binaries, producing blue straggler stars (BSSs). Our results suggest (secular dynamical) constraints for planets around BSSs formed in this manner, where the planets are either in a circumbinary configuration (in case of binary BSSs) or in a circumstellar configuration (in case when the original stellar binary merged to produce a single BSS star). Analogously to the case of short-period binaries, we expect such planets around BSSs to be typically of low mass, and in wide and/or inclined orbits around the BSS or BSS binary.

5.5.5 Approximations in the integrations

In the numerical integrations of Sections 5.3 and 5.4, the ‘cross’ term that appears in the Hamiltonian at the octupole-order, $\overline{H}_{\text{oct,cross}}$ (cf. section 2.4 of HPAPZ15), was neglected. Note that the ‘non-cross terms’, $\overline{H}_{\text{oct,AB}}$, $\overline{H}_{\text{oct,BC}}$ and $\overline{H}_{\text{oct,AC}}$, were always included in Sections 5.3 and

5.4. For the systems considered here, it is unlikely that the cross term $\overline{H}_{\text{oct,cross}}$ has a large effect on the dynamical evolution because its numerical value is generally very small compared to the other terms that appear in the octupole and the next higher, hexadecupole, orders. For example, for the system chosen in Section 5.3.1 (cf. Table 5.1; setting $m_3 = 1 M_J$, $i_B = 0^\circ$ and $i_C = 90^\circ$), the ratio r of the absolute value of the orbit-averaged cross term to the absolute value of all other orbit-averaged terms at octupole and hexadecupole order, defined in equation (10) of HPAPZ15, is $r \approx 6 \times 10^{-8}$ if $a_B = 1 \text{ AU}$, and $r \approx 7 \times 10^{-7}$ if $a_B = 5 \text{ AU}$. For the system of Section 5.3.2 (setting $m_3 = 1 M_J$, $i_B = 0^\circ$ and $i_C = 90^\circ$), $r \approx 6 \times 10^{-7}$ if $a_B = 10 \text{ AU}$, and $r \approx 2 \times 10^{-7}$ if $a_B = 50 \text{ AU}$.

Furthermore, in the integrations we assumed that there is only one planet orbiting the inner binary. Among the currently limited number of *Kepler* transiting circumbinary planets, there is already one system, Kepler 47 (cf. Table 5.2) with three confirmed circumbinary planets orbiting a single binary. Although the case of multiple planets is beyond the scope of this work and left for future work, we note that in this situation, the planet with the tightest orbit with respect to the inner binary likely has the largest potential for shielding the inner binary from the secular torque of the outer orbit. This picture is complicated by the possibility of planet-planet scattering and mean motion resonances if the planets are in close orbits to each other.

5.5.6 Other dissipative effects

In this work, we considered the dissipative effect of tidal friction in the inner binary. Other potentially important dissipative processes that affect the orbital energies are gravitational wave (GW) emission, MB and gas drag in circumbinary discs.

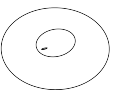
Gravitational wave emission

Shrinkage due to GW emission was not included for numerical reasons, but we note that it is important only in very tight binaries that are beyond the scope of this paper. For circular binaries, the GW inspiral time is $t_{\text{GW}} = 5c^5 a_A^5 / [256G^3 m_1 m_2 (m_1 + m_2)]$, where c is the speed of light (Peters 1964), which is equal to 10 Gyr if $\log_{10}(P_A/d) \approx -0.4$. In our population synthesis, the final orbital periods are typically longer than this value (cf. Fig. 5.7), indicating that shrinkage in the inner orbit due to GW emission is not important (note that the inner orbit is typically nearly circular when the final orbital period is reached).

Magnetic braking

Many uncertainties still exist regarding the efficiency of MB. We evaluated the effect of MB by carrying out the integrations of isolated binaries and triples as in Section 5.4.1, now also including MB for the primary and secondary stars in the inner binary, with standard assumptions for the MB law and its efficiency. Further details are given in Appendix 5.B.

We find that when including the effects of MB, there is a nearly absolute absence of binaries with periods $\lesssim 10 \text{ d}$ (cf. Fig. 5.21), which is grossly inconsistent with observations. The latter, in contrast, show a *peak* in the distribution around 3–6 d (Tokovinin 2014a). This suggests that MB is likely not as efficient as was assumed here. Because MB is not the focus of this work, we chose to exclude it in the integrations of Sections 5.3 and 5.4. Nevertheless, the apparent discrepancy between the canonical efficiency of MB and the period distribution of solar-type MS binaries merits future investigation.



Drag in circumbinary gas discs

If there is a disc around the inner binary, gas drag can cause the circumbinary planet to migrate inwards. Consequently, Q_0 would decrease (cf. equation 5.8), and hence the planet's shielding ability would be increased. However, the lifetime of circumbinary discs is at most a few Myr (Alexander 2012), which is much shorter than the typical duration of our integrations (in our population synthesis, the integration time varied between 1 and 10 Gyr to reflect the typical age of observed *Kepler*-like systems). Therefore, disc migration likely does not affect our conclusions regarding planet shielding in *Kepler*-like systems.

5.6 Conclusions

In recent observations of circumbinary planets, there is an unexpected lack of coplanar circumbinary planets around short-period solar-type MS binaries. This goes *against* observational biases of detecting such planets. We have shown that the lack can be explained by the secular gravitational influence of a circumbinary planet in hierarchical triple systems. Our arguments and conclusions are as follows.

1. Observations show that binaries with periods of $\sim 3-6$ d are nearly always (96%) orbited by a tertiary companion star (Tokovinin et al. 2006). The short orbital period likely resulted from Kozai-Lidov (KL) eccentricity oscillations in the inner orbit induced by the tertiary, combined with tidal friction (Mazeh & Shaham 1979; Eggleton & Kiseleva-Eggleton 2001; Fabrycky & Tremaine 2007; Naoz & Fabrycky 2014). This suggests that many of the short-period *Kepler* eclipsing binaries are triple star systems. The progenitor inner binary was likely wider, with an orbital period of up to $\sim 10^4$ d.
2. We have demonstrated that if there is a circumbinary planet around the progenitor inner binary, then the KL eccentricity oscillations in the inner orbit induced by the tertiary can be quenched through the secular gravitational influence of the circumbinary planet. Thereby, the inner binary is 'shielded' from the torque of the tertiary star, and does not shrink to a tight orbit. However, this only occurs if the circumbinary planet is sufficiently massive, and if its orbit is sufficiently close to and coplanar with the inner binary. In many other cases, the circumbinary orbit is stable but cannot efficiently shield the inner binary, or the circumbinary orbit is destabilized and is most likely ejected.
3. In particular, if a low-mass circumbinary planet is initially inclined with respect to and far away from the inner binary, then planet shielding is typically ineffective, and the inner binary can shrink to a tight orbit. On the other hand, for a more massive planet in an initially approximately coplanar and tight circumbinary orbit, shielding typically prevents the inner orbit from shrinking (cf. Figs 5.14 and 5.12).
4. Consequently, for systems with circumbinary planets surviving through the evolution, short-period inner binaries typically do not have massive ($m_3 \sim M_J$) circumbinary planets on tight ($Q_0 \lesssim 1$, where Q_0 is defined in equation 5.8) and coplanar orbits. Namely, if the latter was present initially, the inner binary would unlikely have shrunk to a short period. In contrast, if the circumbinary planet is less massive ($m_3 \sim 10^{-2} M_J$) and initially inclined with respect to and far away from the inner binary ($Q_0 \gtrsim 1$), then shrinkage is possible. This trend is consistent with the current *Kepler* observations. The transition with respect to the planet mass

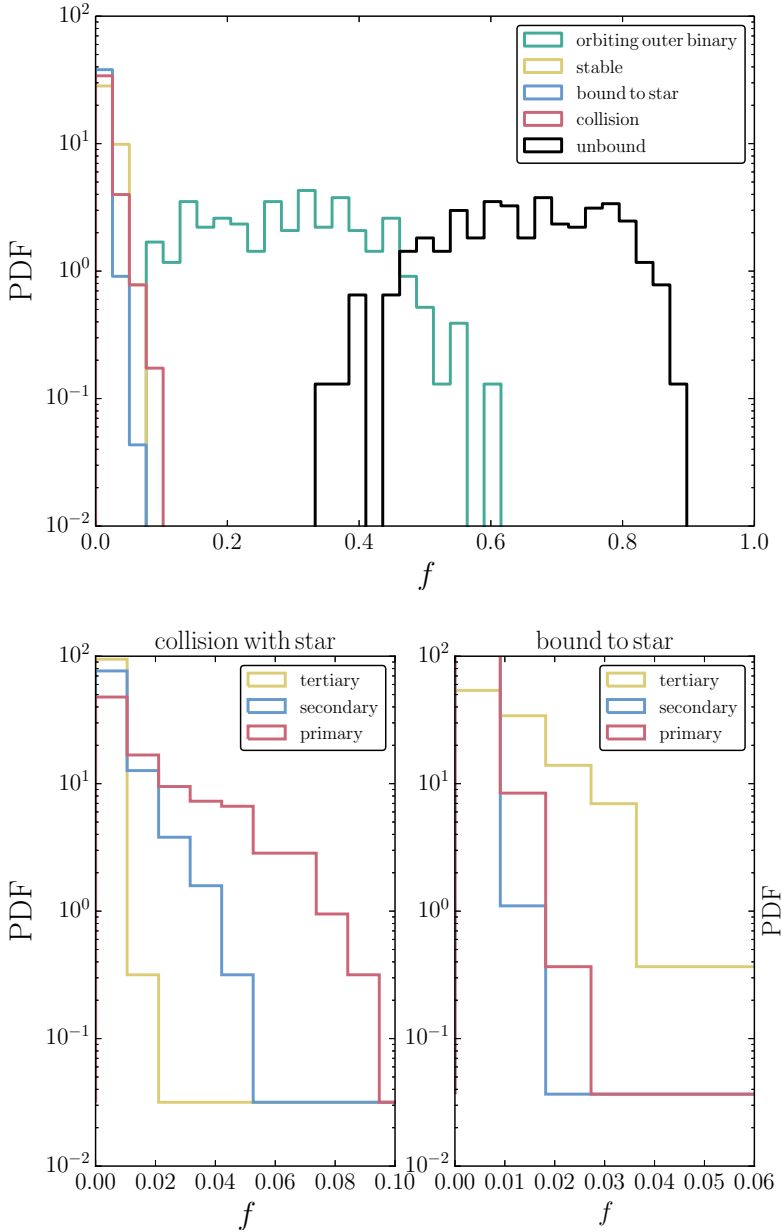
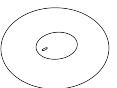


Figure 5.20: The distributions of the fraction of direct N -body integrations with various outcomes, in the case that the orbit of the planet becomes highly eccentric and intersects with the inner binary. The triple system is fixed, whereas a_B , i_B and m_3 are taken from a grid with size 300 as in Section 5.4.1. Top row: overview, with the major channels described in the legend. Bottom row: making a distinction between stars for cases when the planet collides with a star (left column), or when it becomes bound to a star (right column).



and orbital semimajor axis occurs at $Q_0 \sim 1$, which, for a Jupiter-mass planet in a solar-mass triple, corresponds to a circumbinary semimajor axis roughly equal to a tenth of the outer orbit semimajor axis (cf. equation 5.12).

Our results suggest that similar constraints also apply to planets around blue straggler stars.

Acknowledgements

We thank the referee, Smadar Naoz, for providing very helpful comments that led to improvement of the paper. This work was supported by the Netherlands Research Council NWO (grants #639.073.803 [VICI], #614.061.608 [AMUSE] and #612.071.305 [LGM]) and the Netherlands Research School for Astronomy (NOVA). HBP acknowledges support from the ISF I-CORE programme 1829, The European FP-7 CIG programme ‘GRAND’ (333644), the BSF grant number 2012384 and the Asher foundation.

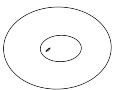
5.A *Kepler* transiting circumbinary planets

In Table 5.2, we give an overview of the currently known *Kepler* transiting circumbinary planets.

System	m_1 $M_\odot$	m_2 $M_\odot$	m_3 M_J	R_1 $R_\odot$	R_2 $R_\odot$	R_3 R_J	a_A AU	P_A d	e_A -	a_B AU	P_B d	e_B -	Ref.
Kepler 16b	0.690	0.203	0.333	0.649	0.226	0.754	0.224	41.1	0.159	0.705	229	0.0069	a
Kepler 34b	1.05	1.02	0.220	1.16	1.09	0.764	0.229	27.8	0.521	1.09	289	0.182	b
Kepler 35b	0.888	0.809	0.127	1.03	0.786	0.728	0.176	20.7	0.142	0.603	131	0.042	c
Kepler 38b	0.949	0.249	<0.384	1.76	0.272	0.388	0.145	18.8	0.103	0.464	106	<0.032	d
Kepler 47b	1.04	0.362		0.964	0.351	0.266	0.0836	7.45	0.0234	0.296	49.5	<0.035	e
Kepler 47c	1.04	0.362		0.964	0.351	0.411	0.0836	7.45	0.0234	0.989	303	<0.411	e
Kepler 47d	1.04	0.362		0.964	0.351		0.0836	7.45	0.0234	0.72	187.3		f
Kepler 64b	1.47	0.37		1.7	0.34	0.55	0.177	20.0	0.204	0.642	139	0.1	g
Kepler 413b	0.820	0.542	0.211	0.776	0.484	0.387	0.101	10.1	0.0365	0.355	66.3	0.118	h
KIC 963289	0.934	0.194	3×10^{-4}	0.833	0.214	0.549	0.185	27.3	0.0510	0.788	241	0.0379	i
ROXs 42B b	0.89	0.36	6-14							~140			j,k
FW Tau b	0.28	0.28	6-14							~330			j,k
Ross 458 C	0.6	0.09	6-12							150-800			k
SR12 C	1.05	0.5	12-15							~1100			k,l
DP Leonis	0.6	0.09	6.39				0.027	0.062	0	8.2	28	0.39	m

^a Doyle et al. (2011)^b Welsh et al. (2012)^c Welsh et al. (2012)^d Orosz et al. (2012b)^e Orosz et al. (2012a)^f Orosz (in prep)^g Kostov et al. (2013)^h Kostov et al. (2014)ⁱ Welsh et al. (2015)^j Kraus et al. (2014)^k Bowler et al. (2014)^l Kuzuhara et al. (2011)^m Schwöpe et al. (2002); Qian et al. (2010); Beuermann et al. (2011)

Table 5.2: Observed circumbinary planets. The binary primary and secondary masses (radii), m_1 (R_1) and m_2 (R_2), respectively, are expressed in $M_\odot$ ($R_\odot$); the circumbinary planet mass (radius), m_3 (R_3), is expressed in M_J (R_J). The binary and circumbinary orbits are denoted with A and B, respectively. The semimajor axes, a_A and a_B , are expressed in AU, and the orbital periods, P_A and P_B , are expressed in days. Note that so far, three circumbinary planets have been detected around the binary in the Kepler 47 system. Kepler 64b is also known as KIC 4862625; the binary+planet system is orbited by another stellar binary. Top rows: *Kepler* transiting circumbinary planets. Bottom rows: circumbinary planets detected through other methods.



5.B Magnetic braking in triples

As discussed in Section 5.5.6, we carried out the integrations of isolated binaries and triples as in Section 5.4.1, now also including the effects of MB for the primary and secondary stars in the inner binary by adopting equation (1) of Barker & Ogilvie (2009). The latter equation is based on the Skumanich relation (Skumanich 1972) and the results of Verbunt & Zwaan (1981); Dobbs-Dixon et al. (2004). Here, we assumed a MB parameter of $\alpha_{\text{MB}} = 1.5 \times 10^{-14} \text{ yr}$ (Barker & Ogilvie 2009).

In Fig. 5.21, we show the distribution of the initial and final inner orbital periods as in Fig. 5.7, but now with MB included. MB has a large effect on the orbital evolution. In contrast to the situation without MB, many systems merge during the evolution even for isolated binary evolution (cf. the dashed lines in Fig. 5.21). Furthermore, when including the tertiary star, no peak is produced in the final period distribution for $P_{\text{A,f}} \sim 1 - 6 \text{ d}$ as in Fig. 5.7. Instead, owing to the large number of merged systems, there is an *absence* of systems with these periods.

For the non-merging systems, we show the distributions of semimajor axis shrink factor f_{shrink} in Fig. 5.22. Although the inner binary still shrinks more in the isolated triple case compared to the isolated binary case, the enhancement in f_{shrink} is not nearly as large compared to the situation without MB (cf. Fig. 5.8). In other words, MB is so efficient at driving an inner binary merger in isolation, that KL cycles with tidal friction hardly affect the evolution.

As also discussed in Section 5.5.6, this suggests that MB is likely not as efficient as was assumed here, and this apparent discrepancy between the canonical efficiency of MB and the period distribution of solar-type MS binaries merits future investigation.

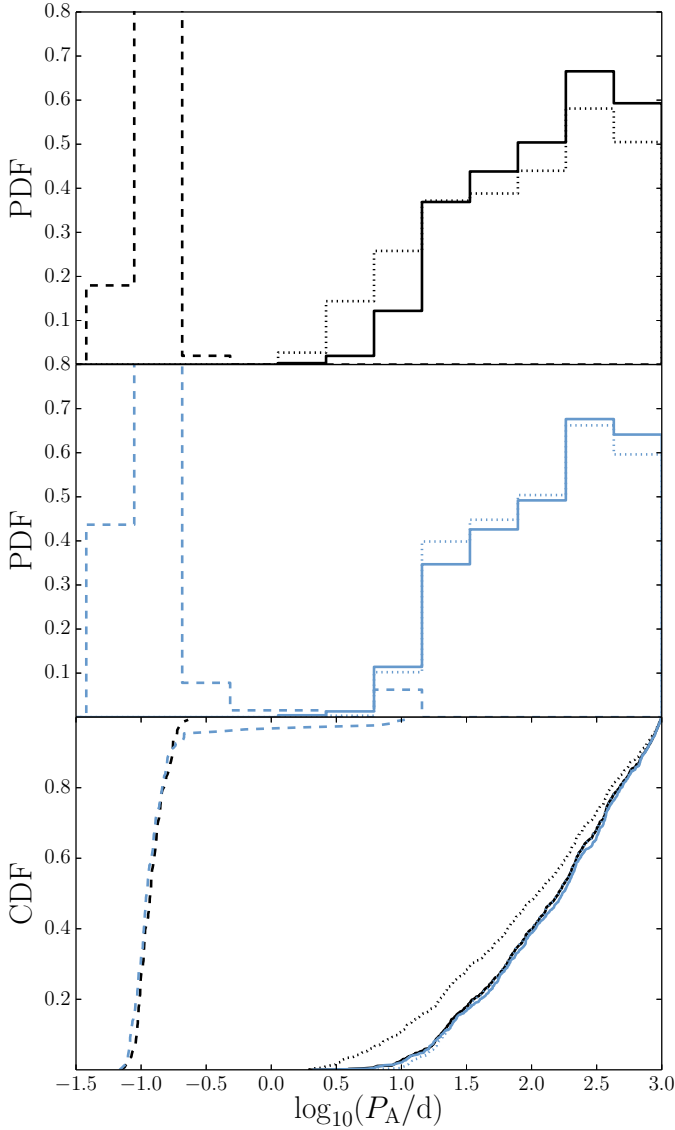
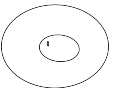


Figure 5.21: The initial (black) and final (blue) inner period distributions for the sampled triples as described in Section 5.4.1 (cf. Fig. 5.21), here also including the effects of MB as discussed in Section 5.5.6 and Appendix 5.B. Distributions corresponding to isolated binary (triple) evolution are shown with black (blue) lines. The initial distributions are shown with dotted lines. Cases when the inner binary merges are shown with dashed lines. Otherwise, the final distributions are shown with solid lines.



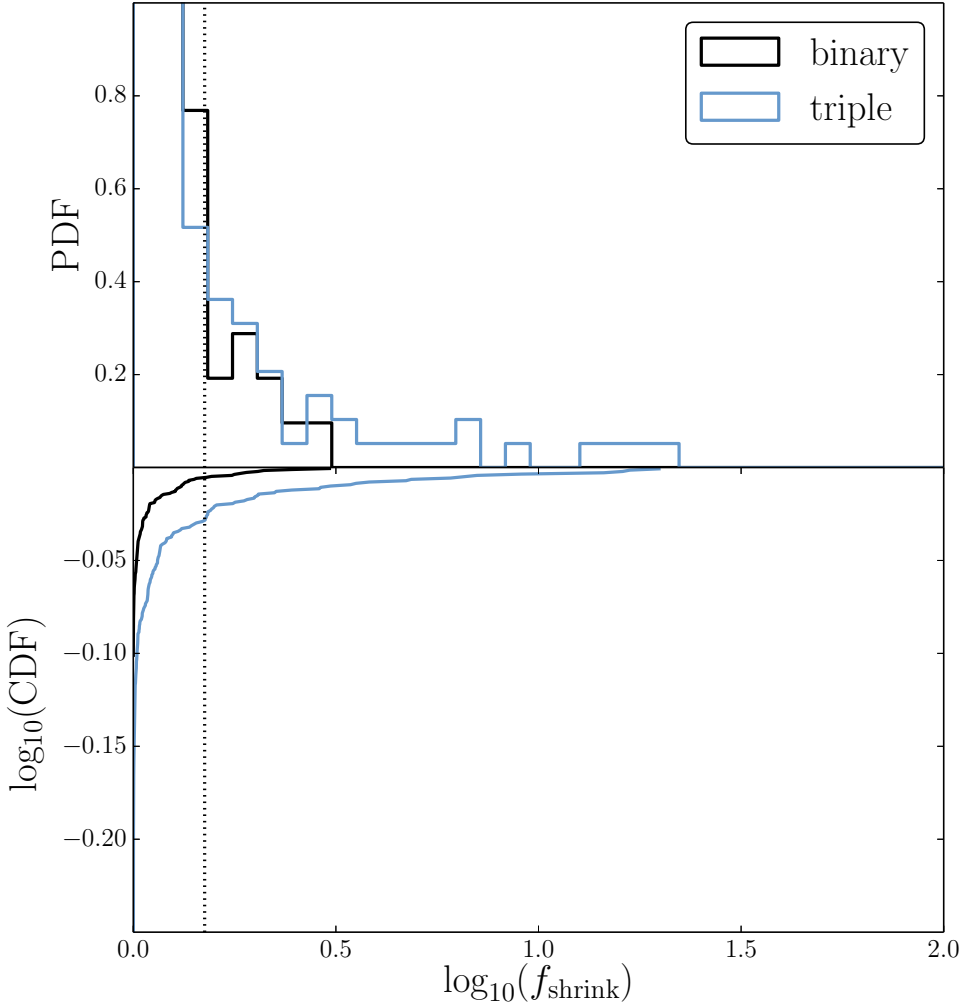
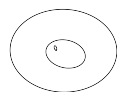


Figure 5.22: The factor $f_{\text{shrink}} \equiv a_{A,i}/a_{A,f}$ with which the inner binary period shrinks for the sampled triples as described in Section 5.4.1, here also including the effects of MB as discussed in Section 5.5.6 and Appendix 5.B. Only systems are included for which the inner binary does not merge. Black (blue) lines: isolated binary (triple) evolution. The black vertical line indicates the cutoff value chosen for our population synthesis study (cf. Section 5.4.1).

6 | Secular dynamics of hierarchical multiple systems composed of nested binaries, with an arbitrary number of bodies and arbitrary hierarchical structure

We present a method for studying the secular gravitational dynamics of hierarchical multiple systems consisting of nested binaries, which is valid for an arbitrary number of bodies and arbitrary hierarchical structure. We derive the Hamiltonian of the system and expand it in terms of the – assumed to be – small ratios x_i of binary separations. At the lowest nontrivial expansion order (quadrupole order, second order in x_i), the Hamiltonian consists of terms which, individually, depend on binary pairs. At higher orders, in addition to terms depending on binary pairs, we also find terms which, individually, depend on more than two binaries. In general, at order n in x_i , individual terms depend on at most $n - 1$ binaries. We explicitly derive the Hamiltonian including all terms up and including third order in x_i (octupole order), and including the binary pairwise terms up and including fifth order in x_i . These terms are orbit averaged, and we present a new algorithm for efficiently solving the equations of motion. This algorithm is highly suitable for studying the secular evolution of hierarchical systems with complex hierarchies, making long-term integrations of such systems feasible. We show that accurate results are obtained for multiplanet systems with semimajor axis ratios as large as ≈ 0.4 , provided that high-order terms are included. In addition to multiplanet systems with a single star, we apply our results to multistar systems with multiple planets.

Hamers, A.S. & Portegies Zwart, S.F. MNRAS 459, 3, 2827-2874 (2016)



6.1 Introduction

Stellar multiple systems are often arranged in a hierarchical configuration composed of binary orbits (Evans 1968; Nugeyre & Bouvier 1981). For most systems, these orbits are nearly Keplerian on short time-scales. The simplest and most common configuration is a hierarchical triple, in which the centre of mass of a binary is orbited by a tertiary companion. Observed triples are, by necessity, dynamically stable on time-scales of the order of the orbital periods. However, on time-scales much longer than the orbital periods, exchanges of torques can give rise to secular oscillations in the eccentricities of the orbits, i.e. Lidov-Kozai (LK) oscillations (Lidov 1962; Kozai 1962). These oscillations have important implications for e.g. short-period binaries (Mazeh & Shaham 1979; Kiseleva et al. 1998; Eggleton & Kiseleva-Eggleton 2001; Eggleton & Kiseleva-Eggleton 2006; Fabrycky & Tremaine 2007; Naoz & Fabrycky 2014).

Similar oscillations can occur in quadruple systems, which are observed in two hierarchical configurations. In the ‘2+2’ or ‘binary-binary’ configuration, two binaries orbit each other’s barycentre (e.g. Pejcha et al. 2013, who studied the secular dynamics using N -body integration methods). In the ‘3+1’ or ‘triple-single’ configuration, a single body orbits the centre of mass of a hierarchical triple (Hamers et al. 2015). Quintuple and even higher-order systems are also observed, although with lower frequency (e.g. Tokovinin 2014a,b). The number of possible configurations of hierarchical N -body systems composed of nested binaries increases with N , and is $N - 2$ for $N \geq 3$ (Nugeyre & Bouvier 1981). This suggests a large variety of high- N systems (e.g. van den Berk et al. 2007).

Even more complex hierarchies arise when considering exoplanets. Since the past decade, planets are being discovered at an exponential rate, and an increasing number of exoplanets are observed in binary, triple and even quadruple systems. In the simplest case of stellar binaries, planets have been observed orbiting an individual star (S-type orbits, in the nomenclature of Dvorak 1982), and orbiting a stellar binary (P-type orbits). Approximately 70 S-type planets in binaries have been found so far (Roell et al. 2012). Fewer circumbinary (or P-type) planets have been found; e.g., nine *Kepler* transiting circumbinary planets have been found so far (e.g. Winn & Fabrycky 2015). However, detecting such planets is generally more difficult compared to S-type planets (Martin & Triaud 2014); it has been estimated that the occurrence rate of P-type planets could be as high as the occurrence rate of exoplanets around single stars (Armstrong et al. 2014).

Although their detection is still severely limited by observational biases, exoplanets have also been found in multistellar systems with more complex hierarchies. Examples include Gliese 667, a triple system with at least two planets orbiting the tertiary star (Bonfils et al. 2013; Feroz & Hobson 2014 and references therein), and 30 Arietis, a quadruple system with an S-type planet around one of the stars (Tokovinin et al. 2006; Guenther et al. 2009; Riddle et al. 2015; Roberts et al. 2015).

These observations indicate that planet formation in stellar multiple systems is not only possible, but may also be a common phenomenon. Multiplanet systems in multistellar systems give rise to complex hierarchies with many layers of binaries, starting from the largest binary, typically a wide stellar binary with periods as long as thousands of years, and ending with tight planetary orbits with periods as short as a few days. Moons enrich the complexity of this picture even more.

So far, the secular, i.e. long-term, dynamics of high-order multiple systems have hardly been explored. This is largely due to the computational complexity of integrating, with suffi-

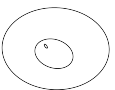
cient precision, dynamical systems with orbital periods differing by many orders of magnitude (Boekholt & Portegies Zwart 2015). An approximate, but nevertheless useful, approach is to average the Hamiltonian over individual orbits, effectively smearing out the point mass moving in each orbit into a ring. The resulting dynamics are that of interacting rings rather than point masses. The typical required time-step of the orbit-averaged equations of motion is a (sufficiently) small fraction of the secular time-scales, which are much longer than the orbital periods. This allows for much faster integration.

This method of averaging has a long history, dating back to the seminal works of Laplace and Lagrange (Lagrange 1781; Laplace 1784; see Laskar 2012 for an historic overview). In Laplace-Lagrange (LL) theory, bodies with similar masses are assumed to orbit a central massive body. The Hamiltonian is expanded in terms of the orbital eccentricities and inclinations which are assumed to be small, and non-secular terms are dropped. This method works well for planetary systems, provided that the eccentricities and inclinations are close to zero. However, on long time-scales (after thousands or more secular oscillations), secular chaotic diffusion can cause high eccentricities and inclinations, at which point the expansion is no longer accurate. For example, in the Solar system, the orbit of Mercury is unstable due to secular chaos on a time-scale of ~ 5 Gyr relative to the current Solar system (Laskar 1994; Ito & Tanikawa 2002; Laskar 2008; Laskar & Gastineau 2009). Furthermore, secular chaos can lead to the production of hot Jupiters, and, more generally, play a crucial role in the long-term evolution of planetary systems (Lithwick & Wu 2011; Wu & Lithwick 2011; Lithwick & Wu 2014). Also, in LL theory, the hierarchy is fixed, and more complex configurations such as moons around the planets, or one or more stellar companions to the central star, are not possible.

An alternative method has been developed and applied by Lidov (1962); Kozai (1962) and later by Harrington (1968); Naoz et al. (2013a) and others. In this method, which has been applied to hierarchical triples, the Hamiltonian is expanded in terms of the ratio of the inner to the outer binary separations, which is assumed to be small. A major advantage of this method is that it is valid for arbitrary eccentricities and inclinations, provided that sufficient hierarchy is maintained (i.e. the bodies in the inner and outer binaries are sufficiently separated even at the closest mutual points in their orbits).

Until recently, this approach has been limited to hierarchical triples. An extension to hierarchical quadruples was presented by Hamers et al. (2015), who derived the expanded and orbit-averaged Hamiltonian for both the ‘3+1’ and ‘2+2’ configurations. Here, we present a generalization to hierarchical multiple systems composed of binary orbits, with an arbitrary number of bodies in an arbitrary hierarchy. The main assumptions are as follows.

- The system is composed of nested binary orbits (see Figure 6.1 for an example quintuple system). Each binary has two ‘children’, which can be either bodies (i.e. point particles) or binaries. If a child is a binary, then its position vector is the centre of mass position vector of all bodies contained within that child. This excludes trapezium-type systems (e.g. Ambartsumian 1954).
- The system is sufficiently hierarchical (see Section 6.2 for a quantitative definition).
- On short time-scales, the binary orbits are well approximated by bound Kepler orbits (i.e. the time-scales for angular momentum-exchange are much longer than the orbital periods).



Although the above assumptions may appear restrictive, it turns out that our method applies to a wide range of astrophysically relevant systems. Integration of the averaged and expanded equations of motion is computationally much less intensive compared to solving the full N -body problem. Therefore, our new method opens up the possibility for studying the long-term dynamics of complex hierarchical systems, without having to resort to costly N -body integrations. Such fast integration is particularly valuable, given the large parameter space involved in these systems,

Moreover, the analytic perturbation method allows for specific effects to be associated with individual terms in the expansion. This allows for a clearer and deeper understanding of the gravitational dynamics compared to direct N -body integration, in which all effects are combined and cannot be disentangled. For example, in hierarchical triples, the octupole-order term, in contrast to the quadrupole-order term, can give rise to orbital flips between prograde and retrograde orbits, potentially leading to very high eccentricities (Lithwick & Naoz 2011; Katz et al. 2011; Naoz et al. 2013a; Teyssandier et al. 2013; Li et al. 2014a,b).

This paper is structured as follows. In Section 6.2, we summarize our derivation and our results of the generalized Hamiltonian for hierarchical multiple systems. A detailed and self-contained derivation is given in Appendix 6.A. We describe a new numerical algorithm to solve the orbit-averaged equations of motion. In the subsequent sections, Sections 6.3 and 6.4, we demonstrate the validity of the method and the algorithm and the potential of their applications, focusing on multiplanet systems in stellar multiple systems. We discuss our method in Section 6.5, and we conclude in Section 6.6.

6.2 The generalized Hamiltonian for hierarchical multiple systems

6.2.1 Definition and description of the system

We consider hierarchical multiple systems that are composed of nested binaries and which can be represented by ‘simplex mobile’ diagrams, introduced by Evans (1968). We note that a similar hierarchical decomposition is made in the KIRA integrator in STARLAB (appendix B of Portegies Zwart et al. 2001). An example for a hierarchical quintuple system ($N = 5$) is given in Figure 6.1. The binaries have two ‘children’, which can be either bodies, or binaries themselves. Conversely, each binary or body has a ‘parent’, with the exception of the unique binary from which all other binaries and bodies originate, i.e. the ‘top’ binary. Each binary can be assigned a ‘level’, defined as the number of binaries that need to be traversed in the mobile diagram to reach that binary, counting from the top binary.

Here, we consider only multiple systems with the ‘simplex’ structure, and not the more general ‘multiplex’ structure, in which members of the system are allowed to have more than two children (Evans 1968). Note, however, that the structure of most observed hierarchical multiple systems is of the simplex type (e.g. Tokovinin 2014a,b). This is also expected on the basis of the requirement of dynamical stability on long time-scales (Nugeyre & Bouvier 1981).

Mathematically, the structure of the system can be specified in terms of a ‘mass ratio matrix’ $\mathbf{A}$. The components of this $N \times N$ -matrix are mass ratios of bodies within the system. The matrix $\mathbf{A}$ is defined such that it gives the relation between the absolute position vectors $\mathbf{R}_i$ of bodies in the system, and the relative binary separation vectors $\mathbf{r}_i$ (see also Figure 6.1). This description

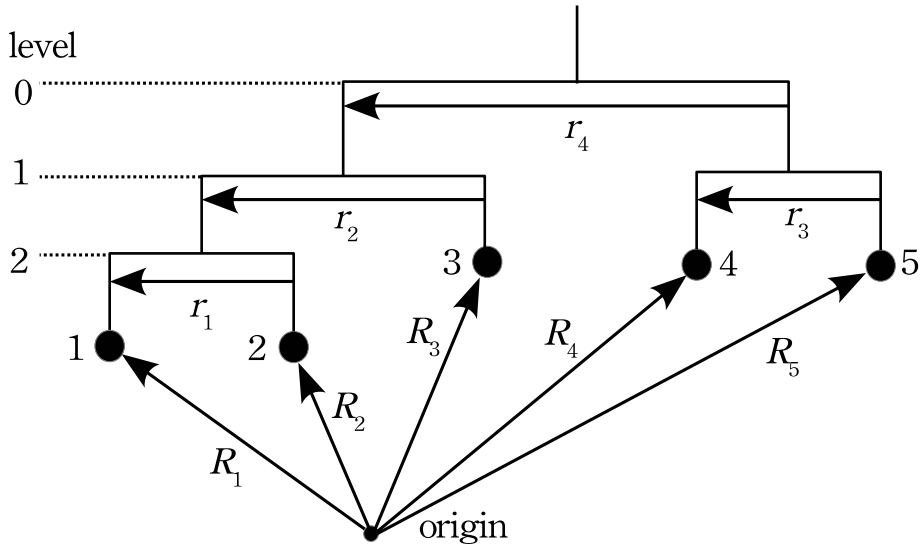
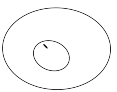


Figure 6.1: A schematic representation of a hierarchical quintuple in a mobile diagram (Evans 1968). Absolute position vectors of bodies with respect to an arbitrary and fixed origin are denoted with capital letters, i.e. with $\mathbf{R}_i$. Binary separation vectors are denoted with small letters, i.e. with $\mathbf{r}_i$. Note that other choices for the directions of the relative vectors are also possible; e.g., $\mathbf{r}_1$ can also be defined as $\mathbf{r}_1 = \mathbf{R}_2 - \mathbf{R}_1$, i.e. pointing from body 1 towards body 2. Consequently, some of the quantities B_{ijk} will change sign (cf. Appendix 6.A.1).



is similar to that used by Nugeyre & Bouvier (1981); Walker (1983); Abad & Docobo (1988). For a precise definition of $\mathbf{A}$ and examples, we refer to Appendix 6.A.1.

6.2.2 The expanded Hamiltonian

In Appendix 6.A, we derive, from first principles, the general (Newtonian) Hamiltonian for hierarchical multiple systems as defined above, and we expand it in terms of ratios x_i of separations of binaries at different levels. These ratios are assumed to be small, i.e. we assume that the system is hierarchical. The general expression is comprised of summations over both binaries and bodies, and is therefore not very useful for our purposes. Fortunately, it is possible to rewrite these summations over both binaries and bodies to summations over only binaries, and we carry out this rewriting up to and including fifth order in x_i , where for the highest two orders, x_i^4 and x_i^5 , we only include binary pair terms. Our main results, at the various orders, are as follows.

1. The terms in the Hamiltonian at the lowest order, i.e. first order in x_i ('dipole order'), vanish identically for *any* hierarchical configuration.
2. The Hamiltonian at second order in x_i ('quadrupole order') consists of terms which, individually, depend on binary *pairs*. For each pair, the binaries must be on different levels, and they must also be connected to each other through their children, i.e. the binary with the highest level must be a descendant of the binary with the lowest level. For example, in Figure 6.1, the included binary pairs are (r_1, r_4) , (r_2, r_4) , (r_3, r_4) and (r_1, r_2) , whereas pairs (r_2, r_3) and (r_1, r_3) are excluded. The individual binary pair terms are mathematically equivalent to the quadrupole-order terms of the hierarchical three-body Hamiltonian.
3. At the 'octupole order', i.e. third order in x_i , the Hamiltonian consists of two types of terms.
 - (a) Terms which, individually, depend on binary *pairs*, with the binaries at different levels. The pairing occurs similarly as for the quadrupole-order terms. Mathematically, the pairwise terms are equivalent to the octupole-order terms of the hierarchical three-body Hamiltonian.
 - (b) Terms which, individually, depend on binary *triplets*, with the binaries at different levels. The triplets to be included are those for which the binary p with the highest level is a 'descendent' of the binary with the intermediate level, u , and for which u is a descendent of the binary k with the lowest level (cf. equation 6.129). In other words, all three binaries must be part of the same branch. In Figure 6.1, there is only one triplet term, which applies to the binary triplet (r_1, r_2, r_4) .

The mathematical form of the individual triplet terms is equivalent to the term depending on three binaries at the octupole order in the hierarchical four-body Hamiltonian ('3+1' configuration; this term was first derived by Hamers et al. 2015, where it was referred to as a 'cross term'). In practice, at least for the type of systems considered in this paper, the triplet term can be safely neglected compared to the binary pair terms, even those of higher orders.

4. At the ‘hexadecupole order’, i.e. fourth order in x_i , the Hamiltonian consists of three types of terms which depend on binary pairs, triplets and quadlets, respectively. Here, we have restricted to explicitly deriving the binary pair terms only.
5. At the ‘dotriacontupole order’, i.e. fifth order in x_i , the Hamiltonian consists of four types of terms which depend on binary pairs, triplets, quadlets and quintlets, respectively. As for the hexadecupole order, we have exclusively derived the binary pair terms explicitly.
6. Generally, at order n in x_i , there are $n - 1$ types of terms. The individual terms depend on at most n binaries. Whether all these terms appear is contingent on the system: if the system does not contain a sufficiently complex hierarchy, then not all different binary terms may appear. For example, in the case of a hierarchical triple, only the binary pair terms appear; a summation with three or more different binaries at high orders does not apply.

In summary, the general hierarchical N -body Hamiltonian can be constructed from the Hamiltonians of smaller subsystems. The higher the order, the larger the number of binaries within these subsystems. To lowest nontrivial order, i.e. the quadrupole order, the N -body Hamiltonian consists entirely of combinations of three-body (i.e. binary pair) Hamiltonians. In other words, at this order, only the interactions between binary pairs are included.

As we show below (cf. Section 6.3), the inclusion of high-order binary pair terms is necessary particularly in planetary systems, where the separation ratios may not be $\ll 1$. For example, in the Solar system (including the planets plus Pluto), the mean of the adjacent ratios of semi-major axes (i.e. $a_{\text{Mercury}}/a_{\text{Venus}}$, $a_{\text{Venus}}/a_{\text{Earth}}$, etc.) is ≈ 0.58 . Good results for this system are obtained only when terms of high orders are included.

6.2.3 Orbit averaging

After expanding the Hamiltonian in terms of the x_i , we orbit average the Hamiltonian assuming that the motion in each binary on suborbital time-scales is exactly Keplerian (cf. Appendix 6.A.5). We employ a vector formalism, where each binary orbit is described by its eccentricity vector, $\mathbf{e}_k$, and $\mathbf{j}_k$. Here, $\mathbf{j}_k$ is the specific angular momentum vector $\mathbf{h}_k$ normalized to the angular momentum of a circular orbit Λ_k , i.e.

$$\mathbf{j}_k \equiv \mathbf{h}_k / \Lambda_k, \quad (6.1)$$

where

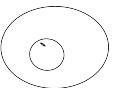
$$\mathbf{h}_k \equiv \mathbf{r}_k \times \dot{\mathbf{r}}_k \quad (6.2)$$

and

$$\Lambda_k = M_{k.C1} M_{k.C2} \sqrt{\frac{G a_k}{M_k}}. \quad (6.3)$$

We define the mass of binary k , M_k , as the sum of the masses of *all bodies* contained within the hierarchy of binary k . The masses of the two children of binary k are denoted with $M_{k.C1}$ and $M_{k.C2}$. Evidently, the latter satisfy

$$M_{k.C1} + M_{k.C2} = M_k. \quad (6.4)$$



For example, in Figure 6.1, $M_2 = m_1 + m_2 + m_3$, $M_{2,C1} = m_1 + m_2$ and $M_{2,C2} = m_3$. Of course, the choice of which child of binary 2 is ‘child 1’ or ‘child 2’ is arbitrary; the alternative, and equally valid, choice is $M_{2,C1} = m_3$ and $M_{2,C2} = m_1 + m_2$.

As described in more detail in Appendix 6.A.5, the averaging procedure (cf. equation 6.133) is not a canonical transformation. However, a transformation of the ‘coordinates’ e_k and j_k can be found, which leads to a transformed Hamiltonian that is equivalent to the old Hamiltonian, and amounts to averaging the Hamiltonian over the orbits (Naoz et al. 2013a). The transformed coordinates differ from the original ones. However, as noted by Naoz et al. (2013a), the untransformed and the transformed coordinates differ by order x_i^2 . Therefore, this difference can usually be neglected. This is borne out by tests of our algorithm with other methods (cf. Section 6.3).

Nevertheless, it is important to keep in mind that the orbit-averaged approximation can break down in situations where the time-scale for changes in the (secular) orbital parameters (in particular, the eccentricities) is shorter than the orbital period (Antonini & Perets 2012; Antonini et al. 2014; Antognini et al. 2014). In our numerical algorithm (cf. Section 6.2.5), we check for this condition during the integrations. In the context of planetary systems, the orbit-averaged approximation breaks down when the planets are sufficiently closely spaced for mean motion resonances to be important, which are not taken into account.

We have explicitly derived all orbit-averaged terms up and including octupole order, and including the pairwise hexadecupole- and dotriacontupole-order terms (cf. Appendix 6.A.5). Furthermore, using complex integration techniques, we have derived a general expression for the averaged pairwise Hamiltonian to order n in terms of quickly-evaluated derivatives of polynomial functions, rather than integrals (cf. Appendix 6.A.5).

6.2.4 General implications

Quadrupole order

An immediate implication of Section 6.2.2 is that to quadrupole order, the qualitative behaviour of the system can be characterized using simple arguments based on the LK time-scales of binary pairs, analogously to the ratio of LK time-scales that was considered for quadruple systems in Hamers et al. (2015).

An order-of-magnitude estimate of the LK time-scale $P_{\text{LK},pk}$, τ_{pk} , for binaries p and k with p a child of k , can be obtained by dividing Λ_p by the associated order-of-magnitude quadrupole-order term in the orbit-averaged Hamiltonian (cf. equations 6.17 and 6.144). This gives

$$\begin{aligned} \tau_{pk} &\equiv \Lambda_p \frac{M_p}{M_{p,C1} M_{p,C2}} \frac{a_k}{GM_{k,CS(p)}} \left(\frac{a_k}{a_p} \right)^2 (1 - e_k^2)^{3/2} \\ &= \frac{1}{2\pi} \frac{P_{\text{orb},k}^2}{P_{\text{orb},p}} \frac{M_k}{M_{k,CS(p)}} (1 - e_k^2)^{3/2}, \end{aligned} \quad (6.5)$$

where we used equation (6.3), and the orbital periods $P_{\text{orb},k}$ are defined through Kepler’s law,

$$P_{\text{orb},k} = 2\pi \sqrt{\frac{a_k^3}{GM_k}}. \quad (6.6)$$

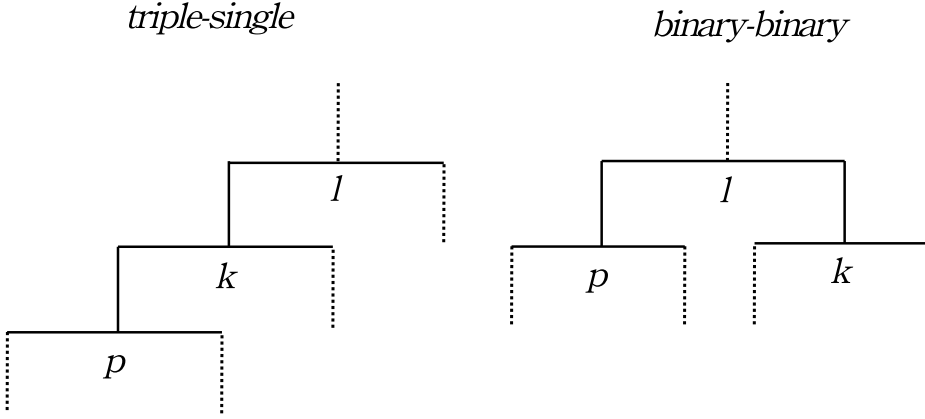


Figure 6.2: Left: mobile diagram of a ‘fully nested’ subsystem consisting of three binaries p , k and l for which p is a child of k , and k is a child of l , used to define the quantity $\mathcal{R}_{pkl}^{(\text{fn})}$ (cf. equation 6.9). Right: mobile diagram of a ‘binary-binary’ subsystem consisting of three binaries p , k and l for which p is a child of l , and k is a child of l , used to define the quantity $\mathcal{R}_{pkl}^{(\text{bb})}$ (cf. equation 6.10).

In the case of isolated triples, the precise LK time-scale to quadrupole order (and with the further assumption of highly hierarchical systems) also depends on the initial e_p and ω_p (i.e. the inner orbit eccentricity and argument of pericentre), as well as the relative inclination i_{pk} between the inner and outer orbits (Kinoshita & Nakai 2007). Formally, we write the LK time-scale for hierarchical N -body systems as

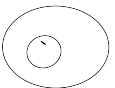
$$P_{\text{LK},pk} = C(e_p, \omega_p, i_{pk}) \frac{P_{\text{orb},k}^2}{P_{\text{orb},p}} \frac{M_k}{M_{k,\text{CS}(p)}} (1 - e_k^2)^{3/2}. \quad (6.7)$$

Here, the dimensionless factor $C(e_p, \omega_p, i_{pk})$ captures any dependence on e_p , ω_p and i_{pk} ; generally, C is of order unity. For example, for a hierarchical triple system with inner binary masses m_1 and m_2 and tertiary mass m_3 , equation (6.7) reduces to the familiar expression (e.g. Innanen et al. 1997; Kiseleva et al. 1998; Antognini 2015)

$$P_{\text{LK};\text{in},\text{out}} = C(e_{\text{in}}, \omega_{\text{in}}, i_{\text{rel}}) \frac{P_{\text{orb},\text{out}}^2}{P_{\text{orb},\text{in}}} \frac{m_1 + m_2 + m_3}{m_3} (1 - e_{\text{out}}^2)^{3/2}. \quad (6.8)$$

The LK time-scale $P_{\text{LK},pk}$ is a proxy for the strength of the secular torque of orbit k on orbit p . Therefore, ratios of LK time-scales applied to different pairs of binaries give a measure for the relative importance of the secular torques.

In the case of three adjacent binaries, p , k and l , for which p is a child of k , and k is a child



of l (i.e. a ‘fully nested’ subsystem, see the left part of Figure 6.2), we consider the ratio

$$\begin{aligned}\mathcal{R}_{pkl}^{(\text{fn})} &\equiv \frac{P_{\text{LK},pk}}{P_{\text{LK},kl}} \\ &= \frac{C(e_p, \omega_p, i_{pk})}{C(e_k, \omega_k, i_{kl})} \frac{P_{\text{orb},k}^3}{P_{\text{orb},p} P_{\text{orb},l}^2} \frac{M_k}{M_l} \frac{M_{l,\text{CS}(k)}}{M_{k,\text{CS}(p)}} \left(\frac{1 - e_k^2}{1 - e_l^2} \right)^{3/2} \\ &= \frac{C(e_p, \omega_p, i_{pk})}{C(e_k, \omega_k, i_{kl})} \left(\frac{a_k^3}{a_p a_l^2} \right)^{3/2} \left(\frac{M_p}{M_k} \right)^{1/2} \frac{M_{l,\text{CS}(k)}}{M_{k,\text{CS}(p)}} \left(\frac{1 - e_k^2}{1 - e_l^2} \right)^{3/2},\end{aligned}\quad (6.9)$$

where in the last line, the expression was written in terms of semimajor axes. Note that for quadruple systems, equation (6.9) reduces to equation (13) of Hamers et al. (2015).

Analogously to the case of fully nested hierarchical quadruple systems (Hamers et al. 2015), the following qualitative behaviour applies.

1. $\mathcal{R}_{pkl}^{(\text{fn})} \ll 1$: the torque of k on p dominates compared to the torque of l on k . LK oscillations can be induced by k on p . Furthermore, induced precession on k from p can quench LK oscillations in k otherwise induced by l .
2. $\mathcal{R}_{pkl}^{(\text{fn})} \gg 1$: the torque of l on k dominates compared to the torque of k on p . LK oscillations can be induced by l on k . Effectively, p can be approximated as a point mass.
3. $\mathcal{R}_{pkl}^{(\text{fn})} \sim 1$: complex, nonregular LK cycles can be induced in p ; chaotic evolution is possible.

In the case of three binaries, p , k and l , for which p is a child of l , and k is a child of l (i.e. a ‘binary-binary’ subsystem, see the right part of Figure 6.2), we consider the ratio

$$\begin{aligned}\mathcal{R}_{pkl}^{(\text{bb})} &\equiv \frac{P_{\text{LK},pl}}{P_{\text{LK},kl}} = \frac{C(e_p, \omega_p, i_{pl})}{C(e_k, \omega_k, i_{kl})} \frac{P_{\text{orb},k}}{P_{\text{orb},p}} \frac{M_p}{M_k} \\ &= \frac{C(e_p, \omega_p, i_{pl})}{C(e_k, \omega_k, i_{kl})} \left(\frac{a_k}{a_p} \right)^{3/2} \left(\frac{M_p}{M_k} \right)^{3/2}\end{aligned}\quad (6.10)$$

(note that $M_{l,\text{CS}(k)} = M_p$ and $M_{l,\text{CS}(p)} = M_k$).

In this case, the following characteristics apply.

1. $\mathcal{R}_{pkl}^{(\text{bb})} \ll 1$: the torque of l on p dominates compared to the torque of l on k . For sufficiently small values of $\mathcal{R}_{pkl}^{(\text{bb})}$, k can effectively be treated as a point mass.
2. $\mathcal{R}_{pkl}^{(\text{bb})} \gg 1$: the torque of l on k dominates compared to the torque of l on p . For sufficiently large values of $\mathcal{R}_{pkl}^{(\text{bb})}$, p can effectively be treated as a point mass.
3. $\mathcal{R}_{pkl}^{(\text{bb})} \sim 1$: complex, nonregular LK cycles can be induced in p and k (Pejcha et al. 2013).

Higher orders

In hierarchical triple systems, the octupole-order term gives rise to generally more complex, and potentially chaotic eccentricity oscillations compared to when including only the quadrupole-order term, especially if the semimajor axis ratio is small and/or the outer orbit eccentricity is high. Orbital flips can occur between prograde and retrograde orbits, potentially leading to very high eccentricities (Lithwick & Naoz 2011; Katz et al. 2011; Naoz et al. 2013a; Teyssandier et al. 2013; Li et al. 2014a,b). The importance of these effects can be evaluated using the octupole parameter, which is essentially the ratio of the order-of-magnitude estimate of the orbit-averaged octupole-order term to the quadrupole-order term. It is given by (Naoz et al. 2013a)

$$\epsilon_{\text{oct, triple}} = \frac{|m_1 - m_2|}{m_1 + m_2} \frac{a_{\text{in}}}{a_{\text{out}}} \frac{e_{\text{out}}}{1 - e_{\text{out}}^2}. \quad (6.11)$$

In the case of hierarchical N -body systems, the situation is more complicated because there will generally be more than one pair of binaries, and because at higher orders (starting with the octupole order), terms appear which individually depend on more than two binaries (cf. Section 6.2.2).

In the applications considered in this paper, the octupole-order triplet terms are found to be negligible compared to the octupole-order binary pair terms, and binary pair terms at higher orders (e.g. Section 6.3.1). This can be understood by estimating the ratio of the orbit-averaged octupole-order binary triplet term, $\bar{S}'_{3,3}$, to the orbit-averaged octupole-order binary pair term, $\bar{S}'_{3,2}$ (cf. equation 6.145). Consider three binaries p , u and k , where $p \in \{k.C\}$, $p \in \{u.C\}$ and $u \in \{k.C\}$ (i.e. a connected binary triplet in which each binary is on a different level). A distinction should be made when evaluating $\bar{S}'_{3,2}$: we apply it to (p, u) and (u, k) (we do not consider (p, k) because by assumption, $a_p/a_u \gg a_p/a_k$). In the case of (p, u) , we find

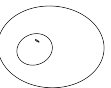
$$\left| \frac{\bar{S}'_{3,3}}{\bar{S}'_{3,2}} \right|_{(p,u)} \sim \frac{e_k}{e_p} \frac{M_{k,\text{CS}(p)}}{|M_{p,C1} - M_{p,C2}|} \frac{M_p}{M_u} \left(\frac{a_u}{a_k} \right)^4 \left(\frac{a_u}{a_p} \right) \left(\frac{1 - e_u^2}{1 - e_k^2} \right)^{7/2}. \quad (6.12)$$

In the case of (u, k) ,

$$\left| \frac{\bar{S}'_{3,3}}{\bar{S}'_{3,2}} \right|_{(u,k)} \sim \frac{M_{u,\text{CS}(p)}}{|M_{u,C1} - M_{u,C2}|} \frac{M_{p,C1} M_{p,C2}}{M_{u,C1} M_{u,C2}} \frac{M_u}{M_p} \left(\frac{a_p}{a_u} \right)^2. \quad (6.13)$$

Note that equations (6.12) and (6.13) diverge as $M_{p,C1} \rightarrow M_{p,C2}$, $M_{u,C1} \rightarrow M_{u,C2}$ and/or $e_p \rightarrow 0$. In those cases, the corresponding term $\bar{S}'_{3,2}$ vanishes. However, additionally, the binary pair hexadecupole-order term does not vanish (cf. equation 6.146). Therefore, one should compare to $\bar{S}'_{4,2}$ (this is not done here).

In the case of a planetary system with order Jupiter-mass planets around an order solar-mass star (with the star in the innermost orbit p), $M_{k,\text{CS}(p)} \sim 10^{-3} M_\odot$, $|M_{p,C1} - M_{p,C2}| \sim 1 M_\odot$, $M_p \sim M_u \sim 1 M_\odot$, $a_u \ll a_k$ and $a_u \gg a_p$ with $a_p/a_u \sim a_u/a_k$. Therefore, the mass ratios in equation (6.12) are of the order of 10^{-3} , whereas the semimajor axis ratio is $\ll 1$. Assuming eccentricities close to zero, this shows that the binary triplet term is negligible compared to the binary pair term in the case of (p, u) . Similarly, $M_{u,\text{CS}(p)} \sim 10^{-3} M_\odot$, $|M_{u,C1} - M_{u,C2}| \sim$



$1 M_{\odot}$ and $M_{p.C1} M_{p.C2} \sim M_{u.C1} M_{u.C2} \sim 10^{-3} M_{\odot}^2$, showing that a similar conclusion applies in the case of (u, k) .

In Appendix 6.A.6, we generalize these estimates for nested planetary systems of the importance of terms other than binary pair terms to arbitrary order n .

When the non-pairwise binary terms are negligible, it suffices to consider the ratios of binary pair terms at sequential orders. We consider the ratio of the order-of-magnitude orbit-averaged binary pair term at order $n+1$, to that of order n . Using the general expression equation (6.162), we find

$$\left| \frac{\overline{S}'_{n+1;2}}{\overline{S}'_{n;2}} \right|_{(p,k)} \sim \left| \frac{M_{p.C2}^n - (-1)^n M_{p.C1}^n}{M_p M_{p.C2}^{n-1} + (-1)^n M_p M_{p.C1}^{n-1}} \right| \frac{a_p}{a_k} \frac{f_n(e_k)}{1 - e_k^2}, \quad (6.14)$$

where $f_n(e_k)$ is a function of e_k . Based on the expressions in equations (6.144), (6.145), (6.146) and (6.147), i.e. for $n \in \{2, 3, 4, 5\}$,

$$f_n(e_k) = \begin{cases} e_k, & n \text{ even}; \\ 1/e_k, & n \text{ odd}, \end{cases} \quad (6.15)$$

and we expect that this applies to any $n \geq 2$. For $n = 2$, equation (6.14) reads

$$\left| \frac{\overline{S}'_{3;2}}{\overline{S}'_{2;2}} \right| \sim \frac{|M_{p.C1} - M_{p.C2}|}{M_{p.C1} + M_{p.C2}} \frac{a_p}{a_k} \frac{e_k}{1 - e_k^2}, \quad (6.16)$$

which reduces to equation (6.11) in the case of a hierarchical triple.

Binary pair terms to high order turn out to be important in multiplanet systems. This is investigated in detail in Section 6.3.1.

6.2.5 Numerical algorithm

We have developed a numerical algorithm written in C++, `SECULARMULTIPLE`, to efficiently integrate the equations of motion for hierarchical multiple systems based on the formalism described in Section 6.2, and presented in detail in Appendix 6.A¹. The equations of motion are derived by taking gradients of the orbit-averaged Hamiltonian in terms of the orbital vectors, $\overline{H}$. Explicitly, the equations of motion are given by the Milankovitch equations (Milankovitch 1939, e.g. Musen 1961; Allan & Ward 1963; Allan & Cook 1964; Breiter & Ratajczak 2005; Tremaine et al. 2009; see Rosengren & Scheeres 2014 for a recent overview), i.e.

$$\frac{d\mathbf{j}_k}{dt} = -\frac{1}{\Lambda_k} [\mathbf{j}_k \times \nabla_{\mathbf{j}_k} \overline{H} + \mathbf{e}_k \times \nabla_{\mathbf{e}_k} \overline{H}]; \quad (6.17a)$$

$$\frac{d\mathbf{e}_k}{dt} = -\frac{1}{\Lambda_k} [\mathbf{e}_k \times \nabla_{\mathbf{j}_k} \overline{H} + \mathbf{j}_k \times \nabla_{\mathbf{e}_k} \overline{H}]. \quad (6.17b)$$

Here, Λ_k is the angular momentum of a circular orbit given by equation (6.3). The system of first-order differential equations for the sets $(\mathbf{j}_k, \mathbf{e}_k)$ for all binaries is solved using the `CVODE` library (Cohen et al. 1996).

¹This algorithm is a generalization of the algorithm that was presented in Hamers et al. (2015).

Our algorithm is interfaced in the `AMUSE` framework (Portegies Zwart et al. 2013; Pelupessy et al. 2013), and uses the `AMUSE` particle datamodel for the user to specify the hierarchical structure and associated parameters in the high-level `PYTHON` language. Binaries and bodies are both members of a single particle set; the type of particle (body or binary) is specified with a Boolean parameter. The system structure is defined by linking the two children of each binary to other members of the same particle set. We have added support for easily switching between the secular approach and direct N -body integration with any of the direct N -body codes available in `AMUSE`.

The `SECULARMULTIPLE` algorithm also supports the inclusion of additional forces, including post-Newtonian (PN) terms to the 1PN and 2.5PN orders, and tidal bulges and tidal friction. The PN terms are implemented by treating the binaries in the system as being isolated, and adding the relevant equations for $\dot{e}_k$ and $\dot{h}_k$. Here, we neglect the contribution of any additional terms which may apply. For example, Naoz et al. (2013b) showed that for hierarchical triples and at the 1PN order, an additional term (an ‘interaction term’) appears that is associated with both inner and outer orbits. Here, we only include the 1PN terms associated with individual binaries (cf. Appendix 6.A.7).

To model the effects of tidal bulges and tidal friction, we adopt the equilibrium tide model of Eggleton et al. (1998) with a constant tidal time lag τ (or, equivalently, a constant viscous time-scale t_V). For tidal friction, we adopted the equations derived by Barker & Ogilvie (2009) which are well defined in the limit $e_k \rightarrow 0$. We implement the equations in binaries where at least one of the children is a body. Here, we treat the ‘companion’ as a point mass, even if the companion is, in reality, a binary. A self-consistent treatment of tides in hierarchical multiple systems is beyond the scope of this work.

As mentioned in Section 6.2.3, the orbit-averaged approximation breaks down if the time-scale for changes of the angular momentum j_k is smaller than the orbital time-scale (Antonini & Perets 2012; Antonini et al. 2014). In `SECULARMULTIPLE`, we allow the user to check for this condition in any binary k at any time in the integration. Using a root-finding algorithm, the integration is stopped whenever $t_{j,k} \leq P_{\text{orb},k}$, where $P_{\text{orb},k}$ is the orbital period of binary k (cf. equation 6.6) and $t_{j,k}$ is the time-scale for the angular momentum of binary k to change by order itself. The latter is given by

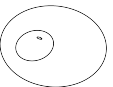
$$t_{j,k} = \left| \frac{1}{j_k} \frac{dj_k}{dt} \right|^{-1} = \left| \frac{e_k}{1 - e_k^2} \frac{de_k}{dt} \right|^{-1}, \quad (6.18)$$

where the time derivatives are obtained numerically from the equations of motion (cf. equations 6.17).

6.3 Tests: S-type multiplanet systems in single and multiple stellar systems

6.3.1 Single-star systems

As a first test and application, we consider single-star multiplanet systems with planets in orbits with small eccentricities and inclinations. This allows for comparison of `SECULARMULTIPLE` with LL theory at relatively low order, i.e. to second order in the eccentricities and the inclinations.



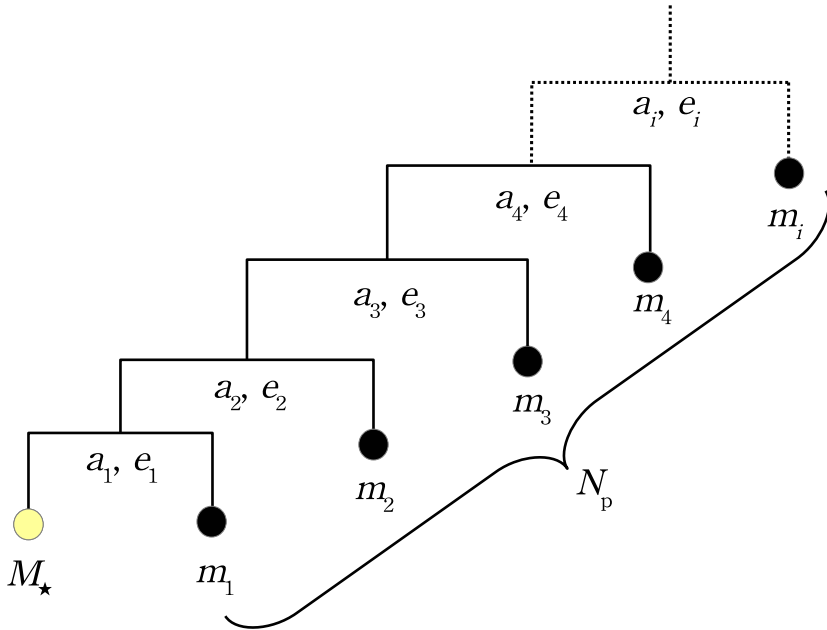


Figure 6.3: A mobile diagram of a hierarchical multiplanet system with a single star (mass M_*). In the framework of the formalism presented in Appendix 6.A, a multiplanet system can be represented as a ‘fully nested’ system. Each planet is assumed to orbit around the centre of mass position of the central star plus the planets contained within its orbit. In practice, the latter position nearly coincides with that of the central star. The planetary masses are m_i , $1 \leq i \leq N_p$, where N_p is the number of planets. The same subindices are used to denote the orbital parameters.

The hierarchy of the system is a ‘fully nested’ configuration, and is depicted schematically in Figure 6.3. Note that in the framework of the formalism presented in Appendix 6.A, a multiplanet system can be represented as system in which each planet is assumed to orbit around the centre of mass of the central star plus the planets contained within its orbit. We consider various planet numbers N_p . The mass of the star is assumed to be $M_\star = 1 M_\odot$. For simplicity, the planetary masses, m_i with $1 \leq i \leq N_p$, are set equal to $m_i = 1 M_J$.

We consider a sequence of systems with different spacing between the planets. The spacing is quantified in terms of K_{ij} , which applies to pairs $(i, j) = (i, i + 1)$ of adjacent planets, and is defined as

$$K_{ij} \equiv \frac{a_j(1 - e_j) - a_i(1 + e_i)}{R_{H;ij}}, \quad (6.19)$$

where $R_{H;ij}$ is the mutual Hill radius, defined by

$$R_{H;ij} \equiv \frac{a_i + a_j}{2C_{H;ij}}; \quad C_{H;ij} \equiv \left(\frac{3M_\star}{m_i + m_j} \right)^{1/3}. \quad (6.20)$$

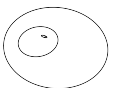
For simplicity, we assume a single initial spacing $K_{ij} = K_0$ for adjacent planets. Furthermore, we choose a single initial eccentricity of $e_i = 0.01$. With these assumptions and with the innermost semimajor axis a_1 specified, equation (6.19) defines the semimajor axes of an N_p -planet system with constant semimajor axis ratios. For reference, we show in Figure 6.4 the relation between the semimajor axis ratio a_i/a_j and K_0 according to equation (6.19).

The initial inclinations are $i_i = 1^\circ$, where the i_i are measured with respect to an arbitrary inertial reference frame. The longitudes of the ascending node, Ω_i , are chosen randomly from a uniform distribution. Note that this implies that the orbits of the planets are (slightly) mutually inclined. The arguments of pericentre ω_i are also sampled randomly from a uniform distribution. Because the latter mainly affect the phases of the secular oscillations in the orbital elements, only a single realization of ω_i is taken for each sequence in K_0 . For Ω_i , five different values are sampled for each K_0 , and results are averaged over the integrations with different Ω_i .

For the fixed parameters, we integrated a sequence of systems with different $K_{ij} = K_0$, Ω_i and N_p . In Figure 6.5, we show an example for $K_0 = 12$ and $N_p = 4$; the lines show the eccentricities (top panel) and the inclinations (bottom panel) of the planets as a function of time. Solid lines are according to `SECULARMULTIPLE` with the inclusion of all terms up and including octupole order, plus the hexadecupole and dotriacontupole binary pair terms. Dashed lines are according to second-order LL theory, computed using the equations given in Murray & Dermott (1999). Note that the latter are valid for any semimajor axis ratio and hence K_0 , provided that the secular approximation is valid (i.e. mean motion resonances are not important). However, because an expansion is made in the eccentricities and inclinations, the latter variables should remain small. For the value of $K_0 = 12$ in Figure 6.5, `SECULARMULTIPLE` is in good agreement with LL theory.

Such agreement is contingent on the inclusion of sufficient high orders in `SECULARMULTIPLE`. To illustrate this, we show in Figure 6.6 the same setup as in Figure 6.5, but now with the solid lines showing the elements of the innermost planet according to `SECULARMULTIPLE` with the inclusion of various terms. We consider the following combinations:

- quadrupole-order terms and octupole-order binary pair terms (i.e. without the octupole-order binary triplet terms);



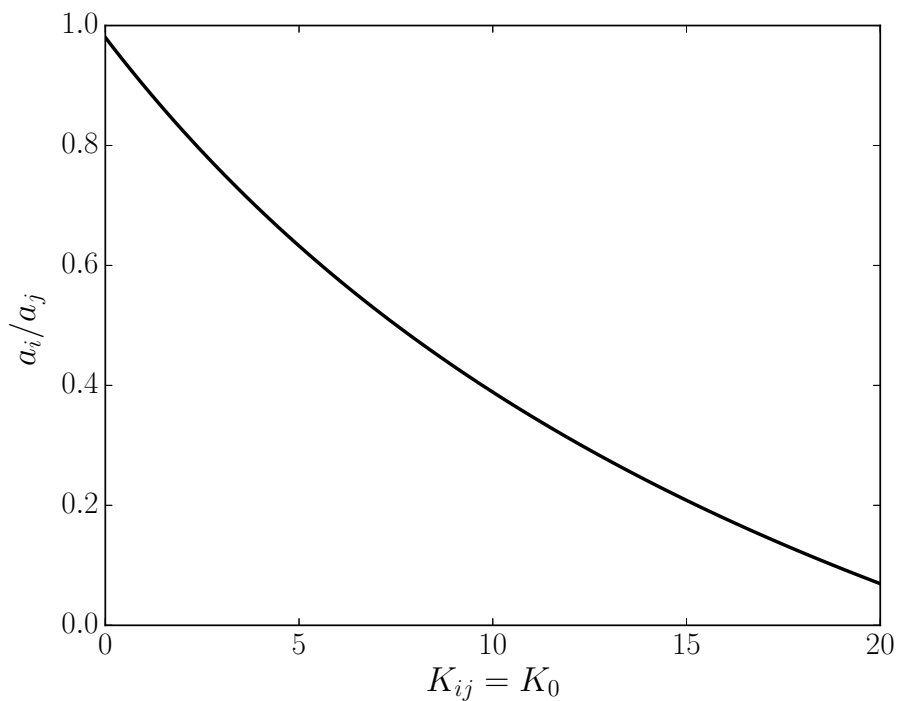


Figure 6.4: The relation between the semimajor axis ratio a_i/a_j and $K_{ij} = K_0$ according to equation (6.19). Assumed parameters are $M_\star = 1 M_\odot$, $m_i = 1 M_J$ and $e_i = 0.01$.

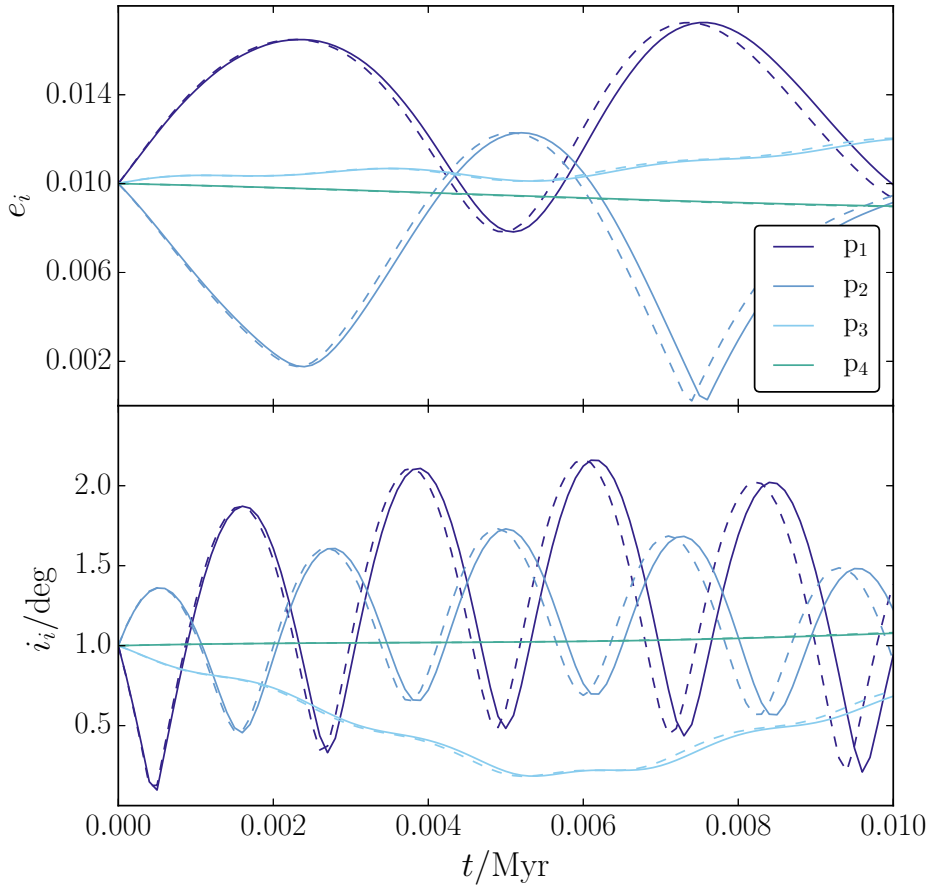
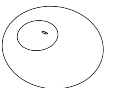


Figure 6.5: The eccentricities (top panel) and inclinations (bottom panel) as a function of time for a four-planet system with spacing $K_0 = 12$ (cf. equation 6.19). The initial eccentricities are $e_i = 0.01$ and inclinations $i_i = 1^\circ$. Solid lines are according to `SECULARMULTIPLE` with all terms up and including octupole order, plus the hexadecupole and dotriacontupole binary pair terms. Dashed lines are according to second-order LL theory. Lines with different colours correspond to the four planets (see the legend).



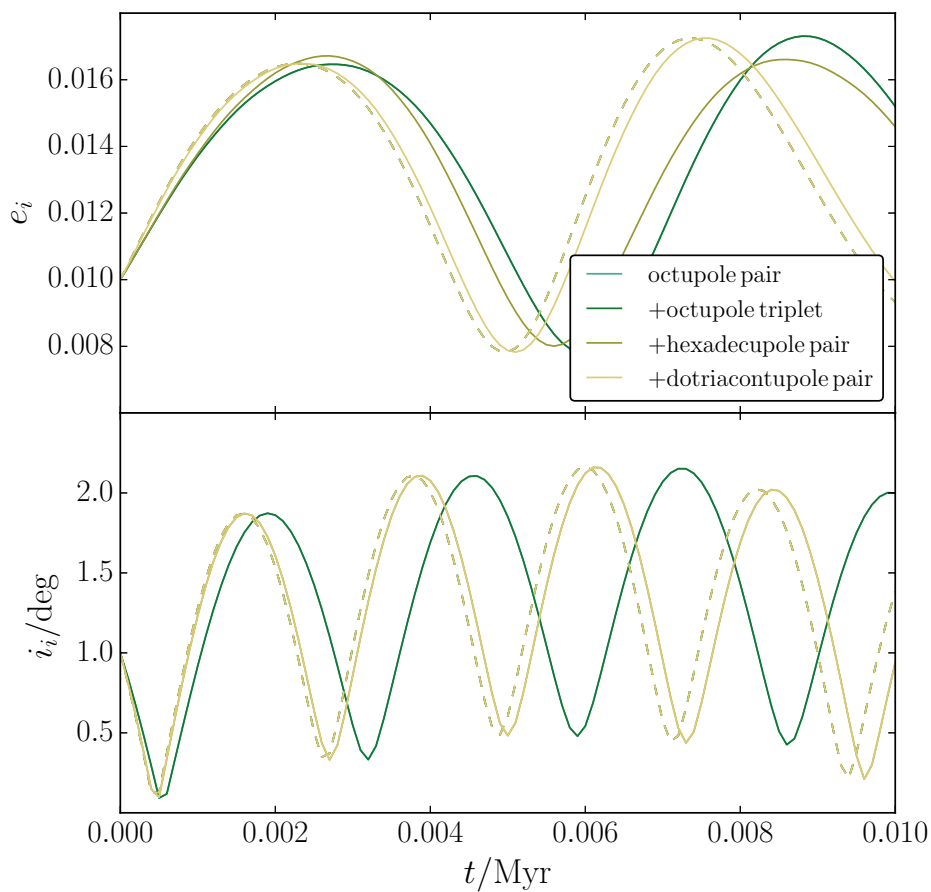


Figure 6.6: As in Figure 6.5, now showing the innermost planet's eccentricity and inclination according to LL theory (dashed lines), and SECULARMULTIPLE with inclusion of various terms (see the legend). Agreement of SECULARMULTIPLE with LL theory increases as higher order terms are included.

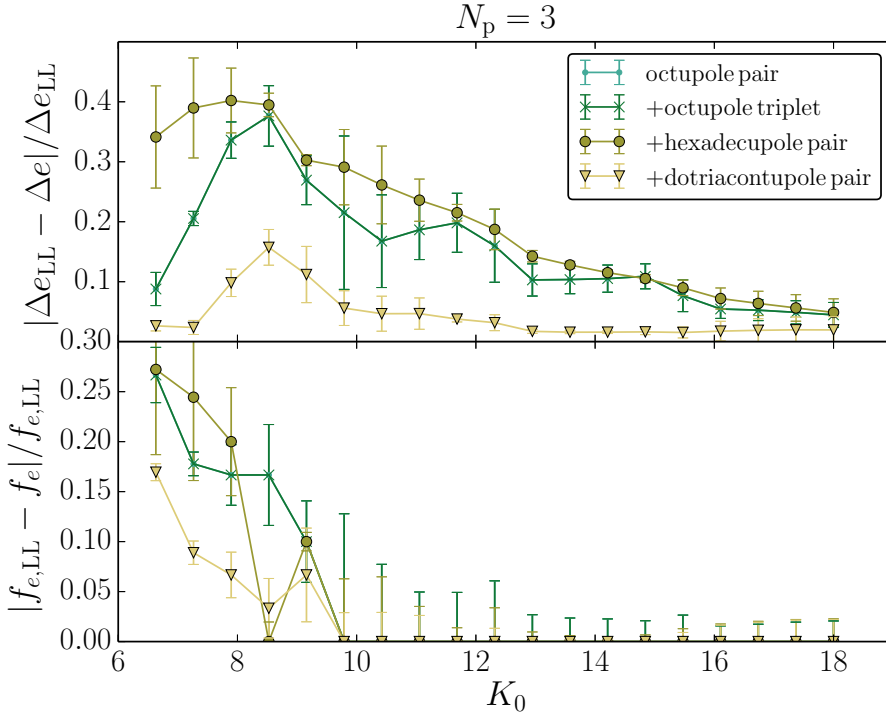
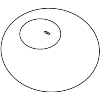


Figure 6.7: Quantitative measures of agreement between `SECULARMULTIPLE` (with various terms, indicated in the legends) and LL theory, assuming $N_p = 3$. We consider the fractional differences between the methods, of the amplitudes of the eccentricity oscillations (upper panel), and of the peak frequencies of the power spectra of the eccentricity oscillations (lower panel). Here, the solid lines represent the mean values of these measures of all planets; error bars show the standard deviations.



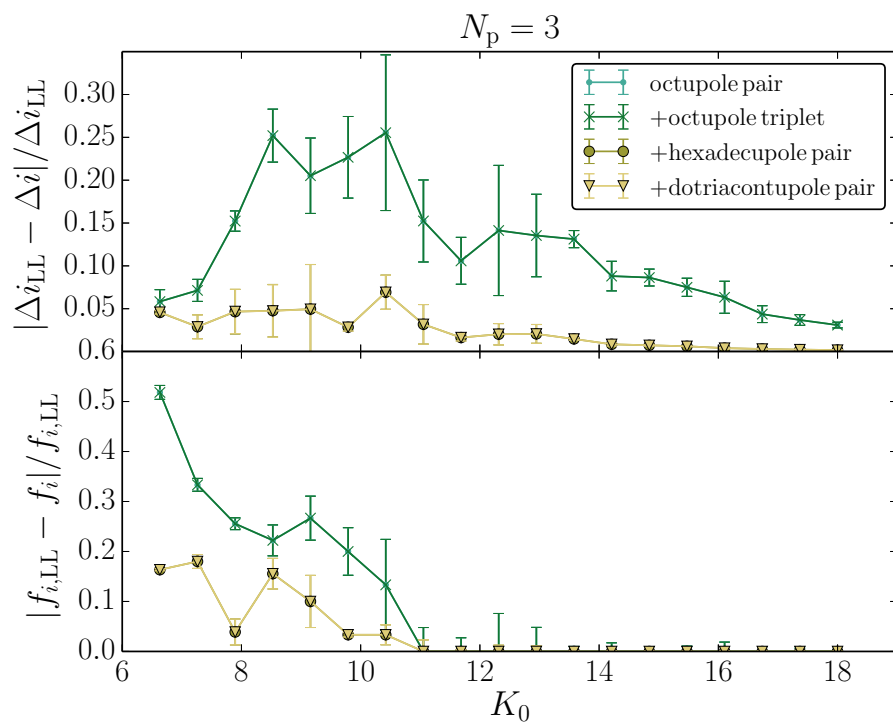


Figure 6.8: Quantitative measures of agreement similar to Figure 6.7, now applied to the oscillations of the inclinations.

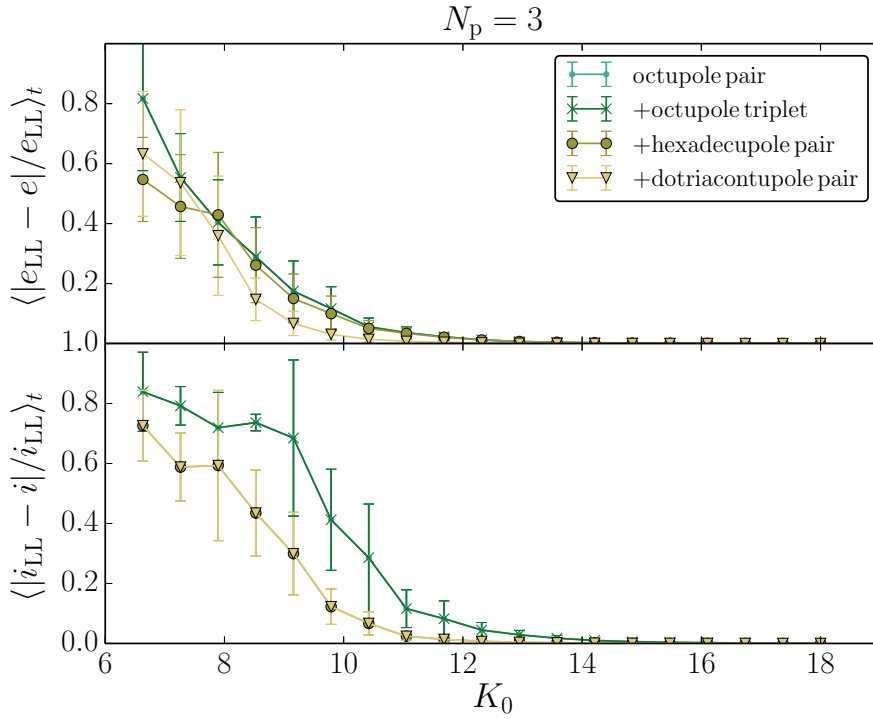
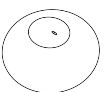


Figure 6.9: Quantitative measures of agreement similar to Figure 6.7, now applied to the time-averaged relative differences between the eccentricities (upper panel) and the inclinations (bottom panel).



- the above, plus the octupole-order binary triplet terms;
- the above, plus the hexadecupole-order binary pair terms;
- the above, plus the dotriacontupole-order binary pair terms.

Clearly, the agreement with LL theory (dashed lines) becomes increasingly better when higher order terms are included. This is because of the expansion in terms of separation ratios, which reduces to ratios of semimajor axes after orbit averaging (note that the eccentricities and inclinations can, in principle, be arbitrarily large as long as orbit averaging is still appropriate and the condition of hierarchy is still satisfied). In the example, the semimajor axis ratio $a_i/a_j \approx 0.31$ is large compared to zero. Therefore, high orders are required for accurate results. Note that the octupole-order triplet terms do not have any noticeable effect. This was estimated before in Section 6.2.4.

Generally, increasing K_0 implies smaller semimajor axis ratios, which improves the agreement of `SECULARMULTIPLE` with LL theory. To study this more quantitatively, and to determine how small K_0 can be chosen for still acceptable agreement, we show in Figures 6.7, 6.8 and 6.9 comparisons between `SECULARMULTIPLE` and LL theory, assuming $N_p = 3$. On the horizontal axes, K_0 is plotted, whereas the vertical axes show various quantitative measures for this agreement. In Figure 6.7, we consider the fractional differences between the methods, of the amplitudes of the eccentricity oscillations (upper panel) and of the peak frequencies of the power spectra of the eccentricity oscillations (lower panel). Here, the solid lines represent the mean values of these measures of all planets; errorbars show the standard deviations. In Figure 6.8, we consider similar differences, now applied to oscillations of the inclinations. Lastly, in Figure 6.9, we consider the time-averaged relative differences between the eccentricities (upper panel) and the inclinations (bottom panel).

In all cases, there is no discernable difference between the inclusion of the octupole-order binary pair terms and the inclusion of all octupole-order terms (i.e. including the binary triplet terms); in Figures 6.7, 6.8 and 6.9, the corresponding curves always overlap. When considering the *amplitude* of the eccentricity oscillations, the octupole-order terms, surprisingly, give better agreement with LL theory compared to the case of including the next higher order terms, the hexadecupole-order terms. However, the amplitude of the inclination oscillations tends to be better with the inclusion of the hexadecupole-order terms. Also, when considering the time-averaged measure (cf. Figure 6.9), including the hexadecupole-order terms also performs better compared to the octupole-order terms only. Note that for the inclination oscillations, there are little to no differences between the hexadecupole- and dotriacontupole-order terms. In contrast, the agreement with respect to eccentricity is typically much better when including dotriacontupole-order terms.

In Figures 6.10 and 6.11, we show similar figures (only considering the time-averaged measure) for $N_p = 4$ and 5, respectively. The trends described above for $N_p = 3$ also apply to larger planet numbers. As the number of planets is increased, the dependence of the measure of agreement as a function of K_0 becomes smoother.

Generally, $\sim 2\%$ discrepancy with respect to LL theory (in terms of the time-averaged measure) is reached with dotriacontupole-order terms if $K_0 \gtrsim 10$. When including only the octupole-order terms, the discrepancy is between ~ 10 and $\sim 20\%$, depending on the number of planets. In particular, the agreement with respect to the eccentricity becomes much better

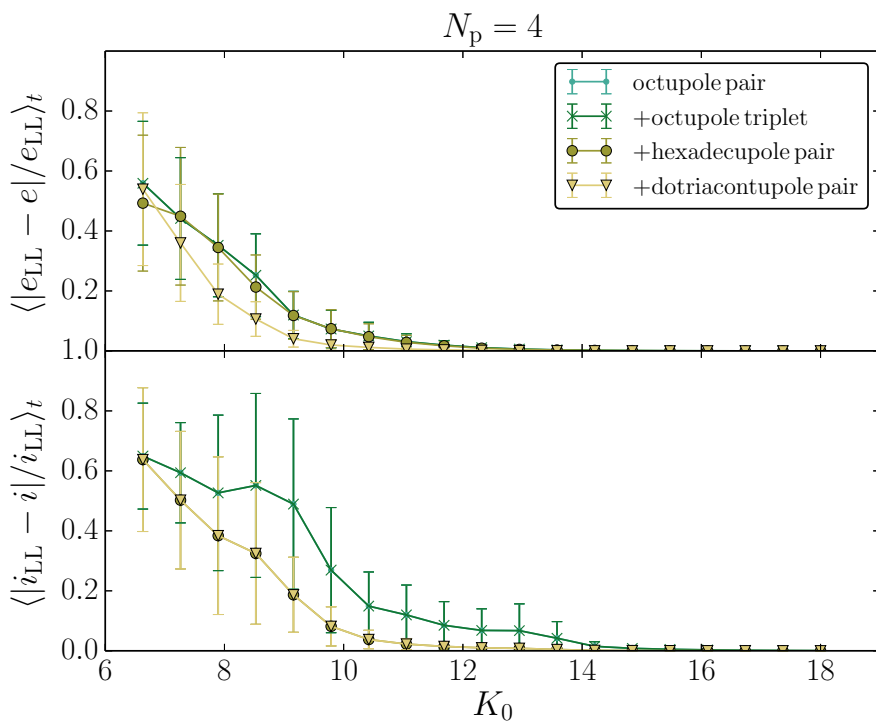
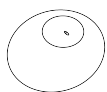


Figure 6.10: Quantitative measures of agreement between `SECULARMULTIPLE` and LL theory as in Figure 6.9, now with four planets. We show the time-averaged relative differences between the eccentricities (upper panel) and the inclinations (bottom panel).



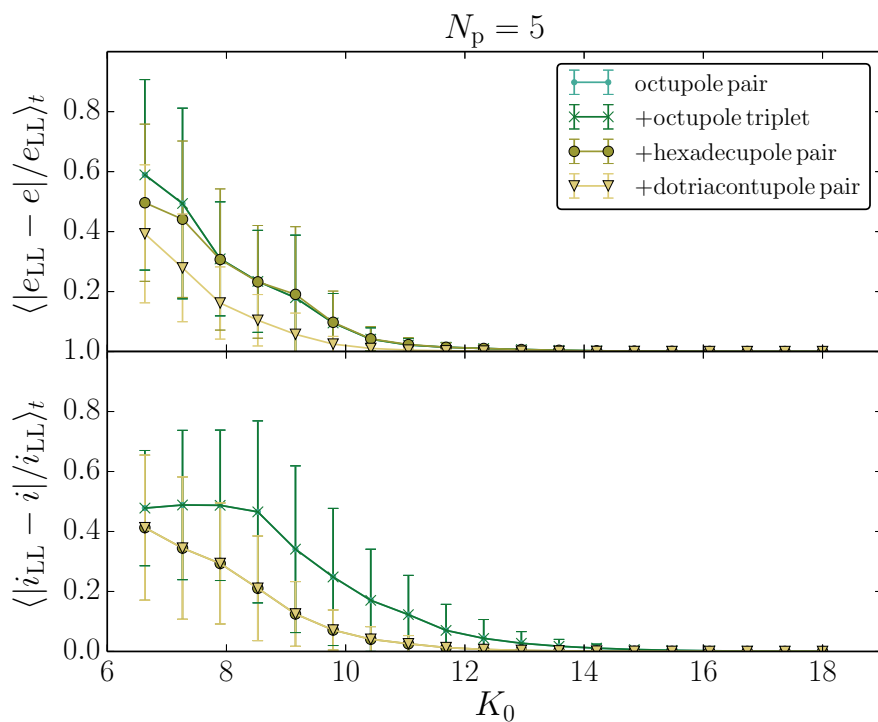


Figure 6.11: Quantitative measures of agreement between SECULARMULTIPLE and LL theory as in Figure 6.10, now with five planets.

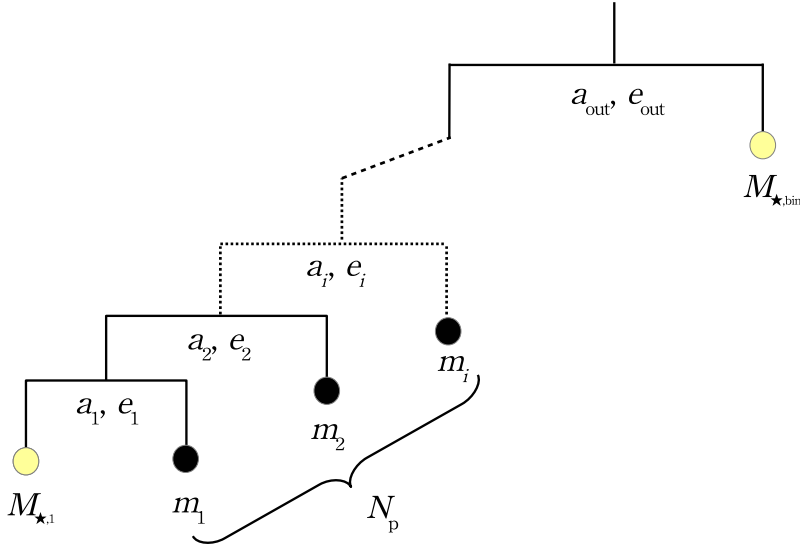


Figure 6.12: A mobile diagram of a hierarchical multiplanet system in a stellar binary (cf. Section 6.3.2). The N_p planets orbit a star with mass $M_{*,1}$. The centre of mass of this subsystem is orbited by a binary companion star with mass $M_{*,bin}$ in an orbit with semimajor axis and eccentricity a_{bin} and e_{bin} , respectively. In Section 6.3.2, we set $N_p = 2$ (to match the setup of Takeda et al. 2008); in Section 6.3.2, we set $N_p = 4$.

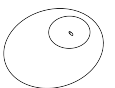
with the inclusion of the dotriacontupole-order terms. This demonstrates the need to include high-order terms in multiplanet systems.

6.3.2 Binary-star systems

As mentioned in Section 6.1, it is possible that a star hosting a multiplanet system is itself part of a multiple stellar system. Here, we consider the case of an S-type multiplanet system in which the host star has a more distant stellar binary companion well beyond the orbits of the planets. A general representation of the system (for N_p planets) is given in Figure 6.12. In Section 6.3.3, we also consider the planet-hosting star to be the tertiary star in a stellar triple system. In the latter case, the stellar ‘binary companion’ to the planet-hosting star is itself a stellar binary (cf. Figure 6.23).

Two-planet system in a stellar binary -- comparison with direct N -body integrations

First, we compare results of `SECULARMULTIPLE` to a number of direct N -body integrations of a system previously studied by Takeda et al. (2008). The latter authors considered two-planet systems, and showed that two distinct dynamical classes exist with respect to the response of the multiplanet system to the secular torque of the binary companion. In one class, the orbits of



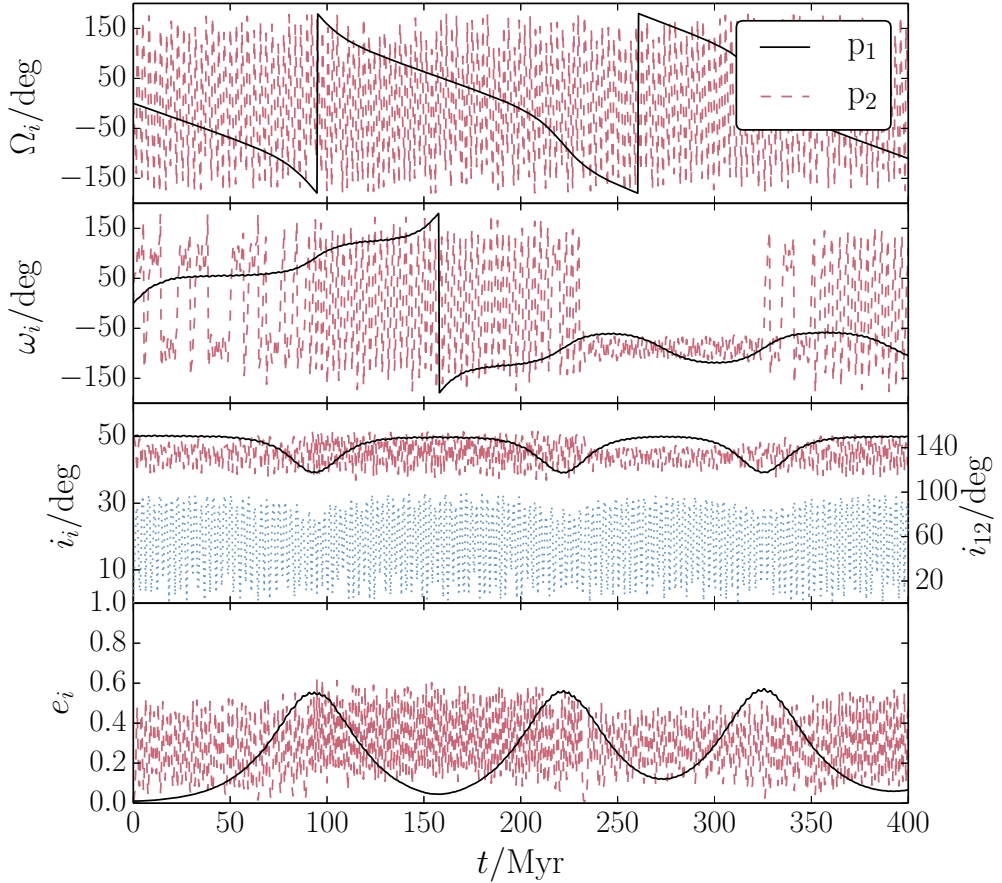


Figure 6.13: Integration of a two-planet system with a stellar binary companion (cf. Figure 6.12). The system parameters are adopted from Fig. 5 of Takeda et al. (2008). The planetary parameters are $m_1 = 1 M_J$, $a_1 = 2$ AU, $e_1 = 0.01$, $i_1 = 50^\circ$, $\omega_1 = 0^\circ$, $\Omega_1 = 0^\circ$ (inner planet; solid black lines), and $m_2 = 0.032 M_J$, $a_2 = 31.6$ AU, $e_2 = 0.01$, $i_2 = 50^\circ$, $\omega_2 = 0^\circ$, $\Omega_2 = 0^\circ$ (outer planet; red dashed lines). The central star has mass $M_{*,1} = 1 M_\odot$, and the binary companion, with mass $M_{*,\text{bin}} = 1 M_\odot$, is in an orbit with $a_{\text{out}} = 750$ AU, $e_{\text{out}} = 0.2$, $i_{\text{out}} = 0^\circ$, $\omega_b = 0^\circ$ and $\Omega_{\text{out}} = 0^\circ$, i.e. the initial mutual inclination between the planetary orbits and the binary orbit is 50° . In the third panel, the blue dotted line shows the mutual inclination i_{12} between the two planets; note the different scale on the vertical axis, indicated to the right of the panel.

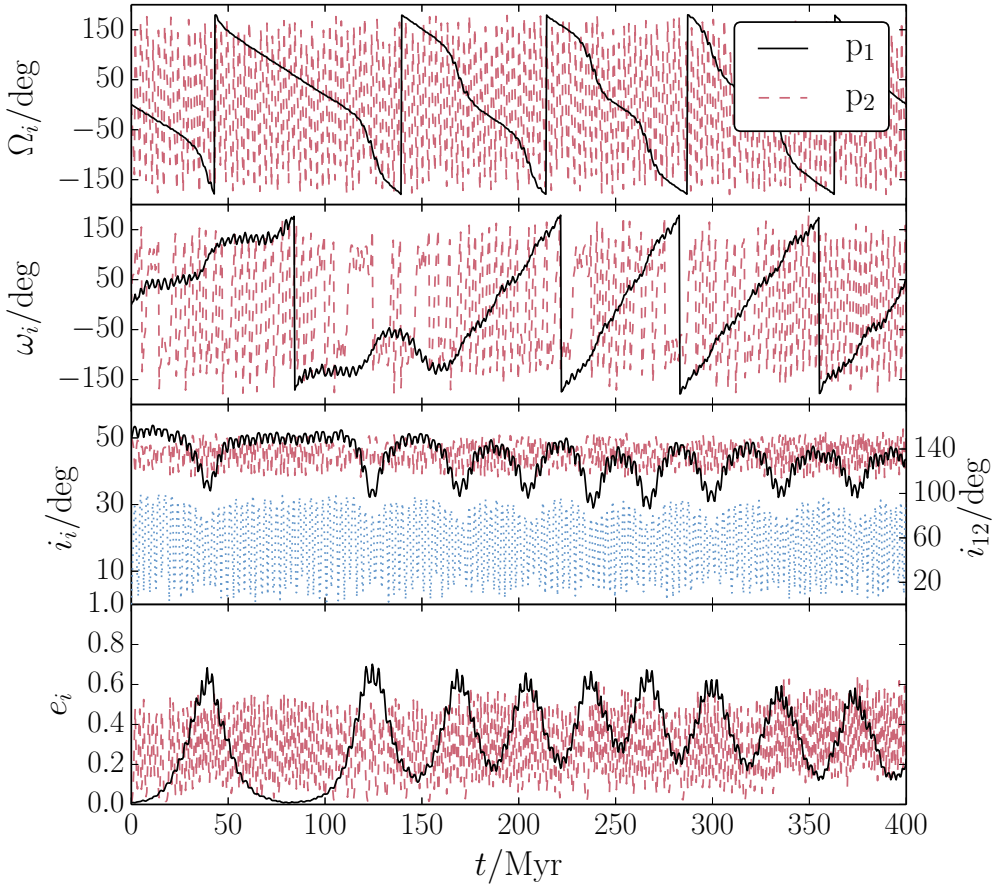
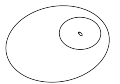


Figure 6.14: As in Figure 6.13, now with $m_2 = 0.32 M_J$ (parameters adopted from Fig. 6 of Takeda et al. 2008).



the planets evolve ‘rigidly’, i.e. closely maintaining their mutual inclinations and nodal angles despite an overall change of the orbital plane of the multiplanet system. In the second class, the orbits are coupled to the secular torque of the binary companion, thereby inducing potentially large mutual inclinations between the planets, and large excitations of the eccentricities.

In Figures 6.13 and 6.14, we show two integrations similar to figs 5 and 6 of Takeda et al. (2008), respectively (refer to the captions for detailed parameters). In these figures, the planets are either decoupled from each other and respond to the secular torque of the binary companion (Figure 6.13, with $m_2 = 0.032 M_J$) or are weakly coupled (Figure 6.14, $m_2 = 0.32 M_J$).

For verification purposes², we carried out a number of direct N -body integrations of these systems with various codes within AMUSE: HERMITE (Hut et al. 1995), MIKKOLA (Mikkola & Merritt 2008), HUAYNO (Pelupessy et al. 2012) and SAKURA (Gonçalves Ferrari et al. 2014). In Figures 6.15 and 6.16, we show, for $m_2 = 0.032 M_J$ and $m_2 = 0.32 M_J$ respectively, the energy errors, the semimajor axes and the eccentricities of the planets for the integrations with these codes. We also include integrations with SECULARMULTIPLE (black lines). We limited the run time of the N -body integrations to approximately three weeks, and only HERMITE and HUAYNO were able to complete the 400 Myr integration within that time. The other codes, MIKKOLA and SAKURA, computed ≈ 42 Myr of the evolution during our run time. Note that the 400 Myr integration with SECULARMULTIPLE takes a few minutes.

For HERMITE and HUAYNO, the energy errors increased to $\mathcal{O}(10^{-1})$ and $\mathcal{O}(10^{-2})$, respectively, and the semimajor axes of the planets started to deviate significantly from their original values (cf. the middle panels of Figures 6.15 and 6.16). Given the large energy errors and the differences in behaviour in the innermost planet semimajor axis evolution between HERMITE and HUAYNO (according to the former, a_1 decreases over several 100 Myr, whereas according to the latter, a_1 increases over the same time-scale), this is likely due to computational errors, and not due to the true dynamical evolution.

The large energy errors in the case of HERMITE and HUAYNO also likely explain why the eccentricity of the innermost planet is not excited to large values of ≈ 0.6 and ≈ 0.7 for $m_2 = 0.032 M_J$ and $m_2 = 0.32 M_J$, respectively, as computed by SECULARMULTIPLE, and according to the integrations of Takeda et al. (2008) (cf. Figures 5 and 6 of the latter paper). In Figures 6.17 and 6.18, we show zoomed-in versions of the eccentricities shown in the bottom panels of Figures 6.15 and 6.16, respectively. For both values of m_2 and the first ≈ 7 Myr, the evolution of e_2 matches well between the various N -body codes and SECULARMULTIPLE, and starts to deviate between the various methods after this time. Note that these deviations occur even for the two most accurate N -body integrations with MIKKOLA and SAKURA. For $m_2 = 0.32 M_J$, the eccentricity evolution of the innermost planet (e_1) agrees between MIKKOLA, SAKURA and SECULARMULTIPLE within a margin of ~ 0.1 . Note that the differences between the secular and direct N -body codes are similar to the differences between the two direct codes MIKKOLA and SAKURA. For $m_2 = 0.032 M_J$, there is a relatively larger discrepancy in e_1 between SECULARMULTIPLE and MIKKOLA of ≈ 0.02 at ≈ 40 Myr. However, SECULARMULTIPLE and SAKURA agree very well. In contrast to MIKKOLA, SAKURA and SECULARMULTIPLE, HERMITE and HUAYNO give a very different evolution of e_1 , and this is likely due to the large energy errors in the integrations with the latter codes.

We conclude that there is good agreement of SECULARMULTIPLE with the most accurate direct N -body integrations. We emphasize that HERMITE and HUAYNO, even though slower compared to SECULARMULTIPLE by a factor of $\sim 10^3$, fail to accurately produce the eccentricity evolution of the

²We were unable to obtain the data from Takeda et al. (2008).

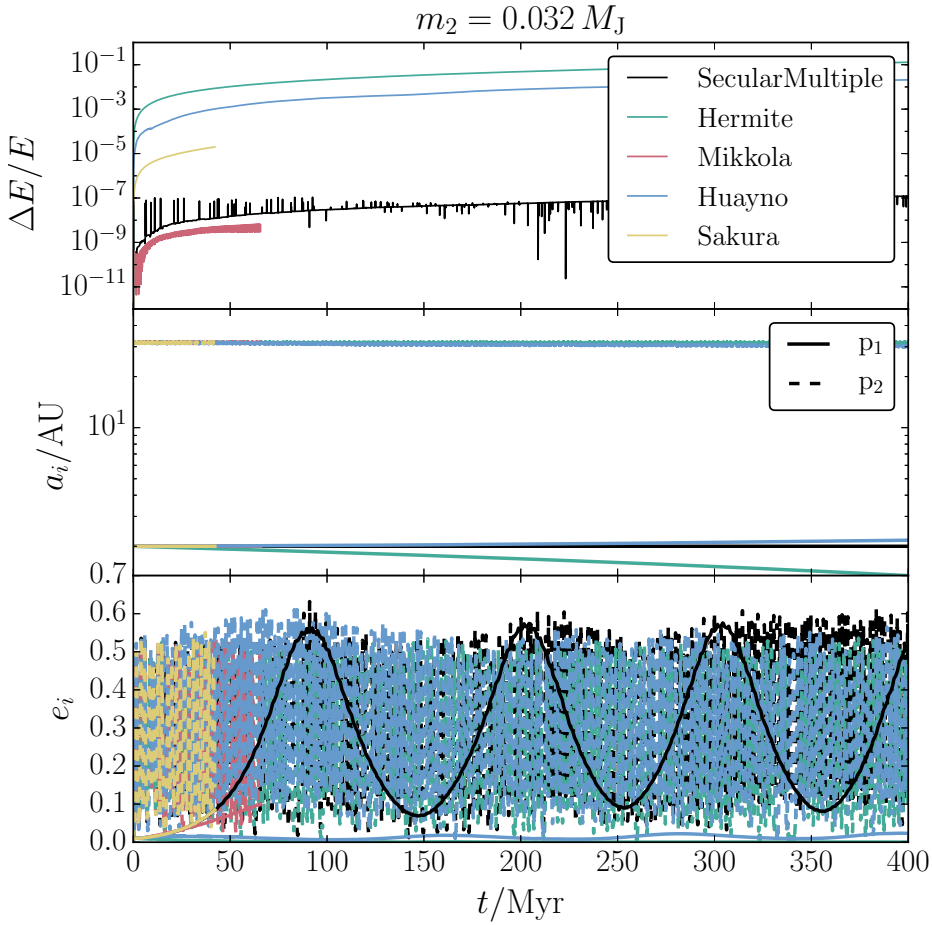
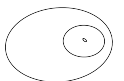


Figure 6.15: Comparison between SECULARMULTIPLE and various direct N -body integrations using different codes (refer to the labels in the top panel) for the system from Figure 6.13 ($m_2 = 0.032 M_J$). Top panel: the energy errors; middle panel: the semimajor axes of the planets (by construction, these remain exactly constant in SECULARMULTIPLE). Bottom panel: the eccentricities of the planets.



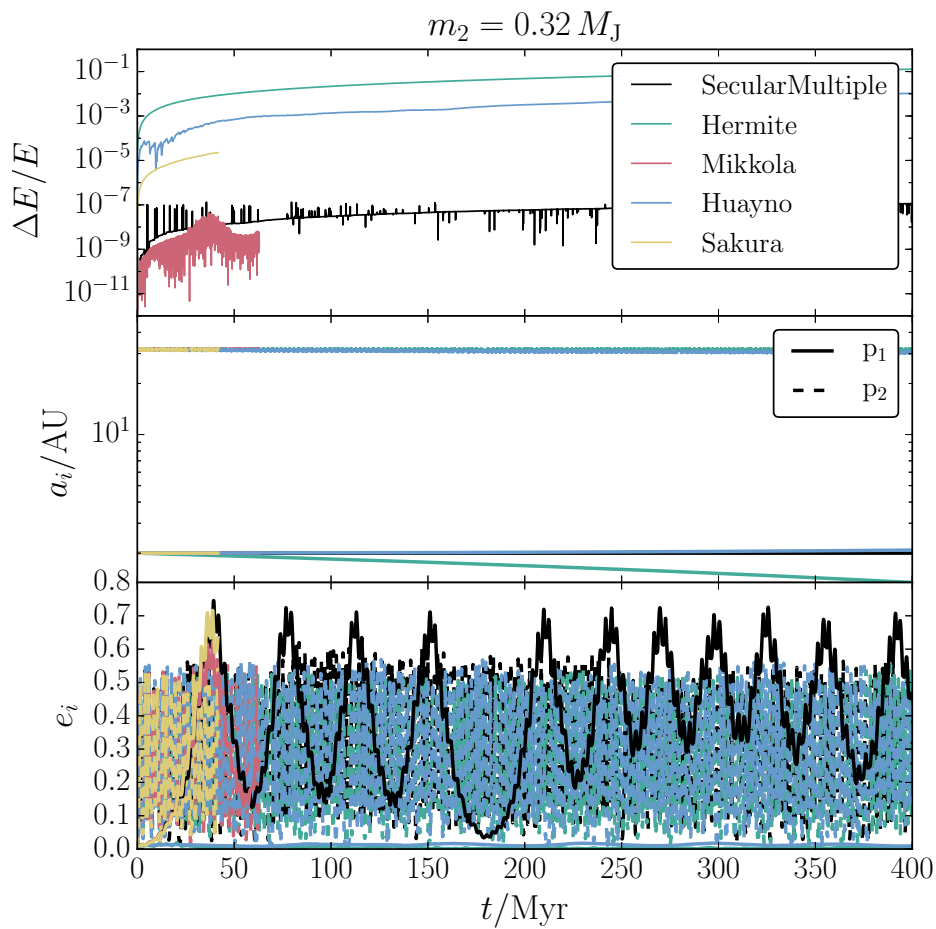


Figure 6.16: As in Figure 6.15, now with $m_2 = 0.32 M_J$.

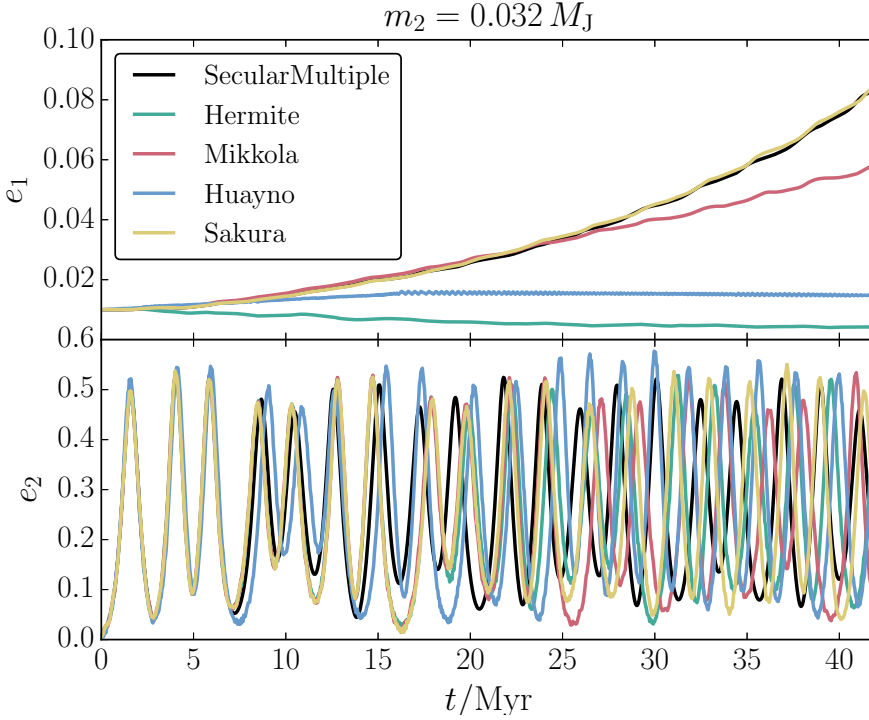


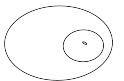
Figure 6.17: A zoomed-in version of the bottom panel of Figure 6.15 (with $m_2 = 0.032 M_J$), showing the eccentricities of planets 1 and 2 in the top and bottom panels, respectively.

innermost planet. To get accurate results on secular time-scales with direct N -body methods, it is crucial that energy errors remain small during the evolution (based on Figure 6.15, no larger than $\sim 10^{-4}$). Note that this minimum energy error is smaller than the value of 10^{-1} suggested by Boekholt & Portegies Zwart (2015); however, this value applies in a statistical sense to an ensemble of systems, and not to individual systems. This requires the use of special integration techniques such as algorithm chain regularization (as in MIKKOLA) or Keplerian-based Hamiltonian splitting (as in SAKURA). Alternatively, for the hierarchical systems considered here, SECULARMULTIPLE agrees well with the most accurate direct N -body codes, and it is faster by several orders of magnitude.

In Section 6.3.2 below, we take advantage of the speed of SECULARMULTIPLE, and study more generally the nature of the eccentricity oscillations as a function of the binary semimajor axis.

Four-planet system with constant spacing in a stellar binary

Here, we consider four-planet systems with fixed spacing and small initial eccentricities and inclinations as in Section 6.3.1, now including a binary companion star with mass $M_{\star, \text{bin}} = 1 M_\odot$. For the multiplanet system, we set $M_{\star, 1} = 1 M_\odot$, $m_i = 1 M_J$, $K_0 = 10$, $a_1 = 0.2 \text{ AU}$,



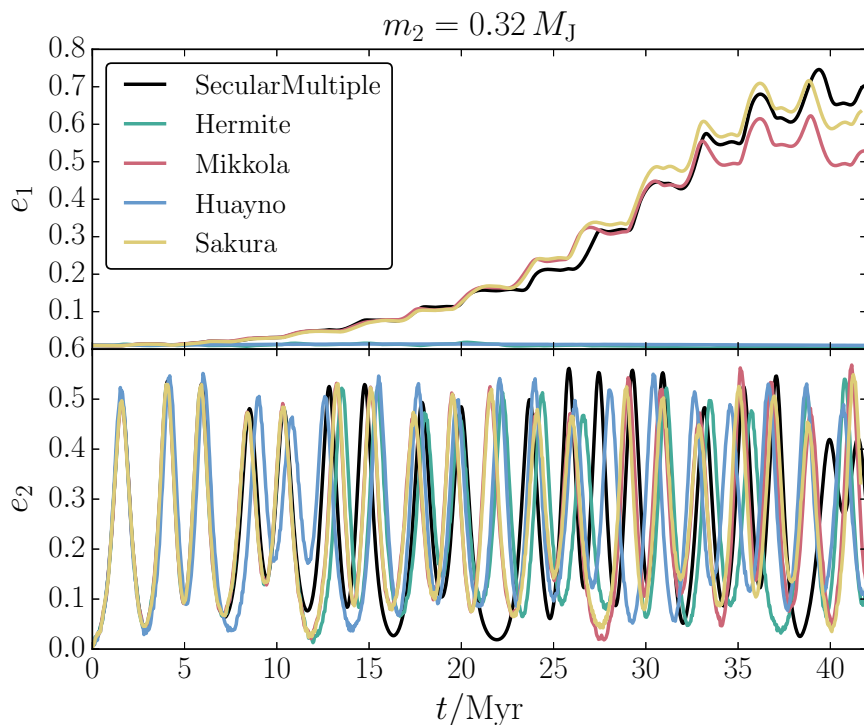


Figure 6.18: As in Figure 6.17, now with $m_2 = 0.32 M_J$.

$e_i = 0.01$ and $i_i = 1^\circ$. A range of semimajor axes for the binary companion orbit a_{out} ³ is assumed: between 50 and 150 AU with $e_{\text{out}} = 0.1$, and between 100 and 300 AU with $e_{\text{out}} = 0.8$. The lower limits on a_{out} are chosen to ensure short-term dynamical stability according to the formulae of Holman & Wiegert (1999). The initial inclination of the binary orbit is set to 80° with $\Omega_{\text{out}} = \omega_{\text{out}} = 0^\circ$. For the two values of e_{out} , we integrated 500 systems with 100 different values of a_{out} , and for each value of a_{out} , five realizations of Ω_i for the planetary system assuming that the Ω_i are randomly distributed. Because the planetary inclinations are small, the initial mutual inclinations of the planets with respect to the binary companion are very similar to each other, and are $\approx 80^\circ$. We integrate each system for a duration of 1 Myr, which is comparable to the LK time-scale of the binary with respect to the outermost planet.

In the top panel of Figure 6.19, we show the resulting maximum values of the eccentricities of the planets as a function of a_{out} , assuming $e_{\text{out}} = 0.1$. Bullets show the average values over Ω_i , whereas the error bars indicate the standard deviations. In the bottom panel of Figure 6.19, we show the mean values of all combinations of the mutual inclinations between the planets, averaged over time.

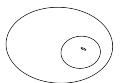
For large a_{out} , the maximum eccentricities are small and so are the mutual inclinations. In this limit, the secular torque of the binary companion is weak compared to the mutual torques exerted by the planets. Although the absolute inclinations and longitudes of the ascending nodes change, the planets remain closely fixed to the mutual plane which is precessing because of torque of the binary companion. In Figure 6.21, we show an example of the evolution of the eccentricities and the inclinations in this regime, with $a_{\text{out}} \approx 100.5$ AU and $e_{\text{out}} = 0.1$. The *absolute* inclinations change by $\approx 150^\circ$ on a time-scale of 1 Myr (cf. the solid lines in the bottom panel of Figure 6.21), whereas the inclinations with respect to the innermost planet do not reach values above $\approx 3^\circ$ (cf. the dotted lines in the bottom panel of Figure 6.21).

As a_{out} is decreased, the torque of the binary companion increases. For sufficiently small a_{out} , the outermost planet is decoupled from the inner planets resulting in larger mutual inclinations; the outermost planet reaches high eccentricity. For $a_{\text{out}} \lesssim 75$ AU, e_4 is high enough for planet 4 to cross its orbit with planet 3; consequently, we stopped the integration. These systems are not included in Figure 6.19. The maximum eccentricities of planets 1 through 3, in particular planet 3, are also affected when decreasing a_{out} . In our systems, however, they do not reach high enough values for orbit crossings to occur. In Figure 6.22, we show an example of the time evolution with $a_{\text{out}} \approx 80.3$ AU and $e_{\text{out}} = 0.1$. The outermost two planets are markedly affected by LK cycles induced by the binary companion. For planet 3, secular oscillations due to planet-planet torques, which have shorter time-scales compared to the LK oscillations, are still clearly present. Note that even though the eccentricity of the outermost planet is strongly increased by the binary companion, the mutual inclination with respect to planet 1 (cf. the green dotted line in the bottom panel of Figure 6.22) remains small, i.e. $\lesssim 5^\circ$.

To estimate the value of a_{out} that separates the two regimes, we use the arguments given by Takeda et al. (2008). We estimate the LK time-scale t_{LK} (cf. equation 6.7) of the binary companion with respect to the outermost orbit by

$$t_{\text{LK}} \approx \frac{1}{3\pi} \left(\frac{P_{\text{orb,out}}^2}{P_{\text{orb,4}}} \right) \frac{M_{\star,1} + M_{\star,\text{out}}}{M_{\star,\text{out}}} (1 - e_{\text{out}}^2)^{3/2}. \quad (6.21)$$

³We use the subscript ‘out’ rather than e.g. ‘bin’, in anticipation of the case of a triple star system discussed in Section 6.3.3.



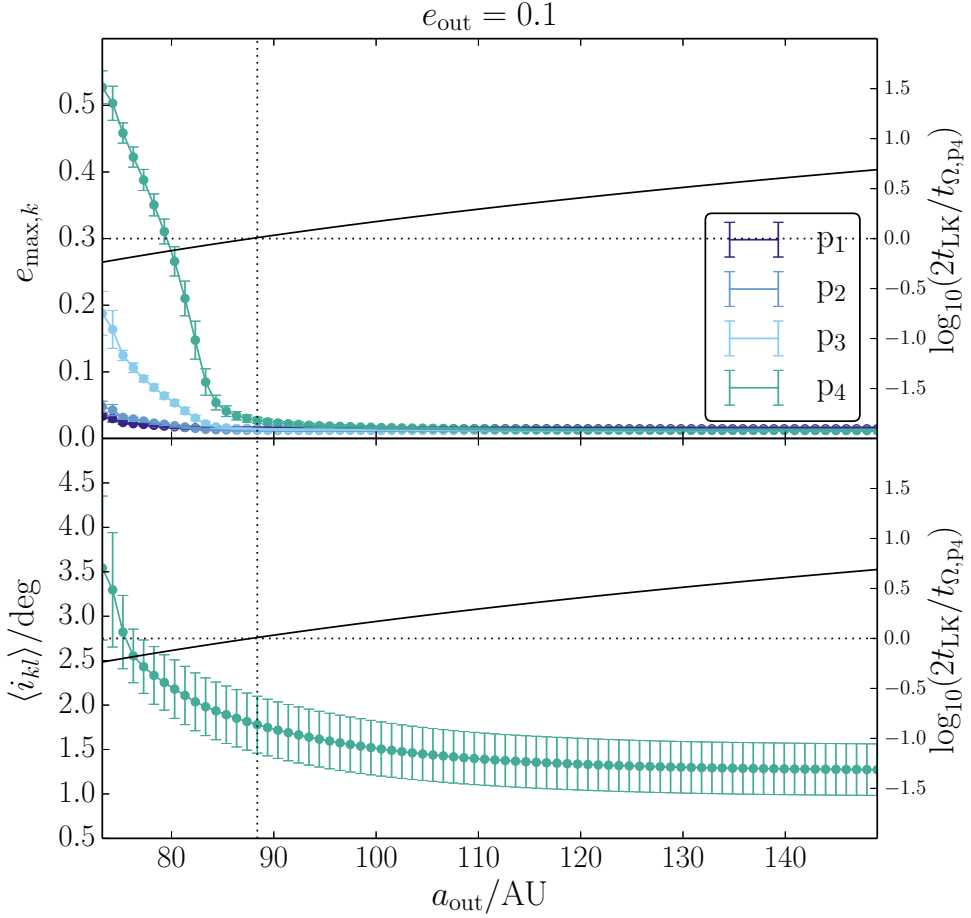


Figure 6.19: For a four-planet system with a binary companion and eccentricity $e_{\text{out}} = 0.1$ (cf. Figure 6.12), the maximum values of the eccentricities of the planets as a function of a_{out} , assuming $e_{\text{out}} = 0.1$. Bullets show the average values over Ω_i , whereas the error bars indicate the standard deviations. In the bottom panel, we show the mean values of all combinations of mutual inclinations between the planets, averaged over time. In each panel, the right axes show the logarithm of the ratio $2t_{\text{LK}}/t_{\Omega,p_4}$, where $t_{\Omega,p_4} \equiv 2\pi/f_4$, of the nodal precession time-scale associated with LK oscillations induced by the binary companion, compared to the nodal precession time-scale associated with the mutual planetary torques.

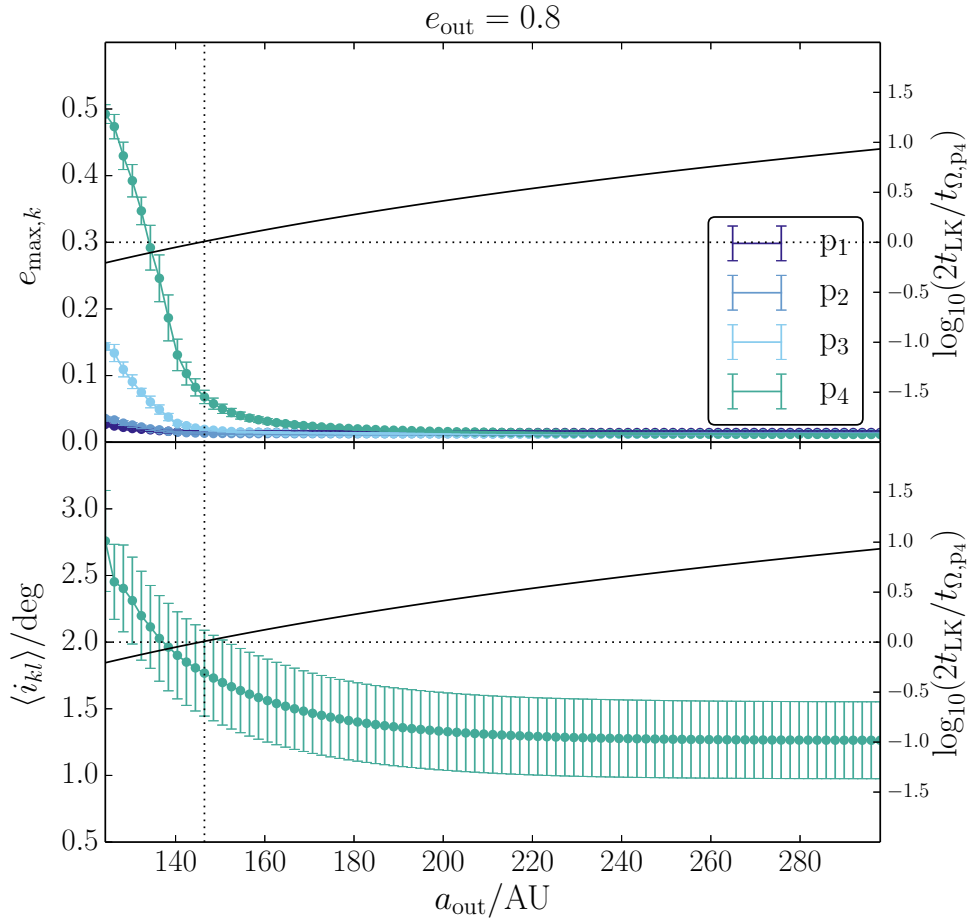
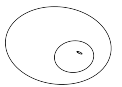


Figure 6.20: Similar to Figure 6.19, now with $e_{\text{out}} = 0.8$.



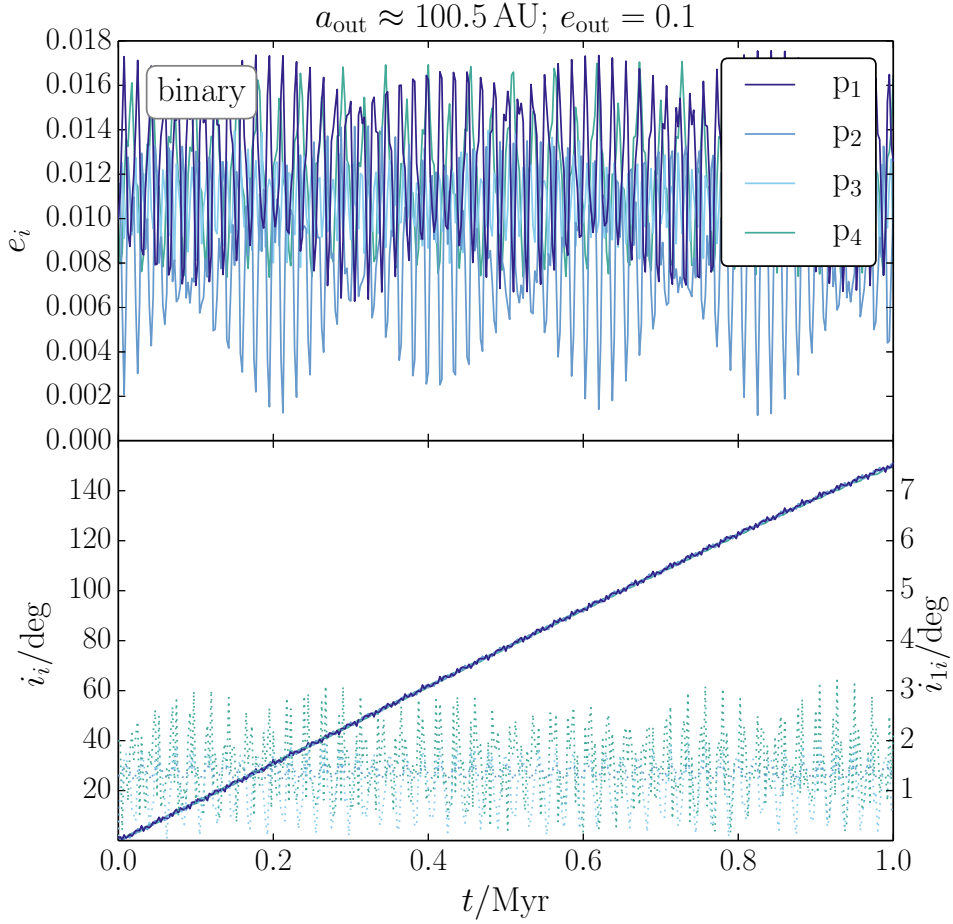


Figure 6.21: The eccentricities and inclinations of a four-planet system in the presence of a binary companion, in the regime in which planet-planet torques dominate compared to the torque of the binary companion. The assumed values of a_{out} and e_{out} are indicated above the top panel. In the bottom panel, the absolute inclinations i_i are shown with solid lines; the inclinations relative to planet 1, i_{1i} (right axis), are shown with dotted lines.

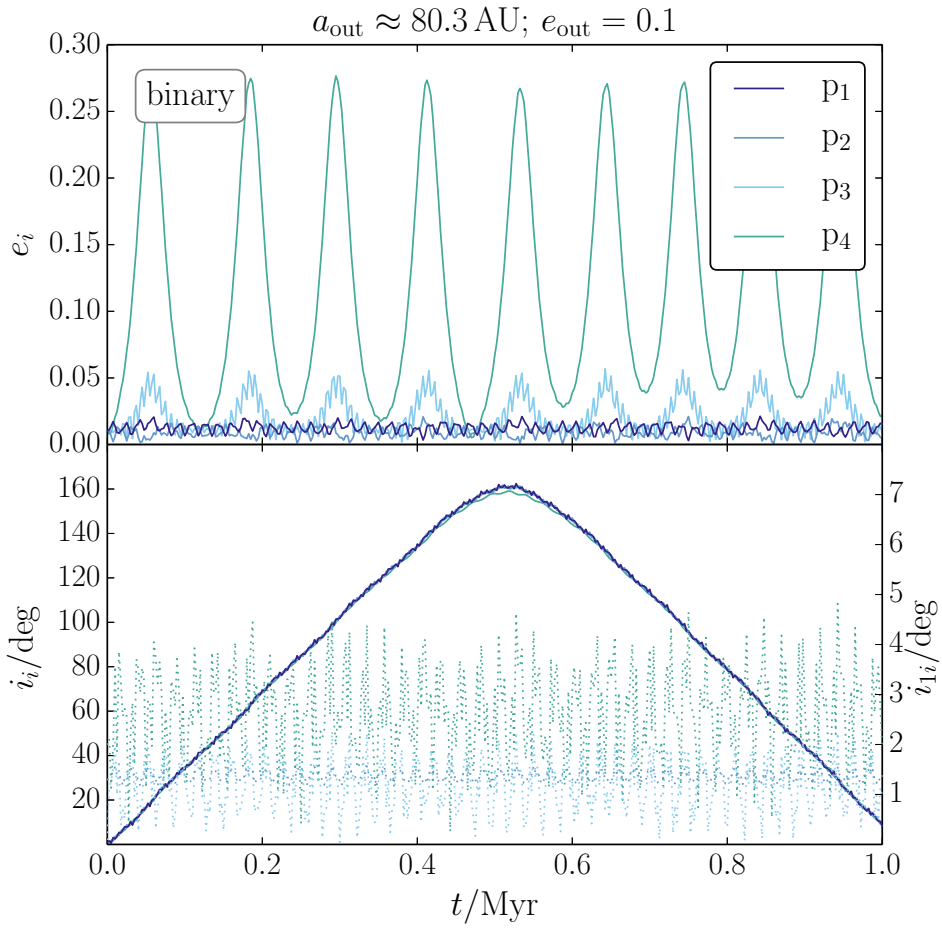
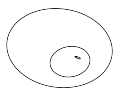


Figure 6.22: Similar to Figure 6.22, now with a smaller binary semimajor axis. In this case, the eccentricities of the planets are excited by the secular torque of the binary companion.



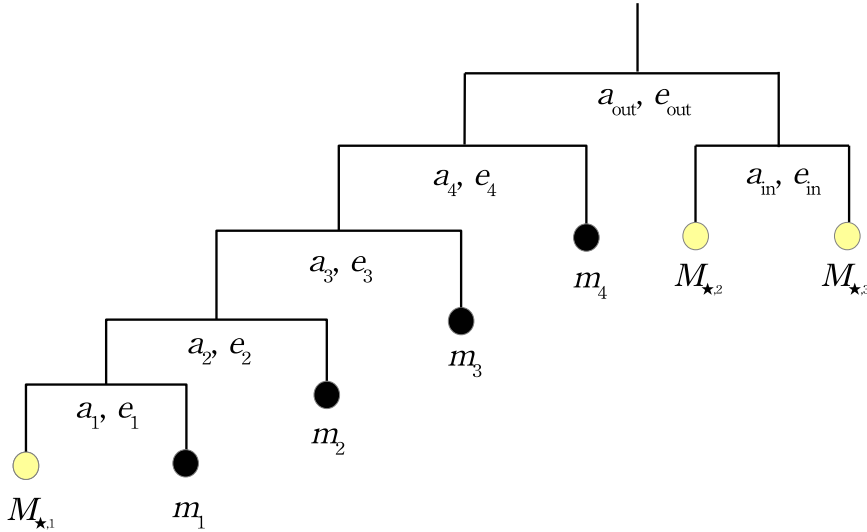


Figure 6.23: A mobile diagram of a hierarchical four-planet system in a stellar triple (cf. Section 6.3.3). The planets are in S-type orbits around the tertiary star ($M_{\star,1}$) of a hierarchical triple system with inner and outer semimajor axes a_{in} and a_{out} , respectively.

We compare $2t_{\text{LK}}$, the approximate time-scale for nodal precession due to the secular torque of the binary companion, to $t_{\Omega, \text{p4}} \equiv 2\pi/f_4$, an estimate of the time-scale associated with nodal oscillation of the fourth planet according to second-order LL theory (Murray & Dermott 1999).

In Figure 6.19, we show with the black solid lines the ratio $2t_{\text{LK}}/t_{\Omega, \text{p4}}$ as a function of a_{out} . The vertical dotted lines indicate the value of a_{out} for which the ratio is ≈ 1 . The latter lines indeed approximate the transition between the LK-dominated and planet-planet-dominated regimes.

In Figure 6.20, we show a similar figure of the maximum eccentricities and averaged inclinations as a function of a_{out} , now with $e_{\text{out}} = 0.8$. The behaviour of the maximum eccentricities and the averaged mutual inclinations as a function of a_{out} is similar as for the case $e_{\text{out}} = 0.1$. As expected, the transition occurs at larger a_{out} , i.e. $a_{\text{out}} \sim 150$ AU compared to ~ 90 AU if $e_{\text{out}} = 0.1$. The method of comparing $2t_{\text{LK}}$ to $t_{\Omega, \text{p4}} \equiv 2\pi/f_4$ again gives an estimate for the transition value of a_{out} .

6.3.3 Triple-star systems

To our knowledge, the secular dynamics of multiplanet systems orbited by a stellar binary have not been explored. Here, we carry out a number of integrations similar to those of Section 6.3.2, but now with the binary companion *star* replaced by a stellar *binary* with semimajor axis a_{in} . The hierarchy of the system is shown in a mobile diagram in Figure 6.23. The system can be viewed as a stellar hierarchical triple system, in which the tertiary star has four S-type planets (effectively making it a heptuple system).

In order to easily compare with the case of a stellar binary, we assume that the sum of the inner binary stellar masses is equal to the mass of the companion star, $M_{\star,\text{bin}} = 1 M_{\odot}$, that was assumed in Section 6.3.2, i.e. $M_{\star,2} + M_{\star,3} = M_{\star,\text{bin}}$. Here, we set $M_{\star,2} = 0.6 M_{\odot}$ and $M_{\star,3} = 0.4 M_{\odot}$. The initial (absolute) inclination of the inner orbit is set to $i_{\text{in}} = 0^\circ$; for the outer orbit, we set $i_{\text{out}} = 80^\circ$. Furthermore, $\Omega_{\text{in}} = \Omega_{\text{out}} = \omega_{\text{in}} = \omega_{\text{out}} = 0^\circ$. Consequently, the mutual inclinations between the multiplanet system and the outer orbit are $\approx 80^\circ$ as in Section 6.3.2, and the inner and outer binaries are inclined by 80° . The initial eccentricities of the stellar orbits are $e_{\text{in}} = e_{\text{out}} = 0.1$. We integrated a grid of systems with various values of a_{in} and a_{out} . As in Sections 6.3.1 and 6.3.2, for each combination of parameters, five integrations were carried out with all parameters fixed, except for the Ω_i of the planets, which were sampled randomly.

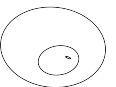
In Figure 6.24, we show the maximum eccentricities (top panel) and the time-averaged mutual inclinations (bottom panel) of the planets as a function of a_{in} , with fixed $a_{\text{out}} = 80 \text{ AU}$ and $e_{\text{out}} = 0.1$. Note that in the equivalent system with the inner binary replaced by a point mass, the maximum eccentricities are approximately 0.02, 0.02, 0.05 and 0.27 for planets 1 through 4, respectively (see e.g. Figure 6.22). These values are indeed attained for small values of a_{in} , i.e. $a_{\text{in}} \lesssim 3 \text{ AU}$, in which case the inner binary is effectively a point mass. However, for $a_{\text{in}} \gtrsim 3 \text{ AU}$, the maximum eccentricities are strongly affected.

For large a_{in} , $a_{\text{in}} \gtrsim 10 \text{ AU}$, the maximum eccentricities are *smaller* compared to the case of a stellar binary. This can be understood by considering the evolution of the nodal angle Ω_{out} of the outer binary. An example of the evolution of the eccentricities, inclinations and nodal angles assuming $a_{\text{in}} \approx 14.0 \text{ AU}$, $a_{\text{out}} \approx 80.0 \text{ AU}$ and $e_{\text{out}} = 0.1$ is shown in Figure 6.25. In the case of a stellar binary, Ω_{out} does not oscillate because the planets are not massive enough to cause precession of the angular momentum vector of the outer binary, $\mathbf{h}_{\text{out}}$. However, in the case of a stellar triple with sufficiently large a_{in} , the inner binary causes a precession of $\mathbf{h}_{\text{out}}$ on a time-scale $2t_{\text{LK,in,out}}$. In the case shown in Figure 6.25, the time-scale for oscillations of Ω_{out} is short compared to the time-scale of LK oscillations of the outer orbit on the outermost planet (cf. Figure 6.22). Consequently, the rapid nodal precession of Ω_{out} reduces the excitation of planetary eccentricities by the secular torque of the outer binary.

However, the opposite effect can occur if the nodal precession time-scale of the inner and outer binaries, $2t_{\text{LK,in,out}}$, is similar to the time-scale t_{Ω,p_4} for precession of the outermost planet nodal angle due to the planet-planet torques. In this case, the (nearly) synchronous nodal precession frequencies of the outer binary and the outermost planets cause an *additional* excitation of the eccentricity oscillations of the innermost planets, compared to the case of a stellar binary.

An example is shown in Figure 6.26, in which $a_{\text{in}} \approx 9.0 \text{ AU}$, $a_{\text{out}} \approx 80.0 \text{ AU}$ and $e_{\text{out}} = 0.1$ are assumed. The eccentricity of the outermost planet, planet 4, is excited to large values; in particular, it evolves in an intricate manner on a time-scale of approximately twice the time-scale of the eccentricity oscillations of the inner binary. The latter indeed occur on time-scales of half the nodal oscillation time-scales, cf. the bottom panel of Figure 6.26. Also, the inclination, of the outermost planet in particular, is strongly excited, reaching values as high as $\approx 30^\circ$ with respect to planet 1 (cf. the green dotted line in the middle panel of Figure 6.26).

In this regime of excitation of the planetary eccentricities coupled to the LK oscillations of the inner and outer stellar binaries, high values of the planetary eccentricities can be reached. For example, in Figure 6.24, for $7 \lesssim a_{\text{in}}/\text{AU} \lesssim 8$, e_4 becomes large enough for orbit crossing with planet 3 (indicated with the red hatched region). In the right axis in the top panel of Figure 6.24, we show with the solid black line the ratio $\mathcal{R}$ of the nodal precession time-scale associated with



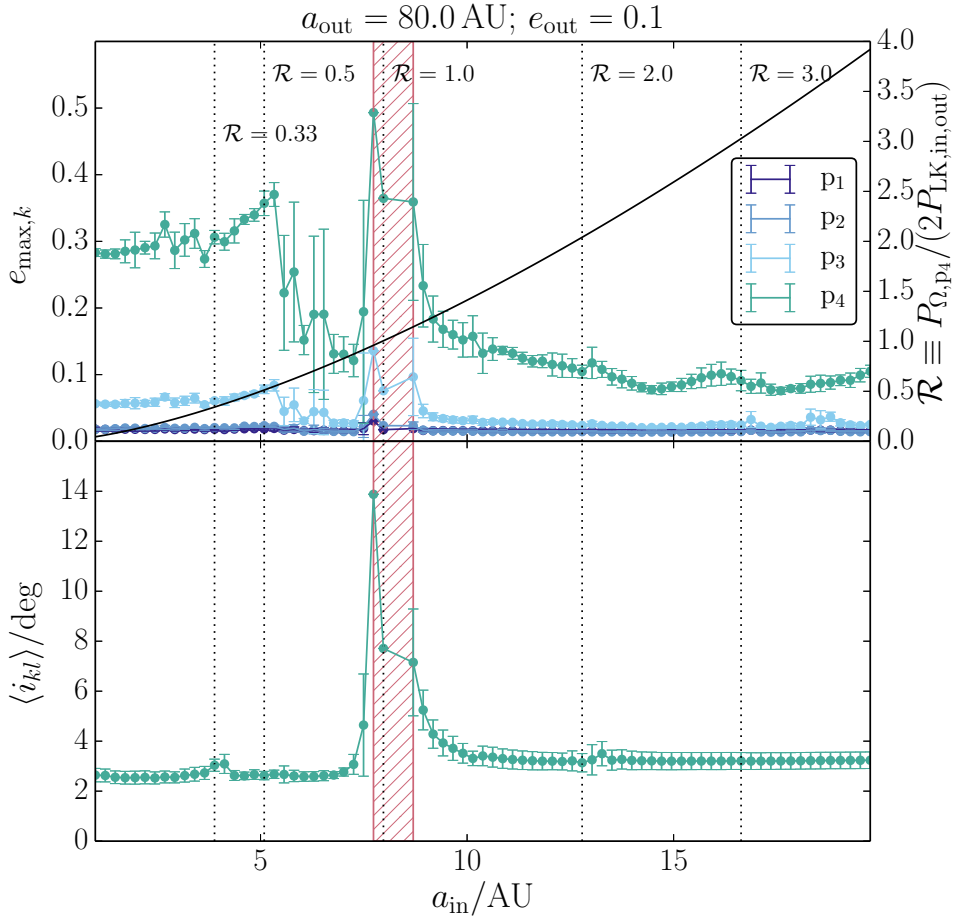


Figure 6.24: For a four-planet system orbited by a stellar binary (cf. Figure 6.23), the maximum eccentricities (top panel) and time-averaged mutual inclinations (bottom panel) of the planets as a function of a_{in} , with fixed $a_{\text{out}} = 80$ AU and $e_{\text{out}} = 0.1$. Each bullet corresponds to five different realizations with respect to Ω_s of the planets; the bullets (error bars) indicate the mean values (standard deviations). In the top panel, the black solid line shows the ratio $\mathcal{R}$ (right axis) of the nodal precession time-scale associated with the outermost planet due to planet-planet torques, to the nodal time-scale associated with inner-outer binary torques. Black vertical dotted lines show values of a_{in} corresponding to various integer ratio values of $\mathcal{R}$, indicated in the top panel. The red hatched region shows the values of a_{in} for which the orbit of the fourth planet is sufficiently eccentric to cross with that of the third planet.

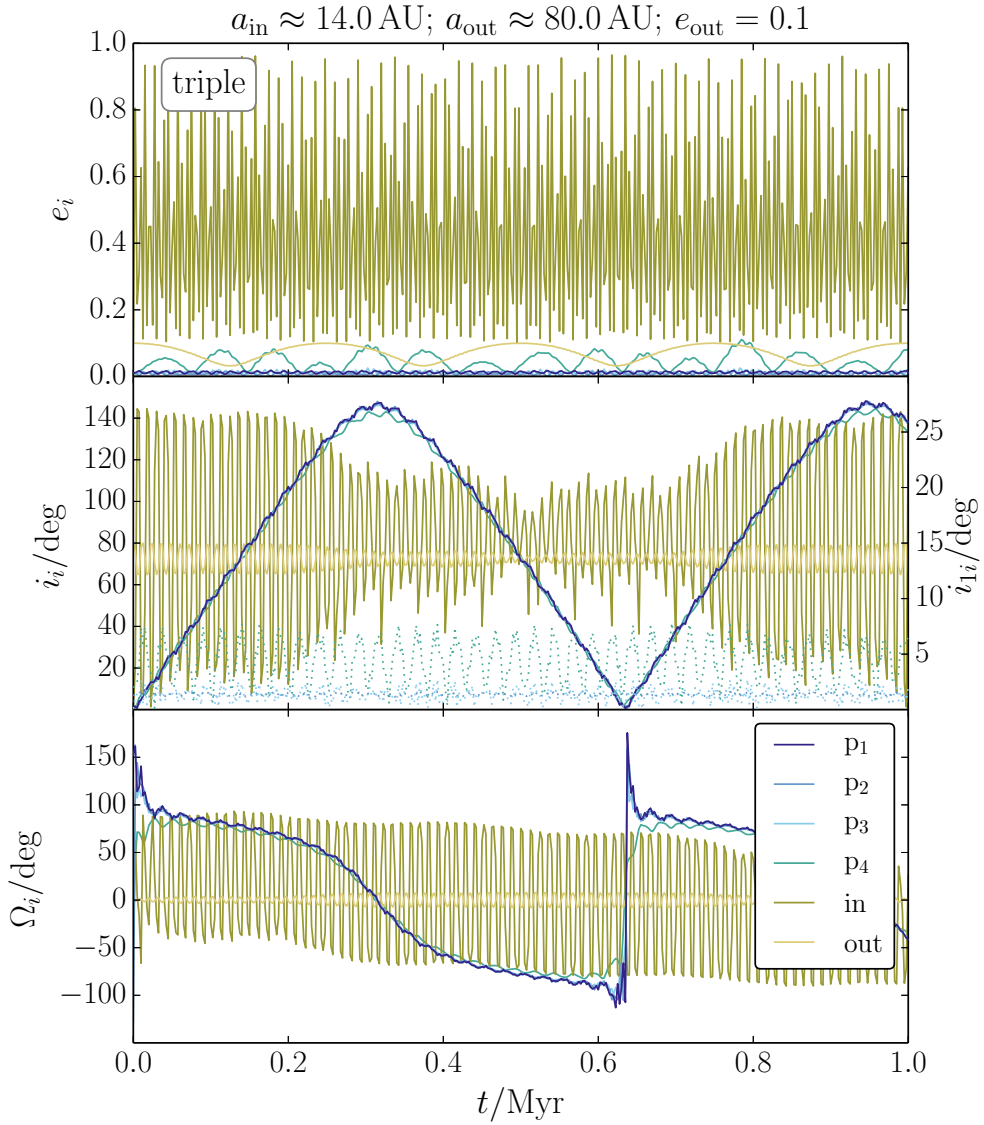
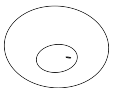


Figure 6.25: Example evolution of a system from Figure 6.24, with $a_{\text{in}} \approx 14.0 \text{ AU}$, $a_{\text{out}} \approx 80.0 \text{ AU}$ and $e_{\text{out}} = 0.1$. Refer to Figure 6.22 for the evolution of the equivalent system in the case of a stellar binary. The top, middle and bottom panels show the eccentricities, inclinations and longitudes of the ascending nodes, respectively, of the four planets and the inner and outer binaries. In the middle panel, the absolute inclinations i_i are shown with solid lines; for the planets, the inclinations relative to planet 1, i_{1i} (right axis), are shown with dotted lines.



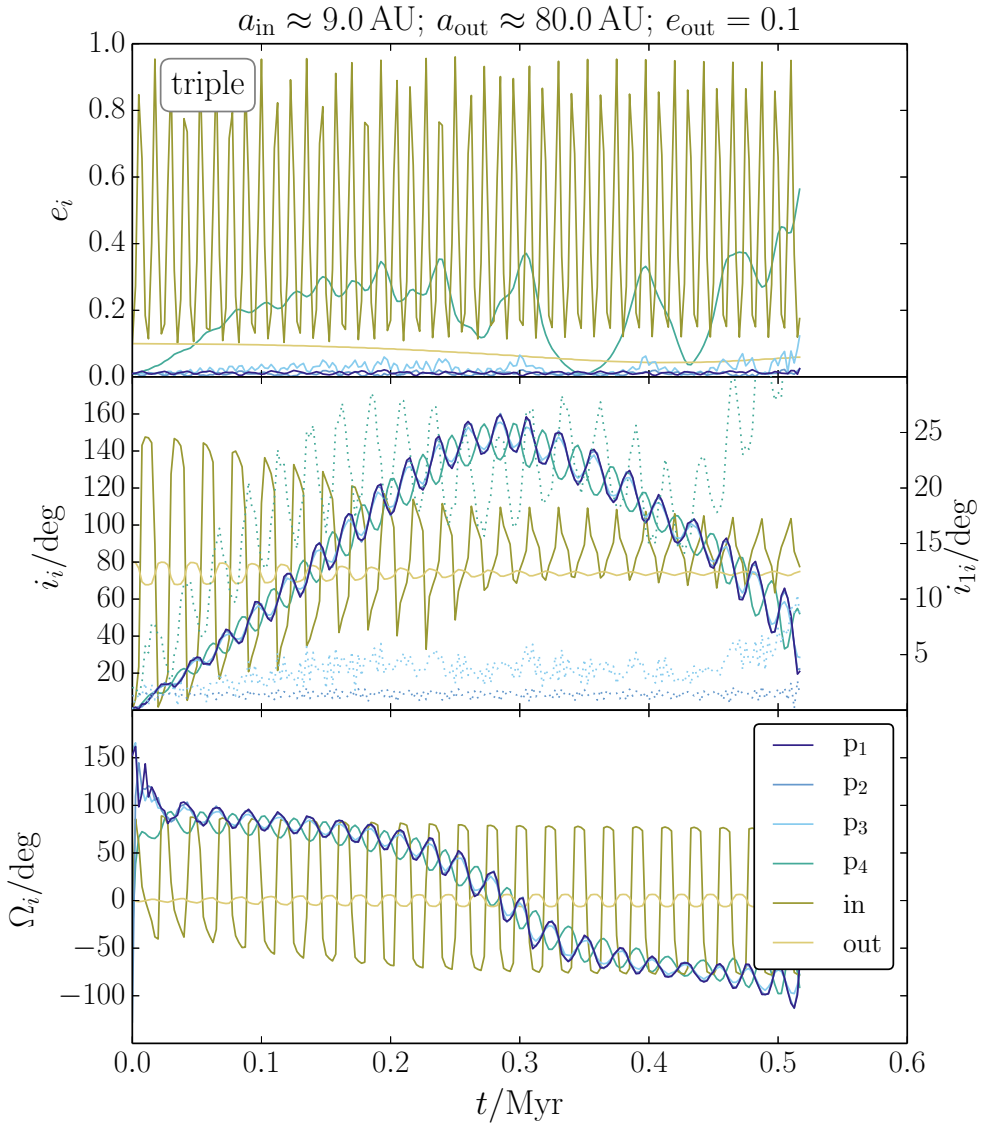


Figure 6.26: Similar to Figure 6.26, now with $a_{\text{in}} \approx 9.0 \text{ AU}$. The integration was stopped at $t \approx 0.5 \text{ Myr}$, when e_3 and e_4 became high enough for the third and fourth orbits to cross.

the outermost planet due to planet-planet torques, to the nodal time-scale associated with the inner-outer binary torques. Note that the latter cause oscillations in Ω_{in} and Ω_{out} on the same time-scale (see e.g. Figure 6.26). Vertical dashed lines indicate values of a_{in} for which $\mathcal{R}$ is equal to various integer ratios. The strongest excitations, likely leading to a destabilization of the planetary system, occur when $\mathcal{R} \sim 1$. Weaker excitations occur near other integer ratios.

In Figures 6.27 and 6.28, we show two figures similar to Figure 6.24 with the maximum eccentricities and time-averaged mutual inclinations as a function of a_{in} , for $a_{\text{out}} \approx 85.0$ AU and $a_{\text{out}} \approx 90.0$ AU, respectively. For $a_{\text{out}} \approx 85.0$ AU, there is a mild excitation of e_4 in the case of a stellar binary, with $e_{\text{max},4} \approx 0.04$ (cf. Figure 6.19). However, in the case of a stellar triple, large values of e_4 can be reached, even leading to orbit crossings, if $\mathcal{R} \sim 1$. Interestingly, there is a high and wide peak in $e_{\text{max},4}$ near $\mathcal{R} = 2/3$; this is not the case for $a_{\text{out}} \approx 80.0$ AU (cf. the top panel of Figure 6.24). For $a_{\text{out}} \approx 90.0$ AU (cf. Figure 6.28), there is no strong excitation of e_4 in the case of a stellar binary (cf. Figure 6.19). In the case of a stellar triple, excitation is still possible if $\mathcal{R} \sim 1$.

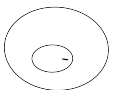
Generally, the locations and relative heights of the peaks with respect to $\mathcal{R}$ change with different parameters of the stellar binary orbit. Further investigation into this behaviour is left for future work. Nevertheless, the results found here already suggest that compared to the case of a single stellar companion, the ‘binary nature’ of the companion in the case of a stellar triple can result in both dynamical *protection* and *destabilization* of the planets against the stellar torques, depending on the parameters. Given that destabilization only occurs for specific values of the ratio $\mathcal{R}$ whereas protection occurs in the other cases, we expect that the typical effect is that of protection.

6.4 Secular constraints in observed systems

6.4.1 30 Arietis

30 Arietis (commonly abbreviated as 30 Ari) is a ‘2+2’ quadruple star system consisting of F-type main-sequence stars, with an age of ≈ 0.9 Gyr (Guenther et al. 2009). The subsystem 30 Ari A is a spectroscopic binary with a period of 1.109526 d and an eccentricity of 0.062 (Morbey & Brosterhus 1974). This subsystem forms a visual binary, with a period of ≈ 34000 yr, with a subsystem containing 30 Ari B, an F6V star (Tokovinin et al. 2006). The latter star harbours 30 Ari Bb, a massive planet ($m \sin i \approx 9.88 M_J$) orbiting 30 Ari B in ≈ 335 d with an eccentricity of 0.289 (Guenther et al. 2009). Recently, Riddle et al. (2015) found an $\approx 0.5 M_\odot$ companion star to 30 Ari B, 30 Ari C, in an orbit with a projected separation of ≈ 22.3 AU (Roberts et al. 2015). In Figure 6.29, we show, in a mobile diagram, the hierarchy of the system, which is effectively a hierarchical quintuple system. To date, 30 Ari is the second confirmed stellar quadruple system known to host at least one exoplanet; the first is Ph1b (Schwamb et al. 2013).

Apart from the masses of the components and (most of) the semimajor axes and eccentricities, the orbital properties of 30 Ari, in particular the relative inclinations, are unknown. Here, we investigate the secular dynamical evolution of the system, focusing on the orbit of the planet 30 Ari Bb (Section 6.4.1). Using secular stability arguments, we constrain the relative inclinations. Furthermore, we investigate the possibility of additional planets around 30 Ari B (Section 6.4.1).



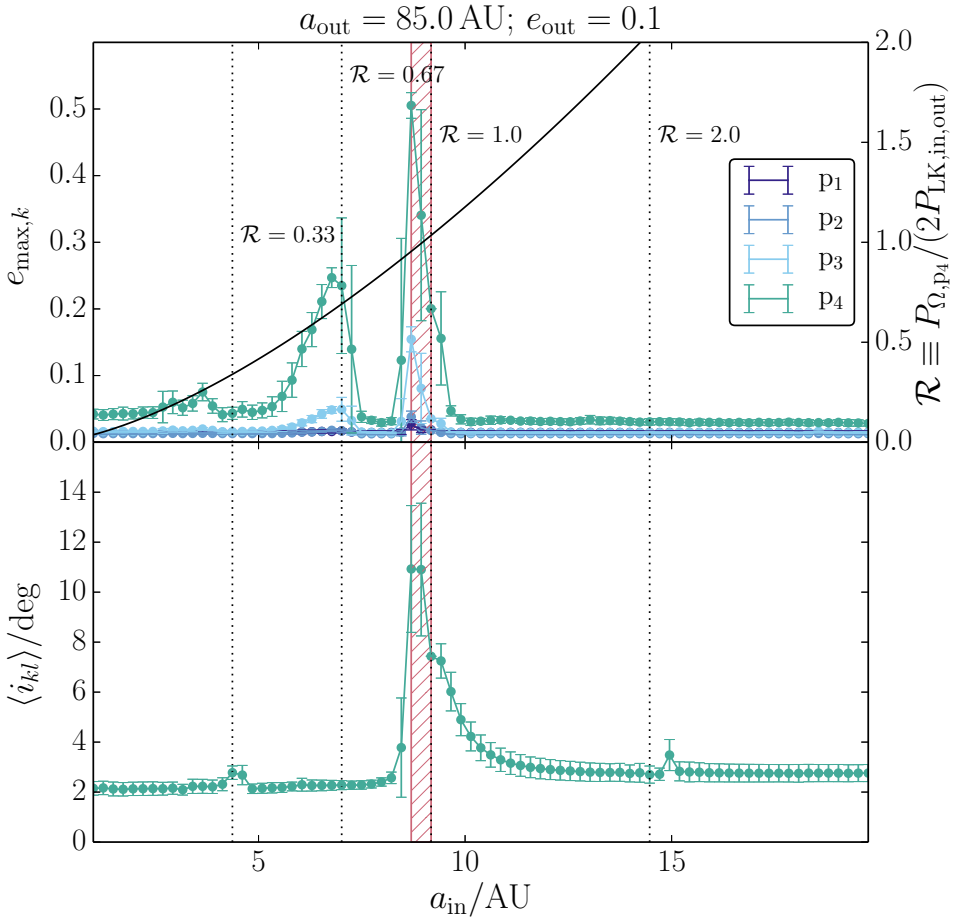


Figure 6.27: Similar to Figure 6.24, now assuming $a_{\text{out}} = 85 \text{ AU}$ and $e_{\text{out}} = 0.1$.

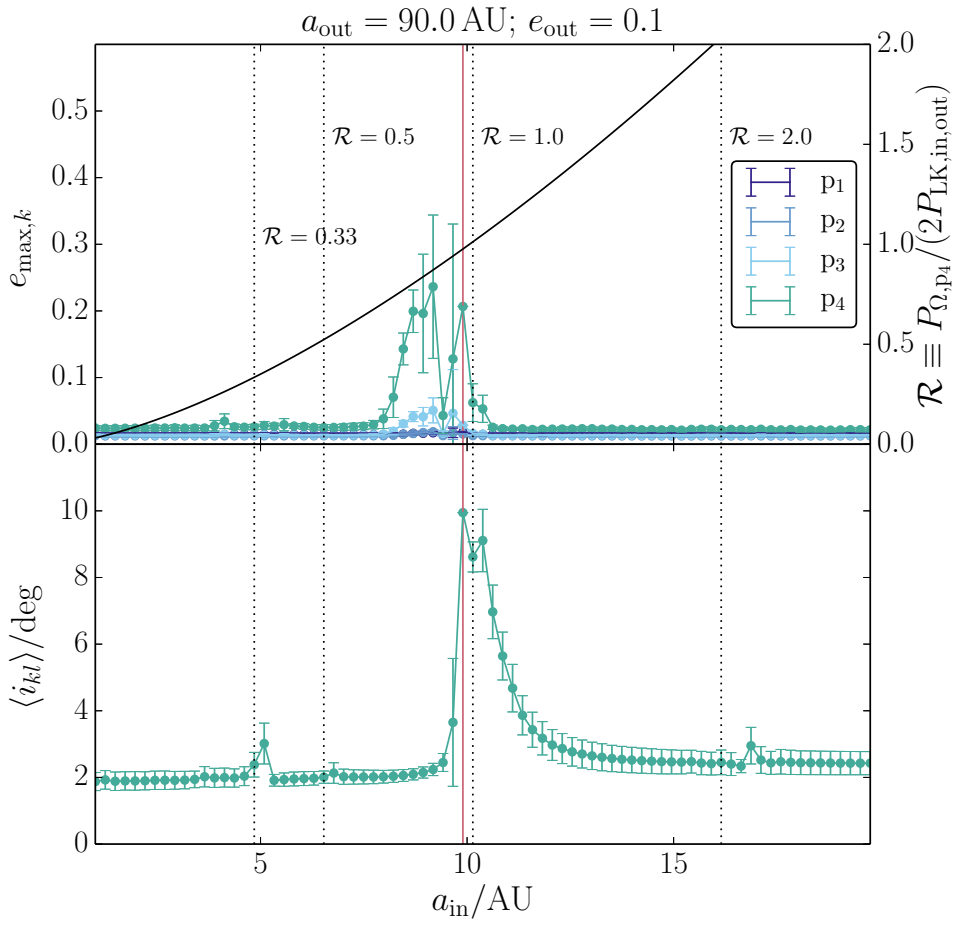
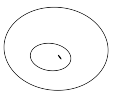


Figure 6.28: Similar to Figure 6.24, now assuming $a_{\text{out}} = 90 \text{ AU}$ and $e_{\text{out}} = 0.1$.



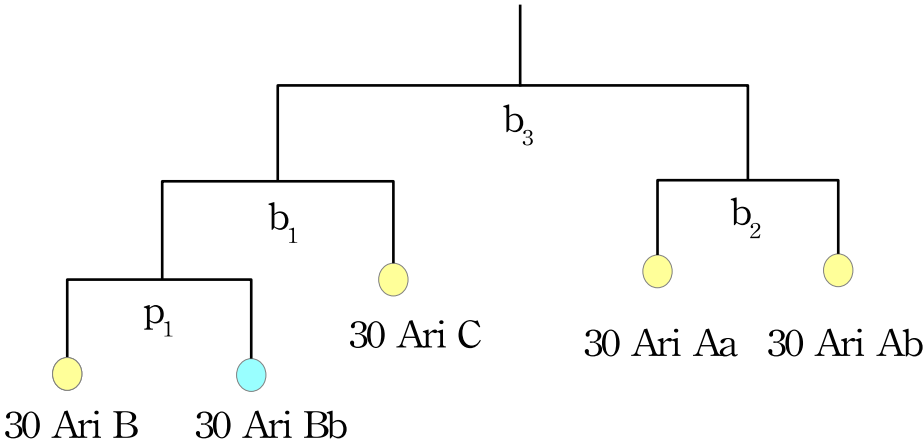


Figure 6.29: 30 Ari represented in a mobile diagram. The assumed orbital parameters are listed in Table 6.1.

Parameter	Value(s)	Parameter	Value(s)
$m_B/M_\odot$	1.13 ^a	P_{b_2}/d	1.109526 ^d
m_{Bb}/M_J	9.88 ^a	e_{b_2}	0.062 ^d
$m_C/M_\odot$	0.5 ^b	P_{b_3}/yr	34,000 ^e
$m_{Aa}/M_\odot$	1.31 ^a	e_{b_3}	0.01 – 0.9 ^c
$m_{Ab}/M_\odot$	0.14 ^a	P_{p_1}/d	335.1 ^a
a_{b_1}/AU	22.3 ^b	e_{p_1}	0.289 ^a
e_{b_1}	0.01 – 0.80 ^c		

^a Guenther et al. (2009)

^b Roberts et al. (2015)

^c Rayleigh distribution $dN/de_i \propto e_i \exp(-\beta e_i^2)$;
 $\langle e_i^2 \rangle^{1/2} = \beta^{-1/2} = 0.3$ (Raghavan et al. 2010)

^d Morbey & Brosterhus (1974)

^e Tokovinin et al. (2006)

Table 6.1: Assumed parameters for 30 Ari. Refer to Figure 6.29 for the definitions of the orbits.

Our assumed parameters are listed in Table 6.1. For 30 Ari C (i.e. orbit b_1), only the projected separation of 22.3 AU is known. Here, we adopt a semimajor axis of $a_{b_1} = 22.3$ AU, and consider a distribution of different eccentricities (see below).

First, we consider the importance of orbit b_2 with regard to the secular evolution of orbit b_1 and its constituents. As shown in Section 6.2.4, this can be estimated by considering the ratio $\mathcal{R}_{b_1;b_2;b_3}^{(bb)}$ of the LK time-scale for the (b_1, b_3) pair, compared to that for the (b_2, b_3) pair. Setting the coefficients C in equation (6.10) equal to unity for simplicity, this ratio is given by

$$\begin{aligned}\mathcal{R}_{b_1;b_2;b_3}^{(bb)} &\sim \left(\frac{a_{b_2}}{a_{b_1}}\right)^{3/2} \left(\frac{M_{b_1}}{M_{b_2}}\right)^{3/2} \\ &\approx \left(\frac{2.4 \times 10^{-2} \text{ AU}}{22.3 \text{ AU}}\right)^{3/2} \left(\frac{1.63 M_\odot}{1.45 M_\odot}\right)^{3/2} \approx 4.1 \times 10^{-5},\end{aligned}\quad (6.22)$$

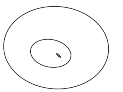
which is $\ll 1$. This shows that binary b_2 , i.e. the 30 Ari A spectroscopic binary, is effectively a point mass from a secular dynamical point of view. Nevertheless, in the numerical integrations below, we do not make this approximation, i.e. we resolve all secular interactions.

Single planet

We carried out an ensemble of integrations of 5000 systems in which the unknown orbital parameters were sampled from assumed distributions. In particular, e_{b_1} and e_{b_3} were sampled from a Rayleigh distribution $dN/de_i \propto e_i \exp(-\beta e_i^2)$ with $\langle e_i^2 \rangle^{1/2} = \beta^{-1/2} = 0.3$ (Raghavan et al. 2010) in the range $0.01 < e_{b_1} < 0.8$ and $0.01 < e_{b_3} < 0.9$. The upper limit on e_{b_1} is motivated by the requirement of (short-term) dynamical stability of the planet with respect to orbit b_1 . Using the formulae of Holman & Wiegert (1999), we find that the maximum value of e_{b_1} such that the planet in an S-type orbit around 30 Ari B is stable against the perturbation of 30 Ari C, is ≈ 0.80 . The upper limit on e_{b_3} is motivated by the requirement of (short-term) dynamical stability of orbit b_1 with respect to orbit b_3 . Using the stability criterion on Mardling & Aarseth (2001), we find that the largest value of e_{b_3} for dynamical stability is ≈ 0.90 . The inclinations of all orbits were sampled from a distribution that is linear in $\cos(i_i)$. The arguments of pericentre ω_i and the longitudes of the ascending nodes Ω_i were sampled assuming random distributions.

We integrated the systems for a duration of $t_{\text{end}} = 1$ Gyr. We also included precession due to relativity in all orbits (cf. Appendix 6.A.7) and precession due to tidal and rotational bulges in the planet (cf. Section 6.2.5). For the planet, we assumed an apsidal motion constant of $k_{\text{AM}} = 0.52$ and a spin frequency of 1 d^{-1} , with the spin vector initially aligned with binary p_1 . Tidal dissipation in the planet was not included in detail; instead, we assumed that strong tidal dissipation would occur as soon as the pericentre of the orbit of the planet reached a value smaller than $3 R_\odot$. In this case, we considered the system to be unstable, in the sense that the semimajor axis of the planet would decrease to a smaller value compared to the observed value. Furthermore, we considered a realization to be unstable if the eccentricities of any the orbits would imply orbit crossing.

In Figure 6.30, we show in the $(i_{p_1-b_1}, i_{b_1-b_3})$ -plane the systems that remained stable (green bullets) and those that became unstable (red crosses). Distributions of these quantities for the two cases are shown in the top and right insets, respectively. There is a strong dependence of



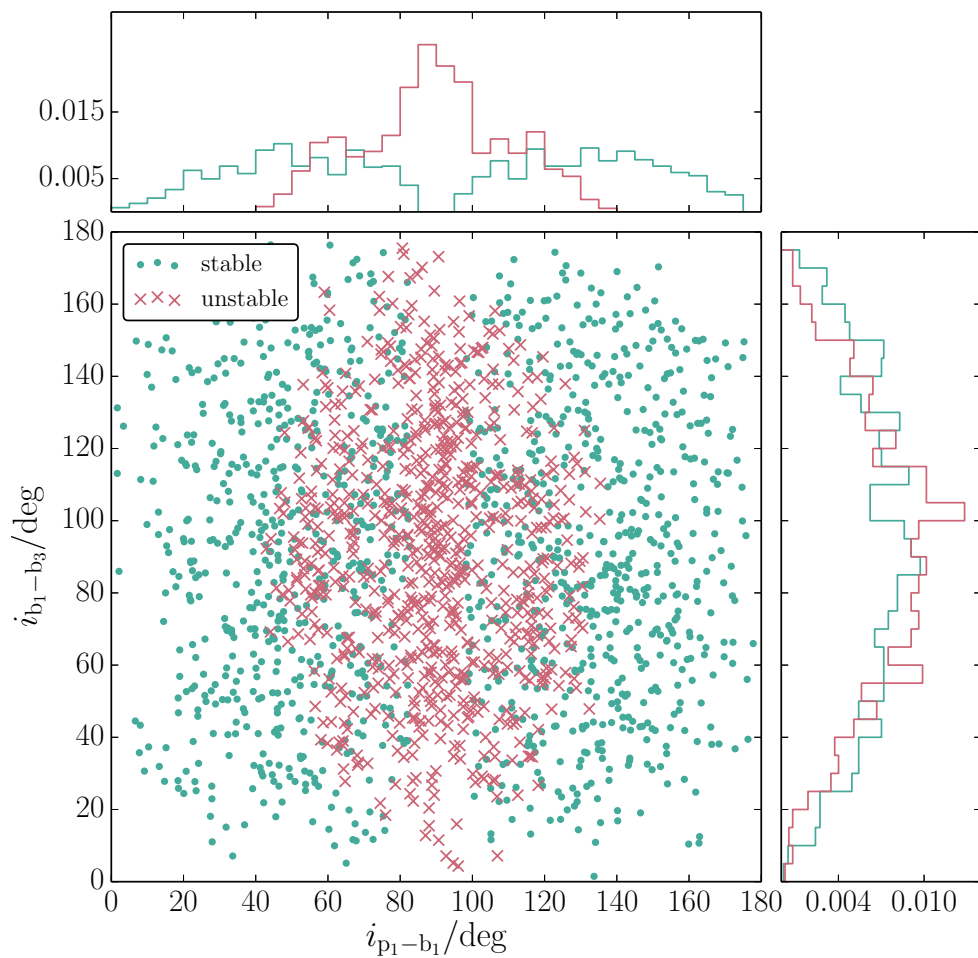


Figure 6.30: Outcomes of integrations of Monte Carlo realizations of 30 Ari in the (i_{p1-b1}, i_{b1-b3}) -plane. Systems that are stable (unstable) are shown with green bullets (red crosses). The insets show distributions of these quantities for these two cases with green and red lines, respectively.

instability on $i_{p_1-b_1}$, with instability occurring between $\approx 40^\circ$ and 140° . For $85^\circ \lesssim i_{p_1-b_1} \lesssim 95^\circ$, all systems are unstable, regardless of $i_{b_1-b_3}$, e_{b_1} or e_{b_3} .

This strong dependence is the result of LK cycles induced in orbit p_1 by orbit b_1 . To describe this more quantitatively, we show in Figure 6.31 the maximum eccentricities reached in orbits p_1 (blue bullets) and b_1 (red crosses) as a function of $i_{p_1-b_1}$. In the bottom panel, the maximum eccentricities are normalized to the initial values; note that the initial value of e_{p_1} is 0.289 for all systems, whereas the initial eccentricities e_{b_1} vary per system (cf. Table 6.1).

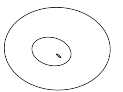
The relative inclination $i_{p_1-b_1}$ should be high enough for eccentricity excitation in orbit p_1 , as expected for LK cycles. The dependence of $e_{\max,p_1}$ on $i_{p_1-b_1}$ is symmetric with respect to $i_{p_1-b_1} = 90^\circ$. When $i_{p_1-b_1}$ approaches 90° , $e_{\max,p_1}$ approaches unity. Consequently, there are no stable systems with $85^\circ \lesssim i_{p_1-b_1} \lesssim 95^\circ$. Depending on the initial argument of pericentre of orbit p_1 , there is a sharp cutoff ($\omega_{p_1} = 90^\circ$) or a smooth transition ($\omega_{p_1} = 0^\circ$) from low to high inclinations. This is illustrated by the two black dashed lines in Figure 6.31, which show an analytic calculation of the maximum eccentricity for these two limiting values of ω_{p_1} for hierarchical triples, computed using the analytic solutions of Kinoshita & Nakai (2007). Note that the latter authors assumed the quadrupole-order limit, and that the ‘outer’ semimajor axis (in this case, a_{b_1}) is much larger than the ‘inner’ semimajor axis (in this case, a_{p_1}). Nevertheless, the analytic solutions generally give good estimates for the largest and smallest values of the numerically calculated $e_{\max,p_1}$ as a function of $i_{p_1-b_1}$.

Exceptions of the latter occur near $i_{p_1-b_1} \approx 60^\circ$ and $i_{p_1-b_1} \approx 120^\circ$ (approximately symmetrically with respect to $i_{p_1-b_1} = 90^\circ$), where $e_{\max,p_1}$ can be higher than the top black analytic line in Figure 6.31. This can be understood by noting that e_{b_1} can be highly excited for $40^\circ \lesssim i_{p_1-b_1} \lesssim 60^\circ$ (see below). In this case, the high eccentricities of orbit b_1 lead to strong perturbations of p_1 .

In Figures 6.31 and 6.32, the dependence of $e_{\max,b_1}$ as a function of $i_{p_1-b_1}$ and $i_{b_1-b_3}$, respectively, is shown with the red crosses. There is a clear dependence of $e_{\max,b_1}$ on $i_{p_1-b_1}$, with excitations of e_{b_1} reaching ≈ 0.8 if $i_{p_1-b_1}$ is around 50° or 130° ; note that $e_{b_1} = 0.8$ is close to the limit of dynamical stability. This dependence can be understood by noting that orbit p_1 can induce precession on orbit b_1 , reducing the amplitude of eccentricity oscillations in b_1 (this effect was discussed previously in Hamers et al. 2015). The torque of orbit p_1 on b_1 is weakest when p_1 and b_1 are inclined by ~ 50 or 130° . Therefore, $e_{\max,b_1}$ shows a maximum near these values.

To understand this more quantitatively, we show in the bottom panel of Figure 6.31 with the dashed lines semi-analytic calculations of the maximum eccentricity in orbit b_1 as a function of $i_{p_1-b_1}$ for various values of $i_{b_1-b_3}$, using a method similar to that of Hamers et al. (2015). Note that for $50^\circ \lesssim i_{p_1-b_1} \lesssim 130^\circ$, this method does not yield a maximum eccentricity (i.e. no solutions can be found). The semi-analytic curves approximately capture the low-inclination dependence of $e_{\max,b_1}$ on $i_{p_1-b_1}$. The larger $i_{b_1-b_3}$, the larger $e_{\max,b_1}$. The dependence of $e_{\max,b_1}$ on $i_{p_1-b_1}$ explains why in Figure 6.32, where $i_{p_1-b_1}$ varies per point, $e_{\max,b_1}$ is typically below the black dashed analytic curves, which assume only a perturbation by b_3 , and no additional precession due to orbit p_1 .

These results show that it is unlikely that orbits p_1 and b_1 are highly inclined (i.e., a mutual inclination close to 90°). If this were the case, then the planet would become highly eccentric and tidally interact with, or collide with 30 Ari B. Furthermore, more moderate inclinations of $40^\circ \lesssim i_{p_1-b_1} \lesssim 60^\circ$ and $120^\circ \lesssim i_{p_1-b_1} \lesssim 140^\circ$ are also not likely because orbit b_1 could be excited to high eccentricity, up to ≈ 0.8 . This would trigger a dynamical instability in the



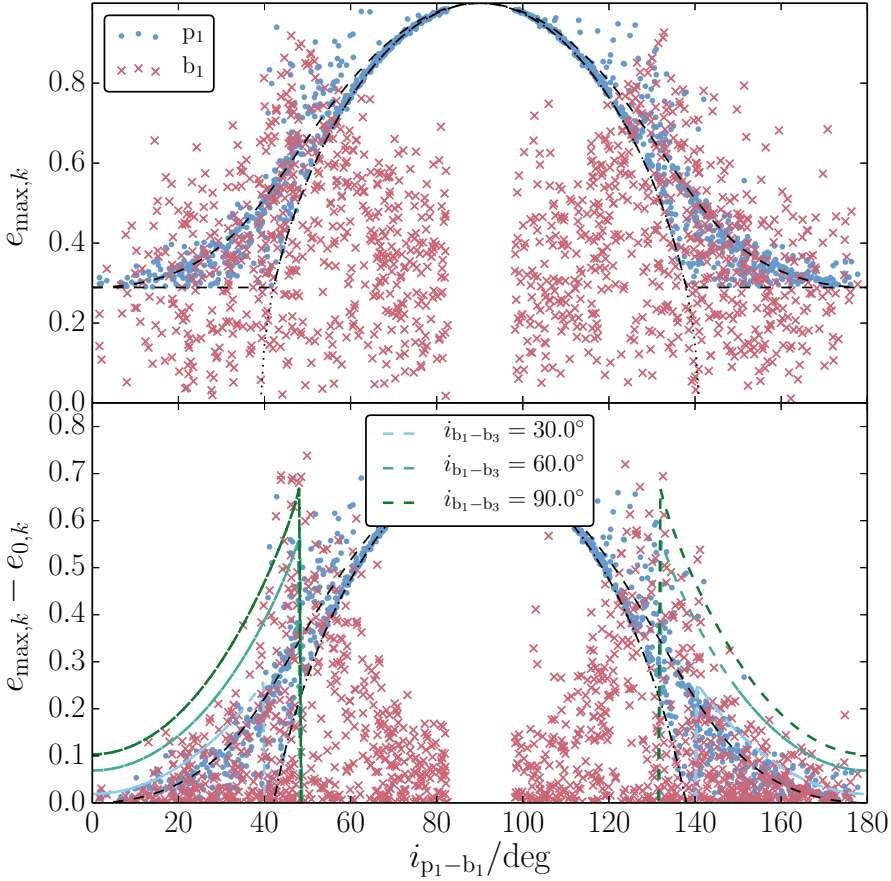


Figure 6.31: For the integrations of Monte Carlo realizations of 30 Ari, the maximum eccentricities reached in orbits p_1 (blue bullets) and b_1 (red crosses) as a function of $i_{p_1-b_1}$. In the bottom panel, the maximum eccentricities are normalized to the initial values; note that in the case of e_{b_1} the initial eccentricities vary per system. In the top panel, the two black dashed lines show an analytic calculation of the maximum eccentricity for two limiting values of ω_{p_1} (Kinoshita & Nakai 2007), $\omega_{p_1} = 0^\circ$ (upper dashed line) and $\omega_{p_1} = 90^\circ$ (lower dashed line). The black dotted line shows the well known relation $e_{\max,p_1} = \sqrt{1 - (5/3) \cos^2(i_{p_1-b_1})}$, in which $\omega_{p_1} = 90^\circ$ and an initial eccentricity of zero are assumed. In the bottom panel, coloured dashed lines show semi-analytic calculations of the maximum eccentricity in orbit b_1 for various values of $i_{b_1-b_3}$, using a method similar to that of Hamers et al. (2015). Note that for $50^\circ \lesssim i_{p_1-b_1} \lesssim 130^\circ$, this method does not yield a maximum eccentricity (i.e. no solutions can be found).

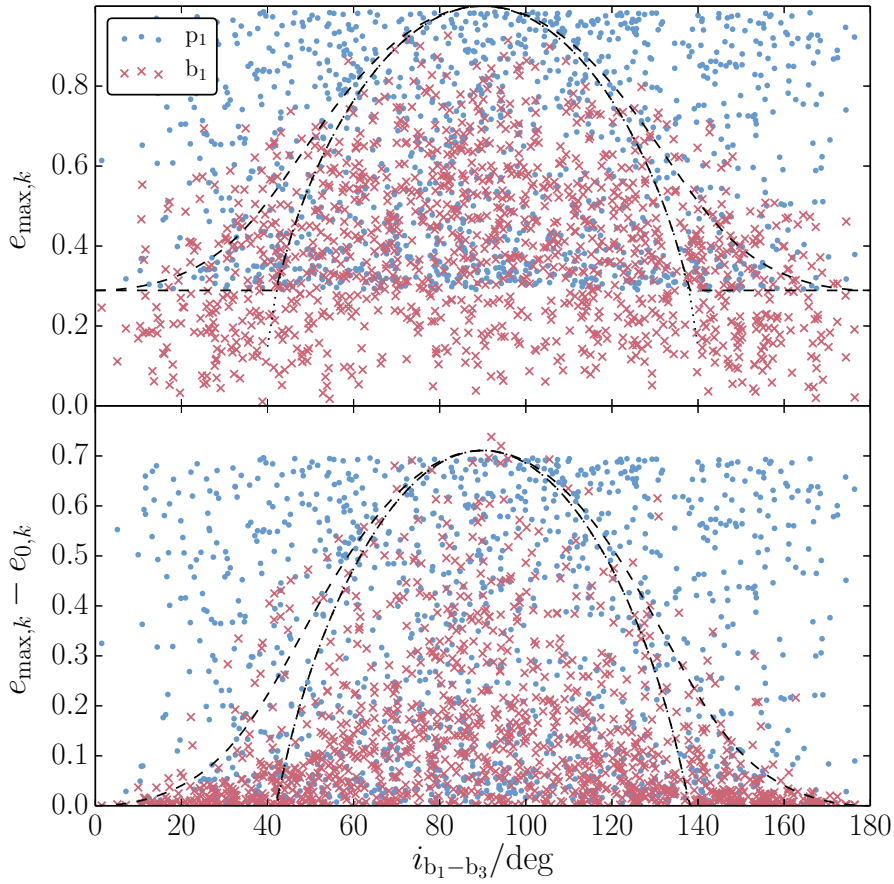
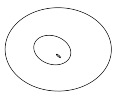


Figure 6.32: Similar to Figure 6.31, now showing the maximum eccentricities as a function of $i_{b_1-b_3}$. In the bottom panel, the maximum eccentricities are normalized to the initial values. The black lines are the same as in Figure 6.31.



planetary orbit, possibly resulting in the ejection of the planet from the system, or a collision with one of the stars. Therefore, low inclinations are more likely, i.e. $0^\circ \lesssim i_{p_1-b_1} \lesssim 40^\circ$ or $140^\circ \lesssim i_{p_1-b_1} \lesssim 180^\circ$ (based on these results, we cannot make constraints with respect to prograde or retrograde orbits). Because of the degeneracy of $e_{\max, b_1}$ with respect to $i_{p_1-b_1}$ and $i_{b_1-b_3}$, no strong constraints can be put on $i_{b_1-b_3}$.

Constraints on additional planets

As a further application, we consider one plausible realization of 30 Ari, and insert an additional planet in orbit around 30 Ari B, outside of the orbit of the observed planet 30 Ari Bb and within the orbit of 30 Ari C. Effectively, the system is then a hierarchical sextuple system. The orbit of the additional planet (p_2) is assumed to be initially coplanar with orbit p_1 , and its eccentricity is $e_{p_2} = 0.01$. Furthermore, we adopt a mutual inclination of p_1 and p_2 with respect to orbit b_1 of $i_{p_1-b_1} = i_{p_2-b_1} = 40^\circ$. The inclination between orbits b_1 and b_3 is also assumed to be $i_{b_1-b_3} = 40^\circ$. We assume $e_{b_1} = 0.01$ and $e_{b_3} = 0.21$.

As free parameters, we consider the mass m_{p_2} and semimajor axis a_{p_2} . Using the formulae of Holman & Wiegert (1999), the lower limit on a_{p_2} based on dynamical stability with respect to 30 Ari B and 30 Ari Bb (i.e. p_1) is $a_{p_2} \gtrsim 2.9$ AU. Similar arguments with respect to 30 Ari C require $a_{p_2} \lesssim 7.6$ AU (assuming $e_{b_1} = 0.01$). Based on this, we take a linear grid in a_{p_2} with $3.5 \leq a_{p_2}/\text{AU} \leq 7.5$. The masses are taken from a linear grid in $\log_{10}(m_{p_2}/M_J)$, with $-3 \leq \log_{10}(m_{p_2}/M_J) \leq 1$.

In Figure 6.33, we show in the $(a_{p_2}, m_{p_2}/M_J)$ -plane the systems that remain stable (green points), and that become unstable (red crosses). Here, unstable systems are due to collisions or tidal interactions of p_1 with the star 30 Ari B (≈ 0.06 of the systems), or orbit crossings of the additional planet p_2 with the first planet (≈ 0.60 of the systems). There is a noticeable dependence of stability on time; to illustrate this, we include two integration times in Figure 6.33: 100 Myr (top panel) and 1 Gyr (bottom panel). As can be expected, more systems become unstable as time progresses. Regardless of age, the additional planet should be placed sufficiently far from the innermost planet. Whereas short-term dynamical stability would allow $3.5 \lesssim a_{p_2}/\text{AU} \lesssim 6$, this is no longer the case when taking into account secular evolution. For $t_{\text{end}} = 1$ Gyr, the minimum semimajor axis, ≈ 6 AU, is a factor of ~ 2 larger compared to the situation when only short-term dynamical stability is taken into account (2.9 AU).

Generally, there are no strong constraints on the mass of the additional planet. This can be understood by noting that the excitation of the eccentricity of p_2 by b_1 is independent of m_{p_2} as long as $m_{p_2} \ll m_C$.

Furthermore, in Figure 6.34, we show the maximum eccentricities in orbits p_1 , p_2 and b_1 as a function of the parameters a_{p_2} (left-hand panel) and m_{p_2}/M_J (right-hand panel), with $t_{\text{end}} = 1$ Gyr. For sufficiently massive planets and sufficiently large semimajor axes, the additional planet can increase the maximum eccentricities of the orbit of p_1 , and, to a minor extent, the orbit of b_1 .

To conclude, assuming a low mutual inclination between orbits p_1 and b_1 and $e_{b_1} = 0.01$ and $e_{b_3} = 0.21$, we find a region of secular dynamical stability of an additional planet around 30 Ari B, outside of and coplanar with the orbit of 30 Ari Bb. The semimajor axis of the orbit of this additional planet should be confined to a narrow range, $6 \lesssim a_{p_2}/\text{AU} \lesssim 7.6$, whereas there are no strong constraints on mass of the planet (cf. the bottom panel of Figure 6.33). Furthermore, the eccentricity of the orbit is likely high, possibly as high as ~ 0.6 (cf. Figure 6.34).

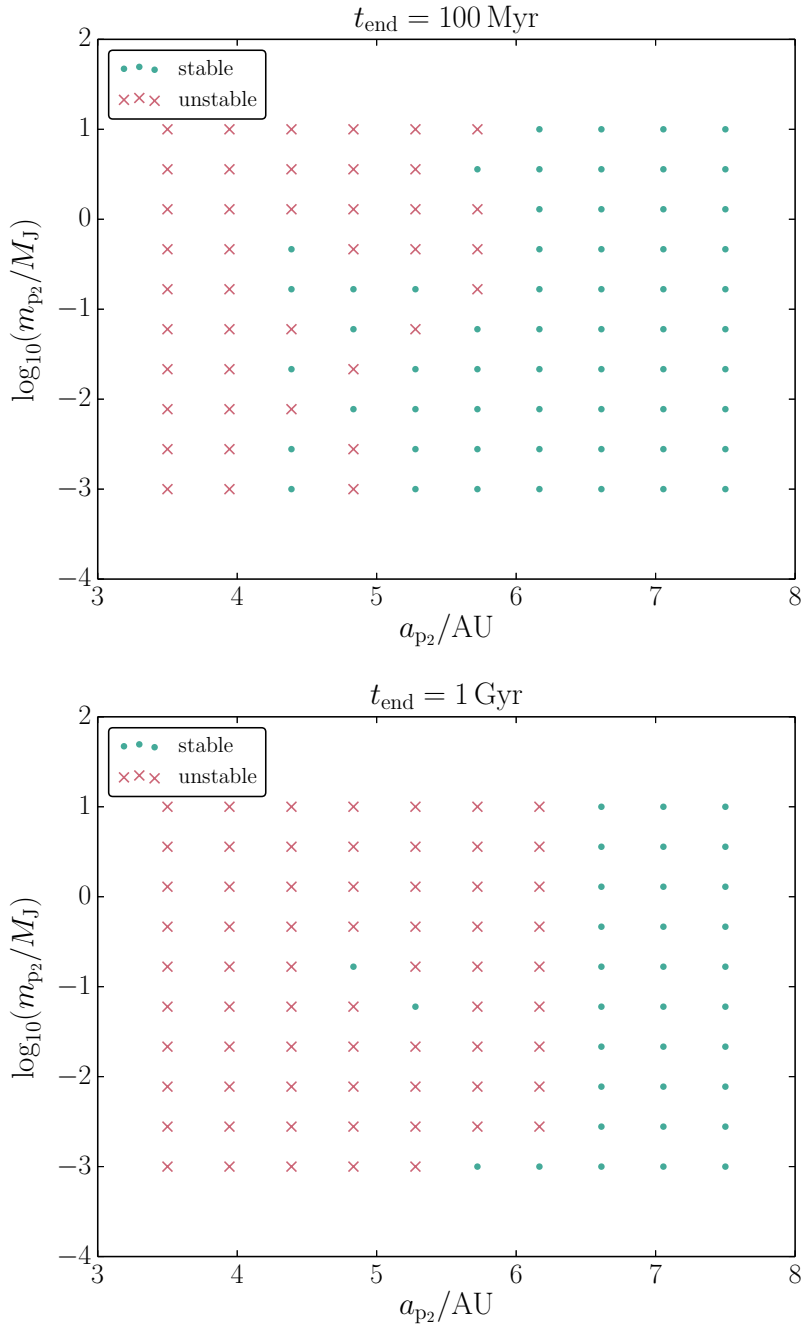
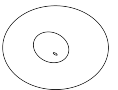


Figure 6.33: Stable (green bullets) and unstable (red crosses) systems in the $(a_{p_2}, m_{p_2}/M_J)$ -plane when taking into account an additional planet p_2 orbiting beyond 30 Ari Bb, and inside the orbit of 30 Ari C. The integration time is either 100 Myr (top panel) or 1 Gyr (bottom panel).



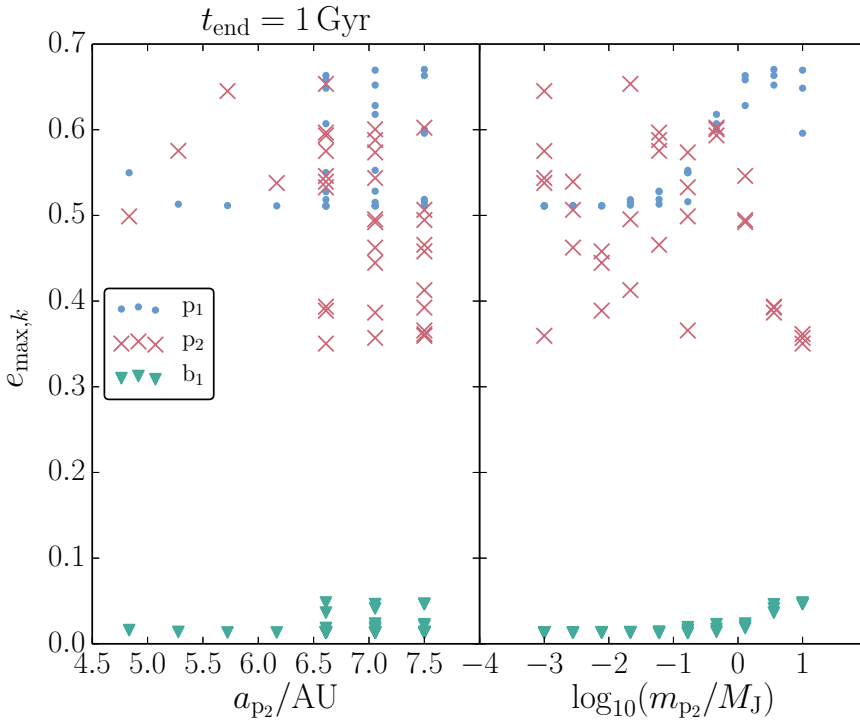


Figure 6.34: The maximum eccentricities in orbits p_1 (blue bullets), p_2 (red crosses) and b_1 (green triangles), as a function of the parameters a_{p_2} (left-hand panel) and m_{p_2}/M_J (right-hand panel). The integration time is $t_{\text{end}} = 1 \text{ Gyr}$.

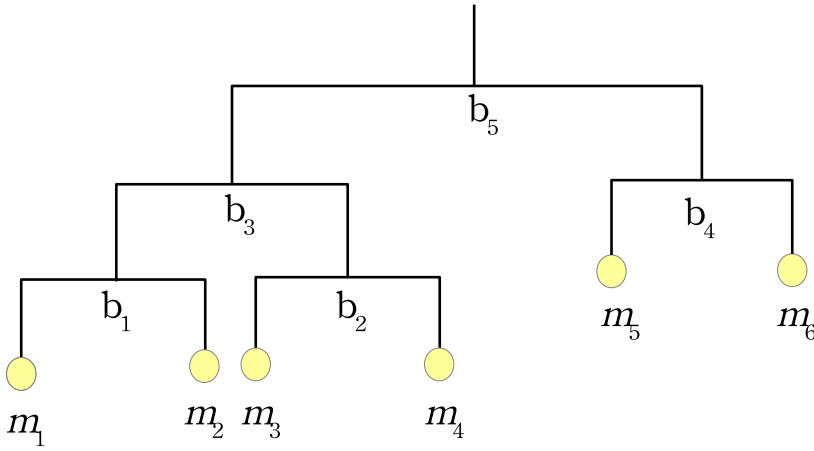


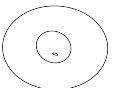
Figure 6.35: A mobile diagram of Mizar (the subsystem with binaries b_1 , b_2 and b_3) and Alcor (the subsystem with binary b_4).

6.4.2 Mizar and Alcor

Mizar and Alcor is perhaps one of the most well-known nearby hierarchical stellar multiple systems. To date, it is known to be a sextuple system. Mizar and Alcor form a visual stellar binary, which is famous for being used as a vision test (Allen 1899). Mizar itself is also a visual binary, known from as early as 1617, when Benedetto Castelli reported resolving it in a letter to Galileo Galilei (Ondra 2004; Siebert 2005). Both components of Mizar, Mizar A and B, are spectroscopic binaries. Mizar A is the first spectroscopic binary known, found by Antonia Maury and reported by Pickering (1890); the spectroscopic binary Mizar B was discovered independently by Frost (1908) and Ludendorff (1908). Recently, Alcor was found to be a binary as well (Mamajek et al. 2010; Zimmerman et al. 2010). It has been disputed whether or not Mizar is gravitationally bound to Alcor. In the recent works of Mamajek et al. (2010); Zimmerman et al. (2010), it was found that this is likely the case. To our knowledge, the secular dynamics of the system have not been explored.

In Figure 6.35, we show the hierarchy of the system in a mobile diagram. Mizar forms a ‘2+2’ quadruple system, orbited by the binary Alcor. Based on the studies mentioned above, we assume masses $m_i/M_\odot = \{2.43, 2.50, 1.6, 1.6, 1.8, 0.3\}$, semimajor axes $a_i/\text{AU} = \{0.25, 0.9, 5.9 \times 10^2, 2.8 \times 10^2, 7.4 \times 10^4\}$ and eccentricities $e_i = \{0.53, 0.46, 0.3, 0.3, 0.3\}$. For the unknown eccentricities of b_3 , b_4 and b_5 , we assume a value of 0.3, a typical value for long-period binaries (Raghavan et al. 2010).

Below, we evaluate the characteristic quantities $\mathcal{R}^{(bb)}$ (cf. equation 6.10) for the relevant binary combinations, setting the quantities C to unity. For the ‘outermost’ binary-binary com-



bination (b_3, b_4, b_5) , we find

$$\mathcal{R}_{345}^{(\text{bb})} \sim \left(\frac{a_{b_4}}{a_{b_3}} \right)^{3/2} \left(\frac{m_1 + m_2 + m_3 + m_4}{m_5 + m_6} \right)^{3/2} \approx 8 \times 10^{-2}, \quad (6.23)$$

indicating that the binary nature of b_4 can typically be neglected. Moreover, even for the pair (b_3, b_5) , the LK time-scale (cf. equation 6.7 with $C = 1$) is $P_{\text{LK};35} \sim 33$ Gyr, which is much longer than the age of the system, ≈ 0.5 Gyr. Therefore, LK cycles in the outermost binaries b_3, b_4 and b_5 are not important during the lifetime of the system.

For the ‘innermost’ combination (b_1, b_2, b_3) , we find

$$\mathcal{R}_{123}^{(\text{bb})} \sim \left(\frac{a_{b_2}}{a_{b_1}} \right)^{3/2} \left(\frac{m_1 + m_2}{m_3 + m_4} \right)^{3/2} \approx 1 \times 10^1, \quad (6.24)$$

showing that LK cycles are typically more important in the (b_2, b_3) pair compared to the (b_1, b_3) pair. Nevertheless, resonant behaviour might be expected in some cases. Using equation (6.7), the individual LK time-scales are

$$P_{\text{LK};13} \sim 980 \text{ Myr}; \quad P_{\text{LK};23} \sim 74 \text{ Myr}, \quad (6.25)$$

showing that during the lifetime of the system, of the order of 10 eccentricity oscillations could have occurred in binary b_2 .

However, the above does not take into account relativistic precession. In fact, the 1PN time-scales in both binaries b_1 and b_2 are much shorter compared to their respective LK time-scales. Using equation (6.173), we find

$$t_{\text{1PN};b_1} \approx 0.1 \text{ Myr}; \quad t_{\text{1PN};b_2} \approx 4 \text{ Myr}, \quad (6.26)$$

showing that LK cycles are likely (nearly) completely quenched.

We show an example in Figure 6.36, where in the top panel, we show the eccentricities assuming only Newtonian terms, and in the bottom panel the pairwise 1PN terms are also included. Even though the initial mutual inclinations $i_{13} = 80^\circ$ and $i_{23} = 85^\circ$ are high, with the inclusion of the 1PN terms, oscillations are severely quenched. We conclude that secular evolution is not important in the Mizar and Alcor system.

6.5 Discussion

6.5.1 The validity of orbit averaging

In our derivations, we orbit averaged the Hamiltonian after the expansion in ratios of separation ratios over *all* orbits (cf. Section 6.2.3). As shown recently by Luo et al. (2016), in the case of hierarchical triple systems with moderately large octupole parameters ($\epsilon_{\text{oct}} \gtrsim 10^{-3}$) and large mass ratios of the tertiary body with respect to the inner binary ($m_3/(m_1 + m_2) \gtrsim 10^2$), the *double* averaging technique can result in errors compared to the true secular evolution (for example, as measured by the fraction in an ensemble of systems in which orbital flips occur). Physically, this discrepancy arises from oscillations of the inner orbital angular momentum vector on time-scales of the outer orbital period (Antonini & Perets 2012; Katz & Dong 2012;

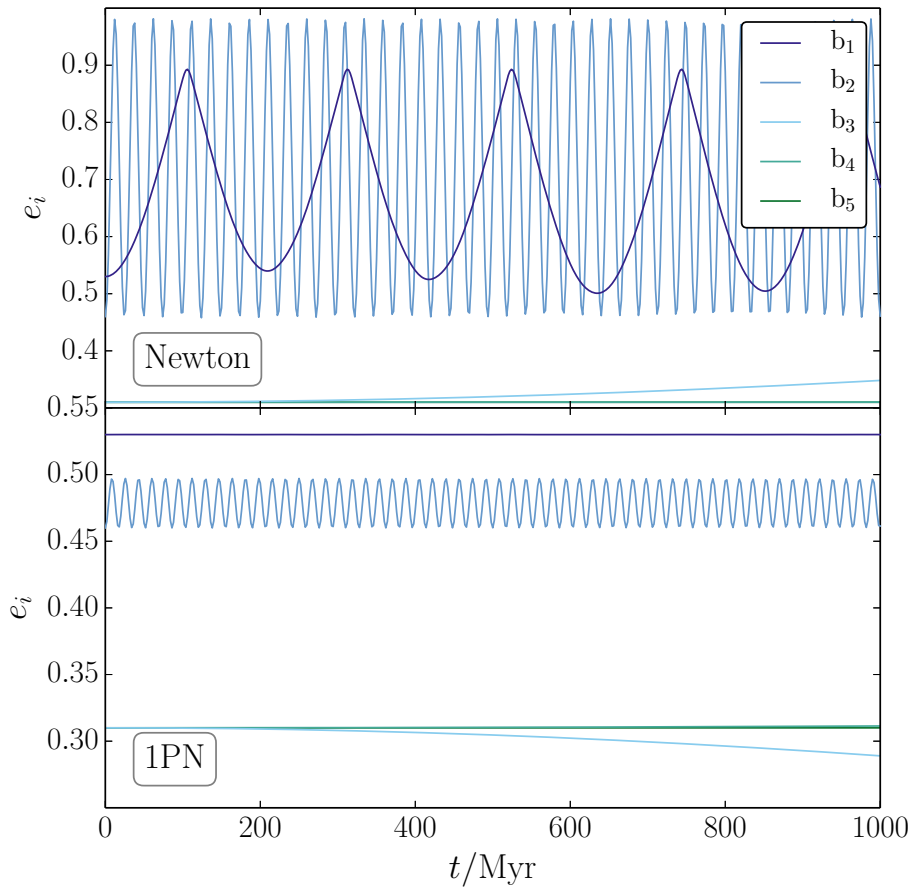
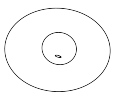


Figure 6.36: Example eccentricity evolution for Mizar and Alcor. In the top panel, we included only Newtonian terms, whereas in the bottom panel, we also included the pairwise 1PN terms. The initial mutual inclinations are $i_{13} = 80^\circ$ and $i_{23} = 85^\circ$.



Antognini et al. 2014), which invalidate the assumption of a fixed inner orbit during the orbit of the tertiary. These effects very likely also apply, to some cases, to the more general hierarchical N -body systems considered in this paper. A similar technique as used by Luo et al. (2016), i.e. to apply a ‘correction’ in the process of averaging over the outer body, might also be applied to hierarchical N -body systems. However, such an analysis is beyond the scope of this work.

6.5.2 Short-term instabilities

The assumptions on which the `SECULARMULTIPLE` algorithm is based (cf. Section 6.1) should be considered not only for the initial conditions of the systems, but also for the subsequent secular evolution. In particular, when eccentricities become high, there is the possibility that dynamical instabilities are triggered on short-term scales, i.e. of the order of the orbital periods. These dynamical instabilities are not taken into account in the secular equations of motion. For example, in multiplanet systems, dynamical stabilities are triggered when the Hill or Laplace boundaries are approached for sufficiently eccentric orbits.

For this reason, a number of stopping conditions are implemented in `SECULARMULTIPLE` that use various stability criteria to evaluate the importance of short-term dynamical instabilities. These criteria are approximate, however. In order to investigate more accurately whether the system is short-term dynamically stable, `SECULARMULTIPLE` allows for the user to easily switch to direct N -body integration. A complication of the latter is that the orbital phases are inherently not modelled in `SECULARMULTIPLE` but evidently, they are required for direct N -body integration. In practice, it is therefore necessary to carry out an ensemble of N -body integrations with e.g. randomly sampled orbital phases.

Typically, when a dynamical instability occurs, this happens on a time-scale that is short compared to secular time-scales. The resulting configuration is likely again hierarchical. Therefore, although `SECULARMULTIPLE` cannot be used to model the short-term phase of dynamical instability itself, once a new stable configuration has been attained, it can again be used to model the subsequent secular evolution. In future work, we intend to use this approach to study the secular stability of multiplanet systems, both in single- and multistar configurations.

6.6 Conclusions

We have presented a new method for studying the secular dynamics of hierarchical N -body systems composed of nested binaries with an arbitrary number of bodies and an arbitrary hierarchical configuration. This algorithm is suitable for studying the secular evolution of hierarchical systems with complex hierarchies, making long-term integrations of such systems feasible. We have applied the method to various multiplanet and multistar systems. Our main conclusions are as follows.

1. We have derived the expanded and orbit-averaged Hamiltonian for the system (a complete self-contained derivation is given in Appendix 6.A). Our main assumptions are that the system is hierarchical, i.e. the ratios x_i of the separation vectors are small, and that orbit averaging is applicable (i.e. the time-scales for angular momentum changes are much longer than orbital time-scales). The results from the expansion are summarized below.

- To first order in x_i (‘dipole order’), all terms vanish identically for *any* hierarchical configuration.

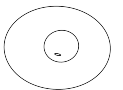
- To order x_i^2 ('quadrupole order'), the Hamiltonian consists of terms which, individually, depend on binary *pairs*. These pairwise terms are mathematically equivalent to the quadrupole-order terms of the hierarchical three-body Hamiltonian.
- To order x_i^3 ('octupole order'), the Hamiltonian consists of two types of terms: terms which, individually, depend on binary *pairs*, and terms which, individually, depend on binary *triplets*. The pairwise terms are mathematically equivalent to the octupole-order terms of the hierarchical three-body Hamiltonian. The triplet terms are mathematically equivalent to the term depending on three binaries at the octupole order in the hierarchical four-body Hamiltonian ('3+1' configuration; this term was first derived by Hamers et al. 2015).
- Generally, to order x_i^n , there are $n - 1$ types of terms (i.e. depending on binary pairs, triplets, etc.). The individual terms depend on at most n binaries. Whether all these terms appear is contingent on the system: if the system does not contain a sufficiently complex hierarchy, then not all types of terms may appear.

In summary, the general hierarchical N -body Hamiltonian can be constructed from the Hamiltonians of smaller subsystems. The higher the order of the expansion, the larger the number of binaries within these subsystems.

2. We have explicitly derived the expanded Hamiltonian up and including octupole-order, and including the pairwise hexadecupole- and dotriacontupole-order terms (cf. Appendix 6.A.5). We orbit-averaged these terms, and also gave a general expression for the orbit-averaged binary pair term to order n in terms of derivatives of polynomial functions (cf. Appendix 6.A.5). Note that for hierarchical triple systems, the Hamiltonian consists exclusively of these pairwise terms. Also, in the case of systems with more complex hierarchies, the pairwise terms are expected to be typically dominant.

3. We presented a new algorithm within the AMUSE framework, SECULARMULTIPLE, to numerically solve the resulting equations of motion. The new algorithm allows for fast long-term integration of hierarchical systems, opening up the possibility for studying the long-term dynamics of complex hierarchical systems without having to resort to costly N -body integrations. Tidal friction and precession due to relativity and tidal bulges are also included in an approximate approach, treating each binary as an isolated system.

4. As a first demonstration, we showed that SECULARMULTIPLE can be used for the secular dynamics of multiplanet systems in both single-star and multiple-star systems. For the case of a single-star system (cf. Section 6.3.1), we compared our results to Laplace-Lagrange (LL) theory in the regime of low eccentricities and inclinations, in which LL theory is essentially exact (barring the effects of mean motion resonances). We found that SECULARMULTIPLE produces results that differ by only $\sim 2\%$ with LL theory, provided that the spacing between the planets is $K_0 \gtrsim 10$ (equivalently, a semimajor axis ratio $\lesssim 0.4$), and that sufficiently high-order pairwise binary terms, as high as the dotriacontupole-order (fifth order in x_i), are included. In the systems that we considered, the octupole-order binary triplet terms do not affect the secular dynamics. More generally, for any order n , we estimated that the non-pairwise binary terms are not important for nested systems with a central massive body, and with the other bodies of comparable mass.



5. Whereas LL theory only applies to planetary systems with low eccentricities and inclinations and with a central massive object, our method can also be used in the case of arbitrary eccentricities (provided that the x_i remain sufficiently small), inclinations and hierarchies. In particular, we showed that `SECULARMULTIPLE` can be used to efficiently study multiplanet systems in binary and triple-star systems (cf. Sections 6.3.2 and 6.3.3, respectively). In the case of triple-star systems in which the tertiary star has multiple S-type planets, we showed that, compared to the case of a binary-star system, the ‘binary nature’ of the companion in the case of a stellar triple can result in both dynamical *protection* and *destabilization* of the planets against the stellar torques, depending on the configuration.

6. We applied our method to the observed stellar quadruple system 30 Arietis (30 Ari), which harbours a massive planet, 30 Ari Bb. The mutual inclinations in this system are unknown. Using secular stability arguments, we showed that the orbit of the planet is likely not highly inclined with respect to 30 Ari C. Furthermore, assuming the latter, we found that there is a narrow region in semimajor axis space ($6 \lesssim a_{p2}/\text{AU} \lesssim 7.6$) which allows for the presence of an additional planet beyond the orbit of 30 Ari Bb, and inside of the orbit of 30 Ari C. The eccentricity of the orbit of the additional planet would likely be high, possibly as high as ~ 0.6 .

We will make `SECULARMULTIPLE` publicly available within `AMUSE`⁴ by 2016.

Acknowledgements

We thank the referee, Will Farr, for providing very useful comments, and for motivating us to further generalize our results. This work was supported by the Netherlands Research Council NWO (grants #639.073.803 [VICI], #614.061.608 [AMUSE] and #612.071.305 [LGM]) and the Netherlands Research School for Astronomy (NOVA).

⁴amusecode.org

6.A Derivation of the expanded and orbit-averaged hierarchical multiple Hamiltonian

In this appendix, we derive the Hamiltonian H for a hierarchical multiple system consisting of nested binaries with an arbitrary number of bodies and hierarchy, expanded in ratios of the binary separations (cf. Appendices 6.A.1, 6.A.2, 6.A.3 and 6.A.4). Subsequently, we orbit average the Hamiltonian (cf. Appendix 6.A.5). Lastly, we give an ad hoc expression for the IPN Hamiltonian (cf. Appendix 6.A.7).

6.A.1 Description of the system structure

Consider a system of N point masses $m_i \in \{m_1 \cdots m_N\}$ with position vectors $\mathbf{R}_i \in \{\mathbf{R}_1 \cdots \mathbf{R}_N\}$, arranged in $N - 1$ binary orbits (see Figure 6.1 for an example quintuple system). The relative position vectors between binaries and/or bodies are denoted by $\mathbf{r}_i$, where generally $i \leq N - 1$; let $\mathbf{B} = \{i : i \leq N - 1\}$ denote the corresponding set of indices. The hierarchy of the system is specified using the mass ratio matrix $\mathbf{A}$ according to

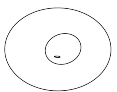
$$\mathbf{r}_i = \sum_{k=1}^N A_{ik} \mathbf{R}_k. \quad (6.27)$$

By definition, the last row in $\mathbf{A}$, i.e. A_{Nk} , corresponds to the centre of mass $\mathbf{r}_N$, and is given by $A_{Nk} = m_k/M$, where

$$M \equiv \sum_{i=1}^N m_i \quad (6.28)$$

is the total mass of the system. For convenience in a number of examples below, we also introduce the following notation for the partial sum over the masses,

$$M_{i,j} \equiv \sum_{n=i}^j m_n. \quad (6.29)$$



$i \rightarrow j$																									
k	1					2					3					4					5				
	2	3	4	5		3	4	5		4	5	4	5		5		4	5		5					
1	1	$\frac{m_2}{M_{1,2}}$	$\frac{m_2}{M_{1,2}}$	$\frac{m_2}{M_{1,2}}$	$-\frac{m_1}{M_{1,2}}$	$-\frac{m_1}{M_{1,2}}$	$-\frac{m_1}{M_{1,2}}$	$-\frac{m_1}{M_{1,2}}$	0	0	0	0	0	0	0	0	0	0	0	0	0				
2	0	1	$\frac{m_3}{M_{1,3}}$	$\frac{m_3}{M_{1,3}}$	1	$\frac{m_3}{M_{1,3}}$	$\frac{m_3}{M_{1,3}}$	$\frac{m_3}{M_{1,3}}$	$-\frac{M_{1,2}}{M_{1,3}}$	$-\frac{M_{1,2}}{M_{1,3}}$	$-\frac{M_{1,2}}{M_{1,3}}$	$-\frac{M_{1,2}}{M_{1,3}}$	$-\frac{M_{1,2}}{M_{1,3}}$	$-\frac{M_{1,2}}{M_{1,3}}$	$-\frac{M_{1,2}}{M_{1,3}}$	$-\frac{M_{1,2}}{M_{1,3}}$	$-\frac{M_{1,2}}{M_{1,3}}$	$-\frac{M_{1,2}}{M_{1,3}}$	$-\frac{M_{1,2}}{M_{1,3}}$	$-\frac{M_{1,2}}{M_{1,3}}$	$-\frac{M_{1,2}}{M_{1,3}}$				
3	0	0	$-\frac{m_5}{M_{4,5}}$	$\frac{m_4}{M_{4,5}}$	0	$-\frac{m_5}{M_{4,5}}$	$\frac{m_4}{M_{4,5}}$	$\frac{m_4}{M_{4,5}}$	$-\frac{m_5}{M_{4,5}}$	$-\frac{m_5}{M_{4,5}}$	$-\frac{m_5}{M_{4,5}}$	$-\frac{m_5}{M_{4,5}}$	$-\frac{m_5}{M_{4,5}}$	$-\frac{m_5}{M_{4,5}}$	$-\frac{m_5}{M_{4,5}}$	$-\frac{m_5}{M_{4,5}}$	$-\frac{m_5}{M_{4,5}}$	$-\frac{m_5}{M_{4,5}}$	$-\frac{m_5}{M_{4,5}}$	$-\frac{m_5}{M_{4,5}}$	$-\frac{m_5}{M_{4,5}}$				
4	0	0	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0				

Table 6.2: The values of B_{ijk} (cf. equation 6.34) for the quintuple system depicted in Figure 6.1. Columns correspond to pairs of bodies (i, j) ; rows correspond to binaries k . The partial mass sums $M_{i,j}$ are defined in equation (6.29). Note that $B_{ijk} = -B_{jik}$ and $B_{iik} = 0$ (cf. equation 6.34).

$k \rightarrow l$									
i	1			2			3		
	1	2	3	4	2	3	3	4	4
1	$\frac{m_2^2}{M_{1,2}^2}$	$\frac{m_2 m_3}{M_{1,2} M_{1,3}}$	0	$\frac{m_2 M_{4,5}}{M_{1,2} M}$	$\frac{m_3^2}{M_{1,3}^2}$	0	$\frac{m_3 M_{4,5}}{M_{1,3} M}$	0	$\frac{M_{4,5}^2}{M^2}$
2	$\frac{m_1^2}{M_{1,2}^2}$	$-\frac{m_1 m_3}{M_{1,2} M_{1,3}}$	0	$-\frac{m_1 M_{4,5}}{M_{1,2} M}$	$\frac{m_3^2}{M_{1,3}^2}$	0	$\frac{m_3 M_{4,5}}{M_{1,3} M}$	0	$\frac{M_{4,5}^2}{M^2}$
3	0	0	0	0	$\frac{M_{1,2}^2}{M_{1,3}^2}$	0	$-\frac{M_{1,2} M_{4,5}}{M_{1,3} M}$	0	$\frac{M_{4,5}^2}{M^2}$
4	0	0	0	0	0	0	0	$\frac{m_3^2}{M_{4,5}^2}$	$\frac{M_{1,3}^2}{M_{4,5}^2}$
5	0	0	0	0	0	0	0	$\frac{m_4^2}{M_{4,5}^2}$	$\frac{M_{1,3}^2}{M_{4,5}^2}$

Table 6.3: The values of C_{ikl} (cf. equation 6.35) for the quintuple system depicted in Figure 6.1. Columns correspond to pairs of binaries (k, l) ; rows correspond to bodies i . The partial mass sums $M_{i,j}$ are defined in equation (6.29). Note that $C_{ikl} = C_{ilk}$ (cf. equation 6.35).

For example, a hierarchical triple system (the simplest hierarchical multiple system) is represented by

$$\mathbf{A} = \begin{pmatrix} 1 & -1 & 0 \\ \frac{m_1}{M_{1,2}} & \frac{m_2}{M_{1,2}} & -1 \\ \frac{m_1}{M} & \frac{m_2}{M} & \frac{m_3}{M} \end{pmatrix}, \quad (6.30)$$

which implies $\mathbf{r}_1 = \mathbf{R}_1 - \mathbf{R}_2$ and $\mathbf{r}_2 = (m_1 \mathbf{R}_1 + m_2 \mathbf{R}_2) / (m_1 + m_2) - \mathbf{R}_3$. With this definition of $\mathbf{A}$, $\mathbf{r}_1$ corresponds to the relative

separation vector of the inner binary between $\mathbf{R}_1$ and $\mathbf{R}_2$, whereas $\mathbf{r}_2$ corresponds to the relative separation vector in the outer binary between the centre of mass of the inner binary and $\mathbf{R}_3$. Another example, which is also used for further illustration below, is a hierarchical quintuple consisting of a triple-binary pair,

$$\mathbf{A} = \begin{pmatrix} 1 & -1 & 0 & 0 & 0 \\ \frac{m_1}{M_{1,2}} & \frac{m_2}{M_{1,2}} & -1 & 0 & 0 \\ 0 & 0 & 0 & 1 & -1 \\ \frac{m_1}{M_{1,3}} & \frac{m_2}{M_{1,3}} & \frac{m_3}{M_{1,3}} & -\frac{m_4}{M_{4,5}} & -\frac{m_5}{M_{4,5}} \\ \frac{m_1}{M} & \frac{m_2}{M} & \frac{m_3}{M} & \frac{m_4}{M} & \frac{m_5}{M} \end{pmatrix}. \quad (6.31)$$

A schematic representation of this system in a mobile diagram is given in Figure 6.1. Note that the order of the rows in $\mathbf{A}$ sets the labels on the binary separation vectors $\mathbf{r}_i$, and the signs determine the directions of the $\mathbf{r}_i$.

The inverse relation between (absolute) position vectors and relative binary separation vectors is given by the inverse matrix of $\mathbf{A}$, i.e.

$$\mathbf{R}_i = \sum_{k=1}^N A_{ik}^{-1} \mathbf{r}_k. \quad (6.32)$$

The potential energy (cf. Appendix 6.A.2) is expressed in terms of distances between pairs of bodies, $\|\mathbf{R}_i - \mathbf{R}_j\|$. Using equation (6.32), the difference vectors can be written as

$$\mathbf{R}_i - \mathbf{R}_j = \sum_{k=1}^N \left(A_{ik}^{-1} - A_{jk}^{-1} \right) \mathbf{r}_k \equiv \sum_{k=1}^N B_{ijk} \mathbf{r}_k, \quad (6.33)$$

where we defined the quantity

$$B_{ijk} \equiv A_{ik}^{-1} - A_{jk}^{-1}. \quad (6.34)$$

Another quantity derived from $\mathbf{A}$, used for the kinetic energy (cf. Appendix 6.A.3), is

$$C_{ikl} \equiv A_{ik}^{-1} A_{il}^{-1}. \quad (6.35)$$

Before proceeding with the expansion of the potential energy in Appendix 6.A.2, we first introduce a number of useful definitions, and discuss some general properties of the three-index quantities B_{ijk} (Appendix 6.A.1) and C_{ikl} (Appendix 6.A.1).

Children, siblings and descendants. We refer to the two members of a binary k as the ‘children’ of that binary, i.e. child 1 and child 2, denoted by ‘ $k.C1$ ’ and ‘ $k.C2$ ’. The children can be either bodies or binaries themselves. Each child has a sibling, i.e. the binary or body with the same parent binary. For a binary pair (k, l) , where l is a direct or indirect child of k , we use ‘ $k.CS(l)$ ’ to denote the sibling in k of the child in k that is connected to l .

We define a descendant of a binary k as a body or binary that is connected to k via one of the children of k , not necessarily directly. For example, in Figure 6.1, bodies 1 and 2 are descendants of binary 4, whereas they are not descendants of binary 3. The set of all descendant *bodies* of a binary k is denoted with ‘ $\{k.C\}$ ’, and the body descendant sets within the children of k are denoted with ‘ $\{k.Cn\}$ ’ for child n . These notations will be used frequently below.

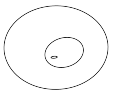
Masses. The mass of binary k , M_k , is defined as the sum of the masses of *all bodies* contained within the hierarchy of binary k . The masses of the children of binary k are denoted with $M_{k.C1}$ and $M_{k.C2}$, which, evidently, satisfy

$$M_{k.C1} + M_{k.C2} = M_k. \quad (6.36)$$

Furthermore,

$$\sum_{i \in \{k.Cn\}} m_i = M_{Cn}. \quad (6.37)$$

Levels. We define the ‘level’ of a binary k as the number of binaries that needs to be traversed to reach k , starting from the ‘absolute top’ binary of the system, i.e. the binary for which all other binaries are children. For example, in Figure 6.1, the top binary is binary 4 with level 0. We denote the level of binary k with ‘ $k.L$ ’.



The ‘sign quantity’ $\alpha(i, j; k)$. For any pair of bodies or binaries (i, j) , there is a unique path, i.e. a set of binaries, connecting i and j . Let this path be denoted by $B_p = B_p(i, j)$. For a given path B_p and a given binary k , we define the quantity $\alpha = \alpha(i, j; k)$ for which

$$\alpha(i, j; k) = \pm 1. \quad (6.38)$$

The positive (negative) sign applies when in binary k , the path B_p is opposed to (directed along) the direction of r_k . In other words, if r_k is defined such that it points from ‘child 2’ to ‘child 1’ in binary k , then the sign is positive if the path in B_p leads from child 1 to child 2 (directed against r_k), and negative if path in B_p leads from child 2 to child 1 (directed along r_k). From this, it also immediately follows that $\alpha(i, j; k) = -\alpha(j, i; k)$. Also, for $j \notin \{k, C1\}$ and $j \notin \{k, C2\}$, $\alpha(k, C1, j; k) = -\alpha(k, C2, j; k)$.

For example, in Figure 6.1, $\alpha(1, 2; 1) = 1$, $\alpha(2, 1; 1) = -1$, $\alpha(1, 5; 4) = 1$ and $\alpha(5, 1; 4) = -1$.

General properties of B_{ijk}

By definition (cf. equation 6.33), the first two indices in B_{ijk} refer to bodies, whereas the third index refers to a binary. Given pair (i, j) with path $B_p = B_p(i, j)$ and given binary k , the following properties of B_{ijk} apply.

- $B_{ijk} = -B_{jik}$ (this follows immediately from equation 6.34).
- $B_{ijk} = 0$ if $k \notin B_p(i, j)$ or $k = N$ (in the latter case, k corresponds to the centre of mass).
- $B_{ijk} = \alpha(i, j; m)$ if binary $k \in B_p$ and both children of k are part of B_p . In other words, k is the lowest level, or ‘top’ binary in B_p . Below, we refer to this binary with the ‘special’ index $m = m(i, j)$. There is a unique binary m for every path B_p between bodies i and j . For example, in Figure 6.1, for $B_p(i, j) = B_p(1, 4)$, the top binary is $m(1, 4) = 4$.
- Otherwise, $B_{ijk} = \alpha(i, j; k) M_{k,C(3-n)} / M_k$ if i and j are connected to each other through child n of k (here, n is either 1 or 2).

To illustrate these properties, we give in Table 6.2 the values of B_{ijk} for our example quintuple system depicted in Figure 6.1, computed directly from the definition equation (6.34).

General properties of C_{ikl}

In C_{ikl} , the first index refers to a body, whereas the second and third indices refer to binaries. From the definition in equation (6.35), it immediately follows that $C_{ikl} = C_{ilk}$. For a pair of binaries $\{k, l\}$, let p denote the binary with the highest level, and q the binary with the lowest level (i.e. binary p is ‘below’ q). The following properties apply.

- $C_{ikl} = 0$ if $k \neq l$ and binaries k and l have the same level.
- $C_{ikl} = 0$ if i is not a descendant of both binaries k and l .
- $C_{ikl} = M_{k,C(3-n)}^2 / M_k^2$ if $k = l$ and i is connected to k through child n of k .
- $C_{ikl} = \alpha(i, q; p) \alpha(i, q; q) M_{p,C(n-3)} M_{q,CS(p)} / (M_p M_q)$ if i is a descendant of p , and i is connected to p through child n of p . The product of two sign quantities implies that C_{ikl} is positive if the path from body i to binary q is either opposed to, or directed along *both* directions of $\mathbf{r}_p$ and $\mathbf{r}_q$. The sign of C_{ikl} is negative if the path in from body i to binary q is opposed to the direction of $\mathbf{r}_p$ and directed along $\mathbf{r}_q$, or if the path in from body i to binary q is directed along to the direction of $\mathbf{r}_p$ and opposed to $\mathbf{r}_q$.

The values of C_{ikl} for our example quintuple system depicted in Figure 6.1, computed from equation (6.35), are given for illustration in Table 6.3.

6.A.2 Expansion of the potential energy

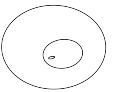
The Newtonian potential energy V is given by

$$V = -G \sum_{i < j} \frac{m_i m_j}{\|\mathbf{R}_i - \mathbf{R}_j\|}. \quad (6.39)$$

Our approach is to write V in terms of the $\mathbf{r}_i$, and to expand it in terms of the (assumed to be) small separation ratios x_i (defined more precisely below).

Using equation (6.33), the norm of the separation vector between two pairs of bodies can be written as

$$\begin{aligned} \|\mathbf{R}_i - \mathbf{R}_j\| &= [(\mathbf{R}_i - \mathbf{R}_j) \cdot (\mathbf{R}_i - \mathbf{R}_j)]^{1/2} = \left[\sum_{k=1}^N \sum_{l=1}^N B_{ijk} B_{ijl} (\mathbf{r}_k \cdot \mathbf{r}_l) \right]^{1/2} = \left[\sum_{k \in \mathbf{B}} \sum_{l \in \mathbf{B}} B_{ijk} B_{ijl} (\mathbf{r}_k \cdot \mathbf{r}_l) \right]^{1/2} \\ &= \left[\sum_{k \in \bar{\mathbf{B}}} \sum_{l \in \bar{\mathbf{B}}} B_{ijk} B_{ijl} (\mathbf{r}_k \cdot \mathbf{r}_l) + \sum_{l \in \bar{\mathbf{B}}} B_{ijm} B_{ijl} (\mathbf{r}_m \cdot \mathbf{r}_l) + \sum_{k \in \bar{\mathbf{B}}} B_{ijk} B_{ijm} (\mathbf{r}_k \cdot \mathbf{r}_m) + B_{ijm} B_{ijm} (\mathbf{r}_m \cdot \mathbf{r}_m) \right]^{1/2} \\ &= r_m \left[\sum_{k \in \bar{\mathbf{B}}} \sum_{l \in \bar{\mathbf{B}}} B_{ijk} B_{ijl} (\hat{\mathbf{r}}_k \cdot \hat{\mathbf{r}}_l) \left(\frac{r_l}{r_m} \right) \left(\frac{r_k}{r_m} \right) + 2 \alpha(i, j; m) \sum_{k \in \bar{\mathbf{B}}} B_{ijk} (\hat{\mathbf{r}}_k \cdot \hat{\mathbf{r}}_m) \left(\frac{r_k}{r_m} \right) + 1 \right]^{1/2}. \end{aligned} \quad (6.40)$$



Here, we defined $\bar{B} \equiv B \setminus \{m\}$ as the set of binaries excluding binary m , where $m = m(i, j)$ is the lowest level, or 'top' binary in $B_p(i, j)$, the path of binaries between bodies i and j . Evidently, $\hat{r}_m \cdot \hat{r}_m = 1$, and as mentioned in Appendix 6.A.1, $B_{ijm} = \alpha(i, j; m)$, hence $B_{ijm}^2 = (\pm 1)^2 = 1$.

By our assumption of hierarchy, r_m is the largest of the separations r_k in the path $B_p(i, j)$, i.e. if $k \in B_p(i, j)$, then $r_k \ll r_m$ if $k \neq m$. Furthermore, B_{ijk} and $\alpha(i, j; m)$ always satisfy $|B_{ijk}| \leq 1$ and $|\alpha(i, j; m)| = 1$. Therefore, for (sufficiently) hierarchical systems, it is appropriate to expand the potential in terms of $x_k \equiv r_k/r_m \ll 1$. The less hierarchical the system, the more terms of higher orders need to be included.

The expression for $\|\mathbf{R}_i - \mathbf{R}_j\|^{-1}$ is a function of a set of small (compared to zero) variables $\{x_n\}$, where $n \in \bar{B}$. The multivariate function that needs to be expanded is

$$f(\{x_n\}) = \frac{1}{r_m} \left[\sum_{k \in \bar{B}} \sum_{l \in \bar{B}} B_{ijk} B_{ijl} (\hat{r}_k \cdot \hat{r}_l) x_k x_l + 2 \alpha(i, j; m) \sum_{k \in \bar{B}} B_{ijk} (\hat{r}_k \cdot \hat{r}_m) x_k + 1 \right]^{-1/2}. \quad (6.41)$$

To $\mathcal{O}(x^5)$, the general Taylor expansion of $f(\{x_n\})$ for all x_n near 0 reads

$$\begin{aligned} f(\{x_n\}) = & f(\{x_n = 0\}) + \sum_{p \in \bar{B}} \frac{\partial f}{\partial x_p} \bigg|_{\{x_n\}=0} x_p + \frac{1}{2} \sum_{p \in \bar{B}} \sum_{q \in \bar{B}} \frac{\partial^2 f}{\partial x_p \partial x_q} \bigg|_{\{x_n\}=0} x_p x_q + \frac{1}{2 \cdot 3} \sum_{p \in \bar{B}} \sum_{q \in \bar{B}} \sum_{u \in \bar{B}} \frac{\partial^3 f}{\partial x_p \partial x_q \partial x_u} \bigg|_{\{x_n\}=0} x_p x_q x_u \\ & + \frac{1}{2 \cdot 3 \cdot 4} \sum_{p \in \bar{B}} \sum_{q \in \bar{B}} \sum_{u \in \bar{B}} \sum_{v \in \bar{B}} \frac{\partial^4 f}{\partial x_p \partial x_q \partial x_u \partial x_v} \bigg|_{\{x_n\}=0} x_p x_q x_u x_v \\ & + \frac{1}{2 \cdot 3 \cdot 4 \cdot 5} \sum_{p \in \bar{B}} \sum_{q \in \bar{B}} \sum_{u \in \bar{B}} \sum_{v \in \bar{B}} \sum_{w \in \bar{B}} \frac{\partial^5 f}{\partial x_p \partial x_q \partial x_u \partial x_v \partial x_w} \bigg|_{\{x_n\}=0} x_p x_q x_u x_v x_w + \mathcal{O}(x^6). \end{aligned} \quad (6.42)$$

Computing the required partial derivatives of f , we find

$$\frac{1}{\|\mathbf{R}_i - \mathbf{R}_j\|} = \frac{1}{r_m} \{S_0 + S_1 + S_2 + S_3 + S_4 + S_5 + \mathcal{O}(x^6)\}, \quad (6.43)$$

where S_n denote terms of order x^n . Up and including fifth order, they are given by

$$S_0 = 1; \quad (6.44)$$

$$S_1 = -\alpha(i, j; m) \sum_{p \in \bar{\mathbf{B}}} B_{ijp} \beta_1(\mathbf{r}_p, \mathbf{r}_m); \quad (6.45)$$

$$S_2 = \frac{1}{2} \sum_{p \in \bar{\mathbf{B}}} \sum_{q \in \bar{\mathbf{B}}} B_{ijp} B_{ijq} \beta_2(\mathbf{r}_p, \mathbf{r}_q, \mathbf{r}_m); \quad (6.46)$$

$$S_3 = -\frac{1}{2} \alpha(i, j; m) \sum_{p \in \bar{\mathbf{B}}} \sum_{q \in \bar{\mathbf{B}}} \sum_{u \in \bar{\mathbf{B}}} B_{ijp} B_{ijq} B_{iju} \beta_3(\mathbf{r}_p, \mathbf{r}_q, \mathbf{r}_u, \mathbf{r}_m) \quad (6.47)$$

$$S_4 = \frac{1}{8} \sum_{p \in \bar{\mathbf{B}}} \sum_{q \in \bar{\mathbf{B}}} \sum_{u \in \bar{\mathbf{B}}} \sum_{v \in \bar{\mathbf{B}}} B_{ijp} B_{ijq} B_{iju} B_{ijv} \beta_4(\mathbf{r}_p, \mathbf{r}_q, \mathbf{r}_u, \mathbf{r}_v, \mathbf{r}_m) \quad (6.48)$$

$$S_5 = -\frac{1}{8} \alpha(i, j; m) \sum_{p \in \bar{\mathbf{B}}} \sum_{q \in \bar{\mathbf{B}}} \sum_{u \in \bar{\mathbf{B}}} \sum_{v \in \bar{\mathbf{B}}} \sum_{w \in \bar{\mathbf{B}}} B_{ijp} B_{ijq} B_{iju} B_{ijv} B_{ijw} \beta_5(\mathbf{r}_p, \mathbf{r}_q, \mathbf{r}_u, \mathbf{r}_v, \mathbf{r}_w, \mathbf{r}_m). \quad (6.49)$$

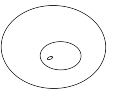
Here, the functions β_n are given by

$$\beta_1(\mathbf{r}_p, \mathbf{r}_m) \equiv \frac{r_p}{r_m} (\hat{\mathbf{r}}_p \cdot \hat{\mathbf{r}}_m); \quad (6.50)$$

$$\beta_2(\mathbf{r}_p, \mathbf{r}_q, \mathbf{r}_m) \equiv \frac{r_p}{r_m} \frac{r_q}{r_m} [3 (\hat{\mathbf{r}}_p \cdot \hat{\mathbf{r}}_m) (\hat{\mathbf{r}}_q \cdot \hat{\mathbf{r}}_m) - (\hat{\mathbf{r}}_p \cdot \hat{\mathbf{r}}_q)]; \quad (6.51)$$

$$\beta_3(\mathbf{r}_p, \mathbf{r}_q, \mathbf{r}_u, \mathbf{r}_m) \equiv \frac{r_p}{r_m} \frac{r_q}{r_m} \frac{r_u}{r_m} [5 (\hat{\mathbf{r}}_p \cdot \hat{\mathbf{r}}_m) (\hat{\mathbf{r}}_q \cdot \hat{\mathbf{r}}_m) (\hat{\mathbf{r}}_u \cdot \hat{\mathbf{r}}_m) - (\hat{\mathbf{r}}_p \cdot \hat{\mathbf{r}}_m) (\hat{\mathbf{r}}_q \cdot \hat{\mathbf{r}}_u) - (\hat{\mathbf{r}}_u \cdot \hat{\mathbf{r}}_m) (\hat{\mathbf{r}}_p \cdot \hat{\mathbf{r}}_q)]; \quad (6.52)$$

$$\begin{aligned} \beta_4(\mathbf{r}_p, \mathbf{r}_q, \mathbf{r}_u, \mathbf{r}_v, \mathbf{r}_m) &\equiv \frac{r_p}{r_m} \frac{r_q}{r_m} \frac{r_u}{r_m} \frac{r_v}{r_m} [35 (\hat{\mathbf{r}}_p \cdot \hat{\mathbf{r}}_m) (\hat{\mathbf{r}}_q \cdot \hat{\mathbf{r}}_m) (\hat{\mathbf{r}}_u \cdot \hat{\mathbf{r}}_m) (\hat{\mathbf{r}}_v \cdot \hat{\mathbf{r}}_m) \\ &\quad - 5 (\hat{\mathbf{r}}_p \cdot \hat{\mathbf{r}}_m) (\hat{\mathbf{r}}_q \cdot \hat{\mathbf{r}}_m) (\hat{\mathbf{r}}_u \cdot \hat{\mathbf{r}}_v) - 5 (\hat{\mathbf{r}}_p \cdot \hat{\mathbf{r}}_m) (\hat{\mathbf{r}}_u \cdot \hat{\mathbf{r}}_m) (\hat{\mathbf{r}}_q \cdot \hat{\mathbf{r}}_v) - 5 (\hat{\mathbf{r}}_p \cdot \hat{\mathbf{r}}_m) (\hat{\mathbf{r}}_v \cdot \hat{\mathbf{r}}_m) (\hat{\mathbf{r}}_q \cdot \hat{\mathbf{r}}_u) \\ &\quad - 5 (\hat{\mathbf{r}}_p \cdot \hat{\mathbf{r}}_m) (\hat{\mathbf{r}}_q \cdot \hat{\mathbf{r}}_m) (\hat{\mathbf{r}}_v \cdot \hat{\mathbf{r}}_u) - 5 (\hat{\mathbf{r}}_q \cdot \hat{\mathbf{r}}_m) (\hat{\mathbf{r}}_u \cdot \hat{\mathbf{r}}_m) (\hat{\mathbf{r}}_p \cdot \hat{\mathbf{r}}_v) - 5 (\hat{\mathbf{r}}_q \cdot \hat{\mathbf{r}}_m) (\hat{\mathbf{r}}_v \cdot \hat{\mathbf{r}}_m) (\hat{\mathbf{r}}_p \cdot \hat{\mathbf{r}}_u) \\ &\quad + (\hat{\mathbf{r}}_p \cdot \hat{\mathbf{r}}_v) (\hat{\mathbf{r}}_q \cdot \hat{\mathbf{r}}_u) + (\hat{\mathbf{r}}_p \cdot \hat{\mathbf{r}}_u) (\hat{\mathbf{r}}_q \cdot \hat{\mathbf{r}}_v) + (\hat{\mathbf{r}}_p \cdot \hat{\mathbf{r}}_q) (\hat{\mathbf{r}}_u \cdot \hat{\mathbf{r}}_v)]; \end{aligned} \quad (6.53)$$



$$\begin{aligned}
\beta_5(\mathbf{r}_p, \mathbf{r}_q, \mathbf{r}_u, \mathbf{r}_v, \mathbf{r}_w, \mathbf{r}_m) \equiv & \frac{r_p}{r_m} \frac{r_q}{r_m} \frac{r_u}{r_m} \frac{r_v}{r_m} \frac{r_w}{r_m} [63(\hat{\mathbf{r}}_p \cdot \hat{\mathbf{r}}_m)(\hat{\mathbf{r}}_q \cdot \hat{\mathbf{r}}_m)(\hat{\mathbf{r}}_u \cdot \hat{\mathbf{r}}_m)(\hat{\mathbf{r}}_v \cdot \hat{\mathbf{r}}_m)(\hat{\mathbf{r}}_w \cdot \hat{\mathbf{r}}_m) \\
& - 7(\hat{\mathbf{r}}_u \cdot \hat{\mathbf{r}}_m)(\hat{\mathbf{r}}_v \cdot \hat{\mathbf{r}}_m)(\hat{\mathbf{r}}_w \cdot \hat{\mathbf{r}}_m)(\hat{\mathbf{r}}_p \cdot \hat{\mathbf{r}}_q) - 7(\hat{\mathbf{r}}_q \cdot \hat{\mathbf{r}}_m)(\hat{\mathbf{r}}_v \cdot \hat{\mathbf{r}}_m)(\hat{\mathbf{r}}_w \cdot \hat{\mathbf{r}}_m)(\hat{\mathbf{r}}_p \cdot \hat{\mathbf{r}}_u) - 7(\hat{\mathbf{r}}_q \cdot \hat{\mathbf{r}}_m)(\hat{\mathbf{r}}_u \cdot \hat{\mathbf{r}}_m)(\hat{\mathbf{r}}_w \cdot \hat{\mathbf{r}}_m)(\hat{\mathbf{r}}_p \cdot \hat{\mathbf{r}}_v) \\
& - 7(\hat{\mathbf{r}}_q \cdot \hat{\mathbf{r}}_m)(\hat{\mathbf{r}}_u \cdot \hat{\mathbf{r}}_m)(\hat{\mathbf{r}}_v \cdot \hat{\mathbf{r}}_m)(\hat{\mathbf{r}}_p \cdot \hat{\mathbf{r}}_w) - 7(\hat{\mathbf{r}}_p \cdot \hat{\mathbf{r}}_m)(\hat{\mathbf{r}}_v \cdot \hat{\mathbf{r}}_m)(\hat{\mathbf{r}}_w \cdot \hat{\mathbf{r}}_m)(\hat{\mathbf{r}}_q \cdot \hat{\mathbf{r}}_u) - 7(\hat{\mathbf{r}}_p \cdot \hat{\mathbf{r}}_m)(\hat{\mathbf{r}}_u \cdot \hat{\mathbf{r}}_m)(\hat{\mathbf{r}}_w \cdot \hat{\mathbf{r}}_m)(\hat{\mathbf{r}}_q \cdot \hat{\mathbf{r}}_v) \\
& - 7(\hat{\mathbf{r}}_p \cdot \hat{\mathbf{r}}_m)(\hat{\mathbf{r}}_u \cdot \hat{\mathbf{r}}_m)(\hat{\mathbf{r}}_v \cdot \hat{\mathbf{r}}_m)(\hat{\mathbf{r}}_q \cdot \hat{\mathbf{r}}_w) - 7(\hat{\mathbf{r}}_p \cdot \hat{\mathbf{r}}_m)(\hat{\mathbf{r}}_q \cdot \hat{\mathbf{r}}_m)(\hat{\mathbf{r}}_w \cdot \hat{\mathbf{r}}_m)(\hat{\mathbf{r}}_u \cdot \hat{\mathbf{r}}_v) - 7(\hat{\mathbf{r}}_p \cdot \hat{\mathbf{r}}_m)(\hat{\mathbf{r}}_v \cdot \hat{\mathbf{r}}_m)(\hat{\mathbf{r}}_w \cdot \hat{\mathbf{r}}_m)(\hat{\mathbf{r}}_q \cdot \hat{\mathbf{r}}_u) \\
& - 7(\hat{\mathbf{r}}_p \cdot \hat{\mathbf{r}}_m)(\hat{\mathbf{r}}_q \cdot \hat{\mathbf{r}}_m)(\hat{\mathbf{r}}_u \cdot \hat{\mathbf{r}}_m)(\hat{\mathbf{r}}_v \cdot \hat{\mathbf{r}}_w) \\
& + (\hat{\mathbf{r}}_w \cdot \hat{\mathbf{r}}_m)(\hat{\mathbf{r}}_p \cdot \hat{\mathbf{r}}_q)(\hat{\mathbf{r}}_u \cdot \hat{\mathbf{r}}_v) + (\hat{\mathbf{r}}_w \cdot \hat{\mathbf{r}}_m)(\hat{\mathbf{r}}_q \cdot \hat{\mathbf{r}}_u)(\hat{\mathbf{r}}_p \cdot \hat{\mathbf{r}}_v) + (\hat{\mathbf{r}}_w \cdot \hat{\mathbf{r}}_m)(\hat{\mathbf{r}}_p \cdot \hat{\mathbf{r}}_u)(\hat{\mathbf{r}}_q \cdot \hat{\mathbf{r}}_v) \\
& + (\hat{\mathbf{r}}_v \cdot \hat{\mathbf{r}}_m)(\hat{\mathbf{r}}_p \cdot \hat{\mathbf{r}}_q)(\hat{\mathbf{r}}_u \cdot \hat{\mathbf{r}}_w) + (\hat{\mathbf{r}}_v \cdot \hat{\mathbf{r}}_m)(\hat{\mathbf{r}}_p \cdot \hat{\mathbf{r}}_u)(\hat{\mathbf{r}}_q \cdot \hat{\mathbf{r}}_w) + (\hat{\mathbf{r}}_v \cdot \hat{\mathbf{r}}_m)(\hat{\mathbf{r}}_q \cdot \hat{\mathbf{r}}_u)(\hat{\mathbf{r}}_p \cdot \hat{\mathbf{r}}_w) \\
& + (\hat{\mathbf{r}}_u \cdot \hat{\mathbf{r}}_m)(\hat{\mathbf{r}}_p \cdot \hat{\mathbf{r}}_q)(\hat{\mathbf{r}}_v \cdot \hat{\mathbf{r}}_w) + (\hat{\mathbf{r}}_u \cdot \hat{\mathbf{r}}_m)(\hat{\mathbf{r}}_p \cdot \hat{\mathbf{r}}_v)(\hat{\mathbf{r}}_q \cdot \hat{\mathbf{r}}_w) + (\hat{\mathbf{r}}_u \cdot \hat{\mathbf{r}}_m)(\hat{\mathbf{r}}_q \cdot \hat{\mathbf{r}}_v)(\hat{\mathbf{r}}_p \cdot \hat{\mathbf{r}}_w) \\
& + (\hat{\mathbf{r}}_q \cdot \hat{\mathbf{r}}_m)(\hat{\mathbf{r}}_p \cdot \hat{\mathbf{r}}_u)(\hat{\mathbf{r}}_v \cdot \hat{\mathbf{r}}_w) + (\hat{\mathbf{r}}_q \cdot \hat{\mathbf{r}}_m)(\hat{\mathbf{r}}_p \cdot \hat{\mathbf{r}}_v)(\hat{\mathbf{r}}_u \cdot \hat{\mathbf{r}}_w) + (\hat{\mathbf{r}}_q \cdot \hat{\mathbf{r}}_m)(\hat{\mathbf{r}}_u \cdot \hat{\mathbf{r}}_v)(\hat{\mathbf{r}}_p \cdot \hat{\mathbf{r}}_w) \\
& + (\hat{\mathbf{r}}_p \cdot \hat{\mathbf{r}}_m)(\hat{\mathbf{r}}_q \cdot \hat{\mathbf{r}}_u)(\hat{\mathbf{r}}_v \cdot \hat{\mathbf{r}}_w) + (\hat{\mathbf{r}}_p \cdot \hat{\mathbf{r}}_m)(\hat{\mathbf{r}}_q \cdot \hat{\mathbf{r}}_v)(\hat{\mathbf{r}}_u \cdot \hat{\mathbf{r}}_w) + (\hat{\mathbf{r}}_p \cdot \hat{\mathbf{r}}_m)(\hat{\mathbf{r}}_u \cdot \hat{\mathbf{r}}_v)(\hat{\mathbf{r}}_q \cdot \hat{\mathbf{r}}_w)].
\end{aligned} \tag{6.54}$$

Generally, we formally write for order n ,

$$S_n = (-1)^n \alpha(i, j; m) \sum_{p_1, \dots, p_n \in \mathbb{B}} \underbrace{B_{ijp_1 \dots p_n}}_{n \times n \times} c_n \beta_n(\underbrace{\mathbf{r}_{p_1}, \dots, \mathbf{r}_{p_n}}_{n \times}, \mathbf{r}_m), \tag{6.55}$$

where c_n are constants. For pairwise terms, i.e. $p_1 = \dots = p_n = p$, we formally write $c_n \beta_n$ as

$$c_n \beta_n(\underbrace{\mathbf{r}_{p_1}, \dots, \mathbf{r}_{p_n}}_{n \times}, \mathbf{r}_m) = c_n \beta_n(\underbrace{\mathbf{r}_p, \dots, \mathbf{r}_p}_{n \times}, \mathbf{r}_m) = \sum_{j=0}^n \frac{(\mathbf{r}_p \cdot \mathbf{r}_m)^j r_p^{n-j}}{r_m^{n+j}} \mathcal{A}_j^{(n)}, \tag{6.56}$$

where $\mathcal{A}_j^{(n)}$ are integer ratio coefficients. They are the same as the coefficients appearing in the Legendre polynomials. The latter can be obtained by e.g. Rodrigues's formula, i.e.

$$\sum_{j=0}^n \mathcal{A}_j^{(n)} x^j = \frac{1}{2^n n!} \frac{d^n}{dx^n} [(x^2 - 1)^n]. \tag{6.57}$$

Note that the terms S_i in which $\hat{\mathbf{r}}_m$ appears an odd number of times (i.e. if n is odd), contain the factor $\alpha(i, j; m)$. This is to be expected: when changing the definition of $\hat{\mathbf{r}}_m$ such that it now points in the opposite direction, i.e. $\hat{\mathbf{r}}'_m = -\hat{\mathbf{r}}_m$, α also changes sign:

$\alpha(i, j; m)' = -\alpha(i, j; m)$. If $\hat{\mathbf{r}}_m$ appears an odd number of times, the resulting even number of minus signs cancel such that the expression of the expansion in terms of $\hat{\mathbf{r}}_m'$ is unaltered. The same property applies to the other separation vectors ($\hat{\mathbf{r}}_p, \hat{\mathbf{r}}_q$, etc.), which always appear only one time in each individual term (e.g., the term $(\hat{\mathbf{r}}_p \cdot \hat{\mathbf{r}}_m)^2$ never appears). When reversing the direction of these vectors, the resulting minus sign is absorbed by the mass ratio quantities (B_{ijp}, B_{ijq} , etc.), which each contain a factor α that also changes sign ($\alpha(i, j, p)$, $\alpha(i, j, q)$, etc.). Physically, this means that the expansion of $\|\mathbf{R}_i - \mathbf{R}_j\|^{-1}$ does not depend on the choice of direction of the relative vectors, as it should be.

6.A.3 Derivation of the kinetic energy term

The Newtonian kinetic energy of the system is given by

$$T = \frac{1}{2} \sum_{i=1}^N m_i (\dot{\mathbf{R}}_i \cdot \dot{\mathbf{R}}_i), \quad (6.58)$$

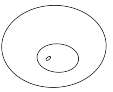
where the dots denote derivatives with respect to time. Using equation (6.32) and assuming constant masses, this can be written in terms of time derivatives of the binary separation vectors as

$$\begin{aligned} T &= \frac{1}{2} \sum_{i=1}^N \sum_{k=1}^N \sum_{l=1}^N m_i A_{ik}^{-1} A_{il}^{-1} (\dot{\mathbf{r}}_k \cdot \dot{\mathbf{r}}_l) = \frac{1}{2} \sum_{i=1}^N \sum_{k=1}^{N-1} \sum_{l=1}^{N-1} m_i A_{ik}^{-1} A_{il}^{-1} (\dot{\mathbf{r}}_k \cdot \dot{\mathbf{r}}_l) = \frac{1}{2} \sum_{i=1}^N \sum_{k \in B} \sum_{l \in B} m_i A_{ik}^{-1} A_{il}^{-1} (\dot{\mathbf{r}}_k \cdot \dot{\mathbf{r}}_l) \\ &\equiv \frac{1}{2} \sum_{i=1}^N \sum_{k \in B} \sum_{l \in B} m_i C_{ikl} (\dot{\mathbf{r}}_k \cdot \dot{\mathbf{r}}_l), \end{aligned} \quad (6.59)$$

where we used that the centre of mass, $\mathbf{r}_N$, satisfies $\dot{\mathbf{r}}_N = \mathbf{0}$, and where we defined $C_{ikl} \equiv A_{ik}^{-1} A_{il}^{-1}$ (cf. equation 6.35). Equation (6.59) can be simplified, as shown below.

The order of the summation over i and k and l can be reversed because B and i are independent (in contrast to $\bar{B}$ and i), i.e.

$$T = \frac{1}{2} \sum_{k \in B} \sum_{l \in B} (\dot{\mathbf{r}}_k \cdot \dot{\mathbf{r}}_l) \sum_{i=1}^N m_i C_{ikl}. \quad (6.60)$$



We split the double summations over binaries into several double summations over binaries, making a distinction between the relative binary levels, and into a single summation corresponding to $k = l$ in equation (6.60),

$$T = \frac{1}{2} \sum_{\substack{p,q \in \mathbb{B} \\ p \neq q \\ p.L > q.L}} (\dot{\mathbf{r}}_p \cdot \dot{\mathbf{r}}_q) \sum_{i=1}^N m_i C_{ipq} + \frac{1}{2} \sum_{\substack{p,q \in \mathbb{B} \\ q.L > p.L \\ p \neq q}} (\dot{\mathbf{r}}_p \cdot \dot{\mathbf{r}}_q) \sum_{i=1}^N m_i C_{ipq} + \frac{1}{2} \sum_{\substack{p,q \in \mathbb{B} \\ p.L = q.L \\ p \neq q}} (\dot{\mathbf{r}}_p \cdot \dot{\mathbf{r}}_q) \sum_{i=1}^N m_i C_{ipq} + \frac{1}{2} \sum_{p \in \mathbb{B}} (\dot{\mathbf{r}}_p \cdot \dot{\mathbf{r}}_p) \sum_{i=1}^N m_i C_{ipp}. \quad (6.61)$$

When $p \neq q$ and the binary levels are the same ($p.L = q.L$), then $C_{ipq} = 0$ (cf. Appendix 6.A.1). Also, $C_{ipq} = 0$ if i is not a descendant of the highest level binary in $\{p, q\}$. Therefore, the summation over bodies i from $i = 1$ to N can be rewritten in terms of a summation over bodies only in the highest level binary, i.e.

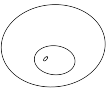
$$T = \frac{1}{2} \sum_{\substack{p,q \in \mathbb{B} \\ p \neq q \\ p.L > q.L}} (\dot{\mathbf{r}}_p \cdot \dot{\mathbf{r}}_q) \left(\sum_{i \in \{p.C1\}} + \sum_{i \in \{p.C2\}} \right) m_i C_{ipq} + \frac{1}{2} \sum_{\substack{p,q \in \mathbb{B} \\ p \neq q \\ q.L > p.L}} (\dot{\mathbf{r}}_p \cdot \dot{\mathbf{r}}_q) \left(\sum_{i \in \{q.C1\}} + \sum_{i \in \{q.C2\}} \right) m_i C_{ipq} \\ + \frac{1}{2} \sum_{p \in \mathbb{B}} (\dot{\mathbf{r}}_p \cdot \dot{\mathbf{r}}_p) \left(\sum_{i \in \{p.C1\}} + \sum_{i \in \{p.C2\}} \right) m_i C_{ipp}. \quad (6.62)$$

We recall that $\{p.Cn\}$ denotes the set of all descendant bodies of child n of binary p . Substituting the explicit expressions for C_{ipq} and C_{ipp}

(cf. Appendix 6.A.1), we find

$$\begin{aligned}
T &= \frac{1}{2} \sum_{\substack{p,q \in B \\ p \neq q \\ p.L > q.L}} (\dot{\mathbf{r}}_p \cdot \dot{\mathbf{r}}_q) \left(\sum_{i \in \{p.C1\}} m_i \frac{\alpha(p.C1, q; p) \alpha(p.C1, q; q) M_{p.C2} M_{q.CS(p)}}{M_p M_q} + \sum_{i \in \{p.C2\}} m_i \frac{\alpha(p.C2, q; p) \alpha(p.C2, q; q) M_{p.C1} M_{q.CS(p)}}{M_p M_q} \right) \\
&+ \frac{1}{2} \sum_{\substack{p,q \in B \\ p \neq q \\ q.L > p.L}} (\dot{\mathbf{r}}_p \cdot \dot{\mathbf{r}}_q) \left(\sum_{i \in \{q.C1\}} m_i \frac{\alpha(q.C1, p; q) \alpha(q.C1, p; p) M_{q.C2} M_{p.CS(q)}}{M_p M_q} + \sum_{i \in \{q.C2\}} m_i \frac{\alpha(q.C2, p; q) \alpha(q.C2, p; p) M_{q.C1} M_{p.CS(p)}}{M_p M_q} \right) \\
&+ \frac{1}{2} \sum_{p \in B} (\dot{\mathbf{r}}_p \cdot \dot{\mathbf{r}}_p) \left(\sum_{i \in \{p.C1\}} m_i \frac{M_{p.C2}^2}{M_p^2} + \sum_{i \in \{p.C2\}} m_i \frac{M_{p.C1}^2}{M_p^2} \right) \\
&= \frac{1}{2} \sum_{\substack{p,q \in B \\ p \neq q \\ p.L > q.L}} (\dot{\mathbf{r}}_p \cdot \dot{\mathbf{r}}_q) \left(M_{p.C1} \frac{\alpha(p.C1, q; p) \alpha(p.C1, q; q) M_{p.C2} M_{q.CS(p)}}{M_p M_q} + M_{p.C2} \frac{-\alpha(p.C1, q; p) \alpha(p.C1, q; q) M_{p.C1} M_{q.CS(p)}}{M_p M_q} \right) \\
&+ \frac{1}{2} \sum_{\substack{p,q \in B \\ p \neq q \\ q.L > p.L}} (\dot{\mathbf{r}}_p \cdot \dot{\mathbf{r}}_q) \left(M_{q.C1} \frac{\alpha(q.C1, p; q) \alpha(q.C1, p; p) M_{q.C2} M_{p.CS(q)}}{M_p M_q} + M_{q.C2} \frac{-\alpha(q.C1, p; q) \alpha(q.C1, p; p) M_{q.C1} M_{p.CS(p)}}{M_p M_q} \right) \\
&+ \frac{1}{2} \sum_{p \in B} (\dot{\mathbf{r}}_p \cdot \dot{\mathbf{r}}_p) \left(\frac{M_{p.C1} M_{p.C2}^2}{M_p^2} + \frac{M_{p.C2} M_{p.C1}^2}{M_p^2} \right) \\
&= \frac{1}{2} \sum_{p \in B} \frac{M_{p.C1} M_{p.C2}}{M_p} (\dot{\mathbf{r}}_p \cdot \dot{\mathbf{r}}_p). \tag{6.63}
\end{aligned}$$

Here, we used that $\alpha(p.C1, q; q) = \alpha(p.C2, q; q)$ and $\alpha(p.C1, q; p) = -\alpha(p.C2, q; p)$ in the summation with $p.L > q.L$. Similarly, in the summation with $p.L < q.L$, $\alpha(q.C1, p; p) = \alpha(q.C2, p; p)$ and $\alpha(q.C1, p; q) = -\alpha(q.C2, p; q)$. The expression in equation (6.63) shows



that the total kinetic energy is just the sum of the kinetic energies of all the individual binaries; all terms depending on velocities of pairs of binaries cancel.

6.A.4 Rewriting summations in the Hamiltonian

The Hamiltonian H is given by the sum of the kinetic and potential energies, $H = T + V$. Equations (6.39), (6.43) and (6.63) give the following expression for the Hamiltonian accurate to fifth order in ratios of the binary separations,

$$\begin{aligned} H &= \frac{1}{2} \sum_{p \in \mathbf{B}} \frac{M_{p,C1} M_{p,C2}}{M_p} (\dot{\mathbf{r}}_p \cdot \dot{\mathbf{r}}_p) - \sum_{i < j} \frac{G m_i m_j}{r_{m(i,j)}} \{ S_0 + S_1 + S_2 + S_3 + S_4 + S_5 + \mathcal{O}(x^6) \} \\ &\equiv \frac{1}{2} \sum_{p \in \mathbf{B}} \frac{M_{p,C1} M_{p,C2}}{M_p} (\dot{\mathbf{r}}_p \cdot \dot{\mathbf{r}}_p) + S'_0 + S'_1 + S'_2 + S'_3 + S'_4 + S'_5 + \mathcal{O}(x^6), \end{aligned} \quad (6.64)$$

where the primed quantities S'_n are defined as

$$S'_n \equiv - \sum_{i < j} \frac{G m_i m_j}{r_m} S_n. \quad (6.65)$$

The terms S'_n in equation (6.64) can be simplified substantially. A key ingredient in the simplifications is to sum over binaries and their children, rather than explicitly over all pairs of bodies in the first summation of equation (6.64), and to apply the properties of the mass ratio quantities B_{ijk} . Below we carry out these simplifications order by order.

Monopole-order term

The ‘monopole-order’ term S'_0 can be simplified by summing over all binaries and, subsequently, the children within those binaries, rather than summing explicitly over all pairs of bodies, i.e.

$$\sum_{i < j} \frac{G m_i m_j}{r_m} = \sum_{p \in \mathbf{B}} \sum_{i \in \{p,C1\}} \sum_{j \in \{p,C2\}} \frac{G m_i m_j}{r_m} = \sum_{p \in \mathbf{B}} \sum_{i \in \{p,C1\}} \sum_{j \in \{p,C2\}} \frac{G m_i m_j}{r_p}. \quad (6.66)$$

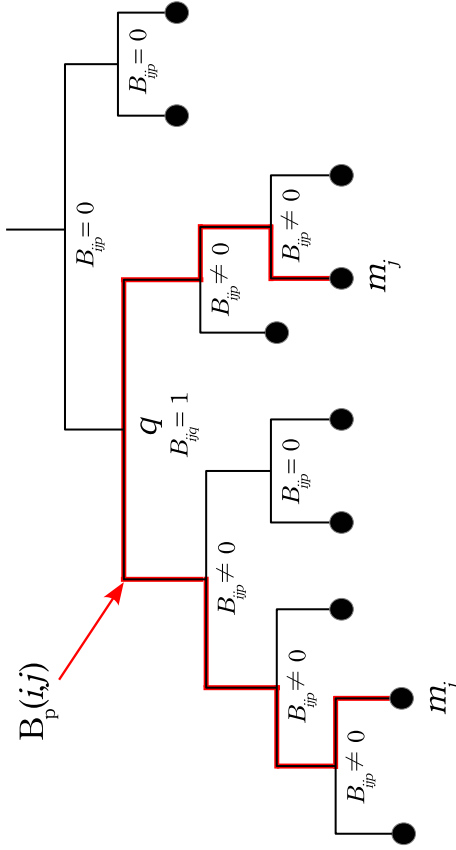


Figure 6.37: A mobile diagram with a particular path $B_p(i, j)$, and a ‘top’ level binary q (cf. Appendix 6.A.4). For each binary, we indicate whether or not B_{ijp} is zero.

The last step in equation (6.66) follows from the constancy of the separation r_m for fixed p and varying i and j , and which is given by r_p . Therefore,

$$-S'_0 \equiv \sum_{i < j} \frac{G m_i m_j}{r_m} = \sum_{p \in B} \frac{G}{r_p} \sum_{i \in \{p, C1\}} m_i \sum_{j \in \{p, C2\}} m_j = \sum_{p \in B} \frac{G}{r_p} \sum_{i \in \{p, C1\}} m_i M_{p, C2} = \sum_{p \in B} \frac{G M_{p, C1} M_{p, C2}}{r_p}. \quad (6.67)$$

Dipole-order term

Here, we show that the ‘dipole-order’ term S'_1 (cf. equation 6.45) vanishes identically for any system. Using equation (6.66) to rewrite the summations of pairs of bodies, S'_1 reads

$$\begin{aligned} S'_1 &\equiv \sum_{i < j} \frac{G m_i m_j}{r_m} \alpha(i, j; m) \sum_{p \in \bar{B}} B_{ijp} \beta_1(\mathbf{r}_p, \mathbf{r}_m) = \sum_{q \in B} \sum_{i \in \{q, C1\}} \sum_{j \in \{q, C2\}} \sum_{p \in \bar{B}} \frac{G m_i m_j}{r_q} \alpha(i, j; q) \sum_{p \in \bar{B}} B_{ijp} \beta_1(\mathbf{r}_p, \mathbf{r}_q) \\ &= \sum_{q \in B} \sum_{i \in \{q, C1\}} \sum_{j \in \{q, C2\}} \sum_{p \in \bar{B}} \frac{G m_i m_j}{r_q} \alpha(i, j; q) \beta_1(\mathbf{r}_p, \mathbf{r}_q), \end{aligned} \quad (6.68)$$

where in the last step we used that m_i , m_j and r_q are independent of p . We recall that $\bar{B} \equiv B \setminus \{q\}$ depends on i and j . Therefore, p runs over all binaries except q . Whenever p is not part of the descendants of q , it is also not within the path $B_p(i, j)$ because i and j run over the children of q . Therefore, $B_{ijp} = 0$ in that case (cf. Appendix 6.A.1, see also Figure 6.37). From this it follows that the summation of p over $\bar{B}$ can be rewritten as a summation of p over all binaries B , where p is a descendant of q , i.e.

$$S'_1 = \sum_{q \in B} \sum_{i \in \{q, C1\}} \sum_{j \in \{q, C2\}} \sum_{\substack{p \in B \\ p \in \{q, C\}}} B_{ijp} \frac{G m_i m_j}{r_q} \alpha(i, j; q) \beta_1(\mathbf{r}_p, \mathbf{r}_q). \quad (6.69)$$

The summation over p is now independent of i and j , and can therefore be placed before the summations over i and j . Also placing other quantities independent of i and j before the summations over i and j , we find

$$\begin{aligned} S'_1 &= \sum_{q \in B} \sum_{\substack{p \in B \\ p \in \{q, C\}}} \frac{G}{r_q} \beta_1(\mathbf{r}_p, \mathbf{r}_q) \sum_{i \in \{q, C1\}} \sum_{j \in \{q, C2\}} \alpha(i, j; q) B_{ijp} m_i m_j \\ &= \sum_{q \in B} \left(\sum_{\substack{p \in B \\ p \in \{q, C1\}}} + \sum_{\substack{p \in B \\ p \in \{q, C2\}}} \right) \frac{G}{r_q} \beta_1(\mathbf{r}_p, \mathbf{r}_q) \sum_{i \in \{q, C1\}} \sum_{j \in \{q, C2\}} \alpha(i, j; q) B_{ijp} m_i m_j. \end{aligned}$$

Here, we explicitly wrote the summations of p over the children of q into summations over child 1 and 2 of q . If p sums over $\{q, C1\}$, then $B_{ijp} = 0$ unless $i \in \{q, C1\}$ is a descendant of p (cf. Appendix 6.A.1). Conversely, if p sums over $\{q, C2\}$, then $B_{ijp} = 0$ unless $j \in \{q, C2\}$

is a descendant of p . Therefore, either summations of i and j over the children of q can be rewritten as summations over the descendants of binary p , i.e.

$$\begin{aligned}
 S'_1 &= \sum_{q \in B} \sum_{\substack{p \in B \\ p \in \{q, C1\}}} \frac{G}{r_q} \beta_1(\mathbf{r}_p, \mathbf{r}_q) \sum_{i \in \{p, C\}} \sum_{j \in \{q, C2\}} \alpha(i, j; q) B_{ijp} m_i m_j + \sum_{q \in B} \sum_{\substack{p \in B \\ p \in \{q, C2\}}} \frac{G}{r_q} \beta_1(\mathbf{r}_p, \mathbf{r}_q) \sum_{i \in \{q, C1\}} \sum_{j \in \{p, C\}} \alpha(i, j; q) B_{ijp} m_i m_j \\
 &= \sum_{q \in B} \sum_{\substack{p \in B \\ p \in \{q, C1\}}} \frac{G}{r_q} \beta_1(\mathbf{r}_p, \mathbf{r}_q) \sum_{j \in \{q, C2\}} m_j \sum_{i \in \{p, C\}} \alpha(i, j; q) B_{ijp} m_i + \sum_{q \in B} \sum_{\substack{p \in B \\ p \in \{q, C2\}}} \frac{G}{r_q} \beta_1(\mathbf{r}_p, \mathbf{r}_q) \sum_{i \in \{q, C1\}} m_i \sum_{j \in \{p, C\}} \alpha(i, j; q) B_{ijp} m_j \\
 &= \sum_{q \in B} \sum_{\substack{p \in B \\ p \in \{q, C1\}}} \frac{G}{r_q} \beta_1(\mathbf{r}_p, \mathbf{r}_q) \sum_{j \in \{q, C2\}} m_j \alpha(p, j; q) \sum_{i \in \{p, C\}} B_{ijp} m_i + \sum_{q \in B} \sum_{\substack{p \in B \\ p \in \{q, C2\}}} \frac{G}{r_q} \beta_1(\mathbf{r}_p, \mathbf{r}_q) \sum_{i \in \{q, C1\}} m_i \alpha(i, p; q) \sum_{j \in \{p, C\}} B_{ijp} m_j.
 \end{aligned}$$

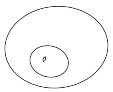
In the last line, we used that in the terms where the outermost summation runs over i , i runs over the descendants of p , therefore $\alpha(i, j; q)$ is equivalent to $\alpha(p, j; q)$; the same applies when the outermost summation runs over j , in which case $\alpha(i, j; q)$ is equivalent to $\alpha(i, p; q)$. Focusing on these outermost summations over children of p , we split the outermost summation over the children of p , using the properties of B_{ijp} (cf. Appendix 6.A.1), we find

$$\begin{aligned}
 \sum_{i \in \{p, C\}} B_{ijp} m_i &= \sum_{i \in \{p, C1\}} B_{ijp} m_i + \sum_{i \in \{p, C2\}} B_{ijp} m_i = \sum_{i \in \{p, C1\}} m_i \frac{\alpha(p, C1, q, C2; p) M_{p, C2}}{M_p} + \sum_{i \in \{p, C2\}} m_i \frac{\alpha(p, C2, q, C2; p) M_{p, C1}}{M_p} \\
 &= \frac{\alpha(p, C1, q, C2; p) M_{p, C2}}{M_p} \sum_{i \in \{p, C1\}} m_i + \frac{-\alpha(p, C1, q, C2; p) M_{p, C1}}{M_p} \sum_{i \in \{p, C1\}} m_i \\
 &= \frac{\alpha(p, C1, q, C2; p) M_{p, C2}}{M_p} M_{p, C1} - \frac{\alpha(p, C1, q, C2; p) M_{p, C1}}{M_p} M_{p, C2} = 0,
 \end{aligned} \tag{6.70}$$

where we used that $\alpha(p, C1, q, C2; p) = -\alpha(p, C2, q, C2; p)$. By the same arguments,

$$\sum_{j \in \{p, C\}} B_{ijp} m_j = 0. \tag{6.71}$$

This shows that $S'_1 = 0$.



Quadrupole-order term

In the summations of the ‘quadrupole-order’ term S'_2 (cf. equation 6.46), the individual terms contain, in principle, three different binaries. We show that only terms with binary pairs remain, significantly simplifying the expression. Using equation (6.66) to rewrite the summations of pairs of bodies, and using similar arguments as in Appendix 6.A.4, we find

$$\begin{aligned}
 -S'_2 &= \sum_{i < j} \frac{Gm_i m_j}{r_m} \frac{1}{2} \sum_{p \in \bar{\mathbf{B}}} \sum_{q \in \bar{\mathbf{B}}} B_{ijp} B_{ijq} \beta_2(\mathbf{r}_p, \mathbf{r}_q, \mathbf{r}_m) = \sum_{k \in \mathbf{B}} \sum_{i \in \{k, \mathbf{C}1\}} \sum_{j \in \{k, \mathbf{C}2\}} \frac{Gm_i m_j}{r_k} \frac{1}{2} \sum_{p \in \bar{\mathbf{B}}} \sum_{q \in \bar{\mathbf{B}}} B_{ijp} B_{ijq} \beta_2(\mathbf{r}_p, \mathbf{r}_q, \mathbf{r}_k) \\
 &= \frac{1}{2} \sum_{k \in \mathbf{B}} \sum_{i \in \{k, \mathbf{C}1\}} \sum_{j \in \{k, \mathbf{C}2\}} \sum_{p \in \bar{\mathbf{B}}} \sum_{q \in \bar{\mathbf{B}}} \frac{Gm_i m_j}{r_k} B_{ijp} B_{ijq} \beta_2(\mathbf{r}_p, \mathbf{r}_q, \mathbf{r}_k) \\
 &= \frac{1}{2} \sum_{k \in \mathbf{B}} \sum_{i \in \{k, \mathbf{C}1\}} \sum_{j \in \{k, \mathbf{C}2\}} \sum_{p \in \bar{\mathbf{B}}} \sum_{q \in \bar{\mathbf{B}}} \frac{Gm_i m_j}{r_k} B_{ijp} B_{ijq} \beta_2(\mathbf{r}_p, \mathbf{r}_q, \mathbf{r}_k) \\
 &= \frac{1}{2} \sum_{k \in \mathbf{B}} \sum_{p \in \bar{\mathbf{B}}} \sum_{q \in \bar{\mathbf{B}}} \frac{G}{r_k} \beta_2(\mathbf{r}_p, \mathbf{r}_q, \mathbf{r}_k) \sum_{i \in \{k, \mathbf{C}1\}} \sum_{j \in \{k, \mathbf{C}2\}} m_i m_j B_{ijp} B_{ijq} \\
 &= \frac{1}{2} \sum_{k \in \mathbf{B}} \sum_{p \in \bar{\mathbf{B}}} \sum_{q \in \bar{\mathbf{B}}} \frac{G}{r_k} \beta_2(\mathbf{r}_p, \mathbf{r}_q, \mathbf{r}_k) \sum_{j \in \{k, \mathbf{C}2\}} m_j \sum_{i \in \{p, \mathbf{C}\}} m_i B_{ijp} B_{ijq} \\
 &\quad \underbrace{\hspace{10em}}_{\equiv R_1} \\
 &\quad + \frac{1}{2} \sum_{k \in \mathbf{B}} \sum_{p \in \bar{\mathbf{B}}} \sum_{q \in \bar{\mathbf{B}}} \frac{G}{r_k} \beta_2(\mathbf{r}_p, \mathbf{r}_q, \mathbf{r}_k) \sum_{i \in \{k, \mathbf{C}1\}} m_i \sum_{j \in \{p, \mathbf{C}\}} m_j B_{ijp} B_{ijq} \cdot \\
 &\quad \underbrace{\hspace{10em}}_{\equiv R_2}
 \end{aligned} \tag{6.72}$$

In the last line, we split the summation of $p \in \{k, \mathbf{C}\}$ into separate summations over child n of k , and we denote the corresponding terms with R_n . Alternatively, because the expression is fully symmetric with respect to p and q (note that $\beta_2(\mathbf{r}_p, \mathbf{r}_q, \mathbf{r}_k) = \beta_2(\mathbf{r}_q, \mathbf{r}_p, \mathbf{r}_k)$), the split can also be applied to the summation of $q \in \{k, \mathbf{C}\}$.

The following arguments apply similarly to R_1 and R_2 ; here, we specify to R_1 . $B_{ijq} = 0$ if $q \notin \mathbf{B}_p(i, j)$. Because $i \in \{p, \mathbf{C}\}$, this implies that for nonzero terms,

1. $q = p$ or

2. $q \neq p$ with

(a) $p \in \{q.C\}$,

(b) $q \in \{p.C\}$ or

(c) $q \in \{k.C2\}$.

We address these cases below individually, starting with the case $q \neq p$ first.

(ii) (a) $q \neq p$ and $p \in \{q.C\}$. In this case, B_{ijq} in the summation of $i \in \{p.C\}$ is constant, and is given by

$$B_{ijq} = \alpha(p, j; q) \frac{M_{q, \text{CS}(p)}}{M_q}. \quad (6.73)$$

Therefore, B_{ijq} can be taken outside of the summation of $i \in \{p.C\}$, i.e.

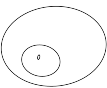
$$\sum_{i \in \{p.C\}} m_i B_{ijp} B_{ijq} = B_{ijq} \sum_{i \in \{p.C\}} m_i B_{ijp} = 0, \quad (6.74)$$

where the last property was shown explicitly in equation (6.70). In other words, the terms associated with case (ii) (a) cancel.

(ii) (b) $q \neq p$ and $q \in \{p.C\}$. If $q \in \{p.C\}$, then $B_{ijq} \neq 0$ only if $i \in \{q.C\}$. Furthermore, B_{ijp} is constant, giving

$$\sum_{i \in \{p.C\}} m_i B_{ijp} B_{ijq} = B_{ijp} \sum_{i \in \{q.C\}} m_i B_{ijq} = 0, \quad (6.75)$$

where we again used equation (6.70). Therefore, the terms associated with case (ii) (b) cancel as well.



(ii) (c) $q \neq p$ and $q \in \{k, C2\}$. Note that in the case of R_1 , $j \in \{k, C2\}$. If $q \in \{k, C2\}$, then $B_{ijq} \neq 0$ only if $j \in \{q, C\}$. Furthermore, B_{ijp} is constant when summing over j , whereas B_{ijq} is constant when summing over $i \in \{p, C\}$. Therefore,

$$\sum_{j \in \{k, C2\}} m_j \sum_{i \in \{p, C\}} m_i B_{ijp} B_{ijq} = \left(\sum_{i \in \{p, C\}} m_i B_{ijp} \right) \left(\sum_{j \in \{q, C\}} m_j B_{ijq} \right) = 0, \quad (6.76)$$

showing that terms associated with case (ii) (c) are zero.

(i) $q = p$. In the case $q = p$, the associated terms do not cancel. The double summation over p and q reduces to a single summation and $B_{ijq} = B_{ijp}$, hence

$$R_1 = \frac{1}{2} \sum_{k \in B} \sum_{p \in B} \sum_{p \in \{k, C1\}} \frac{G}{r_k} \beta_2(\mathbf{r}_p, \mathbf{r}_p, \mathbf{r}_k) \sum_{j \in \{k, C2\}} m_j \sum_{i \in \{p, C\}} m_i B_{ijp}^2. \quad (6.77)$$

The outermost summation of $i \in \{p, C\}$ does not cancel, and evaluates to

$$\begin{aligned} \sum_{i \in \{p, C\}} m_i B_{ijp}^2 &= \left(\sum_{i \in \{p, C1\}} + \sum_{i \in \{p, C2\}} \right) m_i B_{ijp}^2 = \frac{[\alpha(p, C1, k, C2; p)]^2 M_{p, C2}^2}{M_p^2} \sum_{i \in \{p, C1\}} m_i + \frac{[\alpha(p, C2, k, C2; p)]^2 M_{p, C1}^2}{M_p^2} \sum_{i \in \{p, C2\}} m_i \\ &= \frac{M_{p, C2}^2 M_{p, C1} + M_{p, C1}^2 M_{p, C2}}{M_p^2} = \frac{M_{p, C1} M_{p, C2}}{M_p}. \end{aligned} \quad (6.78)$$

The latter expression is independent of j and can therefore be taken in front of the summation of $j \in \{k, C2\}$. Therefore,

$$\begin{aligned} R_1 &= \frac{1}{2} \sum_{k \in B} \sum_{p \in B} \sum_{p \in \{k, C1\}} \frac{M_{p, C1} M_{p, C2}}{M_p} \frac{G}{r_k} \beta_2(\mathbf{r}_p, \mathbf{r}_p, \mathbf{r}_k) \sum_{j \in \{k, C2\}} m_j = \frac{1}{2} \sum_{k \in B} \sum_{p \in B} \sum_{p \in \{k, C1\}} \frac{M_{p, C1} M_{p, C2}}{M_p} \frac{G M_{k, C2}}{r_k} \beta_2(\mathbf{r}_p, \mathbf{r}_p, \mathbf{r}_k) \\ &= \frac{1}{2} \sum_{k \in B} \sum_{p \in B} \sum_{p \in \{k, C1\}} \frac{M_{p, C1} M_{p, C2}}{M_p} \frac{G M_{k, CS(p)}}{r_k} \beta_2(\mathbf{r}_p, \mathbf{r}_p, \mathbf{r}_k), \end{aligned} \quad (6.79)$$

where, as before, we used the notation $M_{k, \text{CS}(p)}$ to denote the mass of the child of k that is the sibling of the child of k that is connected to p (cf. Appendix 6.A.1).

The same arguments apply to R_2 , in which case $p \in \{k, \text{C2}\}$. The result is

$$R_2 = \frac{1}{2} \sum_{\substack{k \in \text{B} \\ p \in \{k, \text{C2}\}}} \sum_{\substack{p \in \text{B} \\ p \in \{k, \text{C2}\}}} \frac{M_{p, \text{C1}} M_{p, \text{C2}}}{M_p} \frac{GM_{k, \text{CS}(p)}}{r_k} \beta_2(\mathbf{r}_p, \mathbf{r}_p, \mathbf{r}_k). \quad (6.80)$$

Adding R_1 and R_2 , writing the separate summations of $p \in \{k, \text{C1}\}$ and $p \in \{k, \text{C2}\}$ again as a single summation of $p \in \{k, \text{C}\}$, and substituting the explicit expression for β_2 (cf. equation 6.51), we find the final simplified expression for the quadrupole-order term,

$$-S'_2 = R_1 + R_2 = \frac{1}{2} \sum_{k \in \text{B}} \sum_{\substack{p \in \text{B} \\ p \in \{k, \text{C}\}}} \frac{M_{p, \text{C1}} M_{p, \text{C2}}}{M_p} \frac{GM_{k, \text{CS}(p)}}{r_k} \left(\frac{r_p}{r_k} \right)^2 \left[3(\hat{\mathbf{r}}_p \cdot \hat{\mathbf{r}}_k)^2 - 1 \right]. \quad (6.81)$$

Therefore, the quadrupole-order term can be written as a double sum over binaries, where one of the binaries is a descendant of the other. As a consequence of the topology of the system and the resulting properties of the quantity B_{ijk} , all the terms depending on three different binaries cancel. Only terms with $p = q$ survive, and the resulting quadrupole-order terms individually only depend on *two* binary separation vectors.

The mathematical form of each of these terms is identical to the quadrupole-order term in the hierarchical three-body system. In particular, the ‘angular part’ $(1/2)[3(\hat{\mathbf{r}}_p \cdot \hat{\mathbf{r}}_k)^2 - 1]$ in equation (6.81) is the same as the second Legendre polynomial $P_2(x) = (1/2)(3x^2 - 1)$ with $x = \cos(i_{pk})$ the cosine of the angle between the ‘inner’ and ‘outer’ separation vectors $\mathbf{r}_p$ and $\mathbf{r}_k$, respectively, that appears in the hierarchical three-body Hamiltonian (Harrington 1968). In other words, at quadrupole order, the Hamiltonian of the hierarchical N -body system can be constructed from the three-body Hamiltonian.

Octupole-order term

For the ‘octupole-order’ term S'_3 (cf. equation 6.47), the individual terms contain, in principle, four different binaries. After rewriting, we find that terms with binary pairs and triplets remain; terms depending on quadrlets cancel. Following similar steps as in Appendix 6.A.4,

$$\begin{aligned}
 S'_3 &= \sum_{i < j} \frac{G m_i m_j}{r_m} \frac{1}{2} \alpha(i, j; m) \sum_{p \in \bar{B}} \sum_{q \in \bar{B}} \sum_{u \in \bar{B}} B_{ijp} B_{ijq} B_{iju} \beta_3(\mathbf{r}_p, \mathbf{r}_q, \mathbf{r}_u, \mathbf{r}_m) \\
 &= \sum_{k \in B} \sum_{i \in \{k, C1\}} \sum_{j \in \{k, C2\}} \frac{G m_i m_j}{r_k} \frac{1}{2} \alpha(i, j; k) \sum_{p \in B} \sum_{q \in \{k, C\}} \sum_{u \in B} B_{ijp} B_{ijq} B_{iju} \beta_3(\mathbf{r}_p, \mathbf{r}_q, \mathbf{r}_u, \mathbf{r}_k) \\
 &= \frac{1}{2} \sum_{k \in B} \sum_{p \in B} \sum_{q \in B} \sum_{u \in B} \frac{G}{r_k} \beta_3(\mathbf{r}_p, \mathbf{r}_q, \mathbf{r}_u, \mathbf{r}_k) \sum_{i \in \{k, C1\}} \sum_{j \in \{k, C2\}} \alpha(i, j; k) m_i m_j B_{ijp} B_{ijq} B_{iju}.
 \end{aligned}$$

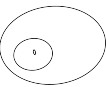
As before, we split the summation of $p \in \{k, C\}$ into two summations, R_1 and R_2 , over the children of k (q and u being equally valid choices). In the case of R_1 ,

$$\begin{aligned}
 R_1 &= \frac{1}{2} \sum_{k \in B} \sum_{p \in B} \sum_{q \in B} \sum_{u \in B} \frac{G}{r_k} \beta_3(\mathbf{r}_p, \mathbf{r}_q, \mathbf{r}_u, \mathbf{r}_k) \sum_{i \in \{p, C\}} \sum_{j \in \{k, C2\}} \alpha(i, j; k) m_i m_j B_{ijp} B_{ijq} B_{iju} \\
 &= \frac{1}{2} \sum_{k \in B} \sum_{p \in B} \sum_{q \in B} \sum_{u \in B} \frac{G}{r_k} \beta_3(\mathbf{r}_p, \mathbf{r}_q, \mathbf{r}_u, \mathbf{r}_k) \sum_{j \in \{k, C2\}} m_j \alpha(p, j; k) \sum_{i \in \{p, C\}} m_i B_{ijp} B_{ijq} B_{iju}, \quad (6.82)
 \end{aligned}$$

where we used that $\alpha(i, j; k) = \alpha(p, j; k)$ is constant in the summation of $i \in \{p, C\}$. The triple summations of p, q and u can be divided into the following five main cases, with several subcases. The subcases are those for which the corresponding $B_{ijp}, B_{ijq},$ etc. are nonzero. Specifying to R_1 (i.e. $p \in \{k, C1\}$), the cases are as follows.

1. $u = q = p$;
2. $q = p$ & $u \neq q$ with
 - (a) $p \in \{u, C\}$,

- (b) $u \in \{p.C\}$ or
- (c) $u \in \{k.C2\}$;
- 3. $u = p$ & $q \neq u$ with
 - (a) $p \in \{q.C\}$,
 - (b) $q \in \{p.C\}$ or
 - (c) $q \in \{k.C2\}$;
- 4. $u = q$ & $p \neq u$ with
 - (a) $p \in \{u.C\}$,
 - (b) $u \in \{p.C\}$ or
 - (c) $u \in \{k.C2\}$;
- 5. $u \neq q$ & $q \neq p$ with
 - (a) $p \in \{q.C\}$ & $p \in \{u.C\}$,
 - (b) $p \in \{q.C\}$ & $u \in \{p.C\}$,
 - (c) $p \in \{q.C\}$ & $u \in \{k.C2\}$,
 - (d) $q \in \{p.C\}$ & $p \in \{u.C\}$,
 - (e) $q \in \{p.C\}$ & $u \in \{p.C\}$,
 - (f) $q \in \{p.C\}$ & $u \in \{k.C2\}$,
 - (g) $q \in \{k.C2\}$ & $p \in \{u.C\}$,
 - (h) $q \in \{k.C2\}$ & $u \in \{p.C\}$,
 - (i) $q \in \{k.C2\}$ & $u \in \{k.C2\}$,
 - (j) $u \in \{p.C\}$ & $q \in \{p.C\}$ or
 - (k) $u \in \{p.C\}$ & $q \in \{k.C2\}$.



(i) $u = q = p$. With $u = q = p$, the corresponding outermost summation in R_1 becomes

$$\begin{aligned}
 \sum_{i \in \{p, C\}} m_i B_{ijp}^3 &= \left(\sum_{i \in \{p, C1\}} + \sum_{i \in \{p, C2\}} \right) m_i B_{ijp}^3 = \frac{[\alpha(p, C1, k, C2; p)]^3 M_{p, C1}^3}{M_p^3} \sum_{i \in \{p, C1\}} m_i + \frac{[\alpha(p, C2, k, C2; p)]^3 M_{p, C1}^3}{M_p^3} \sum_{i \in \{p, C2\}} m_i \\
 &= \frac{[\alpha(p, C1, k, C2; p)] M_{p, C2}^3 M_{p, C1} + [\alpha(p, C2, k, C2; p)] M_{p, C1}^3 M_{p, C2}}{M_p^3} = \alpha(p, C1, k, C2; p) M_{p, C1} M_{p, C2} \frac{M_{p, C2}^2 - M_{p, C1}^2}{M_p^3} \\
 &= \frac{M_{p, C1} M_{p, C2}}{M_p} \alpha(p, C1, k, C2; p) \frac{M_{p, C2} - M_{p, C1}}{M_p}. \tag{6.83}
 \end{aligned}$$

The contribution to R_1 due to case (i) is, therefore,

$$\begin{aligned}
 R_{1(i)} &= \frac{1}{2} \sum_{k \in B} \sum_{\substack{p \in B \\ p \in \{k, C1\}}} \frac{G}{r_k} \beta_3(\mathbf{r}_p, \mathbf{r}_p, \mathbf{r}_k) \sum_{j \in \{k, C2\}} m_j \alpha(p, j; k) \frac{M_{p, C1} M_{p, C2}}{M_p} \alpha(p, C1, k, C2; p) \frac{M_{p, C2} - M_{p, C1}}{M_p} \\
 &= \frac{1}{2} \sum_{k \in B} \sum_{\substack{p \in B \\ p \in \{k, C1\}}} \frac{M_{p, C1} M_{p, C2}}{M_p} \alpha(p, C1, k, C2; p) \frac{M_{p, C2} - M_{p, C1}}{M_p} \alpha(p, k, C2; k) \frac{G M_{k, C2}}{r_k} \beta_3(\mathbf{r}_p, \mathbf{r}_p, \mathbf{r}_k) \\
 &= \frac{1}{2} \sum_{k \in B} \sum_{\substack{p \in B \\ p \in \{k, C1\}}} \frac{M_{p, C1} M_{p, C2}}{M_p} \alpha(p, C1, k, CS(p); p) \frac{M_{p, C2} - M_{p, C1}}{M_p} \alpha(p, k, CS(p); k) \frac{G M_{k, CS(p)}}{r_k} \beta_3(\mathbf{r}_p, \mathbf{r}_p, \mathbf{r}_k). \tag{6.84}
 \end{aligned}$$

A similar result holds for $R_{2(i)}$, i.e.

$$R_{2(i)} = \frac{1}{2} \sum_{k \in B} \sum_{\substack{p \in B \\ p \in \{k, C2\}}} \frac{M_{p, C1} M_{p, C2}}{M_p} \alpha(p, C1, k, CS(p); p) \frac{M_{p, C2} - M_{p, C1}}{M_p} \alpha(p, k, CS(p); k) \frac{G M_{k, CS(p)}}{r_k} \beta_3(\mathbf{r}_p, \mathbf{r}_p, \mathbf{r}_k), \tag{6.85}$$

and, therefore,

$$R_{(i)} \equiv R_{1(i)} + R_{2(i)} = \frac{1}{2} \sum_{k \in \mathbf{B}} \sum_{\substack{p \in \mathbf{B} \\ p \in \{k, \mathbf{C}\}}} \frac{M_{p, \mathbf{C}1} M_{p, \mathbf{C}2}}{M_p} \alpha(p, \mathbf{C}1, k, \mathbf{C}\mathbf{S}(p); p) \frac{M_{p, \mathbf{C}2} - M_{p, \mathbf{C}1}}{M_p} \alpha(p, k, \mathbf{C}\mathbf{S}(p); k) \frac{G M_{k, \mathbf{C}\mathbf{S}(p)}}{r_k} \left(\frac{r_p}{r_k} \right)^3 \times \left[5 (\hat{\mathbf{r}}_p \cdot \hat{\mathbf{r}}_k)^3 - 3 (\hat{\mathbf{r}}_p \cdot \hat{\mathbf{r}}_k) \right], \quad (6.86)$$

where we substituted the explicit expression for β_3 using equation (6.52).

Note the presence of the ‘sign quantities’ α in $R_{1(i)}$ and $R_{2(i)}$. They appear because reversing the definition of the direction of $\mathbf{r}_p$ and $\mathbf{r}_k$, i.e. $\mathbf{r}'_p = -\mathbf{r}_p$ and $\mathbf{r}'_k = -\mathbf{r}_k$, should not change the new expressions in terms of the primed vectors. The minus signs arising from the redefinition are absorbed into the new quantities α' , which also change sign. A similar argument applies to the mass difference term $M_{p, \mathbf{C}2} - M_{p, \mathbf{C}1}$ (see also Appendix 6.A.5).

(ii) (a) $q = p \ \& \ u \neq q \ \& \ p \in \{u, \mathbf{C}\}$. While summing over $i \in \{p, \mathbf{C}\}$, B_{iju} is constant, giving

$$\sum_{i \in \{p, \mathbf{C}\}} m_i B_{ijp}^2 B_{iju} = \frac{\alpha(p, k, \mathbf{C}2; u) M_{u, \mathbf{C}\mathbf{S}(p)}}{M_u} \sum_{i \in \{p, \mathbf{C}\}} m_i B_{ijp}^2 = \frac{\alpha(p, k, \mathbf{C}2; u) M_{u, \mathbf{C}\mathbf{S}(p)}}{M_u} \frac{M_{p, \mathbf{C}1} M_{p, \mathbf{C}2}}{M_p}, \quad (6.87)$$

where we used equation (6.78).

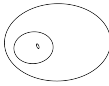
(ii) (b) $q = p \ \& \ u \neq q \ \& \ u \in \{p, \mathbf{C}\}$. While summing over $i \in \{p, \mathbf{C}\}$, B_{ijp} is constant, and $B_{iju} \neq 0$ only if $i \in \{u, \mathbf{C}\}$. Therefore,

$$\sum_{i \in \{p, \mathbf{C}\}} m_i B_{ijp}^2 B_{iju} = B_{ijp}^2 \sum_{i \in \{u, \mathbf{C}\}} m_i B_{iju} = 0, \quad (6.88)$$

where the last step follows from equation (6.70).

(ii) (c) $q = p \ \& \ u \neq q \ \& \ u \in \{k, \mathbf{C}2\}$. While summing over $i \in \{p, \mathbf{C}\}$, B_{iju} is constant, whereas B_{ijp} is constant while summing over $j \in \{k, \mathbf{C}2\}$. Furthermore, $B_{iju} \neq 0$ only if $j \in \{u, \mathbf{C}\}$. This implies

$$\sum_{j \in \{k, \mathbf{C}2\}} m_j \sum_{i \in \{p, \mathbf{C}\}} m_i B_{ijp}^2 B_{iju} = \left(\sum_{j \in \{u, \mathbf{C}\}} m_j B_{iju} \right) \left(\sum_{i \in \{p, \mathbf{C}\}} m_i B_{ijp}^2 \right) = 0. \quad (6.89)$$



Therefore, for R_1 the total contribution due to case (ii) is

$$\begin{aligned}
 R_{1(ii)} &= \frac{1}{2} \sum_{k \in B} \sum_{\substack{p \in B \\ p \in \{k, C1\}}} \sum_{\substack{u \in B \\ u \in \{k, C\}}} \sum_{\substack{p \in B \\ p \in \{u, C\}}} \frac{G}{r_k} \beta_3(\mathbf{r}_p, \mathbf{r}_p, \mathbf{r}_u, \mathbf{r}_k) \sum_{j \in \{k, C2\}} m_j \alpha(p, j; k) \frac{\alpha(p, k, C2; u) M_{u, CS(p)}}{M_u} \frac{M_{p, C1} M_{p, C2}}{M_p} \\
 &= \frac{1}{2} \sum_{k \in B} \sum_{\substack{p \in B \\ p \in \{k, C1\}}} \sum_{\substack{u \in B \\ u \in \{k, C\}}} \sum_{\substack{p \in B \\ p \in \{u, C\}}} \frac{M_{p, C1} M_{p, C2}}{M_p} \alpha(p, k, C2; k) \frac{G M_{k, C2}}{r_k} \frac{\alpha(p, k, C2; u) M_{u, CS(p)}}{M_u} \beta_3(\mathbf{r}_p, \mathbf{r}_p, \mathbf{r}_u, \mathbf{r}_k) \\
 &= \frac{1}{2} \sum_{k \in B} \sum_{\substack{p \in B \\ p \in \{k, C1\}}} \sum_{\substack{u \in B \\ u \in \{k, C\}}} \sum_{\substack{p \in B \\ p \in \{u, C\}}} \frac{M_{p, C1} M_{p, C2}}{M_p} \alpha(p, k, CS(p); k) \frac{G M_{k, CS(p)}}{r_k} \frac{\alpha(p, k, CS(p); u) M_{u, CS(p)}}{M_u} \left(\frac{r_p}{r_k} \right)^2 \left(\frac{r_u}{r_k} \right) \\
 &\quad \times \left[5 (\hat{\mathbf{r}}_p \cdot \hat{\mathbf{r}}_k)^2 (\hat{\mathbf{r}}_u \cdot \hat{\mathbf{r}}_k) - 2 (\hat{\mathbf{r}}_p \cdot \hat{\mathbf{r}}_k) (\hat{\mathbf{r}}_p \cdot \hat{\mathbf{r}}_u) - (\hat{\mathbf{r}}_u \cdot \hat{\mathbf{r}}_k) \right]. \tag{6.90}
 \end{aligned}$$

A similar result applies to $R_{2(ii)}$. This gives

$$\begin{aligned}
 R_{(ii)} \equiv R_{1(ii)} + R_{2(ii)} &= \frac{1}{2} \sum_{k \in B} \sum_{\substack{p \in B \\ p \in \{k, C\}}} \sum_{\substack{u \in B \\ u \in \{k, C\}}} \sum_{\substack{p \in B \\ p \in \{u, C\}}} \frac{M_{p, C1} M_{p, C2}}{M_p} \alpha(p, k, CS(p); k) \frac{G M_{k, CS(p)}}{r_k} \frac{\alpha(p, k, CS(p); u) M_{u, CS(p)}}{M_u} \left(\frac{r_p}{r_k} \right)^2 \left(\frac{r_u}{r_k} \right) \\
 &\quad \times \left[5 (\hat{\mathbf{r}}_p \cdot \hat{\mathbf{r}}_k)^2 (\hat{\mathbf{r}}_u \cdot \hat{\mathbf{r}}_k) - 2 (\hat{\mathbf{r}}_p \cdot \hat{\mathbf{r}}_k) (\hat{\mathbf{r}}_p \cdot \hat{\mathbf{r}}_u) - (\hat{\mathbf{r}}_u \cdot \hat{\mathbf{r}}_k) \right]. \tag{6.91}
 \end{aligned}$$

Note that the sign quantities $\alpha(p, k, CS(p); k)$ and $\alpha(p, k, CS(p); u)$ ensure invariance with respect to the definition of the directions of $\mathbf{r}_k$ and $\mathbf{r}_u$, respectively.

(iii) (a, b, c) $u = p$ & $q \neq u$. This case is completely analogous to case (ii), now with the roles of q and u reversed. Note that β_3 satisfies $\beta_3(\mathbf{r}_p, \mathbf{r}_p, \mathbf{r}_u, \mathbf{r}_k) = \beta_3(\mathbf{r}_p, \mathbf{r}_u, \mathbf{r}_p, \mathbf{r}_k)$. Therefore, $R_{(ii)} = R_{(iii)}$.

The case (iv) also turns out to satisfy $R_{(iv)} = R_{(iii)} = R_{(iii)}$, although the equality of $R_{(iv)}$ to $R_{(ii)}$ and $R_{(iii)}$ is not immediately clear. For this reason, we treat this case in detail.

(iv) (a) $u = q \& p \neq u \& p \in \{u.C\}$. While summing over $i \in \{p.C\}$, B_{iju} is constant, giving

$$\sum_{i \in \{p.C\}} m_i B_{ijp} B_{iju}^2 = B_{iju}^2 \sum_{i \in \{p.C\}} m_i B_{ijp} = 0. \quad (6.92)$$

(iv) (b) $u = q \& p \neq u \& u \in \{p.C\}$. While summing over $i \in \{p.C\}$, $B_{iju} \neq 0$ only if $i \in \{u.C\}$, and B_{ijp} is constant. Therefore,

$$\sum_{i \in \{p.C\}} m_i B_{ijp} B_{iju}^2 = B_{ijp} \sum_{i \in \{u.C\}} m_i B_{iju}^2 = \alpha(u, k.C2; p) \frac{M_{p.CS(u)}}{M_p} \frac{M_{u.C1} M_{u.C2}}{M_u}. \quad (6.93)$$

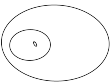
In this case, $\alpha(i, j; k) = \alpha(u, k.C2; k)$.

(iv) (c) $u = q \& p \neq u \& u \in \{k.C2\}$. While summing over $i \in \{p.C\}$, B_{iju} is constant, whereas B_{ijp} is constant while summing over $j \in \{k.C2\}$. Furthermore, $B_{iju} \neq 0$ only if $j \in \{u.C\}$. This implies

$$\sum_{j \in \{k.C2\}} m_j \sum_{i \in \{p.C\}} m_i B_{ijp} B_{iju}^2 = \left(\sum_{j \in \{u.C\}} m_j B_{iju}^2 \right) \left(\sum_{i \in \{p.C\}} m_i B_{ijp} \right) = 0. \quad (6.94)$$

Specifying to $R_{1,1}$, the total contribution due to case (iv) is

$$\begin{aligned} R_{1,1}^{(iv)} &= \frac{1}{2} \sum_{k \in B} \sum_{\substack{p \in B \\ p \in \{k.C1\}}} \sum_{\substack{u \in B \\ u \in \{k.C1\}}} \sum_{\substack{u \in B \\ u \in \{p.C\}}} \frac{G}{r_k} \beta_3(\mathbf{r}_p, \mathbf{r}_u, \mathbf{r}_k) \sum_{j \in \{k.C2\}} m_j \alpha(u, k.C2; k) \frac{\alpha(u, k.C2; p) M_{p.CS(u)}}{M_p} \frac{M_{u.C1} M_{u.C2}}{M_u} \\ &= \frac{1}{2} \sum_{k \in B} \sum_{\substack{p \in B \\ p \in \{k.C1\}}} \sum_{\substack{u \in B \\ u \in \{k.C1\}}} \sum_{\substack{u \in B \\ u \in \{p.C\}}} \frac{M_{u.C1} M_{u.C2}}{M_u} \alpha(u, k.C2; k) \frac{G M_{k.C2}}{r_k} \frac{\alpha(u, k.C2; p) M_{p.CS(u)}}{M_p} \beta_3(\mathbf{r}_p, \mathbf{r}_u, \mathbf{r}_k) \\ &= \frac{1}{2} \sum_{k \in B} \sum_{\substack{p \in B \\ p \in \{k.C1\}}} \sum_{\substack{u \in B \\ u \in \{k.C1\}}} \sum_{\substack{u \in B \\ u \in \{p.C\}}} \frac{M_{u.C1} M_{u.C2}}{M_u} \alpha(u, k.CS(p); k) \frac{G M_{k.CS(p)}}{r_k} \frac{\alpha(u, k.CS(p); p) M_{p.CS(u)}}{M_p} \beta_3(\mathbf{r}_p, \mathbf{r}_u, \mathbf{r}_k). \end{aligned} \quad (6.95)$$



This gives

$$R_{(\text{iv})} \equiv R_{1(\text{iv})} + R_{2(\text{iv})} = \frac{1}{2} \sum_{k \in \mathbf{B}} \sum_{\substack{p \in \mathbf{B} \\ p \in \{k, \mathbf{C}\}}} \sum_{\substack{u \in \mathbf{B} \\ u \in \{k, \mathbf{C}\}}} \frac{M_{u, \mathbf{C}1} M_{u, \mathbf{C}2}}{M_u} \alpha(u, k, \mathbf{CS}(p); k) \frac{GM_{k, \mathbf{CS}(p)}}{r_k} \frac{\alpha(u, k, \mathbf{CS}(p); p) M_{p, \mathbf{CS}(u)}}{M_p} \beta_3(\mathbf{r}_p, \mathbf{r}_u, \mathbf{r}_k). \quad (6.96)$$

In equation (6.96), $p, u \in \{k, \mathbf{C}\}$ and $u \in \{p, \mathbf{C}\}$. Therefore,

$$\begin{aligned} \alpha(u, k, \mathbf{CS}(p); k) &= \alpha(u, k, \mathbf{CS}(u); k); \\ M_{k, \mathbf{CS}(p)} &= M_{k, \mathbf{CS}(u)}; \\ \alpha(u, k, \mathbf{CS}(p); p) &= \alpha(u, k, \mathbf{CS}(u); p). \end{aligned} \quad (6.97)$$

Inserting these relations into equation (6.96) and, subsequently, interchanging p and u (i.e. only changing notation), we find

$$R_{(\text{iv})} = \frac{1}{2} \sum_{k \in \mathbf{B}} \sum_{\substack{p \in \mathbf{B} \\ p \in \{k, \mathbf{C}\}}} \sum_{\substack{u \in \mathbf{B} \\ u \in \{k, \mathbf{C}\}}} \frac{M_{u, \mathbf{C}1} M_{u, \mathbf{C}2}}{M_u} \alpha(u, k, \mathbf{CS}(u); k) \frac{GM_{k, \mathbf{CS}(u)}}{r_k} \frac{\alpha(u, k, \mathbf{CS}(u); p) M_{p, \mathbf{CS}(u)}}{M_p} \beta_3(\mathbf{r}_p, \mathbf{r}_u, \mathbf{r}_k) \quad (6.98)$$

$$\begin{aligned} &= \frac{1}{2} \sum_{k \in \mathbf{B}} \sum_{\substack{u \in \mathbf{B} \\ u \in \{k, \mathbf{C}\}}} \sum_{\substack{p \in \mathbf{B} \\ p \in \{k, \mathbf{C}\}}} \frac{M_{p, \mathbf{C}1} M_{p, \mathbf{C}2}}{M_p} \alpha(p, k, \mathbf{CS}(p); k) \frac{GM_{k, \mathbf{CS}(p)}}{r_k} \frac{\alpha(p, k, \mathbf{CS}(p); u) M_{u, \mathbf{CS}(p)}}{M_u} \beta_3(\mathbf{r}_u, \mathbf{r}_p, \mathbf{r}_k) \\ &= \frac{1}{2} \sum_{k \in \mathbf{B}} \sum_{\substack{p \in \mathbf{B} \\ p \in \{k, \mathbf{C}\}}} \sum_{\substack{u \in \mathbf{B} \\ u \in \{k, \mathbf{C}\}}} \frac{M_{p, \mathbf{C}1} M_{p, \mathbf{C}2}}{M_p} \alpha(p, k, \mathbf{CS}(p); k) \frac{GM_{k, \mathbf{CS}(p)}}{r_k} \frac{\alpha(p, k, \mathbf{CS}(p); u) M_{u, \mathbf{CS}(p)}}{M_u} \left(\frac{r_p}{r_k} \right)^2 \left(\frac{r_u}{r_k} \right) \end{aligned} \quad (6.99)$$

$$\times \left[5 (\hat{\mathbf{r}}_p \cdot \hat{\mathbf{r}}_k)^2 (\hat{\mathbf{r}}_u \cdot \hat{\mathbf{r}}_k) - 2 (\hat{\mathbf{r}}_p \cdot \hat{\mathbf{r}}_k) (\hat{\mathbf{r}}_p \cdot \hat{\mathbf{r}}_u) - (\hat{\mathbf{r}}_u \cdot \hat{\mathbf{r}}_k) \right] = R_{(\text{ii})} = R_{(\text{iii})}. \quad (6.100)$$

Note that $\beta_3(\mathbf{r}_u, \mathbf{r}_p, \mathbf{r}_k) = \beta_3(\mathbf{r}_p, \mathbf{r}_p, \mathbf{r}_k)$.

To conclude, the nonzero terms from cases (ii), (iii) and (iv) are equal, and so adding them gives $R_{(\text{ii})} + R_{(\text{iii})} + R_{(\text{iv})} = 3R_{(\text{ii})}$.

(v) $u \neq q$ & $q \neq p$. In all subcases listed above, two of the quantities in $B_{ijp} B_{ijq} B_{iju}$ are always constant when summing over a particular binary. Consequently, all terms include the term of the form

$$\sum_{i \in \{w, C\}} m_i B_{ijw} = 0, \quad (6.101)$$

where w is any of p, q and u . Therefore, all terms associated with $u \neq q$ and $q \neq p$ are zero, i.e. $R_{(v)} = 0$.

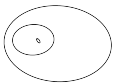
We conclude that the octupole-order term can be written as

$$\begin{aligned} S'_3 &= R_{(i)} + R_{(ii)} + R_{(iii)} + R_{(iv)} + R_{(v)} \\ &= \frac{1}{2} \sum_{k \in B} \sum_{\substack{p \in B \\ p \in \{k, C\}}} \frac{M_{p,C1} M_{p,C2}}{M_p} \alpha(p, C1, k, CS(p); p) \frac{M_{p,C2} - M_{p,C1}}{M_p} \alpha(p, k, CS(p); k) \frac{GM_{k,CS(p)}}{r_k} \left(\frac{r_p}{r_k} \right)^3 \left[5 (\hat{\mathbf{r}}_p \cdot \hat{\mathbf{r}}_k)^3 - 3 (\hat{\mathbf{r}}_p \cdot \hat{\mathbf{r}}_k) \right] \\ &\quad + \frac{3}{2} \sum_{k \in B} \sum_{\substack{p \in B \\ p \in \{k, C\}}} \sum_{\substack{u \in B \\ u \in \{k, C\}}} \frac{M_{p,C1} M_{p,C2}}{M_p} \alpha(p, k, CS(p); k) \frac{GM_{k,CS(p)}}{r_k} \frac{\alpha(p, k, CS(p); u) M_{u,CS(p)}}{M_u} \left(\frac{r_p}{r_k} \right)^2 \left(\frac{r_u}{r_k} \right) \\ &\quad \times \left[5 (\hat{\mathbf{r}}_p \cdot \hat{\mathbf{r}}_k)^2 (\hat{\mathbf{r}}_u \cdot \hat{\mathbf{r}}_k) - 2 (\hat{\mathbf{r}}_p \cdot \hat{\mathbf{r}}_k) (\hat{\mathbf{r}}_p \cdot \hat{\mathbf{r}}_u) - (\hat{\mathbf{r}}_u \cdot \hat{\mathbf{r}}_k) \right]. \end{aligned} \quad (6.102)$$

The first major term consists of binary pair terms, and is mathematically equivalent to the octupole-order term in the hierarchical three-body problem. The second major term consists of binary triplet terms, and is mathematically equivalent to the octupole-order ‘cross term’ in the hierarchical four-body problem (‘3+1’ configuration), which was derived previously by Hamers et al. (2015) (cf. Eq. A7d of the latter paper).

Hexadecupole-order term (pairwise)

Extrapolating the above results, at the ‘hexadecupole’ order, i.e. $\mathcal{O}(x^4)$, terms in the Hamiltonian after rewriting contain terms depending on binary pairs, triplets and quadrlets. For all these combinations, the binaries need to be connected for the terms to be nonvanishing. Furthermore, terms depending on five binaries, i.e. quintlet terms, vanish because of the property equation (6.70).



Using equation (6.66), S'_4 (cf. equation 6.48) is given by

$$\begin{aligned}
 -S'_4 &= \sum_{i < j} \frac{G m_i m_j}{r_m} \frac{1}{8} \sum_{p \in \bar{B}} \sum_{q \in \bar{B}} \sum_{u \in \bar{B}} \sum_{v \in \bar{B}} B_{ijp} B_{ijq} B_{iju} B_{ijv} \beta_4(\mathbf{r}_p, \mathbf{r}_q, \mathbf{r}_u, \mathbf{r}_v, \mathbf{r}_m) \\
 &= \sum_{k \in B} \sum_{i \in \{k, C1\}} \sum_{j \in \{k, C2\}} \frac{G m_i m_j}{r_k} \frac{1}{8} \sum_{p \in B} \sum_{q \in \{k, C\}} \sum_{u \in B} \sum_{v \in B} B_{ijp} B_{ijq} B_{iju} B_{ijv} \beta_4(\mathbf{r}_p, \mathbf{r}_q, \mathbf{r}_u, \mathbf{r}_v, \mathbf{r}_k) \\
 &= \frac{1}{8} \sum_{k \in B} \sum_{p \in B} \sum_{q \in B} \sum_{u \in B} \sum_{v \in B} \frac{G}{r_k} \beta_4(\mathbf{r}_p, \mathbf{r}_q, \mathbf{r}_u, \mathbf{r}_v, \mathbf{r}_k) \sum_{i \in \{k, C1\}} \sum_{j \in \{k, C2\}} m_i m_j B_{ijp} B_{ijq} B_{iju} B_{ijv} \\
 &= -(S'_{4;2} + S'_{4;3} + S'_{4;4}),
 \end{aligned} \tag{6.103}$$

where in the last line, we introduced the notation $S'_{n,l}$ to denote terms of order n that contain summations with individual terms depending on l binaries; generally, $l \leq n$. In this case, these terms are the binary pair terms ($S'_{4;2}$), binary triplet terms ($S'_{4;3}$) and binary quadruplet terms ($S'_{4;4}$).

Rather than deriving all terms at hexadecapole order explicitly, we here restrict to the binary pair terms, which we expect to be typically most important. We refer to Appendix 6.A.4 below for a formal generalization which includes all terms at any order.

Setting $p = q = u = v$ in equation (6.103), we find

$$\begin{aligned}
 -S'_{4;2} &= \frac{1}{8} \sum_{k \in B} \sum_{p \in B} \sum_{q \in \{k, C\}} \frac{G}{r_k} \beta_4(\mathbf{r}_p, \mathbf{r}_p, \mathbf{r}_p, \mathbf{r}_p, \mathbf{r}_k) \sum_{i \in \{k, C1\}} \sum_{j \in \{k, C2\}} m_i m_j B_{ijp}^4 \\
 &= \frac{1}{8} \sum_{k \in B} \sum_{p \in B} \sum_{q \in \{k, C1\}} \frac{G}{r_k} \beta_4(\mathbf{r}_p, \mathbf{r}_p, \mathbf{r}_p, \mathbf{r}_p, \mathbf{r}_k) \sum_{j \in \{k, C2\}} m_j \sum_{i \in \{p, C\}} m_i B_{ijp}^4 \\
 &\quad + \frac{1}{8} \sum_{k \in B} \sum_{p \in B} \sum_{q \in \{k, C2\}} \frac{G}{r_k} \beta_4(\mathbf{r}_p, \mathbf{r}_p, \mathbf{r}_p, \mathbf{r}_p, \mathbf{r}_k) \sum_{i \in \{k, C1\}} m_i \sum_{j \in \{p, C\}} m_j B_{ijp}^4.
 \end{aligned} \tag{6.104}$$

The outermost summation over bodies, in the case of $i \in \{p.C\}$, is given by

$$\begin{aligned} \sum_{i \in \{p.C\}} m_i B_{ijp}^4 &= \left(\sum_{i \in \{p.C1\}} m_i + \sum_{i \in \{p.C2\}} m_i \right) B_{ijp}^4 = \frac{[\alpha(p.C1, k.C2; p)]^4 M_{p.C2}^4}{M_p^4} \sum_{i \in \{p.C1\}} m_i + \frac{[\alpha(p.C2, k.C2; p)]^4 M_{p.C1}^4}{M_p^4} \sum_{i \in \{p.C2\}} m_i \\ &= \frac{M_{p.C2}^4 M_{p.C1} + M_{p.C1}^4 M_{p.C2}}{M_p^4} = M_{p.C1} M_{p.C2} \frac{M_{p.C2}^3 + M_{p.C1}^3}{M_p^4} = \frac{M_{p.C1} M_{p.C2}}{M_p} \frac{M_{p.C1}^2 - M_{p.C1} M_{p.C2} + M_{p.C2}^2}{M_p^2}. \end{aligned} \quad (6.105)$$

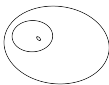
The same result applies to the summation of $j \in \{p.C\}$. We conclude that $S'_{4,2}$ is given by

$$\begin{aligned} -S'_{4,2} &= \frac{1}{8} \sum_{k \in B} \sum_{\substack{p \in B \\ p \in \{k.C\}}} \frac{M_{p.C1} M_{p.C2}}{M_p} \frac{M_{p.C1}^2 - M_{p.C1} M_{p.C2} + M_{p.C2}^2}{M_p^2} \frac{GM_{k.CS(p)}}{r_k} \beta_4(\mathbf{r}_p, \mathbf{r}_p, \mathbf{r}_p, \mathbf{r}_k) \\ &= \frac{1}{8} \sum_{k \in B} \sum_{\substack{p \in B \\ p \in \{k.C\}}} \frac{M_{p.C1} M_{p.C2}}{M_p} \frac{M_{p.C1}^2 - M_{p.C1} M_{p.C2} + M_{p.C2}^2}{M_p^2} \frac{GM_{k.CS(p)}}{r_k} \left(\frac{r_p}{r_k} \right)^4 \left[35 (\hat{\mathbf{r}}_p \cdot \hat{\mathbf{r}}_k)^4 - 30 (\hat{\mathbf{r}}_p \cdot \hat{\mathbf{r}}_k)^2 + 3 \right]. \end{aligned} \quad (6.106)$$

Note that $M_{p.C1}^2 - M_{p.C1} M_{p.C2} + M_{p.C2}^2 > 0$ for any combination of non-zero and positive $M_{p.C1}$ and $M_{p.C2}$. Therefore, unlike the pairwise octupole-order terms (cf. Appendix 6.A.4), the pairwise hexadecupole-order terms never exactly vanish.

Dotriacontupole-order term (pairwise)

As for the hexadecupole-order term, we only explicitly derive the binary pair terms for terms of order x^5 , i.e. the ‘dotriacontupole’-order terms. For formal expressions for all binary interactions, we refer to Appendix 6.A.4.



Using equations (6.49) and (6.66),

$$\begin{aligned}
S'_5 &= \sum_{i < j} \frac{G m_i m_j}{r_m} \frac{1}{8} \alpha(i, j; m) \sum_{p \in \bar{\mathbf{B}}} \sum_{q \in \bar{\mathbf{B}}} \sum_{u \in \bar{\mathbf{B}}} \sum_{v \in \bar{\mathbf{B}}} \sum_{w \in \bar{\mathbf{B}}} B_{ijp} \sum_{w \in \bar{\mathbf{B}}} B_{ijp} B_{ijq} B_{iju} B_{ijv} B_{ijw} \beta_5(\mathbf{r}_p, \mathbf{r}_q, \mathbf{r}_u, \mathbf{r}_v, \mathbf{r}_w, \mathbf{r}_m) \\
&= \sum_{k \in \mathbf{B}} \sum_{i \in \{k, \mathbf{C}1\}} \sum_{j \in \{k, \mathbf{C}2\}} \frac{G m_i m_j}{r_k} \frac{1}{8} \alpha(i, j; k) \sum_{p \in \bar{\mathbf{B}}} \sum_{q \in \bar{\mathbf{B}}} \sum_{u \in \bar{\mathbf{B}}} \sum_{v \in \bar{\mathbf{B}}} \sum_{w \in \bar{\mathbf{B}}} B_{ijp} B_{ijq} B_{iju} B_{ijv} B_{ijw} \beta_5(\mathbf{r}_p, \mathbf{r}_q, \mathbf{r}_u, \mathbf{r}_v, \mathbf{r}_w, \mathbf{r}_k) \\
&= \frac{1}{8} \sum_{k \in \mathbf{B}} \sum_{p \in \bar{\mathbf{B}}} \sum_{q \in \bar{\mathbf{B}}} \sum_{u \in \bar{\mathbf{B}}} \sum_{v \in \bar{\mathbf{B}}} \sum_{w \in \bar{\mathbf{B}}} \sum_{p \in \{k, \mathbf{C}1\}} \sum_{q \in \{k, \mathbf{C}2\}} \sum_{u \in \{k, \mathbf{C}1\}} \sum_{v \in \{k, \mathbf{C}2\}} \sum_{w \in \{k, \mathbf{C}2\}} \frac{G}{r_k} \beta_5(\mathbf{r}_p, \mathbf{r}_q, \mathbf{r}_u, \mathbf{r}_v, \mathbf{r}_w, \mathbf{r}_k) \sum_{i \in \{k, \mathbf{C}1\}} \sum_{j \in \{k, \mathbf{C}2\}} \alpha(i, j; k) m_i m_j B_{ijp} B_{ijq} B_{iju} B_{ijv} B_{ijw} \\
&= S'_{5,2} + S'_{5,3} + S'_{5,4} + S'_{5,5},
\end{aligned} \tag{6.107}$$

with

$$\begin{aligned}
S'_{5,2} &= \frac{1}{8} \sum_{k \in \mathbf{B}} \sum_{p \in \bar{\mathbf{B}}} \sum_{q \in \bar{\mathbf{B}}} \sum_{u \in \bar{\mathbf{B}}} \sum_{v \in \bar{\mathbf{B}}} \sum_{w \in \bar{\mathbf{B}}} \sum_{p \in \{k, \mathbf{C}1\}} \sum_{q \in \{k, \mathbf{C}2\}} \sum_{u \in \{k, \mathbf{C}1\}} \sum_{v \in \{k, \mathbf{C}2\}} \sum_{w \in \{k, \mathbf{C}2\}} \frac{G}{r_k} \beta_5(\mathbf{r}_p, \mathbf{r}_q, \mathbf{r}_u, \mathbf{r}_v, \mathbf{r}_w, \mathbf{r}_k) \sum_{i \in \{k, \mathbf{C}1\}} \sum_{j \in \{k, \mathbf{C}2\}} \alpha(i, j; k) m_i m_j B_{ijp}^5 \\
&= \frac{1}{8} \sum_{k \in \mathbf{B}} \sum_{p \in \bar{\mathbf{B}}} \sum_{q \in \bar{\mathbf{B}}} \sum_{u \in \bar{\mathbf{B}}} \sum_{v \in \bar{\mathbf{B}}} \sum_{w \in \bar{\mathbf{B}}} \sum_{p \in \{k, \mathbf{C}1\}} \sum_{q \in \{k, \mathbf{C}2\}} \sum_{u \in \{k, \mathbf{C}2\}} \sum_{v \in \{k, \mathbf{C}2\}} \sum_{w \in \{k, \mathbf{C}2\}} \frac{G}{r_k} \beta_5(\mathbf{r}_p, \mathbf{r}_q, \mathbf{r}_u, \mathbf{r}_v, \mathbf{r}_w, \mathbf{r}_k) \alpha(p, k, \mathbf{C}2; k) \sum_{j \in \{k, \mathbf{C}2\}} m_j \sum_{i \in \{p, \mathbf{C}1\}} m_i B_{ijp}^5
\end{aligned} \tag{6.108}$$

$$\begin{aligned}
&+ \frac{1}{8} \sum_{k \in \mathbf{B}} \sum_{p \in \bar{\mathbf{B}}} \sum_{q \in \bar{\mathbf{B}}} \sum_{u \in \bar{\mathbf{B}}} \sum_{v \in \bar{\mathbf{B}}} \sum_{w \in \bar{\mathbf{B}}} \sum_{p \in \{k, \mathbf{C}2\}} \sum_{q \in \{k, \mathbf{C}1\}} \sum_{u \in \{k, \mathbf{C}1\}} \sum_{v \in \{k, \mathbf{C}1\}} \sum_{w \in \{k, \mathbf{C}1\}} \frac{G}{r_k} \beta_5(\mathbf{r}_p, \mathbf{r}_q, \mathbf{r}_u, \mathbf{r}_v, \mathbf{r}_w, \mathbf{r}_k) \alpha(p, k, \mathbf{C}1; k) \sum_{i \in \{k, \mathbf{C}1\}} m_i \sum_{j \in \{p, \mathbf{C}2\}} m_j B_{ijp}^5.
\end{aligned} \tag{6.109}$$

The outermost summation over bodies, in the case of $i \in \{p.C\}$, is given by

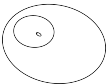
$$\begin{aligned}
 \sum_{i \in \{p.C\}} m_i B_{i,p}^5 &= \left(\sum_{i \in \{p.C1\}} m_i + \sum_{i \in \{p.C2\}} m_i \right) B_{i,p}^5 \\
 &= \frac{[\alpha(p.C1, k.C2; p)]^5 M_{p.C2}^5}{M_p^5} \sum_{i \in \{p.C1\}} m_i + \frac{[\alpha(p.C2, k.C2; p)]^5 M_{p.C1}^5}{M_p^5} \sum_{i \in \{p.C2\}} m_i \\
 &= \frac{[\alpha(p.C1, k.C2; p)] M_{p.C2}^5 M_{p.C1} + [\alpha(p.C2, k.C2; p)] M_{p.C1}^5 M_{p.C2}}{M_p^5} = \alpha(p.C1, k.C2; p) M_{p.C1} M_{p.C2} \frac{M_{p.C2}^4 - M_{p.C1}^4}{M_p^5} \\
 &= \alpha(p.C1, k.C2; p) \frac{M_{p.C1} M_{p.C2}}{M_p} \frac{M_{p.C2}^2 - M_{p.C1}^2}{M_p} \frac{M_{p.C1}^2 + M_{p.C2}^2}{M_p^2}. \tag{6.110}
 \end{aligned}$$

The same result holds for $j \in \{p.C\}$. Therefore,

$$\begin{aligned}
 S'_{5,2} &= \frac{1}{8} \sum_{k \in B} \sum_{\substack{p \in B \\ p \in \{k.C\}}} \frac{M_{p.C1} M_{p.C2}}{M_p} \alpha(p.C1, p.CS(p); p) \frac{M_{p.C2} - M_{p.C1}}{M_p} \frac{M_{p.C1}^2 + M_{p.C2}^2}{M_p^2} \alpha(p, k, CS(p); k) \frac{GM_{k,CS(p)}}{r_k} \\
 &\quad \times \beta_5(\mathbf{r}_p, \mathbf{r}_p, \mathbf{r}_p, \mathbf{r}_p, \mathbf{r}_p, \mathbf{r}_k) \\
 &= \frac{1}{8} \sum_{k \in B} \sum_{\substack{p \in B \\ p \in \{k.C\}}} \frac{M_{p.C1} M_{p.C2}}{M_p} \alpha(p.C1, p.CS(p); p) \frac{M_{p.C2} - M_{p.C1}}{M_p} \frac{M_{p.C1}^2 + M_{p.C2}^2}{M_p^2} \alpha(p, k, CS(p); k) \frac{GM_{k,CS(p)}}{r_k} \left(\frac{r_p}{r_k} \right)^5 \\
 &\quad \times \left[63 (\hat{\mathbf{r}}_p \cdot \hat{\mathbf{r}}_k)^5 - 70 (\hat{\mathbf{r}}_p \cdot \hat{\mathbf{r}}_k)^3 + 15 (\hat{\mathbf{r}}_p \cdot \hat{\mathbf{r}}_k) \right]. \tag{6.111}
 \end{aligned}$$

Generalisation

In the above appendices, we have shown that H can be written in terms of summations of only binaries, and we have derived explicit expressions for all terms up and including octupole order, and including the binary pair terms at hexadecupole and dotriacontupole order. Here, using similar arguments as before, we generalize these results and derive a formal expression for S'_n at *arbitrary* order n including *all* binary interactions (i.e. pairwise terms, triplet terms, etc.).



Equations (6.55) and (6.65) combined read

$$S'_n = (-1)^{n+1} \sum_{i < j} \frac{G m_i m_j}{r_{m(i,j)}} \alpha(i, j; m)^n \sum_{u_1, \dots, u_n \in \bar{B}} c_n \beta_n(\mathbf{r}_{u_1}, \dots, \mathbf{r}_{u_n}, \mathbf{r}_m) B_{ij u_1} \dots B_{ij u_n}, \quad (6.112)$$

where we recall that $m = m(i, j)$ denotes the lowest level binary in $B_p(i, j)$. Using equation (6.66), we rewrite the summation over pairs of bodies to summations over binaries and children within those binaries, giving

$$S'_n = (-1)^{n+1} \sum_{k \in B} \sum_{i \in \{k.C1\}} \sum_{j \in \{k.C2\}} \frac{G m_i m_j}{r_k} \alpha(i, j; k)^n \sum_{u_1, \dots, u_n \in \bar{B}} c_n \beta_n(\mathbf{r}_{u_1}, \dots, \mathbf{r}_{u_n}, \mathbf{r}_k) B_{ij u_1} \dots B_{ij u_n}. \quad (6.113)$$

By definition, none of the binaries u_l in the summations $u_1, \dots, u_n \in \bar{B}$ are the same as binary k . Furthermore, because $i \in \{k.C1\}$ and $j \in \{k.C2\}$, if u_l is not part of the children of binary k (i.e. if $u_l \notin \{k.C\}$), then $B_{ij u_l} = 0$. Therefore, the summations $u_1, \dots, u_n \in \bar{B}$ can be written as summations $u_1, \dots, u_n \in B$ with $u_l \in \{k.C\}$, i.e.

$$\begin{aligned} S'_n &= (-1)^{n+1} \sum_{k \in B} \sum_{i \in \{k.C1\}} \sum_{j \in \{k.C2\}} \frac{G m_i m_j}{r_k} \alpha(i, j; k)^n \sum_{\substack{u_1, \dots, u_n \in B \\ u_1, \dots, u_n \in \{k.C\}}} c_n \beta_n(\mathbf{r}_{u_1}, \dots, \mathbf{r}_{u_n}, \mathbf{r}_k) B_{ij u_1} \dots B_{ij u_n} \\ &= (-1)^{n+1} \sum_{k \in B} \frac{G}{r_k} \sum_{\substack{u_1, \dots, u_n \in B \\ u_1, \dots, u_n \in \{k.C\}}} c_n \beta_n(\mathbf{r}_{u_1}, \dots, \mathbf{r}_{u_n}, \mathbf{r}_k) \sum_{i \in \{k.C1\}} \sum_{j \in \{k.C2\}} \alpha(i, j; k)^n m_i m_j B_{ij u_1} \dots B_{ij u_n}, \end{aligned} \quad (6.114)$$

where in the second line, we used that $u_1, \dots, u_n$ are (no longer) dependent on i and j .

In equation (6.114), each binary u_l in the summation $u_1, \dots, u_n \in B$ is associated with a factor $B_{ij u_l}$. A given binary u_l may occur zero to multiple times, up to n times. For example, the pairwise binary terms correspond to $u_1 = \dots = u_n$, such that there are two unique binaries: k and the equal u 's. For binary u_l , let the number of recurrences be denoted by d_{u_l} ; generally, $0 \leq d_{u_l} \leq n$. Consequently, the power of $B_{ij u_l}$ is d_{u_l} , i.e. the mass ratio factor in S'_n associated with binary u_l is $B_{ij u_l}^{d_{u_l}}$.

In the summations $u_1, \dots, u_n \in B$, the binaries $u_1, \dots, u_n$ occur in combinations with various divisions over the two children of binary k . For a particular combination, let the binaries in child 1 (2) of binary k be denoted by p_l (q_l). Furthermore, for child 1 (2) of k , let p (q) denote the binary with the highest level within that child (see below for the case when the highest level within a child of binary k is shared amongst two or more binaries). The number of recurrences associated with binaries p and q are d_p and d_q , respectively. Note that if $d_p = 0$ ($d_q = 0$), then this implies that there are *no* binaries in child 1 (2) of binary k .

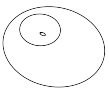
Furthermore, let n_1 (n_2) denote the combined exponent of all mass ratio quantities $B_{ij u_i}$ associated with child 1 (2) of binary k . The allowed values of n_1 and n_2 are $0 \leq n_1 \leq n$ and $0 \leq n_2 \leq n$, respectively, and because the combined power of all mass ratio factors is n , $n_1 + n_2 = n$. For a given nonzero exponent d_p (d_q) of the mass ratio factor associated with binary p (q), there are $n_1 - d_p$ ($n_2 - d_q$) binaries in child 1 (2) distinct from p (q), such that the combined mass ratio exponent in child 1 (2) is n_1 (n_2). Note that these $n_1 - d_p$ ($n_2 - d_q$) binaries are not necessarily distinct; each is associated with a recurrence number d_{p_i} (d_{q_i}). The maximum recurrence number for binary p (q) is n_1 (n_2).

Subsequently, we rewrite the summations $u_1, \dots, u_n \in B$ as separate summations over the binaries $p, p_1, \dots, p_{n_1-d_p}$ and $q, q_1, \dots, q_{n_2-d_q}$ for all possible combinations of n_1, n_2, d_p and d_q . Note that there should only be summations over p and $p_1, \dots, p_{n_1-d_p}$ (q and $q_1, \dots, q_{n_2-d_q}$) if $d_p > 0$ ($d_q > 0$). Also, if both d_p and d_q are zero, then there are no binaries in either child of binary k with the highest level. In that case, the corresponding terms vanish (see below).

When rewriting the summations $u_1, \dots, u_n \in B$ to summations over both children of binary k , it should be taken into account that for given $n_1, n_2, d_p, d_{p_1}, \dots, d_{p_{n_1-d_p}}, d_q$ and $d_{q_1}, \dots, d_{q_{n_2-d_q}}$, there are multiple possible permutations of the binaries $p, p_1, \dots, p_{n_1-d_p}, q, \dots, q_{n_2-d_q}$ from the set $u_1, \dots, u_n \in B$. The number of permutations is determined by n and the number of recurrences of each of the binaries. This is equivalent to placing numbers in n slots, where there are d_p identical numbers, d_{p_1} other identical numbers, d_{p_2} other identical numbers, etc. This gives a number of permutations

$$\mathcal{P}_{d_p, d_{p_1}, \dots, d_{p_{n_1-d_p}}, d_q, d_{q_1}, \dots, d_{q_{n_2-d_q}}}^{n_1, n_2} \equiv \mathcal{P}_{d_p; d_q}^{n_1, n_2} = \frac{(n_1 + n_2)!}{(d_p)!(d_{p_1})! \dots (d_{p_{n_1-d_p}})!(d_q)!(d_{q_1})! \dots (d_{q_{n_2-d_q}})!}, \quad (6.115)$$

where $\mathcal{P}_{d_p, d_{p_1}, \dots, d_{p_{n_1-d_p}}, d_q, d_{q_1}, \dots, d_{q_{n_2-d_q}}}^{n_1, n_2}$ is short-hand notation, and $n_2 = n - n_1$. Note that the quantity $\beta_n(\mathbf{r}_{u_1}, \dots, \mathbf{r}_{u_n}, \mathbf{r}_k)$ is invariant under each permutation because it does not depend on the order of any of the first n arguments (cf. Appendix 6.A.2). Evidently, the same applies to the scalar multiplication $B_{ij u_1} \dots B_{ij u_n}$ in equation (6.114).



We thus arrive at the following rewriting of summations,

$$\begin{aligned}
 & \sum_{\substack{u_1, \dots, u_p \in B \\ u_1, \dots, u_n \in \{k.C\}}} c_n \beta_n(\mathbf{r}_{u_1}, \dots, \mathbf{r}_{u_n}, \mathbf{r}_k) \sum_{i \in \{k.C1\}} \sum_{j \in \{k.C2\}} \alpha(i, j; k)^n m_i m_j B_{ij u_1} \dots B_{ij u_n} \\
 &= \sum_{n_1=0}^n \sum_{d_p=0}^{n_1} \sum_{n_2=n-n_1}^{d_q=0} \sum_{\substack{\text{if } d_p > 0: \\ p, p_1, \dots, p_{n_1-d_p} \in B \\ p, p_1 \in \{k.C1\} \\ p \in \{p_l.C\}}} \sum_{\substack{\text{if } d_q > 0: \\ q, q_1, \dots, q_{n_2-d_q} \in B \\ q, q_1 \in \{k.C2\} \\ q \in \{q_l.C\}}} \mathcal{P}_{d_p, d_q}^{n_1, n_2}(\underbrace{\mathbf{r}_p, \dots, \mathbf{r}_{p_1}, \dots, \mathbf{r}_{p_{n_1-d_p}}}_{d_p \times}, \underbrace{\mathbf{r}_q, \dots, \mathbf{r}_{q_1}, \dots, \mathbf{r}_{q_{n_2-d_q}}}_{d_q \times}, \mathbf{r}_k) \\
 & \quad \times f(n_1, n_2, d_p, d_q) \sum_{i \in \{k.C1\}} \sum_{j \in \{k.C2\}} \alpha(i, j; k)^n m_i m_j \left(B_{ij p_1}^{d_p} B_{ij p_1}^{d_{p_1}} \dots B_{ij p_{n_1-d_p}}^{d_{n_1-d_p}} \right) \left(B_{ij q_1}^{d_q} B_{ij q_1}^{d_{q_1}} \dots B_{ij q_{n_2-d_q}}^{d_{n_2-d_q}} \right). \quad (6.116)
 \end{aligned}$$

Note that if $d_p = 0$ ($d_q = 0$), then also the associated recurrences $d_{p_1}, \dots, d_{n_1-d_p}$ ($d_{q_1}, \dots, d_{n_2-d_q}$) are zero, and the corresponding product $B_{ij p_1}^{d_{p_1}} \dots B_{ij p_{n_1-d_p}}^{d_{n_1-d_p}}$ ($B_{ij q_1}^{d_{q_1}} \dots B_{ij q_{n_2-d_q}}^{d_{n_2-d_q}}$) is unity. The factor $f(n_1, n_2, d_p, d_q)$ takes into account that the terms corresponding to $d_p = d_q = 0$ are zero (see below), and that terms with $d_p = 0$ and $n_1 - d_p > 0$, and $d_q = 0$ and $n_2 - d_q > 0$, should be discarded (by definition of the p and q binaries). Therefore,

$$f(n_1, n_2, d_p, d_q) = \begin{cases} 0, & d_p = 0 \& d_q = 0; \\ 0, & d_p = 0 \& n_1 - d_p > 0; \\ 0, & d_q = 0 \& n_2 - d_q > 0; \\ 1, & \text{otherwise.} \end{cases} \quad (6.117)$$

Note that the use of $f(n_1, n_2, d_p, d_q)$ can be avoided by separating out the terms corresponding to $d_p = 0$ and/or $d_q = 0$. We nevertheless include it, in order to keep the notation relatively compact.

When summing over body $i \in \{k.C1\}$, $B_{ij p_l} = 0$ for any p_l within the binaries $p_1, \dots, p_{n_1-d_p}$, unless $i \in \{p.C\}$ and $p \in \{p_l.C\}$. Similarly, $B_{ij q_l} = 0$ for any q_l within the binaries $q_1, \dots, q_{n_2-d_q}$, unless $j \in \{q.C\}$ and $q \in \{q_l.C\}$. Furthermore, any of the mass ratio quantities $B_{ij p_1}, \dots, B_{ij p_{n_1-d_p}}$, which are nonzero if $p \in \{p_l.C\}$, are constant when summing over $i \in \{p.C\}$ because p is the highest-level binary in child 1 of binary k . Also, for $i \in \{p.C\}$, the quantities $B_{ij q_1}, \dots, B_{ij q_{n_2-d_q}}$ are constant because the ' q '-binaries are, by definition, part of child 2 of binary k , whereas $p \in \{k.C1\}$. Similarly, $B_{ij q_1}, \dots, B_{ij q_{n_2-d_q}}$, $B_{ij p}$ and $B_{ij p_1}, \dots, B_{ij p_{n_1-d_p}}$ are constant when summing over body $j \in \{q.C\}$. Note that when summing $i \in \{p.C\}$ and $j \in \{q.C\}$, $\alpha(i, j; k) = \alpha(p, q; k)$.

Therefore, the summations over bodies i and j in equation (6.116) can be written as

$$\begin{aligned}
 & \sum_{i \in \{k, C1\}} \sum_{j \in \{k, C2\}} \alpha(i, j; k)^n m_i m_j \left(B_{ijp}^{d_p} B_{ijp_1}^{d_{p_1}} \dots B_{ijp_{n_1-d_p}}^{d_{n_1-d_p}} \right) \left(B_{ijq}^{d_q} B_{ijq_1}^{d_{q_1}} \dots B_{ijq_{n_2-d_q}}^{d_{n_2-d_q}} \right) \\
 &= \sum_{i \in \{p, C1\}} \sum_{j \in \{q, C\}} \alpha(p, q; k)^n m_i m_j \left(B_{ijp}^{d_p} B_{ijp_1}^{d_{p_1}} \dots B_{ijp_{n_1-d_p}}^{d_{n_1-d_p}} \right) \left(B_{ijq}^{d_q} B_{ijq_1}^{d_{q_1}} \dots B_{ijq_{n_2-d_q}}^{d_{n_2-d_q}} \right) \\
 &= \alpha(p, q; k)^n \left(B_{ijp_1}^{d_{p_1}} \dots B_{ijp_{n_1-d_p}}^{d_{n_1-d_p}} \right) \left(B_{ijq_1}^{d_{q_1}} \dots B_{ijq_{n_2-d_q}}^{d_{n_2-d_q}} \right) \left(\sum_{i \in \{p, C1\}} m_i B_{ijp}^{d_p} \right) \left(\sum_{j \in \{q, C\}} m_j B_{ijq}^{d_q} \right). \tag{6.118}
 \end{aligned}$$

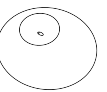
Note that the mass ratio factors outside of the summations of i and j are in fact independent of i and j . One could e.g. write $B_{ijp_1} = B_{ppp_1}$.

Similarly to the previous appendices, the expressions in equation (6.118) with remaining summations over bodies can be simplified by summing separately over the two children of p and q , and by substituting the explicit expression for the mass ratio quantities. For binary p ,

$$\begin{aligned}
 \sum_{i \in \{p, C\}} m_i B_{ijp}^{d_p} &= \left(\sum_{i \in \{p, C1\}} + \sum_{i \in \{p, C2\}} \right) m_i B_{ijp}^{d_p} \\
 &= \left(\frac{\alpha(p, C1, k, CS(p); p) M_{p, C2}^{d_p}}{M_p} \sum_{i \in \{p, C1\}} m_i + \left(\frac{\alpha(p, C2, k, CS(p); p) M_{p, C1}^{d_p}}{M_p} \right) \sum_{i \in \{p, C2\}} m_i \right) \\
 &= \frac{[\alpha(p, C1, qk, CS(p); p)]^{d_p} M_{p, C2}^{d_p} M_{p, C1} + [\alpha(p, C2, k, CS(p); p)]^{d_p} M_{p, C1}^{d_p} M_{p, C2}}{M_p^{d_p}} \\
 &= \alpha(p, C1, k, CS(p); p)^{d_p} \frac{M_{p, C1} M_{p, C2} M_{p, C2}^{d_p-1} + (-1)^{d_p} M_{p, C1}^{d_p-1}}{M_p} \\
 &\equiv \mathcal{M}_{p, k, CS(p)}^{(d_p)}, \tag{6.119}
 \end{aligned}$$

where, for convenience, we introduced the short-hand notation $\mathcal{M}_{p, k, CS(p)}^{(d_p)}$. Analogously,

$$\sum_{j \in \{q, C\}} m_j B_{ijq}^{d_q} = \mathcal{M}_{q, k, CS(q)}^{(d_q)}. \tag{6.120}$$



Note that equation (6.119) is essentially a generalized version of the quantity ‘ M_j ’ that appears at order j in the hierarchical three-body problem as given by equation (1) of Harrington (1968) – in the latter paper, M_j is multiplied by the tertiary mass ‘ m_2 ’, in our case, this corresponds to $\mathcal{M}_{q,k,\text{CS}(q)}^{(d_q)}$. Special values of $\mathcal{M}_{p,k,\text{CS}(p)}^{(d_p)}$ of interest here are

$$\mathcal{M}_{p,k,\text{CS}(p)}^{(0)} = \alpha(p, \text{C1}, k, \text{CS}(p); p)^0 \frac{M_{p,\text{C1}} M_{p,\text{C2}}}{M_p} \frac{M_{p,\text{C2}}^{-1} + M_{p,\text{C1}}^{-1}}{M_p} = M_{p,\text{C1}} + M_{p,\text{C2}} = M_p, \quad (6.121)$$

and

$$\mathcal{M}_{p,k,\text{CS}(p)}^{(1)} = \alpha(p, \text{C1}, k, \text{CS}(p); p) \frac{M_{p,\text{C1}} M_{p,\text{C2}}}{M_p} \frac{M_{p,\text{C2}}^0 - M_{p,\text{C1}}^0}{M_p} = 0. \quad (6.122)$$

Substituting explicit expressions for the remaining mass ratio quantities $B_{ijp_1}, \dots, B_{ijp_{n_1-d_p}}$ and $B_{ijq_1}, \dots, B_{ijq_{n_2-d_q}}$, we find the following general expression for the potential energy,

$$\begin{aligned} S'_n = & (-1)^{n+1} \sum_{k \in \mathbf{B}} \frac{G}{r^k} \sum_{n_1=0}^n \sum_{d_p=0}^{n_1} \sum_{n_2=n-n_1}^{n_1} \sum_{d_q=0}^{n_2} \sum_{\substack{p, p_1, \dots, p_{n_1-d_p} \in \mathbf{B} \\ p, p_1 \in \{k, \text{C1}\} \\ p \in \{p_1, \text{C}\}}} \sum_{\substack{\text{if } d_p > 0: \\ q, q_1 \in \{k, \text{C2}\} \\ q \in \{q_1, \text{C}\}}} \sum_{\substack{\text{if } d_q > 0: \\ q, q_1 \in \{k, \text{C2}\} \\ q \in \{q_1, \text{C}\}}} c_{n_k} \beta_n \left(\underbrace{\mathbf{r}_p, \dots, \mathbf{r}_p}_{d_p \times} \mathbf{r}_{p_1}, \dots, \mathbf{r}_{p_{n_1-d_p}}, \underbrace{\mathbf{r}_q, \dots, \mathbf{r}_q}_{d_q \times} \mathbf{r}_{q_1}, \dots, \mathbf{r}_{q_{n_2-d_q}}, \mathbf{r}_k \right) \\ & \times \alpha(p, q; k)^n f(n_1, n_2, d_p, d_q) \mathcal{P}_{d_p; d_q}^{n_1, n_2} \left\{ \prod_{l=1}^{n_1-d_p} \left[\frac{\alpha(p_l, k, \text{CS}(p); k) M_{p_l, \text{CS}(p)}}{M_{p_l}} \right]^{d_{p_l}} \prod_{l=1}^{n_2-d_q} \left[\frac{\alpha(q_l, k, \text{CS}(q); k) M_{q_l, \text{CS}(q)}}{M_{q_l}} \right]^{d_{q_l}} \right\} \\ & \times \mathcal{M}_{p,k,\text{CS}(p)}^{(d_p)} \mathcal{M}_{q,k,\text{CS}(q)}^{(d_q)}, \end{aligned} \quad (6.123)$$

where $f(n_1, n_2, d_p, d_q)$ was defined in equation (6.117). Note that the products in curly brackets are to be evaluated to unity if the upper limit of the variable l , i.e. $n_1 - d_p$ or $n_2 - d_q$, is zero (this corresponds to no binaries p_l or q_l such that $p \in \{p_l, \text{C}\}$ or $q \in \{q_l, \text{C}\}$).

Each individual term in equation (6.123) depends on binary k and, depending on the values of d_p and d_q , binaries p and at most $n_1 - d_p$ distinct binaries $p_1, \dots, p_{n_1-d_p}$, and binaries q and at most $n_2 - d_q$ distinct binaries $q_1, \dots, q_{n_2-d_q}$. If $d_p \neq 0$ and $d_q \neq 0$, then the number of distinct binaries is at most $N_{\text{bin}, \text{dis}} = 1 + (1 + n_1 - d_p) + (1 + n_2 - d_q) = 3 + n - d_p - d_q$. This number is maximized when d_p and d_q are smallest. Because the product $\mathcal{M}_{p,k,\text{CS}(p)}^{(d_p)} \mathcal{M}_{q,k,\text{CS}(q)}^{(d_q)}$ is zero whenever one or both of d_p and d_q are unity (cf. equation 6.122), the lowest allowed values of $d_p > 0$ and $d_q > 0$ are $d_p = 2$ and $d_q = 2$, which corresponds to $N_{\text{bin}, \text{dis}} = 3 + n - 4 = n - 1$.

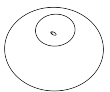
However, the case that either d_p or d_q are zero should also be considered. If e.g. $d_q = 0$ and $d_p \neq 0$, then the occurring binaries are k, p and at most $n_1 - d_p$ distinct binaries $p_1, \dots, p_{n_1 - d_p}$. The corresponding number of distinct binaries is $N_{\text{bin,dis}} = 1 + (1 + n_1 - d_p) = 2 + n_1 - d_p$. The maximum $N_{\text{bin,dis}}$ occurs when $n_1 = n$ and $d_p = 2$ (again, terms corresponding to $d_p = 1$ vanish due to equation 6.122), giving $N_{\text{bin,dis}} = 2 + n - 2 = n$. Evidently, the same result applies when choosing $d_p = 0$ and $d_q \neq 0$.

We conclude that the individual terms in the Hamiltonian depend on at most n different binaries. Because ‘singlets’ (terms depending on just one binary) are not allowed (in addition to binary k , there will always be another binary in either p or q families), the maximum number of types of terms (i.e. binary pairs, triplets, quadruplets, etc.) is $n - 1$.

More than one binary with the highest level in either child of binary k . As mentioned above, it is possible that in either child of binary k , there is no unique binary with the highest level. In this case, let $v_{i1}, \dots, v_s$ denote the s binaries with the common highest level in either of the two children of binary k . When summing over $i \in \{k.C\}$, $B_{ijv_i} = 0$ for binary v_i unless $i \in \{v_i.C\}$. However, when summing $i \in \{v_i.C\}$, the other $B_{ijv_i}, \dots, B_{ijv_s}$ (excluding B_{ijv_i}) are zero, because, by definition, the v -binaries are on the same level. Therefore, if there is more than one binary with the highest level in either child of binary k , then the corresponding terms vanish. This is taken into equation (6.123) with the factor $f(n_1, n_2, d_p, d_q)$ (cf. equation 6.117).

Pairwise binary terms. Equation (6.123) can be made more explicit for the pairwise binary terms. In that equation, the pairwise binary terms correspond to (1) $n_1 = n$ ($n_2 = 0$), $d_p = n_1 = n$ and $d_q = 0$, and (2) $n_1 = 0$ ($n_2 = n$), $d_q = n_2 = n$ and $d_p = 0$. In these cases, $n_1 - d_p = 0$ and $n_2 - d_q = 0$, and there are no summations of the binaries $p_1, \dots, p_{n_1 - d_p}$ and $q_1, \dots, q_{n_2 - d_q}$. Also, in case (1), there is no summation over the q -binaries; in case (2), there is no summation over the p -binaries. Furthermore, the number of permutations is $\mathcal{P}_{d_p, n_1}^{n, n_1} = n!/n! = 1$, and the products in the curly brackets evaluate to unity. The resulting expression is

$$\begin{aligned}
 S'_{n,2} &= (-1)^{n+1} \sum_{k \in B} \frac{G}{r_k} \left[\sum_{\substack{p \in B \\ p \in \{k.C\}}} c_n \beta_n (\underbrace{\mathbf{r}_p, \dots, \mathbf{r}_p}_{n \times}) \mathcal{M}_{p;k,CS(p)}^{(n)} \mathcal{M}_{k,CS(p);p}^{(0)} + \sum_{\substack{q \in B \\ q \in \{k.C2\}}} c_n \beta_n (\underbrace{\mathbf{r}_q, \dots, \mathbf{r}_q}_{n \times}) \mathcal{M}_{q;k,CS(q)}^{(n)} \mathcal{M}_{k,CS(q);q}^{(0)} \right] \\
 &= \sum_{k \in B} \frac{G}{r_k} \sum_{\substack{p \in B \\ p \in \{k.C\}}} c_n \beta_n (\underbrace{\mathbf{r}_p, \dots, \mathbf{r}_p}_{n \times}) \mathcal{M}_{p;k,CS(p)}^{(n)} \mathcal{M}_{k,CS(p);p}^{(0)}
 \end{aligned} \tag{6.124}$$



where in the second line, we changed the variable q in the second term to p , and combined the two separate summations over child 1 and child 2 of k into one summation over all children. Note that $\mathcal{M}_{k,CS(p);p}^{(0)}$ simply evaluates to $M_{k,CS(p)}$ (cf. equation 6.121). This gives

$$\begin{aligned}
 S'_{n;2} &= (-1)^{n+1} \sum_{k \in B} \sum_{\substack{p \in B \\ p \in \{k, C\}}} \frac{M_{p,C1} M_{p,C2}}{M_p} \alpha(p, C1, k, C2; p)^n \frac{M_p^{n-1} + (-1)^n M_{p,C1}^{n-1}}{M_p^{n-1}} \frac{GM_{k,CS(p)}}{r_k} c_n \beta_n \underbrace{(\mathbf{r}_p, \dots, \mathbf{r}_p, \mathbf{r}_k)}_{n \times} \\
 &= (-1)^{n+1} \sum_{k \in B} \sum_{\substack{p \in B \\ p \in \{k, C\}}} \frac{M_{p,C1} M_{p,C2}}{M_p} \alpha(p, C1, k, C2; p)^n \frac{M_p^{n-1} + (-1)^n M_{p,C1}^{n-1}}{M_p^{n-1}} \frac{GM_{k,CS(p)}}{r_k} \sum_{m=0}^n \frac{(\mathbf{r}_p \cdot \mathbf{r}_k)^m r_p^{n-m}}{r_k^{n+m}} \mathcal{A}_m^{(n)}. \quad (6.125)
 \end{aligned}$$

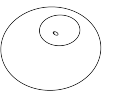
In the second line, we used equation (6.56) (recall that the $\mathcal{A}_m^{(n)}$ are integer ratio coefficients, which are the same coefficients appearing in the n th Legendre polynomials, cf. equations 6.56 and 6.57).

Demonstration of equation (6.123). As a demonstration, we apply equation (6.123) to the case $n = 3$ (octupole order), and show that the resulting expression is the same as the expression that was previously derived in Appendix 6.A.4. For $n = 3$, the combinations of n_1 , n_2 , d_p and d_q are enumerated in Table 6.4. For each combination, we note whether or not it occurs in the summation, either because of the properties of the quantity $f(n_1, n_2, d_p, d_q)$ (cf. equation 6.117), or because of property equation (6.122). For the triplet terms, note that

$\mathcal{P}_{d_p; d_q}^{n_1, n_2} = 3!/2! = 3$. This gives

$$\begin{aligned}
 S'_3 = & \sum_{k \in \mathbb{B}} \sum_{\substack{q \in \mathbb{B} \\ q \in \{k, C2\}}} \frac{G}{r_k} c_3 \beta_3(\mathbf{r}_q, \mathbf{r}_q, \mathbf{r}_q; \mathbf{r}_k) \times 1 \times \alpha(q, k, \text{CS}(q); k) M_{k, \text{CS}(q)} \frac{M_{q, C1} M_{q, C2}}{M_q} \alpha(q, C1, k, \text{CS}(q); q) \frac{M_{q, C2} - M_{q, C1}}{M_q} \\
 & + \sum_{k \in \mathbb{B}} \sum_{\substack{p \in \mathbb{B} \\ p \in \{k, C1\}}} \frac{G}{r_k} c_3 \beta_3(\mathbf{r}_p, \mathbf{r}_p, \mathbf{r}_p; \mathbf{r}_k) \times 1 \times \alpha(p, k, \text{CS}(p); k) M_{k, \text{CS}(p)} \frac{M_{p, C1} M_{p, C2}}{M_p} \alpha(p, C1, k, \text{CS}(p); p) \frac{M_{p, C2} - M_{p, C1}}{M_p} \\
 & + \sum_{k \in \mathbb{B}} \sum_{\substack{q, q_1 \in \mathbb{B} \\ q, q_1 \in \{k, C2\} \\ q \in \{q_1, C\}}} \frac{G}{r_k} c_3 \beta_3(\mathbf{r}_q, \mathbf{r}_q, \mathbf{r}_{q_1}; \mathbf{r}_k) \times 3 \times \frac{\alpha(q, k, \text{CS}(q); k) M_{q_1, \text{CS}(q)}}{M_{q_1}} \alpha(q, k, \text{CS}(q); k) M_{k, \text{CS}(q)} \frac{M_{q, C1} M_{q, C2}}{M_q} \\
 & + \sum_{k \in \mathbb{B}} \sum_{\substack{p, p_1 \in \mathbb{B} \\ p, p_1 \in \{k, C2\} \\ p \in \{p_1, C\}}} \frac{G}{r_k} c_3 \beta_3(\mathbf{r}_p, \mathbf{r}_p, \mathbf{r}_{p_1}; \mathbf{r}_k) \times 3 \times \frac{\alpha(p, k, \text{CS}(p); k) M_{p_1, \text{CS}(p)}}{M_{p_1}} \alpha(p, k, \text{CS}(p); k) M_{k, \text{CS}(p)} \frac{M_{p, C1} M_{p, C2}}{M_p}. \quad (6.126)
 \end{aligned}$$

Taking together the summations over p and q into a single summation of k for the binary pair and triplet terms, relabelling $q_1 = u$ and $p_1 = u$ and substituting the expressions for $c_3 \beta_3$, we recover equation (6.102).



n_1	n_2	d_p	d_q	$n_1 - d_p$	$n_2 - d_q$	Notes
0	3	0	0	0	3	Not included: $d_p = 0$ and $d_q = 0$ ($f = 0$).
0	3	0	1	0	2	Not included: $\mathcal{M}_{p;k,CS(p)}^{(d_p)} \mathcal{M}_{q;k,CS(q)}^{(d_q)} = 0$.
0	3	0	2	0	1	Part of triplet terms (child 2 of k).
0	3	0	3	0	0	Part of pairwise terms (child 2 of k).
1	2	0	0	1	2	Not included: $d_p = 0$ and $d_q = 0$ ($f = 0$).
1	2	0	1	1	1	Not included: $d_p = 0$ and $n_1 - d_p > 0$ ($f = 0$).
1	2	0	2	1	0	Not included: $d_p = 0$ and $n_1 - d_p > 0$ ($f = 0$).
1	2	1	0	0	2	Not included: $d_q = 0$ and $n_2 - d_q > 0$ ($f = 0$).
1	2	1	1	0	1	Not included: $\mathcal{M}_{p;k,CS(p)}^{(d_p)} \mathcal{M}_{q;k,CS(q)}^{(d_q)} = 0$.
1	2	1	2	0	0	Not included: $\mathcal{M}_{p;k,CS(p)}^{(d_p)} \mathcal{M}_{q;k,CS(q)}^{(d_q)} = 0$.
2	1	0	0	2	1	Not included: $d_p = 0$ and $d_q = 0$ ($f = 0$).
2	1	0	1	2	0	Not included: $d_p = 0$ and $n_1 - d_p > 0$ ($f = 0$).
2	1	1	0	1	1	Not included: $\mathcal{M}_{p;k,CS(p)}^{(d_p)} \mathcal{M}_{q;k,CS(q)}^{(d_q)} = 0$.
2	1	1	1	1	0	Not included: $\mathcal{M}_{p;k,CS(p)}^{(d_p)} \mathcal{M}_{q;k,CS(q)}^{(d_q)} = 0$.
2	1	2	0	0	1	Not included: $d_q = 0$ and $n_2 - d_q > 0$ ($f = 0$).
2	1	2	1	0	0	Not included: $\mathcal{M}_{p;k,CS(p)}^{(d_p)} \mathcal{M}_{q;k,CS(q)}^{(d_q)} = 0$.
3	0	0	0	3	0	Not included: $d_p = 0$ and $d_q = 0$ ($f = 0$).
3	0	1	0	2	0	Not included: $\mathcal{M}_{p;k,CS(p)}^{(d_p)} \mathcal{M}_{q;k,CS(q)}^{(d_q)} = 0$.
3	0	2	0	1	0	Part of triplet terms (child 1 of k).
3	0	3	0	0	0	Part of pairwise terms (child 1 of k).

Table 6.4: Combinations of n_1 , n_2 , d_p and d_q occurring in equation (6.123) for $n = 3$. For each combination, we note whether or not it is included in the summations.

Summary

In the previous appendices, we explicitly carried out the rewriting procedure for all terms up and including octupole order, and including the binary pair terms at hexadecupole and dotriacontupole order. We also derived a general, more implicit expression valid for any order n and including all binary interactions (cf. Appendix 6.A.4). For the purposes of giving a useful overview, we repeat the explicit expressions here. The simplified explicit expression reads

$$H = \frac{1}{2} \sum_{k \in B} \left[\frac{M_{k,C1} M_{k,C2}}{M_k} (\hat{\mathbf{r}}_k \cdot \hat{\mathbf{r}}_k) - \frac{G M_{k,C1} M_{k,C2}}{r_k} \right] + S'_2 + S'_3 + S'_4 + S'_5 + \dots \quad (6.127)$$

where

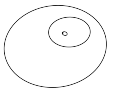
$$S'_2 = -\frac{1}{2} \sum_{k \in B} \sum_{\substack{p \in B \\ p \in \{k,C\}}} \frac{M_{p,C1} M_{p,C2}}{M_p} \frac{G M_{k,CS(p)}}{r_k} \left(\frac{r_p}{r_k} \right)^2 \left[3 (\hat{\mathbf{r}}_p \cdot \hat{\mathbf{r}}_k)^2 - 1 \right]; \quad (6.128)$$

$$S'_3 = \frac{1}{2} \sum_{k \in B} \sum_{\substack{p \in B \\ p \in \{k,C\}}} \frac{M_{p,C1} M_{p,C2}}{M_p} \alpha(p, C1, k, CS(p); p) \frac{M_{p,C2} - M_{p,C1}}{M_p} \alpha(p, k, CS(p); k) \frac{G M_{k,CS(p)}}{r_k} \left(\frac{r_p}{r_k} \right)^3 \left[5 (\hat{\mathbf{r}}_p \cdot \hat{\mathbf{r}}_k)^3 - 3 (\hat{\mathbf{r}}_p \cdot \hat{\mathbf{r}}_k) \right] \quad (6.129)$$

$$+ \frac{3}{2} \sum_{k \in B} \sum_{\substack{p \in B \\ p \in \{k,C\}}} \sum_{\substack{u \in B \\ u \in \{k,C\}}} \frac{M_{p,C1} M_{p,C2}}{M_p} \alpha(p, k, CS(p); k) \frac{G M_{k,CS(p)}}{r_k} \frac{\alpha(p, k, CS(p); u) M_{u,CS(p)}}{M_u} \left(\frac{r_p}{r_k} \right)^2 \left(\frac{r_u}{r_k} \right) \quad (6.130)$$

$$\times \left[5 (\hat{\mathbf{r}}_p \cdot \hat{\mathbf{r}}_k)^2 (\hat{\mathbf{r}}_u \cdot \hat{\mathbf{r}}_k) - 2 (\hat{\mathbf{r}}_p \cdot \hat{\mathbf{r}}_k) (\hat{\mathbf{r}}_p \cdot \hat{\mathbf{r}}_u) - (\hat{\mathbf{r}}_u \cdot \hat{\mathbf{r}}_k) \right];$$

$$S'_4 = -\frac{1}{8} \sum_{k \in B} \sum_{\substack{p \in B \\ p \in \{k,C\}}} \frac{M_{p,C1} M_{p,C2}}{M_p} \frac{M_{p,C1}^2}{M_p} \frac{M_{p,C2} - M_{p,C1}}{M_p} \alpha(p, C1, k, CS(p); p) \frac{G M_{k,CS(p)}}{r_k} \left(\frac{r_p}{r_k} \right)^4 \left[35 (\hat{\mathbf{r}}_p \cdot \hat{\mathbf{r}}_k)^4 - 30 (\hat{\mathbf{r}}_p \cdot \hat{\mathbf{r}}_k)^2 + 3 \right] + S'_{4,3} + S'_{4,4}; \quad (6.131)$$



$$\begin{aligned}
S'_5 = & \frac{1}{8} \sum_{k \in \mathbf{B}} \sum_{\substack{p \in \mathbf{B} \\ p \in \{k, \mathbf{C}\}}} \frac{M_{p, \mathbf{C}1} M_{p, \mathbf{C}2}}{M_p} \alpha(p, \mathbf{C}1, p, \mathbf{C}S(p); p) \frac{M_{p, \mathbf{C}2} - M_{p, \mathbf{C}1}}{M_p} \frac{M_{p, \mathbf{C}2}^2}{M_p^2} \alpha(p, k, \mathbf{C}S(p); k) \frac{GM_{k, \mathbf{C}S(p)}}{r_k} \left(\frac{r_p}{r_k} \right)^5 \\
& \times \left[63 (\hat{\mathbf{r}}_p \cdot \hat{\mathbf{r}}_k)^5 - 70 (\hat{\mathbf{r}}_p \cdot \hat{\mathbf{r}}_k)^3 + 15 (\hat{\mathbf{r}}_p \cdot \hat{\mathbf{r}}_k) \right] + S'_{5,3} + S'_{5,4} + S'_{5,5}.
\end{aligned} \tag{6.132}$$

Here, the dots in equation (6.127) denote higher order terms. Note that in equation (6.127), no approximations are made apart from the expansions in the binary separation ratios. Below, in Appendix 6.A.5, we orbit average equation (6.127).

6.A.5 Orbit averaging

We average the expanded Hamiltonian over the $N - 1$ binary orbits, assuming that in each orbit, the motion on suborbital time-scales can be well approximated by a bound Kepler orbit. We define the orbit-averaged Hamiltonian as

$$\bar{H} \equiv \frac{1}{(2\pi)^{N-1}} \underbrace{\int_0^{2\pi} \cdots \int_0^{2\pi}}_{N-1} H \prod_{k \in \mathbf{B}} dl_k, \tag{6.133}$$

where l_k is the mean anomaly of orbit k . We express the angular momenta and orientations of each binary in terms of the triad of perpendicular orbital state vectors $(\mathbf{j}_k, \mathbf{e}_k, \mathbf{q}_k)$, where $\mathbf{q}_k \equiv \mathbf{j}_k \times \mathbf{e}_k$. Here, $\mathbf{j}_k$ is the specific angular momentum vector with magnitude $j_k = \sqrt{1 - e_k^2}$, where $e_k, 0 \leq e_k < 1$, is the orbital eccentricity. The specific orbital angular momentum vector $\mathbf{j}_k$ is related to the orbital angular momentum vector $\mathbf{h}_k \equiv (M_{k, \mathbf{C}1} M_{k, \mathbf{C}2} / M_k) \mathbf{r}_k \times \dot{\mathbf{r}}_k$ via $\mathbf{h}_k = \Lambda_k \mathbf{j}_k$, where Λ_k is the angular momentum of a circular orbit,

$$\Lambda_k = M_{k, \mathbf{C}1} M_{k, \mathbf{C}2} \sqrt{\frac{Ga_k}{M_k}}, \tag{6.134}$$

with a_k the semimajor axis. Furthermore, $\mathbf{e}_k$ is the eccentricity or Laplace-Runge-Lenz vector that is aligned with the major axis and that has magnitude e_k .

In terms of $\mathbf{e}_k$ and $\mathbf{q}_k$ and the orbital phase (we use either the true anomaly f_k or the eccentric anomaly ε_k as integration variables), the

dot product of $\mathbf{r}_k$ with some other (constant) vector $\mathbf{u}$ is given by two equivalent expressions,

$$\begin{aligned}\mathbf{r}_k \cdot \mathbf{u} &= \frac{a_k (1 - e_k^2)}{1 + e_k \cos(f_k)} [\cos(f_k) (\hat{\mathbf{e}}_k \cdot \mathbf{u}) + \sin(f_k) (\hat{\mathbf{q}}_k \cdot \mathbf{u})] \\ &= a_k [\cos(\varepsilon_k) - e_k] (\hat{\mathbf{e}}_k \cdot \mathbf{u}) + a_k \sqrt{1 - e_k^2} \sin(\varepsilon_k) (\hat{\mathbf{q}}_k \cdot \mathbf{u}).\end{aligned}\quad (6.135)$$

The simplified Hamiltonian, equation (6.127), is expressed in terms of these dot products. Furthermore, the separations are given by the well-known expressions

$$r_k = \frac{a_k (1 - e_k^2)}{1 + e_k \cos(f_k)} = a_k [1 - e_k \cos(\varepsilon_k)]. \quad (6.136)$$

In equation (6.133), it is convenient to use the true anomaly f_k as the integration variable whenever r_k appears with a high negative power, and, otherwise, the eccentric anomaly ε_k . The relevant Jacobians are given by

$$dl_k = \frac{1}{\sqrt{1 - e_k^2}} \left(\frac{r_k}{a_k} \right)^2 df_k = \left(\frac{r_k}{a_k} \right) d\varepsilon_k. \quad (6.137)$$

In order to express the orbit-averaged quantities in terms of dot products of $\mathbf{e}_k$ and $\mathbf{j}_k$ with those of other orbits (i.e. to remove any reference to the auxiliary vector $\mathbf{q}_k$), we repeatedly use the following vector identity for the scalar product of two scalar triple products,

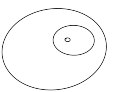
$$(\hat{\mathbf{q}}_k \cdot \mathbf{u}) (\hat{\mathbf{q}}_k \cdot \mathbf{w}) = \left[(\hat{\mathbf{j}}_k \times \hat{\mathbf{e}}_k) \cdot \mathbf{u} \right] \left[(\hat{\mathbf{j}}_k \times \hat{\mathbf{e}}_k) \cdot \mathbf{w} \right] = \mathbf{u} \cdot \mathbf{w} - (\hat{\mathbf{e}}_k \cdot \mathbf{u}) (\hat{\mathbf{e}}_k \cdot \mathbf{w}) - (\hat{\mathbf{j}}_k \cdot \mathbf{u}) (\hat{\mathbf{j}}_k \cdot \mathbf{w}), \quad (6.138)$$

where $\mathbf{u}$ and $\mathbf{w}$ are two arbitrary vectors.

For example, the orbit average of the term $(\mathbf{r}_p \cdot \mathbf{r}_k)^2 / r_k^5$ that appears in the quadrupole-order term of the Hamiltonian (cf. equation 6.127) is given by

$$\left\langle \left\langle \frac{(\mathbf{r}_p \cdot \mathbf{r}_k)^2}{r_k^5} \right\rangle \right\rangle_p = \frac{a_p^2}{2r_k^5} \left[j_{pk}^2 - (\mathbf{j}_p \cdot \mathbf{r}_k)^2 + 5(\mathbf{e}_p \cdot \mathbf{r}_k)^2 \right]; \quad (6.139a)$$

$$\left\langle \left\langle \frac{(\mathbf{r}_p \cdot \mathbf{r}_k)^2}{r_k^5} \right\rangle \right\rangle_{pk} = \frac{a_p^2}{4a_k^3 j_k^5} \left[(1 + 4e_p^2) j_k^2 + (\mathbf{j}_p \cdot \mathbf{j}_k)^2 - 5(\mathbf{e}_p \cdot \mathbf{j}_k)^2 \right]. \quad (6.139b)$$



Note that this expression is well defined in the limit $e_p \rightarrow 0$, in which case $\mathbf{e}_p \rightarrow \mathbf{0}$.

Formally, the averaging of the Hamiltonian over the binary orbits as described by equation (6.133) is not a canonical transformation. However, applying the Von Zeipel transformation technique to the unaveraged Hamiltonian (Brouwer 1959), a canonical transformation can be found that eliminates the short-period terms l_k from the Hamiltonian (cf. appendix A2 of Naoz et al. 2013a; the derivation Hamiltonian there is straightforwardly extended to $N - 1$, rather than two short-period variables). This transformation leads to a transformed Hamiltonian that is equivalent to equation (6.133). The transformed coordinates $\mathbf{e}_k^*$ and $\mathbf{q}_k^*$ differ from the original ones, $\mathbf{e}_k$ and $\mathbf{q}_k$. However, as noted by Naoz et al. (2013a), the untransformed and the transformed coordinates differ by order x^2 . Therefore, this difference can usually be neglected. This is borne out by our tests with other methods (cf. Section 6.3).

Binding energies

If Kepler orbits are assumed, the first two summation terms in equation (6.127) immediately reduce to the sum of the binary binding energies, without having to explicitly average over the orbits. This directly follows from the canonical expressions for the Kepler binary separation (equation 6.136) and the squared velocity,

$$\dot{\mathbf{r}}_k \cdot \dot{\mathbf{r}}_k = \frac{GM_k}{a_k(1 - e_k^2)} [1 + 2e_k \cos(f_k) + e_k^2]. \quad (6.140)$$

Substituting these expressions into the term in equation (6.127) with the single summation of $k \in \mathbf{B}$, we find

$$\begin{aligned} & \sum_{k \in \mathbf{B}} \left[\frac{1}{2} \frac{M_{k,C1} M_{k,C2}}{M_k} (\dot{\mathbf{r}}_k \cdot \dot{\mathbf{r}}_k) - \frac{GM_{k,C1} M_{k,C2}}{r_k} \right] \\ &= \sum_{k \in \mathbf{B}} M_{k,C1} M_{k,C2} \left\{ \frac{1}{2} \frac{GM_k}{M_k a_k (1 - e_k^2)} [1 + 2e_k \cos(f_k) + e_k^2] - \frac{G}{a_k (1 - e_k^2)} [1 + e_k \cos(f_k)] \right\} = \sum_{k \in \mathbf{B}} \frac{GM_{k,C1} M_{k,C2}}{-2a_k}. \end{aligned} \quad (6.141)$$

Each of the individual terms in the latter expression is the Keplerian binary binding energy $E_k < 0$. Because they only depend on the semimajor axes (and not on $\mathbf{e}_k$ and $\mathbf{j}_k$), equation (6.141) does not give rise to secular changes of the orbits (cf. equations 6.17).

Higher order terms

Secular changes arise from the higher-order orbit-averaged terms $\overline{S}_n'$ with $n \geq 2$, which depend on $\mathbf{e}_k$ and $\mathbf{j}_k$. These terms combined form the disturbing function of the system. Given their mathematical equivalence to the corresponding terms in the hierarchical three-body

problem, orbit-averaged expressions for the binary pair terms in the Hamiltonian up and including octupole order can be directly taken from previous studies (e.g. Boué & Fabrycky 2014, where the same vector formalism was adopted). The orbit averaged octupole-order term depending on three binaries and the hexadecupole-order term depending on two binaries were derived previously by Hamers et al. (2015). In the latter work, these terms were expressed in terms of unit vectors $\hat{e}_k$, which are ill-defined in the limit $e_k \rightarrow 0$. Here, we have rederived the orbit average of all terms up and including dotriacontupole order as given explicitly in equation (6.127), where for the hexadecupole and dotriacontupole-order terms, we only included the binary pair terms. We have expressed the result in terms of the non-unit vectors e_k , which are well defined in the limit $e_k \rightarrow 0$, i.e., trivially, $e_k \rightarrow 0$. We have verified the expressions of Hamers et al. (2015) numerically for cases with non-zero eccentricities.

After averaging over the orbits, the Hamiltonian is no longer a function of the phases of the orbits l_k . Consequently, the sign quantities α must be set to unity. A subtlety arises with the binary pair octupole- and dotriacontupole-order terms (or, generally odd order n), which contain the factor

$$\alpha(p, \text{C1}, k, \text{CS}(p); p) (M_{p, \text{C2}} - M_{p, \text{C1}}) = \alpha(p, \text{C2}, k, \text{CS}(p); p) (M_{p, \text{C1}} - M_{p, \text{C2}}). \quad (6.142)$$

After averaging, interchanging the children of p should not change the expression. Therefore, the expression equation (6.142) should be replaced by the absolute mass difference, i.e.

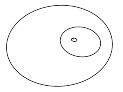
$$\alpha(p, \text{C1}, k, \text{CS}(p); p) (M_{p, \text{C2}} - M_{p, \text{C1}}) \rightarrow |M_{p, \text{C1}} - M_{p, \text{C2}}|. \quad (6.143)$$

The explicit orbit-averaged Hamiltonian, to dotriacontupole order, reads

$$\bar{H} = \sum_{k \in \text{B}} \left(\frac{GM_{k, \text{C1}} M_{k, \text{C2}}}{-2a_k} \right) + \bar{S}'_2 + \bar{S}'_3 + \bar{S}'_4 + \bar{S}'_5 + \dots,$$

where

$$\bar{S}'_2 = \sum_{k \in \text{B}} \sum_{\substack{p \in \text{B} \\ p \in \{k, \text{C}\}}} \frac{M_{p, \text{C1}} M_{p, \text{C2}}}{M_p} \frac{GM_{k, \text{CS}(p)}}{a_k} \left(\frac{a_p}{a_k} \right)^2 \frac{1}{8j_k} \left[(1 - 6e_p^2) j_k^2 + 15 (e_p \cdot j_k)^2 - 3 (j_p \cdot j_k)^2 \right]; \quad (6.144)$$



$$\begin{aligned}
\overline{S}'_3 = & \sum_{k \in B} \sum_{\substack{p \in B \\ p \in \{k, C\}}} \frac{M_{p,C1} M_{p,C2}}{M_p} \frac{|M_{p,C1} - M_{p,C2}|}{M_p} \frac{GM_{k,CS(p)}}{a_k} \left(\frac{a_p}{a_k} \right)^3 \frac{15}{64 j_k^7} \{ 10 (\mathbf{e}_p \cdot \mathbf{j}_k) (\mathbf{e}_k \cdot \mathbf{j}_p) (\mathbf{j}_p \cdot \mathbf{j}_k) \\
& - (\mathbf{e}_p \cdot \mathbf{e}_k) [(1 - 8e_p^2) j_k^2 + 35(\mathbf{e}_p \cdot \mathbf{j}_k)^2 - 5(\mathbf{j}_p \cdot \mathbf{j}_k)^2] \} \\
& + \sum_{k \in B} \sum_{\substack{p \in B \\ p \in \{k, C\}}} \sum_{\substack{u \in B \\ u \in \{k, C\}}} \frac{M_{p,C1} M_{p,C2}}{M_p} \frac{M_{u,CS(p)}}{M_u} \frac{GM_{k,CS(p)}}{a_k} \left(\frac{a_p}{a_k} \right)^2 \left(\frac{a_u}{a_k} \right) \frac{9}{32 j_k^7} \{ 10 (\mathbf{e}_p \cdot \mathbf{e}_u) (\mathbf{e}_p \cdot \mathbf{e}_k) j_k^2 \\
& - 50 (\mathbf{e}_p \cdot \mathbf{e}_k) (\mathbf{e}_p \cdot \mathbf{j}_k) (\mathbf{e}_u \cdot \mathbf{j}_k) - 2 (\mathbf{e}_k \cdot \mathbf{j}_p) (\mathbf{e}_u \cdot \mathbf{j}_p) j_k^2 + 10 (\mathbf{e}_u \cdot \mathbf{j}_k) (\mathbf{e}_k \cdot \mathbf{j}_p) (\mathbf{j}_p \cdot \mathbf{j}_k) \\
& - (\mathbf{e}_u \cdot \mathbf{e}_k) [(1 - 6e_p^2) j_k^2 + 25(\mathbf{e}_p \cdot \mathbf{j}_k)^2 - 5(\mathbf{j}_p \cdot \mathbf{j}_k)^2] \} ;
\end{aligned} \tag{6.145}$$

$$\begin{aligned}
\overline{S}'_4 = & \sum_{k \in B} \sum_{\substack{p \in B \\ p \in \{k, C\}}} \frac{M_{p,C1} M_{p,C2}}{M_p} \frac{M_{p,C1}^2}{M_p} \frac{M_{p,C1} M_{p,C2} + M_{p,C2}^2}{M_p} \frac{GM_{k,CS(p)}}{a_k} \left(\frac{a_p}{a_k} \right)^4 \frac{3}{1024 j_k^{11}} \{ 3 j_k^4 [-6 + e_k^2 + 40e_p^2 (1 + 8e_k^2) \\
& - 20e_p^4 (8 + 15e_k^2)] + 420 (\mathbf{e}_p \cdot \mathbf{e}_k)^2 j_k^2 [- (2 + e_p^2) j_k^2 + 21 (\mathbf{e}_p \cdot \mathbf{j}_k)^2] - 5880 j_k^2 (\mathbf{e}_p \cdot \mathbf{e}_k) (\mathbf{e}_k \cdot \mathbf{j}_p) (\mathbf{e}_p \cdot \mathbf{j}_k) (\mathbf{j}_p \cdot \mathbf{j}_k) \\
& + 5 [28 j_k^2 (1 - 10e_p^2) (4 + 3e_k^2) (\mathbf{e}_p \cdot \mathbf{j}_k)^2 - 6 j_k^2 (8 + 6e_k^2 + e_p^2 (6 + 29e_k^2)) (\mathbf{j}_p \cdot \mathbf{j}_k)^2 - 12 j_k^2 (\mathbf{e}_k \cdot \mathbf{j}_p)^2 ((1 + 13e_p^2) j_k^2 \\
& - 7 (\mathbf{j}_p \cdot \mathbf{j}_k)^2) + 98 j_k^2 (-2 - 3e_k^2 + 4e_p^2 (5 + 3e_k^2)) (\mathbf{e}_p \cdot \mathbf{j}_k)^2 - 441 (2 + e_k^2) (\mathbf{e}_p \cdot \mathbf{j}_k)^4 \\
& + 42 (2 + e_k^2) (j_p^2 j_k + 7 (\mathbf{e}_p \cdot \mathbf{j}_k)^2) (\mathbf{j}_p \cdot \mathbf{j}_k)^2 - 21 (2 + e_k^2) (\mathbf{j}_p \cdot \mathbf{j}_k)^4 \} + \overline{S}'_{4;3} + \overline{S}'_{4;4} ;
\end{aligned} \tag{6.146}$$

$$\begin{aligned}
\overline{S}'_5 = & \sum_{k \in B} \sum_{\substack{p \in B \\ p \in \{k, C\}}} \frac{M_{p,C1} M_{p,C2}}{M_p} \frac{|M_{p,C1} - M_{p,C2}|}{M_p} \frac{M_{p,C2}^2}{M_p^2} \frac{GM_{k,CS(p)}}{a_k} \left(\frac{a_p}{a_k} \right)^5 \frac{105}{4096 j_k^{13}} \left\{ 3024 (\mathbf{e}_p \cdot \mathbf{e}_k)^2 (\mathbf{e}_p \cdot \mathbf{j}_k) \right. \\
& \times (\mathbf{e}_k \cdot \mathbf{j}_p) (\mathbf{j}_p \cdot \mathbf{j}_k) j_k^2 - 28 j_k^2 (\mathbf{e}_p \cdot \mathbf{e}_k)^3 \left[(-26 + 15e_p^2) j_k^2 + 18 (\mathbf{j}_p \cdot \mathbf{j}_k)^2 + 99 (\mathbf{e}_p \cdot \mathbf{j}_k)^2 \right] \\
& - 28 (\mathbf{e}_p \cdot \mathbf{j}_k) (\mathbf{e}_k \cdot \mathbf{j}_p) (\mathbf{j}_p \cdot \mathbf{j}_k) \left[(1 - 4e_p^2) (8 + e_k^2) j_k^2 + 9 (8 + 3e_k^2) (\mathbf{e}_p \cdot \mathbf{j}_k)^2 + 6 j_k^2 (\mathbf{e}_k \cdot \mathbf{j}_p)^2 - 3 (8 + 3e_k^2) (\mathbf{j}_p \cdot \mathbf{j}_k)^2 \right] \\
& - (\mathbf{e}_p \cdot \mathbf{e}_k) \left[\{-8 + e_k^2 + e_p^2 (64 + 748e_k^2) - 4e_p^4 (80 + 179e_k^2)\} j_k^4 \right. \\
& \quad - 693 (8 + 3e_k^2) (\mathbf{e}_p \cdot \mathbf{j}_k)^4 + 42 (\mathbf{e}_p \cdot \mathbf{j}_k)^2 \left\{ (-8 - 19e_k^2 + 6e_p^2 (16 + 5e_k^2)) j_k^2 + 9 (8 + 3e_k^2) (\mathbf{j}_p \cdot \mathbf{j}_k)^2 \right\} \\
& \quad + 14 j_k^2 (8 + e_k^2 - 2e_p^2 (16 + 29e_k^2)) (\mathbf{j}_p \cdot \mathbf{j}_k)^2 - 21 (8 + 3e_k^2) (\mathbf{j}_p \cdot \mathbf{j}_k)^4 \\
& \quad \left. - 28 (\mathbf{e}_k \cdot \mathbf{j}_p)^2 j_k^2 \left((1 + 23e_p^2) j_k^2 - 9 (\mathbf{j}_p \cdot \mathbf{j}_k)^2 \right) \right] \left. \right\} + \overline{S}'_{5;3} + \overline{S}'_{5;4} + \overline{S}'_{5;5}. \tag{6.147}
\end{aligned}$$

General formal expressions for the orbit-averaged binary pair terms at order n

For the binary pair terms, we show how the orbit-averaged integrals at any order can be evaluated using complex integration techniques. We note that Mardling (2013) derived the equivalent orbit-averaged terms for hierarchical three-body systems in terms of Hansen coefficients which are also valid for arbitrary order, but assuming coplanar orbits.

From equation (6.125), the terms depending on binary separations in the unaveraged binary pair Hamiltonian at order n are given by

$$\sum_{m=0}^n \frac{(\mathbf{r}_p \cdot \mathbf{r}_k)^m}{r_k^{n+m+1}} \frac{r_p^{n-m}}{r_p} \mathcal{A}_m^{(n)}, \tag{6.148}$$

where binary $p \in \{k, C\}$ (i.e. $r_p \ll r_k$), and $\mathcal{A}_m^{(n)}$ are coefficients that follow from the Taylor expansion of the potential energy and which are the same as the coefficients appearing in the Legendre polynomials (cf. Appendix 6.A.2, in particular equations 6.56 and 6.57). Generally, $\mathcal{A}_m^{(n)} = 0$ if n is even and m is odd, or if n is odd and m is even.

We first average over the ‘inner’ orbit $\mathbf{r}_p$ with constant $\mathbf{r}_k$ using the eccentric anomaly ε_p , giving (cf. equations 6.133, 6.135, 6.136 and 6.137)

$$\left\langle (\mathbf{r}_p \cdot \mathbf{r}_k)^m \frac{r_p^{n-m}}{r_p} \right\rangle_p = \frac{a_p^n}{2\pi} \int_0^{2\pi} d\varepsilon_p [1 - e_p \cos(\varepsilon_p)]^{n-m+1} \left[\{\cos(\varepsilon_p) - e_p\} (\dot{\mathbf{e}}_p \cdot \mathbf{r}_k) + \sqrt{1 - e_p^2} \sin(\varepsilon_p) (\hat{\mathbf{q}}_p \cdot \mathbf{r}_k) \right]^m. \tag{6.149}$$

Subsequently, we transform to the complex variable $z = \exp(i\varepsilon_p)$ where $i = \sqrt{-1}$, giving

$$\begin{aligned} \langle (\mathbf{r}_p \cdot \mathbf{r}_k)^m r_p^{n-m} \rangle_p &= \frac{a_p^n}{2\pi} \oint_C \frac{dz}{iz} \left[1 - \frac{e_p}{2} (z + z^{-1}) \right]^{n-m+1} \left[\frac{1}{2} (z + z^{-1} - 2e_p) (\hat{\mathbf{e}}_p \cdot \mathbf{r}_k) - \frac{1}{2} i \sqrt{1 - e_p^2} (z - z^{-1}) (\hat{\mathbf{q}}_p \cdot \mathbf{r}_k) \right]^m \\ &= \frac{a_p^n}{2\pi i} \left(\frac{1}{2} \right)^{n+1} \oint_C \frac{dz}{z^{n+2}} (-e_p z^2 + 2z - e_p)^{n-m+1} \left(\frac{\tilde{C}_1}{e_p} z^2 + \tilde{C}_3 z + \frac{\tilde{C}_2}{e_p} \right)^m. \end{aligned} \quad (6.150)$$

Here, the integration path C is the unit circle $z = \exp(i\varepsilon_p)$, $\varepsilon_p \in [0, 2\pi)$ in the complex plane, and the constants $\tilde{C}_1$, $\tilde{C}_2$ and $\tilde{C}_3$ are defined by

$$\tilde{C}_1 \equiv (\mathbf{e}_p \cdot \mathbf{r}_k) - i \sqrt{1 - e_p^2} e_p (\hat{\mathbf{q}}_p \cdot \mathbf{r}_k) = (\mathbf{e}_p \cdot \mathbf{r}_k) - i \sqrt{1 - e_p^2} (\pm 1) \left[e_p^2 r_k^2 - (\mathbf{e}_p \cdot \mathbf{r}_k)^2 - \frac{e_p^2}{1 - e_p^2} (\mathbf{j}_p \cdot \mathbf{r}_k)^2 \right]^{1/2}; \quad (6.151a)$$

$$\tilde{C}_2 \equiv (\mathbf{e}_p \cdot \mathbf{r}_k) + i \sqrt{1 - e_p^2} e_p (\hat{\mathbf{q}}_p \cdot \mathbf{r}_k) = (\mathbf{e}_p \cdot \mathbf{r}_k) + i \sqrt{1 - e_p^2} (\pm 1) \left[e_p^2 r_k^2 - (\mathbf{e}_p \cdot \mathbf{r}_k)^2 - \frac{e_p^2}{1 - e_p^2} (\mathbf{j}_p \cdot \mathbf{r}_k)^2 \right]^{1/2}; \quad (6.151b)$$

$$\tilde{C}_3 \equiv -2(\mathbf{e}_p \cdot \mathbf{r}_k), \quad (6.151c)$$

where we used the vector identity equation (6.138) to eliminate explicit reference to $\hat{\mathbf{q}}_p$.

Taking into account that $0 \leq m \leq n$ and $n > 0$, the integrand in equation (6.150) is a meromorphic function of z with a single pole at $z = 0$. Using a standard Taylor series, the integrand can be written as

$$\frac{1}{z^{n+2}} (-e_p z^2 + 2z - e_p)^{n-m+1} \left(\frac{\tilde{C}_1}{e_p} z^2 + \tilde{C}_3 z + \frac{\tilde{C}_2}{e_p} \right)^m = \frac{1}{z^{n+2}} \sum_{j \geq 0} a_j z^j = \sum_{j \geq 0} a_j \frac{1}{z^{n+2-j}}, \quad (6.152)$$

where the coefficients a_j are given by

$$a_j = \frac{1}{j!} \frac{d^j}{dz^j} \left[(-e_p z^2 + 2z - e_p)^{n-m+1} \left(\frac{\tilde{C}_1}{e_p} z^2 + \tilde{C}_3 z + \frac{\tilde{C}_2}{e_p} \right)^m \right]_{z=0} \quad (6.153)$$

Using Cauchy's theorem,

$$\oint_C \frac{dz}{z^k} = \begin{cases} 0, & k \neq 1; \\ 2\pi i, & k = 1, \end{cases} \quad (6.154)$$

it follows that the integral in equation (6.150) is zero, unless $n + 2 - j = 1$, or $j = n + 1$. Therefore, equation (6.150) can be written as

$$\langle (\mathbf{r}_p \cdot \mathbf{r}_k)^m r_p^{n-m} \rangle_p = a_p^n \left(\frac{1}{2} \right)^{n+1} \frac{1}{(n+1)!} \frac{d^{n+1}}{dz^{n+1}} \left[(-e_p z^2 + 2z - e_p)^{n-m+1} \left(\frac{\tilde{C}_1}{e_p} z^2 + \tilde{C}_3 z + \frac{\tilde{C}_2}{e_p} \right) \right]_{z=0}. \quad (6.155)$$

Alternatively, one can apply the residue theorem to the integral in equation (6.150), leading to the same result.

For any $n > 0$, equation (6.155) reduces to a real-valued function of e_p , r_k , $e_p \cdot \mathbf{r}_k$ and $\mathbf{j}_p \cdot \mathbf{r}_k$. In particular, $\tilde{C}_1$ and $\tilde{C}_2$ (necessarily) always appear in a real-valued combination, and the plus/minus sign in equations (6.151) always reduces to $+1$. Allowed combinations are, for example,

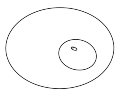
$$\tilde{C}_1 + \tilde{C}_2 = 2(e_k \cdot \mathbf{r}_k); \quad \tilde{C}_1 \tilde{C}_2 = (e_p \cdot \mathbf{r}_k)^2 + (1 - e_p^2) \left[e_p^2 r_k^2 - (e_p \cdot \mathbf{r}_k)^2 - \frac{e_p^2}{1 - e_p^2} (\mathbf{j}_p \cdot \mathbf{r}_k)^2 \right]. \quad (6.156)$$

Furthermore, on dimensional grounds, the maximum combined exponent of r_k , $e_p \cdot \mathbf{r}_k$ and $\mathbf{j}_p \cdot \mathbf{r}_k$ is m . Therefore, equation (6.155) can be formally written as

$$\langle (\mathbf{r}_p \cdot \mathbf{r}_k)^m r_p^{n-m} \rangle_p = a_p^n \sum_{\substack{i_1, i_2 \in \mathbb{N}^0 \\ i_1 + i_2 \leq m}} r_k^{m-i_1-i_2} (e_p \cdot \mathbf{r}_k)^{i_1} (\mathbf{j}_p \cdot \mathbf{r}_k)^{i_2} \mathcal{B}_{i_1, i_2}^{(n, m)}(e_p), \quad (6.157)$$

where $\mathbb{N}^0$ are the natural numbers plus zero, and $\mathcal{B}_{i_1, i_2}^{(n, m)}(e_p)$ are polynomial functions of e_p which can be computed from equation (6.155).

Subsequently, we average over the 'outer' orbit k using the true anomaly f_k . From equations (6.148) and (6.157), it follows that the



terms which need to be averaged are of the form (cf. equations 6.133, 6.135, 6.136 and 6.137)

$$\begin{aligned}
 \left\langle \frac{r_k^{m-i_1-i_2} (e_p \cdot r_k)^{i_1} (j_p \cdot r_k)^{i_2}}{r_k^{n+m+1}} \right\rangle_k &= \left\langle \frac{(e_p \cdot r_k)^{i_1} (j_p \cdot r_k)^{i_2}}{r_k^{n+i_1+i_2+1}} \right\rangle_k \\
 &= \frac{(1-e_k^2)^{\frac{1}{2}-n}}{a_k^{n+1}} \frac{1}{2\pi} \int_0^{2\pi} df_k [1 + e_k \cos(f_k)]^{n-1} [\cos(f_k) (\hat{e}_k \cdot e_p) + \sin(f_k) (\hat{q}_k \cdot e_p)]^{i_1} [\cos(f_k) (\hat{e}_k \cdot j_p) + \sin(f_k) (\hat{q}_k \cdot j_p)]^{i_2} \\
 &= \frac{(1-e_k^2)^{\frac{1}{2}-n}}{a_k^{n+1}} \left(\frac{1}{2} \right)^{n+i_1+i_2-1} \frac{1}{2\pi i} \oint_C \frac{dz}{z^{n+i_1+i_2}} (e_k z^2 + 2z + e_k)^{n-1} \left(\frac{\tilde{C}_{2e}}{e_k} z^2 + \frac{\tilde{C}_{1e}}{e_k} \right)^{i_1} \left(\frac{\tilde{C}_{2j}}{e_k} z^2 + \frac{\tilde{C}_{1j}}{e_k} \right)^{i_2}, \quad (6.158)
 \end{aligned}$$

where we switched again to complex integration with $z = \exp(if_k)$, and the constants are given by

$$\tilde{C}_{1e} \equiv (e_p \cdot e_k) + i e_k (\hat{q}_k \cdot e_p); \quad (6.159a)$$

$$\tilde{C}_{2e} \equiv (e_p \cdot e_k) - i e_k (\hat{q}_k \cdot e_p); \quad (6.159b)$$

$$\tilde{C}_{1j} \equiv (e_k \cdot j_p) + i e_k (\hat{q}_k \cdot j_p); \quad (6.159c)$$

$$\tilde{C}_{2j} \equiv (e_k \cdot j_p) - i e_k (\hat{q}_k \cdot j_p). \quad (6.159d)$$

As before, the integrand in equation (6.155) is a meromorphic function with a single pole at $z = 0$. By the same arguments,

$$\begin{aligned}
 &\left\langle \frac{(e_p \cdot r_k)^{i_1} (j_p \cdot r_k)^{i_2}}{r_k^{n+i_1+i_2+1}} \right\rangle_k \\
 &= \frac{(1-e_k^2)^{\frac{1}{2}-n}}{a_k^{n+1}} \left(\frac{1}{2} \right)^{n+i_1+i_2-1} \frac{1}{(n+i_1+i_2-1)!} \frac{d^{n+i_1+i_2-1}}{dz^{n+i_1+i_2-1}} \left[\frac{(e_k z^2 + 2z + e_k)^{n-1}}{e_k} \left(\frac{\tilde{C}_{2e}}{e_k} z^2 + \frac{\tilde{C}_{1e}}{e_k} \right)^{i_1} \left(\frac{\tilde{C}_{2j}}{e_k} z^2 + \frac{\tilde{C}_{1j}}{e_k} \right)^{i_2} \right]_{z=0} \quad (6.160)
 \end{aligned}$$

The expression in equation (6.160) is generally a function of $e_p, e_k, e_p \cdot e_k, j_p \cdot j_k, e_p \cdot j_k$ and $e_k \cdot j_p$. The combined exponent of $e_p \cdot e_k$

and $e_p \cdot j_k$ is at most i_1 , whereas the combined exponent of $j_p \cdot j_k$ and $e_k \cdot j_p$ is at most i_2 . Formally, equation (6.160) can be written as

$$\left\langle \frac{(e_p \cdot r_k)^{i_1} (j_p \cdot r_k)^{i_2}}{r_k^{n+i_1+i_2+1}} \right\rangle_k = \frac{(1-e_k^2)^{\frac{1}{2}-n}}{a_k^{n+1}} \sum_{\substack{l_1, l_2, l_3, l_4 \in \mathbb{N}^0 \\ l_1+l_3 \leq i_1 \\ l_2+l_4 \leq i_2}} C_{l_1, l_2, l_3, l_4}^{(n, i_1, i_2)} (e_p, e_k) (e_p \cdot e_k)^{l_1} (j_p \cdot j_k)^{l_2} (e_p \cdot j_k)^{l_3} (e_k \cdot j_p)^{l_4}, \quad (6.161)$$

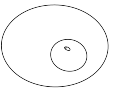
where $C_{l_1, l_2, l_3, l_4}^{(n, i_1, i_2)}(e_p, e_k)$ are polynomial functions of e_p and e_k that can be computed from equation (6.160) using the vector identity equation (6.138).

With these formal expressions for the orbit averages, the averaged pairwise Hamiltonian term to order n is given by

$$\begin{aligned} \overline{S}'_{n,2} &= (-1)^{n+1} \sum_{k \in \mathbf{B}} \sum_{\substack{p \in \mathbf{B} \\ p \in \{k, \mathbf{C}\}}} M_{p, \mathbf{C}1} M_{p, \mathbf{C}2} \left| \frac{M_{p, \mathbf{C}2}^{n-1} + (-1)^n M_{p, \mathbf{C}1}^{n-1}}{M_p} \right| \frac{GM_{k, \mathbf{C}S(p)}}{a_k} \left(\frac{a_p}{a_k} \right)^n \frac{1}{j_k^{2n-1}} \\ &\quad \times \sum_{m=0}^n \sum_{\substack{i_1, i_2 \in \mathbb{N}^0 \\ i_1+i_2 \leq m}} \sum_{\substack{l_1, l_2, l_3, l_4 \in \mathbb{N}^0 \\ l_1+l_3 \leq i_1 \\ l_2+l_4 \leq i_2}} \mathcal{A}_m^{(n)} \mathcal{B}_{i_1, i_2}^{(n, m)}(e_p) C_{l_1, l_2, l_3, l_4}^{(n, i_1, i_2)}(e_p, e_k) (e_p \cdot e_k)^{l_1} (j_p \cdot j_k)^{l_2} (e_p \cdot j_k)^{l_3} (e_k \cdot j_p)^{l_4}, \end{aligned} \quad (6.162)$$

where $\mathcal{A}_m^{(n)}$ are Legendre polynomial coefficients, given implicitly by (cf. Appendix 6.A.2, in particular equations 6.56 and 6.57)

$$\sum_{m=0}^n \mathcal{A}_m^{(n)} x^m = \frac{1}{2^n n!} \frac{d^n}{dx^n} \left[(x^2 - 1)^n \right], \quad (6.163)$$



and $\mathcal{B}_{i_1, i_2}^{(n, m)}(e_p)$ and $\mathcal{C}_{l_1, l_2, l_3, l_4}^{(n, m, i_1, i_2)}(e_p, e_k)$ are polynomial functions of e_p and e_k , given implicitly by

$$\begin{aligned}
 & \sum_{\substack{i_1, i_2 \in \mathbb{N}^0 \\ i_1 + i_2 \leq m}} r_k^{m-i_1-i_2} (e_p \cdot \mathbf{r}_k)^{i_1} (\mathbf{j}_p \cdot \mathbf{r}_k)^{i_2} \mathcal{B}_{i_1, i_2}^{(n, m)}(e_p) \\
 &= \left(\frac{1}{2}\right)^{n+1} \frac{1}{(n+1)!} \frac{\mathbf{d}^{n+1}}{\mathbf{d}z^{n+1}} \left[(-e_p z^2 + 2z - e_p)^{n-m+1} \left(\frac{\tilde{C}_1}{e_p} z^2 + \tilde{C}_3 z + \frac{\tilde{C}_2}{e_p} \right)^m \right]_{z=0}; \\
 & \sum_{\substack{l_1, l_2, l_3, l_4 \in \mathbb{N}^0 \\ l_1 + l_3 \leq i_1 \\ l_2 + l_4 \leq i_2}} \mathcal{C}_{l_1, l_2, l_3, l_4}^{(n, i_1, i_2)}(e_p, e_k) (e_p \cdot \mathbf{e}_k)^{l_1} (\mathbf{j}_p \cdot \mathbf{j}_k)^{l_2} (e_p \cdot \mathbf{j}_k)^{l_3} (e_k \cdot \mathbf{j}_p)^{l_4} \\
 &= \left(\frac{1}{2}\right)^{n+i_1+i_2-1} \frac{1}{(n+i_1+i_2-1)!} \frac{\mathbf{d}^{n+i_1+i_2-1}}{\mathbf{d}z^{n+i_1+i_2-1}} \left[(e_k z^2 + 2z + e_k)^{n-1} \left(\frac{\tilde{C}_{2e}}{e_k} z^2 + \frac{\tilde{C}_{1e}}{e_k} \right)^{i_1} \left(\frac{\tilde{C}_{2j}}{e_k} z^2 + \frac{\tilde{C}_{1j}}{e_k} \right)^{i_2} \right]_{z=0}. \quad (6.164a)
 \end{aligned}$$

Here, the coefficients $\tilde{C}_1$ etc. are given by equations (6.151) and (6.159).

6.A.6 Estimate of the importance of non-binary terms for nested planetary systems

Here, we generalize the estimates of the importance of terms other than pairwise binary terms for nested planetary systems, made in Section 6.2.4, to arbitrary order n . As in Section 6.2.4, we consider three nested binaries p , u and k , where $p \in \{k, C\}$, $p \in \{u, C\}$ and $u \in \{k, C\}$. First, we consider the importance of the pairwise binary terms compared to the triplet term; for the pairwise binary terms, we make a distinction between the pairwise combinations (p, u) and (u, k) . For our purposes of estimating the (orbit-averaged) Hamiltonian, we approximate β_n as

$$\beta_n(\mathbf{r}_{u_1}, \dots, \mathbf{r}_{u_n}; \mathbf{r}_k) \sim \frac{a_{u_1} \dots a_{u_n}}{a_k^n}, \quad (6.165)$$

i.e. we replace the norms of the separation vectors with semimajor axes, and ignore the ‘angular’ dependence, i.e. the dependence on the unit vectors $\hat{\mathbf{r}}_{u_i}$ and $\hat{\mathbf{r}}_k$ (cf. Appendix 6.A.2).

Using equation (6.123) (or equation 6.124), this gives the following estimate for the (orbit-averaged) pairwise terms, applied to the pair (p, u) ,

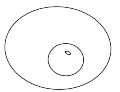
$$\left| \overline{S}_{n;2}' \right|_{(p,u)} \sim \frac{G}{a_u} c_n \left(\frac{a_p}{a_u} \right)^n M_{u,CS(p)} \frac{M_{p,C1} M_{p,C2}}{M_p} \frac{|M_{p,C2}^{n-1} + (-1)^n M_{p,C1}^{n-1}|}{M_p^{n-1}}. \quad (6.166)$$

Applied to the pair (u, k) ,

$$\left| \overline{S}_{n;2}' \right|_{(u,k)} \sim \frac{G}{a_k} c_n \left(\frac{a_u}{a_k} \right)^n M_{k,CS(u)} \frac{M_{u,C1} M_{u,C2}}{M_u} \frac{|M_{u,C2}^{n-1} + (-1)^n M_{u,C1}^{n-1}|}{M_u^{n-1}}. \quad (6.167)$$

The triplet term at order n , applied to (p, u, k) , is found in equation (6.123) by setting (1) $n_1 = n$, $n_2 = 0$, $d_p = n - 1$ and $d_q = 0$, and (2) $n_1 = 0$, $n_2 = n$, $d_p = n - 1$ and $d_q = 0$, giving the estimate

$$\left| \overline{S}_{n;3}' \right|_{(p,u,k)} \sim \frac{G}{a_k} c_n \frac{a_p^{n-1} a_u}{a_k^n} \frac{n!}{(n-1)!} \frac{M_{u,CS(p)}}{M_u} \frac{M_{p,C1} M_{p,C2}}{M_p} \frac{|M_{p,C2}^{n-2} + (-1)^{n-1} M_{p,C1}^{n-2}|}{M_p^{n-2}} M_{k,CS(p)}. \quad (6.168)$$



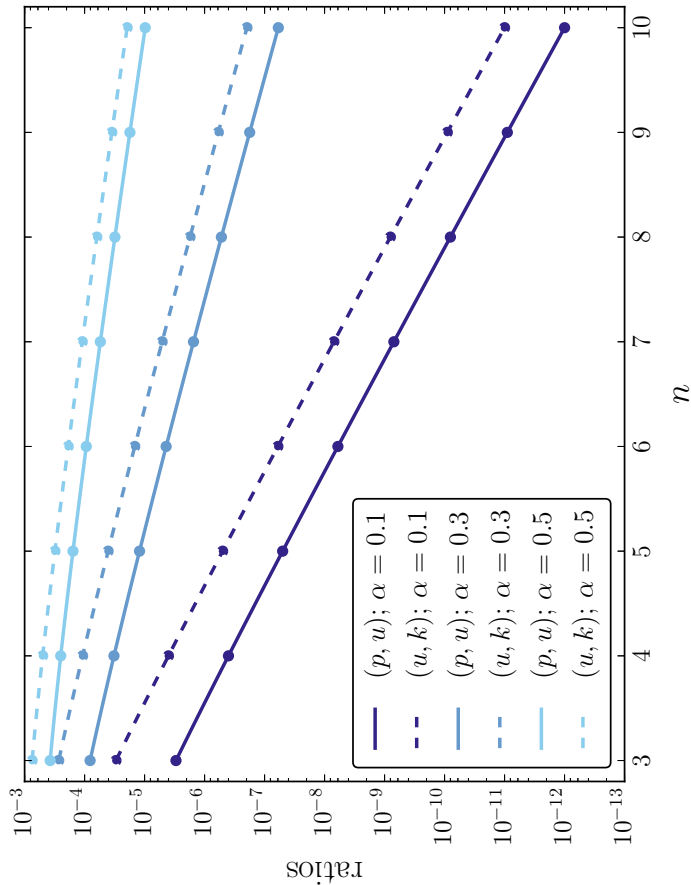


Figure 6.38: The ratios of the binary triplet to the pairwise term in the Hamiltonian, applied to both pairs (p, u) and (u, k) (cf. equation 6.170). We assume $\tilde{q} = 10^{-3}$ and three values of α , indicated in the legend.

Therefore,

$$\frac{|\overline{S}'_{n;3}|_{(p,u,k)}}{|\overline{S}'_{n;2}|_{(p,u)}} \sim n \frac{M_{k,CS(p)}}{M_u} \frac{|M_{p,C2}^{n-2} M_p + (-1)^{n-1} M_{p,C1}^{n-2} M_p|}{|M_{p,C2}^{n-1} + (-1)^n M_{p,C1}^{n-1}|} \frac{a_u}{a_p} \left(\frac{a_u}{a_k} \right)^{n+1}; \quad (6.169a)$$

$$\frac{|\overline{S}'_{n;3}|_{(p,u,k)}}{|\overline{S}'_{n;2}|_{(u,k)}} \sim n \frac{M_{u,CS(p)}}{M_p} \frac{M_{p,C1} M_{p,C2}}{M_{u,C1} M_{u,C2}} \frac{|M_{p,C2}^{n-2} M_p + (-1)^{n-1} M_{p,C1}^{n-2} M_p|}{|M_{u,C2}^{n-1} + (-1)^n M_{u,C1}^{n-1}|} \left(\frac{M_u}{M_p} \right)^{n-1} \left(\frac{a_p}{a_u} \right)^{n-1}. \quad (6.169b)$$

Note that for $n = 3$, these estimates, apart from any dependence on the eccentricities, reduce to equations (6.12) and (6.13). We assume a nested planetary system with stellar mass M_* and planetary mass m_p such that $m_p/M_* = \tilde{q} \ll 1$, and semimajor axis ratios $a_p/a_u = a_u/a_k = \alpha \ll 1$. In this case, $M_{p,C1} = M_*$, $M_{p,C2} = m_p$, $M_{u,C1} \approx M_*$, $M_{u,C2} = m_p$, $M_{k,C1} \approx M_*$ and $M_{k,C2} = m_p$. This gives

$$\frac{|\overline{S}'_{n;3}|_{(p,u,k)}}{|\overline{S}'_{n;2}|_{(p,u)}} \sim n \tilde{q} \alpha^n; \quad (6.170a)$$

$$\frac{|\overline{S}'_{n;3}|_{(p,u,k)}}{|\overline{S}'_{n;2}|_{(u,k)}} \sim n \tilde{q} \alpha^{n-1} \quad (6.170b)$$

(note that $|M_{p,C2}^{n-2} M_p + (-1)^{n-1} M_{p,C1}^{n-2} M_p| \approx |m_p^{n-2} M_* + (-1)^{n-1} M_*^{n-1}| \approx M_*^{n-1}$).

In Figure 6.38, we show the ratios given by equations (6.170) as a function of n , assuming $\tilde{q} = 10^{-3}$ (i.e. order Jupiter-mass planets around a solar-mass star), and various values of α . The Hamiltonian ratios decrease strongly for increasing n and decreasing α . For the smallest applicable n , $n = 3$, and $\alpha = 0.5$, the ratios are $< 10^{-3}$, showing that the triplet terms are small.

More generally, and less quantitatively, equation (6.123) shows that for nested planetary systems, the non-binary terms, in contrast to pairwise binary terms, are multiplied by mass ratio factors $[M_{p_l,CS(p)}/M_{p_l}]^{d_{p_l}} \approx (m_p/M_*)^{d_{p_l}} = \tilde{q}^{d_{p_l}}$ with $1 \leq d_{p_l} \leq n-2$. Also, note that for planetary systems, the quantity $\mathcal{M}_{p,k,CS(p)}^{(d_p)}$ is approximately the same for pairwise and non-pairwise terms. Furthermore, the function β_n (cf. equation 6.165) is generally smaller for non-binary pairs compared to adjacent binary pairs, because in addition to ratios of semimajor axes of adjacent pairs, the non-binary pair terms also include ratios of semimajor axes of non-adjacent pairs (e.g. equation 6.168

contains the ratio $a_p/a_k \ll a_u/a_k$). We conclude that for nested planetary systems, the non-binary pair terms are generally small compared to the pairwise binary terms applied to adjacent binaries.

6.A.7 Ad hoc expression for the 1PN Hamiltonian

In the post-Newtonian (PN) approximation, corrections to Newtonian gravity due to general relativity are included by adding terms of order $(v/c)^n$, where v is the orbital velocity and c is the speed of light. To lowest order in (v/c) , the N -body equations of motion are known and have been derived by Lorentz & Droste (1917) and Einstein et al. (1938). They give rise to a 1PN Hamiltonian that is conserved to order $(v/c)^2$. For the hierarchical three-body problem, the 1PN Hamiltonian was expanded in binary separation ratios and orbit averaged by Naoz et al. (2013b). In the latter paper, it was found that, in addition to terms corresponding to separate precession in the inner and outer orbits, there is also a term associated with both inner and outer binaries, i.e. an ‘interaction’ term.

Here, we do *not* attempt to derive, from first principles, the generalized 1PN Hamiltonian H_{1PN} for hierarchical N -body systems. Rather, we construct an ad hoc expression by assuming that the dominant terms in H_{1PN} are given by terms which, individually, depend on only one binary, and which give rise to the well-known rate of precession in the 1PN two-body problem. Extrapolating from e.g. the results of Naoz et al. (2013b), these terms for a binary k should have the form

$$H_{\text{1PN},k} = -\frac{3G^2 M_{k,C1} M_{k,C2} M_k}{c^2 a_k^2 (1 - e_k^2)^{1/2}}. \quad (6.171)$$

From the equations of motion (cf. equation 6.17), it follows that the precession associated with equation (6.171), which only depends on $\mathbf{e}_k$, is

$$\frac{d\mathbf{e}_k}{dt} = \frac{2\pi}{t_{\text{1PN}}} \hat{\mathbf{j}}_k \times \mathbf{e}_k, \quad (6.172)$$

i.e. $\mathbf{e}_k$ precesses around $\hat{\mathbf{j}}_k$ with time-scale t_{1PN} . The latter is given by

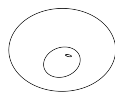
$$t_{\text{1PN}} = \frac{1}{3} P_{\text{orb},k} (1 - e_k^2) \frac{a_k}{r_{g,k}}, \quad (6.173)$$

where the orbital period is $P_{\text{orb},k} = 2\pi\sqrt{a_k^3/(GM_k)}$ and the gravitational radius is $r_{g,k} = GM_k/c^2$.

Therefore, our assumed 1PN Hamiltonian is

$$H_{\text{1PN}} \approx H - \sum_{k \in \text{B}} \frac{3G^2 M_{k,C1} M_{k,C2} M_k}{c^2 a_k^2 (1 - e_k^2)^{1/2}}, \quad (6.174)$$

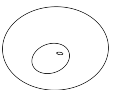
where H is the Newtonian part given by equation (6.127). Evidently, in this form, orbit averaging of equation (6.174) only affects H .



7 | Implications of secular evolution on high eccentricity migration giving rise to hot Jupiters in multiplanet systems

Hot Jupiters (HJs) are Jupiter-like planets that reside very closely to their host star, within ~ 0.05 AU. Their formation is not well understood. It is typically believed that they cannot have formed in situ, implying that some form of migration must have occurred after their formation. We study the production of HJs through secular evolution in multiplanet systems with 3 to 5 planets. In this variant of high- e migration, the eccentricity of the orbit of the innermost planet is excited on secular time-scales, triggering orbital migration due to tidal dissipation. We use a secular dynamics code to study the conditions for producing high eccentricities through secular evolution, and we carry out a population synthesis study of multiplanet systems. We find that an angular momentum deficit (AMD) satisfying $1 \lesssim \text{AMD}/(m_1 \sqrt{a_1}) \lesssim 2$ is required to produce HJs. In addition, the ratio of the Lidov-Kozai (LK) time-scales applied to the innermost two orbits should be close to unity. We find that HJs are only produced if the viscous time-scale is very short (≈ 0.014 yr). In contrast, up to ≈ 0.3 of systems lead to tidal disruption of the innermost planet. Although the semimajor axis distribution of the HJs is consistent with observations, we do not find any warm Jupiters in our simulations, i.e. planets with semimajor axes between 0.1 and 1 AU.

Hamers, A.S., Antonini, F., Lithwick, Y., Perets, H.B. & Portegies Zwart, S.F. In preparation.



7.1 Introduction

In the past decades, radial velocity and transit methods have revealed a population of gas giant planets of order Jupiter mass around solar-type stars with orbital periods downward of 10 d, i.e. hot Jupiters (HJs). Currently, the main consensus is that HJs could not have formed *in situ* in the protoplanetary disc phase, because of an insufficient amount of disc material and/or too high temperatures at these close regions to the host star (e.g. Lin et al. 1996; however, recently it has been suggested that *in situ* formation might be possible through core accretion, Batygin et al. 2015). If HJs were formed at larger separations, i.e. beyond the snow line of one to a few AU, then this implies that they must have experienced strong inwards migration after their formation, by two orders of magnitude in separation. Two main migration scenarios have been proposed: (1) migration induced by orbital energy dissipation due to gas drag in the protoplanetary disc phase (e.g. Goldreich & Tremaine 1980; Lin & Papaloizou 1986; Bodenheimer et al. 2000; Tanaka et al. 2002), and (2) migration induced by tidal dissipation in the HJ, requiring high orbital eccentricity (commonly known as ‘high-*e*’ migration).

For high-*e* migration, various subscenarios have been proposed to drive the high eccentricities needed to produce the observed short orbital periods through tidal dissipation. They include (1) eccentricity excitation because of close encounters between planets (Rasio & Ford 1996; Chatterjee et al. 2008; Ford & Rasio 2008; Jurić & Tremaine 2008), (2) excitation of the eccentricity because of secular Lidov-Kozai (LK) oscillations (Lidov 1962; Kozai 1962) induced by a distant binary companion star or an additional (massive) planet on an inclined orbit (Wu & Murray 2003; Fabrycky & Tremaine 2007; Naoz et al. 2012; Petrovich 2015a; Anderson et al. 2016; Petrovich & Tremaine 2016), (3) secular eccentricity excitation induced by a close and coplanar, but eccentric planetary companion (Petrovich 2015b), and (4) eccentricity excitation induced by secular chaos in multiplanet systems with at least 3 planets, and with mildly inclined and eccentric orbits (Wu & Lithwick 2011; Lithwick & Wu 2011, 2014).

It is currently unclear which of the two main migration scenarios (1) and (2) applies to the observed HJs, or if a combination of the two is responsible for forming HJs. Both scenarios have successes and failures in describing properties of observed HJs. The observed period distribution of HJs peaks around $\sim 3 - 5$ d, and their eccentricities are close to zero (e.g. Santerne et al. 2016). Most high-*e* migration scenarios predict that the final orbit of the HJ should indeed be circular and pile up with a short period, around ~ 3 d. In contrast, disc migration scenarios cannot explain a specific peak in the period distribution. On the other hand, high-*e* migration through LK cycles in three-body systems (subscenarios 2 and 3) require the presence of a stellar binary or compact planetary companion. However, not all HJs are observed to have such a companion. In recent surveys, $\sim 50\%$ of HJs are found to have a stellar companion (e.g. Knutson et al. 2014; Ngo et al. 2015).

The fourth high-*e* scenario involves three or more planets around a single star, i.e. $N > 3$ systems (Wu & Lithwick 2011; Lithwick & Wu 2011, 2014). Similarly to the case of three-body systems (i.e. a binary companion or two planets), secular interactions can change the eccentricities of the orbits, in particular the innermost orbit, on long time-scales (i.e. much longer than the orbital periods). For $N = 3$, high relative inclinations (typically $\gtrsim 40^\circ$) and/or tight and eccentric outer orbits are required to produce high eccentricities. However, in the case of $N > 3$, the conditions for producing high eccentricities in the innermost orbit are less stringent. In particular, the initial eccentricities and relative inclinations can be much smaller, and the planets need not be very closely spaced. For these conditions, suborbital effects such

as close encounters and mean motion resonances are typically unimportant, and the long-term evolution is driven mainly by secular interactions. In contrast to $N = 3$, for $N > 3$ the secular evolution is typically chaotic, giving rise to highly irregular eccentricity oscillations. Over the course of $\sim 0.1 - 10$ Gyr, high eccentricities can be reached in the inner orbit, potentially leading to HJs. A well known example is the Solar system, in which secular chaos can drive the orbit of Mercury to become unstable (i.e. lead to an ejection of the planet, or a collision of the planet with the Sun) on a time-scale of 5 Gyr, and with a probability of $\approx 1\%$ (Laskar & Gastineau 2009; Lithwick & Wu 2014).

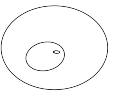
The less stringent conditions for secular chaos in $N > 3$ systems are compatible with current observations of HJs which exclude close-in companions for a subset of HJs, whereas detections of further-away companions (at $\gtrsim 5 - 10$ AU) are still largely incomplete because of observational limitations. This argument can also be reversed: the production of a HJ through secular chaos *requires* further-away companions, therefore these companions are expected to be observed in HJ systems in the future.

Despite the potential importance of secular evolution in multiplanet systems for the production of HJs, extensive parameter space studies and/or Monte Carlo studies to quantify observed properties of the HJs (in particular, the final period and obliquity distribution) are still lacking. The long secular time-scales compared with the short orbital periods mean that direct N -body integrations, such as those carried out by Wu & Lithwick (2011), are computationally very expensive to carry out, especially considering the large number of parameters for $N > 3$ systems. Until recently, secular, orbit-averaged methods valid for high eccentricities and inclinations were limited to three bodies and to systems that are not too compact, as a consequence of the expansion of the Hamiltonian in terms of ratios of binary separations (or, after averaging, in terms of semimajor axes ratios).

In recent work (Hamers & Portegies Zwart 2016), we presented a generalisation of the secular, orbit-averaged method previously used for hierarchical three-body systems (e.g. Lidov 1962; Kozai 1962; Harrington 1968; Naoz et al. 2013a), to systems composed of nested binary orbits, with an arbitrary number of bodies and an arbitrary hierarchy. In the latter paper, we explicitly derived the orbit-averaged equations of motion to fifth order in the separation ratios for binary pairs, and showed that such high orders are required to accurately capture the orbital evolution in planetary systems. We implemented the equations of motion in a code, `SECULARMULTIPLE`. Here, we apply `SECULARMULTIPLE` to study the formation of HJs through secular evolution in multiplanet systems with $N > 3$. The main advantage of this method is that compared to direct N -body integration, the evolution can be computed much faster, allowing to quickly explore large regions of the parameter space.

Our main focus are the conditions for producing high eccentricities and HJs through secular evolution, and the implied properties of the HJs. Using a combination of population synthesis with grid sampling, we study the dependence of the HJ properties on various parameters, including the efficiency of tidal dissipation, the number of planets, the width of the initial mutual inclination and eccentricity distribution, and the radius of the innermost planet. These parameters, in particular the efficiency of tidal dissipation, are highly uncertain. In the context of the scenario of secular evolution in multiplanet systems, they can be constrained by comparing our simulated distributions to the observed distributions.

The structure of this paper is as follows. In Section 7.2, we describe the secular method and other assumptions. We verify the secular method by comparing to a (limited) number of direct N -body integrations in Section 7.3. Subsequently, we apply `SECULARMULTIPLE` to explore the



conditions for producing HJs in Section 7.4. In Section 7.5, we carry out a population synthesis study. We note that Sections 7.3 and 7.4 are of a more technical nature, and could be omitted at a first reading. A summary of the latter sections is given at the beginning of Section 7.5. We discuss our results in Section 7.6, and we conclude in Section 7.7.

7.2 Methods and assumptions

7.2.1 Notations and overview

In Table 7.1, we give a list of quantities used in this work with a description. Where applicable, we give for reference the values of the (initial) parameters that are assumed in the various sections of this paper.

7.2.2 Secular dynamics

To model the long-term gravitational dynamics of the multiplanet system, we use the algorithm `SECULARMULTIPLE` (Hamers & Portegies Zwart 2016) within the `AMUSE` framework (Portegies Zwart et al. 2013; Pelupessy et al. 2013). `SECULARMULTIPLE` applies to self-gravitating systems composed of nested binaries with an arbitrary hierarchy. In particular, a multiplanet system is represented as a ‘nested’ hierarchical multiple system of point particles; the star is contained within the ‘innermost’ binary system (i.e. Jacobian coordinates; see e.g. Figure 3 of Hamers & Portegies Zwart 2016). The algorithm is based on an expansion of the Hamiltonian in terms of binary separation ratios, which are assumed to be small. The resulting Hamiltonian is orbit averaged, and the equations of motion, defining a system of ordinary differential equations, are solved numerically in terms of the orbital vectors $\mathbf{e}_k$ and $\mathbf{h}_k$ for all binaries k .

As shown in Hamers & Portegies Zwart (2016), depending on the compactness of the system, high orders are required to accurately describe the orbital evolution. Here, we include terms corresponding to pairwise binary interactions up and including fifth order in the separation ratios. To third order (‘octupole order’), we include the triplet binary terms, although these terms are unimportant for multiplanet systems with roughly equal-mass planets (Hamers & Portegies Zwart 2016). At higher orders, terms that individually correspond to more than two binaries, are not included; as shown in Hamers & Portegies Zwart (2016), these terms are likewise unimportant compared to the pairwise binary terms.

Relativistic corrections are included in the algorithm by adding the orbit-averaged precession rates of the line of apsides to the orbits. Potential ‘interaction’ terms between binaries (e.g. Naoz et al. 2013b) are neglected (see also Hamers & Portegies Zwart 2016).

7.2.3 Tidal evolution

The tidal evolution of the innermost planet and the star is modeled with the equilibrium tide model of Eggleton et al. (1998). This model also includes the effect of precession of the orbit of the innermost planet due to tidal bulges, and the spin evolution of both the star and the planet. The equilibrium tide model is described in terms of the tidal time lag τ or viscous time-scale t_V , the apsidal motion constant k_{AM} , the gyration radius r_g and the initial spin period P_s , for both the star and the innermost planet. For reference, our assumed values are given in Table 7.1.

Symbol	Description	Values in section		
		7.3	7.4	7.5
N_p	number of planets	3	3-5	3-5
M_*	stellar mass	$1 M_\odot$	$1 M_\odot$	$1 M_\odot$
m_i	mass of planet i	$3 \times 10^{-4} M_\odot$ ($i = 1$) $1 \times 10^{-4} M_\odot$ ($i > 1$)	$1 M_J$	$0.5\text{-}5 M_J$
R_*	stellar radius	$1 R_\odot$	$1 R_\odot$	$1 R_\odot$
R_i	radius of planet i	$1 R_J$	$1 R_J$	$1\text{-}1.5 R_J$
η	tidal disruption factor (cf. equation 7.7)	—	2.7	2.7
$t_{v,*}$	stellar viscous time-scale	—	—	5 yr
$t_{v,1}$	planet 1 viscous time-scale	≈ 4.8 yr	—	0.0137, 0.137, 1.37 yr
$k_{AM,*}$	stellar apsidal motion constant	—	0.014	0.014
$k_{AM,1}$	planet 1 apsidal motion constant	0.19	0.25	0.25
$r_{g,*}$	stellar gyration radius	—	0.08	0.08
$r_{g,1}$	planet 1 gyration radius	—	0.25	0.25
$P_{s,*}$	stellar spin period	—	10 d	10 d
$P_{s,1}$	planet 1 spin period	—	10 hr	10 hr
θ_*	stellar obliquity (stellar spin-orbit 1 angle)	—	0°	0°
θ_1	planet 1 obliquity (planet 1 spin-orbit 1 angle)	—	0°	0°
a_i	planet i orbital semimajor axis	1 AU ($i = 1$) 6 AU ($i = 2$) $12\text{-}62 \text{ AU}$ ($i = 3$)	1 AU ($i = 1$) 6 AU ($i = 2$) $\approx 12\text{-}68 \text{ AU}$ ($i = 3$) $\approx 25\text{-}1067 \text{ AU}$ ($i = 4$) $\approx 52\text{-}2.0 \times 10^4 \text{ AU}$ ($i = 5$)	$1\text{-}4 \text{ AU}$ ($i = 1$) $6\text{-}10 \text{ AU}$ ($i = 2$) $15\text{-}30 \text{ AU}$ ($i = 3$) $35\text{-}50 \text{ AU}$ ($i = 4$) $60\text{-}100 \text{ AU}$ ($i = 5$)
e_i	planet i orbital eccentricity	$\approx 0.4\text{-}0.6$	0-0.5	0-0.8
i_i	planet i orbital inclination	—	0-0.5 rad	0-30°
ω_i	planet i argument of pericenter	0-360°	0-360°	0-360°
Ω_i	planet i longitude of ascending node	0-360°	0-360°	0-360°
β	width of inclination and eccentricity distribution	—	—	8.2, 14.6, 32.8

Table 7.1: Notations of the quantities used in this work. Where applicable, we give the values of the (initial) parameters that are assumed in the simulations in various sections.

We assumed a constant tidal time lag τ , or, equivalently, a constant viscous time-scale t_V . The latter are related according to (e.g. Socrates et al. 2012)

$$t_V = 3 \left(1 + \frac{1}{k_L} \right) \frac{R^3}{GM\tau}, \quad (7.1)$$

where k_L is the potential Love number (i.e. twice the apsidal motion constant, $k_L = 2k_{AM}$), G is the gravitational constant, and R and M are the radius and mass of the body. Apart from its simplicity, a constant τ follows from the equations of motion with a number of physically-motivated assumptions (Socrates & Katz 2012). For high- e migration, Socrates et al. (2012) derive a value of $\tau_1 \approx 0.66$ s, or $t_{V,1} \approx 1.2 \times 10^4$ hr, for gas giant planets (as expressed by equation 7.1, the precise relation between τ and t_V depends on k_L , R and M).

We note that our assumption of a constant τ is in contrast to Wu & Murray (2003), who assumed a τ which is dependent on the orbital period (Socrates et al. 2012). A constant τ is also assumed in Fabrycky & Tremaine (2007); however, the latter authors assumed $t_{V,1} = 10^{-3}$ yr, which is three orders of magnitude smaller than the value derived by Socrates et al. (2012). As mentioned in Section 7.1, in light of the large uncertainties of the strength of tidal dissipation, for the integrations in this paper we assume various values of $t_{V,1}$ differing by various orders of magnitude compared to the value derived by (Socrates et al. 2012).

7.3 Verification of the secular method with N -body integrations

As mentioned in Section 7.1, the secular evolution of multiplanet systems can be chaotic, particularly when the number of planets is $N_p \geq 2$ (or, equivalently, $N > 3$). For direct N -body integrations, this implies that changing the initial orbital phases or the accuracy of the integration can lead to a completely different outcome after some time during which the eccentricities have changed by, say, order unity. This implies that it is not very meaningful to compare integrations from direct N -body integrations to integrations with `SECULARMULTIPLE`, in which the orbits are averaged over, on a one-to-one basis. However, our expectation is that `SECULARMULTIPLE` produces the correct secular dynamical evolution in a statistical sense, i.e. for an ensemble of systems.

In this section, we investigate this expectation by comparing `SECULARMULTIPLE` to integrations with the direct N -body code `ARCHAIN` (Mikkola & Merritt 2008). The latter code uses algorithmic chain regularisation to integrate the equations of motion with high precision, in particular systems with well defined hierarchies. In addition to relativistic corrections, tidal interactions are taken into account with the same model assumed for the secular integrations, i.e. the (non-averaged) equilibrium tide model (Hut 1981). For more details regarding the direct N -body code, we refer to Antonini et al. (in preparation).

7.3.1 Initial conditions

The `ARCHAIN` code was used to integrate 100 3-planet systems for ≈ 120 Myr. The stellar mass was set to $M_\star = 1 M_\odot$, and the planetary masses were assumed to be $m_1 = 3 \times 10^{-4} M_\odot \approx 0.314 M_J$ and $m_2 = m_3 = 1 \times 10^{-3} M_\odot \approx 1.05 M_J$. The initial semimajor axes were assumed

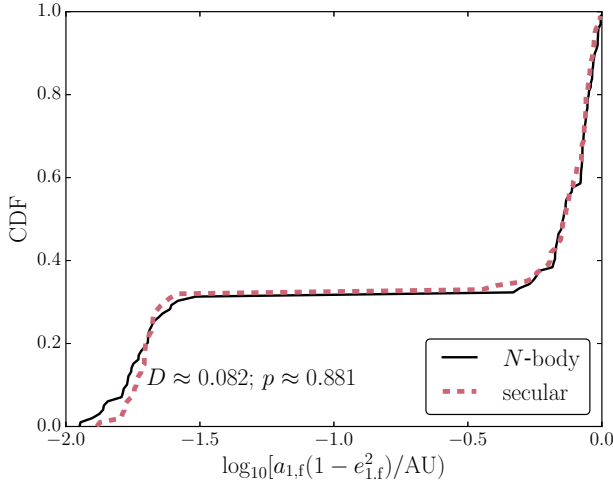


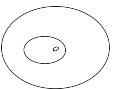
Figure 7.1: The CDF of the innermost planet’s orbit semilatus rectum, $r_{1,\text{slr},f} \equiv a_{1,f}(1 - e_{1,f}^2)$, after 120 Myr of evolution according to ARCHAIN (black solid line) and to SECULARMULTIPLE (red dashed line). The D and p statistics for the two-sided K-S test between the distributions are indicated in the panel.

to be $a_1 = 1$ AU, $a_2 = 6$ AU, and a_3 was varied between 12 and 62 AU with increments of 0.5 AU. The eccentricities and inclinations for all orbits were set equal to each other, and values were assumed ranging between ≈ 0.4 and 0.6 (with the inclinations measured in radians). The other orbital angles, ω_i and Ω_i , were sampled from linear distributions between 0 and 360° . The innermost planet was assumed to have a radius $R_1 = 1 R_J$, a viscous time-scale $t_{V,1} \approx 4.8$ yr, and an apsidal motion constant $k_{\text{AM},1} = 0.19$ (this corresponds to a time lag of $\tau = 0.66$ s, cf. equation 7.1).

7.3.2 Results

With ARCHAIN, a fraction of 0.32 of the systems became HJ systems by 120 Myr, i.e. the final semilatus rectum of the orbit of the innermost planet, $r_{1,\text{slr},f} \equiv a_{1,f}(1 - e_{1,f}^2)$, reached $r_{1,\text{slr},f} < 0.091$ AU, corresponding to the semimajor axis of a 10-d planet in a circular orbit around a Solar-mass star. Here, we use the semilatus rectum to define HJs, because after 120 Myr, some systems were still decaying tidally (i.e. a_1 and e_1 decreasing), while decoupled from the secular oscillations. In the latter case, when tidal evolution dominates, the final result (i.e. after $\gg 120$ Myr) is a circular orbit with semimajor axis $a_{1,f} = a_1(1 - e_1^2) = r_{1,\text{slr}}$. The HJ fraction after 120 Myr of evolution obtained with SECULARMULTIPLE is 0.33.

In Figure 7.1, we show the distribution of $r_{1,\text{slr},f}$ (i.e. after 120 Myr of evolution) in terms of the cumulative density function (CDF) according to ARCHAIN (black solid line) and according to SECULARMULTIPLE (red dashed line). There is a pileup of systems around $\log_{10}(r_{1,\text{slr},f}/\text{AU}) \approx -1.7$ corresponding to the HJ systems. The number of systems subsequently stalls at ≈ 0.32 with increasing $r_{1,\text{slr},f}$, until at $\log_{10}(r_{1,\text{slr},f}/\text{AU}) \approx -0.4$, this number increases again. The latter correspond to non- (or weakly) migrating planets; either their semimajor axes have de-



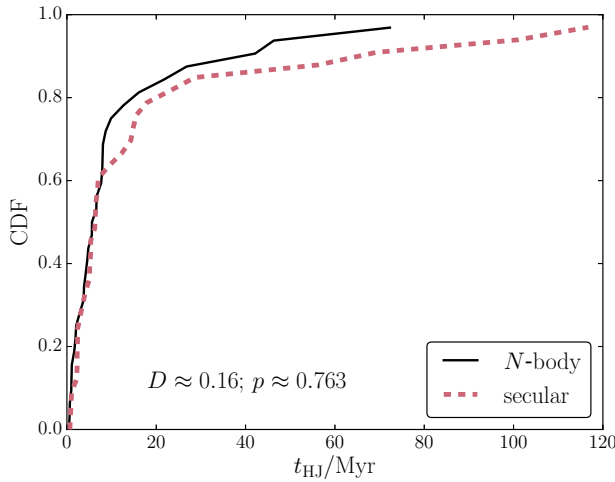


Figure 7.2: The CDF of the HJ formation times according to ARCHAIN (black solid line) and to SECULARMULTIPLE (red dashed line). K-S test statistics are indicated.

creased slightly due to tidal evolution, and/or their eccentricities are excited because of secular evolution.

Overall, in Figure 7.1 there is good agreement between ARCHAIN and SECULARMULTIPLE; the two-sided Kolmogorov-Smirnov (K-S; Kolmogorov 1933; Smirnov 1948) test statistics are $D \approx 0.082$ and $p \approx 0.881$. There is a tendency for SECULARMULTIPLE to slightly underestimate the smallest semilatus recta, or the smallest orbital periods of the HJs, compared to ARCHAIN. We do not believe that this discrepancy strongly affects our conclusions in the integrations below. In particular, when considering WJs, this implies that if no WJs are produced in the integrations with SECULARMULTIPLE, then it is unlikely that any would be produced in more accurate N -body calculations.

In Figure 7.2, we show another comparison between the two methods in terms of the distributions of the HJ formation times. With both methods, most (≈ 0.7) HJs are produced early in the evolution, i.e. within the first 10 Myr. The K-S test yields $D \approx 0.16$ and $p \approx 0.763$, showing that the two distributions are statistically consistent. Note, however, that the formation times in SECULARMULTIPLE after 10 Myr tend to be slightly longer.

We show in Figure 7.3 for both methods the normalised joint distributions of a_1 and e_1 for all combined output times and systems in the $(a_1, 1 - e_1)$ -plane (left panel: ARCHAIN; right panel: SECULARMULTIPLE). For reference, we also included a number of analytic curves, which are obtained using approximate analytic expressions for various time-scales associated with the secular evolution, and precession due to tidal bulges and GR (i.e. Schwarzschild precession, SP). For more details regarding these curves, we refer to Antonini et al. (in preparation).

Note that in Figure 7.3, the number of output times is not the same for both methods. In particular, in the case of SECULARMULTIPLE, not all tidally decaying tracks are resolved. Therefore, when comparing the two methods, one should consider primarily the relative distributions within the same $(a, 1 - e)$ -plane. Also note that the color coding is in logarithmic scale. Taking

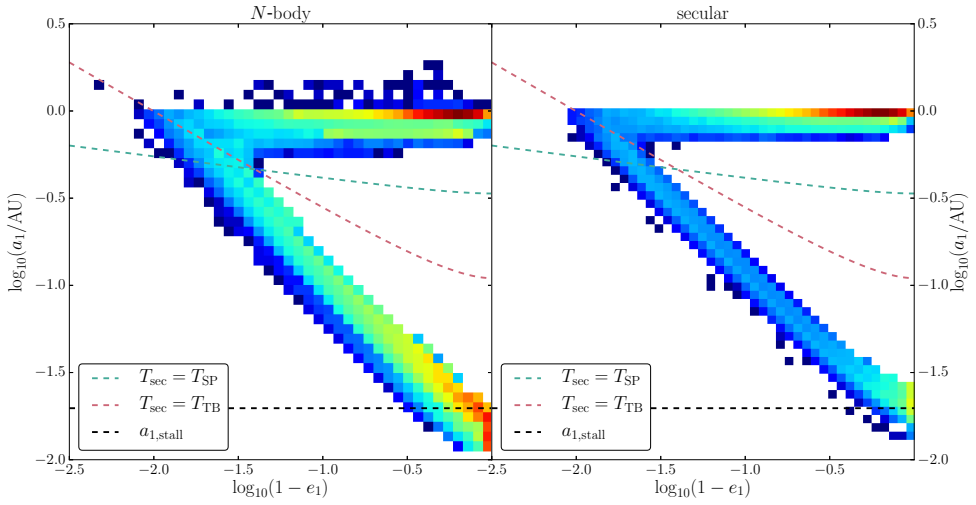
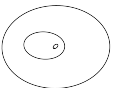


Figure 7.3: The normalised joint distributions of a_1 and e_1 for all combined output times and systems in the $(a_1, 1 - e_1)$ -plane, according to ARCHAIN (left panel) and SECULARMULTIPLE (right panel). The color coding is in log scale, with the lowest and highest occurrences shown with blue and red, respectively. The dashed lines show a number of analytic curves (cf. Antonini et al., in preparation) – the red dashed and green dashed lines show the conditions for which the secular time-scale is equal to the tidal bulges and SP time-scales, respectively. The black horizontal line shows an estimate of the stalling semimajor axis assuming quenching of secular oscillations by precession induced by tidal bulges.



this into account, the distributions in the $(a, 1 - e)$ -planes shown in Figure 7.3 look similar.

A distinct difference between the two methods appears at semimajor axes $a_1 > 1$ AU. In some of the direct N -body integrations, a_1 exceeds the initial $a_1 = 1$ AU. This increase can be attributed to strong (non-secular) interactions between the planets that affect the orbital energies, and which are not included in the secular method. However, as shown in the left panel of Figure 7.3, these non-secular interactions are not very important, given that these outwardly migrating excursions occur only in a few systems.

7.3.3 Orbit crossings

In multiplanet systems such as those considered here, it is possible that orbits cross at various moments in the evolution as the orbital eccentricities are excited, i.e. the pericentre of an outer orbit can become smaller than the apocentre of an inner orbit. For an adjacent orbit pair $(i, j) = (i, i + 1)$ with $a_i < a_j = a_{i+1}$, this condition is given by $K_{ij} \leq 0$, where the orbital separation K_{ij} (in units of the mutual Hill radius) is given by

$$K_{ij} \equiv \frac{a_j(1 - e_j) - a_i(1 + e_i)}{R_{H;ij}}, \quad (7.2)$$

with

$$R_{H;ij} \equiv \frac{a_i + a_j}{2} \left(\frac{m_i + m_j}{3M_\star} \right)^{1/3} \quad (7.3)$$

the mutual Hill radius. Note that this condition does not necessarily imply physical intersections of the orbits. Regardless, if $K_{ij} \leq 0$, then the system can no longer be considered as being hierarchical. Strictly speaking, this implies that the assumptions on which the secular method is based, break down.

The `SECULARMULTIPLE` code includes a stopping condition for orbit crossings. However, when enabling this condition, in many of the systems that would otherwise produce HJs, the integration is terminated at an early moment in the evolution due to orbit crossings. Consequently, the results from the secular code are no longer in agreement with the direct N -body integrations.

In order to obtain useful results, orbit crossing stopping conditions were disabled in the integrations with `SECULARMULTIPLE` above. This can be justified a posteriori by considering the consistency of the results of the integrations between the direct and secular methods shown above. Also, when the orbit crossing stopping condition is disabled, the condition of orbit crossing occurred for a fraction of ≈ 0.00016 and ≈ 0.021 of all output times for orbit pairs (1, 2) and (2, 3), respectively. These fractions are consistent with `ARCHAIN`; in the latter code, the orbit crossing fractions were ≈ 0.0002 and ≈ 0.015 , for orbit pairs (1, 2) and (2, 3), respectively.

However, it is important to consider that the above conclusions regarding the agreement between the secular and direct N -body methods, in particular without orbit crossing stopping conditions, only apply to a limited set of systems. Generally, it should not be assumed a priori that the secular method still gives accurate results when orbit crossings occur. Apart from the breakdown of the expansion in terms of separation ratios, orbit crossings might be followed by short-term instabilities leading to unstable configurations and ejections and/or collisions of planets, which are not taken into account in the secular method. Although such instabilities are not important for the evolution in the systems considered here, this could be different in other situations.

In order to investigate these cases, `SECULARMULTIPLE` allows for convenient switching to any of the direct N -body codes available in `AMUSE`. This method can e.g. be used to evaluate the short-term stability at the onset of orbit crossings, and we will apply it in several cases in the secular integrations in subsequent sections. However, a detailed investigation into the conditions for which the secular approach breaks down (in particular, as a function of K_{ij}), is beyond the scope of this paper.

7.4 Conditions for producing high eccentricities due to secular evolution

It was shown previously by Wu & Lithwick (2011); Lithwick & Wu (2011, 2014), that high eccentricities can be induced in multiplanet systems, particularly in the orbit of the inner planet, if the angular-momentum deficit (AMD) is sufficiently large. Here, the AMD is defined as

$$\text{AMD} = \sum_{i=1}^{N_p} m_i \sqrt{a_i} \left[1 - \sqrt{1 - e_i^2} \cos(i_{\text{var},i}) \right], \quad (7.4)$$

where $i_{\text{var},i}$ is the inclination with respect to the invariable plane, i.e. the plane perpendicular to the total orbital angular-momentum vector of the system. Physically, the AMD is the difference in the z -component of the total orbital angular-momentum vector, $L_{\text{tot},z}$, between the case of either coplanar and circular orbits (in which case $L_{\text{tot},z}$ is maximal), and the actual configuration. Here, the orbital angular-momentum vectors are defined with respect to the central star, which is assumed to be much more massive than the other bodies. Note that in equation (7.4), the AMD is defined as this difference in $L_{\text{tot},z}$ in units of $\sqrt{GM_\star}$.

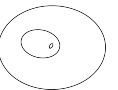
In terms of the AMD, high eccentricities and/or chaotic motion can be achieved if $\text{AMD} \gtrsim m_1 \sqrt{a_1}$ (Wu & Lithwick 2011; Lithwick & Wu 2011, 2014). This can be understood broadly from equation (7.4) by requiring that $e_1 \rightarrow 1$. In the minimum extreme case, $e_1 \rightarrow 1$ corresponds to $e_i \rightarrow 0$ and $i_{\text{var},i} \rightarrow 0$ for all other orbits ($i > 1$), i.e. all angular momentum has been transferred from the inner orbit to the outer orbits. In this case, $1 - \sqrt{1 - e_1^2} \cos(i_{\text{var},1}) = 1$ and $1 - \sqrt{1 - e_i^2} \cos(i_{\text{var},i}) = 0$ for $i > 1$, implying that the $\text{AMD} = m_1 \sqrt{a_1}$.

In this section, we investigate in more detail the conditions for exciting high eccentricities through secular evolution, and, therefore, potentially producing HJs, using numerical integrations with `SECULARMULTIPLE`. We will not only investigate the dependence on the AMD, but also on another quantity related to the periods of LK oscillations applied to different pairs of orbits.

7.4.1 Dependence of the maximum eccentricities on the AMD with varying eccentricities and inclinations

Initial conditions

Firstly, we consider the dependence of the maximum eccentricities attained as a function of the AMD with varying eccentricities and inclinations, whereas the semimajor axes are kept fixed. Depending on the number of planets ($N_p = 3, 4$ or 5), the latter were assumed to be $a_1 = 1$, $a_2 = 6$, $a_3 \approx 12.84$, $a_4 \approx 28.98$ and $a_5 \approx 62.00$ AU, respectively (in relation to Section 7.4.2 and equation 7.5 in particular, these choices correspond to $\mathcal{R}_0^{(123)} \approx 1.5$, $\mathcal{R}_0^{(234)} \approx 0.27$ and



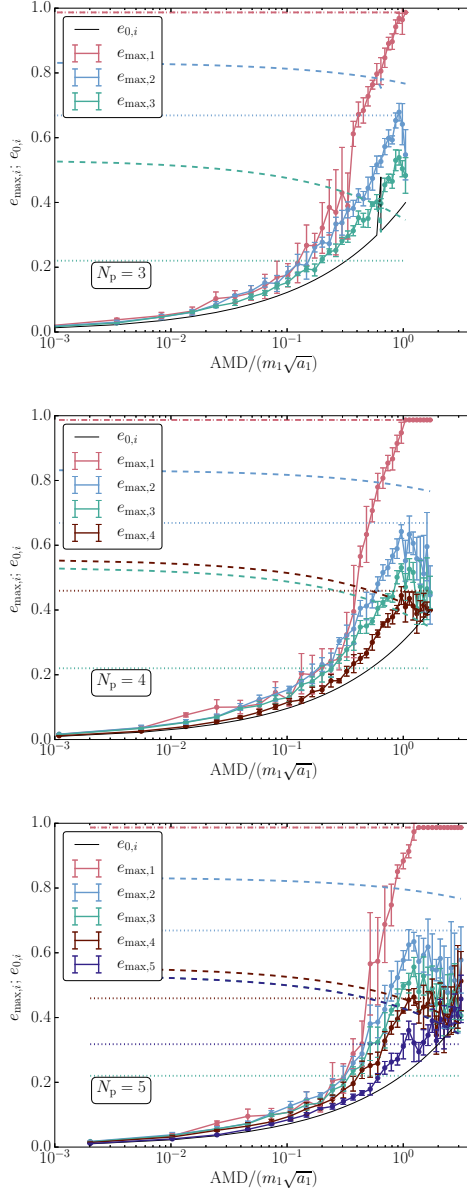


Figure 7.4: The maximum eccentricities attained in the N_p orbits as a function of the AMD, for $N_p = 3$, 4 and 5 in the top, middle and bottom panels, respectively (cf. Section 7.4.1). The errorbars indicate the standard deviation of the spread in $e_{\max,i}$ due to different initial values of the ω_i . The black solid lines show the initial values $e_{0,i}$. For each orbit, two limiting eccentricities related to tidal disruption or orbit crossing are shown with colored dashed and dotted lines, the color corresponding to each orbit. For the innermost orbit, these lines correspond to the tidal disruption of the planet (assuming equation 7.7). For other orbits i , the dotted line corresponds to orbit crossing of orbit i with orbit $i - 1$ ($K_{ij} = 0$; cf. equation 7.2), assuming e_{i-1} is equal to the value corresponding to disruption (if $i - 1 = 1$), or orbit crossing of orbit $i - 1$ with orbit $i - 2$ (otherwise). The dashed lines correspond to orbit crossing of orbit i with orbit $i - 1$, assuming $e_{i-1} = e_{0,i}$ is equal to the initial value (cf. the black solid lines).

$\mathcal{R}_0^{(345)} \approx 0.35$). In each simulation, the initial eccentricities e_i and the inclinations i_i (with respect to an arbitrary inertial reference frame) of all the planets were set to the same value $e_i = i_i = e_{0,i}$, and values were considered of $e_{0,i} = 0.01$ to 0.5 (with the inclinations measured in radians), corresponding to a range of AMDs of $\approx 10^{-3}$ to ≈ 5 (depending on the number of planets). The planetary masses were set to $m_i = 1 M_J$, and the stellar mass was $M_\star = 1 M_\odot$. The longitudes of the ascending nodes were fixed for each system, assuming random values between 0 and 2π for the N_p orbits. For each value of $e_{0,i}$, five realisations were carried out with different arguments of pericenter ω_i in order to take into account any possible sensitivities on the initial conditions due to chaos.

Because we are interested in the dependence of the maximum eccentricities as a function of the physical parameters due to secular evolution alone, tidal friction was not included in the integrations here. The integration time for each system was set to $t_{\text{end}} = 1$ Gyr. Note that due to the typically chaotic nature of the evolution (see e.g. Figures 7.7 and 7.10), it is possible that the maximum eccentricity attained is larger over a longer time-scale. Here, our interest is to compare the relative effects of the parameters for the same integration time. In Section 7.5, we consider longer integration times of up to 10 Gyr.

The integrations were stopped if the innermost planet was tidally disrupted by the star (cf. equation 7.7, with $\eta = 2.7$). Orbit crossing stopping conditions were not imposed; see Section 7.4.1 for a discussion on orbit crossings.

Results

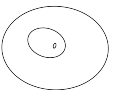
In Figure 7.4, we show the maximum eccentricities attained in the N_p orbits as a function of the AMD, for $N_p = 3, 4$ and 5 in the top, middle and bottom panels, respectively. The errorbars indicate the standard deviation of the spread in $e_{\text{max},i}$ due to different initial values of the ω_i . The latter are therefore expected to be a measure for the chaotic nature of the system. As expected, the maximum eccentricities increase with increasing AMD, and high values of $e_{\text{max},1}$ in particular are reached for $\text{AMD} \gtrsim m_1 \sqrt{a_1}$. The higher the orbit number, the lower $e_{\text{max},i}$; generally, $e_{\text{max},i+1} < e_{\text{max},i}$.

The significance of orbit crossings

In Figure 7.4, we show for all orbits $i > 1$ two curves (colored dashed and dotted lines) corresponding to orbit crossings with orbit $i - 1$. The dotted line corresponds to orbit crossing of orbit i with orbit $i - 1$ assuming e_{i-1} is equal to the value corresponding to tidal disruption (if $i - 1 = 1$), or orbit crossing of orbit $i - 1$ with orbit $i - 2$ (otherwise). The dashed lines correspond to orbit crossing of orbit i with orbit $i - 1$, assuming $e_{i-1} = e_{0,i}$ is equal to the initial value (cf. the black solid lines). The actual moment of orbit crossing lies between these two extreme cases.

For large AMDs, the outermost orbits, in particular, cross. To investigate these cases, we carried out a number of direct N -body integrations starting from the onset of the first crossing of any orbit pair. We integrated the system for 10^{-2} Myr after this time, and carried out different realisations with different orbital phases (sampled randomly from a uniform distribution of the mean anomaly), in order to gain insight into the sensitivity on the initial conditions.

In Figure 7.5, we show for $N_p = 3$ the evolution of the semimajor axes (top panels), eccentricities (middle panels) and the separations K_{ij} (bottom panels), for two values of the AMD



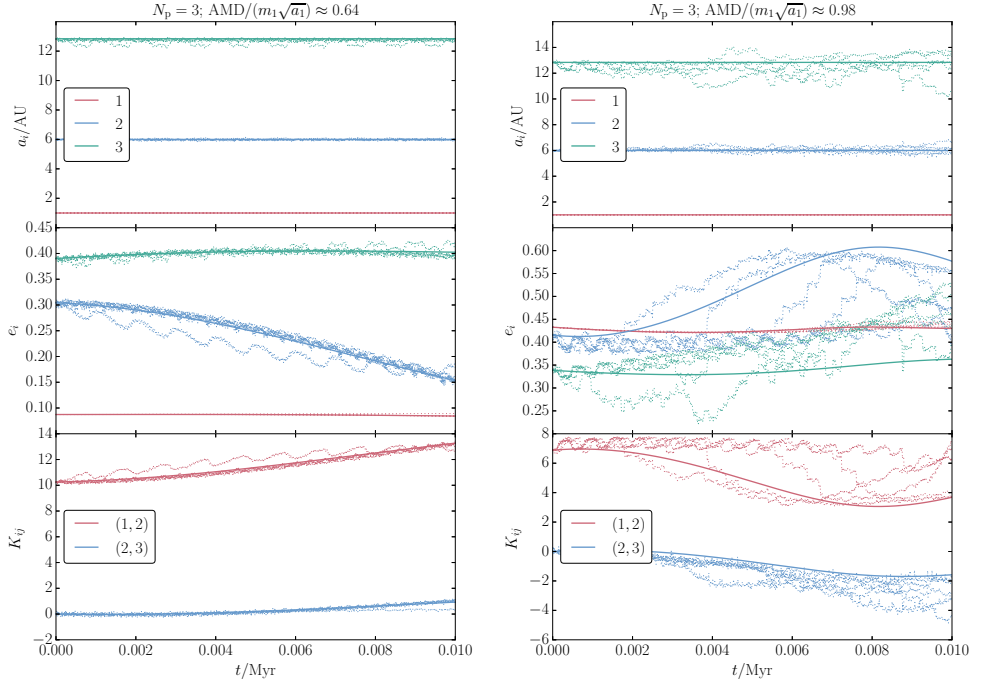


Figure 7.5: The evolution of the semimajor axes (top panels), the eccentricities (middle panels) and the orbital separations K_{ij} (bottom panels; cf. equation 7.2), for direct (dotted lines) and secular (solid lines) integrations, following the onset of orbit crossing in any of the orbits in the set of integrations shown in Figure 7.4. The left and right columns correspond to two values of the AMD with $N_p = 3$, indicated at the top. In the case of the direct N -body integrations, several realisations are shown.

(left and right columns). Dotted lines show results according to the various direct N -body realisations; solid lines show the evolution according to `SECULARMULTIPLE`. In the left column, $\text{AMD} \approx 0.64 m_1 \sqrt{a_1}$, orbits 2 and 3 initially crossed; by 0.01 Myr, they are no longer crossing ($K_{23} > 0$), which is the case in both the secular and direct integrations. The semimajor axes remain nearly constant in the direct integrations, showing that the system remains stable during the integration, and no ejections and/or collisions occur. For most N -body realisations, the secular and direct results agree within a few per cent.

In the right column of Figure 7.5, the AMD is larger, $\text{AMD} \approx 0.98 m_1 \sqrt{a_1}$, and therefore the maximum attained eccentricities are larger as well. In the direct integrations, the interactions are clearly stronger, which is reflected by the larger variations in the semimajor axes, and the larger deviations in the semimajor axes and eccentricities between the different realisations. Despite the stronger interactions which are of non-secular nature (i.e. which also affect the semimajor axes), the secular method still appears to give a approximation of the long-term evolution averaged over the various direct N -body realisations.

Similar results also apply to the simulations with $N_p = 4$ and 5: the orbits remain stable and no longer cross after ~ 0.01 Myr. Also, the secular method approximately captures the long-term evolution even after the moment of orbit crossing.

These integrations show that, for the systems considered here, orbit crossings do not necessarily imply a complete breakdown of the secular method. Rather, comparison of the left and right columns of Figure 7.5 suggests that the greater the ‘orbit crossing penetration’, i.e. the smaller K_{ij} , the larger the discrepancies between the secular and direct methods. As mentioned before, a detailed investigation into this phenomenon is beyond the scope of this paper.

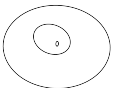
7.4.2 Dependence of the maximum eccentricities on the semimajor axes ratios

In Section 7.4.1, we showed that there is a strong dependence of the maximum eccentricities on the AMD, with high eccentricities reached for $\text{AMD} \gtrsim m_1 \sqrt{a_1}$. Here, we show that there is another condition for producing high eccentricities, which is related, to a strong degree, to the semimajor axes of the planets, and which, to our knowledge, has not been pointed out in this context before.

Specifically, we consider the dependence of the maximum eccentricities on the ratio $\mathcal{R}_0^{(ijk)}$ of the LK time-scales associated with the orbit pairs (i, j) and (j, k) . A *three*-planet system can be considered as a hierarchical quadruple system in the ‘fully nested’ configuration, with three hierarchical orbits corresponding to planets 1-3. As shown by Hamers et al. (2015), for such systems, chaotic motion can be induced if the LK time-scales of the (1,2) and (2,3) pairs are comparable. For (more general) fully nested hierarchical N -body systems, this condition is equivalent to comparing the LK time-scales of three adjacent orbits (Hamers & Portegies Zwart 2016). Therefore, for three adjacent (non-intersecting) nested orbits (i, j, k) with semimajor axes $a_i < a_j < a_k$, we consider the ratio

$$\mathcal{R}_0^{(ijk)} \equiv \frac{P_{\text{LK},ij,0}}{P_{\text{LK},jk,0}} \sim \left(\frac{a_j^3}{a_i a_k^2} \right)^{3/2} \frac{m_k}{m_j} \left(\frac{1 - e_{0,j}^2}{1 - e_{0,k}^2} \right)^{3/2}, \quad (7.5)$$

where $P_{\text{LK},ij,0}$ is the (initial) LK time-scale associated with orbits i and j , and $e_{0,i}$ is the initial eccentricity of orbit i . Here, we assumed that $m_i \ll M_*$.



To investigate the dependence on $\mathcal{R}_0^{(ijk)}$, we carried out a suite of integrations with `SECULARMULTIPLE` with the semimajor axes of the first two planets fixed at $a_1 = 1$ AU and $a_2 = 6$ AU. The semimajor axes of the other planets were generated assuming a linear spacing in $\log_{10}(\mathcal{R}_0^{(ijk)})$, and with the condition that the initial semimajor axis ratio of adjacent orbits was $a_j/a_i > 2$. Approximate ranges of the semimajor axes are indicated in Table 7.1. The initial eccentricities and inclinations were fixed at $e_i = i_i = e_{0,i}$ (with the inclinations measured in radians), with either $e_{0,i} = 0.3$ or $e_{0,i} = 0.4$. As in Section 7.4.1, $m_i = 1 M_J$ and $M_\star = 1 M_\odot$ were assumed, the longitudes of the ascending nodes were fixed for each system assuming random values between 0 and 2π , and for each semimajor axis combination, five realisations were carried out with different arguments of pericenter ω_i . With these simplified initial conditions, $\mathcal{R}_0^{(ijk)}$ is given by the simple relation

$$\mathcal{R}_0^{(ijk)} \sim \left(\frac{a_j^3}{a_i a_k^2} \right)^{3/2}. \quad (7.6)$$

Three planets

In Figure 7.6, we show for $N_p = 3$ the dependence of the maximum eccentricities in orbits 1, 2 and 3 as a function of $\mathcal{R}_0^{(123)}$ with the red, blue and green lines, respectively. The initial eccentricities and inclinations are $e_{0,i} = 0.3$ and $e_{0,i} = 0.4$ in the top and bottom panels, respectively. Note that the largest value of $\mathcal{R}_0^{(123)}$ in our integrations is set by the smallest value of a_3 , $a_3 \approx 12.31$ AU (cf. equation 7.5), and is $\mathcal{R}_0^{(123)} \approx (6^3/12.31^2)^{3/2} \approx 1.7$.

For $\mathcal{R}_0^{(123)} \lesssim 10^{-1}$, there is a mild excitation of e_1 with $e_{\max,1} \sim 0.5$ and $e_{\max,1} \sim 0.6$ for the respective values of $e_{0,i}$, and little excitation of e_2 and e_3 . An example of the eccentricities and inclinations as a function of time (for $e_{0,i} = 0.3$) is given in the top row of Figure 7.7, in which case $\mathcal{R}_0^{(123)} \approx 0.01$. The inner orbit eccentricity oscillates on a time-scale of ~ 0.7 Myr, with an amplitude of ≈ 0.2 ; the intermediate orbit eccentricity e_2 oscillates with a slightly lower amplitude of ≈ 0.15 . The mutual inclinations (shown in Figure 7.7 in the second column) do not exceed $\approx 35^\circ$.

For $e_{0,i} = 0.3$ in Figure 7.6, around $\mathcal{R}_0^{(123)} \approx 0.3$, the maximum eccentricities of all orbits increase strongly with $\mathcal{R}_0^{(123)}$, with values of ≈ 0.8 , 0.5 and 0.4 for orbits 1, 2 and 3, respectively. In the region $0.3 \lesssim \mathcal{R}_0^{(123)} \lesssim 1$, there is no strong dependence of the maximum eccentricities on $\mathcal{R}_0^{(123)}$. An example illustrating this regime of $\mathcal{R}_0^{(123)}$ is given in the bottom row of Figure 7.7, where $\mathcal{R}_0^{(123)} \approx 1.0$. The behaviour of all three eccentricities is more erratic and less repetitive compared to the case $\mathcal{R}_0^{(123)} \approx 0.01$. In particular e_1 shows temporary excursions to lower and higher values on time-scales of ~ 100 Myr, suggestive of chaotic evolution. In contrast to the case $\mathcal{R}_0^{(123)} \approx 0.01$, the outermost orbit eccentricity e_3 no longer remains (very nearly) constant at $e_3 \approx 0.3$, also but also oscillates erratically with an amplitude of ≈ 0.3 . The mutual inclinations in this case reach values of $\approx 65^\circ$.

The AMDs are shown as a function of $\mathcal{R}_0^{(123)}$ in Figure 7.6 with the black solid lines (right-hand axes). For $\mathcal{R}_0^{(123)} = 10^{-2}$ and $e_{0,i} = 0.3$ ($e_{0,i} = 0.4$), the AMD is $\approx 0.9 m_1 \sqrt{a_1}$ ($\approx 1.6 m_1 \sqrt{a_1}$). Although the AMD decreases with increasing $\mathcal{R}_0^{(123)}$ (the latter corresponds

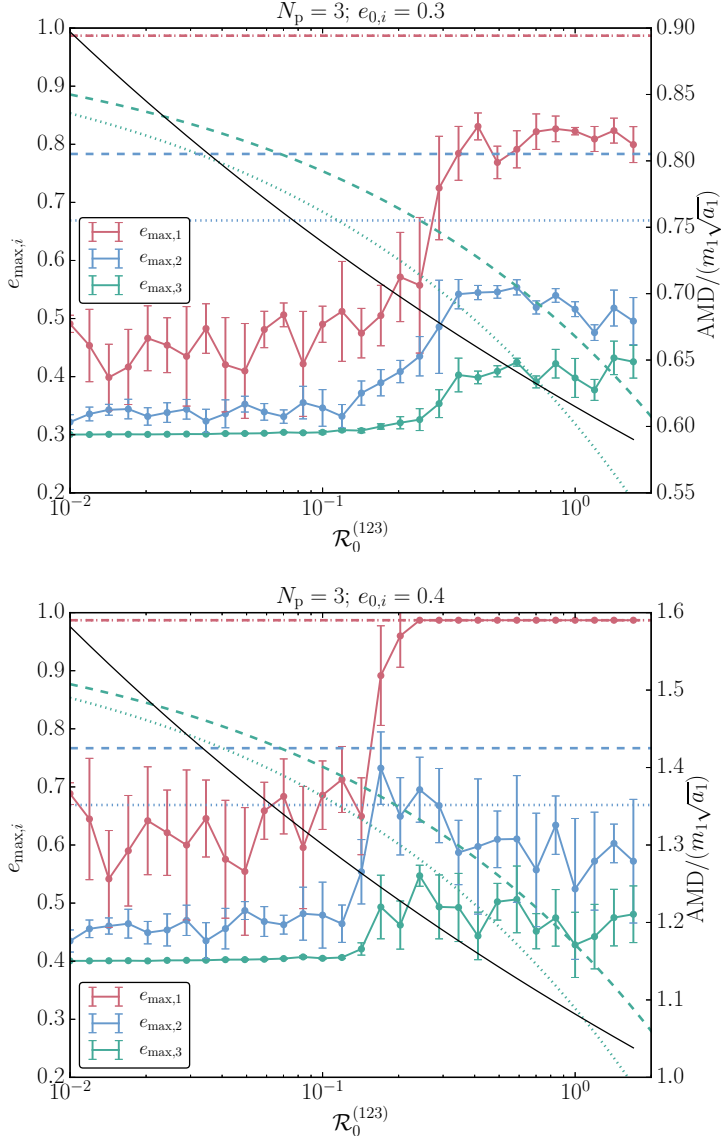
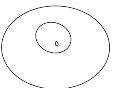


Figure 7.6: Dots and solid colored lines: the maximum eccentricities of the orbits of three-planet systems as a function of $\mathcal{R}_0^{(123)}$ (cf. equation 7.5), for the secular integrations described in Section 7.4. Quantities pertaining to orbits 1, 2 and 3 are depicted with red, blue and green colors, respectively. The black solid line shows the AMD (right-hand axis). Top panel: assuming initial eccentricities and inclinations $e_{0,i} = i_{0,i} = 0.3$; bottom panel: $e_{0,i} = i_{0,i} = 0.4$. Similarly to Figure 7.4, for each orbit, two limiting eccentricities related to tidal disruption or orbit crossing are shown with colored dashed and dotted lines, the color corresponding to each orbit. For the innermost orbit, these lines correspond to the tidal disruption of the planet (assuming equation 7.7). For other orbits i , the dotted line corresponds to orbit crossing of orbit i with orbit $i - 1$ ($K_{ij} = 0$; cf. equation 7.2), assuming e_{i-1} is equal to the value corresponding to disruption (if $i - 1 = 1$), or orbit crossing of orbit $i - 1$ with orbit $i - 2$ (otherwise). The dashed lines correspond to orbit crossing of orbit i with orbit $i - 1$, assuming $e_{i-1} = e_{0,i}$ is equal to the initial value.



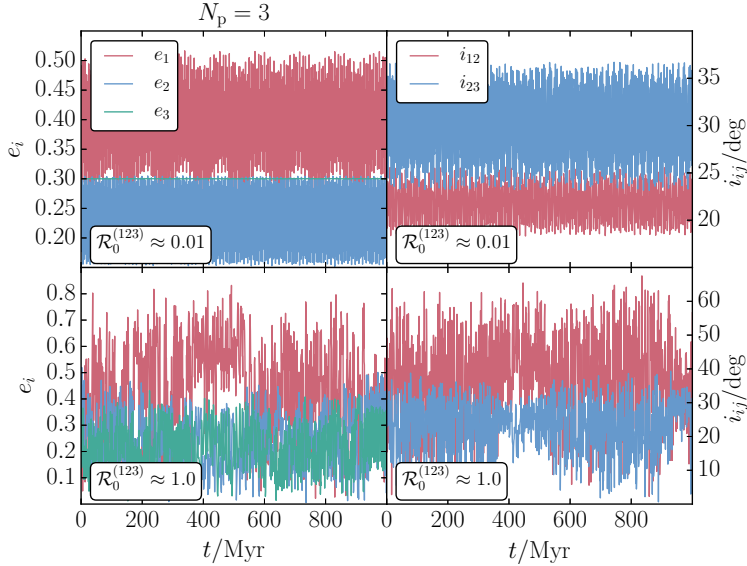


Figure 7.7: Two example integrations (left column: the eccentricities e_i ; right column: the mutual inclinations i_{ij} between two orbits i and j), taken from the systems included in Figure 7.6 with $e_{0,i} = 0.3$. Top row: $\mathcal{R}_0^{(123)} \approx 0.01$; bottom row: $\mathcal{R}_0^{(123)} \approx 1.0$.

to smaller a_3 whereas a_1 and a_2 are constant, implying a smaller AMD), the maximum eccentricities *increase*. This shows that the AMD does not give a complete condition for high eccentricities due to secular evolution. In fact, eccentricity excitation to $e_{\max,1} \approx 0.8$ is possible for $e_{0,i} = 0.3$ and $\text{AMD}/(m_1\sqrt{a_1}) < 1$, provided that $\mathcal{R}_0^{(123)}$ is close to unity. The AMD is, nevertheless, an important quantity. This is illustrated in the bottom panel of Figure 7.6, where $e_{0,i} = 0.4$ and the AMD is $> m_1\sqrt{a_1}$. Here, the eccentricities are more highly excited compared to the case $e_{0,i} = 0.3$, with values of $e_{\max,1}$ reaching values high enough for the tidal disruption of the innermost planet (cf. the horizontal red dashed line).

Four planets

Results for $N_p = 4$ planets and $e_{0,i} = 0.3$ are shown in Figure 7.8. In each panel, we show the maximum eccentricities of the four orbits as a function of $\mathcal{R}_0^{(234)}$, assuming various values of $\mathcal{R}_0^{(123)}$ (plotted with different colors, as indicated in the legend at the bottom of the figure). Depending on $\mathcal{R}_0^{(123)}$, the eccentricities of the innermost orbit can be higher compared to the case $N_p = 3$. For small $\mathcal{R}_0^{(123)}$, $\log_{10}(\mathcal{R}_0^{(123)}) \approx -2.0$ and ≈ -1.25 , there is no strong excitation of e_1 , irrespective of the value of $\mathcal{R}_0^{(234)}$. When $\mathcal{R}_0^{(123)}$ is larger, $\log_{10}(\mathcal{R}_0^{(123)}) \approx -0.5$ and ≈ 0.25 , $e_{\max,1} \approx 0.8$ is similar to the value that was obtained for $N_p = 3$ and with $\mathcal{R}_0^{(123)} \sim 1$ if $\mathcal{R}_0^{(234)} \sim 10^{-2}$ is small. The latter corresponds to a relatively large value of a_4 , i.e. the semi-major axis of the fourth planet is too large to affect the innermost orbit (for the integration time

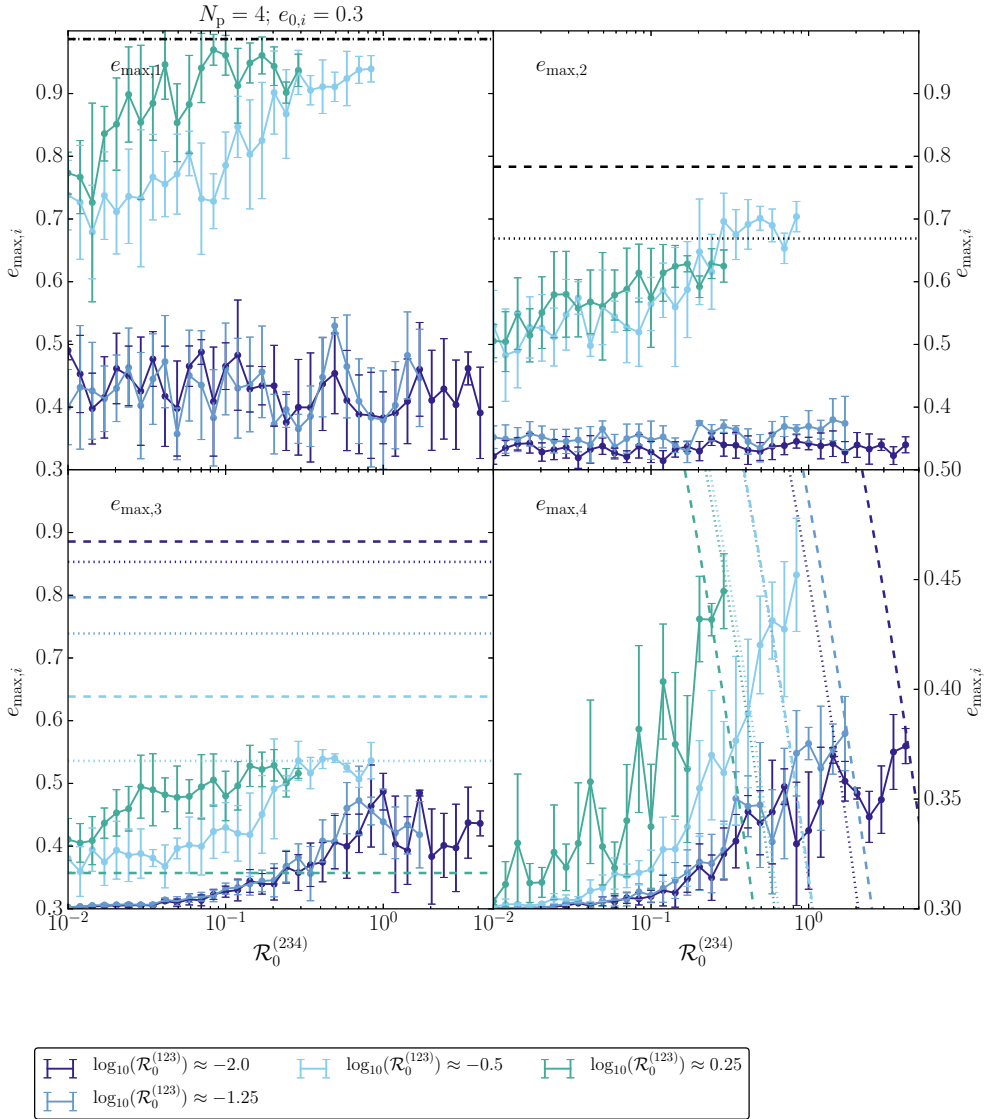


Figure 7.8: For $N_p = 4$ and $e_{0,i} = 0.3$, the maximum eccentricities of the orbits as a function of $\mathcal{R}_0^{(234)}$, assuming various values of $\mathcal{R}_0^{(123)}$ (plotted with different colors, as indicated in the legend at the bottom of the figure). Each panel corresponds to a different orbit, indicated in the top-left corner. As in Figure 7.6, for each orbit, two limiting eccentricities related to tidal disruption or orbit crossing are shown with colored dashed and dotted lines, the color corresponding to each orbit. For the innermost orbit, these lines correspond to the tidal disruption of the planet (assuming equation 7.7). For other orbits i , the dotted line corresponds to orbit crossing of orbit i with orbit $i-1$ ($K_{ij} = 0$; cf. equation 7.2), assuming e_{i-1} is equal to the value corresponding to disruption (if $i-1 = 1$), or orbit crossing of orbit $i-1$ with orbit $i-2$ (otherwise). The dashed lines correspond to orbit crossing of orbit i with orbit $i-1$, assuming $e_{i-1} = e_{0,i}$ is equal to the initial value.

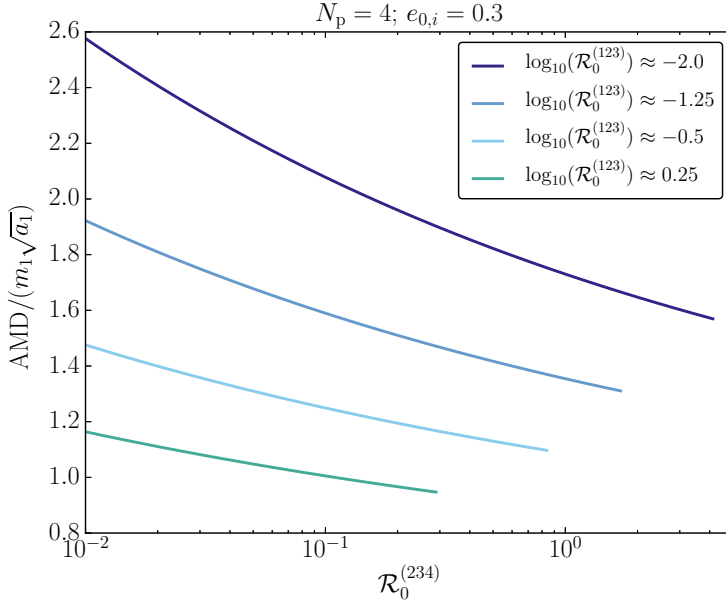


Figure 7.9: For $N_p = 4$ and $e_{0,i} = 0.3$, the AMD as a function of $\mathcal{R}_0^{(234)}$ for the various values of $\mathcal{R}_0^{(123)}$ included in Figure 7.8.

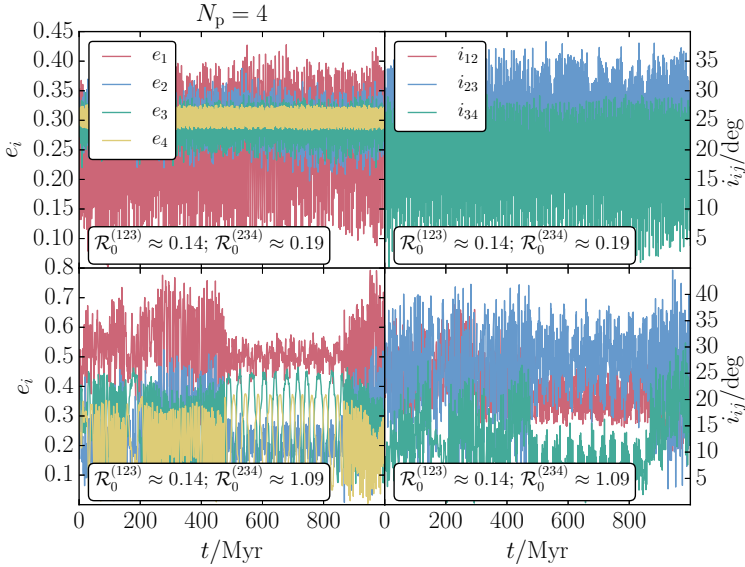


Figure 7.10: Two example integrations (left column: the eccentricities e_i ; right column: the mutual inclinations i_{ij} between two orbits i and j), taken from the systems included in Figure 7.8. Top row: $\mathcal{R}_0^{(123)} \approx 0.14$ and $\mathcal{R}_0^{(234)} \approx 0.19$; bottom row: $\mathcal{R}_0^{(123)} \approx 0.14$ and $\mathcal{R}_0^{(234)} \approx 1.09$.

of 1 Gyr). As $\mathcal{R}_0^{(234)}$ is increased, $e_{\max,1}$ typically increases to ≈ 0.98 , approaching the value for tidal disruption.

The dependence on $\mathcal{R}_0^{(234)}$ can be understood qualitatively by noting that if $\mathcal{R}_0^{(234)} \sim 1$, the second orbital eccentricity e_2 tends to be excited due to the commensurability of the LK time-scales associated with orbit pairs (2, 3) and (3, 4). This is shown in the top-right panel of Figure 7.8, where $e_{\max,2}$ increases as $\mathcal{R}_0^{(234)}$ approaches unity. Consequently, as e_2 is excited, this results in larger excitations of e_1 .

The quantity $\mathcal{R}_0^{(234)}$ also has a discernible effect on the third and fourth orbital eccentricities. The latter quantities are excited above their initial values of 0.3, reaching values as high as $e_{\max,3} \approx 0.55$ and $e_{\max,4} \approx 0.45$, respectively, when both $\mathcal{R}_0^{(123)}$ and $\mathcal{R}_0^{(234)}$ are close to unity.

Two examples for $e_{0,i} = 0.3$ are shown in Figure 7.10, where $\mathcal{R}_0^{(123)} \approx 0.14$ and $\mathcal{R}_0^{(234)} \approx 0.19$ (top row) and $\mathcal{R}_0^{(123)} \approx 0.14$ and $\mathcal{R}_0^{(234)} \approx 1.09$ (bottom row). In the top row, $\mathcal{R}_0^{(123)}$ and $\mathcal{R}_0^{(234)}$ are not close to unity, and there is no strong excitation of the innermost orbital eccentricity ($e_{\max,1} \approx 0.4$). In the bottom row, although the value of $\mathcal{R}_0^{(123)}$ is the same, owing to a value of $\mathcal{R}_0^{(234)}$ which is near unity, there is more significant excitation of e_1 with $e_{\max,1} \approx 0.8$, and generally more erratic behaviour of all eccentricities and mutual inclinations as a function of time.

The AMD is shown as a function of $\mathcal{R}_0^{(234)}$ for the various values of $\mathcal{R}_0^{(123)}$ in Figure 7.9. Similarly to the case $N_p = 3$, the AMD decreases with increasing $\mathcal{R}_0^{(123)}$ and $\mathcal{R}_0^{(234)}$, yet the maximum eccentricities increase because the commensurability conditions of the LK time-scales are approached more closely. In particular, note that the AMD is largest if $\log_{10}(\mathcal{R}_0^{(123)}) \approx -2.0$, whereas there is no large excitation of e_1 (for any value of $\mathcal{R}_0^{(234)}$).

Five planets

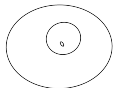
In Figure 7.11, we show the maximum eccentricities for $N_p = 5$ and $e_{0,i} = 0.3$ as a function of $\mathcal{R}_0^{(345)}$, for various values of $\mathcal{R}_0^{(123)}$ and $\mathcal{R}_0^{(234)}$. Each color corresponds to a different value of $\mathcal{R}_0^{(123)}$, whereas the various symbols represent different values of $\mathcal{R}_0^{(234)}$. The corresponding AMDs are shown as a function of $\mathcal{R}_0^{(345)}$ in Figure 7.12.

The largest values of $e_{\max,1}$ (corresponding to tidal disruption) are attained for $\log_{10}(\mathcal{R}_0^{(123)}) \approx 0.25$ and $\log_{10}(\mathcal{R}_0^{(234)}) \approx -0.5$, and $\mathcal{R}_0^{(345)} \gtrsim 0.3$. Such high values were not obtained for $N_p = 4$ for similar values of $\mathcal{R}_0^{(123)}$ and $\mathcal{R}_0^{(234)}$, again showing that the innermost orbit can be further excited with the addition of more planets.

Nevertheless, to attain high values of $e_{\max,1}$, $\mathcal{R}_0^{(123)}$ and $\mathcal{R}_0^{(234)}$ should be close to unity. The most sensitive dependence is clearly on $\mathcal{R}_0^{(123)}$. For example, for $\log_{10}(\mathcal{R}_0^{(123)}) \approx -2.0$ (dark blue symbols), $e_{\max,1}$ is no larger than ≈ 0.5 regardless of $\mathcal{R}_0^{(234)}$ and $\mathcal{R}_0^{(345)}$, whereas for $\log_{10}(\mathcal{R}_0^{(234)}) \approx -2.0$ (triads of any color), $e_{\max,1} \sim 0.8$ provided that $\log_{10}(\mathcal{R}_0^{(234)}) \approx 0.25$.

7.4.3 Tidal disruption versus tidal dissipation

In the previous sections, we considered the conditions for exciting high eccentricities in the inner orbit due to secular evolution (including precession due to tidal bulges, rotation and rela-



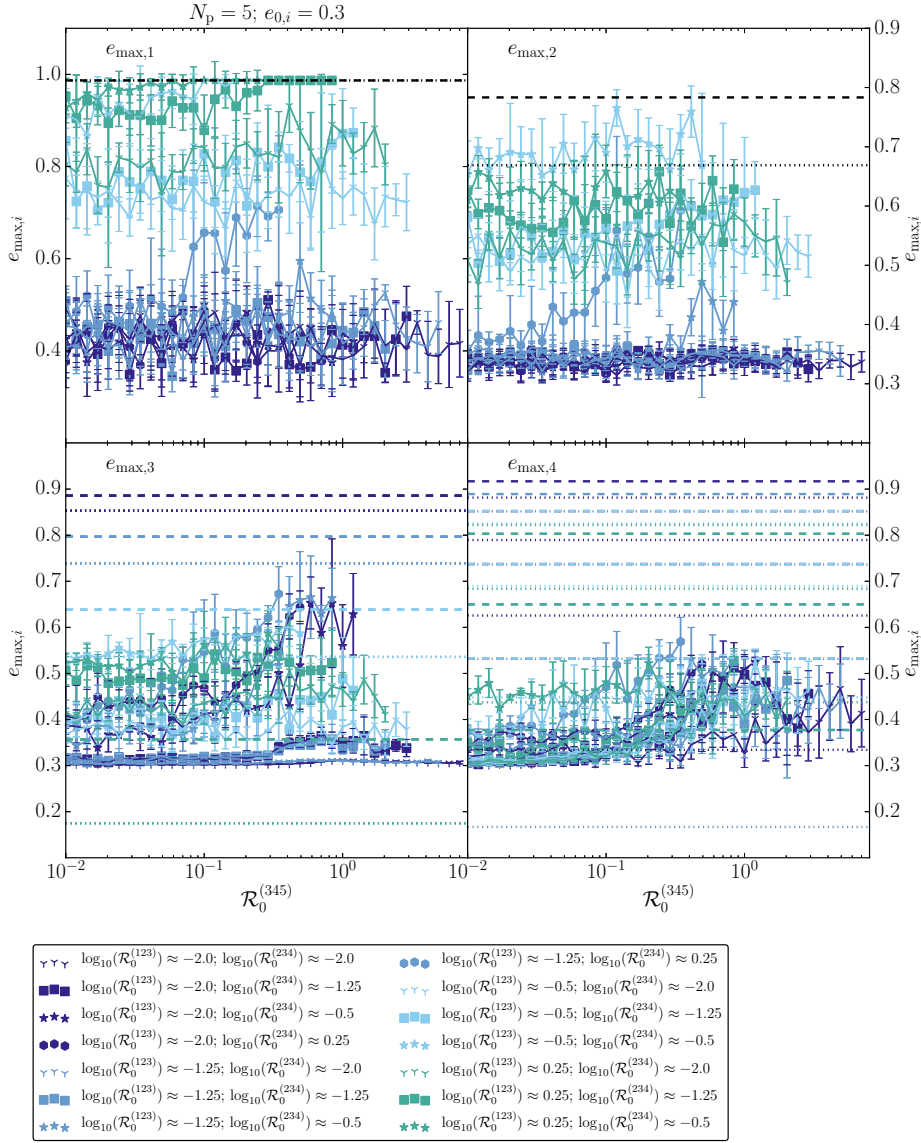


Figure 7.11: For $N_p = 5$ and $e_{0,i} = 0.3$, the maximum eccentricities of the orbits as a function of $\mathcal{R}_0^{(345)}$, assuming various values of $\mathcal{R}_0^{(123)}$ and $\mathcal{R}_0^{(234)}$ (plotted with different colors and symbols, as indicated in the legend at the bottom of the figure). Each panel corresponds to a different orbit, indicated in the top-left corner. As in Figure 7.8, for each orbit, two limiting eccentricities related to tidal disruption or orbit crossing are shown with colored dashed and dotted lines, the color corresponding to each orbit. For the innermost orbit, these lines correspond to the tidal disruption of the planet (assuming equation 7.7). For other orbits i , the dotted line corresponds to orbit crossing of orbit i with orbit $i - 1$ ($K_{ij} = 0$; cf. equation 7.2), assuming e_{i-1} is equal to the value corresponding to disruption (if $i - 1 = 1$), or orbit crossing of orbit $i - 1$ with orbit $i - 2$ (otherwise). The dashed lines correspond to orbit crossing of orbit i with orbit $i - 1$, assuming $e_{i-1} = e_{0,i}$ is equal to the initial value.

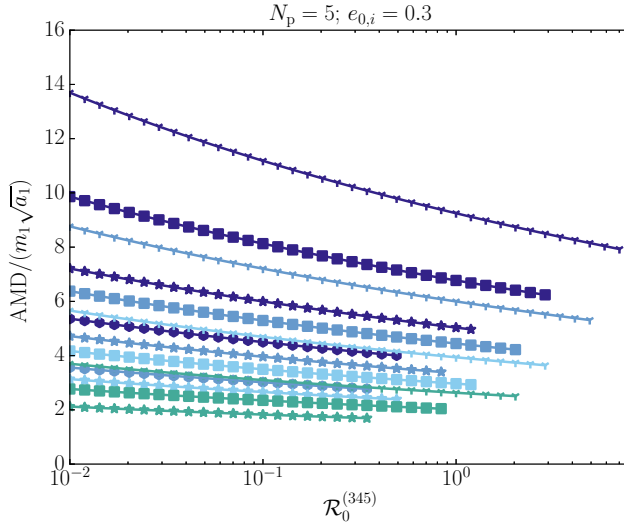


Figure 7.12: For $N_p = 5$ and $e_{0,i} = 0.3$, the AMD as a function of $\mathcal{R}_0^{(345)}$ for the various values of $\mathcal{R}_0^{(123)}$ and $\mathcal{R}_0^{(234)}$ included in Figure 7.11 (refer to the legend at the bottom of the latter figure for the meaning of the colors and symbols).

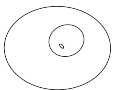
tivity). Here, we investigate the implications of these high eccentricities when both tidal dissipation in the innermost planet, and the possible tidal disruption of the planet are also taken into account. We assume that the innermost planet is tidally disrupted if the innermost orbit satisfies $r_{p,1} = a_1(1 - e_1) < r_t$, where the tidal disruption radius r_t is defined as

$$r_t = \eta \left(\frac{M_\star}{m_1} \right)^{1/3}. \quad (7.7)$$

Here, η is a dimensionless parameter; throughout, we assume $\eta = 2.7$ (Guillochon et al. 2011).

We consider secular integrations with an example system similar to those in the previous sections. We set the semimajor axes to 1, 6, 12, 30 and 50 AU for the (up to) five orbits, respectively. The initial eccentricities and inclinations are again assumed to be equal for all orbits, and we assume $e_i = i_i = e_{0,i} = 0.4$. All other parameters are the same as in Sections 7.4.1 and 7.4.2.

First, we (still) do not take into account tidal dissipation in the innermost planet, nor are the integrations stopped if $r_{p,1} \leq a_1(1 - e_1)$. Also, integrations were not stopped when orbits crossed. To investigate the behaviour of the eccentricity evolution, we show in Figure 7.13 the absolute values of the time derivative $d\ell_1/dt$ as a function of ℓ_1 , where $\ell_1 \equiv \sqrt{1 - e_1^2}$, for three, four and five planets in the top, middle and bottom panels, respectively. The integration time was set to 100 Myr, and data is shown for 10^4 equally-timed snapshots. We found that these output data are insufficiently sampled to numerically compute the derivatives. Therefore, we obtained the derivatives analytically from the equations of motion, thereby obtaining the ‘true’ derivatives.



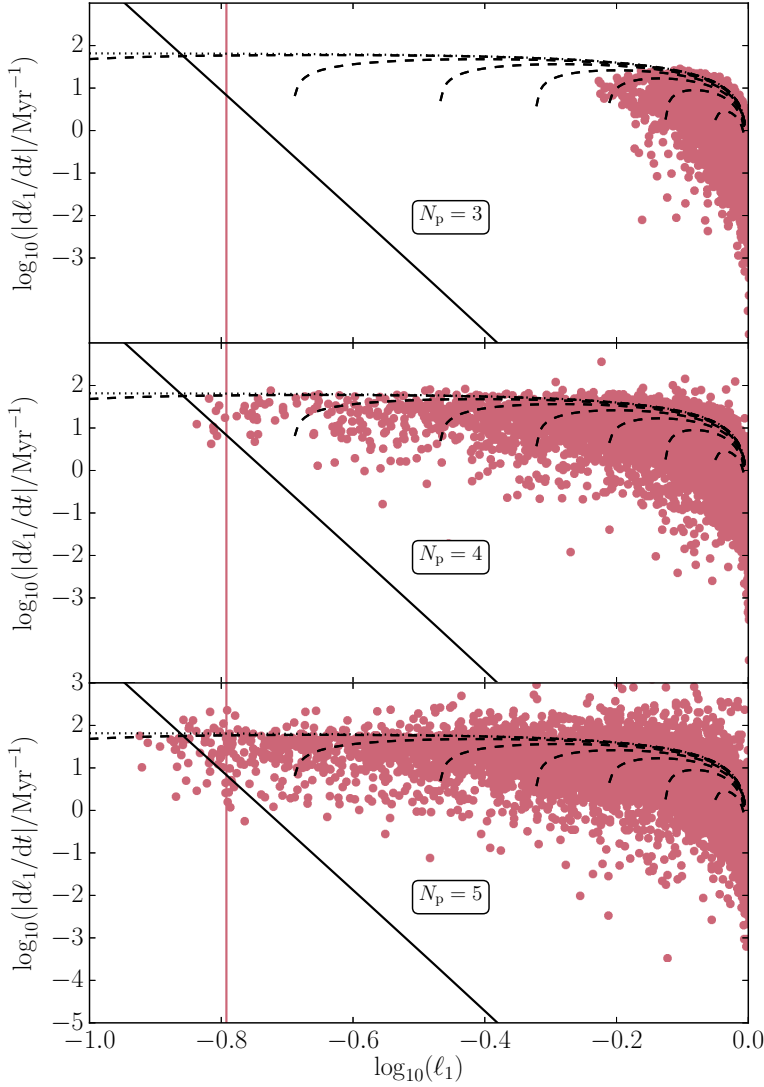


Figure 7.13: The absolute values of the time derivative $d\ell_1/dt$ as a function of ℓ_1 , where $\ell_1 \equiv \sqrt{1 - e_1^2}$, for integrations with three (top panel), four (middle panel) and five (bottom panel) planets (cf. Section 7.4.3). The black dotted lines show the analytic relation equation (7.8) valid in the limit $\ell_1 \rightarrow 1$. The black dashed lines show more general relations based on the three-body quadrupole-order Hamiltonian for different initial inclinations. The vertical red line shows the value corresponding to tidal disruption (cf. equation 7.7). The solid black line shows the time derivative of ℓ_1 for tidal dissipation, assuming $\tau = 1.4 \times 10^{-2}$ yr.

The addition of planets results in larger eccentricity excitation (smaller minimum values of ℓ_1), and the maximum values of $|d\ell_1/dt|$ are also larger. The latter is qualitatively expected given the typically more ‘spiked’ behaviour of the eccentricities as a function of time when more planets are included (e.g. compare Figs. 7.7 and 7.10).

To zeroth order, the dependence of $d\ell_1/dt$ as a function of ℓ_1 due to secular evolution alone can be described by analytic relations that apply to the lowest-order (quadrupole) expansion of the three-body Hamiltonian (e.g. Kinoshita & Nakai 2007). In the limit $\ell_1 \rightarrow 0$,

$$\left| \frac{d\ell_1}{dt} \right|_{\text{sec}} \approx \frac{12\sqrt{6}}{\tau_{\text{sec}}} (1 - \ell_1^2), \quad (7.8)$$

where the time-scale τ_{sec} is defined as

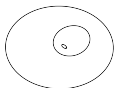
$$\tau_{\text{sec}} \equiv \frac{8}{\pi} \frac{P_{\text{orb},2}^2}{P_{\text{orb},1}} \frac{M_\star + m_1 + m_2}{m_2} (1 - e_2^2)^{3/2}, \quad (7.9)$$

and $P_{\text{orb},i}$ is the period of orbit i . This relation is shown in Figure 7.13 with the black dotted lines. In addition, we show with black dashed lines analytic relations valid for any ℓ_1 (within the quadrupole limit, and assuming that $\ell_1 \cos(i_{12})$ is exactly constant, which is the case if $a_2 \gg a_1$). The latter depend also on other parameters, such as the initial eccentricity and the argument of pericenter. The strongest dependence is on the initial mutual inclination i_{12} ; we show curves corresponding to 20 values of $\cos(i_{12})$ between -1 and 1. Note that for a given initial $\cos(i_{12}) \equiv \theta_0$ and initial $\ell_1 = \ell_{1,0}$, the minimum values of ℓ_1 are $\approx \sqrt{5/3}\theta_0\ell_{1,0}$, and the latter set the ranges of the analytic black dashed curves in Figure 7.13.

Despite the ‘crudeness’ of the analytic relations, the behaviour of $|d\ell_1/dt|$ as a function of ℓ_1 can be reasonably approximated by them in the limit $\ell_1 \rightarrow 0$, and this can be used to gain more insight into the behaviour when tidal dissipation and tidal disruptions are also taken into account. In Figure 7.14, we show results from integrations of the five-planet system, where we now included tidal dissipation in the innermost planet, and the integrations were stopped when the tidal disruption criterion was reached. In the figure, the final innermost semimajor axes are plotted as a function of the viscous time-scale $t_{V,1}$. Systems in which the innermost planet was tidally disrupted, are indicated with red crosses. Other systems, in which the innermost planet migrated (eventually reaching a circular orbit), are indicated with blue bullets.

The value $t_{V,1} \approx 1.4$ yr that was inferred for high- e migration by Socrates et al. (2012), is indicated with the vertical blue line. This value of $t_{V,1}$ results in the tidal disruption of the planet (assuming $\eta = 2.7$). As $t_{V,1}$ is decreased (and tidal dissipation is made more efficient), the final semimajor axis first decreases slightly because of some orbital shrinkage due to tidal dissipation; however, the planet is still tidally disrupted. Around $t_{V,1} \lesssim 10^{-3}$ yr, there is a sudden transition between tidally disrupted and migrating planets. The final semimajor axis in this regime is ~ 0.04 AU, comparable to the semimajor axes of HJs. Note that this value corresponds to tidal dissipation which is 1000 times more efficient compared to the value obtained by Socrates et al. (2012). For even smaller $t_{V,1}$, the final semimajor axis decreases. Values similar to those of WJs, i.e. ~ 0.1 AU, are attained for $t_{V,1} \sim 10^{-9}$ yr. The associated tidal dissipation strength is extreme, however, and not very plausible.

To understand the dependence of the final semimajor axis on $t_{V,1}$, in particular the transition between tidally disrupted and migrating planets, we consider a simple model based on the



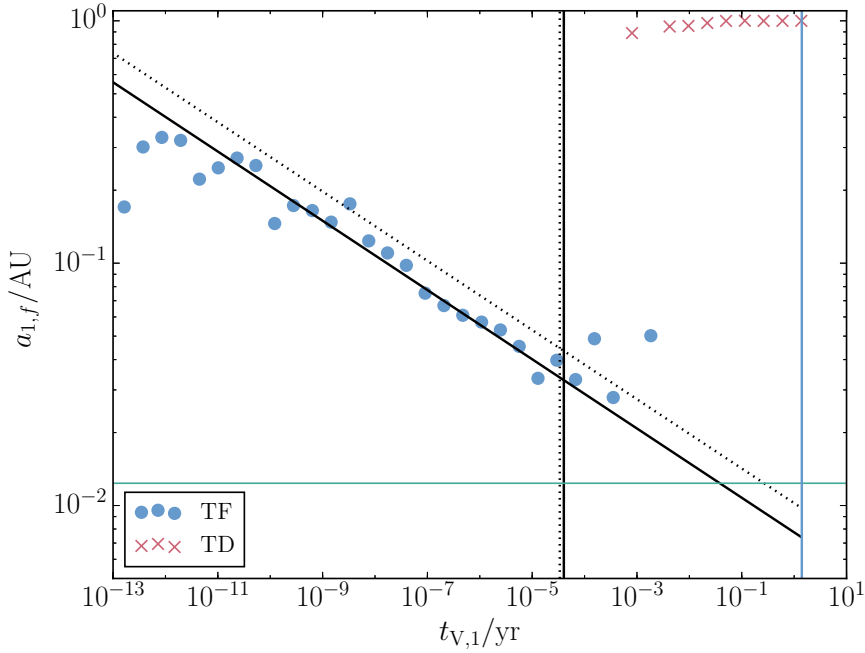


Figure 7.14: Final semimajor axis versus inner planet viscous time-scale for integrations of a five-planet system (cf. Section 7.4.3), taking into account tidal dissipation and tidal disruptions. Systems in which the innermost planet was tidally disrupted, are indicated with red crosses. Other systems, in which the innermost planet migrated, are indicated with blue bullets. The vertical blue line shows the value $t_{V,1} \approx 1.4$ yr that was inferred for high- e migration by Socrates et al. (2012). The solid black vertical line shows equation (7.13), an estimate of the transition value of $t_{V,1}$ between migration and tidal disruption. An approximate and simpler version of equation (7.13), equation (7.14), is shown with the vertical dotted black line. The horizontal green line shows the stalling semimajor axis due to tidal bulges according to equation (7.15). The approximations of the stalling semimajor axis equations (7.21) and (7.22) are shown with the inclined black dotted and solid lines, respectively.

analytic relation equation (7.8). In the limit $\ell_1 \rightarrow 0$, the time derivative of ℓ_1 due to tidal dissipation, assuming pseudosynchronisation (i.e. the spin frequency is equal to the orbital frequency at pericenter), is given approximately by (Hut 1981)

$$\left| \frac{d\ell_1}{dt} \right|_{\text{TF}} \approx \frac{451}{320} \frac{1}{\tau_{\text{TF}}} \left(\frac{R_1}{a_1} \right)^8 \frac{1 - \ell_1^2}{\ell_1^{14}}, \quad (7.10)$$

where the time-scale τ_{TF} is defined as

$$\tau_{\text{TF}} \equiv \frac{1}{27} \frac{t_{\text{V},1}}{3(k_{\text{AM},1} + 2)} \left(\frac{m_1}{M_\star} \right)^2. \quad (7.11)$$

Equation (7.10) is plotted in Figure 7.13 with the black solid line, assuming $\tau = 1.4 \times 10^{-2}$ yr. Also, we show with the vertical solid red line the value of ℓ_1 corresponding to tidal disruption (cf. equation 7.7).

Using that the dependence of $|d\ell_1/dt|$ as a function of ℓ_1 due to secular evolution is approximately described by equation (7.8) in the limit $\ell_1 \rightarrow 0$, we characterise the behaviour of the evolution in the $(\ell_1, |d\ell_1/dt|)$ -plane in this limit $\ell_1 \rightarrow 0$ as follows.

1. Above the black solid line, the evolution is dominated by secular torques, and ℓ_1 oscillates back and forth on the black dotted lines between minimum and maximum values of ℓ_1 .
2. If, due to the secular torques, ℓ_1 reaches the black solid line, tidal dissipation will dominate, circularising (ℓ_1 increasing to unity) and shrinking the orbit.
3. However, if the intersection between the dotted and solid black lines, $\ell_{1,\text{TF,sec}}$, occurs at a value of ℓ_1 that is smaller than the value $\ell_{1,\text{TD}}$ at which the planet is tidally disrupted, then the planet will be disrupted at a value ℓ_1 *before* tidal dissipation is sufficiently strong to circularise and shrink the orbit.

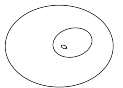
The intersection between the dotted and solid black lines, $\ell_{1,\text{TF,sec}}$, is found by equating equations (7.8) and (7.10), giving

$$\ell_{1,\text{TF,sec}}^{14} = \frac{451}{320} \frac{1}{12\sqrt{6}} \frac{\tau_{\text{sec}}}{\tau_{\text{TF}}} \left(\frac{R_1}{a_1} \right)^8. \quad (7.12)$$

Based on the above model, we expect that the transition between tidal disruption and migration occurs when $\ell_{1,\text{TF,sec}} \approx \ell_{1,\text{TD}}$, where $\ell_{1,\text{TD}} = \sqrt{1 - e_{1,\text{TD}}^2}$ and $e_{1,\text{TD}} = 1 - r_t/a_1$. The corresponding value of $t_{\text{V},1}$ is

$$t_{\text{V},1} \approx \frac{\tau_{\text{sec}}}{12\sqrt{6}} \frac{451}{320} \ell_{1,\text{TD}}^{-14} \frac{27}{3} (k_{\text{AM},1} + 2) \left(\frac{M_\star}{m_1} \right)^2 \left(\frac{R_1}{a_1} \right)^8. \quad (7.13)$$

Note that for simplicity, we neglect any shrinkage of the orbit due to tidal friction at the moment when $\ell_{1,\text{TF,sec}} \approx \ell_{1,\text{TD}}$. We relax this assumption below. Equation (7.13) is plotted in Figure 7.14 with the solid black vertical line. Despite the crudeness of the model, it gives a reasonable estimate of the transition value of ℓ_1 between the migration and tidal dissipation regimes.



Since $e_{1,\text{TD}}$ is close to unity, $\ell_{1,\text{TD}}$ can be approximated as $\ell_{1,\text{TD}} \approx \sqrt{2(1 - e_{1,\text{TD}})}$. This gives

$$t_{\text{V},1} \approx P_{\text{orb},1} \frac{3}{\pi\sqrt{6}} \frac{451}{320} (2\eta)^{-7} \left(\frac{a_2}{a_1}\right)^3 (1 - e_2^2)^{3/2} \times \left(\frac{M_\star}{m_2}\right) \left(\frac{M_\star}{m_1}\right)^{-1/3} \left(\frac{R_1}{a_1}\right). \quad (7.14)$$

The latter relation, which is expressed in terms of the orbital period, is plotted in Figure 7.14 with the vertical dotted black line. Equation (7.14) is indeed a good approximation of equation (7.13). Note that equation (7.14) is very sensitive on η , i.e. the tidal disruption ‘threshold factor’. For $\eta = 2.7$ (which is assumed in Figure 7.14), $t_{\text{V},1}/P_{\text{orb},1} \approx 3.4 \times 10^{-5}$, whereas for $\eta = 1$, $t_{\text{V},1}/P_{\text{orb},1} \approx 3.5 \times 10^{-2}$.

In Section 7.3.2, we mentioned the ‘stalling’ semimajor axis $a_{1,\text{stall}}$ based on the quenching of the secular oscillations due to tidal bulges at a minimum ℓ_1 . The subsequent evolution is assumed to be due to tidal evolution alone (decoupled from the secular oscillations), i.e. with constant semilatus rectum, resulting in a circularised orbit with semimajor axis $a_{1,\text{stall}}$. Following an analysis similar to that of Wu & Lithwick (2011), this gives

$$a_{1,\text{stall}} \approx a_1 \ell_{1,\text{TB}}^2 \approx \left(\frac{15}{4\pi} \frac{21}{8} k_{\text{L},1}\right)^{2/9} R_1 \left(\frac{a_2}{a_1}\right)^{2/3} \left(\frac{M_\star^2}{m_1 m_2}\right)^{2/9} \left(\frac{R_1}{a_1}\right)^{1/9}. \quad (7.15)$$

(note that some of the exponents in equation (7.15) are slightly different from those of Wu & Lithwick (2011); this is because in their Eq. (1), Wu & Lithwick (2011) did not include a factor of ℓ_1 to take into account the dependence of the secular time-scale on ℓ_1 in the limit $\ell_1 \rightarrow 0$).

In Section 7.3.2, tidal disruptions were not taken into account (i.e. formally $\eta = 0$). Consequently, with the assumed $t_{\text{V},1} \approx 4.8$ yr, the considered systems were in the regime where precession due to tidal bulges is important. This can be inferred indirectly from Figure 7.14 by extrapolating the blue bullets to values of $t_{\text{V},1} \approx 4.8$ yr (i.e. without taking into account tidal disruptions). In Figure 7.14, the stalling semimajor axis due to tidal bulges is shown with the horizontal green line. For $t_{\text{V},1} \approx 4.8$ yr, the blue bullets would indeed lie in the regime where the stalling semimajor axis is limited by tidal bulges. However, with nonzero η , migration occurs only for sufficiently small $t_{\text{V},1}$. In this case, it is shown in Figure 7.14 that the stalling semimajor axis is no longer mediated by tidal bulges.

We consider a simple model with the aim of capturing the dependence of the final semimajor axis on $t_{\text{V},1}$ in the case of migration. For notational convenience, we drop below the subscript ‘1’ to denote quantities associated with the innermost planet or orbit. We assume that tidal bulges and tidal disruptions are not important (we also neglect relativistic precession). In this case, taking the analytic approximation for the secular contribution to $d\ell_1/dt$ in the limit $\ell_1 \rightarrow 0$ (cf. equation 7.8), and taking into account tidal dissipation, the equations of motion read

$$\begin{cases} \dot{\ell} & \approx \beta_e \frac{1}{\tau_{\text{TF}}} \left(\frac{R}{a}\right)^8 \frac{1-\ell^2}{\ell^{14}} \mp \frac{12\sqrt{6}}{\tau_{\text{sec}}} (1 - \ell^2); \\ \dot{a} & \approx -\beta_a \frac{1}{\tau_{\text{TF}}} \left(\frac{R}{a}\right)^8 \frac{1-\ell^2}{\ell^{15}}. \end{cases} \quad (7.16)$$

Here, dots denote derivatives with respect to time, and the minus (plus) sign in the second term of $\dot{\ell}$ corresponds to a decreasing (increasing) value of ℓ in the LK cycle (depending on the

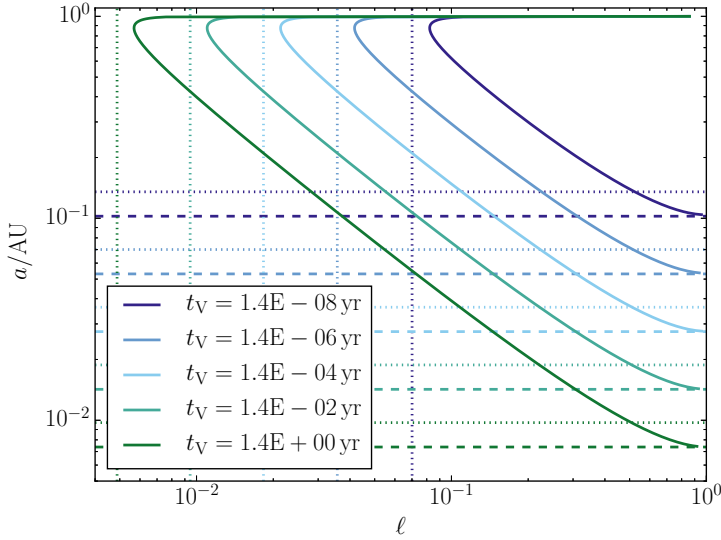


Figure 7.15: Solid lines: equation (7.19) in the (ℓ, a) -plane for various values of $t_{V,1}$. The vertical dotted lines show an approximation of ℓ_c where $a_c = a_0 = 1$ AU was assumed. The horizontal dotted lines show the rough approximation equation (7.21) of the stalling semimajor axis; the horizontal dashed lines show the more accurate approximation equation (7.22).

phase, i.e. the argument of pericenter ω ; in this simple model, we do not consider the precise functional dependence of ℓ on ω). As before, we assume pseudosynchronisation, for which $\beta_e = 451/320$, and $\beta_a = 2\beta_e$. Division of $\dot{a}$ by $\dot{\ell}$ eliminates reference to time, and gives the following differential equation for a_1 and ℓ_1 :

$$\frac{da}{d\ell} \approx -2\frac{a}{\ell} \left[1 \mp \alpha \left(\frac{a}{R} \right)^8 \ell^{14} \right]^{-1}, \quad (7.17)$$

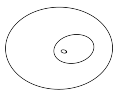
where the dimensionless quantity α is defined as

$$\alpha \equiv \frac{12\sqrt{6}}{\beta_e} \frac{\tau_{\text{TF}}}{\tau_{\text{sec}}}. \quad (7.18)$$

The solution of equation (7.17), setting $a = a_0$ when $\ell = \ell_0$, is given by

$$\ell(a) \approx \left(\frac{R}{a} \right)^{1/2} \left[\left(\frac{R}{a_0} \right)^7 \ell_0^{-14} + 7\alpha \left(\frac{a_0}{R} - \frac{a}{R} \right) \right]^{-1/14}. \quad (7.19)$$

For illustration, we show the curve defined by equation (7.19) in the (ℓ, a) -plane for various values of $t_{V,1}$.



We obtain two approximations of the stalling semimajor axis. First, let ℓ_c denote the value of ℓ for which $\dot{\ell} = 0$ assuming the negative sign for the secular term (cf. equation 7.16), i.e.

$$\ell_c^{14} = \frac{1}{\alpha} \left(\frac{R}{a_c} \right)^8, \quad (7.20)$$

where a_c is the associated value of a . For a rough approximation, we set $a_c = a_0$ to the initial value of a (in general, $a_c < a_0$), and approximate the subsequent evolution with neglect of secular torques, i.e. with constant semilatus rectum. This implies that the final semimajor axis (at circularisation) is given by

$$a_f \approx a_c \ell_c^2 \approx a_0 \ell_c^2 \approx a_0 \alpha^{-1/7} \left(\frac{R}{a_0} \right)^{8/7}. \quad (7.21)$$

Note that a_f is weakly dependent on a_0 and approximately linearly dependent on R , similar to equation (7.15). Equation (7.21) is shown with the horizontal dotted lines in Figure 7.15.

A more accurate approximation can be obtained from the solution equation (7.19). Setting the final $\ell_f = \ell_f(a_f) = 1$, assuming $a_0 \gg R$ and $a_f \ll a_0$, this condition gives

$$a_f \approx a_0 (7\alpha)^{-1/7} \left(\frac{R}{a_0} \right)^{8/7}, \quad (7.22)$$

which is the same as equation (7.21) modulo a factor of $7^{-1/7} \approx 0.76$. Equation (7.22) is shown with the horizontal dashed lines in Figure 7.15; it gives a good approximation of the stalling semimajor axes implied by equation (7.19).

The approximations of the stalling semimajor axis equations (7.21) and (7.22) are shown in Figure 7.14 with the inclined black dotted and solid lines, respectively. Equation (7.22) in particular gives a good description of the stalling semimajor axis as a function of $t_{V,1}$. The slope of the relation is fairly small; both equations (7.21) and (7.22) scale with $t_{V<1}$ according to $a_{1,f} \propto t_{V,1}^{-1/7}$.

7.5 Population synthesis

In Section 7.3, we verified that the properties of HJs formed through secular dynamical and tidal interactions in multiplanet systems, as computed with `SECULARMULTIPLE`, are statistically consistent with more accurate, but much more computationally expensive direct N -body integrations. Subsequently, we investigated in Section 7.4 the conditions for exciting high eccentricities, and thereby possibly producing HJs through high- e migration, in multiplanet systems with between 3 and 5 planets. In addition to a sufficiently large angular-momentum deficit ($\text{AMD} > 1 m_1 \sqrt{a_1}$, cf. equation 7.4), the ratios $\mathcal{R}_0^{(ijk)}$ of Lidov-Kozai (LK) time-scales (cf. equation 7.5) associated with three adjacent orbits (i, j, k) should be near unity, in particular the ratio applied to the innermost orbits. If $\mathcal{R}_0^{(ijk)} \sim 1$, commensurate LK-times-scales for the pairs (i, j) and (j, k) can give rise to coupled and resonant secular oscillations, potentially driving high eccentricities.

We emphasize that despite the statistical consistency of the secular method with the direct N -body integrations as found in Section 7.3, it should be taken into account that orbit crossings

occurred in the integrations presented in that section, both with the secular and direct methods. In such cases, the assumptions on which the secular method is based formally break down, and results should be interpreted critically. We addressed this concern further in Section 7.4.1, where we carried out direct N -body integrations starting at the onset of orbit crossings, and found that the secular method approximately captures the subsequent long-term evolution. This is, however, not necessarily always the case. In our population synthesis study as described below, integrations were stopped at orbit crossings. To investigate whether this affects our conclusions, some of the integrations were not stopped at such occurrences.

7.5.1 Initial conditions

To generate the initial conditions of our systems, we use a combination of grid and Monte Carlo sampling. We define a grid with the following parameters: (1) the viscous time-scale $t_{V,1}$ of the innermost planet (or, equivalently, the tidal time lag τ_1), (2) the number of planets N_p , (3) β , which is related to the rms width of the initial inclination and eccentricity distribution (see below), and (4) the radius of the innermost planet, R_1 . The following values are considered:

$$\begin{cases} t_{V,1} & \in \{10^{-2}, 10^{-1}, 10^0\} t_{V,\text{SKD}}; \\ N_p & \in \{3, 4, 5\}; \\ \beta & \in \left\{ \left(20 \frac{\pi}{180}\right)^{-2}, \left(15 \frac{\pi}{180}\right)^{-2}, \left(10 \frac{\pi}{180}\right)^{-2} \right\} \\ & \approx \{8.2, 14.6, 32.8\}; \\ R_1 & \in \{1.0, 1.5\} R_J. \end{cases} \quad (7.23)$$

Here, $t_{V,\text{SKD}} = 1.2 \times 10^4 \text{ hr} \approx 1.37 \text{ yr}$ is the viscous time-scale that was determined for high- e migration by Socrates et al. (2012).

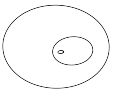
For each combination of grid parameters, we sampled $N_{\text{MC}} = 1000$ systems using the following approach. The masses m_i of the N_p planets were sampled from linear distributions with $0.5 M_J < m_i < 5 M_J$. The inclinations i_i (as measured in radians, and defined with respect to an arbitrary inertial reference frame) and eccentricities e_i of the planetary orbits are both sampled from a Rayleigh distribution,

$$\frac{dN}{dx} \propto x \exp(-\beta x^2), \quad (7.24)$$

where $\beta = \langle x^2 \rangle^{-1}$ (assuming $0 < x < \infty$) characterises the width of the distribution. Here, the sampling limits are $0 < e_i < 0.8$ and $0^\circ < i_i < 30^\circ$. The arguments of pericenter ω_i and the longitudes of the ascending nodes Ω_i were sampled from linear distributions between 0° and 360° .

The semimajor axes of the planets a_i were sampled from linear distributions, with the fixed ranges

$$\begin{cases} 1.0 \text{ AU} < a_1 < 4.0 \text{ AU}; \\ 6.0 \text{ AU} < a_2 < 10.0 \text{ AU}; \\ 15.0 \text{ AU} < a_3 < 30.0 \text{ AU}; \\ 35.0 \text{ AU} < a_4 < 50.0 \text{ AU}; \\ 60.0 \text{ AU} < a_5 < 100.0 \text{ AU}. \end{cases} \quad (7.25)$$



These choices are somewhat arbitrary. The semimajor axes of planets beyond ~ 5 AU are still poorly constrained by observations. Therefore, some assumption must be made on the semimajor axes, in particular for planets outside of the innermost planet. Note that the ranges of a_1 , a_2 and a_3 are similar to the values that were assumed by Wu & Lithwick (2011) (1, 6 and 16 AU, respectively). The choice of a linear distribution, apart from its simplicity, is motivated by the ability to easily disentangle any dependence of the results on the semimajor axes. At any rate, it should be taken into account that our choice of the initial semimajor axes is not unique, and likely affects the results of the population synthesis.

We rejected a sampled combination of m_i , a_i and e_i if $K_{ij} < K_0 \equiv 2$ for any adjacent pair $(i, j) = (i, i + 1)$ of orbits (cf. equation 7.2). This value is also somewhat arbitrary. Choosing a value of $K_0 = 0$ would produce a large fraction of initially orbit-crossing systems, whereas a large value of K_0 , say $K_0 = 12$, would produce too-well-separated systems in which the eccentricities would hardly evolve. As a compromise, we have set a value of $K_0 = 2$.

In Figure 7.16, we show for these assumptions the initial distributions of the AMD (top panel) and $\mathcal{R}_0^{(ijk)}$ (bottom panel) with either 3, 4 or 5 planets. The median value of $\mathcal{R}_0^{(ijk)}$ is $\approx 10^{-0.5} \approx 0.3$. The fraction of systems with $\mathcal{R}_0^{(ijk)} \gtrsim 1$ is ≈ 0.2 . The AMD is strongly dependent on both the number of planets and the width of the eccentricity and inclination distribution. With $N_p = 3$ and $\beta \approx 32.8$ (corresponding to the smallest rms value of e_i and i_i), only a fraction of ≈ 0.01 of systems has an AMD which is $> m_1 \sqrt{a_1}$; with $N_p = 5$ and $\beta \approx 8.2$, this fraction is ≈ 0.4 .

7.5.2 Stopping conditions

The secular integrations were stopped if one of the following conditions was met.

1. The integration time reached 10 Gyr.
2. A HJ was formed, i.e. $P_1 < 10$ d and $e_1 < 10^{-3}$. In this case, the orbit is well within the regime of decoupling of tidal evolution from secular evolution (see e.g. Figure 7.3), and there is no further tidal evolution since the orbit is (essentially) circularised.
3. The innermost planet was tidally disrupted by the star (cf. equation 7.7, with $\eta = 2.7$; Guillochon et al. 2011).
4. Any of the orbits intersected, i.e. $K_{ij} < 0$ for any pair of adjacent orbits (i, j) (cf. equation 7.2), and the secular approximation formally breaks down. In order to investigate the sensitivity of our results on this stopping condition, we also carried out additional simulations for some parameter combinations in which we did not check for orbit crossings.
5. The run time of the simulation exceeded 4 CPU hours (this was imposed for practical reasons).

7.5.3 Results

Overview

Our results are summarised in Table 7.2, where, for each combination of $t_{V,1}$, N_p , β and R_1 , we show the fractions of the $N_{MC} = 1000$ Monte Carlo-sampled systems for the following outcomes.

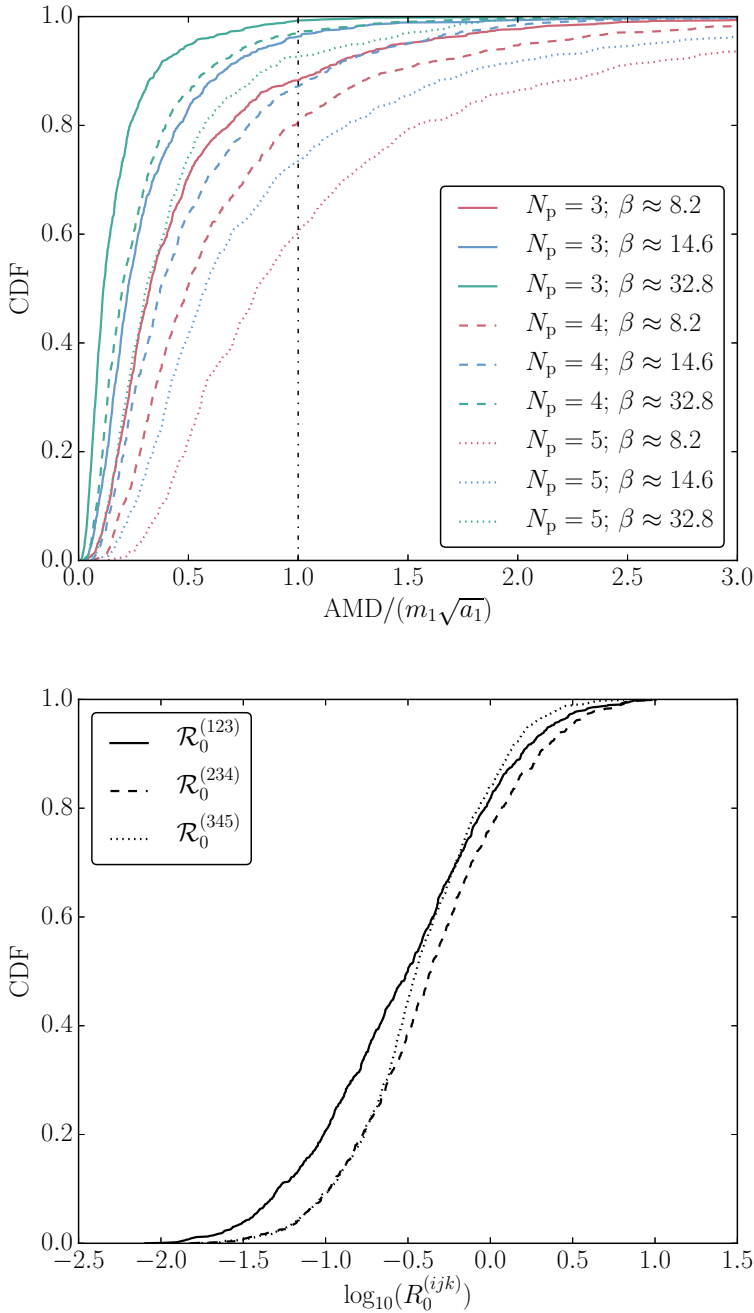
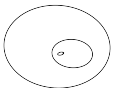


Figure 7.16: The initial cumulative distributions of the AMD (top panel) and $\mathcal{R}_0^{(ijk)}$ (bottom panel) for the Monte Carlo simulations of Section 7.5 with either 3, 4 or 5 planets. In the bottom panel, we show distributions for all combinations of N_p and β . In the bottom panel, $\beta \approx 8.2$ is assumed; note that the dependence of $\mathcal{R}_0^{(ijk)}$ on β is weak.



1. No migration occurred, i.e. the final orbital period $P_1 > 100$ d ($f_{\text{no migration}}$).
2. A HJ was formed (f_{HJ}).
3. The innermost planet was tidally disrupted (f_{TD}).
4. An orbit crossing occurred ($f_{\text{orbit crossing}}$).
5. The maximum run time was exceeded ($f_{\text{run time exceeded}}$).

These fractions are given after either 5 or 10 Gyr of integration, or by sampling a random time between 100 Myr and 10 Gyr for each of the systems (output times were separated by 100 Myr), corresponding to a constant star formation rate (indicated with t_x in the table). For $t_{\text{V},1} = 10^{-2} t_{\text{V,SKD}}$ and $R_1 = 1 R_J$, we also carried out simulations without the orbit crossing stopping condition. These can be recognised in Table 7.2 with the entries with $f_{\text{orbit crossing}}$ marked as $-$.

The fractions of HJs formed are typically low; the largest fraction is 0.007, obtained after 10 Gyr for the set of simulations with $t_{\text{V},1} = 0.01 t_{\text{V,SKD}}$, $N_p = 4$, $\beta \approx 8.2$ and $R_1 = 1 R_J$. In contrast, the fraction of tidal disruptions is much larger, typically a few per cent, and reaching values of $\approx 0.2 - 0.3$ for $N_p = 5$. We discuss the reasons for this difference between the tidal disruption and HJ fractions in Section 7.5.3.

In our population synthesis runs, the number of WJs formed (defined as planets with a period between 10 and 100 days) is even smaller than the number of HJs, and the associated fractions are not included in Table 7.2. Among the 54,000 integrations, 34 HJs were formed after 10 Gyr, whereas the number of WJs at that time is 2 (the number of WJs and HJs at 5 Gyr is 2 and 26, respectively). Moreover, the semimajor axes of those 2 WJs are ≈ 0.1 AU, i.e. these WJs are on the very ‘hot’ end of the WJ spectrum, and are on the (not very well defined) boundary between between WJs and HJs.

For larger values of $t_{\text{V},1}$, i.e. for weaker tidal dissipation strength in the innermost planet, no HJs are formed at all, for any of the combinations of the other grid parameters. It should, however, be taken into account that the number of Monte Carlo realisations per parameter combination was limited to $N_{\text{MC}} = 1000$, implying that the HJ fractions could be less than 0.001, but nonzero. Also, uncertainties associated with the occurrences of orbit crossings and premature terminations of the integrations because of the exceeding of the maximum run time, should be taken into account.

A comparison of the HJ fractions between runs with and without the orbit crossing stopping condition enabled in Table 7.2, shows that disabling the stopping condition results in different fractions. Typically, the fraction increases, which can be understood by noting that systems in which orbit crossings occur are likely to produce high eccentricities in the inner orbit, and therefore HJs. However, despite this effect, the fractions remain small, not reaching values larger than 0.007. Therefore, we believe that our finding that the HJ fraction is small is not contingent on uncertainties associated with our orbit crossing stopping condition.

Furthermore, the fraction of systems in which the run time was exceeded typically increases with β . However, the HJ fractions tend to be largest for the smallest β , and decrease rapidly with increasing β . Therefore, we also believe that the HJ fractions would not be very different if this stopping condition were not imposed.

$t_{V,1}$	N_p	β	R_1	$f_{\text{no migration}}$									f_{HJ}			f_{TD}			$f_{\text{orbit crossing}}$			$f_{\text{run time exceeded}}$		
t_{end} (Gyr) $\rightarrow$				5	10	t_x	5	10	t_x	5	10	t_x	5	10	t_x	5	10	t_x	5	10	t_x	5	10	t_x
0.01	3	8.2	1.0	0.825	0.805	0.834	0.001	0.001	0.0	0.055	0.056	0.052	0.119	0.125	0.112	0.0	0.013	0.002						
0.01	3	8.2	1.0	0.924	0.91	0.926	0.003	0.003	0.003	0.071	0.073	0.07	—	—	—	0.001	0.014	0.001						
0.01	3	8.2	1.5	0.828	0.793	0.826	0.003	0.004	0.004	0.036	0.038	0.034	0.133	0.145	0.133	0.0	0.02	0.003						
0.01	3	14.6	1.0	0.91	0.886	0.914	0.005	0.006	0.005	0.027	0.028	0.023	0.057	0.06	0.052	0.001	0.02	0.006						
0.01	3	14.6	1.0	0.974	0.945	0.966	0.0	0.001	0.001	0.025	0.026	0.025	—	—	—	0.001	0.028	0.008						
0.01	3	14.6	1.5	0.935	0.912	0.932	0.001	0.002	0.002	0.018	0.019	0.018	0.046	0.049	0.044	0.0	0.018	0.004						
0.01	3	32.8	1.0	0.984	0.964	0.981	0.0	0.0	0.0	0.003	0.003	0.003	0.011	0.011	0.009	0.002	0.022	0.007						
0.01	3	32.8	1.0	0.996	0.968	0.991	0.0	0.0	0.0	0.004	0.004	0.004	—	—	—	0.0	0.028	0.005						
0.01	3	32.8	1.5	0.989	0.95	0.982	0.0	0.0	0.0	0.002	0.002	0.002	0.009	0.009	0.009	0.0	0.038	0.007						
0.01	4	8.2	1.0	0.345	0.282	0.371	0.0	0.001	0.001	0.058	0.062	0.055	0.589	0.614	0.56	0.008	0.041	0.013						
0.01	4	8.2	1.0	0.772	0.617	0.771	0.006	0.007	0.004	0.173	0.18	0.164	—	—	—	0.049	0.196	0.061						
0.01	4	8.2	1.5	0.305	0.234	0.323	0.002	0.004	0.002	0.068	0.07	0.064	0.612	0.63	0.588	0.013	0.062	0.023						
0.01	4	14.6	1.0	0.566	0.446	0.573	0.001	0.001	0.001	0.041	0.045	0.041	0.38	0.406	0.364	0.012	0.102	0.021						
0.01	4	14.6	1.0	0.821	0.61	0.817	0.002	0.002	0.002	0.107	0.112	0.095	—	—	—	0.07	0.276	0.086						
0.01	4	14.6	1.5	0.562	0.451	0.564	0.002	0.002	0.002	0.039	0.04	0.04	0.385	0.416	0.365	0.011	0.09	0.028						
0.01	4	32.8	1.0	0.843	0.726	0.824	0.003	0.003	0.003	0.01	0.01	0.01	0.124	0.133	0.12	0.02	0.129	0.043						
0.01	4	32.8	1.0	0.961	0.853	0.952	0.0	0.0	0.0	0.019	0.02	0.02	—	—	—	0.02	0.127	0.028						
0.01	4	32.8	1.5	0.875	0.765	0.867	0.004	0.005	0.003	0.006	0.006	0.006	0.096	0.108	0.094	0.017	0.114	0.029						
0.01	5	8.2	1.0	0.093	0.046	0.111	0.0	0.0	0.0	0.044	0.046	0.044	0.843	0.851	0.817	0.02	0.056	0.028						
0.01	5	8.2	1.0	0.531	0.286	0.543	0.003	0.003	0.003	0.318	0.326	0.3	—	—	—	0.148	0.386	0.155						
0.01	5	8.2	1.5	0.081	0.041	0.112	0.002	0.003	0.003	0.053	0.053	0.05	0.838	0.845	0.809	0.026	0.058	0.026						
0.01	5	14.6	1.0	0.201	0.105	0.227	0.001	0.001	0.001	0.056	0.058	0.055	0.691	0.715	0.673	0.051	0.121	0.044						
0.01	5	14.6	1.0	0.674	0.406	0.673	0.002	0.002	0.002	0.203	0.21	0.186	—	—	—	0.12	0.382	0.139						
0.01	5	14.6	1.5	0.206	0.112	0.24	0.002	0.002	0.001	0.055	0.057	0.054	0.695	0.713	0.662	0.042	0.116	0.043						
0.01	5	32.8	1.0	0.594	0.397	0.613	0.0	0.0	0.0	0.031	0.033	0.03	0.288	0.305	0.263	0.087	0.265	0.094						
0.01	5	32.8	1.0	0.863	0.618	0.821	0.0	0.0	0.0	0.051	0.06	0.051	—	—	—	0.087	0.322	0.128						
0.01	5	32.8	1.5	0.589	0.363	0.581	0.0	0.0	0.0	0.015	0.015	0.015	0.322	0.359	0.3	0.074	0.263	0.104						
0.1	3	8.2	1.0	0.826	0.805	0.835	0.0	0.0	0.0	0.03	0.031	0.029	0.144	0.149	0.133	0.0	0.016	0.003						
0.1	3	8.2	1.5	0.819	0.782	0.824	0.0	0.0	0.0	0.042	0.043	0.04	0.139	0.147	0.13	0.0	0.028	0.006						
0.1	3	14.6	1.0	0.916	0.888	0.914	0.0	0.0	0.0	0.023	0.025	0.022	0.061	0.063	0.057	0.0	0.024	0.007						
0.1	3	14.6	1.5	0.917	0.887	0.915	0.0	0.0	0.0	0.019	0.019	0.017	0.063	0.068	0.059	0.001	0.026	0.009						
0.1	3	32.8	1.0	0.99	0.963	0.986	0.0	0.0	0.0	0.004	0.004	0.004	0.006	0.006	0.005	0.0	0.027	0.005						
0.1	3	32.8	1.5	0.989	0.967	0.988	0.0	0.0	0.0	0.004	0.004	0.004	0.006	0.007	0.005	0.001	0.022	0.004						
0.1	4	8.2	1.0	0.304	0.24	0.325	0.0	0.0	0.0	0.071	0.072	0.068	0.611	0.636	0.589	0.015	0.052	0.017						
0.1	4	8.2	1.5	0.307	0.228	0.326	0.0	0.0	0.0	0.059	0.06	0.058	0.619	0.638	0.594	0.015	0.074	0.022						
0.1	4	14.6	1.0	0.544	0.438	0.559	0.0	0.0	0.0	0.051	0.054	0.049	0.389	0.408	0.363	0.016	0.101	0.029						
0.1	4	14.6	1.5	0.52	0.349	0.525	0.0	0.0	0.0	0.04	0.043	0.04	0.39	0.401	0.37	0.05	0.207	0.065						
0.1	4	32.8	1.0	0.825	0.682	0.818	0.0	0.0	0.0	0.015	0.015	0.013	0.11	0.11	0.103	0.05	0.193	0.068						
0.1	4	32.8	1.5	0.869	0.746	0.853	0.0	0.0	0.0	0.006	0.008	0.005	0.107	0.121	0.101	0.018	0.126	0.041						
0.1	5	8.2	1.0	0.082	0.042	0.116	0.0	0.0	0.0	0.054	0.056	0.05	0.852	0.865	0.814	0.012	0.038	0.019						
0.1	5	8.2	1.5	0.104	0.045	0.119	0.0	0.0	0.0	0.078	0.078	0.077	0.799	0.811	0.776	0.019	0.066	0.029						
0.1	5	14.6	1.0	0.206	0.121	0.233	0.0	0.0	0.0	0.057	0.059	0.055	0.701	0.724	0.677	0.036	0.096	0.036						
0.1	5	14.6	1.5	0.217	0.121	0.246	0.0	0.0	0.0	0.064	0.064	0.061	0.684	0.705	0.65	0.035	0.11	0.043						
0.1	5	32.8	1.0	0.621	0.427	0.614	0.0	0.0	0.0	0.021	0.024	0.021	0.292	0.322	0.278	0.066	0.227	0.087						
0.1	5	32.8	1.5	0.559	0.33	0.562	0.0	0.0	0.0	0.027	0.03	0.03	0.364	0.418	0.33	0.051	0.222	0.077						
1.0	3	8.2	1.0	0.811	0.787	0.82	0.0	0.0	0.0	0.046	0.047	0.044	0.143	0.15	0.136	0.0	0.015	0.0						
1.0	3	8.2	1.5	0.808	0.787	0.814	0.0	0.0	0.0	0.04	0.041	0.039	0.152	0.159	0.145	0.0	0.013	0.002						
1.0	3	14.6	1.0	0.914	0.889	0.919	0.0	0.0	0.0	0.022	0.023	0.022	0.063	0.069	0.057	0.001	0.019	0.002						
1.0	3	14.6	1.5	0.924	0.9	0.921	0.0	0.0	0.0	0.021	0.021	0.02	0.054	0.055	0.053	0.001	0.024	0.006						
1.0	3	32.8	1.0	0.992	0.965	0.985	0.0	0.0	0.0	0.002	0.003	0.003	0.006	0.007	0.006	0.0	0.025	0.006						
1.0	3	32.8	1.5	0.986	0.961	0.982	0.0	0.0	0.0	0.0	0.0	0.0	0.013	0.013	0.012	0.001	0.026	0.006						
1.0	4	8.2	1.0	0.345	0.277	0.365	0.0	0.0	0.0	0.06	0.063	0.058	0.587	0.604	0.567	0.008	0.056	0.01						
1.0	4	8.2	1.5	0.367	0.287	0.38	0.0	0.0	0.0	0.049	0.05	0.047	0.576	0.6	0.561	0.008	0.063	0.012						
1.0	4	14.6	1.0	0.538	0.436	0.542	0.0	0.0	0.0	0.042	0.043	0.041	0.411	0.436	0.393	0.009	0.085	0.024						
1.0	4	14.6	1.5	0.536	0.433	0.534	0.0	0.0	0.0	0.036	0.037	0.036	0.409	0.434	0.399	0.019	0.096	0.031						
1.0	4	32.8	1.0	0.888	0.771	0.87	0.0	0.0	0.0	0.009	0.009	0.009	0.092	0.103	0.087	0.011	0.117	0.034						
1.0	4	32.8	1.5	0.864	0.745	0.864	0.0	0.0	0.0	0.014	0.015	0.014	0.103	0.114	0.092	0.019	0.125	0.031						
1.0	5	8.2	1.0	0.097	0.065	0.117	0.0	0.0	0.0	0.067	0.067	0.067	0.823	0.829	0.8	0.012	0.038	0.016						
1.0	5	8.2	1.5	0.112	0																			

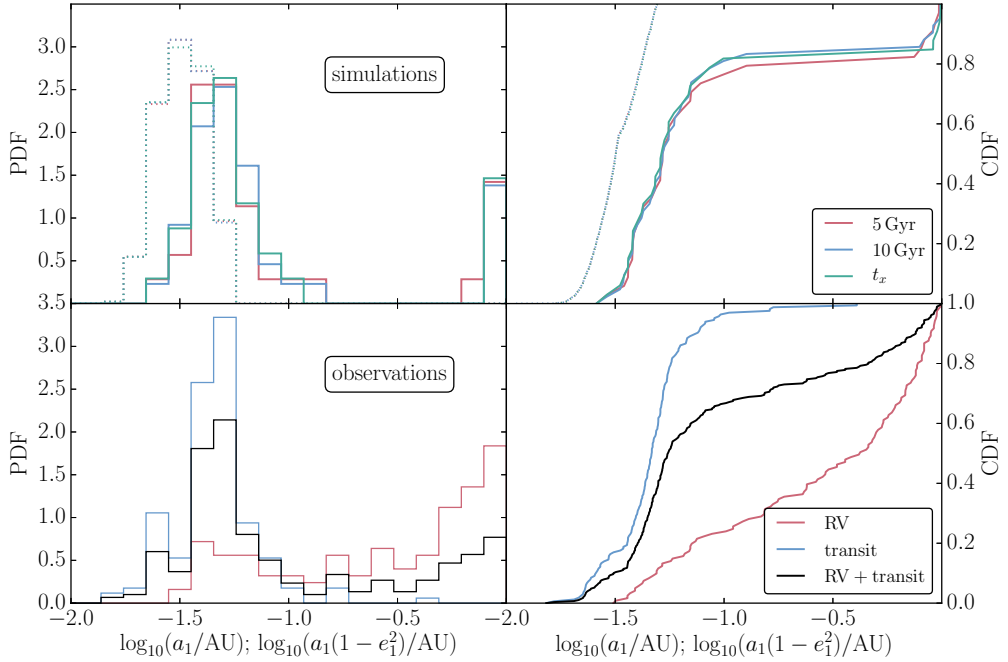


Figure 7.17: The distributions of the semimajor axes of the innermost planet at various times (top row), compared with observed distributions of the semimajor axes of planets with $m_1 \sin i > 0.1 M_J$ obtained from the *Open Exoplanet Catalogue* (bottom row). For the simulated distributions, results from all parameter combinations have been combined. Solid lines correspond to systems in which a WJ or HJ was formed, or no migration occurred (excluding orbit crossing systems). Dashed lines correspond to systems in which the innermost planet was tidally disrupted; in this case, the abscissa is the logarithm of the semilatus rectum $a_1 (1 - e_1^2)$, cf. Section 7.5.3). Results from the simulations are shown for three times: after 5 Gyr (red lines), 10 Gyr (blue lines), and assuming a random time between 100 Myr and 10 Gyr (green lines). The first (second) column shows the PDF (CDF).

Final semimajor axis distributions

In the top row of Figure 7.17, we show the distributions of the semimajor axes of the innermost planet at various times. Here, we have combined the results from all parameter combinations. With the solid lines, we show the semimajor axes of the ‘non-disruptive’ systems in which a WJ or HJ was formed, or no migration occurred (excluding orbit crossing systems). We consider the distributions after 5 Gyr (red lines), 10 Gyr (blue lines), and assuming a random time for each system between 100 Myr and 10 Gyr (t_x ; green lines). With the dotted lines, we show the distributions of the semilatus recti $a_1 (1 - e_1^2)$ of the orbits of tidally disrupted planets. The orbits of the tidally disrupted planets are wide, with semimajor axes $\gtrsim 1$ AU, and they must therefore be highly eccentric in order for the planets to be tidally disrupted. In the absence of further secular perturbations and assuming no actual tidal disruption, tidal dissipation would circularise and shrink the orbit to a semimajor axis of $a_1 (1 - e_1^2)$.

As expected, the semilatus recti for the tidally disrupted planets are typically smaller than the semimajor axes of the HJs (whose orbits are invariably circularised). There is some overlap because of the different radii that were assumed, ($R_1 = 1$ or $1.5 R_J$), implying different tidal disruption distances. The simulated semimajor axis distribution of the HJs is peaked around $\approx 10^{-1.3} \approx 0.05$ AU. This value is consistent with previous studies of high- e migration assuming three bodies (i.e. a binary or planetary companion, e.g. Petrovich & Tremaine 2016).

In the bottom row of Figure 7.17, we show the distributions of observed giant planets ($m_1 \sin i > 0.1 M_J$) obtained from the *Open Exoplanet Catalogue*¹. We included data from radial velocity (RV) and/or transit methods. The distributions from both RV and transit methods show a peak at a value of $\approx 10^{-1.3} \approx 0.05$ AU which is consistent with the simulations. However, similarly to other studies of high- e migration (e.g. Petrovich & Tremaine 2016; Antonini et al. 2016), the simulations fail to produce the large population of WJs in the region between ≈ 0.1 and 1 AU. When considering only observations of transiting planets, the simulated and observed distributions seem very similar. However, such observations are heavily biased to small semimajor axis, which is demonstrated by comparing with the RV observations, which are less biased towards short periods.

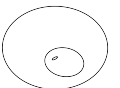
Dependence on the grid parameters

The observed semimajor axis distributions suggest that at least some HJs arise from secular evolution in multiplanet systems. In Figure 7.17, we combined results from all parameters. Here, we consider in more detail the dependence of the results on the assumed parameters.

In Figure 7.18, the innermost orbit semilatus rectum distributions after 10 Gyr (or after a stopping condition was met) are plotted for different slices of the parameter space. In each panel, distributions are shown for all parameters combined, except one (different linestyles). In addition, we distinguish between the different types of systems: systems with no migration (black), HJs (light red), WJs (blue), tidally disrupted inner planets (yellow), orbit crossings (dark blue) or exceeding of run time (dark red).

In the top panel, we consider the dependence on $t_{V,1}$, the viscous time-scale of the innermost planet. As expected, the distributions for non-HJ and WJ forming systems are essentially independent on $t_{V,1}$. In contrast, HJs (and WJs) are only formed for the smallest viscous time-scale of $\approx 1.4 \times 10^{-2}$ yr, which is 100 smaller (i.e. 100 times more efficient tides) compared

¹<http://openexoplanetcatalogue.com/>. Date of retrieval: May 3 2016.



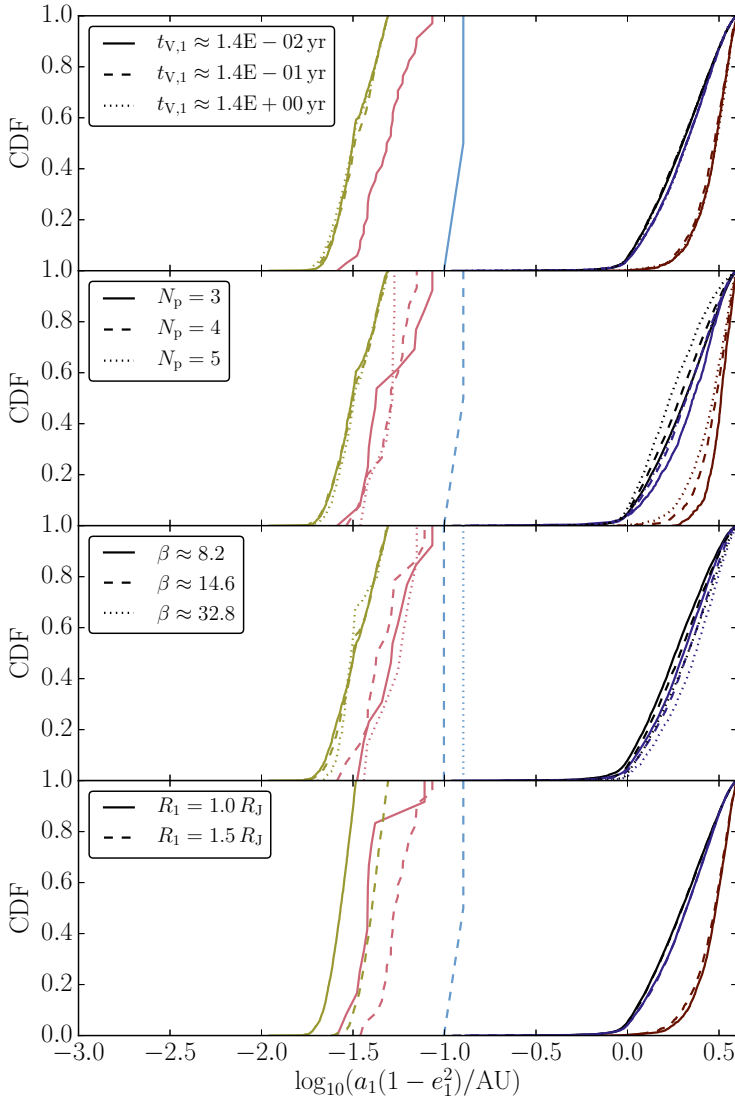


Figure 7.18: The innermost orbit semilatus rectum distributions after 10 Gyr (or after a stopping condition was met) are plotted for different slices of the parameter space. In each panel, distributions are shown for all parameters combined, except one (different linestyles). The varying parameters are $t_{V,1}$, N_p , β and R_1 in the top through bottom panels, respectively. Different types of systems are indicated with different colors: systems with no migration (black), HJs (light red), WJs (light blue), tidally disrupted inner planets (yellow), orbit crossings (dark blue) or exceeding of run time (dark red).

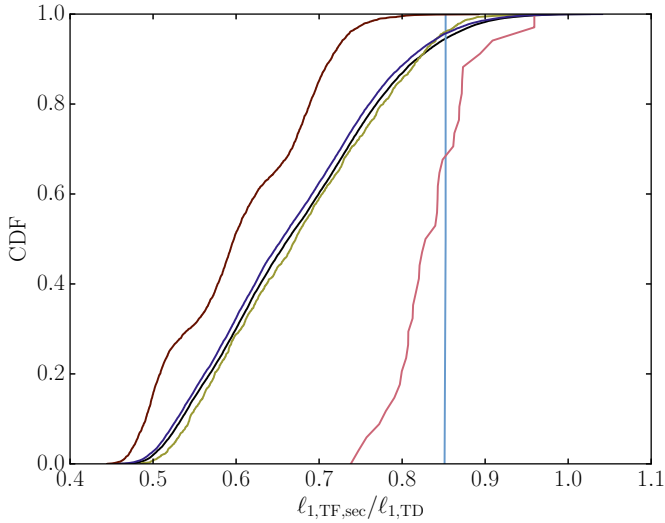


Figure 7.19: The initial distributions of the ratio $\ell_{1,TF,sec}/\ell_{1,TD}$ (cf. equation 7.12) for the various outcomes of the Monte Carlo simulations. Colors indicate systems with no migration (black), HJs (light red), WJs (light blue), tidally disrupted inner planets (yellow), orbit crossings (dark blue) or exceeding of run time (dark red).

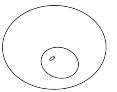
to the value that was obtained by Socrates et al. (2012) (this was also shown in Table 7.2). We note that this requirement for highly efficient tides was also found for other high- e migration scenarios, in particular in stellar binaries (Petrovich 2015a). At face value, this might suggest that high- e migration cannot explain HJs. However, there are still large uncertainties associated with tidal dissipation in stars and giant planets, therefore it may be premature to exclude values of $t_{v,1}$ as short as $\approx 1.4 \times 10^{-2}$ yr.

The dependence on the number of planets and β is shown in the second and third panels of Figure 7.18, respectively. Despite the propensity of exciting higher eccentricities with a larger number of planets and/or smaller β shown previously in Section 7.4, there is no strong dependence of the semilatus rectum distributions on these parameters.

Lastly, the dependence on the radius of the innermost planet is shown in the bottom panel of Figure 7.18. For HJs, the final semilatus rectum is smaller for smaller radii, consistent with the estimates in Section 7.4.3. For the tidally disrupted planets, a larger radius corresponds to a larger tidal disruption radius (cf. equation 7.7), and therefore a larger semilatus rectum at the moment of disruption.

Tidal disruptions

The large proportion of tidal disruptions compared to HJs can be understood with the arguments given in Section 7.4.3. In the latter section, it was shown that the tidal dissipation strength plays an important role in determining whether high eccentricities result in migration, or tidal disruption. In particular, the transition between the two in terms of $t_{v,1}$ was shown to be reasonably



approximated by equation (7.14). Using the same arguments, for a given system in which the innermost orbital eccentricity is highly excited, we expect tidal disruption if $\ell_{1,\text{TF,sec}} < \ell_{1,\text{TD}}$ (cf. equation 7.12). The boundary is expected to correspond to approximately $\ell_{1,\text{TF,sec}} = \ell_{1,\text{TD}}$.

In Figure 7.19, we show the distributions of the ratio $\ell_{1,\text{TF,sec}}/\ell_{1,\text{TD}}$ for the various outcomes of the Monte Carlo simulations. These ratios are invariably < 1 , suggesting that tidal disruption is more likely than migration. This is consistent with our finding in the Monte Carlo simulations that the fractions of HJs are small compared to the fractions of tidal disruptions. As might be expected, the tidal disruption systems show a preference for smaller ratios compared to the HJs (and WJs). The latter show a strong preference for the largest ratios, closest to unity.

Dependence on the initial orbital properties

In Section 7.4, we investigated the conditions for exciting the eccentricity of the innermost orbit to high values, and concluded that this can be the case if the AMD is $\gtrsim 1 m_1 \sqrt{a_1}$ and $\mathcal{R}_0^{(ijk)}$ is close to unity (for either one or more triplet of orbits). The initial distributions of the AMD and $\mathcal{R}_0^{(ijk)}$ for our Monte Carlo simulations were shown in Figure 7.16. In Figure 7.20, we show how the various outcomes of the simulations depend on the AMD (top panel) and the $\mathcal{R}_0^{(ijk)}$ (lower three panels).

With regard to the AMD, there are clear differences between the outcomes. Of the non-migrating systems (black line), nearly all (≈ 0.99) have an AMD which is $< m_1 \sqrt{a_1}$. In contrast, all disrupted systems have an AMD $\gtrsim 1 m_1 \sqrt{a_1}$, and the HJ systems have an AMD $\gtrsim 0.8 m_1 \sqrt{a_1}$, with the majority (~ 0.8) of systems having AMD $> 1 m_1 \sqrt{a_1}$. The systems in which orbit crossings occurred (blue lines) have higher AMD compared to the non-migrating systems, which can be attributed to higher eccentricities with higher AMDs, therefore more likely leading to an orbit crossing. However, for these systems ≈ 0.8 have $< 1 m_1 \sqrt{a_1}$, indicating that the majority of integrations of systems with orbit crossings would not produce HJs or tidally disrupted planets if not stopping at the onset of orbit crossings. Similarly, the systems with exceeded run times (dark red lines) have low AMD comparable to the non-migrating systems, indicating that these systems would produce few to no HJs or tidally disrupted planets if the integrations were not stopped before reaching 10 Gyr.

There are also clear differences with respect to the AMD between the HJ and tidal disruption systems. Most (≈ 0.9) of the HJs have AMD $< 1.5 m_1 \sqrt{a_1}$, whereas for the tidal disruption systems this fraction is markedly lower, ≈ 0.4 . The preference for the tidal disruption systems for higher AMDs can be explained by the higher eccentricities reached in the innermost orbit, therefore more likely resulting in the (immediate) tidal disruption of the innermost planet, rather than tidal dissipation, which requires a certain amount of time to dissipate energy and reduce the eccentricity. This implies that there is a ‘window’ for producing HJs through secular evolution: the AMD should be large enough to excite high eccentricities (i.e. AMD $\gtrsim 1 m_1 \sqrt{a_1}$, but small enough to prevent violent excitation of the eccentricities that lead to tidal disruption before tidal dissipation can be effective.

With regard to $\mathcal{R}_0^{(ijk)}$, differences between the outcomes arise primarily with respect to $\mathcal{R}_0^{(123)}$. This is consistent with our finding in Section 7.4.2 that for ≤ 5 planets, $\mathcal{R}_0^{(123)}$ is the most crucial of the $\mathcal{R}_0^{(ijk)}$ quantities insofar as to drive high eccentricities. The HJ and tidal disruption systems show a preference for values of $\mathcal{R}_0^{(123)}$ close to unity, whereas the other outcomes show preferentially lower values of $\mathcal{R}_0^{(123)}$. The systems with exceeded run times

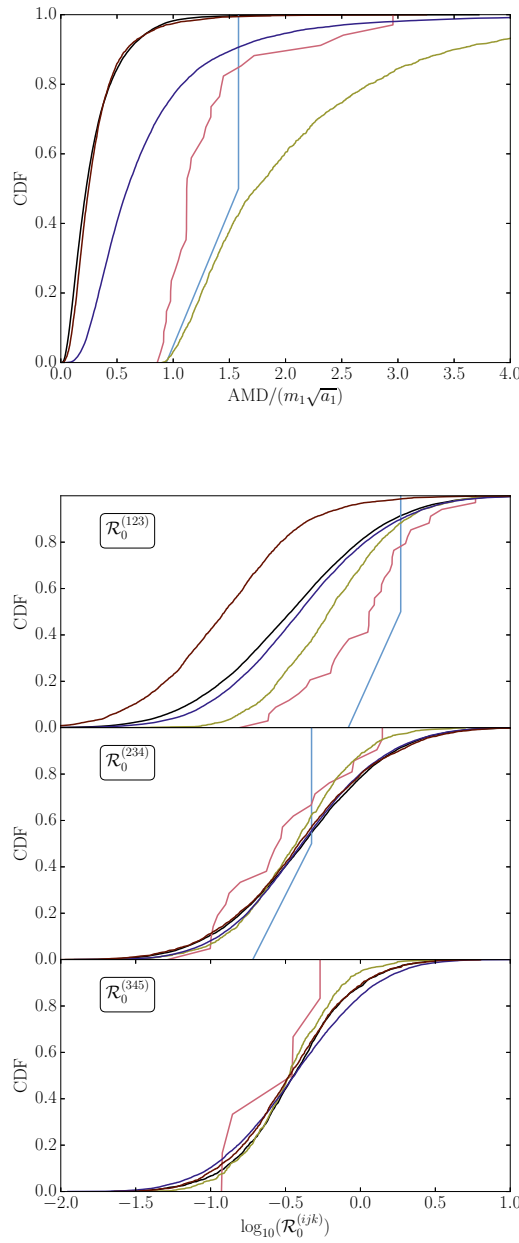
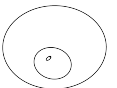


Figure 7.20: The initial distributions of the AMD (top panel) and the quantity $\mathcal{R}_0^{(ijk)}$ (lower three panels) for the various outcomes of the Monte Carlo simulations (cf. Figure 7.16, in which these initial distributions were shown for all sampled systems). Colors indicate systems with no migration (black), HJs (light red), WJs (light blue), tidally disrupted inner planets (yellow), orbit crossings (dark blue) or exceeding of run time (dark red).



typically have the lowest values of $\mathcal{R}_0^{(123)}$, and these can be identified as systems with large a_1 and small a_2 (see below). The exceeding of the run time can be understood by noting that in this case, the LK time-scale associated with the innermost pair (1-2) is particularly short, implying that the associated number of oscillations in 10 Gyr is very large, and requiring a long time to integrate (evidently, longer than our maximum allowed CPU time of 4 hr). With regard to $\mathcal{R}_0^{(234)}$ and $\mathcal{R}_0^{(345)}$, we find a slightly larger preference for HJs and tidal disruption systems with $-0.5 \lesssim \log_{10}(\mathcal{R}_0^{(234)}) \lesssim 0$ and $-0.5 \lesssim \log_{10}(\mathcal{R}_0^{(345)}) \lesssim 0$ compared to the other outcomes, consistent with Section 7.4.2.

Both the AMD and $\mathcal{R}_0^{(ijk)}$ are useful quantities for theoretical considerations. However, for the purpose of comparing to observations, it is more convenient to consider the semimajor axes and eccentricities. In Figure 7.21, we show how the various outcomes in our simulations depend on the initial semimajor axes (left column) and the eccentricities (right column), for the first three orbits. We recall that the semimajor axes were sampled linearly from fixed ranges, whereas the eccentricities were sampled from a Rayleigh distribution with various widths expressed by the grid parameter β (cf. Section 7.5.1).

As also reflected by the $\mathcal{R}_0^{(ijk)}$, the largest differences in the initial semimajor axes between the various outcomes occur in the innermost orbit. The distribution of the initial a_1 for the HJ and tidal disruption systems is skewed towards small values compared to the other systems, with $a_1 \lesssim 2$ AU for most of the systems (≈ 0.95 and ≈ 0.8 , respectively). This can be attributed to two effects: $\mathcal{R}_0^{(123)}$ is larger, in particular closer to unity, if a_1 is smaller (cf. equation 7.5). Furthermore, the required eccentricities for small pericenter distances (and hence tidal dissipation or tidal disruption) are lower for smaller semimajor axes. The systems with exceeded run times (red lines) preferentially have large a_1 , whereas a_2 is preferentially small. As mentioned above, this can be attributed to the short LK time-scales for the (1-2) pair in these cases. With regard to the other orbits and outcomes, no clear differences can be discerned in the initial distributions of the semimajor axes.

HJ and tidal disruption systems typically show a preference for initially higher values of the eccentricities, which can be explained by the larger associated values of the AMD. With regard to e_2 , there appears to be a somewhat stronger dependence on higher values of HJ and tidal disruption systems. This can be explained by the associated larger value of $\mathcal{R}_0^{(123)}$. Note that the dependence on $\mathcal{R}_0^{(234)}$ and $\mathcal{R}_0^{(345)}$ (and therefore e_3 , e_4 and e_5) is weaker (cf. Figure 7.20).

Observations of companion planets to HJs are currently still strongly limited. In the latest surveys, detections of Jupiter-mass planets are only 100% complete for planets out to ≈ 10 AU (Bryan et al. 2016). In our simulations, except for the innermost two planets, the orbits span a much larger range in semimajor axis. Despite the incompleteness of the observations, we show in Figure 7.22 observed distributions of the semimajor axes (top panel) and the eccentricities (bottom panel) for planets with $m_1 \sin(i) > 0.1 M_J$ obtained from the *Open Exoplanet Catalogue*. We consider multiplanet systems with at least two observed planets, and make a distinction between systems with the innermost observed orbit $a_1 < 1$ AU (solid lines) and $a_1 > 1$ AU (dashed lines). Some difference can be seen in the distributions of a_2 for the two populations with $a_1 < 1$ AU and $a_1 > 1$ AU – a_2 tends to be smaller for the former population. This trend is not reflected in our simulations (cf. the second row of Figure 7.21). With regard to a_3 , there seems to be large difference in the distribution of a_3 for the two populations. However, our observational sample only includes 5 systems with at least 3 planets, so with this low number of systems this difference cannot be considered significant.

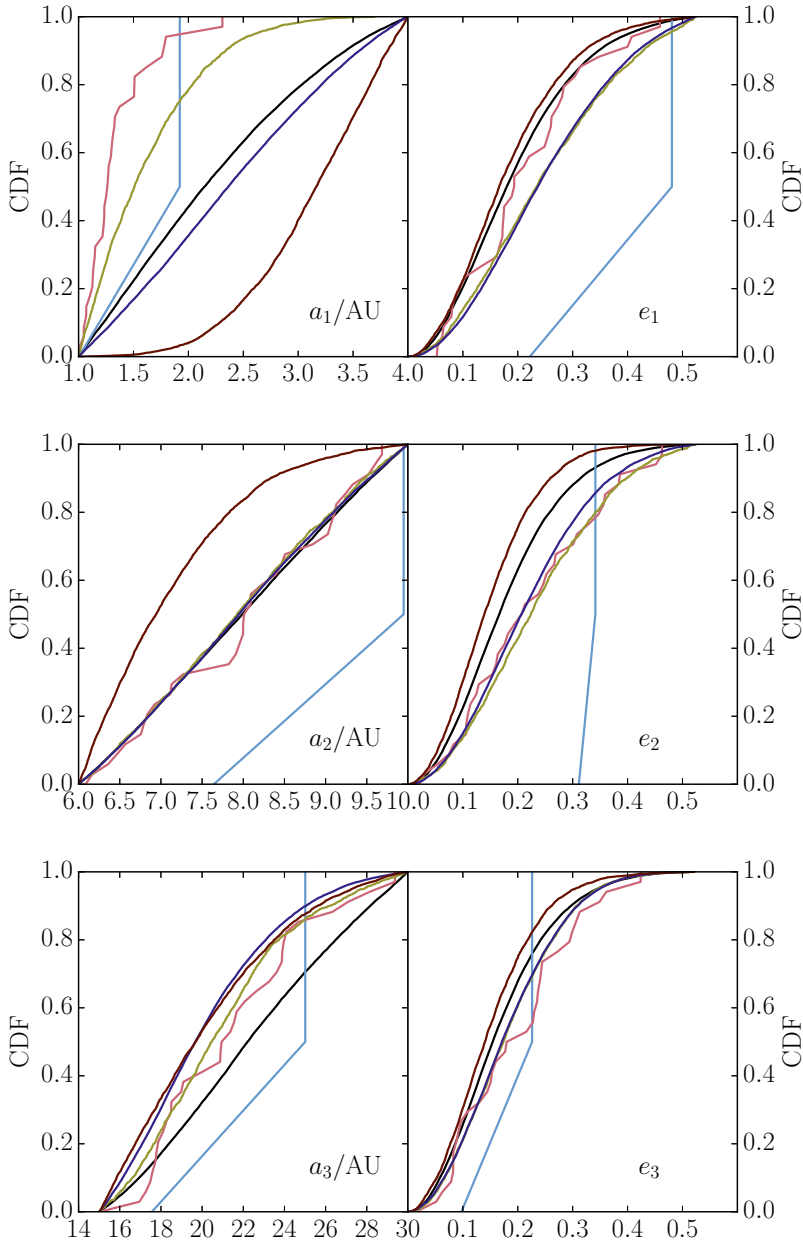
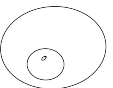


Figure 7.21: The initial distributions of the semimajor axes (left column) and the eccentricities (right column) for the various outcomes of the Monte Carlo simulations. Colors indicate systems with no migration (black), HJs (light red), WJs (blue), tidally disrupted inner planets (yellow), orbit crossings (dark blue) or exceeding of run time (dark red).



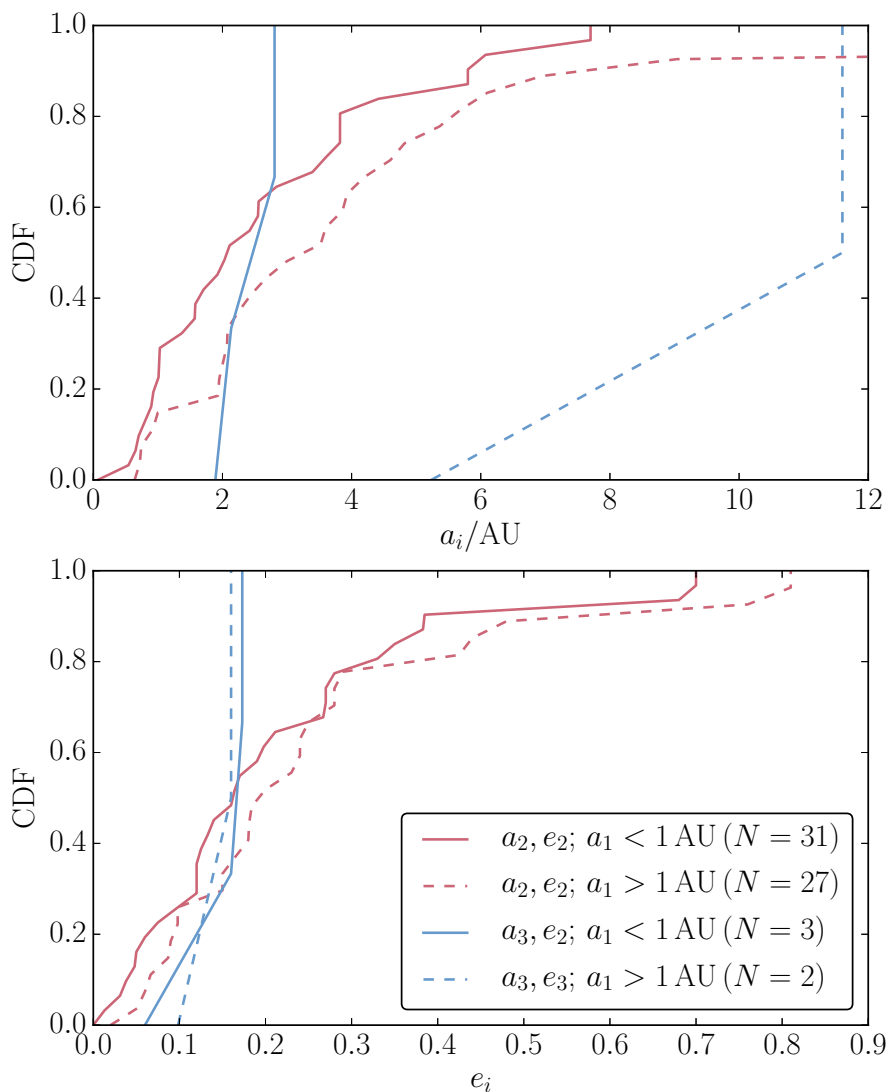


Figure 7.22: Observed distributions of the semimajor axes (top panel) and the eccentricities (bottom panel) for planets with $m_1 \sin(i) > 0.1 M_J$ obtained from the *Open Exoplanet Catalogue*. We consider multiplanet systems with at least two observed planets, and make a distinction between systems with the innermost observed orbit $a_1 < 1$ AU (solid lines) and $a_1 > 1$ AU (dashed lines).

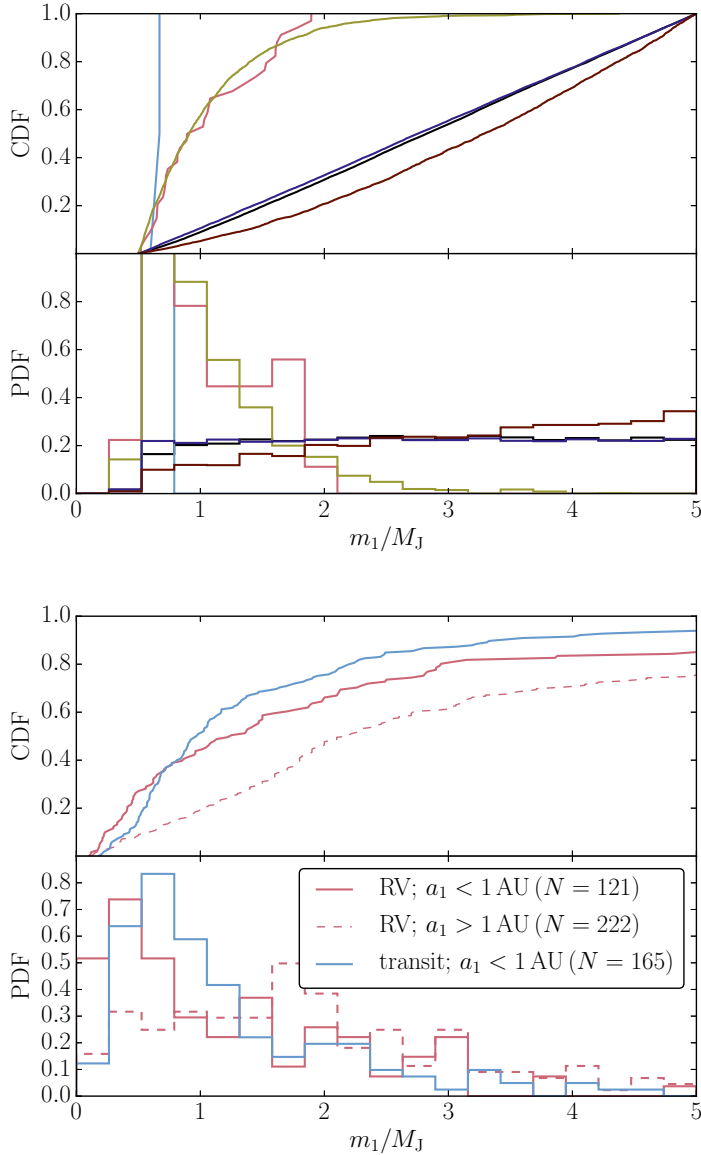
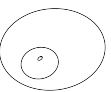


Figure 7.23: Top panels: distributions of the initial mass of the innermost planet for the Monte Carlo simulations with all grid parameters combined. Different types of systems are indicated with different colors: systems with no migration (black), HJs (light red), WJs (blue), tidally disrupted inner planets (yellow), orbit crossings (dark blue) or exceeding of run time (dark red). Bottom panels: the observed mass distributions of planets with $m_1 \sin(i) > 0.1 M_J$ (obtained from the *Open Exoplanet Catalogue*). Two discovery methods are included: RV (red) and transit (blue). Also, a distinction is made between the semimajor axis: $a_1 < 1$ AU (solid lines) and $a_1 > 1$ AU (dashed lines). At the time of writing, the database did not contain transiting planets with $m_1 \sin(i) > 0.1 M_J$ and $a_1 > 1$ AU.



Mass dependence

In our synthesized populations, we assumed a flat distribution of the mass of the innermost planet between 0.5 and 5 M_J . The orbits of lower-mass planets carry less orbital angular momentum compared to higher-mass counterparts, making the former more susceptible to angular-momentum exchanges with other orbits. Therefore, secular eccentricity excitation is expected to be more pronounced if the outer planets are more massive than the innermost planet.

In the top panel of Figure 7.23, we show the distributions of the mass of the innermost planet for the various types of systems. For the non-migrating systems (black line), the mass distribution is consistent with a flat distribution (a linear CDF), reflecting the initial distribution, and showing no mass preferences. For HJ and tidal disruption systems, the mass distributions are different. For the latter groups, there is a clear preference for lower-mass planets, with ≈ 0.5 of the systems having a mass $m_1 \lesssim 0.9 M_J$. There are no HJs, and few tidal disruption systems with $m_1 \gtrsim 2 M_J$. This implies a clear prediction for high- e migration in multiplanet systems: the HJ should have a typical (median) mass of $\sim 1 M_J$, and not be more massive than $\sim 2 M_J$.

The observed mass distributions of planets with $m_1 \sin(i) > 0.1 M_J$ (obtained from the *Open Exoplanet Catalogue*) are shown in the bottom panel of Figure 7.23. Here, we made a distinction between discovery method (RV or transit) and semimajor axis ($a_1 < 1$ AU and $a_1 > 1$ AU). Note that the time of writing, the database did not contain transiting planets $m_1 \sin(i) > 0.1 M_J$ and $a_1 > 1$ AU. The planets from the RV observations within 1 AU are typically of lower mass compared to planets at > 1 AU. This trend is consistent with the predictions of secular evolution as described above. However, one should be cautious when attributing this mass difference to secular evolution alone, given that the latter unlikely produces all HJs, and the observed mass distribution is also likely affected by other processes, such as primordial ‘mass segregation’, whereas in the simulations we assumed an initially linear distribution. Moreover, the RV observations are likely biased, because planets at > 1 AU are more easily detected if they are more massive.

HJ formation times

As mentioned in Section 7.1, evolution due to secular evolution occurs on potentially Gyr-timescales, implying that HJs formed through this mechanism could be formed at relatively late stages in the MS life-time of the host star. In Figure 7.24 we show the ‘final’ times associated with the various outcomes in our population synthesis (with all parameters combined). For systems in which a stopping condition occurred, these final times are the age of the system when the integration was stopped. In the case of HJs (light red line), these times are the times of HJ formation. As shown in Figure 7.24, the HJs in our simulations are indeed formed relatively late, with a median formation time of ≈ 2 Gyr, and ≈ 0.2 of the HJs formed after ≈ 6 Gyr. These times are much longer compared to high- e migration due to close encounters (Rasio & Ford 1996; Chatterjee et al. 2008; Ford & Rasio 2008; Jurić & Tremaine 2008), and somewhat longer compared to those typically found for high- e migration in stellar binaries (Anderson et al. 2016).

Stellar obliquity

In Figure 7.25, we show, for all grid parameters combined, the distributions after 10 Gyr (or after a stopping condition) of the stellar obliquity, i.e. the angle between the stellar spin and

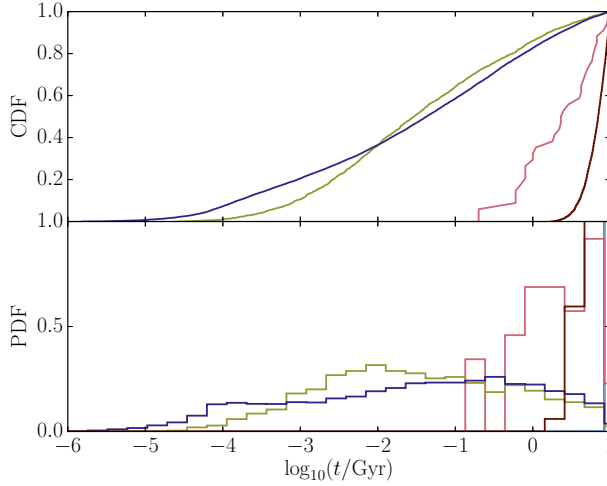


Figure 7.24: The final times associated with the various outcomes in our population synthesis (with all parameters combined). For systems in which a stopping condition occurred, these final times are the age of the system when the integration was stopped. In the case of HJs (light red line), these times are the times of HJ formation. Different types of systems are indicated with different colors: systems with no migration (black), HJs (light red), WJs (blue), tidally disrupted inner planets (yellow), orbit crossings (dark blue) or exceeding of run time (dark red).

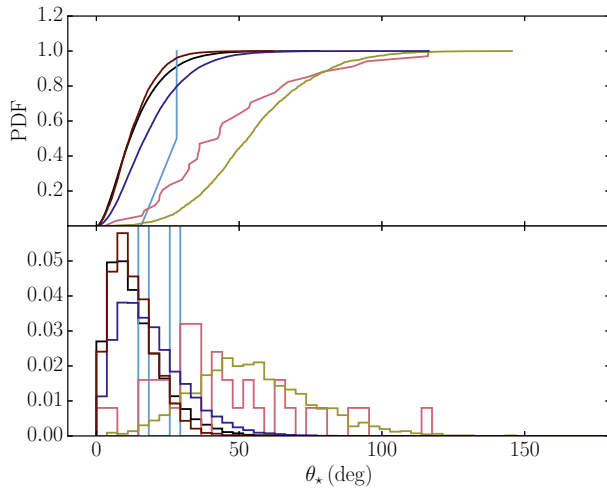
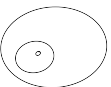


Figure 7.25: Distributions after 10 Gyr (or after a stopping condition) of the stellar obliquity θ_* for the Monte Carlo simulations with all grid parameters combined. As in Figure 7.18, different types of systems are indicated with different colors: systems with no migration (black), HJs (light red), WJs (blue), tidally disrupted inner planets (yellow), orbit crossings (dark blue) or exceeding of run time (dark red).



the orbit of the innermost planet. For high- e migration scenarios with three bodies (or, more generally, LKCTF), it has been well established that the obliquity for HJ systems should be clustered around 40° and 130° (Fabrycky & Tremaine 2007; Anderson et al. 2016). This can be explained broadly by noting that during LK cycles, the eccentricity maxima occur at mutual inclinations of $\approx 40^\circ$ or 130° , and this angle tends to be ‘locked’ after the onset of strong tidal dissipation (this is assuming that the stellar spin vector itself does not change due to spin-planet coupling, see Storch et al. 2014; Storch & Lai 2015; Anderson et al. 2016).

In the case of multiplanet systems, we find that for the HJs and tidally disrupted planets, the obliquity distributions are distinct from the distribution of non-migrating systems. The former distributions are broadly distributed between $\sim 0^\circ$ and $\sim 120^\circ$, with some preference for obliquities around 40° (for HJs) or 50° (for tidal disruption systems). There is, however, not a strongly peaked distribution around 40° , and no peak around 130° , as found e.g. by Anderson et al. (2016), who considered high- e migration in stellar binaries. The absence of a strongly peaked distribution around 40° can be understood qualitatively by noting that in the case of secular evolution in multiplanet systems, there is no characteristic minimum angle for exciting high eccentricities (of course, higher inclinations do imply higher eccentricities, as shown in Section 7.4.1). Also, in the case of high- e migration in stellar binaries, the inclinations are usually assumed to be random, implying a roughly equal occurrence of prograde and retrograde orbits, and explaining a pileup at both $\sim 40^\circ$ and $\sim 130^\circ$. This is not the case in our simulations, however, in which the planets were always initially prograde. Nevertheless, we still find that ~ 0.1 of the HJs systems have retrograde obliquities.

7.6 Discussion

7.6.1 HJ contribution and comparisons to other variants of high- e migration

In our population synthesis simulations, the highest HJ fraction obtained was 0.007, assuming a high tidal dissipation efficiency which is 100 times larger than the value obtained for high- e migration by Socrates et al. (2012). For the purposes of this section, we round this fraction to $f_{\text{HJ},\text{sim}} = 0.01$, and adopt this value as a maximum, and ‘optimistic’ fraction. We remark that higher HJ fractions would be obtained in simulations with even smaller values of $t_{\text{V},1}$ (cf. Section 7.4.3), but here we focus on the results of our population synthesis simulations only, and so do not consider this possibility.

Assuming a giant planet occurrence rate of $f_{\text{GP}} = 0.1$ and a multiplanet fraction (three or more planets) of $f_{\text{MP}} = 0.1$, the fraction of systems with HJs due to high- e migration is $f_{\text{HJ}} = f_{\text{HJ},\text{sim}} f_{\text{GP}} f_{\text{MP}} = 10^{-4}$. The observed HJ fraction is $f_{\text{HJ},\text{obs}} \sim 0.01$ (e.g. Wright et al. 2012), implying that at most 0.01 of the HJs can be explained through secular evolution in multiplanet systems. We remark, however, that the assumed fraction of $f_{\text{MP}} = 0.1$ is based on current RV observations (from the *Open Exoplanet Database*), which are not complete for masses $\gtrsim 1 M_J$ and semimajor axes $\gtrsim 10 \text{ AU}$ (e.g. Bryan et al. 2016). The true fraction is likely higher, implying a larger contribution of secular evolution in multiplanet systems to the HJ population.

Our HJ contribution fraction estimate f_{HJ} is up to ~ 10 lower compared to studies of high- e migration in 2-planet or stellar binary systems. For example, Anderson et al. (2016) find $f_{\text{HJ}} \sim$

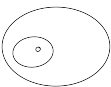
0.15 for high- e migration in stellar binaries (their adopted simulated HJ fraction is $f_{\text{HJ},\text{sim}} = 0.03$, and the binary companion fraction is assumed to be 0.5). In Petrovich & Tremaine (2016), the 2-planet case is considered, and a HJ fraction of $f_{\text{HJ},\text{sim}} = 0.051$ is found.

7.7 Conclusions

We have studied the formation of hot Jupiters (HJs) through secular excitation of eccentricities in multiplanet systems (≥ 3 planets). In this variant of high- e migration, the eccentricity of the orbit of the innermost planet in a multiplanet system (at least three planets) is excited, on secular time-scales, to high values. Combined with tidal dissipation, which is highly effective at high eccentricities, this can result in a Jupiter-like planet in a close orbit around its star.

Our conclusions are as follows.

1. For a set of three-planet systems taking into account tidal dissipation in the innermost planet, we have compared results from our secular code `SECULARMULTIPLE` (Hamers & Portegies Zwart 2016) to the highly accurate direct N -body code `ARCHAIN` (Mikkola & Merritt 2008) (Section 7.3). Statistically, the two methods yield similar results. Specifically, with regard to the semilatus rectum distributions of the HJs formed in these simulations, the D and p -values of the K-S test are ≈ 0.082 and ≈ 0.881 , respectively. Because of the much greater speed of `SECULARMULTIPLE` (100-Myr integrations which require ~ 1 week of integration, can be completed in several minutes with `SECULARMULTIPLE`), the latter code allows a detailed exploration of the large parameter space of multiplanet systems.
2. Using `SECULARMULTIPLE`, we studied the conditions for producing high eccentricities in the innermost orbit in multiplanet systems with 3, 4 or 5 planets (Section 7.4). Such high eccentricities are required to produce HJs through high- e migration. We find that the requirement $\text{AMD} > 1 m_1 \sqrt{a_1}$, where AMD is the angular-momentum deficit of the system, is a necessary, but not complete condition for exciting high eccentricities. In addition to the AMD, we have shown that the ratio $\mathcal{R}_0^{(ijk)}$ of Lidov-Kozai (LK) time-scales (Lidov 1962; Kozai 1962) associated with three adjacent orbits (i, j, k) is another important parameter for eccentricity excitation. This parameter is analogous to the quantity $\mathcal{R}_0$ that applies to ‘3+1’ hierarchical quadruple systems (Hamers et al. 2015). Values of $\mathcal{R}_0^{(ijk)}$ close to unity correspond to commensurate LK-times-scales for the pairs (i, j) and (j, k) , which give rise to coupled and resonant secular oscillations, potentially driving high eccentricities. In contrast, for values of $\mathcal{R}_0^{(ijk)}$ much less than, or exceeding unity, the secular oscillations tend to be ‘detuned’. For planetary systems with 4 or 5 planets, we have found that the strongest dependence of eccentricity excitation is on $\mathcal{R}_0^{(123)}$. A weaker, but still distinguishable dependence, can be observed on the quantities $\mathcal{R}_0^{(234)}$ and $\mathcal{R}_0^{(345)}$.
3. We carried out a population synthesis study of multiplanet systems with `SECULARMULTIPLE`, taking into account tidal dissipation in the innermost planet and the central star (Section 7.5). For 54 combinations of the viscous time-scale of the innermost planet $t_{\text{V},1}$, the number of planets $N_p \in \{3, 4, 5\}$, the width of the inclination and eccentricity distribution β and the radius of the innermost planet R_1 , we integrated 1000 Monte Carlo realisations of systems for up to 10 Gyr and with planetary masses sampled linearly between 0.5 and 5 M_J , and semimajor axes sampled linearly in separate ranges for each planet (up to 100 AU).



4. In our population synthesis computations, we have found HJ fractions of at most 0.007, assuming a tidal dissipation strength that is a factor 100 higher compared to the value inferred for high- e migration by Socrates et al. (2012). For weaker tidal dissipation (higher $t_{V,1}$), we find no HJs among the 1000 integrations for each parameter combination. These HJ fractions can be considered low, given that other variants of high- e migration (in particular, an S-type planet in a stellar binary) typically give HJ fractions of ~ 0.01 to ~ 0.1 . Assuming a giant planet fraction of 0.1 and an observed HJ fraction of 0.01, our results suggest that secular evolution in multiplanet systems can explain at most ~ 0.01 of observed HJ systems (assuming an ‘optimistic’ HJ fraction of 0.01). We note that larger fractions would be obtained for even lower values of the innermost planet viscous time-scale, i.e. $t_{V,1} < 1.4 \times 10^{-2}$ yr.

5. On the other hand, we find that in up to ~ 0.3 of the systems, the innermost planet is tidally disrupted. The large proportion of tidally disrupted planets to HJs can be explained by noting that the eccentricity of the innermost orbit can be violently excited if the AMD is sufficiently large, implying that the planet is rapidly tidally disrupted before tidal dissipation is able to shrink and circularise the orbit. This implies that there is a ‘window’ in terms of the AMD for which HJs can be produced, in our simulations between $1 \gtrsim \text{AMD}/(m_1\sqrt{a_1}) \gtrsim 2$ (cf. Figure 7.20). In addition to a specific AMD, we find that HJs are preferentially produced for systems with $\mathcal{R}_0^{(123)}$ close to unity. In our simulations, this translates to a preference for initial values of the semimajor axis of the innermost orbit of between 1 and 2 AU. This preference is also due to the smaller eccentricity required for small pericenter distances. The large fractions of tidal disruptions in some of our simulations suggest that tidal disruptions in multiplanet systems, even if not extremely compact, can be fairly common. For non-compact planetary systems, this suggests a possible difference in metallicity between stars with two or fewer planets, compared to stars with three or more planets.

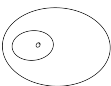
6. The semimajor axis distribution of the HJs in our simulations is strongly peaked around 0.05 AU, and this peak coincides with the peak in the observed semimajor axis distribution of exoplanets with $m_1 \sin(i) > 0.1 M_J$. However, in our simulations, no planets are produced with final semimajor axes in the ‘period valley’ between 0.1 and 1 AU, whereas observations show a significant population of planets with these semimajor axes, i.e. warm Jupiters (WJs). It is unlikely that these WJs are produced through secular evolution in multiplanet systems, unless the tidal dissipation is implausibly strong (cf. Section 7.4.3). Other high- e migration scenarios also fail to produce WJs in the observed quantities (Antonini et al. 2016; Petrovich & Tremaine 2016). WJs are therefore likely produced via other mechanisms, such as *in situ* formation, or disk migration.

7. In our population synthesis, we found that HJs and tidally disrupted planets are preferentially not massive, i.e. $m_1 \lesssim 2 M_J$, with a median value of $\approx 1 M_J$ (cf. Section 7.5.3). The stellar obliquity distribution is fairly uniform between $\sim 0^\circ$ and $\sim 120^\circ$ and does not show a strong peak at $\sim 40^\circ$ and no peak at $\sim 130^\circ$, as opposed to high- e migration in stellar binary or 2-planet systems (cf. Section 7.5.3).

8. Another characteristic of HJs formed in our simulations and which is relevant for observations, is the relatively late formation time of up to ~ 10 Gyr (cf. Section 7.5.3). This is in stark contrast to disk migration, for which formation is expected to occur within the first few Myr. Furthermore, this aspect can potentially distinguish between other variants of high- e migration, which typically predict shorter formation times.

Acknowledgements

This work was supported by the Netherlands Research Council NWO (grants #639.073.803 [VICI], #614.061.608 [AMUSE] and #612.071.305 [LGM]) and the Netherlands Research School for Astronomy (NOVA). We acknowledge the use of the astric computer cluster of the national Israeli astrophysics I-CORE center. HBP acknowledges support from European union career integration grant ‘GRAND’, the Minerva center for life under extreme planetary conditions and the Israel science foundation excellence center I-CORE grant 1829.



Samenvatting

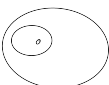
We hebben de gravitationele dynamische evolutie van hiërarchische systemen met verschillende hiërarchieën bestudeerd in verschillende astrofysische konteksten. We begonnen met ‘multiplexsystemen’ met een supermassief zwart gat. Daarna hebben we ‘simplexsystemen’ bekeken, beginnend met de algemene dynamica van quadruplesystemen (Hoofdstuk 4), en later met een toepassing op planeten rondom dubbelsterren, ofwel *circumbinary planets* (Hoofdstuk 5). Ten slotte hebben wij onze methoden gegeneraliseerd naar hiërarchische simplexsystemen met een willekeurige structuur en willekeurig aantal lichamen (Hoofdstuk 6), en hebben wij deze toegepast op de vorming van *hot Jupiters* via migratie met hoge excentriciteit in multiplaneet-systemen (Hoofdstuk 7).

Hoofdstuk 2 – Relativistische dynamica rondom een supermassieve zwart gat

In dit hoofdstuk hebben we de gravitationele langetermijnevolutie bestudeerd van banen rondom een supermassief zwart gat. We hebben een specifieke N -bodycode ontwikkeld, TPI , welke ons in staat heeft gebracht om dergelijke banen efficiënt te integreren met meeneming van relativistische correcties, en interacties met een groot aantal ($\sim 10^3$) ‘veldsterren’, welke per aanname in (vrijwel) vaste Keplerbanen bewegen. We hebben laten zien dat de resultaten van TPI overeenkomen met andere meer nauwkeurige, tevens langzamere N -bodycodes.

Met behulp van TPI hebben we simulaties uitgevoerd van het Galactisch centrum (GC) om modellen van de oorsprong van de S-sterren te testen. Hierbij zijn 4800 veldsterren aangenomen, een aantal dat voorheen niet haalbaar was met andere N -bodycodes. We namen aan dat de S-sterren oorspronkelijk in sterk excentrische banen waren, consistent met de ontbinding van dubbelsterren door de getijdenkracht van het supermassieve zwarte gat. We hebben gevonden dat de cumulatieve excentriciteitsverdeling van de S-sterren evolueert tot $N(e) \sim e^{2.6}$ op een tijdschaal van 7 ± 0.1 Myr, wat consistent is met waarnemingen. Onze resultaten suggereren een onderlimiet van de typische leeftijd van de S-sterren van ~ 7 Myr en ~ 25 Myr, aangenomen dat de S-sterren respectievelijk in een korte tijdspanne, of geleidelijk zijn gevormd.

Verder hebben wij uit onze simulaties eerste- en tweede-orde diffusiecoëfficiënten bepaald in termen van de genormaliseerde impulsmomentvariable $\ell \equiv \sqrt{1 - e^2}$. Deze coëfficiënten beschrijven de langetermijnevolutie van het (baan)impulsmoment. We identificeerden drie impulsmomentregimes, waarin de diffusiecoëfficiënten verschillend afhangen van ℓ . Voor lage ℓ kunnen de diffusiecoëfficiënten goed beschreven worden in termen van *non-resonant re-*



laxation, dat wil zeggen dynamische relaxatie vanwege ontmoetingen tussen twee lichamen. Voor hoge ℓ kunnen de diffusiecoëfficiënten goed beschreven worden in termen van *resonant relaxation*, gerelateerd aan krachtmomenten werkend op de banen, welke afkomstig zijn van gecorreleerde ontmoetingen tussen lichamen. In het tussenregime is de zogenaamde ‘Schwarzschildbarrière’, waarin relativistische precessie van de banen belangrijk is. In dit regime hebben we een nieuwe vorm van relaxatie geïdentificeerd, namelijk *anomalous relaxation*, waarin diffusie wordt gedreven door een combinatie van *resonant relaxation* en snelle relativistische precessie. In het bijzonder hebben we laten zien dat de Schwarzschildbarrière geassocieerd kan worden met een sterke daling van de diffusiecoëfficiënten met afnemende ℓ . We hebben onze resultaten toegepast om de *steady-state*-verdeling (de verdeling geëxtrapoleerd tot het oneindige) van ℓ te bepalen.

Hoofdstuk 3 – Planetesimalen in het GC

Verdergaand met het vorige onderwerp van banen rondom een supermassief zwart gat, hebben we de dynamica van planetesimalen in het GC bestuurd. Hierbij hebben we onze aandacht gevestigd op de vernietiging van planetesimalen door de getijdenkracht van het supermassieve zwarte gat, welke mogelijk een waarneembare flits kan produceren in nabij-infrarood- en röntgengolflengtes. Dergelijke flitsen worden op ruwweg dagelijkse basis waargenomen vanuit het supermassieve zwarte gat in het GC, Sgr A*. We hebben aangenomen dat planetesimalen gevormd zijn ofwel in een grote-schaal wolk, en daarbij gebonden aan het supermassieve zwarte gat (scenario 1), ofwel in schijven rondom sterren (scenario 2). In het laatste scenario hebben we gevonden dat de getijdenkrachten van het supermassieve zwarte gat de planetesimalen van hun moederster wegtrekken op afstanden van $r \lesssim 0.5$ pc van het supermassieve zwarte gat. Planetesimalen worden ook van hun moederster ontbonden vanwege gravitationele interacties met andere sterren, en dit proces verwijdert bijna alle planetesimalen binnen de invloedssfeer van het supermassieve zwarte gat, ≈ 4 pc, na ~ 100 Myr.

We hebben de Fokker-Planckvergelijking in de energieruimte toegepast om de frequentie van vernietigingen van planetesimalen door het supermassieve zwarte gat uit te rekenen, en we hebben daarbij deze ontbindingseffecten van planetesimalen en andere potentieel belangrijke effecten meegenomen, zoals relaxatie vanwege interacties met zware objecten (bijvoorbeeld zware gaswolken). Onze vernietigingsfrequentie na $t = 10$ Gyr is $\approx 0.6 \text{ d}^{-1}$, en dit is ruwweg consistent met de waargenomen frequentie van flitsen van eenmaal per dag. Onze resultaten zijn ongevoelig voor de aannames van de modellen, in het bijzonder de beginverdeling van planetesimalen (dat wil zeggen scenario's 1 of 2). Voor scenario (2) is het aangenomen aantal planetesimalen per ster $N_{a/*} = 2 \times 10^7$ afgeleid van waarnemingen van schijven rondom sterren in de buurt van de Zon. Daarentegen impliceert ons resultaat dat in scenario (1) het aantal planetesimalen in de grote-schaal wolk rondom het supermassieve zwarte gat sterk gecorreleerd is met het aantal sterren, en dit vereist *finetuning* van $N_{a/*}$. Wij verkiezen de meer natuurlijke verklaring dat planetesimalen in galactische centra zoals het GC niet veel anders gevormd worden dan planetesimalen rondom sterren dichtbij de Zon.

Hoofdstuk 4 – Seculaire evolutie van hiërarchische quadruplesystemen

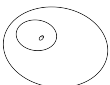
In de resterende hoofdstukken hebben we onze aandacht verlegd naar hiërarchische simplexsystemen, dat wil zeggen systemen welke bestaan uit geneste dubbelsterbanen. We hebben eerst hiërarchische quadruplesystemen in de ‘3+1’-configuratie bestudeerd, welke bestaan uit een hiërarchisch triplesysteem waaromheen een vierde lichaam beweegt. We namen aan dat het systeem voldoende hiërarchisch is, dat wil zeggen dat de verhoudingen x_i van de baanseparaties klein zijn, en we expandeerden de Hamiltoniaan in term van de x_i . Vervolgens hebben we de Hamiltoniaan gemiddeld over de banen, en hebben de bewegingsvergelijkingen geïmplementeerd in een computercode, `SECULARQUADRUPL`.

Daarna hebben we de seculaire evolutie van sterk hiërarchische systemen bestudeerd, welke goed beschreven worden door de laagste-orde termen in de Hamiltoniaan, en we hebben de evolutie van het systeem gekarakteriseerd in termen van de verhoudingen van Lidov-Kozai (LK) timeschalen, toegepast op verschillende baanparen. Laat de binnenste, middelste en buitenste banen aangeduid worden met respectievelijk A, B en C. Voor deze systemen hebben we gevonden dat de globale dynamica kwalitatief beschreven kan worden in termen van de initiële verhouding van de LK tijdschalen toegepast op de AB en BC baanparen, $\mathcal{R}_0 \equiv P_{\text{LK},\text{AB},0}/P_{\text{LK},\text{BC},0}$.

Als $\mathcal{R}_0 \ll 1$, dan domineert het krachtmoment van baan B op A ten opzichte van het krachtmoment van baan C op B. Als banen A en B initieel coplanair waren, dan blijven in dit geval deze banen coplanair. Als banen A en B initieel geïnclineerd ten opzichte van elkaar waren, dan zijn de LK-oscillaties in baan A niet veel veranderd door de aanwezigheid van het vierde lichaam. Excentriciteitsoscillaties in baan B vanwege het seculaire krachtmoment van baan C worden efficiënt uitgedoofd vanwege precessie op korte tijdschaal geïnduceerd door baan A op baan B.

Als $\mathcal{R}_0 \gg 1$, dan domineert het krachtmoment van baan C op baan B ten opzichte van het krachtmoment van baan B op baan A. Aanvankelijk verandert de inclinatie van baan B, terwijl dit niet het geval is voor baan A. Dit brengt een relatieve inclinatie tussen banen A en B teweeg, zelfs als deze aanvankelijk niet relatief geïnclineerd waren. Precessie van baan B, op een korte tijdschaal ten opzichte van de LK-tijdschaal toegepast op de AB banen, voorkomt echter significante excentriciteitsoscillaties in baan A, en dooft zelfs LK-oscillaties uit als banen A en B initieel geïnclineerd zijn.

Ten slotte, als $\mathcal{R}_0 \sim 1$, dan kunnen complexe LK-oscillaties voorkomen in baan A en welke sterk gekoppeld zijn met de LK-oscillaties in baan B. Deze laatste oscillaties zijn tevens anders ten opzichte van de situatie waarin baan A vervangen is met een puntdeeltje, hoewel dit typisch een minder sterk effect is. Zelfs als banen A en B initieel coplanair zijn, dan kan de geïnduceerde inclinatie resulteren in excentriciteitsoscillaties in baan A met grote amplitude. De extreem hoge excentriciteiten kunnen significante implicaties hebben voor sterke interacties zoals getijden, gravitatiegolven, en het botsen en/of samensmelten van sterren en compacte objecten.



Hoofdstuk 5 – Het verklaren van de afwezigheid van planeten rondom dubbelsterren met korte baanperiodes

In dit hoofdstuk hebben we SECULARQUADRUPLE toegepast om een recente curieuze eigenschap te verklaren in waarnemingen door *Kepler* van planeten rondom dubbelsterren, waargenomen door hun overgang over de sterren. Aangenomen dat het voorkomen van dergelijke planeten niet gerelateerd is aan de baanperiode van de dubbelster, dan zouden ongeveer tweemaal zoveel planeten ontdekt moeten zijn rondom dubbelsterren met korte baanperiodes ($\sim 3 - 6$ d), vergeleken met planeten rondom dubbelsterren met langere baanperiodes. Er zijn echter tot nu toe geen enkele planeten gevonden rondom deze dubbelsterren met korte baanperiodes, terwijl er 10 zijn ontdekt rondom dubbelsterren met langere baanperiodes. Het is daarom waarschijnlijk dat er een intrinsieke afwezigheid is van (waarneembare) planeten rondom dubbelsterren met korte baanperiodes.

We hebben laten zien dat deze afwezigheid verklaard kan worden door de seculaire invloed van een planeet rondom een dubbelster in hiërarchische *triplestersystemen*. Men denkt dat nauwe dubbelsterren in dergelijke triplestersystemen zijn gevormd via LK-oscillaties in de binnenste baan, veroorzaakt door de derde ster, gecombineerd met getijdenwerking. De baan van de binnenste dubbelster was in dit geval aanvankelijk wijder, met een baanperiode tot $\sim 10^4$ yr. We hebben laten zien dat met een planeet rondom de binnenste dubbelster, LK-oscillaties in de binnenste baan, veroorzaakt door de derde ster, uitgedoofd kunnen worden door additionele precessie geïnduceerd door de planeet. Dit betekent dat de binnenste dubbelster wordt ‘beschermd’ tegen het krachtmoment van de baan van de derde ster, en tengevolge krimpt de binnenste baan niet. Dit beschermende effect is echter alleen het geval als de planeet voldoende massief is, en als de baan voldoende dichtbij en geïnclineerd is ten opzichte van de binnenste dubbelster. In andere gevallen is ofwel de baan van de planeet stabiel maar kan de planeet de binnenste dubbelster niet beschermen, ofwel de baan van de planeet is instabiel, wat waarschijnlijk leidt tot een ontsnapping van de planeet van het systeem. In het bijzonder, als een lage-massaplaneet initieel geïnclineerd is van en sterk verwijderd is de binnenste dubbelster, dan is er typisch geen beschermingseffect, en kan de baan van de binnenste dubbelster krimpen tot een nauwe baan. Aan de andere kant, een meer massieve planeet in een initieel coplaire en dichte baan ten opzichte van de binnenste dubbelster zorgt typisch voor een beschermend effect, waardoor de binnenste baan niet krimpt.

Dit impliceert dat nauwe dubbelsterren typisch geen massieve ($m_3 \sim M_J$) planeten hebben op nauwe en coplaire banen ten opzichte van de dubbelster, welke krimp van de baan hadden voorkomen. Als in tegenstelling de planeet minder massief is ($m_3 \sim 10^{-2} M_J$) en initieel geïnclineerd ten opzichte van en ver verwijderd was van de dubbelster, dan is vernauwing van de baan van de dubbelster mogelijk. Deze trend is consistent met de huidige waarnemingen van *Kepler*, en suggereert dat er zich planeten kunnen bevinden rondom nauwe dubbelsterren, welke op het moment niet waargenomen kunnen worden.

Hoofdstuk 6 – Seculaire evolutie van hiërarchische multi- plesystemen

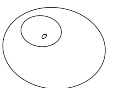
We hebben de methode en het algoritme van Hoofdstuk 4 gegeneraliseerd naar systemen bestaande uit geneste dubbelsterbanen (dat wil zeggen simplexbanen) met een willekeurig aantal lichamen, en een willekeurige configuratie. We hebben de Hamiltoniaan geëxpandeerd in termen van de verhoudingen van baanseparaties, welke per aanname klein zijn, en een middeling over de banen uitgevoerd. Het formalisme hebben we geïmplementeerd in een algoritme, `SECULARMULTIPLE`, waarmee langetermijnintegraties van hiërarchische systemen met complexe hiërarchieën mogelijk zijn, en we hebben eerste toepassingen gepresenteerd. In het bijzonder hebben wij onze methode toegepast op de hoge-multipliciteitssystemen Mizar en Alcor, en 30 Ari.

Hoofdstuk 7 – *Hot Jupiters* in multiplaneetsystemen

We hebben `SECULARMULTIPLE` toegepast op de vorming van *hot Jupiters* (HJs; Jupiterachtige planeten die zich bevinden in zeer nauwe banen ten opzichte van hun ster) via migratie met hoge-excentriciteit in multiplaneetsystemen (systemen met tenminste drie planeten). In dit scenario wordt aangenomen dat HJs gevormd zijn als Jupiterachtige planeten op ~ 1 AU afstand van de ster. Vervolgens werd de excentriciteit van de baan van de planeet geëxiteerd vanwege seculaire interacties met de andere planeten, waardoor de mate van getijdenwerking sterk toenam, en uiteindelijk een HJ gevormd is in een circulaire baan met een halve-lange $a \sim 0.05$ AU.

We hebben eerst geverifieerd dat ons seculaire algoritme statistisch gezien correcte resultaten geeft vergeleken met directe N -bodyintegraties. De laatste zijn computationeel gezien echter veel duurder om uit te voeren, en daarom stelt de seculaire methode ons in staat om efficiënt de grote parameterruimte van multiplaneetsystemen te verkennen. We hebben de condities bestudeerd waarvoor hoge excentriciteiten opgewekt kunnen worden, en daarmee potentieel HJs gevormd kunnen worden. We hebben gevonden dat niet alleen de parameter de *angular momentum deficit* groot genoeg moet zijn, maar ook moet de verhouding van LK-tijdschalen $\mathcal{R}_0^{(ijk)}$, toegepast op naburige baanparen (i, j) and (j, k) , in de buurt van één moet zijn. De sterkste afhankelijkheid geldt voor $\mathcal{R}_0^{(123)}$, dat wil zeggen de verhouding toegepast op de binnenste banen.

Vervolgens hebben we een populatiesynthese uitgevoerd waarin we gebruik hebben gemaakt van een combinatie van grid en Monte Carlo sampling. We hebben gevonden dat het aantal gevormde HJs typisch klein is, en sterk afhankelijk is van de aangenomen sterkte van de getijdenwerkingen, dat wil zeggen de *viscous time-scale* (of, de equivalente *time lag*). Met de *viscous time-scale* van Socrates et al. (2012) hebben we geen enkele HJs gevonden, terwijl ten hoogste ~ 0.01 van de systemen tot HJs leiden als de aangenomen getijdenkracht 100 keer sterker is. Dit kan verklaard worden door de ‘gewelddadige’ eigenschap van de excitatie van de excentriciteit door seculaire evolutie, wat typisch leidt tot een snelle vernietiging van de planeet, in tegenstelling tot een langzamer proces van getijdenwerking waarbij de baan nauwer en meer cirkelvormig wordt. Onze resultaten impliceren dat migratie met hoge-excentriciteit in multiplaneetsystemen ten hoogste 0.01 van de waargenomen HJs kan verklaren, aangenomen dat de getijdenwerking in de planeet 100 maal sterker is ten opzichte van Socrates et al. (2012).



Summary

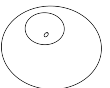
We have studied the gravitational dynamical evolution of hierarchical systems with various hierarchies, in various astrophysical contexts. We began by considering systems of the ‘multiplex’ type, dominated by an SBH (Chapters 2 and 3). Subsequently, we focussed on simplex-type systems, beginning with the general dynamics of quadruple systems (Chapter 4), and continuing with an application to circumbinary planets (Chapter 5). Finally, we generalised our methods to hierarchical simplex-type systems with an arbitrary hierarchical structure and an arbitrary number of bodies (Chapter 6), and applied these methods to the production of HJs through high-eccentricity migration in multiplanet systems (Chapter 7).

Chapter 2 – Relativistic dynamics around an SBH

In this chapter, we studied the long-term gravitational dynamics of orbits around a supermassive black hole (SBH). We developed a special-purpose N -body code, `TPI`, that allowed us to efficiently integrate these orbits including general relativistic (GR) corrections, and taking into account interactions with a large ($\sim 10^3$) number of ‘field stars’, which are assumed to move in (nearly) fixed Keplerian orbits. We have shown that `TPI` agrees with other, more accurate but slower codes.

Using `TPI`, we performed simulations of the Galactic Centre (GC) to test models for the origin of the S-stars with 4800 field stars close to the SBH. Such simulations were previously not feasible with other direct N -body codes. We assumed that the S-stars were deposited initially on to orbits of very high eccentricity, which is expected for the tidal disruption of a stellar binary. We have found that the cumulative eccentricity distribution of the S-stars evolves to $N(e) \sim e^{2.6}$ on a time-scale of 7 ± 0.1 Myr, which is consistent with observations. Our results suggest a lower limit on the typical age of the S-stars of ~ 7 Myr and ~ 25 Myr assuming burst and continuous formation of the S-stars, respectively.

In addition, from our simulations we extracted first- and second-order diffusion coefficients in the normalized angular-momentum variable $\ell \equiv \sqrt{1 - e^2}$. These coefficients describe the long-term angular momentum evolution. We identified three angular-momentum regimes, in which the diffusion coefficients depend in functionally different ways on ℓ . In the regime of low ℓ , the diffusion coefficients are well described in terms of non-resonant relaxation, i.e. relaxation due to two-body encounters. In the regime of high ℓ , the diffusion coefficients are well described in terms of resonant relaxation (RR), i.e. torques arising from correlated encounters. In the intermediate regime, near the ‘Schwarzschild barrier’ (SB) where relativistic precession



is important, we identified a new regime of ‘anomalous relaxation’, in which diffusion is driven by a combination of RR and rapid relativistic precession. In particular, we showed that the SB can be associated with a rapid drop of the diffusion coefficients with decreasing ℓ . We presented analytic expressions, in terms of physical parameters, that describe the diffusion coefficients in all three angular-momentum regimes. Also, we applied our results to obtain the steady-state distribution (i.e. extrapolated to infinity) of the orbital angular momentum.

Chapter 3 – Planetesimals in the GC

Continuing with the previous topic of orbits around an SBH, we studied the dynamics of planetesimals in the GC, and focussed on the tidal disruption of planetesimals by the SBH, possibly producing near-infrared and X-ray flares, which are observed from the SBH in the GC, Sgr A*, on an approximately daily basis. We assumed that the planetesimals were either formed in a large-scale cloud bound to the SBH (scenario 1), or in debris discs around stars (scenario 2). In the latter case, we found that the tidal force of the SBH is effective at stripping the planetesimals from their parent stars at distances $r \lesssim 0.5$ pc from the SBH. Stripping by gravitational encounters with other stars is effective at stripping nearly all planetesimals within the radius of influence, ≈ 4 pc, after ~ 100 Myr.

We used the Fokker-Planck equation in energy space taking into account these stripping effects and other possibly important effects such as relaxation due to massive perturbers, and determined the disruption rate of planetesimals by the SBH. We found that the disruption rate at $t = 10$ Gyr is $\approx 0.6 \text{ d}^{-1}$, which is roughly consistent with the observed rate of the flares of once per day. Our rates are insensitive to the model assumptions, in particular the initial distribution of planetesimals, i.e. scenarios (1) or (2). In the case of scenario (2), the assumed number of planetesimals per star $N_{a/\star} = 2 \times 10^7$, is inferred from observations of debris discs around stars in the solar neighbourhood. In contrast, in the case of scenario (1), i.e. formation in a large-scale cloud, this implies that the number of planetesimals formed is strongly correlated with the number of stars, which requires finetuning of $N_{a/\star}$. We favoured the more natural explanation that planetesimals in galactic nuclei similar to the GC are formed no differently than planetesimals around stars in the solar neighbourhood.

Chapter 4 – Secular evolution of hierarchical quadruple systems

In the remaining chapters, we changed tack to hierarchical systems of the simplex type, i.e. systems consisting of nested binary orbits. We first considered hierarchical quadruple systems in the ‘3+1’ configuration, i.e. a hierarchical triple orbited by a fourth body. We assumed that the system is sufficiently hierarchical, i.e. that the ratios x_i of the binary separations are small, and we expanded the Hamiltonian in terms of the x_i . Subsequently, we orbit-averaged the expanded Hamiltonian, and implemented the equations of motion into a computer code, `SECULARQUADRUPLE`.

Subsequently, we studied the secular evolution of highly hierarchical systems that are well described by the lowest order terms in the Hamiltonian, and we characterised the evolution in terms of ratios of Lidov-Kozai (LK) time-scales applied to different binary pairs. Let the

innermost, intermediate and outermost binaries by denoted with A, B and C, respectively. For these systems, we found that the global dynamics can be qualitatively described in terms of the (initial) ratio of the LK time-scales applied to the AB binary and BC binary pairs, $\mathcal{R}_0 \equiv P_{\text{LK,AB},0}/P_{\text{LK,BC},0}$.

If $\mathcal{R}_0 \ll 1$, then the torque of binary B on A dominates compared to the torque of binary C on binary B, and therefore binaries A and B remain coplanar if this was initially the case. If binaries A and B are initially inclined, LK eccentricity oscillations in binary A are not much affected by the presence of the fourth body. Eccentricity oscillations in binary B due to the secular torque of binary C are efficiently quenched due to short time- scale precession induced on binary B by binary A.

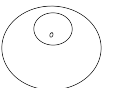
If $\mathcal{R}_0 \gg 1$, then the torque of binary C on binary B dominates compared to the torque of binary B on binary A. Initially, the inclination of binary B changes, whereas this is not the case for binary A. This induces a mutual inclination between binaries A and B, even if they are initially not inclined. However, rapid precession of binary B compared to the LK time-scale for the AB pair prevents any significant eccentricity oscillations in binary A, and even quenches LK oscillations if binaries A and B are initially inclined.

Lastly, if $\mathcal{R}_0 \sim 1$, complex LK eccentricity oscillations can occur in binary A that are strongly coupled with the LK eccentricity oscillations in binary B. The latter are also affected compared to the situation in which binary A were replaced by a point mass, although this is typically a much smaller effect. Even if binaries A and B are initially coplanar, the induced inclination can result in high-amplitude eccentricity oscillations in binary A. These extreme eccentricities could have significant implications for strong interactions such as tidal interactions, gravitational wave dissipation, and collisions and mergers of stars and compact objects.

Chapter 5 – Explaining the lack of circumbinary planets around short-period binaries

In this chapter, we applied `SECULARQUADRUPEL` to explain a recent curious result in observations of transiting circumbinary planets by the *Kepler* spacecraft. Assuming that the occurrence rate of circumbinary planets is not related to the binary period, then about twice more planets are expected to have been observed around short-period binaries ($\sim 3 - 6$ d) compared to their longer period counterparts. However, so far, no planets have been observed around these short-period binaries, whereas 10 have been found around longer-period binaries. Therefore, it there is likely an intrinsic lack of (observable) circumbinary planets around short-period binaries.

We have shown that this lack can be explained by the secular influence of a circumbinary planet in hierarchical *triple star* systems. Short-period binaries are believed to have formed in stellar triple systems through LK eccentricity oscillations in the inner orbit induced by the tertiary, combined with tidal friction. In this case, the progenitor inner binary was wider, with an orbital period of up to $\sim 10^4$ d. We have shown that if there is a circumbinary planet around the progenitor inner binary, then the LK eccentricity oscillations in the inner orbit induced by the tertiary can be quenched by additional precession induced by the circumbinary planet. Thereby, the inner binary is ‘shielded’ from the torque of the tertiary star, and does not shrink to a tight orbit. However, this only occurs if the circumbinary planet is sufficiently massive, and if its orbit is sufficiently close to and coplanar with the inner binary. In other cases, the circumbinary orbit is stable but cannot efficiently shield the inner binary, or the circumbinary orbit is destabilized



and is most likely ejected. In particular, if a low-mass circumbinary planet is initially inclined with respect to and far away from the inner binary, then planet shielding is typically ineffective, and the inner binary can shrink to a tight orbit. On the other hand, for a more massive planet in an initially approximately coplanar and tight circumbinary orbit, shielding typically prevents the inner orbit from shrinking.

Consequently, short-period binaries typically do not have massive ($m_3 \sim M_J$) circumbinary planets on tight and coplanar orbits, which would have prevented orbital shrinkage. In contrast, if the circumbinary planet not massive ($m_3 \sim 10^{-2} M_J$) and initially inclined with respect to and far away from the inner binary, then shrinkage is possible. This trend is consistent with the current *Kepler* observations, and suggests that the latter type of currently unobservable planets may be present around short-period binaries.

Chapter 6 – Secular evolution of hierarchical multiple systems

We generalised the method and the algorithm developed in Chapter 4 to systems consisting of nested binary (i.e. simplex) orbits, with an *arbitrary* number of bodies, and an *arbitrary* hierarchical configuration. We expanded the Hamiltonian in terms of ratios of binary separations, which are assumed to be small, and carried out orbital averaging. We implemented the formalism in an algorithm, `SECULARMULTIPLE`, making long-term integrations of hierarchical systems with complex hierarchies feasible, and presented first applications. In particular, we applied our method to the high-multiplicity systems Mizar and Alcor, and 30 Ari.

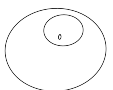
Chapter 7 – Hot Jupiters in multiplanet systems

We applied `SECULARMULTIPLE` to the formation of hot Jupiters (HJs) through high-eccentricity migration in multiplanet systems (systems with at least three planets). In this scenario, hot Jupiters are assumed to have been formed as Jupiter-like planets at ~ 1 AU. Subsequently, the orbital eccentricity was excited due to secular interactions with the other planets, strongly increasing the rate of tidal dissipation, and eventually producing a HJ in a circular orbit with a semimajor axis of ~ 0.05 AU.

First, we verified that our secular algorithm produces statistically correct results compared to direct N -body simulations. The latter are computationally much more expensive, however, therefore the secular method allows for an efficient exploration of the large parameter space of multiplanet systems. We investigated the conditions for exciting high eccentricities in multiplanet systems, and therefore potentially HJs. In addition to a sufficiently large angular momentum deficit (AMD), we found that a necessary condition for eccentricity excitation is that the ratio of LK time-scales $\mathcal{R}_0^{(ijk)}$ applied to adjacent orbit pairs (i, j) and (j, k) should be near unity. The strongest dependence applies to $\mathcal{R}_0^{(123)}$, i.e. the ratio applied to the innermost orbits.

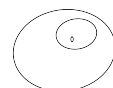
Subsequently, we carried out a population synthesis study in which we combined grid and Monte Carlo sampling. We found that the number of HJs formed is typically small, and is strongly dependent on the assumed tidal dissipation strength, i.e. the viscous time-scale (or, equivalently, the time lag). Assuming the viscous time-scale derived previously by Socrates et al. (2012), we found no HJs at all, whereas at most ~ 0.01 of the systems lead to HJs if

the assumed tidal dissipation strength is assumed to be 100 times more efficient. This can be explained by the ‘violent’ nature of secular eccentricity driving of the innermost planet, which tends to lead to the rapid tidal disruption of the planet, rather than the slower process of orbital shrinkage and circularisation due to tidal dissipation. Our results imply that high- e migration in multiplanet systems can explain at most 0.01 of the observed HJs, assuming 100 stronger tidal dissipation compared to Socrates et al. (2012).



Bibliography

- Abad A. J., Docobo J. A., 1988, *Celestial Mechanics*, 41, 333
- Abbott B. P., et al., 2016, *Physical Review Letters*, 116, 061102
- Aharon D., Perets H. B., 2015, *ApJ*, 799, 185
- Alexander T., 2005, *Phys. Rep.*, 419, 65
- Alexander R., 2012, *ApJ*, 757, L29
- Alexander T., Pfuhl O., 2014, *ApJ*, 780, 148
- Allan R. R., Cook G. E., 1964, *Royal Society of London Proceedings Series A*, 280, 97
- Allan R. R., Ward G. N., 1963, *Proceedings of the Cambridge Philosophical Society*, 59, 669
- Allen R. H., 1899, *Star-names and their meanings*
- Amaro-Seoane P., 2012, preprint, (arXiv:1205.5240)
- Amaro-Seoane P., Sopuerta C. F., Freitag M. D., 2013, *MNRAS*, 429, 3155
- Ambartsumian V. A., 1954, *Contr. Obs. Byurakan*, 15, 3
- Anderson K. R., Storch N. I., Lai D., 2016, *MNRAS*, 456, 3671
- Antognini J. M. O., 2015, *MNRAS*, 452, 3610
- Antognini J. M., Shappee B. J., Thompson T. A., Amaro-Seoane P., 2014, *MNRAS*, 439, 1079
- Antoniadis J., et al., 2013, *Science*, 340, 448
- Antonini F., 2014, *ApJ*, 794, 106
- Antonini F., Merritt D., 2013, *ApJ*, 763, L10
- Antonini F., Perets H. B., 2012, *ApJ*, 757, 27
- Antonini F., Lombardi Jr. J. C., Merritt D., 2011, *ApJ*, 731, 128
- Antonini F., Murray N., Mikkola S., 2014, *ApJ*, 781, 45
- Antonini F., Hamers A. S., Lithwick Y., 2016, preprint, (arXiv:1604.01781)
- Armstrong D. J., Osborn H. P., Brown D. J. A., Faedi F., Gómez Maqueo Chew Y., Martin D. V., Pollacco D., Udry S., 2014, *MNRAS*, 444, 1873
- Baganoff F. K., et al., 2001, *Nature*, 413, 45
- Baganoff F. K., et al., 2003, *ApJ*, 591, 891
- Bahcall J. N., Wolf R. A., 1976, *ApJ*, 209, 214
- Bahcall J. N., Wolf R. A., 1977, *ApJ*, 216, 883
- Bar-Or B., Kupi G., Alexander T., 2013, *ApJ*, 764, 52
- Barker A. J., Ogilvie G. I., 2009, *MNRAS*, 395, 2268
- Barrière N. M., et al., 2014, *ApJ*, 786, 46
- Bartko H., et al., 2009, *ApJ*, 697, 1741
- Bartko H., et al., 2010, *ApJ*, 708, 834
- Batygin K., Bodenheimer P. H., Laughlin G. P., 2015, preprint, (arXiv:1511.09157)



- Belorizsky D., 1930, *Bulletin Astronomique* 6, pp 417–434
- Berukoff S. J., Hansen B. M. S., 2006, *ApJ*, 650, 901
- Beuermann K., et al., 2011, *A&A*, 526, A53
- Binney J., Tremaine S., 2008, *Galactic Dynamics: Second Edition*. Princeton University Press
- Blaes O., Lee M. H., Socrates A., 2002, *ApJ*, 578, 775
- Blum R. D., Ramírez S. V., Sellgren K., Olsen K., 2003, *ApJ*, 597, 323
- Bodenheimer P., Hubickyj O., Lissauer J. J., 2000, *Icarus*, 143, 2
- Boekholt T., Portegies Zwart S., 2015, *Computational Astrophysics and Cosmology*, 2, 2
- Bonfils X., et al., 2013, *A&A*, 549, A109
- Boué G., Fabrycky D. C., 2014, *ApJ*, 789, 110
- Bowler B. P., Liu M. C., Kraus A. L., Mann A. W., 2014, *ApJ*, 784, 65
- Breiter S., Ratajczak R., 2005, *MNRAS*, 364, 1222
- Brem P., Amaro-Seoane P., Sopena C. F., 2014, *MNRAS*, 437, 1259
- Bromley B. C., Kenyon S. J., 2015, *ApJ*, 806, 98
- Brouwer D., 1959, *AJ*, 64, 378
- Bryan M. L., et al., 2016, *ApJ*, 821, 89
- Buchholz R. M., Schödel R., Eckart A., 2009, *A&A*, 499, 483
- Burgay M., et al., 2003, *Nature*, 426, 531
- Chatterjee S., Ford E. B., Matsumura S., Rasio F. A., 2008, *ApJ*, 686, 580
- Chen X., Amaro-Seoane P., 2014, *ApJ*, 786, L14
- Chirikov B. V., 1979, *Phys. Rep.*, 52, 263
- Cohen S. D., Hindmarsh A. C., Dubois P. F., 1996, *Computers in Physics*, 10, 138
- Cohn H., 1979, *ApJ*, 234, 1036
- Cohn H., Kulsrud R. M., 1978, *ApJ*, 226, 1087
- Conroy K. E., Prša A., Stassun K. G., Orosz J. A., Fabrycky D. C., Welsh W. F., 2014, *AJ*, 147, 45
- Correia A. C. M., Laskar J., Farago F., Boué G., 2011, *Celestial Mechanics and Dynamical Astronomy*, 111, 105
- Damour T., Deruelle N., 1981, *Physics Letters A*, 87, 81
- Do T., Ghez A. M., Morris M. R., Lu J. R., Matthews K., Yelda S., Larkin J., 2009, *ApJ*, 703, 1323
- Do T., et al., 2013, *ApJ*, 779, L6
- Dobbs-Dixon I., Lin D. N. C., Mardling R. A., 2004, *ApJ*, 610, 464
- Dodds-Eden K., et al., 2011, *ApJ*, 728, 37
- Doyle L. R., et al., 2011, *Science*, 333, 1602
- Duquennoy A., Mayor M., 1991, *A&A*, 248, 485
- Dvorak R., 1982, *Oesterreichische Akademie Wissenschaften Mathematisch naturwissenschaftliche Klasse Sitzungsberichte Abteilung*, 191, 423
- Eckart A., Genzel R., Hofmann R., Sams B. J., Tacconi-Garman L. E., 1993, *ApJ*, 407, L77
- Eddington A. S., 1916, *MNRAS*, 76, 572
- Eggleton P. P., Kiseleva-Eggleton L., 2001, *ApJ*, 562, 1012
- Eggleton P. P., Kiseleva-Eggleton L., 2006, *Ap&SS*, 304, 75
- Eggleton P. P., Kiseleva L. G., Hut P., 1998, *ApJ*, 499, 853
- Eilon E., Kupi G., Alexander T., 2009, *ApJ*, 698, 641
- Einstein A., Infeld L., Hoffmann B., 1938, *Annals of Mathematics*, 39, 65
- Eisenhauer F., et al., 2005, *ApJ*, 628, 246
- Evans D. S., 1968, *QJRAS*, 9, 388
- Fabrycky D., Tremaine S., 2007, *ApJ*, 669, 1298

- Feroz F., Hobson M. P., 2014, MNRAS, 437, 3540
- Ferrarese L., Merritt D., 2000, ApJ, 539, L9
- Ford E. B., Rasio F. A., 2008, ApJ, 686, 621
- Ford E. B., Kozinsky B., Rasio F. A., 2000, ApJ, 535, 385
- Frank J., Rees M. J., 1976, MNRAS, 176, 633
- Frost E. B., 1908, *Astronomische Nachrichten*, 177, 171
- Fujii M., Iwasawa M., Funato Y., Makino J., 2010, ApJ, 716, L80
- Funato Y., Hut P., McMillan S., Makino J., 1996, AJ, 112, 1697
- Gaburov E., Harfst S., Portegies Zwart S. F., 2009, New A, 14, 630
- Genzel R., Thatte N., Krabbe A., Kroker H., Tacconi-Garman L. E., 1996, ApJ, 472, 153
- Genzel R., Schödel R., Ott T., Eckart A., Alexander T., Lacombe F., Rouan D., Aschenbach B., 2003a, *Nature*, 425, 934
- Genzel R., et al., 2003b, ApJ, 594, 812
- Genzel R., Eisenhauer F., Gillessen S., 2010, *Reviews of Modern Physics*, 82, 3121
- Ghez A. M., et al., 2008, ApJ, 689, 1044
- Gillessen S., Eisenhauer F., Trippe S., Alexander T., Genzel R., Martins F., Ott T., 2009, ApJ, 692, 1075
- Goldreich P., Tremaine S., 1980, ApJ, 241, 425
- Goldstein H., 1975, *American Journal of Physics*, 43, 737
- Goldstein H., 1976, *American Journal of Physics*, 44, 1123
- Goldstein H., Poole C., Safko J., 2002, *Classical mechanics*
- Gomes R., Levison H. F., Tsiganis K., Morbidelli A., 2005, *Nature*, 435, 466
- Gonçalves Ferrari G., Boekholt T., Portegies Zwart S. F., 2014, MNRAS, 440, 719
- Goodman J., 1983, ApJ, 270, 700
- Guenther E. W., Hartmann M., Esposito M., Hatzes A. P., Cusano F., Gandolfi D., 2009, A&A, 507, 1659
- Guillochon J., Ramirez-Ruiz E., Lin D., 2011, ApJ, 732, 74
- Gürkan M. A., Hopman C., 2007, MNRAS, 379, 1083
- Guyer J. E., Wheeler D., Warren J. A., 2009, *Computing in Science & Engineering*, 11, 6
- Hamers A. S., Portegies Zwart S. F., 2016, MNRAS,
- Hamers A. S., Pols O. R., Claeys J. S. W., Nelemans G., 2013, MNRAS, 430, 2262
- Hamers A. S., Portegies Zwart S. F., Merritt D., 2014, MNRAS, 443, 355
- Hamers A. S., Perets H. B., Antonini F., Portegies Zwart S. F., 2015, MNRAS, 449, 4221
- Hamilton D. P., Burns J. A., 1992, *Icar*, 96, 43
- Hansen B. M. S., 2010, ApJ, 723, 285
- Hansen B. M. S., Milosavljević M., 2003, ApJ, 593, L77
- Harrington R. S., 1968, AJ, 73, 190
- Harrington R. S., 1969, *Celestial Mechanics*, 1, 200
- Hayashi C., 1981, *Progress of Theoretical Physics Supplement*, 70, 35
- Heggie D. C., 1975, MNRAS, 173, 729
- Heggie D. C., Hut P., McMillan S. L. W., 1996, ApJ, 467, 359
- Hills J. G., 1988, *Nature*, 331, 687
- Holman M. J., Wiegert P. A., 1999, AJ, 117, 621
- Holman M., Touma J., Tremaine S., 1997, *Nature*, 386, 254
- Hopman C., Alexander T., 2006a, ApJ, 645, 1152
- Hopman C., Alexander T., 2006b, ApJ, 645, L133
- Hulse R. A., Taylor J. H., 1975, ApJ, 195, L51
- Hurley J. R., Pols O. R., Tout C. A., 2000, MNRAS, 315, 543

- Hut P., 1981, *A&A*, 99, 126
- Hut P., 1983, *ApJ*, 268, 342
- Hut P., 1993, *ApJ*, 403, 256
- Hut P., Bahcall J. N., 1983, *ApJ*, 268, 319
- Hut P., Makino J., McMillan S., 1995, *ApJ*, 443, L93
- Innanen K. A., Zheng J. Q., Mikkola S., Valtonen M. J., 1997, *AJ*, 113, 1915
- Ito T., Tanikawa K., 2002, *MNRAS*, 336, 483
- Iwasawa M., An S., Matsubayashi T., Funato Y., Makino J., 2011, *ApJ*, 731, L9
- Jurić M., Tremaine S., 2008, *ApJ*, 686, 603
- Katz B., Dong S., 2012, preprint, (arXiv:1211.4584)
- Katz B., Dong S., Malhotra R., 2011, *Physical Review Letters*, 107, 181101
- Kepler J., 1609, *Astronomia nova*.
- Kidder L. E., 1995, *Phys. Rev. D*, 52, 821
- Kinoshita H., Nakai H., 1999, *Celestial Mechanics and Dynamical Astronomy*, 75, 125
- Kinoshita H., Nakai H., 2007, *Celestial Mechanics and Dynamical Astronomy*, 98, 67
- Kiseleva L. G., Eggleton P. P., Mikkola S., 1998, *MNRAS*, 300, 292
- Knutson H. A., et al., 2014, *ApJ*, 785, 126
- Kolmogorov A., 1933, *G. Ist. Ital. Attuari*, 4, 83
- Kostov V. B., McCullough P. R., Hinse T. C., Tsvetanov Z. I., Hébrard G., Díaz R. F., Deleuil M., Valenti J. A., 2013, *ApJ*, 770, 52
- Kostov V. B., et al., 2014, *ApJ*, 784, 14
- Kozai Y., 1962, *AJ*, 67, 591
- Kraus A. L., Ireland M. J., Cieza L. A., Hinkley S., Dupuy T. J., Bowler B. P., Liu M. C., 2014, *ApJ*, 781, 20
- Kustaanheimo P., Stiefel E. L., 1965, *J. Reine Angew. Math.*, 218, 204
- Kuzuhara M., Tamura M., Ishii M., Kudo T., Nishiyama S., Kandori R., 2011, *AJ*, 141, 119
- Lagrange J. L., 1781, *Première partie, Nouveaux Mémoires de l'Académie des Sciences et Belles-Lettres de Berlin*, p. 125
- Laplace P. S., 1784, *Mem. Acad. royale des Sciences de Paris, Oeuvres complètes*, XI 49
- Laskar J., 1994, *A&A*, 287, L9
- Laskar J., 2008, *Icarus*, 196, 1
- Laskar J., 2012, preprint, (arXiv:1209.5996)
- Laskar J., Boué G., 2010, *A&A*, 522, A60
- Laskar J., Gastineau M., 2009, *Nature*, 459, 817
- Lee M. H., Peale S. J., 2006, *Icarus*, 184, 573
- Lestrade J.-F., Morey E., Lasso A., Phou N., 2011, *A&A*, 532, A120
- Leung G. C. K., Lee M. H., 2013, *ApJ*, 763, 107
- Li G., Naoz S., Kocsis B., Loeb A., 2014a, *ApJ*, 785, 116
- Li G., Naoz S., Holman M., Loeb A., 2014b, *ApJ*, 791, 86
- Lidov M. L., 1962, *Planet. Space Sci.*, 9, 719
- Lin D. N. C., Papaloizou J., 1986, *ApJ*, 309, 846
- Lin D. N. C., Bodenheimer P., Richardson D. C., 1996, *Nature*, 380, 606
- Lithwick Y., Naoz S., 2011, *ApJ*, 742, 94
- Lithwick Y., Wu Y., 2011, *ApJ*, 739, 31
- Lithwick Y., Wu Y., 2014, *Proceedings of the National Academy of Science*, 111, 12610
- Lorentz H. A., Droste J., 1917, *Versl. K. Akad. Wetensch. Amsterdam*, 26, 392

- Lu J. R., Ghez A. M., Hornstein S. D., Morris M. R., Becklin E. E., Matthews K., 2009, *ApJ*, 690, 1463
- Ludendorff H., 1908, *Astronomische Nachrichten*, 177, 7
- Luo L., Katz B., Dong S., 2016, *MNRAS*, 458, 3060
- Madigan A.-M., Hopman C., Levin Y., 2011, *ApJ*, 738, 99
- Makino J., 1991, *ApJ*, 369, 200
- Mamajek E. E., Kenworthy M. A., Hinz P. M., Meyer M. R., 2010, *AJ*, 139, 919
- Maness H., et al., 2007, *ApJ*, 669, 1024
- Mardling R. A., 2013, *MNRAS*, 435, 2187
- Mardling R. A., Aarseth S. J., 2001, *MNRAS*, 321, 398
- Martin D. V., Triaud A. H. M. J., 2014, *A&A*, 570, A91
- Martin D. V., Triaud A. H. M. J., 2015, *MNRAS*, 449, 781
- Martin D. V., Mazeh T., Fabrycky D. C., 2015, *MNRAS*, 453, 3554
- Mazeh T., Shaham J., 1979, *A&A*, 77, 145
- Meibom S., Mathieu R. D., 2005, *ApJ*, 620, 970
- Merritt D., 2004, *Physical Review Letters*, 92, 201304
- Merritt D., 2010, *ApJ*, 718, 739
- Merritt D., 2013a, *Dynamics and Evolution of Galactic Nuclei*
- Merritt D., 2013b, *Dynamics and Evolution of Galactic Nuclei*
- Merritt D., Vasiliev E., 2012, *Physical Review D*, 86, 102002
- Merritt D., Harfst S., Bertone G., 2007, *Phys. Rev. D*, 75, 043517
- Merritt D., Alexander T., Mikkola S., Will C. M., 2011, *Phys. Rev. D*, 84, 044024
- Michael E., Perets H. B., Grishin E., 2015, preprint, (arXiv:1506.08818)
- Mikkola S., Aarseth S. J., 1993, *Celestial Mechanics and Dynamical Astronomy*, 57, 439
- Mikkola S., Aarseth S., 2002, *Celestial Mechanics and Dynamical Astronomy*, 84, 343
- Mikkola S., Merritt D., 2008, *AJ*, 135, 2398
- Mikkola S., Tanikawa K., 1999, *MNRAS*, 310, 745
- Milankovitch M., 1939, *Bull. Serb. Acad. Math. Nat.*, 6
- Milgrom M., 1983, *ApJ*, 270, 365
- Miller M. C., Hamilton D. P., 2002, *ApJ*, 576, 894
- Morbey C. L., Brosterhus E. B., 1974, *PASP*, 86, 455
- Morris M., 1993, *ApJ*, 408, 496
- Muñoz D. J., Lai D., 2015, *Proceedings of the National Academy of Science*, 112, 9264
- Murray C. D., Dermott S. F., 1999, *Solar system dynamics*
- Murray-Clay A., Loeb A., 2012, *Nature*, 3
- Musen P., 1961, *J. Geophys. Res.*, 66, 2797
- Naoz S., Fabrycky D. C., 2014, *ApJ*, 793, 137
- Naoz S., Farr W. M., Lithwick Y., Rasio F. A., Teyssandier J., 2011, *Nature*, 473, 187
- Naoz S., Farr W. M., Rasio F. A., 2012, *ApJ*, 754, L36
- Naoz S., Farr W. M., Lithwick Y., Rasio F. A., Teyssandier J., 2013a, *MNRAS*, 431, 2155
- Naoz S., Kocsis B., Loeb A., Yunes N., 2013b, *ApJ*, 773, 187
- Nayakshin S., Sazonov S., Sunyaev R., 2012, *MNRAS*, 419, 1238
- Neilsen J., et al., 2013, *ApJ*, 774, 42
- Newton I., 1687, *Philosophiae Naturalis Principia Mathematica*. Auctore Js. Newton, doi:10.3931/e-rara-440.
- Ngo H., et al., 2015, *ApJ*, 800, 138
- Nikolažuk M., Walter R., 2013, *A&A*, 552, A75

- Nitadori K., Makino J., 2008, *New A*, 13, 498
- Nugeyre J. B., Bouvier P., 1981, *Celestial Mechanics*, 25, 51
- Ogilvie G. I., Lin D. N. C., 2007, *ApJ*, 661, 1180
- Oh S., Kim S. S., Figer D. F., 2009, *Journal of Korean Astronomical Society*, 42, 17
- Ondra L., 2004, *S&T*, 108, 74
- Orosz J. A., et al., 2012a, *Science*, 337, 1511
- Orosz J. A., et al., 2012b, *ApJ*, 758, 87
- Paumard T., et al., 2006, *Journal of Physics Conference Series*, 54, 199
- Pejcha O., Antognini J. M., Shappee B. J., Thompson T. A., 2013, *MNRAS*, 435, 943
- Pelupessy F. I., Jänes J., Portegies Zwart S., 2012, *New A*, 17, 711
- Pelupessy F. I., van Elteren A., de Vries N., McMillan S. L. W., Drost N., Portegies Zwart S. F., 2013, *A&A*, 557, A84
- Perets H. B., 2009, *ApJ*, 690, 795
- Perets H. B., Fabrycky D. C., 2009, *ApJ*, 697, 1048
- Perets H. B., Gualandris A., 2010, *ApJ*, 719, 220
- Perets H. B., Naoz S., 2009, *ApJ*, 699, L17
- Perets H. B., Hopman C., Alexander T., 2007, *ApJ*, 656, 709
- Perets H. B., Gualandris A., Kupi G., Merritt D., Alexander T., 2009, *ApJ*, 702, 884
- Peters P. C., 1964, *Physical Review*, 136, 1224
- Petrovich C., 2015a, *ApJ*, 799, 27
- Petrovich C., 2015b, *ApJ*, 805, 75
- Petrovich C., Tremaine S., 2016, preprint, (arXiv:1604.00010)
- Pfuhl O., et al., 2011, *ApJ*, 741, 108
- Pickering E. C., 1890, *The Observatory*, 13, 80
- Porquet D., Predehl P., Aschenbach B., Grosso N., Goldwurm A., Goldoni P., Warwick R. S., Decourchelle A., 2003, *A&A*, 407, L17
- Portegies Zwart S. F., Verbunt F., 1996, *A&A*, 309, 179
- Portegies Zwart S. F., McMillan S. L. W., Hut P., Makino J., 2001, *MNRAS*, 321, 199
- Portegies Zwart S., McMillan S. L. W., van Elteren E., Pelupessy I., de Vries N., 2013, *Computer Physics Communications*, 183, 456
- Prodan S., Murray N., Thompson T. A., 2013, preprint, (arXiv:1305.2191)
- Qian S.-B., Liao W.-P., Zhu L.-Y., Dai Z.-B., 2010, *ApJ*, 708, L66
- Raghavan D., et al., 2010, *ApJS*, 190, 1
- Rappaport S., Deck K., Levine A., Borkovits T., Carter J., El Mellah I., Sanchis-Ojeda R., Kalomeni B., 2013, *ApJ*, 768, 33
- Rasio F. A., Ford E. B., 1996, *Science*, 274, 954
- Rauch K. P., Tremaine S., 1996, *New A*, 1, 149
- Riddle R. L., et al., 2015, *ApJ*, 799, 4
- Roberts Jr. L. C., Tokovinin A., Mason B. D., Riddle R. L., Hartkopf W. I., Law N. M., Baranec C., 2015, *AJ*, 149, 118
- Roell T., Neuhauser R., Seifahrt A., Mugrauer M., 2012, *A&A*, 542, A92
- Rosenbluth M. N., MacDonald W. M., Judd D. L., 1957, *Phys. Rev.*, 107, 1
- Rosengren A. J., Scheeres D. J., 2014, *Celestial Mechanics and Dynamical Astronomy*, 118, 197
- Rubin V. C., Ford W. K. J., Thonnard N., 1980, *ApJ*, 238, 471
- Salpeter E. E., 1955, *ApJ*, 121, 161
- Santerne A., et al., 2016, *A&A*, 587, A64

- Schödel R., et al., 2007, *A&A*, 469, 125
- Schödel R., Merritt D., Eckart A., 2009, *A&A*, 502, 91
- Schödel R., Feldmeier A., Kunneriath D., Stolovy S., Neumayer N., Amaro-Seoane P., Nishiyama S., 2014, *A&A*, 566, A47
- Schwamb M. E., et al., 2013, *ApJ*, 768, 127
- Schwarzschild K., 1916, *Abh. Konigl. Preuss. Akad. Wissenschaften Jahre 1906,92*, Berlin, 1907, 1916, 189
- Schwope A. D., Hambaryan V., Schwarz R., Kanbach G., Gänsicke B. T., 2002, *A&A*, 392, 541
- Sepinsky J. F., Willems B., Kalogera V., 2007, *ApJ*, 660, 1624
- Shen Z.-Q., Lo K. Y., Liang M.-C., Ho P. T. P., Zhao J.-H., 2005, *Nature*, 438, 62
- Siebert H., 2005, *Journal for the History of Astronomy*, 36, 251
- Sigurdsson S., Rees M. J., 1997, *MNRAS*, 284, 318
- Skumanich A., 1972, *ApJ*, 171, 565
- Smirnov N., 1948, *Annals of Mathematical Statistics*, 19, 279
- Socrates A., Katz B., 2012, preprint, (arXiv:1209.5723)
- Socrates A., Katz B., Dong S., 2012, preprint, (arXiv:1209.5724)
- Spitzer L., 1987, *Dynamical evolution of globular clusters*
- Storch N. I., Lai D., 2015, *MNRAS*, 448, 1821
- Storch N. I., Anderson K. R., Lai D., 2014, *Science*, 345, 1317
- Sundman K. F., 1912, *Acta Math.*, 36, 105
- Takeda G., Rasio F. A., 2005, *ApJ*, 627, 1001
- Takeda G., Kita R., Rasio F. A., 2008, *ApJ*, 683, 1063
- Tanaka H., Takeuchi T., Ward W. R., 2002, *ApJ*, 565, 1257
- Teyssandier J., Naoz S., Lizarraga I., Rasio F. A., 2013, *ApJ*, 779, 166
- Thompson T. A., 2011, *ApJ*, 741, 82
- Tokovinin A., 2014a, *AJ*, 147, 86
- Tokovinin A., 2014b, *AJ*, 147, 87
- Tokovinin A., Thomas S., Sterzik M., Udry S., 2006, *A&A*, 450, 681
- Tokovinin A., Gorynya N. A., Morrell N. I., 2014, *MNRAS*, 443, 3082
- Toonen S., Nelemans G., Portegies Zwart S., 2012, *A&A*, 546, A70
- Touma J. R., Tremaine S., Kazandjian M. V., 2009, *MNRAS*, 394, 1085
- Tremaine S., Touma J., Namouni F., 2009, *AJ*, 137, 3706
- Trippe S., et al., 2008, *A&A*, 492, 419
- Verbunt F., Zwaan C., 1981, *A&A*, 100, L7
- Vika M., Driver S. P., Graham A. W., Liske J., 2009, *MNRAS*, 400, 1451
- Wahba G., ed. 1990, *Spline models for observational data*
- Walker I. W., 1983, *Celestial Mechanics*, 29, 149
- Wang Q.-D., 1991, *Celestial Mechanics and Dynamical Astronomy*, 50, 73
- Weinberg S., 1972, *Gravitation and Cosmology: Principles and Applications of the General Theory of Relativity*
- Welsh W. F., et al., 2012, *Nature*, 481, 475
- Welsh W. F., et al., 2015, *ApJ*, 809, 26
- Will C. M., 2012, *Classical and Quantum Gravity*, 29, 217001
- Will C. M., 2014a, *Classical and Quantum Gravity*, 31, 244001
- Will C. M., 2014b, *Phys. Rev. D*, 89, 044043
- Winn J. N., Fabrycky D. C., 2015, *ARA&A*, 53, 409

- Wright J. T., Marcy G. W., Howard A. W., Johnson J. A., Morton T. D., Fischer D. A., 2012, *ApJ*, 753, 160
- Wu Y., Lithwick Y., 2011, *ApJ*, 735, 109
- Wu Y., Murray N., 2003, *ApJ*, 589, 605
- Wu Y., Murray N. W., Ramsahai J. M., 2007, *ApJ*, 670, 820
- Wyatt M. C., 2008, *ARA&A*, 46, 339
- Zimmerman N., et al., 2010, *ApJ*, 709, 733
- Zubovas K., Nayakshin S., Markoff S., 2012, *MNRAS*, 421, 1315
- Zwicky F., 1933, *Helvetica Physica Acta*, 6, 110
- van Kampen N. G., 1992, *Stochastic Processes in Physics and Chemistry*. Elsevier Science, Amsterdam
- van den Berk J., Portegies Zwart S. F., McMillan S. L. W., 2007, *MNRAS*, 379, 111

Curriculum Vitae

I was born on December 27, 1988 in Utrecht, the Netherlands, and was raised in Wijk bij Duurstede, a small town to the South-East of Utrecht. From 2001 to 2007, I attended secondary gymnasium education at the Revis Lyceum in Wijk bij Duurstede and Doorn, with a specialisation in the natural sciences.

In 2007, I enrolled in the bachelor programme Physics and Astronomy at Utrecht University. I graduated cum laude in 2010; I carried out my BSc thesis on the mergers on massive main-sequence binaries under supervision of Onno Pols and Selma de Mink at the Sterrekundig Instituut Utrecht (SIU). I continued my education at Utrecht University with the master programme Astrophysics and Space Research. For my MSc research, I was again supervised by Onno Pols, and I studied the long-term dynamical and stellar evolution of hierarchical triple star systems. Unfortunately, the SIU was shut down in 2011 due to reorganisation and budget cuts at Utrecht University, and I continued my MSc research with Onno Pols at the Radboud University Nijmegen, now also supervised by Gijs Nelemans. I graduated cum laude in 2012 (with grade 10/10 for the MSc research), as one of the last astrophysics MSc graduates of Utrecht University.

In 2012, I was fortunate to get offered a PhD position in the computational astrophysics group at Leiden Observatory under supervision of Simon Portegies Zwart. I started with this position in September 2012. I started a collaboration with David Merritt, professor at the Rochester Institute of Technology in Rochester NY, USA, where I visited twice; David and I also met at various other occasions. In 2014, I attended the ISIMA summer school on gravitational dynamics in Toronto, Canada, and initiated collaborations with Hagai Perets, professor at the Technion at the Israel Institute of Technology in Haifa, Israel, and Fabio Antonini, currently a postdoctoral scholar at the Center for Interdisciplinary Exploration and Research in Astrophysics (CIERA) in Evanston IL, USA. In 2015, I visited Haifa and Evanston for conferences and work visits, and also started collaboration with Yoram Lithwick, professor at CIERA.

Shortly after the defence of my thesis, I will participate as a young scientist at the 2016 Lindau Nobel Laureate Meeting in Physics in Lindau, Germany. In September 2016, I will continue my research in astrophysics as a member of the Institute for Advanced Study in Princeton, NJ, USA.

List of publications

First author

1. **Hamers, A.S.**, Antonini, F., Lithwick, Y., Perets, H.B., Portegies Zwart, S.F. *Implications of secular evolution on high eccentricity migration giving rise to hot Jupiters in multiplanet systems*. In preparation.
2. **Hamers, A.S.**, Portegies Zwart, S.F. *Secular dynamics of hierarchical multiple systems composed of nested binaries, with an arbitrary number of bodies and arbitrary hierarchical structure. First applications to multiplanet and multistar systems*. 2016, MNRAS, Volume 459, Issue 3, p.2827-2874 (ADS).
3. **Hamers, A.S.**, Perets, H.B., Portegies Zwart, S.F. *A triple origin for the lack of tight coplanar circumbinary planets around short-period binaries*. 2016, MNRAS, Volume 455, Issue 3, p.3180-3200 (ADS).
4. **Hamers, A.S.**, Perets, H.B., Antonini, F., Portegies Zwart, S.F. *Secular dynamics of hierarchical quadruple systems: the case of a triple system orbited by a fourth body*. 2015, MNRAS, Volume 449, Issue 4, p.4221-4245 (ADS).
5. **Hamers, A.S.**, Portegies Zwart, S.F. *Probing the formation of planetesimals in the Galactic Centre using Sgr A* flares*. 2015, MNRAS, Volume 446, Issue 1, p.710-729 (ADS).
6. **Hamers, A.S.**, Portegies Zwart, S.F., Merritt, D. *Relativistic dynamics of stars near a supermassive black hole*. 2014, Volume 443, Issue 1, p.355-387 (ADS).
7. **Hamers, A.S.**, Pols, O.R., Claeys, J.S.W., Nelemans, G. *Population synthesis of triple systems in the context of mergers of carbon-oxygen white dwarfs*. 2013, MNRAS, Volume 430, Issue 3, p.2262-2280 (ADS).

Co-author

1. Toonen, S., **Hamers, A.S.**, Portegies Zwart, S.F. *The evolution of hierarchical triple star-systems*. Submitted to Computational Astrophysics and Cosmology.
2. Antonini, F., **Hamers, A.S.**, Lithwick, Y. *Dynamical constraints on the origin of hot and warm Jupiters with close friends*. Submitted to ApJ.

3. Jílková, L., **Hamers, A.S.**, Hammer, M., Portegies Zwart, S.F. *Mass transfer between debris discs during close stellar encounters*. 2016, MNRAS, Volume 457, Issue 4, p.4218-4235 (ADS).

Acknowledgements

This thesis would not have been possible without the continuing support of my supervisors, collaborators, colleagues, friends and family, and I would like to dedicate these last pages to their acknowledgement.

First and foremost, I am greatly indebted to my supervisor, Simon Portegies Zwart. Simon, in 2012, when I applied for a PhD position at Leiden Observatory, you put faith in my abilities when you created an additional PhD position out of your VICI grant specifically for me, and with a lot of freedom for pursuing my own interests. You have always been a constant source for new ideas, yet have also given me the time to finish existing projects. Through your guidance and support, I believe that I have matured scientifically, not only by developing my own abilities to find new interesting astrophysical problems and to solve them, but also to establish new collaborations and to develop a scientific career.

I am also much indebted to my co-promotor, Hagai Perets, who I first met at the ISIMA summer school in Toronto in 2014. Hagai, you have been like a second mentor during my PhD. Like Simon, you always have good ideas which are both interesting astrophysically, but also feasible to tackle in practice. You have been very supportive in my scientific career by supporting my effort to find a postdoc position in the US, despite your own offer to come to Haifa. I hope we can continue our productive collaboration in the future.

I would like to thank David Merritt, with whom I wrote my first PhD paper on the dynamics of orbits around a supermassive black hole. David, as an expert you introduced me to this topic which was completely new to me at the time, and I think I cannot have asked for a better teacher. You helped me realise the potential of the TPI code, that would otherwise likely not have seen useful applications.

One of David's former graduate students is Fabio Antonini, who I met already in 2011 during MODEST-11 in Leiden. Fabio, I am very grateful for inviting me to visit CIERA in September 2015, where you introduced me to Yoram Lithwick, and with whom we started collaborating. I would also like to thank Yoram for his supervision and collaboration, and for offering me a position as a postdoctoral fellow at CIERA. In the end I decided to accept the IAS offer, but it was a hard choice to make.

I would also like to thank Erez Michaely, a PhD student in Haifa supervised by Hagai Perets. Erez, thank you for your generous hospitality during my visit in Haifa. I enjoyed our discussions, and our collaboration on the pre-main-sequence evolution of triple stars.

I am also indebted to collaborators from the CASTLE group in Leiden, in particular Silvia Toonen, Lucie Jílková and Tjarda Boekholt. Silvia, I enjoyed the process of developing a code together, and I think we both benefitted from each other's expertise (stellar/binary evolution and

secular evolution). Lucie, I enjoyed our collaboration on stellar flybys, and all the other related discussions. Thank you for giving me a problem to tackle analytically (as you know, those are my favourite types of problems). Your numerical simulations were very motivating for me to find useful analytical relations, and I found it very satisfying to see that these matched your numerical simulations (fairly) well. Tjarda, thank you for integrating an exotic triple system with your code back in 2012. I enjoyed our later discussions on chaos and the Solar system, and our project to find the importance of post-Newtonian cross terms on the precession of Mercury. I hope we can continue working together as postdocs in the future.

I also want to thank the other CASTLE group members, including Arjen, Carmen, Dan, Edwin, Inti, Jeroen, Martha, Michiko, Nathan, Nora, Steven, Thomas and Tjibbaria, many with whom I have had the pleasure of interesting discussions throughout the years. I always enjoyed the group meetings with their interesting discussions on papers or plots that people brought, and also on less scientific topics. Arjen, as our `AMUSE` developer and expert, you have helped me on countless occasions. In particular, when I was working on implementing `SECULARMULTIPLE` in `AMUSE`, I found it very useful to be sharing an office with you. Dan, I enjoyed our discussions on the Galactic Centre and dynamics in general. Edwin, I enjoyed our many conversations and discussions, both on scientific and non-scientific topics, and thank you for inviting me to your home on several occasions. Jeroen, thank you for helping me with implementing GPU acceleration of `TPI` with `SAPPORO`, which you had to partly modify in order to get the code working, and which sped up the code by a factor of 20. Nathan, thank you for helping with the sometimes unexplainable behaviour of `MI6`, and with `AMUSE` in general. I enjoyed our discussions on secular triple dynamics.

I have had the privilege of supervising two very talented students, Huibert het Lam and Emiel Por. Huibert and Emiel, I enjoyed being part of your research projects, and wish you good luck in your future careers.

Last, but not least, I would like to thank my family and (non-scientific) friends, who helped to remind me that life is not only about science and work. Arne, Rozemarijn, Jesse en Anne-Lotte, jullie zijn altijd erg betrokken geweest bij mijn studie en promotie. Ik hoop dat wij onze spel- en gourmettradies kunnen voortzetten, ondanks de voorlopig groter zijnde afstand. Jan Olav, ik kijk altijd uit naar onze regelmatige gamesessies, en hoop dat wij deze kunnen voortzetten. Ronald en Sjanneke, jullie hebben mij door de jaren heen enorm gesteund, niet alleen met morele steun in het soms moeilijke proces van het doen van onderzoek, maar ook in muzikaal opzicht. Ook ben ik Nick en Annelies ontzettend dankbaar dat zij het mogelijk hebben gemaakt dat ik beschik over een prachtige Bechsteinvleugel, waarmee ik mijn muzikale ambities heb kunnen waarmaken. En natuurlijk ook Dieuwertje, voor de pianolessen, al 20 jaar lang, en voor de grote interesse in mijn opleiding en carrière.

Heidi en Arie, jullie zijn net als Ronald en Sjanneke een enorme steun geweest. Arie, jouw buitengewone nieuwsgierigheid voor de astronomie (als niet-astronoom) is voor mij een grote stimulans geweest, en was tevens een grote motivatie om ook populaire stukken te schrijven. Heidi, du hast mich immer gestützt, und du warst immer da für mich in Wijk wenn die Arbeit (oder der Lärm von den Bauarbeiten an der Kooilaan) zuviel für mich wurden.