

Secular Dynamics of Binaries in Stellar Clusters



Chris Hamilton
Emmanuel College
University of Cambridge

This thesis is submitted for the degree of
Doctor of Philosophy

October 2021

Declaration

This thesis is the result of my own work and includes nothing which is the outcome of work done in collaboration except as declared in the Preface and specified in the text. It is not substantially the same as any that I have submitted, or, is being concurrently submitted for a degree or diploma or other qualification at the University of Cambridge or any other University or similar institution except as declared in the Preface and specified in the text. I further state that no substantial part of my thesis has already been submitted, or, is being concurrently submitted for any such degree, diploma or other qualification at the University of Cambridge or any other University or similar institution except as declared in the Preface and specified in the text. It does not exceed the (unlimited) word limit for the Mathematics Degree Committee.

Summary

Chris Hamilton | Secular Dynamics of Binaries in Stellar Clusters

The orbital evolution of two bound point masses (a ‘binary’) perturbed by external tidal forces represents one of the oldest problems in celestial mechanics. Most obviously, tidal perturbations may arise due to an external point mass bound to the binary, as in the Lidov-Kozai (LK) theory of hierarchical triples, but they can also stem from the gravitational field of an extended stellar system (e.g. galaxy or globular cluster) in which the binary resides. Due to the weakness of the external perturbation, the resulting orbital evolution is usually *secular* in nature, i.e. it occurs on timescales much longer than any characteristic orbital period. This thesis is concerned with the secular dynamical evolution of tidally perturbed binary systems.

If problems of this sort are centuries old, what motivation is there to further study them now? In fact, interest in the problem of tidally perturbed binaries has surged recently due to the discovery of various exotic astrophysical phenomena, not least the mergers of compact object (black hole and/or neutron star) binaries by the LIGO/Virgo collaboration. The question of how these binaries shrink rapidly enough to merge within a Hubble time is still an open one, but tidal perturbations may provide the answer. For instance, LK oscillations driven by a tertiary companion can naturally drive a binary orbit to become highly eccentric, boosting gravitational wave emission and substantially speeding up binary coalescence. Similar ideas (with different sources of dissipation at pericentre) have been previously considered for explaining the origin of other exotic objects, such as hot Jupiters, blue stragglers, and Type 1a supernovae. Thus, understanding the tidally-forced eccentricity evolution and possible mergers of binary systems has become a central focus of modern research in astrophysical dynamics.

In this thesis we consider the secular evolution of binaries driven by the tidal gravitational field of an arbitrary axisymmetric host system (‘cluster’) in which the binary moves. We formulate the most general possible theory of tide-driven secular evolution of two bound point masses, applicable to a wide variety of astrophysical systems. Our secular Hamiltonian theory (averaged over both the inner Keplerian orbit of the binary and its outer orbit within the cluster) reproduces classical results — such as LK evolution and the effect of the Galactic tide on Oort Cloud comets — in appropriate limits, but is more general. We then investigate the secular dynamics in detail, uncovering new dynamical characteristics that are far removed from the canonical LK behaviour. We

also extend the secular theory by accounting for the important non-Newtonian effects of general relativistic (GR) perihelion precession and gravitational wave (GW) emission, and the non-secular effect of short-timescale fluctuations in the perturbing torque. These three effects, unavoidably important in many practical applications, add further levels of complexity and richness to the binary dynamics.

The central result of the theory is that the mean-field gravitational tidal potential of a star cluster is often sufficient to torque a binary so that it performs large-amplitude eccentricity oscillations. This result has significant consequences for the dynamical evolution of compact object binaries, many of which reside in stellar clusters. We show that it leads to mergers of compact object binaries which could not have merged if they were isolated, and calculate the resulting observable merger rate.

In summary, then, the purpose of this thesis is three-fold: to formulate a general unified theory of binary dynamical evolution; to propose a possible origin for LIGO/Virgo compact object merger events; and to uncover and explain a range of new, important and beautiful dynamical phenomena.

To Abbey

Acknowledgements

I am lucky to have had several superb mentors. To name just a few I am grateful to:

Roman Rafikov, my supervisor;

Gordon Ogilvie, my advisor;

Henrik Latter;

Alexander Schekochihin;

James Binney;

John Magorrian;

Jean-Baptiste Fouvry;

Tobias Heinemann;

Eugene Vasiliev.

I have also benefited from the generosity of the Department of Applied Mathematics and Theoretical Physics and Emmanuel College, Cambridge; the Rudolf Peierls Centre for Theoretical Physics and Merton College, Oxford; and the Science and Technology Facilities Council (STFC), to which I owe my PhD funding.

Contents

| | |
|--|-----------|
| List of commonly used symbols | x |
| 1 Introduction | 1 |
| 1.1 Motivation | 3 |
| 1.2 Outline of this thesis | 13 |
| Appendices | 17 |
| 1.A The two-body problem in Newtonian gravity | 17 |
| 2 General formulation | 20 |
| 2.1 Introduction | 21 |
| 2.2 Hamiltonian with cluster tides | 21 |
| 2.3 Averaging the tidal Hamiltonian | 25 |
| 2.4 Time-averaging in axisymmetric potentials | 28 |
| 2.5 The secular Hamiltonian | 31 |
| 2.6 Dependence of Hamiltonian coefficients A and Γ on the cluster potential and binary orbit | 33 |
| 2.7 Validity of secular Hamiltonian | 47 |
| 2.8 Discussion | 49 |
| 2.9 Summary | 53 |
| Appendices | 55 |
| 2.A The singly-averaged Hamiltonian in orbital elements | 55 |
| 2.B Recovering the Lidov-Kozai quadrupole Hamiltonian | 55 |
| 2.C Epicyclic orbits | 56 |
| 2.D Signs and sizes of A and Γ | 57 |
| 2.E Octupole Hamiltonian | 60 |
| 2.F Numerical prescription for computing time-averages | 64 |
| 3 Secular dynamics | 65 |
| 3.1 Introduction | 66 |
| 3.2 General aspects of secular dynamics | 67 |
| 3.3 Secular dynamics in the case $\Gamma > 1/5$ | 76 |
| 3.4 The case $0 < \Gamma \leq 1/5$ | 79 |
| 3.5 The case $-1/5 < \Gamma \leq 0$ | 84 |

| | | |
|----------|---|------------|
| 3.6 | The case $\Gamma \leq -1/5$ | 85 |
| 3.7 | Accuracy of the doubly-averaged approximation | 85 |
| 3.8 | Effect of short-range forces on the cluster-tide driven evolution | 97 |
| 3.9 | Discussion | 99 |
| 3.10 | Summary | 106 |
| | Appendices | 107 |
| 3.A | Detailed characteristics of the $-1/5 < \Gamma \leq 0$ regime. | 107 |
| 3.B | Detailed characteristics of the $\Gamma \leq -1/5$ regime. | 108 |
| 4 | The effect of general relativistic precession | 110 |
| 4.1 | Introduction | 111 |
| 4.2 | Dynamical framework | 113 |
| 4.3 | Phase space behaviour | 115 |
| 4.4 | High eccentricity behaviour | 128 |
| 4.5 | Discussion | 139 |
| 4.6 | Summary | 145 |
| | Appendices | 146 |
| 4.A | Mathematical details of phase space behaviour for $\Gamma > 0$ | 146 |
| 4.B | High-eccentricity behaviour for orbits whose eccentricity maxima are found at $\omega = 0$ | 149 |
| 4.C | Analytic solution for orbital elements at high eccentricity | 150 |
| 4.D | Phase space behaviour and maximum eccentricity in $\Gamma \leq 0$ regimes | 159 |
| 5 | The effect of gravitational wave emission | 167 |
| 5.1 | Introduction | 168 |
| 5.2 | High eccentricity results without GW emission | 170 |
| 5.3 | Gravitational wave emission | 174 |
| 5.4 | Numerical examples | 193 |
| 5.5 | Discussion | 203 |
| 5.6 | Summary | 207 |
| | Appendices | 209 |
| 5.A | Relation to Randall & Xianyu (2018) | 209 |
| 6 | Short-timescale fluctuations | 214 |
| 6.1 | Introduction | 214 |
| 6.2 | Dynamical framework | 217 |
| 6.3 | Short-timescale fluctuations and high eccentricity behaviour | 218 |
| 6.4 | The effect of GR precession | 231 |
| 6.5 | Discussion | 242 |
| 6.6 | Summary | 246 |

| | |
|--|------------|
| Appendices | 247 |
| 6.A Singly-averaged equations of motion | 247 |
| 6.B Fluctuating Hamiltonian | 249 |
| 6.C Phase dependence of fluctuating behaviour | 250 |
| 7 Compact object binary mergers in stellar clusters | 253 |
| 7.1 Introduction | 253 |
| 7.2 Dynamical framework | 254 |
| 7.3 Calculation of the merger fractions | 256 |
| 7.4 Merger rates | 261 |
| 7.5 Discussion | 264 |
| 8 Conclusion | 269 |
| 8.1 Summary | 269 |
| 8.2 Future work | 271 |
| References | 278 |

List of commonly used symbols

| | |
|---|--|
| G | Gravitational constant, $6.674 \times 10^{-11} \text{ m}^3 \text{ kg}^{-1} \text{ s}^{-2}$ |
| c | Speed of light, $2.998 \times 10^8 \text{ m s}^{-1}$ |
| M_{\odot} | Solar mass, $1.988 \times 10^{30} \text{ kg}$ |
| pc | Parsec, $3.086 \times 10^{16} \text{ m}$ |
| AU | Astronomical Unit, $1.496 \times 10^{11} \text{ m}$ |
| Φ | Cluster potential |
| ρ | Density distribution corresponding to cluster potential, satisfying the Poisson equation $\nabla^2 \Phi = 4\pi G \rho$ |
| \mathcal{M} | Cluster mass |
| b | Cluster scale radius (for spherically symmetric Φ) |
| (X, Y, Z) | Fixed Cartesian coordinates in cluster rest frame; Z is cluster symmetry axis |
| (R, ϕ, Z) | Fixed cylindrical coordinates, defined through $R \equiv \sqrt{X^2 + Y^2}$ and $\phi \equiv \arctan(Y/X)$ |
| $\mathbf{R}_g(t)$ | Binary's outer (barycentric) orbit around cluster centre |
| \overline{Q} | Value of time-dependent quantity $Q(t)$ time-averaged over the outer orbit $\mathbf{R}_g(t)$ |
| r_p, r_a | Peri/apocentre of binary's outer orbit (for spherically symmetric Φ) |
| a_g, e_g | Generalised semimajor axis and eccentricity of binary's outer orbit (for spherically symmetric Φ) |
| \mathcal{E}, \mathcal{L} | Specific energy and angular momentum of binary outer orbit in the potential Φ |
| T_R, T_ϕ, T_Z | Radial, azimuthal and vertical periods of binary outer orbit |
| $\Omega_R, \Omega_\phi, \Omega_Z$ | Radial, azimuthal and vertical angular frequencies of binary outer orbit |
| A, Γ | Parameters setting strength and characteristics of the doubly-averaged cluster perturbation; they depend on Φ and the choice of outer orbit |
| A^* | Dimensionless measure of A in spherical potentials, $A^* \equiv A/(GM/b^3)$ |

| | |
|----------------------------------|--|
| m_1, m_2 | Binary component masses |
| μ | Shorthand for $G(m_1 + m_2)$ |
| $\mathbf{r} = (x, y, z)$ | Position vector of m_1 relative to m_2 |
| $\Phi_{\alpha\beta}(t)$ | Partial derivative of Φ with respect to r_α, r_β directions (a.k.a. tidal tensor), evaluated at position $\mathbf{R}_g(t)$ |
| a | Binary's inner orbital semimajor axis |
| n_K | Inner orbital Keplerian mean motion, $n_K \equiv \sqrt{\mu/a^3}$ |
| T_b | Inner orbital period |
| e | Inner orbital eccentricity |
| i | Inner orbital inclination, relative to (X, Y) plane |
| ω | Inner orbital argument of pericentre |
| Ω | Inner orbital longitude of ascending node, relative to X axis |
| M | Inner orbital mean anomaly |
| $\langle Q \rangle_M$ | Value of quantity $Q(M)$ averaged over mean anomaly M |
| E | Inner orbital eccentric anomaly, defined through $M = E - e \sin E$ |
| p | Inner orbital periastron distance, $p \equiv a(1 - e)$ |
| L | First Delaunay action, $L \equiv \sqrt{\mu a}$ |
| J | Inner orbital angular momentum, $J = L\sqrt{1 - e^2}$ |
| j | Dimensionless angular momentum, $j \equiv J/L = \sqrt{1 - e^2}$ |
| J_z | Projection of J on Z axis, $J_z = J \cos i$ |
| j_z | Projection of J_z on Z axis, $j_z = j \cos i$ |
| H | Hamiltonian governing the dynamics of the inner binary orbit |
| C | Shorthand for $Aa^2/8$ |
| H^* | Dimensionless Hamiltonian, $H^* \equiv H/C$ |
| H_0 | Kepler Hamiltonian, $H_0 = -\mu/2a$ |
| H_1 | Perturbing cluster potential, to quadrupolar order in tidal approximation |
| H_{GR} | Perturbing Hamiltonian that encodes 1PN GR apsidal precession |
| Θ | Constant of motion in doubly-averaged dynamics, $\Theta = j_z^2 = (1 - e^2) \cos^2 i$ |
| Σ, D | Alternative constants of motion for doubly-averaged dynamics |
| $e_{\text{max}}, j_{\text{min}}$ | Maximum e , minimum j for given doubly-averaged secular cycle |

| | |
|----------------------------------|--|
| i_{\min} | Minimum inclination in doubly-averaged dynamics (coincides with $e = e_{\max}$) |
| e_{\lim} | Upper limit on possible values of e in doubly averaged dynamics, $e_{\lim} \equiv \sqrt{1 - \Theta}$ |
| t_{sec} | Period of secular eccentricity oscillations |
| t_1 | Rough approximation of t_{sec} |
| t_{\min} | Time the binary spends with $j \in (j_{\min}, 2j_{\min})$ in each secular eccentricity cycle |
| ϵ_{GR} | Dimensionless measure of strength of 1PN GR precession relative to cluster tides |
| Δa | Change in semimajor axis due to GW emission over one secular cycle |
| T_{m} | Time required for binary to merge via GW emission |
| δj | Short-timescale (i.e. singly-averaged) fluctuation in j |

1

Introduction

Contents

| | |
|---|-----------|
| 1.1 Motivation | 3 |
| 1.1.1 LIGO/Virgo gravitational wave detections: why do compact objects merge? | 3 |
| 1.1.2 Compact object binaries in stellar clusters | 7 |
| 1.1.3 Blue stragglers and millisecond pulsars in globular clusters | 9 |
| 1.1.4 The impact of cluster tides | 12 |
| 1.2 Outline of this thesis | 13 |
| Appendices | 17 |
| 1.A The two-body problem in Newtonian gravity | 17 |
| 1.A.1 Orbital elements | 17 |
| 1.A.2 Delaunay variables and Hamilton's equations | 18 |

Take two point masses in relative motion, and let them attract each other with a force that is inversely proportional to the square of the distance between them. What happens?

This is the Newtonian (or Keplerian) two-body problem, and it has formed the basis of celestial mechanics since it was first successfully tackled mathematically by Newton in 1687. As every undergraduate physicist knows, the solution to the Newtonian two-body problem is very simple: the bodies orbit their common barycentre, and their relative motion describes a ‘Keplerian ellipse’ with that barycentre at one focus (Figure 1.5; see §1.A for a review). For the purposes of this thesis, any such bound two-body system will be called a ‘binary’, whether it involves two stars, two black holes, the Moon-Earth system, a comet-Sun system, or whatever else.

With that settled, a natural followup question, which turns out to be of huge astrophysical importance, is: *what happens to the internal Keplerian orbit of a binary under the influence of weak perturbations?* It is from this question that the central theme

of this thesis shall emerge. Of course as it stands it is an extremely broad question since, depending on the context, weak perturbations to binaries can be extremely various in character and duration. To illustrate this, let us consider a wide (semimajor axis $a \gtrsim 10^4 \text{AU}$) stellar binary in the Galactic field, somewhere in the vicinity of the Solar neighbourhood. Let us also be rather informal for now — there will be plenty of time for exhaustive referencing of the literature hereafter. Treat the stars as pure point masses with no internal structure, and consider only Newtonian gravity. What kind of weak perturbations will the binary feel?

First, there will be a number of incoherent, transient flyby perturbations from passing stars, molecular clouds, clumps of dark matter, and other bits of Galactic detritus. The great majority of these will be weak perturbations, with impact parameters $\gg a$. These flybys will affect the binary’s barycentric orbit around the galaxy, but also torque the inner Keplerian orbit, potentially modifying the orbital elements. Second, there will be a coherent (as opposed to transient) torque on the binary due to the Galactic tide. This torque will have a periodicity of the order of the Sun’s orbital period around the Galaxy¹, i.e. $\sim 200 \text{Myr}$. Moreover, whereas the aforementioned flyby torques will occur from more or less random directions, the Galactic tidal torque will be strongly biased in a particular direction, owing to the extreme aspect ratio of the Galactic disk². Third, if the binary also happens to have a bound tertiary companion then it will feel a torque due to that companion. This torque, while coherent, will have a periodicity far shorter than that of the Galactic tide; it will not be strongly biased in one direction, but nor will it be isotropic; and depending on the tertiary’s mass and semimajor axis it may or may not dominate the binary’s orbital dynamics.

Of course, if we drop the point-mass and Newtonian assumptions then there exists a slew of weak non-Newtonian and non-secular effects that can potentially modify a stellar binary’s internal Keplerian orbit. These effects can become important should the binary somehow achieve a sufficiently small periastron distance. For instance, when stars pass close enough to one another, their internal fluid motions can lead to tidal friction and an associated orbital decay; general relativistic (GR) precession can cause rapid advance of the binary’s apsidal angle; gravitational wave (GW) emission can cause orbital shrinkage and circularisation; and stellar evolution, mass loss from stellar winds, precession due to rotational bulges, etc. can all play a potentially crucial role in steering the binary’s dynamics. This is not to mention the possibility of strong perturbations, e.g. from passing

¹Though it does not alter our qualitative argument, this is of course not the only time-scale in the problem; for instance, stars in the solar neighbourhood are also oscillating vertically with respect to the Galactic plane with period $\sim 87 \text{Myr}$ (Binney & Tremaine 2008). For binaries in a thin disk the existence of such vertical oscillations tends not to affect the secular dynamics much (see §2.8).

²In technical language, the resulting tidal tensor $\partial^2 \Phi / \partial R_\alpha \partial R_\beta$ will have one strongly dominant component — see §2.C.

stars with impact parameters $\sim a$, from a natal kick of order $\sim 10^2$ km/s when one of the stars goes supernova, and so forth. The list of potential complications multiplies even further if one generalises the problem to include more diverse types of ‘binary’, e.g. an exoplanet orbiting a star, a satellite orbiting the Earth, a pair of tightly-bound white dwarfs, or a supermassive black hole binary at the centre of a galaxy.

What we learn from this brief thought experiment is that the general problem of perturbed astrophysical binaries is one of huge complexity. To make progress on the timescale of a single PhD thesis we must therefore narrow the scope of our investigation. In order to motivate our choice of priorities, let us now provide some observational context. We will begin by discussing the recent LIGO/Virgo detections of gravitational waves emanating from the mergers of compact object binaries.

1.1 Motivation

1.1.1 LIGO/Virgo gravitational wave detections: why do compact objects merge?

In 1916, Einstein predicted that any two massive objects in orbit around one another must radiate gravitational waves (GWs). He also predicted that the radiation was far too weak ever to be detected. Though the first of these predictions was accurate, the second was not: a century on, the LIGO/Virgo Collaboration finally detected a gravitational wave burst produced in a cataclysmic merger of two black holes (Abbott et al. 2016). By 2020, LIGO/Virgo was detecting a compact object (black hole (BH) and/or neutron star (NS)) binary merger roughly every week (The LIGO Scientific Collaboration et al. 2020). As illustrated in Figure 1.1, by the end of the first half of LIGO/Virgo’s Third Observing Run (termed ‘O3a’), there was already a populous ‘graveyard’ of 50 known compact object binary mergers. Hundreds of new merger detections are expected in the coming years.

This high detection rate, however, raises a tantalising theoretical question: what is making all these compact objects collide? Consider that every compact object binary is on some elliptical orbit with semimajor axis a and eccentricity e . Moreover, every such binary is constantly dissipating energy via Einstein’s gravitational radiation, meaning a is getting smaller, and so given enough time they will all eventually merge. The fundamental issue is that a large fraction of massive stellar binaries — potential progenitors of compact object binaries — are born with semimajor axes of $a_0 \sim 1 - 10$ AU (Duquennoy & Mayor 1991). The problem is that 1-10AU is far too large if that compact object binary is then required to merge via GW emission within a Hubble time. More precisely, the

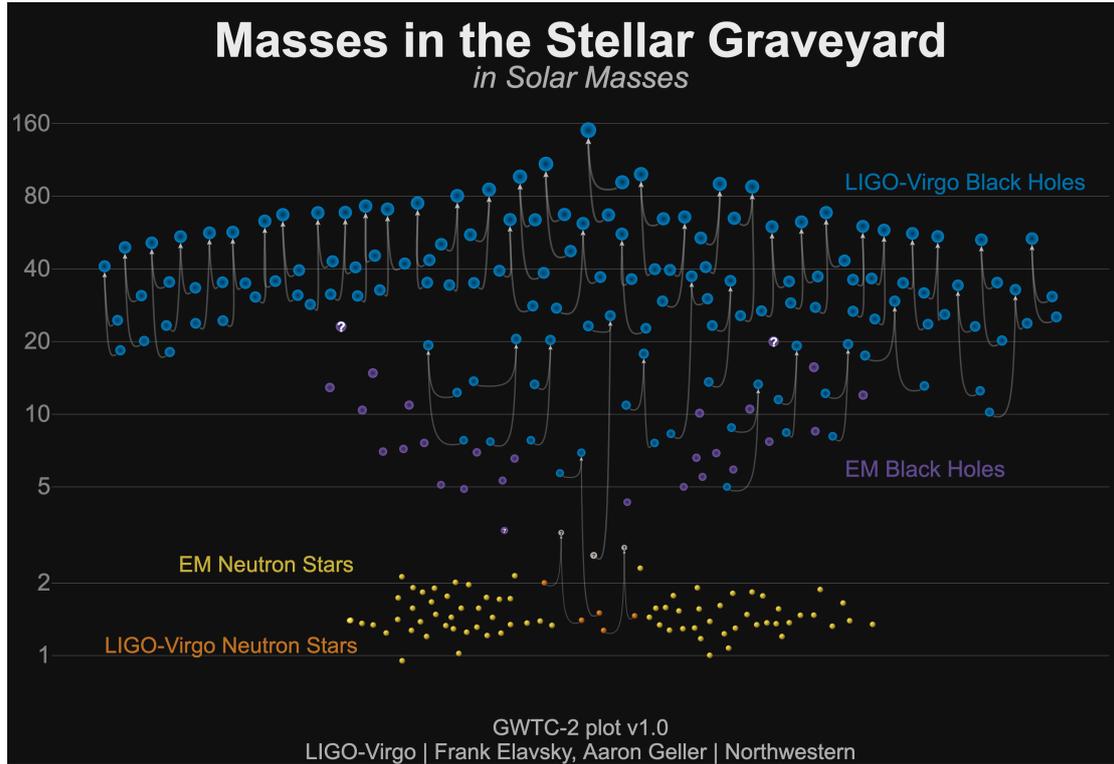


Figure 1.1: Masses of all compact binaries detected by LIGO/Virgo after observing run O3a. Black holes are shown in blue and neutron stars in orange. Also shown are stellar mass black holes (purple) and neutron stars (yellow) detected via electromagnetic observations. [Image credit: LIGO/Virgo/Northwestern Univ./Frank Elavsky; downloaded from <https://www.ligo.org/detections/03acatalog.php> on 14 April 2021].

merger time for an isolated binary with constituent masses m_1 and m_2 on a circular orbit ($e = 0$) is (Peters 1964, equation (5.10)):

$$T_m^{\text{circ}}(a_0) = \frac{5c^5 a_0^4}{256G^3(m_1 + m_2)m_1 m_2} \quad (1.1)$$

$$\approx 10 \text{ Gyr} \times \left(\frac{m}{30M_\odot}\right)^{-3} \left(\frac{a_0}{0.2 \text{ AU}}\right)^4, \quad (1.2)$$

where in the numerical estimate we assumed $m_1 = m_2 = m$. Thus if an isolated BH-BH binary with total mass $60M_\odot$ is to merge within the age of the Universe, it must have a very small initial semimajor axis, $a_0 \lesssim 0.2\text{AU}$. And yet, somehow, black hole binaries *do* merge. Clearly some mechanism is required that can either form compact object binaries with a very small initial separation, or that can greatly diminish the separation distance of compact object binaries after they are formed, boosting the efficiency of their GW emission. Either way, small separation is key. We now discuss some potential mechanisms by which compact object binaries may achieve small separation.



To begin with, we should not immediately rule out the formation and subsequent merger of compact object binaries from massive stellar binary progenitors. We know that massive stars expand to radii of several AU before going supernova. Thus any binary consisting of two massive stars which is not significantly wider than this will inevitably undergo a phase of *common envelope evolution* in which the gaseous envelope of one star engulfs the other (Paczynski 1971; Tutukov & Yungelson 1973; Iben & Livio 1993; Taam & Sandquist 2000; Kalogera et al. 2007; Belczynski et al. 2016). The idea is that provided neither star loses so much mass during this phase that it fails to go supernova, much of the orbital angular momentum of the binary can be transferred to the gaseous envelope that surrounds it. This allows the binary to shrink its orbit significantly and ultimately to merge. However, in practice the common-envelope evolution is extremely difficult to model — it involves orbital dynamics, (magneto)hydrodynamics, stellar evolution, mass transfer, etc. — and as a result it is far from clear that this constitutes a prevalent (or even viable) merger channel in reality. A second, alternative mechanism, which avoids the common envelope phase entirely, is that of chemically homogeneous evolution, or CHE (Mandel & de Mink 2016; de Mink & Mandel 2016). CHE relies on the fact that if a massive star rotates rapidly, the material inside the star is constantly mixed. This mixing means that rather than stellar evolution proceeding in radial shells — first with only hydrogen burning in the core, then with helium burning in the core and a hydrogen-burning shell outside the core, and so on — it instead proceeds homogeneously throughout the stellar core. This prevents the build-up of a radial chemical gradient which in turn means the star does not expand to a problematically large size before it collapses into a compact remnant. If this mixing process occurs for both stars in the binary then an initially small-separation stellar binary turns into a small-separation compact object binary, and hence merges relatively rapidly. Unfortunately, like the mechanism of common-envelope evolution, the CHE mechanism is replete with uncertainties; in particular, most stellar evolution models are one-dimensional (i.e. they assume spherical symmetry), and so the mixing caused by rotation cannot be modelled self-consistently. Finally, a third distinct mechanism that can lead to small separation is the formation and hardening of compact object binaries directly within the dense gaseous disks surrounding Active Galactic Nuclei (e.g. Stone et al. 2017; Bartos et al. 2017; Tagawa et al. 2021). In this case, the binary’s dense environment is advantageous in speeding up the merger: the binary orbit can harden due to gaseous drag and/or flyby stellar encounters³. Once again, this is a very difficult problem which couples star formation, orbital dynamics, and the fluid dynamics of a turbulent magnetized disk.

³In fact, strong encounters between already-formed compact object binaries and passing stars in dense stellar clusters (as opposed to AGN disks) represent yet another popular channel for producing LIGO/Virgo mergers — see §1.1.2.

Thus, just like the other two mechanisms mentioned previously, the errors on the predicted merger rate from this channel encompass several orders of magnitude.

An altogether different way to achieve small separation is by somehow driving an initially wide compact object binary (which may have formed from a wide stellar binary progenitor, or formed dynamically in a cluster core — see §1.1.2) to very high eccentricity e . The key idea here is that for a given semimajor axis a , if the binary’s eccentricity $e \rightarrow 1$ then its pericentre distance $p \equiv a(1 - e)$ is greatly diminished. As a result the components of the binary repeatedly pass very close to one another, allowing significant energy to be dissipated in bursts of GWs, efficiently shrinking the binary orbit and hastening the merger. Mathematically, the merger time for an isolated binary with very large eccentricity $e_0 \approx 1$ is related to the formula (1.1) for circular binaries by (Peters 1964):

$$T_m^{\text{iso}}(a_0, e_0) \approx T_m^{\text{circ}}(a_0) \times \frac{768}{425}(1 - e_0^2)^{7/2}. \quad (1.3)$$

In other words the merger time is smaller than that for circular binaries by an amount $\sim (1 - e_0^2)^{7/2} \ll 1$. Because of this, it becomes possible to merge a much wider binary (i.e. a BH-BH binary with $a_0 \gg 0.2\text{AU}$) simply by exciting sufficiently high eccentricity. Therefore, in recent years a great deal of effort has gone into searching for mechanisms by which compact object binaries might achieve very high eccentricity.

A broad category of proposed mechanisms consists of secular eccentricity excitation of binaries by some perturbing tidal potential⁴. The most famous tidal effect that can result in high eccentricities is the Lidov-Kozai (LK) mechanism (Lidov 1962; Kozai 1962; see Naoz 2016 for a review)⁵. LK theory applies to hierarchical triple systems⁶, i.e. systems where an inner binary is orbited by — and hence torqued by — a bound tertiary perturber. It depends crucially on the the inclination angle between the binary’s inner orbital plane and that of the tertiary perturber around the binary’s barycentre, which takes values $i_0 \in (-\pi, \pi)$. If $|i_0|$ is sufficiently large, then LK theory tells us that the binary’s eccentricity e can be driven periodically to large values. This greatly reduces the pericentre distance p , potentially leading to more rapid mergers via the mechanism described above (see e.g. Blaes et al. 2002; Wen 2003; Antonini & Perets 2012; Antognini et al. 2014; Antonini, Chatterjee, et al. 2016; Silsbee & Tremaine 2017; Liu & Lai 2017 and references therein). From a purely practical point of view, this merger channel has several advantages over those mentioned above: the most basic LK theory can be solved (semi-)analytically,

⁴Note that for the remainder of this thesis, except where explicitly stated, the word ‘tidal’ refers to the tidal gravitational force acting upon a binary due to an external companion (star, stellar cluster, etc.), and not to e.g. the internal fluid tides of a star.

⁵It has recently been appreciated that the LK mechanism was essentially already known to Hugo von Zeipel in the very early 20th Century — see Ito & Ohtsuka (2019).

⁶Similar ideas also apply to quadruple and other multibody systems, and indeed these too have been invoked to explain LIGO/Virgo merger events (see e.g. Fragione & Kocsis 2019).

and there is no complicated fluid dynamics or stellar evolution to contend with, meaning the governing physics is relatively simple and computation is rather inexpensive. LK is also a ubiquitous mechanism in the sense that it is available to any binary that has a tertiary companion. In fact, the simplicity and ubiquity of the LK mechanism means that LK-induced eccentricity excitation has also been invoked as an explanation for the formation of other exotic objects such as hot Jupiters (Fabrycky & Tremaine 2007), Type Ia supernovae (Thompson 2011) and blue stragglers (Antonini, Chatterjee, et al. 2016).

However, despite the considerable appeal of the LK mechanism, it alone cannot solve the problem of compact object merger progenitors. The fact remains that we need to understand how and where the compact object binaries are formed in the first place, and (if we are to appeal to the LK channel) how they came to acquire a bound tertiary companion. Such considerations naturally lead us to think about the densest stellar environments in the Universe — globular and nuclear star clusters.

1.1.2 Compact object binaries in stellar clusters

Globular and nuclear star clusters provide several avenues for the formation of compact object binaries. Three- and four-body encounters in the dense environments of clusters greatly enhance the binary NS formation rate *dynamically*: the abundance per unit mass of low-mass X-ray binaries is around 10^2 times higher in globulars, and 10^3 times higher in the central parsec of the Galaxy, than it is in the Galactic field (Katz 1975; Clark 1975; Generozov et al. 2018). Similarly, BH-BH binaries should form dynamically in cluster cores provided the BHs are retained in their clusters at birth (Portegies Zwart & McMillan 2000; O’Leary et al. 2006; Rodriguez, Chatterjee, et al. 2016; Antonini, Chatterjee, et al. 2016). This possibility is supported by the recent discovery of a detached binary consisting of a BH and a main-sequence turnoff star in the globular cluster NGC 3201 (Giesers, Dreizler, et al. 2018; Giesers, Kamann, et al. 2019).

As the majority of dynamically-formed compact object binaries are too wide to merge via GW emission within a Hubble time, it is not enough to explain how they form: one must also explain how they shrink. Luckily, dense stellar systems represent ideal environments for *perturbing* compact object binaries as well as forming them. Frequent stellar encounters can harden binaries in cluster cores, leading to eventual mergers that might occur after the binary is ejected from the cluster (Antonini & Rasio 2016; Leigh et al. 2018). Cluster cores can also act as factories for the dynamical production of black hole triples, since massive objects naturally sink to the centres of clusters through dynamical friction (Martinez et al. 2020). The tertiary companion to any binary may then drive eccentricity excitation via LK coupling (Antonini, Chatterjee, et al. 2016). For binaries in nuclear star clusters, a central supermassive black hole (SMBH), if present,

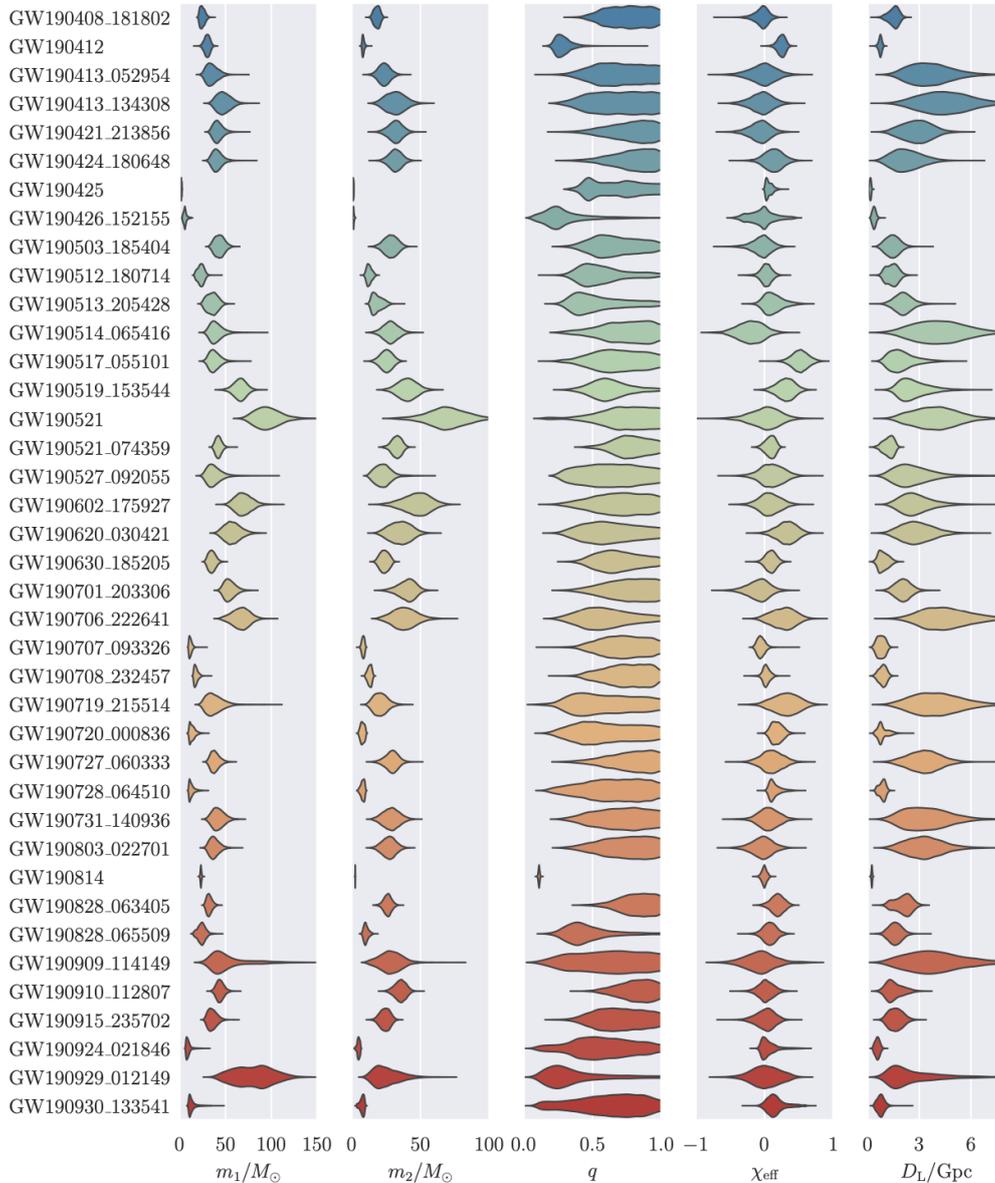


Figure 1.2: Taken from Abbott et al. (2020). Columns show the estimated probability distributions for primary mass m_1 , secondary mass m_2 , mass ratio q , effective spin χ_{eff} , and luminosity distance d_L for all candidate events in the first half of LIGO/Virgo’s third observing run (O3a). The vertical size of each coloured blob is proportional to the corresponding probability estimate.

can play the role of the tertiary driving LK oscillations and orbital decay (e.g. Antonini & Perets 2012; Petrovich & Antonini 2017; Hamers, Bar-Or, et al. 2018a), similar to triples in the field. An additional formation channel comes from four-body effects and exchange interactions which are very common in these dense systems (Miller & Hamilton 2002).

Thus, clusters provide an optimal environment for both forming compact object binaries and merging them via dynamical mechanisms. Moreover, these dynamical mechanisms may help solve some other problems posed by the LIGO/Virgo data, which

contains much more information than the basic *rate* at which compact objects merge in the local universe. To illustrate this, in Figure 1.2 (taken from Abbott et al. (2020)), we show the likelihood estimates for the primary and secondary masses m_1 , m_2 , the mass ratio $q \equiv m_2/m_1$, the effective spin χ_{eff} and the luminosity distance D_L for all merger candidates found in run O3a. Several PhD theses could be written about these data alone; for simplicity we will focus only on the values of the the effective spin χ_{eff} , defined as:

$$\chi_{\text{eff}} \equiv \frac{m_1\chi_1 + m_2\chi_2}{m_1 + m_2}. \quad (1.4)$$

Here $\chi_i \in (-1, 1)$ is the normalised projection of the spin angular momentum vector of the i th binary component (mass m_i) along the binary’s inner orbital angular momentum axis. Thus χ_{eff} represents a mass-weighted combination of these projected spins, and itself takes values $\in (-1, 1)$. If the binary was formed from a massive stellar binary progenitor (and perhaps then merged via one of the first three mechanisms outlined in §1.1.1), one would expect the spins to align with each other and with the orbital angular momentum axis because of mass transfer and/or tidal alignment. Thus in this ‘isolated binary’ scenario we would predict $\chi_{\text{eff}} \approx 1$ (Farr et al. 2018). However, we see from Figure 1.2 that many of the spin measurements are consistent with $\chi_{\text{eff}} = 0$ or even $\chi_{\text{eff}} < 0$. Such measurements suggest the prevalence of dynamical formation channels, for which the spins ought to be more or less randomly aligned (Rodríguez, Zevin, et al. 2016; Liu & Lai 2018). As mentioned in §1.1.1 dynamical merger channels are to be found in abundance in globular and nuclear clusters.

1.1.3 Blue stragglers and millisecond pulsars in globular clusters

In fact, the motivation to study tidally perturbed binaries in stellar clusters extends far beyond the mergers of compact objects. For instance, the formation of the Solar System itself may have been strongly influenced by the cluster environment in which the Sun was presumably born (Brasser et al. 2006; Batygin et al. 2020). Moreover, stellar clusters are known to host an abundance of other exotica thought to have been formed through high-eccentricity migration and/or mergers of two-body systems. Here we touch briefly on two such types of exotic object found in clusters, namely blue stragglers and millisecond pulsars.

First, consider blue stragglers (e.g. Sandage 1953; Boffin et al. 2015). These are stars that sit to the left of the main-sequence turnoff in their host cluster’s colour-magnitude diagram, and hence are bluer than normal main sequence stars of equivalent luminosity — see Figure 1.3 for illustration. Blue stragglers are believed to be formed when a main sequence star receives a fresh supply of hydrogen. Such a replenishment of hydrogen can happen either via mass transfer in a short-period binary system, or via direct merger of the main sequence star with another main sequence star (see Boffin et al. 2015 for

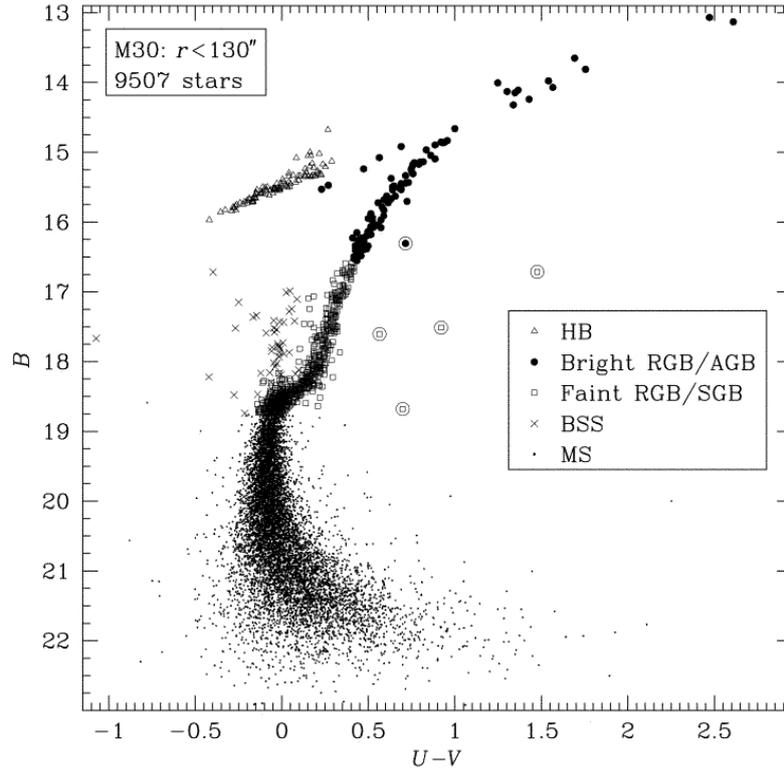


Figure 1.3: Colour-magnitude diagram for 9507 stars in the core of the globular cluster M30. Blue straggler stars (BSS) are indicated with crosses. Figure taken from Guhathakurta et al. (1998).

a modern review). Of course, if one appeals to the latter mechanism then one requires some dynamical channel by which stars can merge efficiently — the probability of chance collisions between individual main sequence stars in a typical globular cluster is far too low (Hills & Day 1976). Unsurprisingly, as we saw with compact objects above, LK oscillations in triple systems have been invoked to explain these supposed mergers (Perets & Fabrycky 2009; Antonini, Chatterjee, et al. 2016)⁷. One may also appeal to direct collisions as a result of chaotic three- and four-body interactions in cluster cores (e.g. Zevin et al. 2019). While in the latter case one would intuitively expect that the relative abundance of blue stragglers would peak in the cluster core and decay with radius thereafter, it is interesting to note that this is not what is observed. Instead, as shown in Figure 1.4 (see also Mapelli, Sigurdsson, Ferraro, et al. (2006)), in many clusters the blue stragglers are distributed bimodally, with a prominent rise in their relative abundance in the cluster’s outskirts. Explaining this radial distribution may require a combination of the binary mass transfer channel (perhaps between primordial binaries in the cluster

⁷Indirect evidence for this channel may perhaps be found in the recent detection of a high binary fraction among blue stragglers in globular clusters; in the cluster NGC3201, a large fraction ($\sim 40\%$) of blue straggler stars are themselves in binary systems (Giesers, Dreizler, et al. 2018).

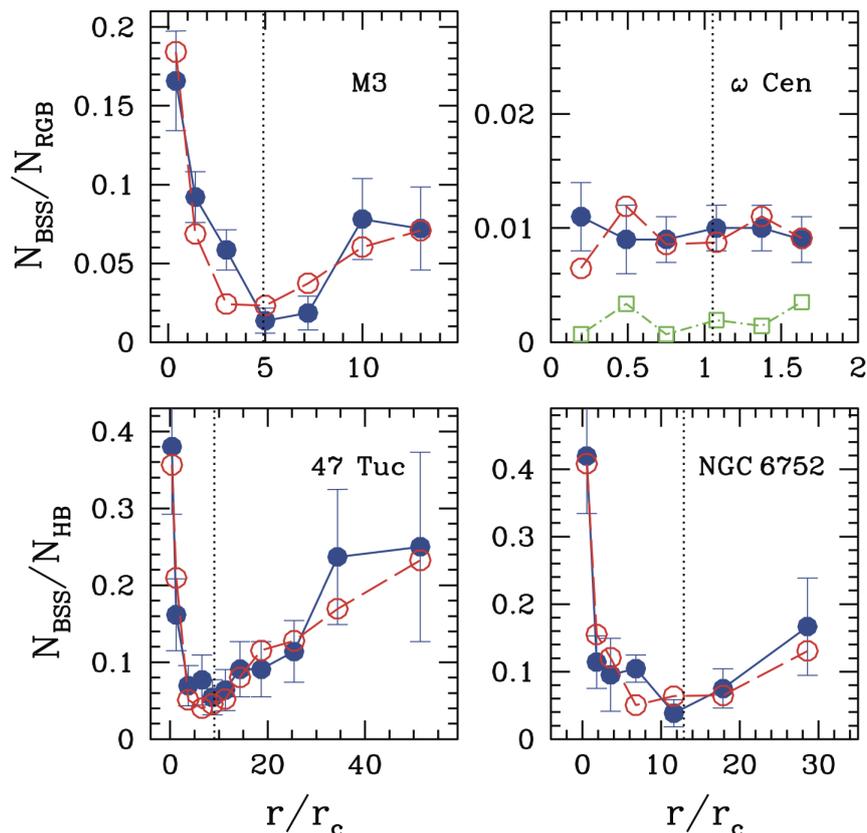


Figure 1.4: From Mapelli, Sigurdsson, Ferraro, et al. (2006). The radial distribution of observed blue straggler stars in various Galactic globular clusters is shown in blue (red symbols are from simulations). Here $N_{\text{BSS}}(r)$ is the number of blue stragglers in a given radial bin around radius r , and r_c is the characteristic radius of the best fit King model (typically $r_c \sim 1\text{pc}$). In each panel, N_{BSS} is normalised either by the radial distribution of red giant branch stars N_{RGB} or horizontal branch stars N_{HB} . Apart from in the (extremely massive, and in several other ways very peculiar) cluster ω Cen, the radial distribution of blue stragglers in each cluster is clearly bimodal; in particular the relative abundance of blue stragglers prominently in the cluster outskirts.

outskirts, see Mapelli, Sigurdsson, Colpi, et al. 2004) and a chaotic collisional channel (Knigge et al. 2009), or something new altogether.

Finally, we consider millisecond pulsars. These are extremely rapidly spinning neutron stars, believed to have been spun up via accretion of matter from a companion star (e.g. Bhattacharya & van den Heuvel 1991; Phinney & Kulkarni 1994). Naively, it is not clear how such systems can be formed since a typical stellar binary with at least one massive component will inevitably undergo a common envelope phase, and it seems unlikely that a compact object-main sequence binary will be the result. Note, however, that out of a total of ~ 2900 pulsars that are currently documented (Manchester et al. 2005)⁸, around 8% reside in globular clusters⁹. This is a very large fraction, given

⁸ATNF pulsar catalogue, <https://www.atnf.csiro.au/people/pulsar/psrcat>, accessed 14 April 2021.

⁹According to the catalogue at <https://www3.mpifr-bonn.mpg.de/staff/pfreire/GCpsr.html>, ac-

that all clusters combined contain much less than 1% of the total stellar mass in our Galaxy¹⁰. It is natural therefore to wonder about dynamical mechanisms for forming millisecond pulsars in such clusters. In principle the neutron star can form in isolation somewhere in the cluster, and then capture a stellar companion during one of the chaotic encounters that are prevalent in the dense cluster core. Subsequently, repeated encounters of the resulting binary system with other cluster stars can gradually ‘harden’ the binary, decreasing its semimajor axis until it is tight enough to allow accretion (Heggie 1975; Pooley et al. 2003), spinning up the neutron star. Moreover, some millisecond pulsar binaries are known to possess finite (though very small) eccentricities. An isolated binary that was undergoing mass transfer would be circularised, whereas the repeated weak flyby encounters experienced by a binary in a stellar cluster are capable of exciting this eccentricity (Heggie & Rasio 1996). Thus, dense stellar clusters are an excellent environment for producing millisecond pulsars with the properties we observe.

1.1.4 The impact of cluster tides

The main thing that has been overlooked in all studies mentioned above — except those on Solar System formation (Gaidos 1995; Morbidelli & Levison 2004; Brassier et al. 2006; Batygin et al. 2020) — is the effect of the smooth component of the cluster’s mean gravitational field on the internal dynamics of the binary. On the other hand a similar effect, that of the smooth Galactic tide, is accounted for routinely in studies of Oort Cloud comet dynamics, and of the evolution of wide binaries in the Milky Way disk (Heisler & Tremaine 1986; Jiang & Tremaine 2010; Veras & Evans 2013a). In both cases it is known that the Galactic tide can drive binaries (i.e. the wide stellar binary or the comet-Sun system) periodically to high eccentricity in a manner very reminiscent of the LK mechanism. This begs the question: might the mean gravitational field of stellar clusters do the same to compact object binaries, stellar binaries, and so on?

Given this motivation, we now declare an interest in the following question:

- what happens to (idealised, point-mass, Newtonian) binary systems when they are perturbed weakly by an external body (e.g. tertiary companion, cluster, or galaxy) to which they are gravitationally bound?

Answering this question will be the primary focus of Chapters 2 and 3. From that investigation it will emerge that the similarity between LK dynamics and the Galactic tide-driven dynamics is not a coincidence. Instead, they are both special cases of a more

cessed 14 April 2021.

¹⁰The total mass in stellar clusters in our galaxy is $\sim 10^8 M_\odot$ while the the remaining stellar population weighs $\sim 10^{11} M_\odot$ (Binney & Merrifield 1998).

general theory of binaries orbiting arbitrary axisymmetric potentials — including cluster potentials — which we derive and explore in detail.

Having found that cluster potentials can torque binaries to high eccentricity, and with a view to investigating compact object mergers, a natural followup question is:

- how is the resulting dynamical evolution impacted by non-Newtonian and non-secular effects, such as GR precession, GW emission and short-timescale fluctuations in the torque?

This question will be addressed in Chapters 4, 5 and 6. There we will uncover several new results, some of which shed new light on the LK and Galactic tide-driven mechanisms, and some of which show markedly different behaviour peculiar to cluster-tide driven systems.

Finally, these different strands of investigation naturally culminate in the question

- can tide-driven eccentricity excitation account for (some of) the compact object mergers currently being detected by LIGO/Virgo?

This question we address quantitatively in Chapter 7.

As we will see over the following 260 or so pages, answering these three questions will involve a wide array of analytical, semi-analytical and numerical calculations, and will conjure up a number of new and non-intuitive physical surprises. Moreover, in several instances we will uncover new results that pertain to the LK scenario and yet had not been appreciated before, despite several decades of work on the LK theory and its applications.

1.2 Outline of this thesis

Above we have given a broad-brush review of the literature and observational motivation for this thesis, as well as an overview of the questions to be addressed. In addition to this each of the following six Chapters will include a subsection entitled ‘Relation to existing literature’ or similar, in which we discuss the scholarly context of that Chapter in detail. Let us now provide a more specific outline of what the main Chapters consist of.

Chapter 2 [based upon Hamilton & Rafikov (2019b)]. In this Chapter we formulate the most general possible theory of tide-driven secular evolution of two bound point masses in arbitrary axisymmetric host systems. We derive a secular Hamiltonian (averaged over both the inner Keplerian orbit of the binary and its outer orbit within the cluster), valid to quadrupole order in the tidal expansion, for an arbitrary cluster potential and explore its characteristics. This doubly-averaged ‘tidal’ Hamiltonian depends on just two parameters, which fully absorb the information about the background cluster potential and the binary’s orbit within it: a dimensional parameter A setting the secular timescale,

and a dimensionless parameter Γ which determines the phase portrait of the binary's inner orbital evolution. We examine the dependence of A and Γ on the cluster potential (both spherical and axisymmetric) and on the binary orbit within the cluster. Our theory reproduces known secular results — such as Lidov-Kozai (LK) theory and the theory of Galactic tide-driven evolution of Oort Cloud comets — in appropriate limits, but is more general. It provides a universal framework for understanding dynamical evolution of various types of binaries driven by the smooth tidal field of any axisymmetric potential.

Chapter 3 [based upon Hamilton & Rafikov (2019c)]. In this Chapter we provide a thorough exploration of the phase-space of the secular Hamiltonian derived above as Γ is varied. We find that for $\Gamma > 1/5$ the phase-space structure and the resulting dynamical evolution of binary orbital elements are qualitatively similar to the LK problem. However, this is only one of four possible regimes, because the dynamics are qualitatively changed by bifurcations at $\Gamma = 1/5, 0, -1/5$. These bifurcations are important because, for instance, binaries in the cores of clobular clusters tend to have $0 < \Gamma \leq 1/5$. We show how the dynamics are altered in each regime and calculate characteristics such as secular evolution timescale, maximum possible eccentricity, etc. We verify the predictions of our doubly-averaged formalism numerically and find it to be very accurate when its underlying assumptions are fulfilled, typically meaning that the secular timescale should exceed the period of the binary around the cluster by $\gtrsim 10 - 10^2$ (depending on the cluster potential and binary orbit).

Chapter 4 [based upon Hamilton & Rafikov (2021)]. For many applications, general-relativistic (GR) apsidal precession is very important, and has been accounted for in various LK calculations. In this Chapter we generalise and extend these LK studies by exploring in detail the effect of GR precession on (quadrupole-level) tidal evolution of binaries orbiting in arbitrary axisymmetric potentials (which includes LK theory as a special case). We study the (doubly-averaged) orbital dynamics for arbitrary strengths of GR and binary initial conditions and uncover entirely new phase space morphologies with important implications for the binary orbital evolution. We also explore how GR precession affects secular evolution of binary orbital elements when the binary reaches high eccentricity ($e \rightarrow 1$), and delineate several different dynamical regimes. Finally we provide an analytic solution to the DA problem at high eccentricity (with or without GR), which has not been presented before even in LK theory.

Chapter 5. The results of Chapters 2-4 are applicable to a variety of astrophysical systems. In particular, they can be used to understand the high-eccentricity behaviour of (cluster) tide-driven compact object mergers, i.e. LIGO/Virgo gravitational wave (GW) sources. In this Chapter we add GW emission into our calculations and aim to understand the physics of cluster tide-driven mergers, thereby expanding upon and generalising the well-known LK-driven merger scenario. We provide for the first time an analytical understanding of the different evolutionary stages of the binary’s semimajor axis, maximum eccentricity, secular oscillation timescale, etc., all the way to merger.

Chapter 6. Everything up to this stage has been tackled using ‘doubly-averaged’ (DA) secular theory. DA theories involve averaging the dynamics both over the binary’s internal Keplerian orbit and its ‘outer’ barycentric orbit relative to the perturber. At the test particle quadrupole level, a binary’s DA maximum eccentricity e_{\max} is always limited by the initial relative inclination i_0 between the two orbits; precisely, one has $e_{\max} \leq e_{\lim} \equiv (1 - \Theta)^{1/2}$ where $\Theta = (1 - e_0^2) \cos^2 i_0$ and e_0 is the initial eccentricity. However, DA theories do not account for fluctuations in the torque on the timescale of the outer orbital period. These short-timescale fluctuations can increase a binary’s maximum eccentricity beyond e_{\lim} , and can also accumulate over time such that the secular behaviour diverges entirely from the initial DA prediction. In this Chapter we consider the impact of these short-timescale fluctuations and derive an approximate expression for their magnitude. We also uncover a new effect, relativistic phase space diffusion (RPSD), which arises from an interplay between extremely high eccentricity secular behaviour, short-timescale fluctuations, and general relativistic pericentre precession. RPSD occurs at eccentricity peaks and is essentially instantaneous compared to the secular timescale, but can change the subsequent secular evolution dramatically. This process is present even in classical LK dynamics, but it has not been uncovered until now.

Chapter 7 [based upon Hamilton & Rafikov (2019a)]. In this Chapter we bring together many results from the previous Chapters to explore quantitatively a new channel for the production of binary mergers in clusters. In this channel, the tidal field of the cluster secularly drives the binary to high eccentricity until gravitational wave emission becomes important. We compute present day merger rates for BH-BH, NS-BH and NS-NS binaries, varying the cluster potential and central concentration of the binary population (but ignoring cluster evolution and stellar flybys). Unlike other mechanisms, this new dynamical channel can produce a significant number of mergers out to cluster-centric distances of several pc. We find merger rates that can contribute to the observed LIGO/Virgo rate at the level of several per cent.

Most of these Chapters contain supplementary material in the form of Appendices, which may be skipped on a first reading. Finally, in Chapter 8 we summarise the results of this thesis and discuss future directions for research.

Appendices

1.A The two-body problem in Newtonian gravity

To solve the problem of perturbed binaries, we must first understand unperturbed binaries. Therefore in this section we will review the most basic problem in celestial mechanics — the problem of two point masses orbiting each other under Newtonian gravity. This will allow us to introduce a few key concepts and establish the notation that we will use throughout this thesis.

1.A.1 Orbital elements

Vital to any study of astrophysical binaries is the set of *orbital elements* which describe the relative elliptical motion. Let the binary components have masses m_1 and m_2 respectively, and let the relative displacement of m_1 from m_2 be \mathbf{r} . Then $\mathbf{r}(t)$ describes a Keplerian ellipse, as illustrated in Figure 1.5a. This diagram is sufficient to define:

- the semimajor axis a of the ellipse;
- its eccentricity e ;
- the mean anomaly M of the binary motion, measured relative to the pericentre line in the orbital plane.

To complete the set of orbital elements we need three more angles, which follow from defining a fixed reference plane (X, Y) and reference direction (X) , as illustrated in Figure 1.5b. This diagram suffices to define:

- the inclination i of the binary orbital plane relative to the (X, Y) plane;
- the longitude of ascending node Ω , relative to the X axis;
- the binary's pericentre¹¹ angle ω , measured relative to the line of nodes.

In practise, the definition of the (X, Y) plane will be related to the symmetry of the potential in which the binary orbits — see §2.2.

The set of six orbital elements $(a, e, i, \omega, \Omega, M)$ specify uniquely the relative position \mathbf{r} and velocity $\dot{\mathbf{r}}$ of the binary components at a given time t . The great advantage of using orbital elements rather than $\mathbf{r}(t)$ and $\dot{\mathbf{r}}(t)$ is that for an isolated binary, five orbital elements — namely $(a, e, i, \omega, \Omega)$ — remain constant, while the mean anomaly evolves linearly with time,

$$M(t) = M(0) + n_K t, \quad (1.5)$$

¹¹also known as the 'perihelion', 'periastron' or 'apsidal' angle.

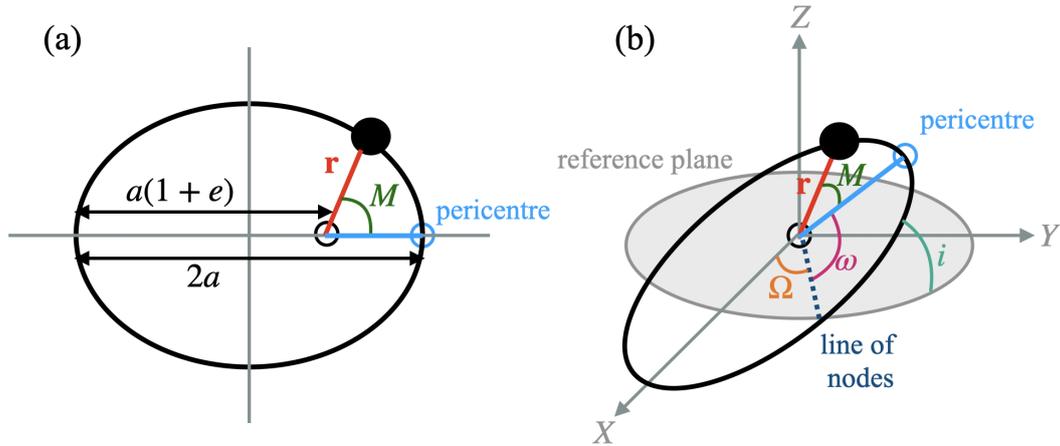


Figure 1.5: Diagrams illustrating the definition of binary orbital elements. (a) In the binary's orbital plane, where \mathbf{r} is the relative separation of the binary components. This diagram is sufficient to define a , e and M . (b) The same ellipse as in (a) viewed from a fixed Cartesian (X, Y, Z) reference frame. The fixed (X, Y) reference plane allows us to define the remaining orbital elements Ω , ω and i .

where $n_K \equiv \sqrt{G(m_1 + m_2)/a^3}$ is the (constant) Keplerian mean motion. The formulae that convert between the orbital elements and the components of \mathbf{r} are given in §2.2.2.

A central focus of this thesis will be on calculating the time evolution of the other five orbital elements when the binary is (weakly) perturbed. This will involve deriving dynamical equations of motion. However, it turns out that the orbital elements are not the best coordinates with which to derive the equations of motion of a perturbed binary, because they are not *canonically conjugate*. A more convenient set of canonically conjugate variables are the *Delaunay variables*.

1.A.2 Delaunay variables and Hamilton's equations

The Delaunay variables are a set of canonically conjugate variables which allow us to take advantage of the Hamiltonian approach to mechanics and thus derive dynamical equations of motion with ease (for much more detail see Murray & Dermott 1999; Arnold 1989). Let us define the canonical Delaunay coordinates $\boldsymbol{\psi}$ and canonical Delaunay momenta \mathbf{I} as follows:

$$\boldsymbol{\psi} = (M, \omega, \Omega), \quad \mathbf{I} = (L, J, J_z), \quad (1.6)$$

where

$$L \equiv \sqrt{G(m_1 + m_2)a}, \quad J = L\sqrt{1 - e^2}, \quad J_z = J \cos i. \quad (1.7)$$

Here J is just the magnitude of a vector \mathbf{J} , which is the total internal orbital angular momentum of the binary divided by its reduced mass $m_1 m_2 / (m_1 + m_2)$, while J_z is the component of \mathbf{J} along the Z axis (Figure 1.5b).

Then for binaries whose motion is governed by an arbitrary Hamiltonian $h(\boldsymbol{\psi}, \mathbf{I})$ (with dimensions of energy per unit mass), the equations of motion are simply Hamilton's equations:

$$\frac{d\boldsymbol{\psi}}{dt} = \frac{\partial h}{\partial \mathbf{I}}, \quad \frac{d\mathbf{I}}{dt} = -\frac{\partial h}{\partial \boldsymbol{\psi}}. \quad (1.8)$$

Once we have derived the equations of motion using Delaunay variables $(M, \omega, \Omega, L, J, J_z)$, it is then trivial to transform back to orbital elements via (1.7).

For an unperturbed binary the governing Hamiltonian is simply $h = -G^2(m_1 + m_2)^2/2L^2$, so that $d\mathbf{I}/dt = \mathbf{0}$ and $d\boldsymbol{\psi}/dt = (n_K, 0, 0)$, as expected.

The work presented in this Chapter has been published
in *Monthly Notices of the Royal Astronomical Society* as
Hamilton & Rafikov (2019b).

2

General formulation

Contents

| | | |
|------------|--|-----------|
| 2.1 | Introduction | 21 |
| 2.2 | Hamiltonian with cluster tides | 21 |
| 2.2.1 | Tidal approximation | 22 |
| 2.2.2 | Orbital elements and Delaunay variables | 24 |
| 2.3 | Averaging the tidal Hamiltonian | 25 |
| 2.3.1 | Singly-averaged Hamiltonian: averaging over the mean anomaly M | 26 |
| 2.3.2 | Doubly-averaged Hamiltonian: averaging over time | 27 |
| 2.4 | Time-averaging in axisymmetric potentials | 28 |
| 2.4.1 | $\Phi_{\alpha\beta}$ in cylindrical coordinates | 28 |
| 2.4.2 | Orbit families and non-commensurable frequencies | 29 |
| 2.4.3 | Time-averages in spherical potentials | 30 |
| 2.4.4 | Time-averages in axisymmetric potentials | 30 |
| 2.5 | The secular Hamiltonian | 31 |
| 2.5.1 | Orbits in a Kepler potential: link to the Lidov-Kozai mechanism | 32 |
| 2.5.2 | Epicyclic orbits in a disk: link to Heisler & Tremaine (1986) | 33 |
| 2.6 | Dependence of Hamiltonian coefficients A and Γ on the cluster potential and binary orbit | 33 |
| 2.6.1 | General properties of A and Γ | 34 |
| 2.6.2 | Behavior of Hamiltonian characteristics in some spherical potentials | 35 |
| 2.6.3 | Behavior of Hamiltonian characteristics in axisymmetric potentials | 40 |
| 2.7 | Validity of secular Hamiltonian | 47 |
| 2.7.1 | Spherical potentials | 47 |
| 2.7.2 | Axisymmetric potentials | 48 |
| 2.8 | Discussion | 49 |
| 2.8.1 | Commensurable frequencies | 50 |
| 2.8.2 | Relation to previous work | 51 |
| 2.9 | Summary | 53 |
| | Appendices | 55 |
| 2.A | The singly-averaged Hamiltonian in orbital elements | 55 |
| 2.B | Recovering the Lidov-Kozai quadrupole Hamiltonian | 55 |
| 2.C | Epicyclic orbits | 56 |
| 2.D | Signs and sizes of A and Γ | 57 |

| | | |
|------------|---|-----------|
| 2.D.1 | Spherical potentials | 58 |
| 2.D.2 | Axisymmetric potentials | 59 |
| 2.E | Octupole Hamiltonian | 60 |
| 2.E.1 | Time-averaging over an axisymmetric torus | 61 |
| 2.E.2 | Link to the test particle octupole LK Hamiltonian | 62 |
| 2.F | Numerical prescription for computing time-averages | 64 |

2.1 Introduction

Given the motivation in Chapter 1 we now seek to investigate, in the most general way possible, the long-term evolution of binary systems moving in the smooth global potential of a much more massive host system. As we will see, it will turn out that the only additional constraint we have to place upon the host system in order to develop a predictive secular theory is that it is axisymmetric. Thus, we devote this Chapter to the derivation of the general Hamiltonian governing secular orbital evolution of a binary in an arbitrary axisymmetric ‘cluster’ potential. We also explore how the characteristics of this Hamiltonian depend on the properties of the cluster potential and the binary’s orbit within it.

The Chapter is structured as follows. In §2.2 we derive the tidal Hamiltonian for the dynamical evolution of binary orbital elements due to any tidal perturbation when expanded to quadrupole order. In §§2.3-2.5 we average the tidal potential over both the binary’s inner orbit and then over many orbits of the binary around the (assumed axisymmetric) cluster, arriving at a simple doubly-averaged (secular) Hamiltonian which describes long-term evolution of the binary’s orbital elements. The coefficients entering this secular Hamiltonian depend on the potential of the host system and the binary’s barycentric orbit within this potential, and we explore this dependence in detail in §2.6. We verify the time-averaging procedure numerically in §2.7. In §2.8 we discuss the limitations of our theory, and show how our general results are connected with various special cases already explored by others (see also Appendices 2.B & 2.C).

2.2 Hamiltonian with cluster tides

Let us consider a binary system with semi-major axis a and eccentricity e , consisting of point masses m_1 and m_2 . The binary components interact gravitationally with each other and with a fixed smooth background potential Φ of a much more massive system, which we will later take to be axisymmetric. The application we have most readily in mind is that of binary stars in the mean field potential of a globular or nuclear star cluster, and for this reason we will frequently refer to m_1 and m_2 as ‘stars’ and to the background

system as ‘the cluster’. However it should be borne in mind that our analysis works for any system of two gravitationally bound objects (binary black holes, comet-Sun system, etc.) moving in any axisymmetric potential (galaxy, open cluster, young stellar cluster, etc.).

Throughout this thesis we will refer to the binary’s orbit around the cluster as the ‘outer orbit’, while the orbit of the binary components about their common barycentre will be called the ‘inner orbit’, to coincide with the standard terminology in Lidov-Kozai (LK) studies (e.g. Naoz 2016). To describe the outer and inner orbits we set up two coordinate systems — see Figure 2.1 for illustration.

The first, given by $\mathbf{R} = (X, Y, Z)$, has its origin at the centre of the cluster. In this coordinate system, the radius vector of the outer orbit, i.e. from the cluster centre to the barycentre of the binary is given by $\mathbf{R}_b = (X_b, Y_b, Z_b)$. The second (non-inertial) coordinate system has its origin at \mathbf{R}_b , and its axes are fixed to be aligned with those of the first system, so only its origin moves. The position of star $i = 1, 2$ in the non-inertial system is then given by $\mathbf{r}_i = (x_i, y_i, z_i)$. The position of star i relative to the centre of the cluster is $\mathbf{R}_i = \mathbf{R}_b + \mathbf{r}_i = (X_b + x_i, Y_b + y_i, Z_b + z_i)$ and the barycentre is at $\mathbf{R}_b = (m_1\mathbf{R}_1 + m_2\mathbf{R}_2)/(m_1 + m_2)$.

The equation of motion of star $i = 1, 2$ is then

$$\frac{d^2(\mathbf{r}_i + \mathbf{R}_b)}{dt^2} = -(\nabla\Phi)_{\mathbf{R}_b + \mathbf{r}_i} - \frac{Gm_j}{|\mathbf{r}_i - \mathbf{r}_j|^3}(\mathbf{r}_i - \mathbf{r}_j), \quad (2.1)$$

for $i \neq j$, where the subscript on derivatives means that we evaluate the derivative at $\mathbf{R}_b + \mathbf{r}_i$.

Defining the relative position $\mathbf{r} = (x, y, z) \equiv \mathbf{r}_1 - \mathbf{r}_2$, and $\mu \equiv G(m_1 + m_2)$, one obtains from (2.1)

$$\frac{d^2\mathbf{r}}{dt^2} = -[(\nabla\Phi)_{\mathbf{R}_b + \mathbf{r}_1} - (\nabla\Phi)_{\mathbf{R}_b + \mathbf{r}_2}] - \frac{\mu\mathbf{r}}{r^3}, \quad (2.2)$$

which is the general equation of relative motion of the binary components.

2.2.1 Tidal approximation

We now employ the tidal approximation, which means that in equation (2.2) we expand the potential $\Phi(\mathbf{R}_b + \mathbf{r}_i)$ around \mathbf{R}_b . The Cartesian components of the vector $\nabla\Phi$ at position $\mathbf{R}_b + \mathbf{r}_i$ are

$$[(\nabla\Phi)_{\mathbf{R}_b + \mathbf{r}_i}]_\alpha = \left(\frac{\partial\Phi}{\partial R_\alpha}\right)_{\mathbf{R}_b} + \sum_\beta \left(\frac{\partial^2\Phi}{\partial R_\alpha\partial R_\beta}\right)_{\mathbf{R}_b} r_{i,\beta} + O(r_i^2), \quad (2.3)$$

where α, β are the coordinate indices, so that e.g. R_α and $r_{i,\alpha}$ represent the components of $\mathbf{R} = (X, Y, Z)$ and $\mathbf{r}_i = (x_i, y_i, z_i)$ respectively. We expect that terms $O(r_i^2)$ will be

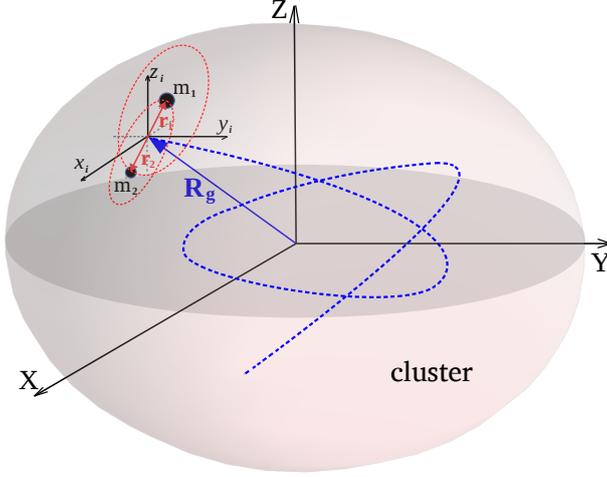


Figure 2.1: Illustration of the binary within the axisymmetric cluster. The binary’s barycentre \mathbf{R}_b coincides, to sufficient accuracy, with a ‘guide’ radius vector \mathbf{R}_g moving as a test particle in the cluster-centric coordinate system (X, Y, Z) . The symmetry axis of the cluster is Z . Binary inclination i is measured relative to the (X, Y) plane and the longitude of the ascending node Ω of the binary is measured with respect to the X axis (§2.2.2). Note that the trajectory of \mathbf{R}_g (illustrated with a blue dashed line) is not closed in a general axisymmetric cluster potential.

subdominant because the distance to the centre of the cluster ($\sim |\mathbf{R}_b|$) is much greater than the binary separation r . In this approximation we find

$$\frac{d^2\mathbf{r}}{dt^2} = -\sum_{\beta} r_{\beta} \frac{\partial}{\partial R_{\beta}} (\nabla\Phi)_{\mathbf{R}_b} - \frac{\mu\mathbf{r}}{r^3}, \quad (2.4)$$

with r_{α} the components of $\mathbf{r} = (x, y, z)$. Keeping in mind that, since the axes of the two coordinate systems are aligned, we may interchange $\partial/\partial R_{\alpha}$ with $\partial/\partial r_{\alpha}$, we can also write this as

$$\frac{d^2\mathbf{r}}{dt^2} = -(\mathbf{r} \cdot \nabla) (\nabla\Phi)_{\mathbf{R}_b} - \frac{\mu\mathbf{r}}{r^3}. \quad (2.5)$$

With x, y, z as our canonical coordinates and $p_x = \dot{x}, p_y = \dot{y}, p_z = \dot{z}$ as the corresponding momenta, these equations of motion may be derived from the time-dependent Hamiltonian

$$H = H_0 + H_1, \quad (2.6)$$

where

$$H_0 = \frac{1}{2}\mathbf{p}^2 - \frac{\mu}{r}, \quad (2.7)$$

$$H_1 = \left(\frac{\partial^2\Phi}{\partial x^2}\right)_{\mathbf{R}_b} \frac{x^2}{2} + \left(\frac{\partial^2\Phi}{\partial y^2}\right)_{\mathbf{R}_b} \frac{y^2}{2} + \left(\frac{\partial^2\Phi}{\partial z^2}\right)_{\mathbf{R}_b} \frac{z^2}{2} \\ + \left(\frac{\partial^2\Phi}{\partial x\partial y}\right)_{\mathbf{R}_b} xy + \left(\frac{\partial^2\Phi}{\partial x\partial z}\right)_{\mathbf{R}_b} xz + \left(\frac{\partial^2\Phi}{\partial y\partial z}\right)_{\mathbf{R}_b} yz, \quad (2.8)$$

and $H_1 \ll H_0$. To compress the notation we write $(\partial^2\Phi/\partial r_\alpha\partial r_\beta)_{\mathbf{R}_b} \equiv \Phi_{\alpha\beta}$ so that the perturbing (‘tidal’) Hamiltonian reads

$$H_1 = \frac{1}{2} \sum_{\alpha\beta} \Phi_{\alpha\beta}(\mathbf{R}_b) r_\alpha r_\beta, \quad (2.9)$$

where we sum over $\alpha, \beta = x, y, z$.

The dominant part of the Hamiltonian H_0 corresponds to the motion of an isolated binary star about its own barycentre, and has no explicit time dependence. The perturbing term H_1 takes into account the tidal effects of the external potential, which will drive the secular evolution of the binary orbital elements. It implicitly depends on time through $\mathbf{R}_b(t)$, which we look at next.

According to equation (2.1) the evolution of \mathbf{R}_b is governed by

$$\begin{aligned} \frac{d^2\mathbf{R}_b}{dt^2} &= -\frac{m_1(\nabla\Phi)_{\mathbf{R}_b+\mathbf{r}_1} + m_2(\nabla\Phi)_{\mathbf{R}_b+\mathbf{r}_2}}{m_1 + m_2} \\ &= -(\nabla\Phi)_{\mathbf{R}_b} \left[1 + O\left(r^2/|\mathbf{R}_b|^2\right) \right]. \end{aligned} \quad (2.10)$$

The small correction terms on the right hand side of this equation mean that, in general, the motion of \mathbf{R}_b in the cluster does not coincide exactly with that of a test particle. However, at the level of accuracy needed in this work we can neglect this difference and assume that \mathbf{R}_b coincides with the ‘guide’ radius vector \mathbf{R}_g , which evolves according to the equation of motion of a test particle in the cluster potential,

$$d^2\mathbf{R}_g/dt^2 = -(\nabla\Phi)_{\mathbf{R}_g}. \quad (2.11)$$

In other words, in the following we set $\mathbf{R}_b = \mathbf{R}_g$ and calculate $\mathbf{R}_g(t)$ using equation (2.11).

Our neglect of the terms quadratic and higher order in r_i in equation (2.3) is equivalent to the so-called ‘quadrupole approximation’ in the hierarchical three-body problem. Keeping the next (quadratic) term in the expansion would correspond to the ‘octupole approximation’, and so on. In Appendix 2.E we describe the extension of our tidal Hamiltonian to octupole order and provide a connection to the LK problem in the octupole approximation.

2.2.2 Orbital elements and Delaunay variables

We now introduce standard orbital elements in the frame of the binary (see also §1.A). The reference direction is taken to be the X direction and the reference plane the (X, Y) plane (see Figure 2.1; we will later take the Z axis to be the symmetry axis of the potential but the assumption of axisymmetry is not needed at the moment). We define binary argument of pericentre ω , inclination i , longitude of ascending node Ω and mean anomaly

M relative to this reference plane and direction. When written in orbital elements the relative coordinates $\mathbf{r} = (x, y, z)$ become (Murray & Dermott 1999):

$$\begin{aligned} x = & a \left(\cos \Omega \left[(\cos E - e) \cos \omega - \sqrt{1 - e^2} \sin E \sin \omega \right] \right. \\ & \left. - \cos i \sin \Omega \left[(\cos E - e) \sin \omega + \sqrt{1 - e^2} \sin E \cos \omega \right] \right), \end{aligned} \quad (2.12)$$

$$\begin{aligned} y = & a \left(\sin \Omega \left[(\cos E - e) \cos \omega - \sqrt{1 - e^2} \sin E \sin \omega \right] \right. \\ & \left. + \cos i \cos \Omega \left[(\cos E - e) \sin \omega + \sqrt{1 - e^2} \sin E \cos \omega \right] \right), \end{aligned} \quad (2.13)$$

$$z = a \sin i \left[(\cos E - e) \sin \omega + \sqrt{1 - e^2} \sin E \cos \omega \right], \quad (2.14)$$

and $M = E - e \sin E$ where E is the eccentric anomaly. It is important that the orbital elements are defined with respect to a reference frame with axis directions fixed in time (c.f. Brassler 2001; Veras & Evans 2013c; Correa-Otto et al. 2017). In the limit of the cluster tide going to zero these orbital elements stay constant.

For dynamical studies it is often more convenient to use Delaunay variables (§1.A), in which the actions

$$L = \sqrt{\mu a}; \quad J = L\sqrt{1 - e^2}; \quad J_z = J \cos i, \quad (2.15)$$

are complemented by their conjugate angles M, ω, Ω . We will use them extensively in Chapter 3. Since Delaunay variables are angle-action variables, we can easily identify the conserved quantities in the Hamiltonian. The dominant part of the Hamiltonian (2.7) reads

$$H_0 = -\frac{\mu}{2a} = -\frac{\mu^2}{2L^2}, \quad (2.16)$$

while the perturbing Hamiltonian is given by equation (2.9) with x, y and z given by equations (2.12), (2.13) and (2.14) respectively (or their Delaunay equivalents).

2.3 Averaging the tidal Hamiltonian

Dynamics of binaries in stellar clusters benefits from a natural separation of scales. For example, a Solar mass binary with $a = 20$ AU has an inner orbital period of ~ 100 years, while its outer orbit around a globular cluster might have a period of $\sim 10^5$ years. As we show in Chapter 3, the resulting secular evolution of the binary's orbital elements due to the tidal potential of the cluster may take $\sim 10^8$ years.

This naturally allows us to simplify our Hamiltonian (2.9), first by integrating out the fast evolution of the mean anomaly M of the inner orbit ('single-averaging', see §2.3.1), and then by also integrating over many (outer) orbits of the binary around the cluster ('double averaging', see §2.3.2).

2.3.1 Singly-averaged Hamiltonian: averaging over the mean anomaly M

We begin by averaging over the shortest timescale in the problem, namely over the inner orbital motion of the binary components around their common barycentre. Our singly-averaged Hamiltonian is

$$\langle H \rangle_M = H_0 + \langle H_1 \rangle_M, \quad (2.17)$$

where the average of a quantity \mathcal{F} over the mean anomaly M is defined as

$$\langle \mathcal{F} \rangle_M \equiv \frac{1}{2\pi} \int_0^{2\pi} \mathcal{F} dM = \frac{1}{2\pi} \int_0^{2\pi} (1 - e \cos E) \mathcal{F} dE. \quad (2.18)$$

The coefficients $\Phi_{\alpha\beta}$ depend on time only through $\mathbf{R}_g(t)$, which is a ‘slow’ variable, so

$$\langle H_1 \rangle_M = \frac{1}{2} \sum_{\alpha\beta} \Phi_{\alpha\beta} \langle r_\alpha r_\beta \rangle_M. \quad (2.19)$$

For reference, the full algebraic expressions for $\langle r_\alpha r_\beta \rangle_M$ in terms of orbital elements are given in Appendix 2.A. The singly-averaged Hamiltonian (2.19) incorporating these expressions is completely general and can be used to describe orbital evolution of binaries moving in an arbitrary external potential. As we see later in this Chapter, and study in detail in Chapter 6, the DA approximation can break down at very high binary eccentricity, and one must describe the system using the SA theory instead. The explicit SA equations of motion are written down in §6.A (equations (6.33)-(6.36)).

Obviously we have eliminated the angle M , therefore the conjugate action $L = \sqrt{\mu a}$ is conserved, and so the binary’s semi-major axis a is constant. When written in Delaunay variables, the singly-averaged Hamiltonian $\langle H_1 \rangle_M$ is a function of J, J_z, ω, Ω and the time t through the time-dependent coefficients $\Phi_{\alpha\beta}(\mathbf{R}_g(t))$.

Example: orbits in a harmonic potential

For illustration, as well as to connect to subsequent results, we consider a binary orbiting in a globular cluster with a triaxial constant-density core. For orbits in this core the potential is that of a three-dimensional harmonic oscillator with frequencies κ_α , namely $\Phi = \sum_\alpha \frac{1}{2} \kappa_\alpha^2 R_\alpha^2$. Then $\Phi_{\alpha\beta} = \kappa_\alpha^2 \delta_{\alpha\beta}$ so the singly-averaged Hamiltonian (2.19) becomes

$$\langle H_1 \rangle_M = \frac{1}{2} \left[\kappa_x^2 \langle x^2 \rangle_M + \kappa_y^2 \langle y^2 \rangle_M + \kappa_z^2 \langle z^2 \rangle_M \right]. \quad (2.20)$$

Let us now consider an axisymmetric core where the symmetry axis is the Z axis. Then $\kappa_x = \kappa_y$ and the binary’s outer orbit fills a section of a cylindrical surface with elliptical cross-section (aligned with the Z axis). Using equations (2.59)-(2.61) we end up with

$$\langle H_1 \rangle_M = \frac{\kappa_+^2 a^2}{8} \times \left[(2 + 3e^2) \left(1 + \frac{\kappa_-^2}{\kappa_+^2} \cos^2 i \right) + 5 \frac{\kappa_-^2}{\kappa_+^2} e^2 \sin^2 i \cos 2\omega \right], \quad (2.21)$$

where $\kappa_{\pm}^2 \equiv \kappa_x^2 \pm \kappa_z^2$.

Note that the dependence on the longitude of the ascending node Ω has dropped out of this Hamiltonian, so the z -component of binary angular momentum J_z is conserved. That is, in the reference frame of a binary orbiting an axisymmetric harmonic potential, the perturbation due to the tidal field of the cluster effectively becomes axisymmetric after averaging only over the inner orbit of the binary (single-averaging), not its outer orbit around the cluster (double-averaging). This is despite the fact that the outer orbit itself is not axisymmetric in general, even after averaging over a long time interval (its projection onto the (X, Y) plane is an ellipse centred at the origin). This property does not hold for arbitrary potentials.

Things simplify further if we assume the core to be spherically symmetric. Without loss of generality we can then assume the outer orbit of the binary to be in the $Z = 0$ plane, and put all frequencies equal to κ (so that $\kappa_- = 0$). We find that

$$\langle H_1 \rangle_M = \frac{\kappa^2 a^2}{4} (2 + 3e^2) = \frac{\kappa^2 L^4}{4\mu^2} \left(5 - 3 \frac{J^2}{L^2} \right). \quad (2.22)$$

This singly-averaged Hamiltonian is now also independent of the argument of pericentre ω . As a result, in this case there is no evolution of eccentricity or inclination of the binary. The only variation of its orbital elements is apsidal precession at the rate

$$\dot{\omega} = \frac{\partial \langle H_1 \rangle_M}{\partial J} = -\frac{3}{2} \frac{\kappa^2}{n_K} \sqrt{1 - e^2}, \quad (2.23)$$

independent of the orientation of the binary orbit (i.e. ω, Ω, i). Here $n_K = \sqrt{\mu/a^3}$ is the Keplerian mean motion of the binary.

2.3.2 Doubly-averaged Hamiltonian: averaging over time

As we already mentioned, binary orbital elements change significantly on timescales that are much longer than the outer orbital period of the binary around its host system. For that reason, it makes sense to average $\langle H \rangle_M$ over many outer orbits. Indicating such time-averages with an over bar, we write:

$$\overline{\langle H \rangle}_M = -\frac{\mu^2}{2L^2} + \overline{\langle H_1 \rangle}_M, \quad (2.24)$$

where the doubly-averaged perturbing Hamiltonian $\overline{\langle H_1 \rangle}_M$ differs from its singly-averaged predecessor $\langle H_1 \rangle_M$ (equation (2.19)) only in that each $\Phi_{\alpha\beta}$ is now replaced by its time-averaged value $\overline{\Phi}_{\alpha\beta}$:

$$\overline{\langle H_1 \rangle}_M = \frac{1}{2} \sum_{\alpha\beta} \overline{\Phi}_{\alpha\beta} \langle r_{\alpha} r_{\beta} \rangle_M. \quad (2.25)$$

This works because the outer orbit $\mathbf{R}_g(t)$ only enters $\Phi_{\alpha\beta}$ and not $r_\alpha r_\beta$.

Equation (2.25) is the doubly-averaged perturbing Hamiltonian and is the main result of this section. It describes the secular evolution of the orbital elements of any binary perturbed by a smooth external potential Φ . However in its current abstract form it is not of much use. In the following section we show how the time-averages $\overline{\Phi}_{\alpha\beta}$ may be calculated for cluster potentials possessing certain symmetries, culminating in the expressions (2.42), (2.43).

2.4 Time-averaging in axisymmetric potentials

So far we did not need to specify anything about the potential Φ . However, we will now focus on binaries orbiting in fixed axisymmetric potentials (§§2.4.1-2.4.2). We then describe how the time-averaging procedure works in practice for binaries in spherical clusters (§2.4.3) and then for general axisymmetric potentials (§2.4.4).

2.4.1 $\Phi_{\alpha\beta}$ in cylindrical coordinates

In an axisymmetric cluster we can choose the symmetry axis of the potential to be the Z axis (like in Figure 2.1). Then it makes sense to write down the derivatives $\Phi_{\alpha\beta}$ in the cylindrical (R, ϕ, Z) coordinate system with origin at the centre of the cluster, where $R = \sqrt{X^2 + Y^2}$ and $\phi = \tan^{-1}(Y/X)$. The axisymmetric potential may then be expressed as $\Phi(R, Z)$, and we find

$$\Phi_{xx} = \frac{1}{2} \left[\left(\frac{\partial^2 \Phi}{\partial R^2} \right)_{\mathbf{R}_g} + \left(\frac{1}{R} \frac{\partial \Phi}{\partial R} \right)_{\mathbf{R}_g} \right] + \frac{1}{2} \cos 2\phi_g \left[\left(\frac{\partial^2 \Phi}{\partial R^2} \right)_{\mathbf{R}_g} - \left(\frac{1}{R} \frac{\partial \Phi}{\partial R} \right)_{\mathbf{R}_g} \right], \quad (2.26)$$

$$\Phi_{yy} = \frac{1}{2} \left[\left(\frac{\partial^2 \Phi}{\partial R^2} \right)_{\mathbf{R}_g} + \left(\frac{1}{R} \frac{\partial \Phi}{\partial R} \right)_{\mathbf{R}_g} \right] - \frac{1}{2} \cos 2\phi_g \left[\left(\frac{\partial^2 \Phi}{\partial R^2} \right)_{\mathbf{R}_g} - \left(\frac{1}{R} \frac{\partial \Phi}{\partial R} \right)_{\mathbf{R}_g} \right], \quad (2.27)$$

$$\Phi_{zz} = \left(\frac{\partial^2 \Phi}{\partial Z^2} \right)_{\mathbf{R}_g}, \quad (2.28)$$

$$\Phi_{xy} = \frac{1}{2} \sin 2\phi_g \left[\left(\frac{\partial^2 \Phi}{\partial R^2} \right)_{\mathbf{R}_g} - \left(\frac{1}{R} \frac{\partial \Phi}{\partial R} \right)_{\mathbf{R}_g} \right], \quad (2.29)$$

$$\Phi_{xz} = \cos \phi_g \left(\frac{\partial^2 \Phi}{\partial R \partial Z} \right)_{\mathbf{R}_g}, \quad (2.30)$$

$$\Phi_{yz} = \sin \phi_g \left(\frac{\partial^2 \Phi}{\partial R \partial Z} \right)_{\mathbf{R}_g}. \quad (2.31)$$

Here ϕ_g is the azimuthal coordinate of \mathbf{R}_g , namely $\tan^{-1}(Y_g/X_g)$, and again the subscripts on derivatives mean ‘evaluated at position $\mathbf{R}_g(t)$ ’. The coefficients $\Phi_{\alpha\beta}$ have certain symmetry properties which will become important when we consider their time-averaged values in §2.4.2.

2.4.2 Orbit families and non-commensurable frequencies

We now consider which orbit families are possible in general axisymmetric potentials. Numerical orbit integration confirms that most orbits in axisymmetric potentials are regular and appear to respect a third integral of motion I_3 alongside energy \mathcal{E} and the Z -component of angular momentum \mathcal{L}_Z (Binney & Tremaine 2008; Merritt 2013). The vast majority of these regular orbits form a tube, or ‘torus’, around the symmetry axis: in an oblate potential they are short-axis tube orbits, while in a prolate potential they are (inner- or outer-) long axis tube orbits. Other possibilities include near-resonant regular orbits and irregular (chaotic) orbits, but both of these are typically scarce compared to the tubes.

We will ignore chaotic orbits in what follows since they are very rare in axisymmetric potentials (Regev 2006). We are left with tube orbits and (near-)resonant non-tube orbits. The resonant family corresponds to $\mathbf{R}_g(t)$ having commensurable frequencies. Mathematically, if we denote the frequencies of motion of $\mathbf{R}_g(t)$ in each direction by the vector $\boldsymbol{\Omega} = (\Omega_R, \Omega_\phi, \Omega_Z)$, we must consider whether there exists any triple of integers $\mathbf{n} = (n_1, n_2, n_3)$ such that

$$\mathbf{n} \cdot \boldsymbol{\Omega} = 0. \quad (2.32)$$

The role of commensurabilities and near-commensurabilities will be discussed in §§2.7-2.8.

If the frequencies are non-commensurable (i.e. the relation (2.32) does not hold), then the outer orbit of the binary will trace out a non-repeating path around the cluster. Over time this path will densely fill a 3D axisymmetric torus whose symmetry axis is Z . We may therefore replace the time-average of a function following an orbit $\mathbf{R}_g(t)$ with a weighted (by the time the orbit spends at a given point) volume-average over the torus.

Since the torus is axisymmetric, time-averaging over non-commensurable orbits inevitably involves integrating the expressions (2.26)-(2.31) over azimuthal angle ϕ_g . As a result, time-averages of $\Phi_{\alpha\beta}$ become

$$\overline{\Phi_{xx}} = \overline{\Phi_{yy}} = \frac{1}{2} \left[\overline{\left(\frac{\partial^2 \Phi}{\partial R^2} \right)_{\mathbf{R}_g}} + \overline{\left(\frac{1}{R} \frac{\partial \Phi}{\partial R} \right)_{\mathbf{R}_g}} \right], \quad (2.33)$$

$$\overline{\Phi_{zz}} = \overline{\left(\frac{\partial^2 \Phi}{\partial Z^2} \right)_{\mathbf{R}_g}}, \quad (2.34)$$

$$\overline{\Phi_{xy}} = 0, \quad (2.35)$$

$$\overline{\Phi_{xz}} = \overline{\Phi_{yz}} = 0, \quad (2.36)$$

see equations (2.26)-(2.31). In practice, vanishing of $\overline{\Phi_{xy}}$, $\overline{\Phi_{xz}}$, $\overline{\Phi_{yz}}$ typically requires averaging over many outer orbits — see §2.7.

2.4.3 Time-averages in spherical potentials

In spherical potentials the outer orbit of the binary remains in the same plane, which can be chosen to coincide with the (X, Y) plane. In this plane the coefficients Φ_{xz} and Φ_{yz} vanish identically. In other words, equation (2.36) holds true even without averaging over the outer orbit. At the same time, $\bar{\Phi}_{xy}$ asymptotically tends to zero only upon averaging over many orbits, as we will see later in §2.7.1.

In the (X, Y) plane the path of \mathbf{R}_g is a rosette, assuming it has non-commensurable radial and azimuthal frequencies; see Figure 2.7 for illustration. Over time the rosette densely fills an axisymmetric annulus with inner and outer radii corresponding to the pericentre r_p and apocentre r_a of the outer orbit $\mathbf{R}_g(t)$. When discussing spherical potentials we will sometimes refer to this as the ‘axisymmetric annulus approximation’.

In this case it is easy to write down an analytical formula for the averages $\bar{\Phi}_{\alpha\beta}$ in terms of r_p and r_a , as averaging over dt can be replaced with averaging over dR via $dt = v_R^{-1}dR$, where $v_R = [2(\mathcal{E} - \Phi(R)) - \mathcal{L}^2/R^2]^{1/2}$ is the radial velocity. Specific energy \mathcal{E} and angular momentum \mathcal{L} of the outer orbit $\mathbf{R}_g(t)$ in a spherical potential Φ can be explicitly expressed as function of r_p and r_a as follows:

$$\mathcal{E}(r_p, r_a) = \frac{r_a^2\Phi(r_a) - r_p^2\Phi(r_p)}{r_a^2 - r_p^2}, \quad (2.37)$$

$$\mathcal{L}(r_p, r_a) = \sqrt{\frac{2[\Phi(r_a) - \Phi(r_p)]}{r_p^{-2} - r_a^{-2}}}. \quad (2.38)$$

With this in mind, we can write the time-average of any radially-dependent function $\mathcal{F}(R)$ as

$$\bar{\mathcal{F}} = \frac{\int_{r_p}^{r_a} dR \mathcal{F}(R) [2(\mathcal{E} - \Phi(R)) - \mathcal{L}^2/R^2]^{-1/2}}{\int_{r_p}^{r_a} dR [2(\mathcal{E} - \Phi(R)) - \mathcal{L}^2/R^2]^{-1/2}}. \quad (2.39)$$

Remembering that only the azimuthally-averaged versions of $\Phi_{\alpha\beta}$ provide non-zero contributions (see §2.4.2 and equations (2.33)-(2.36)), we see that in spherical potentials the time-averages $\bar{\Phi}_{\alpha\beta}$ can be calculated in a straightforward fashion via integration over radius R .

2.4.4 Time-averages in axisymmetric potentials

We would like to generalise the approach of §2.4.3 to axisymmetric potentials $\Phi(R, Z)$. This would involve averaging $\Phi_{\alpha\beta}$ over the (R, Z) cross-section of an axisymmetric torus filled by the outer orbit of the binary — see the central columns of Figures 2.10 & 2.11 for examples of such cross-sections. However, there are several difficulties with this approach.

First, each $dR dZ$ element of the cross-section enters the averaging procedure with a certain weight proportional to the time the orbit spends in it. Calculating this weight is

not trivial and involves the knowledge of the meridional velocity (v_R, v_Z) structure. For a general axisymmetric potential this calculation cannot be done analytically.

Second, even the shape of the outer boundary of the cross-section cannot be determined analytically for a general axisymmetric potential. The difficulty is that the knowledge of \mathcal{E} and the Z -component of angular momentum \mathcal{L}_Z (which are integrals of motion in a general axisymmetric potential) is not sufficient to determine the shape of the meridional cross section of the torus: one also needs to know a third integral of motion I_3 . The exact shape of the torus cross-section is known only for orbits in Staeckel potentials (Binney & Tremaine 2008), since only for those do we have analytic expression for the third integral I_3 . Even then, writing down a formula for the time-averaged coefficients $\overline{\Phi}_{\alpha\beta}$ is tiresome because of the awkward confocal-ellipsoidal coordinate system involved (Binney & Tremaine 2008) and the complicated functional form of the third integral.

For these reasons, in this thesis we usually¹ compute time-averages over the outer orbit in axisymmetric potentials by directly integrating the orbit of a guide $\mathbf{R}_g(t)$ numerically using equation (2.11) for a given set of initial conditions $(R, v_R, \phi, v_\phi, Z, v_Z)$, where v_R is the velocity in the direction of increasing R , etc. This orbit is then used to carry out the time-average of $\Phi_{\alpha\beta}$ using a method outlined in Appendix 2.F.

Note that, unlike in the spherically symmetric case, $\overline{\Phi}_{xz}$ and $\overline{\Phi}_{yz}$ no longer vanish identically due to a symmetry of the potential. Nevertheless, equations (2.35)-(2.36) are still fulfilled upon averaging over many outer orbits.

2.5 The secular Hamiltonian

Despite the fact that in general axisymmetric potentials we cannot write down a useful analytic expression for time-averages, we can still continue our derivation of the secular Hamiltonian owing to the symmetries of the $\overline{\Phi}_{\alpha\beta}$ coefficients (equations (2.33)-(2.36)). These symmetry properties allow us to greatly simplify the doubly-averaged perturbing Hamiltonian (2.25) so that it reads:

$$\overline{\langle H_1 \rangle}_M = \frac{1}{2} \overline{\Phi}_{xx} \langle x^2 + y^2 \rangle_M + \frac{1}{2} \overline{\Phi}_{zz} \langle z^2 \rangle_M. \quad (2.40)$$

Let us define the quantities

$$A \equiv \overline{\Phi}_{zz} + \overline{\Phi}_{xx}, \quad B \equiv \overline{\Phi}_{zz} - \overline{\Phi}_{xx}, \quad \Gamma \equiv B/3A, \quad (2.41)$$

¹There are special cases in certain axisymmetric potentials where we can compute time-averages (semi-)analytically, see §2.6.1 & §2.6.3.

(note that the constants A and B are *not* related to the Oort constants!) and write x, y, z in terms of orbital elements using equations (2.59)-(2.61). Then the Hamiltonian (2.40) becomes

$$\overline{\langle H_1 \rangle}_M = CH_1^* \quad \text{where} \quad C = Aa^2/8, \quad (2.42)$$

and H_1^* is the ‘dimensionless Hamiltonian’

$$H_1^* = (2 + 3e^2)(1 - 3\Gamma \cos^2 i) - 15\Gamma e^2 \sin^2 i \cos 2\omega. \quad (2.43)$$

Note that H_1^* involves only a single dimensionless parameter Γ , while C depends on A (which has units of (frequency)²). In Chapter 3 we will see that Γ determines the phase space structure of the Hamiltonian while A sets the timescale for secular evolution. All the information about the cluster properties and the characteristics of the (outer) orbit of the binary enter the Hamiltonian through these two parameters only. In §2.6 we investigate in detail how these key quantities depend on the form of the background potential and the outer orbit of the binary within the potential.

The dependence of the Hamiltonian upon the longitude of ascending node Ω has dropped out under time-averaging and so the z -component of angular momentum J_z is conserved, as we would expect for an axisymmetric perturbation (which the cluster potential looks like from the binary frame upon averaging over many outer orbits). The dimensionless quantity $J_z/L = \sqrt{1 - e^2} \cos i$ is an integral of motion, which implies that variations of eccentricity must come at the expense of changes in inclination and vice versa, just as in the LK mechanism (Lidov 1962; Kozai 1962; Naoz, Farr, et al. 2013).

Finally, we note that the doubly-averaged perturbing Hamiltonian (2.42) appears very similar to the singly-averaged one derived for the example of an axisymmetric harmonic potential in §2.3.1 (equation (2.21)). Indeed, comparing equations (2.21) and (2.42) one might be tempted to say that axisymmetric harmonic potentials have $\Gamma = -\kappa_-^2/3\kappa_+^2$. However, this correspondence is a mathematical coincidence: the assumption of non-commensurability (§2.4.2) does not apply to harmonic potentials, for which all orbits are closed non-precessing ellipses. Despite their similarities the Hamiltonians (2.21) and (2.42) are different objects derived under different approximations.

2.5.1 Orbits in a Kepler potential: link to the Lidov-Kozai mechanism

Another example of such a mathematical coincidence is represented by the well known test particle quadrupole Lidov-Kozai (LK) problem (Lidov 1962; Kozai 1962). The Hamiltonian for this problem takes the form (2.43) if we were to set $\Gamma = 1$. However, we have derived this Hamiltonian under the assumption that \mathbf{R}_g fills an axisymmetric torus, while in the LK case \mathbf{R}_g moves in a Keplerian ellipse, which in the test particle limit does

not precess. Nevertheless, it is known that for elliptical orbits the time-averaged tidal Keplerian potential is exactly axisymmetric to quadrupole order (e.g. Katz, Dong, & Malhotra 2011; Naoz, Farr, et al. 2011), and so (2.42) *does* in fact hold.

In Appendix 2.B we show explicitly how the LK Hamiltonian is recovered in the ‘test particle quadrupole’ approximation from the singly-averaged equation (2.19) in the limit that the background potential Φ in which the binary orbits arises from a point mass at the origin. We recover the LK Hamiltonian exactly if we set $\Gamma = 1$ in (2.43).

2.5.2 Epicyclic orbits in a disk: link to Heisler & Tremaine (1986)

For wide binaries in the solar neighbourhood, the tidal potential of the Galactic disk can provide the dominant torque, as shown by Heisler & Tremaine (1986) for the Oort Cloud comets. Averaged over many orbits of the Sun around the Galaxy, the Galactic disk provides an axisymmetric tide onto the binary. In Appendix 2.C we show how to calculate $\bar{\Phi}_{\alpha\beta}$ in the case where \mathbf{R}_g performs epicyclic motion in a disk. It is then easy to recover the tidal Hamiltonian of Heisler & Tremaine (1986) from (2.42). We reproduce the dimensionless version of Heisler & Tremaine (1986)’s Hamiltonian by setting $\Gamma = 1/3$ in (2.43).

2.6 Dependence of Hamiltonian coefficients A and Γ on the cluster potential and binary orbit

All of the information about the effect of the tidal potential on secular dynamics of the binary is contained in the two crucial quantities A and Γ , which are constructed from the time-averages $\bar{\Phi}_{zz}$ and $\bar{\Phi}_{xx}$, see equation (2.41).

The quantity A is a direct measure of the strength of the potential. Its influence on the dynamics is trivial: it enters the problem only as a proportionality constant of the Hamiltonian (equation (2.42)), and therefore merely sets the (inverse of the) secular timescale. In addition, A is also a fairly intuitive quantity: if a binary is in a strong tidal potential we expect it will have a large A .

The meaning of Γ is less intuitive than A although its influence upon the system is quite profound. In Chapter 3 we will see that the phase portrait of the Hamiltonian H_1^* undergoes several bifurcations as we change the value of Γ , altering the dynamics completely. In particular, we show that there are four qualitatively different regimes — (i) $\Gamma > 1/5$, (ii) $0 < \Gamma \leq 1/5$, (iii) $-1/5 < \Gamma \leq 0$, and (iv) $\Gamma \leq -1/5$. The value of Γ is so important because, for instance, high-eccentricity excitation is ubiquitous for binaries in regime (i), whereas it is much less readily achieved by binaries in regime (ii). Hence, most effort in this section will be directed towards understanding which cluster potentials

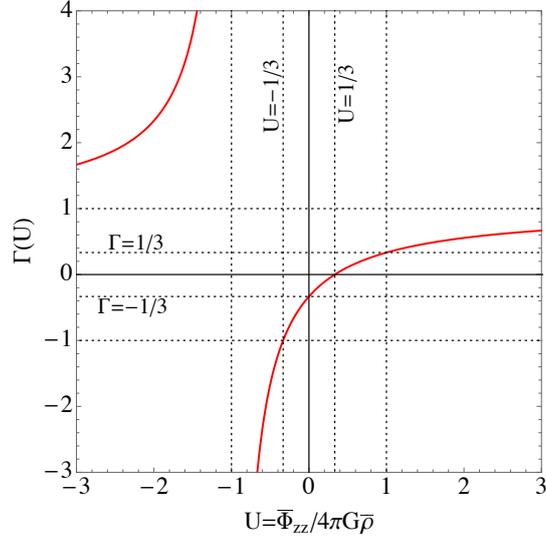


Figure 2.2: Plot of the function $\Gamma(U)$ defined by equation (2.47) with $U = \bar{\Phi}_{zz}/(4\pi G\bar{\rho})$ given by equation (2.46). See text for details.

and outer binary orbits give rise to which values of Γ . So far we have seen that $\Gamma = 1$ for any orbit in a Keplerian potential, and that $\Gamma = 1/3$ for epicyclic orbits in a thin disk. In this section we explore in more detail how the values of Γ (and A) depend on the form of the background potential Φ and the binary's outer orbit \mathbf{R}_g within it.

We start by stating some general properties of A and Γ in §2.6.1. We then discuss the behavior of these parameters in certain spherical (§2.6.2) as well as axisymmetric (§2.6.3) potentials.

2.6.1 General properties of A and Γ

Writing down Poisson's equation in cylindrical coordinates

$$\frac{\partial^2 \Phi}{\partial R^2} + \frac{1}{R} \frac{\partial \Phi}{\partial R} + \frac{\partial^2 \Phi}{\partial Z^2} = 4\pi G\rho, \quad (2.44)$$

and using equations (2.33)-(2.34) & (2.41), one can easily show that in a general axisymmetric potential

$$A = \frac{1}{2}(\bar{\Phi}_{zz} + 4\pi G\bar{\rho}), \quad B = \frac{1}{2}(3\bar{\Phi}_{zz} - 4\pi G\bar{\rho}), \quad (2.45)$$

where $\bar{\rho}$ is the cluster density in the vicinity of the binary time-averaged over many outer orbital periods. Then, defining the dimensionless ratio

$$U \equiv \frac{\bar{\Phi}_{zz}}{4\pi G\bar{\rho}} = -1 + \frac{A}{2\pi G\bar{\rho}}, \quad (2.46)$$

we can write Γ quite generally as

$$\Gamma(U) = \frac{3U - 1}{3(U + 1)}. \quad (2.47)$$

The function $\Gamma(U)$ is plotted in Figure 2.2.

In principle there are no limits on the values U can take, although in practice, achieving values of U less than -1 (and hence $\Gamma > 1$) may require rather contrived orbits. An example of such an orbit with $U < -1$ and $\Gamma > 1$ is given in Appendix 2.D (see Figure 2.13). Note that $U < -1$ necessarily implies that $A < 0$, see equation (2.46).

Somewhat stronger statements can be formulated for realistic spherically symmetric cluster potentials, as we show in Appendix 2.D. In particular, one can demonstrate that in such potentials $A > 0$, $B \geq 0$, and $0 \leq \Gamma \leq 1$. In non-spherical potentials negative values of Γ become possible for certain binary orbits as we will show in §2.6.3.

It is instructive to consider the values of Γ for some specific potentials Φ .

- In the case of a Keplerian cluster potential, i.e. a spherical point mass potential with vanishingly small density ρ outside the centre, one has $\Phi_{zz} > 0$, $\bar{\rho} \rightarrow 0$, $U \rightarrow +\infty$ and $\Gamma \rightarrow 1$ (see §2.5.1).
- In a spherical harmonic potential, symmetry dictates that $\Phi_{zz} = (1/3)\nabla^2\Phi = (4\pi/3)G\rho$ so that $U = 1/3$ and $\Gamma = 0$ (see §2.3.1).
- In a spherical cluster with a cusped density distribution $\rho \propto r^{-\beta}$ with $0 < \beta < 3$ (having finite mass at the centre) we have $\Gamma = \beta/[3(4 - \beta)]$, see Appendix 2.D.1.
- In a thin galactic disk, assuming that Φ_{zz} dominates over other spatial derivatives in Poisson's equation, one has $\Phi_{zz} \approx 4\pi G\rho$; hence we find $U = 1$ and $\Gamma = 1/3$ (see §2.5.2).
- In the opposite limit of a 'cylindrical' (or highly prolate) potential $\Phi = \Phi(R)$ with no Z -dependence, one has $\Phi_{zz} = 0$, $U = 0$ and $\Gamma = -1/3$.

The values (or ranges) of U and $\Gamma(U)$ for these and some other types of cluster potential are summarized in Table 2.4. We stress again that even though applying the 'axisymmetric annulus' approximation gives the correct results for Keplerian and spherical harmonic potentials, this is a mathematical coincidence unless the outer orbit of the binary in these potentials is circular (see §2.3.1 & 2.5.1).

2.6.2 Behavior of Hamiltonian characteristics in some spherical potentials

In spherical potentials the values of $\bar{\Phi}_{\alpha\beta}$ that enter A and Γ are computed using the analytic expression (2.39), which for a fixed potential depends only on the peri/apocentre

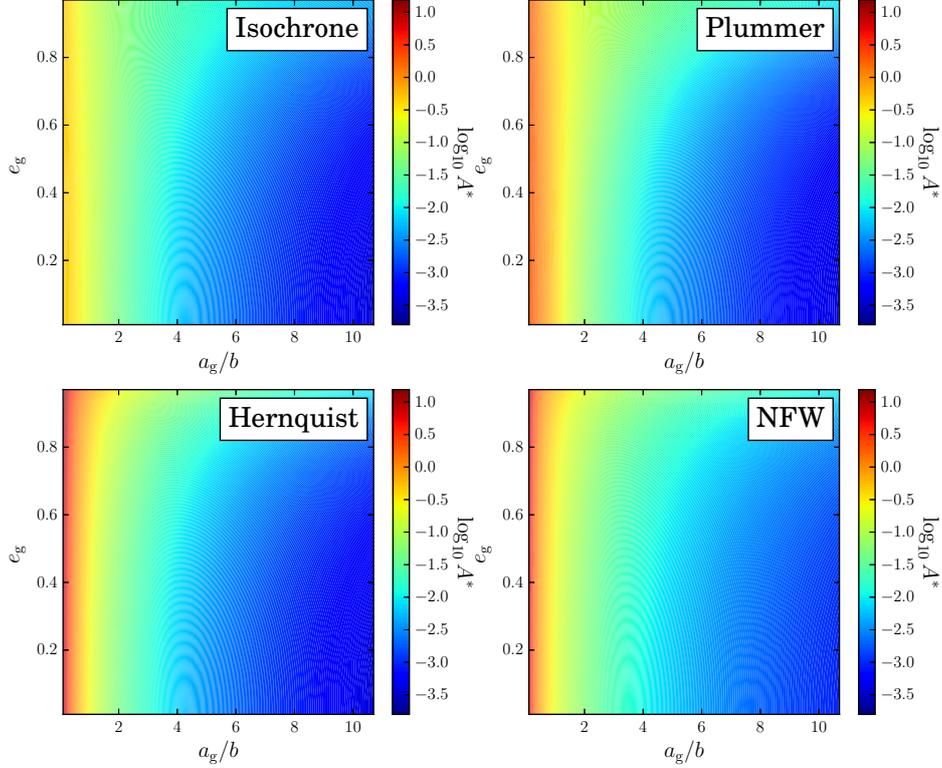


Figure 2.3: Contour plots of $\log_{10} A^*$, where $A^* \equiv A/(GM/b^3)$, as a function of generalized semi-major axis a_g and eccentricity e_g (see equation (2.48)) for binary orbits in four spherical potentials (equations (2.50)-(2.53)), each with scale radius b . The value of A^* in general depends both on a_g and e_g .

(r_p, r_a) of the binary's outer orbit $\mathbf{R}_g(t)$. We can define the outer orbit's generalised semi-major axis a_g and generalised eccentricity e_g in terms of the peri/apocentre as

$$a_g \equiv \frac{1}{2}(r_a + r_p); \quad e_g \equiv \frac{r_a - r_p}{r_a + r_p}. \quad (2.48)$$

These reduce to the usual orbital elements in the case of a Keplerian potential. These variables fully characterize the outer orbit of the binary in a given spherical potential.

In any spherical potential with scale radius b and total mass \mathcal{M} we can also construct the dimensionless parameter $A^* \equiv A/(GM/b^3)$; this normalization arises because A is constructed from the second derivatives of the potential, which are of order² GM/b^3 , see equation (2.41). This allows us to estimate

$$A = 226 \text{ Myr}^{-2} \times \left(\frac{A^*}{0.5}\right) \left(\frac{\mathcal{M}}{10^5 M_\odot}\right) \left(\frac{b}{\text{pc}}\right)^{-3}. \quad (2.49)$$

²Note that $2\pi/\sqrt{GM/b^3}$ is roughly $\sim T_\phi$, the characteristic azimuthal orbital period of the binary around the cluster, so that $A \sim 4\pi^2 A^*/T_\phi^2$.

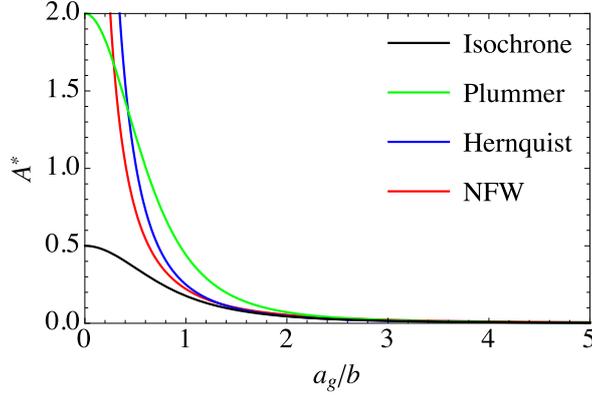


Figure 2.4: Plots of A^* for circular outer orbits ($e_g = 0$) as a function of a_g/b (where b is a scale radius) in the same four potentials as in Figure 2.3.

Both A^* and Γ are dimensionless numbers which, for a given potential, depend only on a_g and e_g . In the following we will explore the dependence of A^* (rather than A , which also depends on the cluster mass and size) and Γ on the shape of the potential and the binary orbit in it.

We use the following four potentials, which give a representative overview of the possible behaviours of A in realistic spherically symmetric clusters:

(i) the isochrone potential (which has a constant-density core and half mass radius $r_h = 3.06b$)

$$\Phi_{\text{iso}}(r) = -G\mathcal{M}/(b + \sqrt{b^2 + r^2}), \quad (2.50)$$

(ii) the Plummer potential (also has a core and $r_h = 1.31b$)

$$\Phi_{\text{Plum}}(r) = -G\mathcal{M}/\sqrt{b^2 + r^2}, \quad (2.51)$$

(iii) the Hernquist potential (has no core, and $r_h = 2.41b$)

$$\Phi_{\text{Hern}}(r) = -G\mathcal{M}/(b + r), \quad (2.52)$$

(iv) the NFW potential (has no core and has a divergent mass profile)

$$\Phi_{\text{NFW}}(r) = -G\mathcal{M}r^{-1} \ln(1 + r/b). \quad (2.53)$$

The NFW potential arises from a density distribution

$$\rho(r) \propto \left(\frac{r}{b}\right)^{-1} \left(1 + \frac{r}{b}\right)^{-2}. \quad (2.54)$$

In equation (2.53) the quantity \mathcal{M} is *not* the mass of the model (which is formally infinite), just a constant with units of mass.

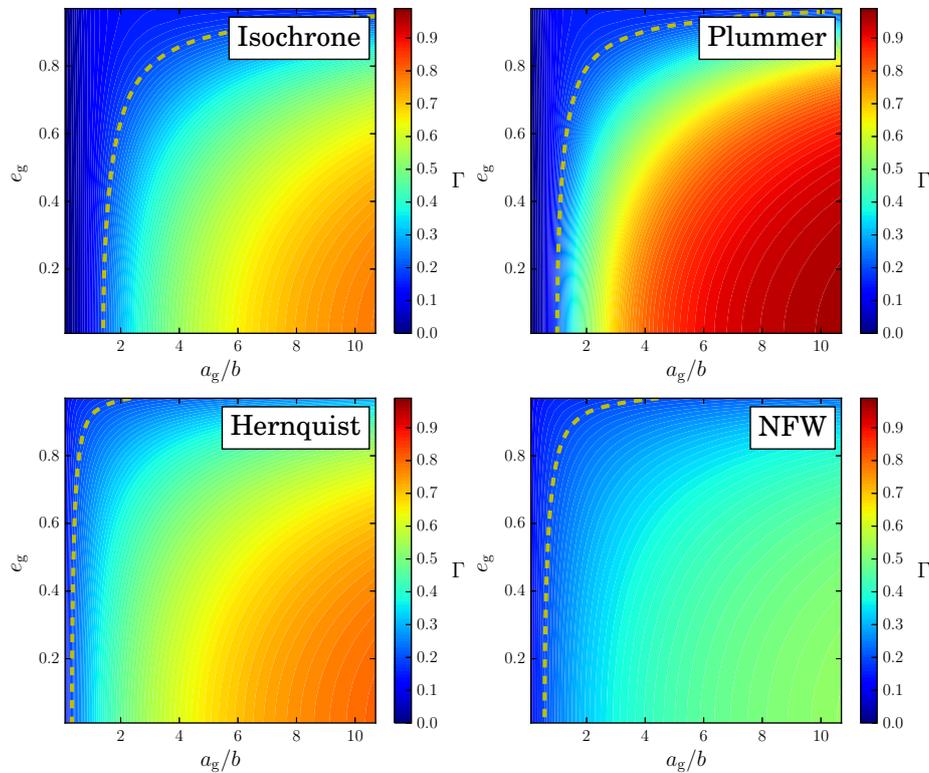


Figure 2.5: Same as Figure 2.3 but showing contour plots of the parameter Γ . Note that $\Gamma \rightarrow 0$ at the centre of the cored potentials (isochrone and Plummer). The dashed yellow in each plot corresponds to $\Gamma = 1/5$, which is the location of an important bifurcation in the dynamical phase portrait, as we show in Chapter 3.

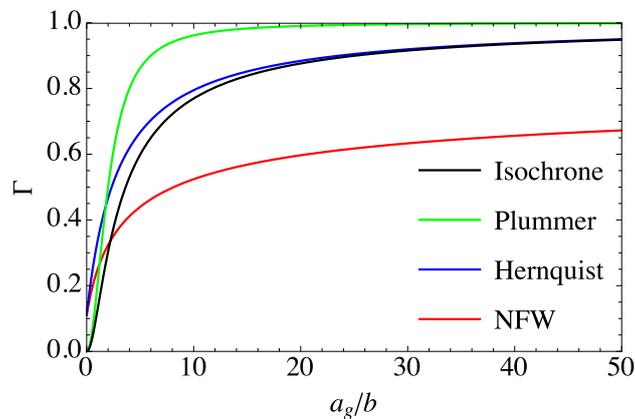


Figure 2.6: Behavior of the parameter Γ for circular orbits ($e_g = 0$) as a function of a_g/b (where b is a scale radius) in the same four potentials as in Figure 2.5. This plot demonstrates that Γ converges to 1 for potentials that are Keplerian at large radii. In the opposite limit $a_g/b \rightarrow 0$, we see that $\Gamma \rightarrow 0$ for the cored potentials and $\Gamma \rightarrow 1/9$ for NFW, as expected (§2.6.2).

For illustration, in Figure 2.7 (left panels) we show three examples of numerically integrated orbits in some of these potentials. (Each orbit was integrated for 100 azimuthal periods; the first few periods are highlighted in red). The first two (‘I’ and ‘II’) orbit the isochrone potential (2.50), which has a constant density core for $r \lesssim b$. The third (‘III’) orbits the Hernquist potential (2.52), which is coreless. In Table 2.1 we list the peri/apocentre $r_{p/a}$, semi-major axis a_g , eccentricity e_g , azimuthal period T_ϕ , and the values of A^* and Γ calculated using equation (2.39). We also provide values of $A^*_{\text{num}}, \Gamma_{\text{num}}$ obtained by direct averaging of $\Phi_{\alpha\beta}$ along each numerically integrated outer orbit (see Appendix 2.F), to which we will return when discussing the validity of the axisymmetric averaging approximation in §2.7.

Behavior of A^*

In Figure 2.3 we plot $\log_{10} A^*$ in the (a_g, e_g) plane for the potentials (2.50)-(2.53). We see that A^* is a strong function of a_g but a weaker function of e_g . The dependence on e_g emerges predominantly for orbits with $e_g \gtrsim 0.5$; it is rather weak at all e_g for orbits with $a_g \lesssim b$, where b is the scale radius of the potential in question. The difference in radial A^* behavior between different potentials is most pronounced for orbits with $a_g \lesssim b$. In this region there is a sharp increase in A^* in the uncored (Hernquist and NFW) potentials, but a much shallower gradient in the cored potentials (isochrone and Plummer).

We can make the comparison more quantitative by examining the radial profile of A^* for circular outer orbits (of radius a_g and eccentricity $e_g = 0$). Then A^* is a function of a_g/b only, and is plotted in Figure 2.4. We see that $A^*(a_g/b)$ becomes significantly larger than 1 for $a_g \ll b$ in the case of non-cored potentials, but reaches a maximum of 0.5 in the isochrone case. For those potentials with finite total mass \mathcal{M} we can construct the density-weighted average

$$\langle A^* \rangle_\rho = \frac{1}{\mathcal{M}} \int_0^\infty 4\pi r^2 \rho(r) A^*(r) dr, \quad (2.55)$$

still assuming a circular outer orbit. We find $\langle A^* \rangle_\rho = 0.0617, 0.4234$ and 0.65 for isochrone, Plummer and Hernquist potentials respectively. The isochrone model has by far the smallest $\langle A^* \rangle_\rho$.

Behavior of Γ

Figure 2.5 shows Γ in the (a_g, e_g) plane for the same four potentials. We see that $\Gamma \rightarrow 0$ for $a_g \ll b$ in cored potentials, because in the cluster core the potential is close to the spherical harmonic potential (for which Γ is effectively zero, see §2.3.1 and §2.6.1). For the coreless potentials Γ is always positive, as expected. We see that for $a_g \gtrsim b$, the value of Γ is quite sensitive to the outer orbit eccentricity e_g in all four potentials. At

fixed a_g , increasing e_g corresponds to a decrease in Γ . For the cored potentials this is because high- e_g orbits start probing the cluster core where again the potential is roughly spherical harmonic, which tends to lower Γ .

Meanwhile, increasing a_g at fixed e_g tends to increase Γ . At large a_g all finite mass potentials reduce to a Keplerian potential for which $\Gamma = 1$. In the NFW potential, the Γ profile is shallow because the potential decays slowly, namely as $\Phi_{\text{NFW}} \sim r^{-1} \ln(r/b)$ for $r \gg b$. Hence it never becomes sufficiently Keplerian to reach $\Gamma \sim 1$.

To better illustrate this convergence at large a_g/b , in Figure 2.6 we show $\Gamma(a_g/b)$ for circular outer orbits ($e_g = 0$) in the same four potentials as in Figure 2.5. We see that the $\Gamma \rightarrow 1$ convergence does occur for all potentials that are asymptotically Keplerian as $r \rightarrow \infty$, although in some cases one has to go to radii $a_g \gtrsim 50b$ to observe it satisfactorily. Additionally, at very small radii the NFW density profile can be approximated as a power-law cusp, $\rho \propto r^{-1}$, see equation (2.54). Using the result listed in §2.6.1 (and derived in Appendix 2.D.1) we expect to find $\Gamma = 1/9$ as $a_g \rightarrow 0$, and indeed this is reflected in Figure 2.6.

We note that some of the orbits in the potentials (2.50)-(2.53) will have commensurable (or almost commensurable) radial and azimuthal frequencies. For these orbits, i.e. at some points in (a_g, e_g) space, equation (2.39) is not valid, because its derivation relies upon orbits densely filling an axisymmetric annulus, see §2.4.2. This is particularly true of potentials with a core at small a_g , where the potential is close to harmonic (c.f. §2.7). Nevertheless, Figures 2.3 and 2.5 give a good idea of how A and Γ change as we consider different orbits within the cluster.

2.6.3 Behavior of Hamiltonian characteristics in axisymmetric potentials

For axisymmetric potentials it is difficult to make rigorous mathematical statements about A and Γ . In Appendix 2.D.2 we show that in this case, in principle, Γ can take any value $\in (-\infty, \infty)$; however, to achieve extreme negative values of Γ , or $\Gamma > 1$, may require very unusual orbits. (Two such examples are given in Appendix 2.D.2). Meanwhile, in this section we focus on the most typical orbits in axisymmetric potentials via numerical examples. The A^* and Γ values in this section are calculated numerically using the procedure outlined in Appendix 2.F, and are therefore denoted $A_{\text{num}}^*, \Gamma_{\text{num}}$.

We will use two axisymmetric potentials in our numerical examples. The first is the flattened power-law potential (Evans 1994):

$$\Phi_{\text{FPL}}(R, Z) = -\Phi_0 \frac{b^\beta}{(R^2 + (Z/q)^2 + b^2)^{\beta/2}}, \quad (2.56)$$

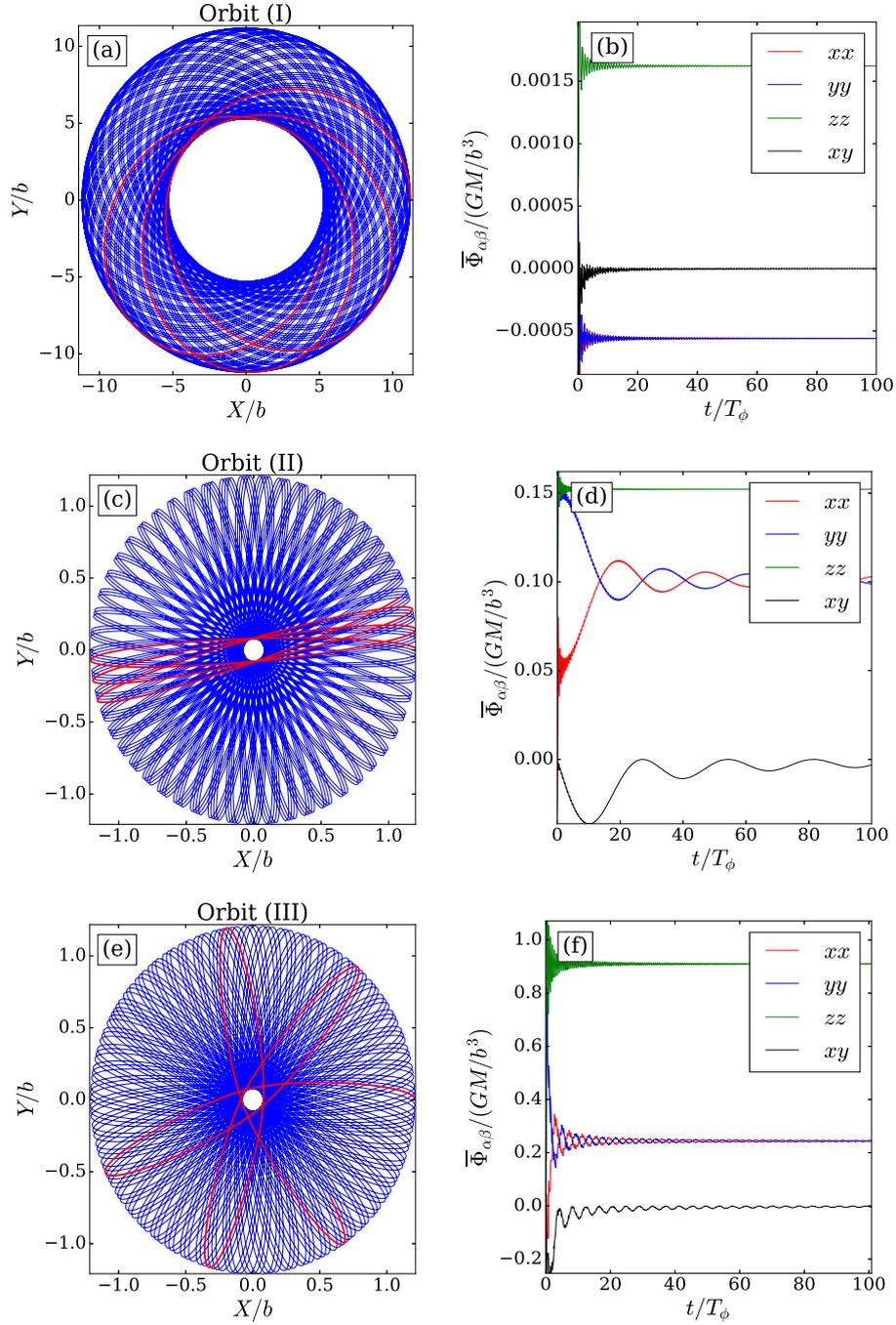


Figure 2.7: Orbits in spherical potentials (see Table 2.1) used for demonstrating convergence to the ‘axisymmetric annulus’ approximation. Left panels show the outer orbit \mathbf{R}_g in the (X, Y) plane, while right panels demonstrate the convergence of time-averaged coefficients $\bar{\Phi}_{\alpha\beta}$ (insets illustrate the color scheme for each $\alpha\beta$ coordinate pair). Orbits (I) and (II) were integrated in the spherical isochrone potential (2.50), while Orbit (III) is in the Hernquist potential (2.52). All panels show 100 azimuthal periods’ worth of data. In the left panels we highlight the first few azimuthal periods of the orbit in red. Convergence is much slower for Orbit (II) because it spends most of its time in a constant density core, so radial and azimuthal frequencies are almost commensurable, leaving large unfilled gaps in the annulus even after $100T_\phi$.

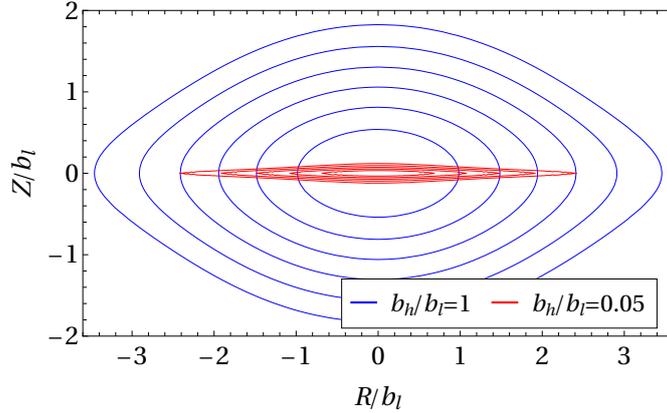


Figure 2.8: Contours of constant $\log_{10}(\rho_{\text{MN}}/\rho_{\text{MN},0})$ where $\rho_{\text{MN}} \equiv \nabla^2\Phi_{\text{MN}}/(4\pi G)$ is the density distribution corresponding to the Miyamoto-Nagai potential (2.57), and $\rho_{\text{MN},0}$ is the central density at $R = Z = 0$. We show the cases $b_h/b_\ell = 1$ (blue) and $b_h/b_\ell = 0.05$ (red). In both cases, contours are spaced linearly from -1.5 to 0 .

where $-\Phi_0$ is the central potential, b is a core radius and q is the oblateness parameter: $q < 1$ corresponds to an oblate potential which can be used to model elliptical galaxies and galactic bulges (Evans 1994). The natural definition of A^* in this case is $A^* \equiv A/(|\Phi_0|/b^2)$. We choose $\beta = 1/2$ and $q = 0.94$, meaning that this potential is only slightly flattened. One can derive a number of useful analytical results in such weakly non-spherical potentials; we defer this investigation to a future study. Here we simply demonstrate that even in the case of a weakly flattened potential large departures from the behaviour typical for purely spherical potentials described in §2.6.2 become possible.

The other potential we will use is the Miyamoto-Nagai potential (Miyamoto & Nagai 1975):

$$\Phi_{\text{MN}}(R, Z) = -\frac{GM}{\sqrt{R^2 + \left(b_\ell + \sqrt{Z^2 + b_h^2}\right)^2}}, \quad (2.57)$$

where b_ℓ is the scale length and b_h is the scale height. As one changes the value of b_h/b_ℓ , the Miyamoto-Nagai potential smoothly transitions from the Kuzmin potential of a razor thin disk ($b_h \ll b_\ell$) to the spherical Plummer potential frequently used to model globular clusters ($b_h \gg b_\ell$) (Binney & Tremaine 2008). In Figure 2.8 we plot contours of constant density $\rho_{\text{MN}} \equiv \nabla^2\Phi_{\text{MN}}/(4\pi G)$ in the (R, Z) plane for two Miyamoto-Nagai models used in this Chapter, namely $b_h/b_\ell = 1$ and $b_h/b_\ell = 0.05$. The natural definition of A^* in this potential is $A^* \equiv A/(GM/b_\ell^3)$.

Orbits in the midplane of an axisymmetric potential

The simplest non-spherical case to consider is when the binary's outer orbit \mathbf{R}_g is confined to the (X, Y) midplane of an axisymmetric potential. Then \mathbf{R}_g still traces a planar rosette with a fixed peri/apocentre (r_p, r_a) just as in the spherical case, so we can easily compute

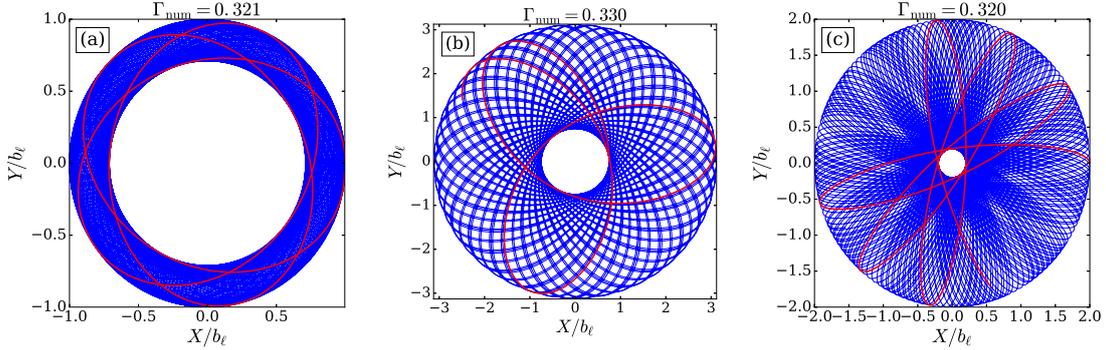


Figure 2.9: Examples of orbits in the midplane of the thin disk represented by the Miyamoto-Nagai potential (2.57) with $b_h/b_\ell = 0.05$. We integrate the orbits for $100T_\phi$ (the first few T_ϕ are highlighted red). The resulting numerically determined Γ_{num} values are all very close to the value $\Gamma = 1/3$ that was predicted simply on the basis of the disk being very thin, so that the vertical curvature of the potential is much larger than the radial curvature. See §2.6.3 for details.

A and Γ as in §2.4.3. In Appendix 2.C we show how to compute A, B, Γ in the case of epicyclic outer orbits in a disk-like potential. We find $A = B = \nu^2$, where ν is the vertical epicyclic frequency at the guiding radius; therefore $\Gamma = 1/3$.

In fact, we already deduced in §2.6.1 that $\Gamma = 1/3$ will hold for *any* orbit \mathbf{R}_g which is confined to the plane of a very thin axisymmetric disk. This follows from the fact that the curvature of the potential is by far greatest in the Z direction at any given position in the disk, so that $\Phi_{zz}(R, 0) \gg \Phi_{xx}(R, 0) \forall R$. Then $A \approx B \approx \overline{\Phi_{zz}}$ and so $\Gamma \approx 1/3$. In Figure 2.9 we confirm this prediction using three very different orbits in the (X, Y) midplane of a thin ($b_h/b_\ell = 0.05$) Miyamoto-Nagai potential. The Γ_{num} values are (a) 0.321, (b) 0.330 and (c) 0.320, all very close to $\Gamma = 1/3$.

Orbits that are far from coplanar

As we show in Appendix 2.D, we always have $\Gamma \geq 0$ in realistic, finite-mass spherical potentials. For Γ to fall below zero the potential must be non-spherical, but also, according to definitions (2.41), the outer orbit of the binary must have $|\overline{\Phi_{xx}}| > |\overline{\Phi_{zz}}|$. Qualitatively, this implies that the average ‘radial curvature’ of the potential over the orbit needs to be greater than the average ‘vertical curvature’. This is not going to be the case while the orbit is confined near a single plane, as we have just seen. However, this situation is naturally realised in potentials that are highly prolate in the Z direction (asymptotically ‘cylindrical’, with $\Phi(R, Z) = \Phi(R)$). In such potentials $\Gamma \approx -1/3$ (see §2.6.1). Also, to probe the negative Γ regime we can consider orbits in non-spherical potentials that make large excursions ‘out of the plane’, i.e. in the Z direction.

This is demonstrated in Figure 2.10, in which we plot four Orbits (‘IV’-‘VII’) in the flattened power-law potential (2.56) with $q = 0.94$ and $\beta = 1/2$. These four Orbits are

Table 2.1: Details of the outer orbits \mathbf{R}_g used for numerical verification of the ‘axisymmetric annulus’ approximation in §2.6.2, §2.7.1 and Figure 2.7. An orbit in a spherical potential (with mass \mathcal{M} and scale radius b) is uniquely specified by its peri/apocentre distances $r_{p/a}$, or equivalently by its generalised semi-major axis a_g and eccentricity e_g . We also provide the orbit’s azimuthal period T_ϕ around the cluster, its analytical A^* and Γ values, and the corresponding numerically computed values $A_{\text{num}}^*, \Gamma_{\text{num}}$, all to 3 significant figures.

| Orbit of \mathbf{R}_g | Potential | $(r_p/b, r_a/b)$ | a_g/b | e_g | $T_\phi \sqrt{\frac{GM}{b^3}}$ | A^* | A_{num}^* | Γ | Γ_{num} |
|-------------------------|----------------------|------------------|---------|-------|--------------------------------|---------|--------------------|----------|-----------------------|
| (I) | Φ_{iso} | (5.29, 11.2) | 8.2 | 0.36 | 171 | 0.00106 | 0.00106 | 0.685 | 0.686 |
| (II) | Φ_{iso} | (0.08, 1.21) | 0.65 | 0.88 | 18.0 | 0.253 | 0.255 | 0.0676 | 0.0650 |
| (III) | Φ_{Hern} | (0.08, 1.21) | 0.65 | 0.88 | 11.0 | 1.15 | 1.15 | 0.192 | 0.195 |

Table 2.2: Properties of Orbits (IV)-(VII) integrated in the flattened power-law potential (2.56) with $\beta = 1/2$ and $q = 0.94$ (c.f. Figure 2.10). The initial conditions of Orbits (IV)-(VII) are identical except for the initial azimuthal velocity v_ϕ . We take initial $(R, v_R, \phi, Z, v_Z) = (b, 0.1\sqrt{GM/b}, 0.1, 0.2b, 0.5\sqrt{GM/b})$, and initial v_ϕ is given below.

| Orbit | Potential | $v_\phi/\sqrt{GM/b}$ | A_{num}^* | Γ_{num} |
|-------|---------------------|----------------------|--------------------|-----------------------|
| (IV) | Φ_{FPL} | 1.35 | 0.0332 | 0.243 |
| (V) | Φ_{FPL} | 0.95 | 0.182 | 0.192 |
| (VI) | Φ_{FPL} | 0.35 | 0.552 | 0.016 |
| (VII) | Φ_{FPL} | 0.05 | 0.626 | -0.085 |

initiated with exactly the same initial conditions except for their initial azimuthal velocity v_ϕ ; the full details of the initial conditions, as well as the resulting A_{num}^* and Γ_{num} values, are given in Table 2.2. In the top row of Figure 2.10 we have Orbit (IV), with initial $v_\phi = 1.35\sqrt{GM/b}$. Orbit (IV) is certainly not planar but the typical excursions in Z are fairly small compared to the excursions in R . As a result the $\Gamma_{\text{num}} = 0.243$ value is less than $1/3$ but still significantly greater than zero. As we move down the page we decrease the initial azimuthal velocity each time, so that Orbits (V)-(VII) initially have $v_\phi/\sqrt{GM/b} = 0.95, 0.35, 0.05$ respectively (while keeping all other initial conditions the same). The radial excursions decrease as the initial azimuthal velocity decreases, until they become comparable to the vertical excursions. Eventually Γ moves below zero in Orbit (VII), see Table 2.2. The A^* values grow as we move down the page since the binary samples a stronger potential when it is closer to the origin.

For our final set of examples we compare four more Orbits (‘VIII’-‘XI’) in the Miyamoto-Nagai potential (2.57) with $b_h/b_\ell = 1$, the initial conditions for which are given alongside

Table 2.3: Properties of Orbits (VIII)-(XI) integrated in the Miyamoto-Nagai potential (2.57) with $b_h/b_\ell = 1$ (c.f. Figure 2.11). The initial conditions of Orbits (VIII)-(XI) are identical except for the initial vertical coordinate Z . We take initial $(R, v_R, \phi, v_\phi, v_Z) = (b_\ell, 0, 0, 0.25\sqrt{GM/b_\ell}, 0)$, and initial Z is given below.

| Orbit | Potential | Z/b_ℓ | A_{num}^* | Γ_{num} |
|--------|------------------------------------|------------|--------------------|-----------------------|
| (VIII) | $\Phi_{\text{MN}}, b_h/b_\ell = 1$ | 0.1 | 0.256 | 0.153 |
| (IX) | $\Phi_{\text{MN}}, b_h/b_\ell = 1$ | 1.0 | 0.122 | 0.042 |
| (X) | $\Phi_{\text{MN}}, b_h/b_\ell = 1$ | 2.0 | 0.0392 | -0.163 |
| (XI) | $\Phi_{\text{MN}}, b_h/b_\ell = 1$ | 3.0 | 0.0140 | -0.384 |

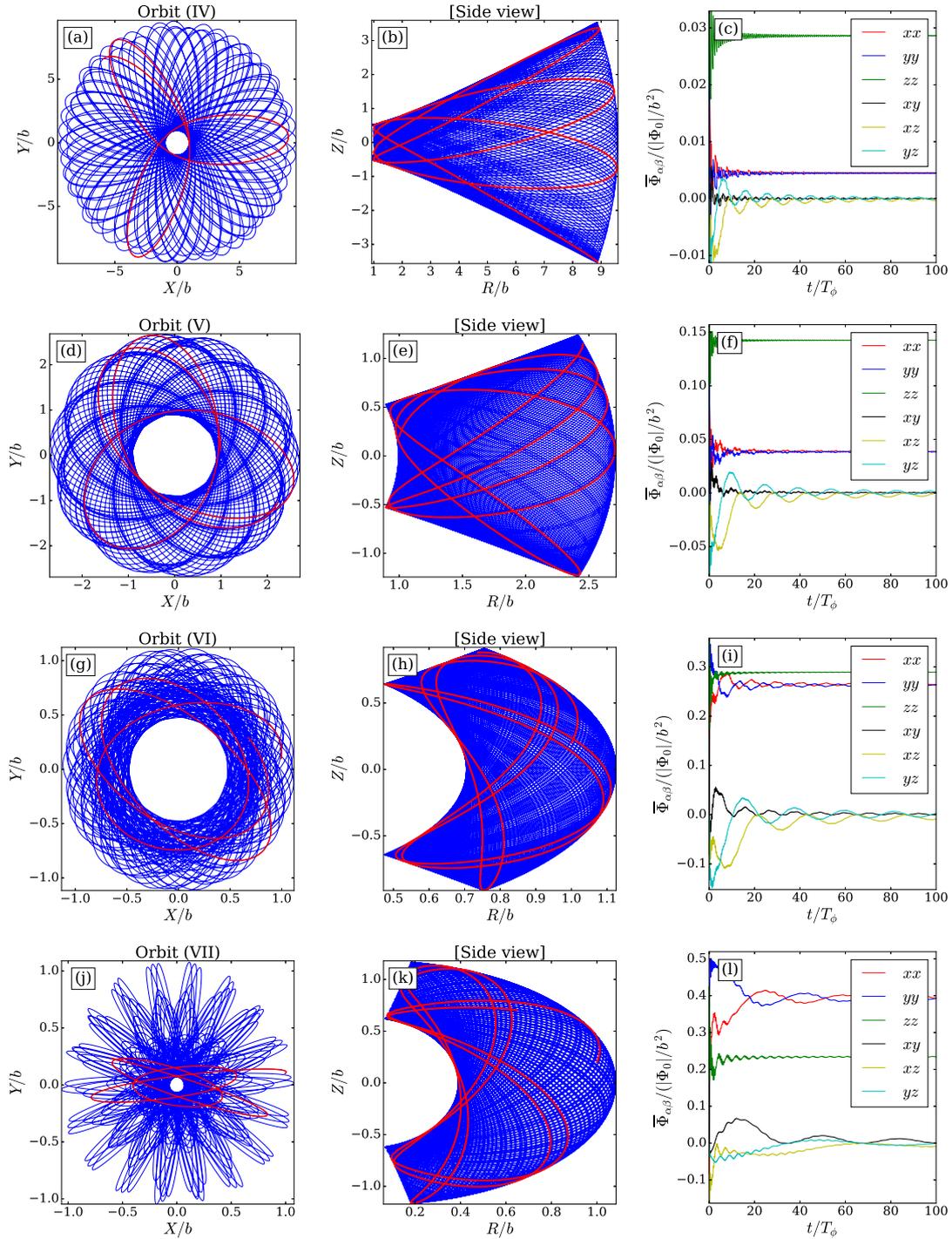


Figure 2.10: Similar to Fig. 2.7, but for Orbits (IV)-(VII) demonstrating convergence to the ‘axisymmetric torus’ approximation in non-spherical potentials. Middle panels show the meridional projection of $\mathbf{R}_g(t)$. Also, the right panels show the convergence of Φ_{xz} and Φ_{yz} , which are not identically zero for non-coplanar orbits (they should vanish only upon outer orbit averaging). These Orbits with properties listed in Table 2.2 were integrated for 100 azimuthal periods in a flattened power law potential (2.56) with $\beta = 1/2$ and $q = 0.94$. Initially, the Orbits differ only in their azimuthal velocity v_ϕ , see Table 2.2. (Note the different axis scales for different orbits).

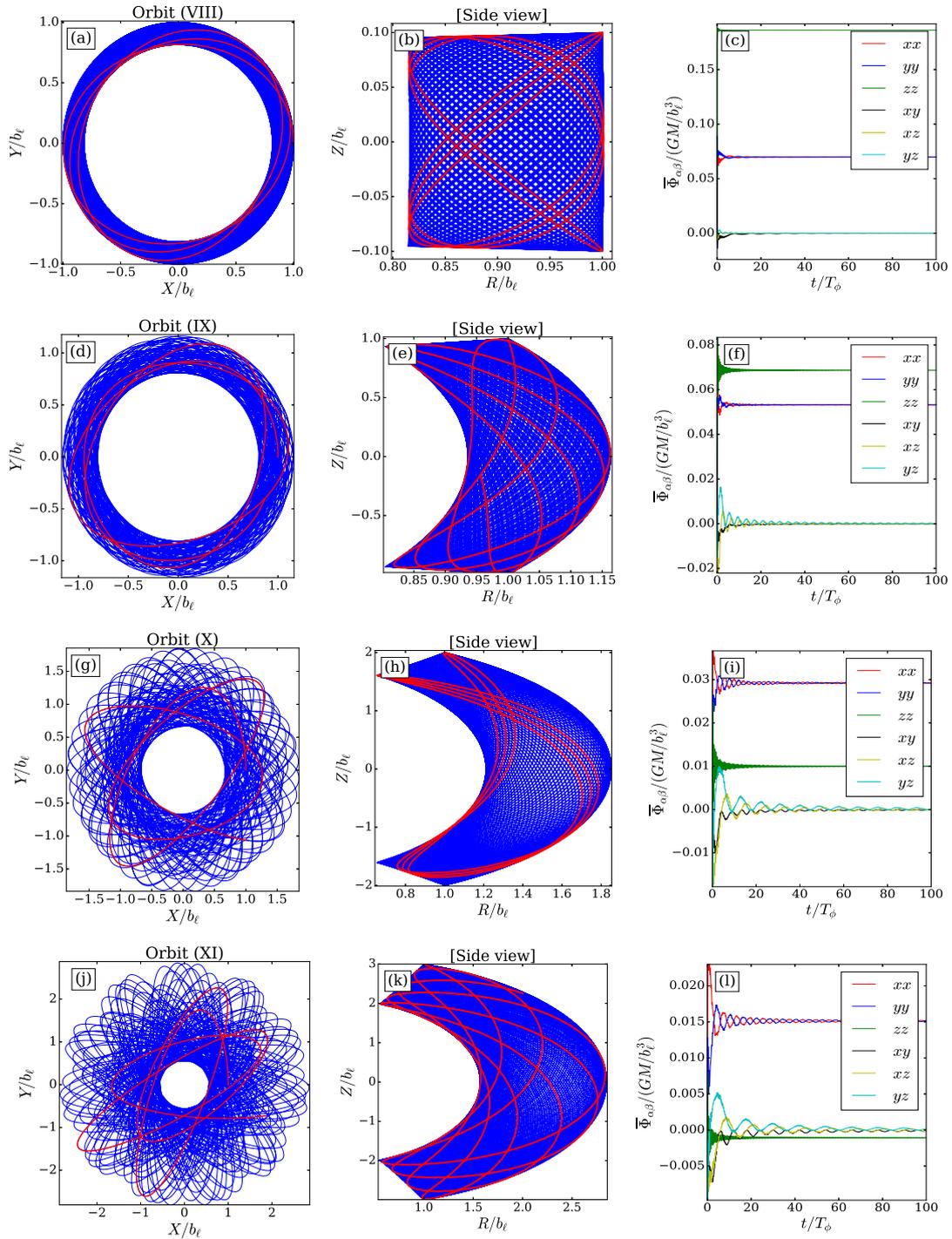


Figure 2.11: Similar to Figure 2.10, but now for Orbits (VIII)-(XI) with characteristics given in Table 2.3 integrated in a Miyamoto-Nagai potential (2.57) with $b_h/b_\ell = 1$. These orbits differ only in their initial vertical coordinate Z , with initial $Z/b_\ell = 0.1, 1.0, 2.0, 3.0$ respectively. All panels show 100 azimuthal periods' worth of data. (Note the drastically changing vertical scale in the central column, particularly panel (b)).

their resulting $A_{\text{num}}^*, \Gamma_{\text{num}}$ values in Table 2.3. At large distances this potential is significantly flatter than the $q = 0.94$ flattened power-law potential. Orbits (VIII)-(XI) are plotted in the left hand column of Figure 2.11. Each Orbit has exactly the same initial conditions except we change the initial vertical coordinate Z , using $Z = 0.1b_\ell, b_\ell, 2b_\ell$ and $3b_\ell$ respectively. Increasing the initial Z thickens the orbit. Orbit (VIII) (top row, initial $Z = 0.1b_\ell$) is almost coplanar and has $\Gamma_{\text{num}} = 0.153$. The Γ_{num} value decreases as we move down the page, reaching a minimum value of $\Gamma_{\text{num}} = -0.384$ for Orbit (XI) (bottom row, $Z = 3b_\ell$) which is thicker vertically than it is radially. Meanwhile, as we move from top to bottom the A_{num}^* value decreases because the orbit spends more time away from the midplane where the tidal potential is strongest. An even more extreme orbit in this potential, with the same initial conditions except $Z = 4b_\ell$ (resulting in $\Gamma_{\text{num}} = -1.4$), is presented in Figure 2.12.

Finally, note that Orbit (VIII) is very similar in appearance to the orbit in Figure 2.9a: both are roughly epicyclic, so they should obey equations (2.75) and (2.76). The difference between them is that in the case of Orbit (VIII) the potential felt by the binary is significantly less flattened, since this Orbit resides predominantly in the quasi-spherical core of the potential ($b_h/b_\ell = 1$). As a result, $\bar{\Phi}_{zz}$ does not dominate over $\bar{\Phi}_{xx}$ and hence Γ ends up being significantly less than $1/3$ (unlike the orbit in Figure 2.9a which explores a much more flattened version of the Miyamoto-Nagai potential with $b_h/b_\ell = 0.05$).

2.7 Validity of secular Hamiltonian

In §2.6 we focused on understanding the typical values of A, Γ for various types of orbit in different potentials. However the Hamiltonian (2.42) is only valid if the symmetry conditions (2.33)-(2.36) for the time-averages $\bar{\Phi}_{\alpha\beta}$ are satisfied. In addition it is reasonable to require that $\bar{\Phi}_{\alpha\beta}$ converge to fixed values (say to within a few percent) on timescales significantly shorter than the timescale for secular evolution (which will be derived in Chapter 3). In this section we check the validity of these assumptions numerically. The procedure for calculating $\bar{\Phi}_{\alpha\beta}$ numerically is given in Appendix 2.F.

2.7.1 Spherical potentials

In a spherical potential, orbits $\mathbf{R}_g(t)$ that have non-commensurable frequencies densely fill an axisymmetric annulus. If this is true then the time-averaged coefficients $\bar{\Phi}_{\alpha\beta}$ should obey the symmetry properties (2.33), (2.35).

To verify this we use Orbits (I)-(III). Their initial conditions are given alongside their $A^*, A_{\text{num}}^*, \Gamma, \Gamma_{\text{num}}$ values in Table 2.1. The right panels in Figure 2.7 show the corresponding running average (from $t = 0$ to current time) of numerically computed

$\bar{\Phi}_{\alpha\beta}$. As the number of completed orbits grows, the time-averaged derivatives of the potential tend to converge towards fixed values.

Orbit (I) has rather large semi-major axis and small eccentricity, so that it stays far from the core at all times, filling its annulus densely. The ‘axisymmetric annulus’ approximation works very well in this case, so the (semi-)analytic and numerically computed values agree: $A^* = A_{\text{num}}^*$ and $\Gamma = \Gamma_{\text{num}}$ to within 1% accuracy.

We have picked rather extreme examples in Orbits (II) and (III) in order to demonstrate behaviour of orbits \mathbf{R}_g that are both very radial and tightly bound near the centre of the cluster. Orbit (II) spends a lot of time in the isochrone potential’s constant density core where its frequencies are almost commensurable ($\Omega_R \approx 2\Omega_\phi$); as a result it precesses slowly, so that there are unfilled gaps left in its annulus even after $t = 100T_\phi$. This issue does not arise in the uncored Hernquist potential, so Orbit (III) fills its annulus more efficiently than Orbit (II). Nevertheless, the axisymmetric approximation is still very successful in both cases, with a maximum discrepancy of $\sim 4\%$ arising between the Γ and Γ_{num} values of Orbit (II).

However, we notice that while the converged symmetry properties of the $\bar{\Phi}_{\alpha\beta}$ (see equations (2.33), (2.35)) are well established after $\gtrsim 15T_\phi$ for Orbits (I) and (III), they are less well established for Orbit (II) even at $t \gtrsim 45T_\phi$. Again this is because Orbit (II) does not fill its annulus efficiently. This can be problematic because if the $\bar{\Phi}_{\alpha\beta}$ fail to converge on a timescale shorter than the secular evolution timescale, the doubly-averaged theory can break down, as we will see in Chapter 3.

2.7.2 Axisymmetric potentials

In a (non-spherical) axisymmetric potential, orbits $\mathbf{R}_g(t)$ that have non-commensurable frequencies densely fill an axisymmetric torus and so the time-averaged coefficients $\bar{\Phi}_{\alpha\beta}$ should obey the symmetry properties (2.33), (2.35), and (2.36). To verify this numerically we use Orbits (IV)-(VII) in the flattened power-law potential (2.56) with $q = 0.94$ and $\beta = 1/2$ (see Table 2.2 and Figure 2.10), and Orbits (VIII)-(XI) in the Miyamoto-Nagai potential (2.57) with $b_h/b_\ell = 1$ (see Table 2.3 and Figure 2.11).

Some features of the $\bar{\Phi}_{\alpha\beta}$ convergence plots are similar to the spherical case. For example, in Figure 2.10 the derivatives $\bar{\Phi}_{\alpha\beta}$ converge rather slowly in the bottom panel because the Orbit (VII) fills its torus rather sparsely, owing to the large fraction of time it spends in the almost-harmonic potential of the core.

Note that the rightmost columns in Figures 2.10 and 2.11 also show the convergence of $\bar{\Phi}_{xz}$ and $\bar{\Phi}_{yz}$. This is different from Figure 2.7 since now we are dealing with non-planar orbits so that these derivatives are no longer identically zero. Although the corresponding time-averages $\bar{\Phi}_{xz}$, $\bar{\Phi}_{yz}$ do indeed converge to zero in all of our axisymmetric examples

as expected, in most cases their convergence takes significantly longer than that of the other $\bar{\Phi}_{\alpha\beta}$ coefficients. This is what we would expect from looking at the ϕ_g dependence of equations (2.26)-(2.31): the derivatives Φ_{xx} , Φ_{yy} and Φ_{xy} fluctuate twice as rapidly with respect to ϕ_g compared to Φ_{xz} , Φ_{yz} . Slower convergence of $\bar{\Phi}_{xz}$ and $\bar{\Phi}_{yz}$ seems to be especially apparent for strongly non-coplanar orbits, i.e. orbits which make large excursions in the Z direction. Orbits that inefficiently fill their torus (i.e. on timescales longer than the secular evolution timescale) can render the doubly-averaged theory inaccurate, as discussed in detail in §3.7.

2.8 Discussion

In deriving the secular Hamiltonian (equations (2.42), (2.43)) we relied on several approximations. First, we assumed that the outer orbit-averaging procedure used for computing potential derivatives $\bar{\Phi}_{\alpha\beta}$ converges rapidly when compared to the timescale for secular evolution (to be derived in §3.2.6). The rate of convergence of various $\bar{\Phi}_{\alpha\beta}$ components was explored in §2.7. In Chapter 3 we will use direct numerical integrations of binaries orbiting in stellar cluster potentials to study how well this double averaging procedure works in practice.

Second, we truncated our expansion of the tidal Hamiltonian in §2.2.1 at the quadrupole order. However, studies of the LK mechanism have shown the importance of higher order — ‘octupole’ — terms for the dynamics of triples in certain situations, particularly when the outer orbit is highly eccentric (Lithwick & Naoz 2011; Li et al. 2014). This raises a question of whether octupole terms can be important for the secular dynamics of binaries in external tidal fields. While we derive octupole-level corrections to our doubly-averaged Hamiltonian in Appendix 2.E, in practice they are unlikely to be important for our purposes. This is because in realistic situations the ratio of the semi-major axis of the inner binary orbit ($a \lesssim 100$ AU) is much smaller than the size of its outer orbit ($|\mathbf{R}_g| \sim 1$ pc, comparable to the cluster size), rendering the timescale on which octupole-level effects may manifest themselves too long (we will see in Chapter 3 that a characteristic timescale of secular evolution driven by quadrupole terms in a typical globular cluster is at least tens of Myrs)³.

Third, our calculation assumes a spatially smooth and time-invariant tidal potential. This approximation neglects the granularity and stochastic variability of the cluster potential caused by encounters with other stars, which are very important in dense environments of clusters (Heggie 1975; Hut & Bahcall 1983; Hut 1983; Heggie & Rasio

³On the other hand, octupole terms could have some relevance at high outer orbital eccentricities, where the pericentre distance of the binary’s outer orbit is much smaller than 1pc. Additionally, octupole effects might accumulate to a significant level on timescales of many tens or hundreds of secular periods, which could be relevant for cluster tide-driven mergers (Chapter 5).

1996; Collins & Sari 2008). The cumulative effect of a large number of such encounters is what eventually contributes to the smooth tidal field of the cluster (Collins & Sari 2010); thus, one hopes that in the long run our framework should provide a qualitatively accurate picture of binary evolution in clusters. We discuss stellar encounters further in §3.9.3 and §8.2.1.

An effect that can modify the binary’s *outer* orbit in a stellar cluster is resonant relaxation. Rauch & Tremaine (1996) showed that in quasi-Keplerian systems (such as nuclear clusters dominated by a central super-massive black hole) angular momentum is efficiently exchanged between stellar orbits that have commensurable frequencies. When applied to a binary in a quasi-Keplerian cluster, precession of the binary’s outer orbit due to resonant interactions with other stars (so-called ‘vector resonant relaxation’) can alter its inclination relative to the inner orbital plane, potentially bringing an initially low-inclination binary into a high-inclination regime and triggering LK oscillations (Hamers, Bar-Or, et al. 2018b). While this effect has not been explored for binaries in non-Keplerian potentials, vector resonant relaxation can indeed operate in non-Keplerian systems such as globular clusters (Meiron & Kocsis 2019).

Additionally, the fact that a binary is typically heavier than the average star in a cluster means that it will tend to sink towards the centre of the cluster via dynamical friction (Binney & Tremaine 2008). Moreover, the global properties of the cluster itself may evolve as a result of two-body relaxation leading to core collapse. All of these effects can change the values of A and Γ for a given binary over long time intervals. We defer an exploration of their impact upon binary evolution to future work (see also §8.2).

Finally, an important assumption that lies at the foundation of our time-averaging procedure is that different frequencies characterizing binary motion in the cluster are not commensurable with each other (see §2.4.2). If this condition is violated, the outer orbit \mathbf{R}_g no longer fills an axisymmetric torus inside the cluster uniformly in azimuth, rendering the equations (2.33)-(2.36) formally invalid. This issue is addressed in more detail next.

2.8.1 Commensurable frequencies

Orbits in realistic spherical potentials obey the following relation between the radial (Ω_R) and azimuthal (Ω_ϕ) frequencies (Binney & Tremaine 2008):

$$1/2 \leq \Omega_\phi/\Omega_R \leq 1. \quad (2.58)$$

Thus, in spherical potentials there are infinitely many rational values of Ω_ϕ/Ω_R in the interval $(1/2, 1)$. As a result, a small number of resonant points in the (a_g, e_g) plane will have $\Omega_\phi/\Omega_R = p/q$ with integer p, q . Strictly speaking, binaries on these resonant outer orbits will not satisfy the axisymmetric (‘filled annulus’) averaging approximation. On

the other hand we expect that these orbits should still be described (at least roughly) by the filled annulus approximation as long as the integers $p, q \gg 1$. In fact, numerical experiments in spherical potentials (not shown here) suggest that even for binaries with Ω_ϕ/Ω_R very close to $2/3, 3/4$, etc., good convergence to this approximation occurs within several outer orbital periods. One has to choose highly pathological initial conditions in order to get this approximation to fail.

The most important cases of outer orbits with commensurable frequencies are those in the harmonic ($\Omega_\phi/\Omega_R \equiv 1/2$) and Keplerian ($\Omega_\phi/\Omega_R \equiv 1$) potentials. Even here we can continue to describe the dynamics straightforwardly: if the potential is purely harmonic we can treat it as in §2.3.1, and we show in Appendix 2.B that Keplerian potentials are described by our doubly-averaged formalism with $\Gamma = 1$. However, it should be stressed that we have *not* used the axisymmetric averaging approximation in either of these cases: the harmonic potential just happens to be effectively axisymmetric after single-averaging, and the Keplerian potential is known to be axisymmetric under double averaging to quadrupole order. In neither case do orbits ‘fill their annulus’⁴.

In practice the most problematic case is when the binary experiences a potential that is *almost* harmonic or *almost* Keplerian, so that the outer orbit precesses apsidally, but not quickly enough to fill a circular annulus on a secular timescale and thereby qualify for an axisymmetric treatment (see also Bub & Petrovich 2020). For example, orbits that spend a lot of time in a constant-density core of the cluster potential experience an *almost* harmonic potential and so tend to fill their annulus very slowly (Figure 2.7c).

2.8.2 Relation to previous work

Many previous studies have looked at secular evolution of binaries perturbed by external potentials. The effect of an arbitrary quadrupole perturbation upon a binary has been briefly considered by Mikkola & Nurmi (2006). In particular, their equation (20) gives the quadrupole potential experienced by a binary in a star cluster consisting of a large number of point masses m_k . Our perturbing Hamiltonian H_1 is recovered from their result in the mean-field limit (i.e. by replacing the exact potential of the cluster, $-G \sum_k m_k / |\mathbf{R} - \mathbf{R}_k|$, with the smooth potential $\Phi(\mathbf{R})$). However, Mikkola & Nurmi (2006) did not explicitly convert to orbital elements, perform any averaging, or develop any secular theory as we do here. In a similar vein, a short paper by Katz & Dong (2011) considered the secular dynamics of a binary perturbed by a generic quadratic potential and included axisymmetric potentials as a special case. They *did* convert to orbital elements but did not go much further; in particular they did not provide any prescription for computing the coefficients of the averaged perturbing potential.

⁴The Keplerian potential is not axisymmetric to octupole order — see Appendix 2.E.

Studies of tidal effects of the Galactic disk on wide binaries (Heisler & Tremaine 1986; Byl 1986; Yabushita 1989) represent an important limit ($\Gamma \rightarrow 1/3$) of our general theory, see §2.5.2 and Appendix 2.C. Since Heisler & Tremaine (1986), Galactic tides have been included in many studies of cometary orbits (e.g. Matese & Whitman 1989; Matese & Whitmire 1996; Breiter, Dybczynski, et al. 1996; Wiegert & Tremaine 1999; Brassier 2001; Fouchard 2004; Fouchard, Froeschlé, Matese, et al. 2005; Fouchard, Froeschlé, Valsecchi, et al. 2006; Breiter, Fouchard, et al. 2007), as well as planetesimal orbits (e.g. Higuchi et al. 2007).

Veras & Evans (2013c) considered a very general form of the perturbed two-body problem, allowing for both position- and velocity-dependent tidal forces to act upon the binary. Their equations (25)-(29) are more general versions of our singly-averaged equations (c.f. our singly-averaged Hamiltonian (2.19)), and our equations are recovered if one sets the velocity-dependent forces to zero. However they did not derive any analogues of our doubly-averaged equations. Veras & Evans (2013b) noted that Galactic forces may impact the evolution of exoplanetary systems around stars near the bulge of the Galaxy where the Galactic tide is much stronger than it is in the Solar neighbourhood.

Another interesting and obvious limit of our theory, $\Gamma = 1$ — which is, however, rather distinct, see §2.5.1 and Appendix 2.B — has been explored in numerous studies of Lidov-Kozai dynamics (Lidov 1962; Kozai 1962; Fabrycky & Tremaine 2007; Naoz 2016) and its extensions. One interesting extension to the LK problem was made by Petrovich & Antonini (2017), who considered the effect of an axisymmetric (non-spherical) nuclear cluster potential on compact object binaries that are themselves orbiting a central super-massive black hole (SMBH). The non-spherical part of the cluster potential was considered to drive nodal precession of the binary’s quasi-Keplerian outer orbit around the SMBH (continuously changing the relative inclination in the triple system composed of the binary and SMBH, which is important for the operation of LK cycles in this sub-system), while the dominant spherical part drove apsidal precession of the outer orbit. Our doubly-averaged formalism covers this problem in the case where the characteristic timescales for nodal and apsidal precession of the outer orbit are much shorter than the secular timescale, so that the outer orbit fills its torus. Our singly-averaged equations cover it in all cases. However, unlike Petrovich & Antonini (2017), we also account for the *direct* effect of the tidal torque due to the potential of the cluster on the orbital elements of the inner orbit of the binary.

Several authors have considered the problem of a star in orbit around a SMBH in a nuclear cluster (e.g. Sridhar & Touma 1999; Ivanov et al. 2005; Löckmann et al. 2008; Šubr et al. 2009; Chang 2009; Haas et al. 2011; Merritt 2013; Li et al. 2015; Iwasa & Seto 2016; Iwasa & Seto 2017). The SMBH-star system effectively forms a binary. The binary’s

Keplerian orbital elements may then evolve on secular timescales due to some combination of (i) the mean field nuclear cluster potential, (ii) GR pericentre precession, (iii) an infalling massive black hole on a slowly decaying circular orbit, (iv) a circumnuclear ring of material, etc. While this class of problems is reminiscent of our work, it is not quite the same because in the SMBH-star case the barycentre of the binary does not move, and so there is no clean separation between single- and double-averaging. In some cases, e.g. for a binary that sits at the centre of a spherical cusp, there is not even a well-defined tidal expansion of the potential. Instead, averaging of the potential is incorporated into the averaging over the stellar Keplerian orbit around the SMBH, which is different from our approach.

One of the most interesting recent applications of secular dynamics has been the possibility of substantial shrinking of binary orbits by LK cycles with dissipative effects. Such applications include the origin of hot Jupiters (Fabrycky & Tremaine 2007; Naoz, Farr, et al. 2011; Petrovich 2015; Hamers 2017), formation of blue stragglers (Perets & Fabrycky 2009; Knigge et al. 2009), white-dwarf mergers (Thompson 2011; Katz, Dong, & Malhotra 2011; Toonen et al. 2018), and compact object binary mergers in globular or nuclear star clusters (Antognini et al. 2014; Rodriguez, Morscher, et al. 2015; Naoz 2016; Silsbee & Tremaine 2017; Petrovich & Antonini 2017; Leigh et al. 2018). Binary evolution driven by cluster tides explored in our work represents a different evolutionary scenario that may lead to similar outcomes (without invoking a nearby third companion). The possibility that this mechanism can lead to compact object mergers is studied in Chapters 5 and 7.

2.9 Summary

In this Chapter we explored secular evolution of binary systems orbiting in axisymmetric stellar clusters. We derived a Hamiltonian describing this evolution for an arbitrary form of the smooth cluster potential, averaged it over the (inner) orbital motion of the binary, and then averaged it again over the (outer) orbit of the binary around the cluster assuming the potential is axisymmetric. Our results can be summarized as follows.

- When the doubly-averaged Hamiltonian is cast in dimensionless form, all the information about the tidal potential is contained in a single dimensionless parameter Γ , which depends on the background potential Φ and the orbit of the binary in the cluster. The value of this parameter determines the phase portrait of the binary evolution, to be explored in Chapter 3.
- The timescale of secular evolution is set by another (dimensional) parameter A , which, like Γ , depends on the cluster potential and the binary's outer orbit.

- In certain cases A and Γ can be calculated (semi-)analytically. Such cases include (a) orbits in spherical potentials, (b) orbits confined to the midplane of an axisymmetric potential, and (c) epicyclic orbits near the midplane of an axisymmetric potential. We demonstrate how our calculations reproduce the known results for Lidov-Kozai evolution, evolution of Oort Cloud comets due to the Galactic tide, and so on.
- We map out the behavior of A and Γ in different spherically symmetric potentials as a function of size and radial elongation of the binary orbit. We find that Γ is small in the central regions of clusters with cored potentials, but tends to unity in clusters with finite mass as the orbit size increases. In general, $0 \leq \Gamma \leq 1$ in realistic finite-mass spherical potentials.
- In general axisymmetric potentials, Γ can easily attain negative values, in particular for highly inclined (i.e. non-coplanar) orbits.
- The accuracy with which our doubly-averaged Hamiltonian characterizes binary evolution deteriorates for highly non-coplanar orbits in axisymmetric potentials. Commensurability of orbital frequencies in the cluster potential may also present a problem for application of our theory at a quantitative level.

These results will be used extensively in Chapter 3, where we systematically explore the dynamics of binaries driven by the tidal field of clusters for different values of Γ . There we verify numerically the predictions of the secular theory based on our doubly-averaged Hamiltonian, and derive timescales for secular eccentricity oscillations.

Appendices

2.A The singly-averaged Hamiltonian in orbital elements

In this Appendix we give the full algebraic expressions in terms of orbital elements for the terms $\langle r_\alpha r_\beta \rangle_M$ that enter the singly-averaged Hamiltonian (2.19). They are

$$\begin{aligned} \langle x^2 \rangle_M = \frac{a^2}{16} & \left(\cos 2i(2 + 3e^2 + 5e^2 \cos 2\omega \cos 2\Omega) + 5e^2 \cos 2\omega(3 \cos 2\Omega + 2 \sin^2 i) \right. \\ & \left. + (2 + 3e^2)(3 + 2 \cos 2\Omega \sin^2 i) - 20e^2 \cos i \sin 2\omega \sin 2\Omega \right), \end{aligned} \quad (2.59)$$

$$\begin{aligned} \langle y^2 \rangle_M = \frac{a^2}{16} & \left(\cos 2i(2 + 3e^2 - 10e^2 \cos 2\omega \cos^2 \Omega) + 5e^2 \cos 2\omega(1 - 3 \cos 2\Omega) \right. \\ & \left. + (2 + 3e^2)(3 - 2 \cos 2\Omega \sin^2 i) + 20e^2 \cos i \sin 2\omega \sin 2\Omega \right), \end{aligned} \quad (2.60)$$

$$\langle z^2 \rangle_M = \frac{a^2}{4} \sin^2 i \left(2 + 3e^2 - 5e^2 \cos 2\omega \right), \quad (2.61)$$

$$\begin{aligned} \langle xy \rangle_M = \frac{a^2}{16} & \left(20e^2 \cos i \cos 2\Omega \sin 2\omega \right. \\ & \left. + (5e^2(3 + \cos 2i) \cos 2\omega + 2(2 + 3e^2) \sin^2 i) \sin 2\Omega \right), \end{aligned} \quad (2.62)$$

$$\langle xz \rangle_M = \frac{a^2}{4} \sin i \left(5e^2 \cos \Omega \sin 2\omega - \cos i(2 + 3e^2 - 5e^2 \cos 2\omega) \sin \Omega \right), \quad (2.63)$$

$$\langle yz \rangle_M = \frac{a^2}{4} \sin i \left(5e^2 \sin \Omega \sin 2\omega + \cos i(2 + 3e^2 - 5e^2 \cos 2\omega) \cos \Omega \right). \quad (2.64)$$

2.B Recovering the Lidov-Kozai quadrupole Hamiltonian

To derive the LK Hamiltonian we take equation (2.19) and average it over time using $\Phi(r) = -GM/r$, and with $\mathbf{R}_g(t)$ describing a Keplerian ellipse with the focus at the origin.

First, it is obvious that for this potential, $\Phi''(R_g) \propto \Phi'(R_g)/R_g \propto R_g^{-3}$. Then we must average the right hand sides of equations (2.26)-(2.29) which requires that we average the quantities R_g^{-3} , $R_g^{-3} \cos 2\phi_g$ and $R_g^{-3} \sin 2\phi_g$.

Without loss of generality we may choose the argument of pericentre ω_g of the ellipse to be zero, so that $\phi_g = f_g$, the true anomaly. Then $R_g = a_g(1 - e_g \cos E_g)$ with E_g being the eccentric anomaly, so for an arbitrary function $\mathcal{S}(\phi_g)$ we can write

$$\overline{R_g^{-3} \mathcal{S}(\phi_g)} = \frac{1}{2\pi} \int_0^{2\pi} dE_g a_g^{-3} (1 - e_g \cos E_g)^{-2} \mathcal{S}(f_g), \quad (2.65)$$

and we can convert between f_g and E_g using

$$\cos f_g = \frac{\cos E_g - e_g}{1 - e_g \cos E_g}. \quad (2.66)$$

The answers are

$$\overline{R_g^{-3}} = \frac{4}{3} \overline{R_g^{-3} \cos 2\phi_g} = a_g^{-3} (1 - e_g^2)^{-3/2}, \quad (2.67)$$

$$\overline{R_g^{-3} \sin 2\phi_g} = 0. \quad (2.68)$$

Using these identities, we find a remarkably simple relation between the time-averaged coefficients:

$$\overline{\Phi_{xx}} = \overline{\Phi_{yy}} = -\frac{1}{2} \overline{\Phi_{zz}}, \quad \text{with } \overline{\Phi_{zz}} = GMa_g^{-3} (1 - e_g^2)^{-3/2}, \quad (2.69)$$

and as expected $\overline{\Phi_{xy}} = \overline{\Phi_{xz}} = \overline{\Phi_{yz}} = 0$. Note that the regime $\overline{\Phi_{xx}} = \overline{\Phi_{yy}}$ is exactly that of an axisymmetric perturbing potential (see e.g. Katz, Dong, & Malhotra 2011). The resulting perturbing Hamiltonian will therefore be the same as (2.42), with the added simplification that $A = (GM/2)a_g^{-3}(1 - e_g^2)^{-3/2}$ and $\Gamma = 1$. Making these substitutions we find

$$\overline{\langle H_1 \rangle}_M = -\frac{GMa^2}{16a_g^3(1 - e_g^2)^{3/2}} [(2 + 3e^2)(3 \cos^2 i - 1) + 15e^2 \sin^2 i \cos 2\omega]. \quad (2.70)$$

This is precisely the dimensionless test particle quadrupole Lidov-Kozai Hamiltonian (Lidov 1962; Kozai 1962; Kinoshita & Nakai 2007; Lithwick & Naoz 2011; Antognini et al. 2014). It describes the secular evolution of a hierarchical triple system in which the outer orbit dominates the angular momentum budget.

2.C Epicyclic orbits

In this Appendix we look at the behavior of A and Γ in the case of a binary performing epicyclic motion in an axisymmetric disk, to connect with the results of Heisler & Tremaine (1986), who calculated the secular effect of the Galactic tide on the Oort Cloud comets.

Let the guiding centre of the binary's orbit be a circle of radius R_c in the $Z = 0$ plane of the potential. The potential experienced by the binary can then be approximated as⁵

$$\begin{aligned} \Phi(R, Z) = & \Phi(R_c, 0) + \left(\frac{\partial \Phi}{\partial R} \right)_{(R_c, 0)} (R - R_c) \\ & + \frac{1}{2} \left(\frac{\partial^2 \Phi}{\partial R^2} \right)_{(R_c, 0)} (R - R_c)^2 + \frac{1}{2} \left(\frac{\partial^2 \Phi}{\partial Z^2} \right)_{(R_c, 0)} Z^2. \end{aligned} \quad (2.71)$$

⁵We implicitly assume that the disk is symmetric about its midplane, $\Phi(R, Z) = \Phi(R, -Z)$, so that $\partial^2 \Phi / \partial R \partial Z = 0$ at $Z = 0$. Otherwise the binary would not remain in the midplane anyway.

Using this expression and equations (2.33)-(2.34) it is easy to show that

$$\bar{\Phi}_{xx} = \bar{\Phi}_{yy} = \frac{1}{2}(\kappa^2 - 2\Omega_c^2); \quad \bar{\Phi}_{zz} = \nu^2, \quad (2.72)$$

while all other $\bar{\Phi}_{\alpha\beta} = 0$; here

$$\Omega_c^2 = \left(\frac{1}{R} \frac{\partial \Phi}{\partial R} \right)_{(R_c, 0)}, \quad (2.73)$$

is the angular frequency of the guiding centre, while

$$\kappa^2 \equiv \left(\frac{\partial^2 \Phi}{\partial R^2} + \frac{3}{R} \frac{\partial \Phi}{\partial R} \right)_{(R_c, 0)}, \quad \nu^2 \equiv \left(\frac{\partial^2 \Phi}{\partial Z^2} \right)_{(R_c, 0)}, \quad (2.74)$$

are the radial and vertical epicyclic frequencies of \mathbf{R}_g respectively. Hence

$$A = \nu^2 + \frac{1}{2}(\kappa^2 - 2\Omega_c^2), \quad (2.75)$$

$$\Gamma = \frac{\nu^2 - \frac{1}{2}(\kappa^2 - 2\Omega_c^2)}{3[\nu^2 + \frac{1}{2}(\kappa^2 - 2\Omega_c^2)]}. \quad (2.76)$$

Near the midplane of a galactic disk, and in particular in the Heisler & Tremaine (1986) case of the solar neighbourhood of the Milky Way, it is almost always the case that $\Omega_c \sim \kappa \ll \nu$. Thus to a very good approximation $A = \nu^2$ and $\Gamma = 1/3$. Plugging these results into our doubly-averaged Hamiltonian (2.42), written in Delaunay variables, we find

$$\begin{aligned} \overline{\langle H_1 \rangle}_M &= \frac{\nu^2 L^2}{4\mu^2 J^2} (J^2 - J_z^2) [J^2 + 5(L^2 - J^2) \sin^2 \omega] \\ &= \frac{\pi G \rho_0 L^2}{\mu^2 J^2} (J^2 - J_z^2) [J^2 + 5(L^2 - J^2) \sin^2 \omega], \end{aligned} \quad (2.77)$$

where we have eliminated ν in favour of the density in the Solar neighbourhood ρ_0 using Poisson's equation $4\pi G \rho_0 = (\nabla^2 \Phi)_{\mathbf{R}_g} \approx (\partial^2 \Phi / \partial z^2)_{\mathbf{R}_g} \equiv \nu^2$. This is precisely the Hamiltonian arrived at by Heisler & Tremaine (1986) (c.f. their equation (14)) when considering the effect of the Galactic tide on the Oort Cloud comets.

2.D Signs and sizes of A and Γ

Here we provide some technical details about the statements on the signs and values of A and Γ made in §2.6.1. Also, in Table 2.4 we summarize some information about these coefficients for certain potentials. We provide two examples of orbits with extreme values of Γ .

Table 2.4: Summary of U and $\Gamma(U)$ ranges that are possible for orbits \mathbf{R}_g in different classes of potential Φ .

| Type of the potential/orbit \mathbf{R}_g | U range | Γ range |
|--|------------------------------|-----------------------------------|
| General axisymmetric potential | $-\infty \leq U \leq \infty$ | $-\infty \leq \Gamma \leq \infty$ |
| Spherical potential (assuming $d\rho/dr \leq 0$ and finite mass) | $1/3 \leq U \leq \infty$ | $0 \leq \Gamma \leq 1$ |
| Midplane of a thin disk | $U = 1$ | $\Gamma = 1/3$ |
| Vertical cylindrical potential | $U = 0$ | $\Gamma = -1/3$ |
| Axisymmetric harmonic potential | $0 \leq U \leq 1$ | $-1/3 \leq \Gamma \leq 1/3$ |
| Spherical harmonic potential | $U = 1/3$ | $\Gamma = 0$ |
| Keplerian potential | $U \rightarrow \infty$ | $\Gamma \rightarrow 1$ |
| Spherical cusp potential (density $\rho \propto r^{-\beta}$) | $U = 1/(3 - \beta)$ | $\Gamma = \beta/[3(4 - \beta)]$ |

2.D.1 Spherical potentials

Consider a spherically symmetric potential $\Phi = \Phi(r)$, where $r \equiv \sqrt{R^2 + Z^2}$ is the spherical radius. According to our convention, the outer orbit of the binary always lies in $Z = 0$ plane of the associated cylindrical (R, ϕ, Z) coordinate system. Then it is a simple matter to show that

$$\left(\frac{\partial^2 \Phi}{\partial Z^2}\right)_{\mathbf{R}_g} = \left(\frac{1}{R} \frac{\partial \Phi}{\partial R}\right)_{\mathbf{R}_g} = \left(\frac{1}{r} \frac{d\Phi}{dr}\right)_{\mathbf{R}_g}, \quad (2.78)$$

$$\left(\frac{\partial^2 \Phi}{\partial R^2}\right)_{\mathbf{R}_g} = \left(\frac{d^2 \Phi}{dr^2}\right)_{\mathbf{R}_g}. \quad (2.79)$$

Using these conversions as well as equations (2.33), (2.34), (2.41) we find

$$A = \frac{1}{2} \left[\left(\frac{d^2 \Phi}{dr^2}\right)_{\mathbf{R}_g} + 3 \left(\frac{1}{r} \frac{d\Phi}{dr}\right)_{\mathbf{R}_g} \right], \quad (2.80)$$

$$B = -\frac{1}{2} \left[\left(\frac{d^2 \Phi}{dr^2}\right)_{\mathbf{R}_g} - \left(\frac{1}{r} \frac{d\Phi}{dr}\right)_{\mathbf{R}_g} \right]. \quad (2.81)$$

We can now prove that $A > 0$ for (almost) any realistic spherical potential, and thereby show that for such systems $0 \leq \Gamma \leq 1$. In a spherical potential $\Phi(r)$ we have

$$\frac{d\Phi}{dr} = \frac{GM(r)}{r^2}, \quad (2.82)$$

where $\mathcal{M}(r)$ is the cluster mass enclosed inside radius r . Also, Poisson's equation reads

$$\frac{d^2 \Phi}{dr^2} + \frac{2}{r} \frac{d\Phi}{dr} = 4\pi G\rho, \quad (2.83)$$

allowing us to rewrite equation (2.80) as

$$A = \frac{1}{2} \left[4\pi G\overline{\rho(\mathbf{R}_g)} + \left(\frac{GM(r)}{r^3}\right)_{\mathbf{R}_g} \right]. \quad (2.84)$$

Since $\rho > 0$ and $\mathcal{M} > 0$ at all radii, this inevitably results in $A > 0$.

Finally, since

$$\frac{d^2\Phi}{dr^2} - \frac{1}{r} \frac{d\Phi}{dr} = r \frac{d}{dr} \left(\frac{1}{r} \frac{d\Phi}{dr} \right), \quad (2.85)$$

equation (2.81) can be rewritten as

$$B = -\frac{1}{2} \left[r \frac{d}{dr} \left(\frac{GM(r)}{r^3} \right) \right]_{\mathbf{R}_g}. \quad (2.86)$$

For any spherical system in which the density is a non-increasing function of radius, $d(\mathcal{M}/r^3)/dr \leq 0$ for any r and hence $B \geq 0$.

Let us now focus on spherical systems with $d\rho/dr \leq 0$. If the cluster has a constant density core, then $\mathcal{M}(r) \sim r^3$ as $r \rightarrow 0$ and so $B \rightarrow 0$ (equation (2.86)). Hence if \mathbf{R}_g orbits entirely inside the constant density region, $\Gamma = 0$. A potential without a core will always have a non-zero value of Γ for orbits at small radii.

At the other extreme, as $r \rightarrow \infty$ we have $\rho \rightarrow 0$, and usually the enclosed mass $\mathcal{M}(r) \rightarrow \text{const}$ (although see below for potentials arising from power-law cusp density profiles). Hence $d(\mathcal{M}(r)/r^3)/dr \rightarrow -3\mathcal{M}/r^4$ where \mathcal{M} is the total mass of the cluster, and in this limit we get $B = 3A$. Thus an orbit \mathbf{R}_g that spends its time exclusively at very large radii r compared to the scale radius of the cluster will have $\Gamma \rightarrow 1$. This is precisely the Lidov-Kozai limit: for potentials that are Keplerian as $r \rightarrow \infty$ (i.e. those with finite mass), orbiting far from the core is equivalent to orbiting a point mass at the origin.

Finally, for any orbit in a spherical cluster with a power-law density cusp $\rho(r) \propto r^{-\beta}$ with $0 < \beta < 3$ (so that the mass is finite at the centre) one naturally has $\mathcal{M}(r) = 4\pi\rho r^3/(3 - \beta)$. Then we find from equations (2.84), (2.86) that

$$A = 2\pi G\bar{\rho} \frac{4 - \beta}{3 - \beta}, \quad B = 2\pi G\bar{\rho} \frac{\beta}{3 - \beta}, \quad (2.87)$$

and so $U = 1/(3 - \beta)$ and $\Gamma = \beta/[3(4 - \beta)]$.

2.D.2 Axisymmetric potentials

In a general axisymmetric potential there is no constraint on how negative the parameter U (defined by equation (2.46) and plotted in Figure 2.2) can be. Non-spherical potentials naturally feature regions with $\Phi_{zz} < 0$, especially near the poles; choosing a highly inclined (with respect to the equatorial plane of the potential) orbit with large radius so that $\bar{\rho}$ is vanishingly small, one can drive strongly negative U , thereby achieving extreme (positive or negative) values of Γ .

In Figure 2.12 we give an example of an orbit with $\Gamma_{\text{num}} = -1.4$. We use the Miyamoto-Nagai potential with $b_h/b_\ell = 1$ and precisely the same initial conditions as Orbits (VIII)-(XI) in the main text (see Table 2.3), except that we now take the initial Z

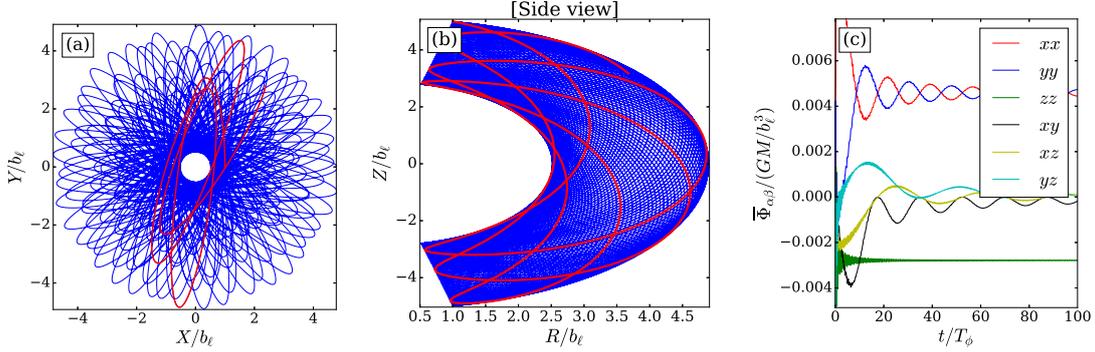


Figure 2.12: Example of an orbit with $\Gamma_{\text{num}} = -1.4$. We use the Miyamoto-Nagai potential with $b_h/b_\ell = 1$ and the same initial conditions as Orbits (VIII)-(XI) in the main text, except the initial Z value is $5.0b_\ell$.

coordinate to be $5.0b_\ell$. All three panels show $100T_\phi$ of data. The large initial value of Z means that the orbit spends a lot of time near the poles of the potential where $\bar{\Phi}_{zz} < 0$.

In Figure 2.13 we provide an example of a very polar orbit resulting in $\Gamma_{\text{num}} = 2.0$ ($U = -2.3$). We use the Miyamoto-Nagai potential $b_h/b_\ell = 0.1$, and initial conditions

$$(R, v_R, \phi, v_\phi, Z, v_Z) = (0.011b_\ell, 0, 0.04, 0.4\sqrt{GM/b_\ell}, 4.0b_\ell, 0). \quad (2.88)$$

In Figure 2.13a we display only the first 5 vertical periods T_Z of the orbit in the meridional (R, Z) plane. When integrated for a long time, the orbit remains almost polar but precesses very slowly around the Z axis until it eventually fills an axisymmetric torus after a few thousand T_Z . Figure 2.13a shows that the convergence of the $\bar{\Phi}_{\alpha\beta}$ coefficients in this case is very slow and takes $\gtrsim 5000T_Z$. In practice, unless the binary is very tight, secular theory is unlikely to work well for such an orbit. Indeed, for a relatively wide binary the secular evolution timescale is likely to be much shorter than $5000T_Z$, meaning that our assumption that the binary fills its torus (and hence the $\Phi_{\alpha\beta}$ converge) in much less than a secular timescale is violated. We will explore this issue in more detail in §3.7.

2.E Octupole Hamiltonian

The Hamiltonian derived in §2.2 is correct to quadratic order in $a/|\mathbf{R}_g|$, the so-called ‘quadrupole approximation’. We can attempt to derive a more accurate Hamiltonian by keeping the higher order terms in the series expansion of equation (2.3). The next (‘octupole’) term that we would include is

$$\frac{1}{2!} \sum_{\beta\gamma} \left(\frac{\partial^2 \Phi}{\partial R_\alpha \partial R_\beta \partial R_\gamma} \right)_{\mathbf{R}_g} r_{i,\beta} r_{i,\gamma}. \quad (2.89)$$

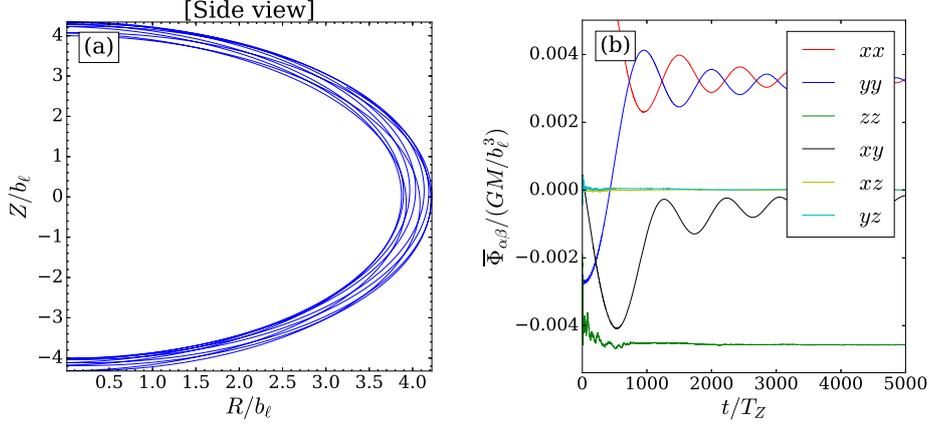


Figure 2.13: Example of an orbit with $\Gamma_{\text{num}} = 2.0$. We use the Miyamoto-Nagai potential with $b_h/b_\ell = 0.1$ and initial conditions $(R, v_R, \phi, v_\phi, Z, v_Z) = (0.011b_\ell, 0, 0.04, 0.4\sqrt{GM/b_\ell}, 4b_\ell, 0)$. In panel (a) we show the meridional (R, Z) plane for only the first 5 vertical periods T_Z of the integration. The orbit precesses very slowly in azimuth, filling its torus after several thousand T_Z .

We then use the fact that $\mathbf{r} \equiv \mathbf{r}_1 - \mathbf{r}_2$ and $m_1\mathbf{r}_1 + m_2\mathbf{r}_2 = \mathbf{0}$ to write

$$r_{1,\alpha} = \frac{m_2}{m_1 + m_2} r_\alpha; \quad r_{2,\alpha} = -\frac{m_1}{m_1 + m_2} r_\alpha. \quad (2.90)$$

As a result, the equation for the relative motion $d^2\mathbf{r}/dt^2$ (equation (2.5)) can be written purely in terms of \mathbf{r} . The next order (‘octupole’) correction to the Hamiltonian (2.42) is

$$H_{\text{oct}} = \frac{1}{3!} \left(\frac{m_2 - m_1}{m_2 + m_1} \right) \sum_{\alpha\beta\gamma} \Phi_{\alpha\beta\gamma} r_\alpha r_\beta r_\gamma. \quad (2.91)$$

Note that the octupole term vanishes for equal-mass binaries ($m_1 = m_2$).

The corresponding doubly-averaged perturbing octupole term is then simply

$$\overline{\langle H_{\text{oct}} \rangle}_M = \frac{1}{3!} \left(\frac{m_2 - m_1}{m_2 + m_1} \right) \sum_{\alpha\beta\gamma} \overline{\Phi}_{\alpha\beta\gamma} \langle r_\alpha r_\beta r_\gamma \rangle_M. \quad (2.92)$$

2.E.1 Time-averaging over an axisymmetric torus

As in the quadrupole case, it turns out that when time-averaged over an axisymmetric torus the coefficients $\overline{\Phi}_{\alpha\beta\gamma}$ satisfy various symmetry properties. After a little algebra one can show that

$$\overline{\Phi}_{xxz} = \overline{\Phi}_{yyz}, \quad (2.93)$$

and all other $\overline{\Phi}_{\alpha\beta\gamma} = 0$ except for $\overline{\Phi}_{zzz}$. Hence

$$\overline{\langle H_{\text{oct}} \rangle}_M = \frac{1}{3!} \left(\frac{m_2 - m_1}{m_2 + m_1} \right) \left[3\overline{\Phi}_{xxz} \langle (x^2 + y^2)z \rangle_M + \overline{\Phi}_{zzz} \langle z^3 \rangle_M \right]. \quad (2.94)$$

Writing out the $\langle \cdot \rangle_M$ factors in terms of orbital elements we have

$$\langle (x^2 + y^2)z \rangle_M = -\frac{5}{128}a^3e \left[28e^2 \sin^3 i \sin 3\omega + (4 + 3e^2)(7 \sin i + 3 \sin 3i) \sin \omega \right], \quad (2.95)$$

$$\langle z^3 \rangle_M = -\frac{5}{16}a^3e(6 + e^2 - 7e^2 \cos 2\omega) \sin^3 i \sin \omega. \quad (2.96)$$

Equations (2.94)-(2.96) provide a general framework for accounting for the octupole contribution to the tidal Hamiltonian in an arbitrary axisymmetric potential. Note there is no Ω dependence in the octupole Hamiltonian, so J_z is still conserved to octupole order.

2.E.2 Link to the test particle octupole LK Hamiltonian

Note that one *cannot* recover the test particle octupole term of the doubly-averaged Lidov-Kozai Hamiltonian by putting $\Phi = -GM/r$ in (2.94), because equation (2.94) is derived under the axisymmetric approximation. The time-averaged potential of a perturber on a Keplerian orbit is only axisymmetric at the *quadrupole* level, and the symmetry is broken by octupole terms. Instead one must integrate over the outer Keplerian orbit exactly, as in Appendix 2.B.

In general there are 10 independent time-averaged coefficients $\bar{\Phi}_{\alpha\beta\gamma}$ to consider. We choose the outer orbit to be in the $Z = 0$ plane so we can immediately eliminate four of these, $\bar{\Phi}_{zzz} = \bar{\Phi}_{xxz} = \bar{\Phi}_{yyz} = \bar{\Phi}_{xyz} = 0$. We can also choose the pericentre of the outer orbit to be on the X axis without loss of generality, so that the ellipse traced by the outer orbit is symmetric under $Y \rightarrow -Y$. Then all $\bar{\Phi}_{\alpha\beta\gamma}$ that contain an odd number of Y derivatives will be antisymmetric under $Y \rightarrow -Y$, so their time-averages over this ellipse will vanish: $\bar{\Phi}_{xxy} = \bar{\Phi}_{zzy} = \bar{\Phi}_{yyy} = 0$. This leaves us with only three non-zero terms in the doubly-averaged octupole LK Hamiltonian:

$$\overline{\langle H_{\text{Oct}} \rangle}_M = \frac{1}{3!} \left(\frac{m_2 - m_1}{m_2 + m_1} \right) \left[\bar{\Phi}_{xxx} \langle x^3 \rangle_M + 3\bar{\Phi}_{yyx} \langle y^2 x \rangle_M + 3\bar{\Phi}_{zzx} \langle z^2 x \rangle_M \right]. \quad (2.97)$$

For reference we now write down the terms that make up equation (2.97). First we write down the necessary $\bar{\Phi}_{\alpha\beta\gamma}$ coefficients in terms of cylindrical coordinates R_g and ϕ_g :

$$\bar{\Phi}_{xxx} = \frac{3GM}{R_g^4} \cos \phi_g (5 \cos^2 \phi_g - 3), \quad (2.98)$$

$$\bar{\Phi}_{yyx} = \frac{3GM}{R_g^4} \cos \phi_g (5 \sin^2 \phi_g - 1), \quad (2.99)$$

$$\bar{\Phi}_{zzx} = -\frac{3GM}{R_g^4} \cos \phi_g. \quad (2.100)$$

The time-averages of these coefficients are

$$\overline{\Phi}_{xxx} = GMa_g^{-4}(1 - e_g^2)^{-5/2} \times (9e_g/4), \quad (2.101)$$

$$\overline{\Phi}_{yyx} = GMa_g^{-4}(1 - e_g^2)^{-5/2} \times (3e_g/4), \quad (2.102)$$

$$\overline{\Phi}_{zzx} = GMa_g^{-4}(1 - e_g^2)^{-5/2} \times (-3e_g). \quad (2.103)$$

The mean-anomaly averaged quantities are

$$\begin{aligned} \langle x^3 \rangle_M &= - (5/64)a^3 e (\cos \omega \cos \Omega - \cos i \sin \omega \sin \Omega) \\ &\quad \times \left[\cos 2i(6 + e^2 + 7e^2 \cos 2\omega \cos 2\Omega) \right. \\ &\quad \left. + 7e^2 \cos 2\omega(3 \cos 2\Omega + 2 \sin^2 i) \right. \\ &\quad \left. + (6 + e^2)(3 + 2 \cos 2\Omega \sin^2 i) \right. \\ &\quad \left. - 28e^2 \cos i \sin 2\omega \sin 2\Omega \right], \end{aligned} \quad (2.104)$$

$$\begin{aligned} \langle y^2 x \rangle_M &= (5/256)a^3 e \\ &\quad \times \left[2 \cos 2i \cos \omega (-2 - 5e^2 + 7e^2 \cos 2\omega) \right. \\ &\quad \times (\cos \Omega + 3 \cos 3\Omega) \\ &\quad \left. + 2 \cos \omega (-7(2 + e^2 + e^2 \cos 2\omega) \cos \Omega \right. \\ &\quad \left. + (6 - 13e^2 + 35e^2 \cos 2\omega) \cos 3\Omega) \right. \\ &\quad \left. + 4 \cos 3i(6 + e^2 - 7e^2 \cos 2\omega) \cos^2 \Omega \sin \omega \sin \Omega \right. \\ &\quad \left. + \cos i \sin \omega [(26 + 23e^2 + 7e^2 \cos 2\omega) \sin \Omega \right. \\ &\quad \left. - 3(2 + 19e^2 + 35e^2 \cos 2\omega) \sin 3\Omega] \right], \end{aligned} \quad (2.105)$$

$$\begin{aligned} \langle z^2 x \rangle_M &= (5/16)a^3 e \sin^2 i \\ &\quad \times \left[\cos \omega (-2 - 5e^2 + 7e^2 \cos 2\omega) \cos \Omega \right. \\ &\quad \left. + \cos i(6 + e^2 - 7e^2 \cos 2\omega) \sin \omega \sin \Omega \right]. \end{aligned} \quad (2.106)$$

Plugging the results (2.106)-(2.101) in to (2.97), the resulting doubly-averaged test particle octupole Lidov-Kozai Hamiltonian is

$$\begin{aligned} \overline{\langle H_{\text{Oct}} \rangle}_M &= \left(\frac{m_2 - m_1}{m_2 + m_1} \right) \times \frac{15}{128} GMa_g^{-4} e_g (1 - e_g^2)^{-5/2} a^3 \\ &\quad \times \left\{ \left(e + \frac{3e^3}{4} \right) \left[(1 - 11\theta - 5\theta^2 + 15\theta^3) \cos(\omega - \Omega) \right. \right. \\ &\quad \left. \left. + (1 + 11\theta - 5\theta^2 - 15\theta^3) \cos(\omega + \Omega) \right] \right. \\ &\quad \left. - \frac{35}{4} e^3 \left[(1 - \theta - \theta^2 + \theta^3) \cos(3\omega - \Omega) \right. \right. \\ &\quad \left. \left. + (1 + \theta - \theta^2 - \theta^3) \cos(3\omega + \Omega) \right] \right\}, \end{aligned} \quad (2.107)$$

where $\theta \equiv \cos i$. Equation (2.107) is precisely the result found in standard LK literature (e.g. Ford, Kozinsky, et al. 2000; Lithwick & Naoz 2011; Naoz 2016).

2.F Numerical prescription for computing time-averages

To calculate the time-averages $\overline{\Phi_{\alpha\beta}}$ numerically we use the orbit integrator in `galpy` (Bovy 2015). Given the initial position $\mathbf{R}_g(0)$ and velocity of the binary's outer orbit around the cluster, we integrate its equation of motion (2.11) numerically in the smooth cluster potential Φ . We use a constant timestep Δt so that after k timesteps the time elapsed is $t_k = k\Delta t$. Then the running time-average of a quantity $\mathcal{F}(\mathbf{R}_g(t))$ is

$$\overline{\mathcal{F}}(t) = \frac{1}{t} \int_0^t dt' \mathcal{F}(t') \approx \frac{\Delta t}{t} \sum_{k=0}^{t/\Delta t} \mathcal{F}(t_k). \quad (2.108)$$

In nearly all numerical examples shown in this Chapter we used $\Delta t \approx T_\phi/100$ where T_ϕ is the azimuthal period of \mathbf{R}_g , and integrated the outer orbit for approximately $100T_\phi$. The exception is Figure 2.13, where we used $\Delta t \approx T_Z/10$ (T_Z is the vertical period of \mathbf{R}_g) and integrated the outer orbit for $5000T_Z$.

The work presented in this Chapter has been published
in *Monthly Notices of the Royal Astronomical Society* as
Hamilton & Rafikov (2019c).

3

Secular dynamics

Contents

| | | |
|------------|---|-----------|
| 3.1 | Introduction | 66 |
| 3.2 | General aspects of secular dynamics | 67 |
| 3.2.1 | Phase portrait | 68 |
| 3.2.2 | Fixed points and orbit families | 68 |
| 3.2.3 | Does a given orbit librate or circulate? | 69 |
| 3.2.4 | Maximum and minimum eccentricities | 70 |
| 3.2.5 | Range of parameter values | 71 |
| 3.2.6 | Timescales of eccentricity oscillations | 72 |
| 3.3 | Secular dynamics in the case $\Gamma > 1/5$ | 76 |
| 3.3.1 | Fixed points and orbit families | 77 |
| 3.3.2 | Range of parameter values | 78 |
| 3.3.3 | Maximum and minimum eccentricities | 78 |
| 3.3.4 | Timescales of eccentricity oscillations | 79 |
| 3.4 | The case $0 < \Gamma \leq 1/5$ | 79 |
| 3.4.1 | Fixed points and orbit families | 80 |
| 3.4.2 | Range of parameter values | 80 |
| 3.4.3 | Maximum and minimum eccentricities | 82 |
| 3.4.4 | Timescales of eccentricity oscillations | 84 |
| 3.5 | The case $-1/5 < \Gamma \leq 0$ | 84 |
| 3.6 | The case $\Gamma \leq -1/5$ | 85 |
| 3.7 | Accuracy of the doubly-averaged approximation | 85 |
| 3.7.1 | Method | 87 |
| 3.7.2 | Accuracy of the doubly-averaged approximation for $\Gamma > 1/5$. | 87 |
| 3.7.3 | Accuracy of the doubly-averaged approximation for $0 < \Gamma \leq 1/5$. | 91 |
| 3.7.4 | Accuracy of the doubly-averaged approximation in the cases $-1/5 < \Gamma \leq 0$ and $\Gamma \leq -1/5$ | 94 |
| 3.7.5 | Discussion: validity of the doubly-averaged secular theory | 95 |
| 3.8 | Effect of short-range forces on the cluster-tide driven evolution | 97 |
| 3.9 | Discussion | 99 |
| 3.9.1 | Critical inclination for the existence of fixed points | 100 |
| 3.9.2 | High eccentricity behaviour | 102 |
| 3.9.3 | Stellar scattering and other non-ideal effects | 103 |

| | | |
|-------------|---|------------|
| 3.9.4 | Relation to previous work | 105 |
| 3.10 | Summary | 106 |
| | Appendices | 107 |
| 3.A | Detailed characteristics of the $-1/5 < \Gamma \leq 0$ regime. | 107 |
| 3.A.1 | Fixed points and orbit families | 107 |
| 3.A.2 | Range of parameter values | 107 |
| 3.A.3 | Maximum and minimum eccentricities | 107 |
| 3.A.4 | Timescales of eccentricity oscillations | 107 |
| 3.B | Detailed characteristics of the $\Gamma \leq -1/5$ regime. | 108 |
| 3.B.1 | Fixed points and orbit families | 108 |
| 3.B.2 | Range of parameter values | 108 |
| 3.B.3 | Maximum and minimum eccentricities | 108 |
| 3.B.4 | Timescales of eccentricity oscillations | 109 |

3.1 Introduction

In Chapter 2 we developed a general Hamiltonian framework to describe the evolution of the orbital elements of a binary driven by the tidal field of a cluster in which it orbits. In so doing we arrived at a secular (‘doubly-averaged’) perturbing Hamiltonian (equations (2.42)-(2.43)), which we repeat here for convenience:

$$\overline{\langle H_1 \rangle}_M = CH_1^*, \quad \text{where} \quad C = \frac{Aa^2}{8} \quad (3.1)$$

is a constant with dimensions of energy per unit mass and H_1^* is the ‘dimensionless Hamiltonian’

$$H_1^* = (2 + 3e^2) \left(1 - 3\Gamma \cos^2 i \right) - 15\Gamma e^2 \sin^2 i \cos 2\omega. \quad (3.2)$$

Chapter 2 explored in detail the dependence of A and Γ upon the shape of the background potential and the binary’s orbit within it.

The goal of this Chapter is to systematically explore the dynamics that result from the general secular theory based on the Hamiltonian (3.1)-(3.2). As we will see, central to this investigation is the orbital *phase space*, i.e. the (ω, e) plane within which a binary moves on a one-dimensional trajectory. It turns out that the binary dynamics and phase space structure in the range $\Gamma > 1/5$ are qualitatively very similar to those arising in the LK case, but that bifurcations occur when $\Gamma = 1/5, 0$ and $-1/5$, which change the picture significantly. As a result, in this Chapter we treat separately four distinct dynamical regimes:

$$\Gamma > 1/5, \quad (3.3)$$

$$0 < \Gamma \leq 1/5, \quad (3.4)$$

$$-1/5 < \Gamma \leq 0, \quad (3.5)$$

$$\Gamma \leq -1/5. \quad (3.6)$$

This Chapter is structured as follows. In §3.2 we derive general results that hold for all Γ : conditions for the existence of fixed points in the phase space (§3.2.2), the criteria for phase-space trajectories to librate or circulate (§3.2.3), the values of maximum and minimum eccentricities (§3.2.4), the timescale of eccentricity oscillations (§3.2.6), etc. Then, in §§3.3-3.6, we explore the details of each of the Γ regimes (3.3)-(3.6) separately. The validity of the doubly-averaged (secular) theory is verified numerically in §3.7. The impact of short-range forces, in particular general relativistic (GR) pericentre precession, on the cluster tide-driven secular evolution is explored in §3.8 (a much more detailed study is deferred to Chapter 4). We collect our results in §3.9, discuss them in light of the existing literature and comment on the applicability of our formalism. We summarise our findings and discuss potential applications in §3.10.

3.2 General aspects of secular dynamics

Our goal is to understand evolution of the orbital elements of the inner orbit of the binary — eccentricity, inclination, etc. — as it moves in the cluster potential. We do this by carefully investigating the phase-space of the dimensionless Hamiltonian H_1^* (equation (3.2)).

The DA Hamiltonian (3.1) does not depend on the mean anomaly M , so the action L is conserved. Hence we can choose to work with dimensionless versions of our variables (J, J_z); following the notation of Antognini (2015) we define

$$\Theta \equiv J_z^2/L^2 = (1 - e^2) \cos^2 i, \quad j \equiv J/L = \sqrt{1 - e^2}, \quad (3.7)$$

and we clearly have $0 \leq \Theta \leq j^2 \leq 1$. Obviously j is just the dimensionless angular momentum. Then

$$H_1^* = \frac{1}{j^2} \left[(j^2 - 3\Gamma\Theta)(5 - 3j^2) - 15\Gamma(j^2 - \Theta)(1 - j^2) \cos 2\omega \right]. \quad (3.8)$$

Since H_1^* is independent of the longitude of the ascending node Ω , the dimensionless quantity Θ is also an integral of motion (the analog of the ‘Kozai constant’). This simply reflects conservation of the Z -component of angular momentum since the doubly-averaged potential is axisymmetric. Since Θ is conserved, we can always infer the time evolution of binary inclination from the behavior of its eccentricity via $\cos^2 i = (1 - e^2)^{-1}\Theta$. Finally, definitions (3.7) imply that e and j must obey

$$0 \leq e \leq \sqrt{1 - \Theta}, \quad \Theta^{1/2} \leq j \leq 1, \quad (3.9)$$

to be physically meaningful for a fixed Θ .

Given these considerations, the secular Hamiltonian is a function of the dimensionless angle-action variables ω and j . The equations of motion fully describing their evolution are

$$\frac{d\omega}{dt} = \frac{C}{L} \frac{\partial H_1^*}{\partial j} = \frac{6C}{L} \frac{[5\Gamma\Theta - j^4 + 5\Gamma(j^4 - \Theta) \cos 2\omega]}{j^3}, \quad (3.10)$$

$$\frac{dj}{dt} = -\frac{C}{L} \frac{\partial H_1^*}{\partial \omega} = -\frac{30\Gamma C}{L} \frac{(j^2 - \Theta)(1 - j^2)}{j^2} \sin 2\omega. \quad (3.11)$$

Our subsequent analysis of binary dynamics is largely based on these evolution equations.

3.2.1 Phase portrait

Since the dimensionless doubly-averaged Hamiltonian (3.8) ends up being a function of $j = \sqrt{1 - e^2}$ and ω , one can get a good perception of the secular dynamics by plotting contours of H_1^* in the (ω, e) plane. We do this in Figures 3.4, 3.5, 3.6 & 3.7 for the Γ ranges (3.3)-(3.6) respectively, and for varying Θ . In each panel the limiting eccentricity $e_{\text{lim}} = \sqrt{1 - \Theta}$ is represented by a dashed black line. The direction in which orbits traverse their trajectories is indicated with green arrows in Figures 3.4a, 3.5d, 3.6a & 3.7d.

We will explain the features of these phase portraits as we discuss each of them individually in §§3.3-3.6. An observation that we would like to make now is that all phase-space trajectories are split into two families: librating and circulating. The librating orbits are closed contours of H_1^* which loop around a fixed point (always located at $\omega = \pm\pi/2$ as explained in §3.2.2), whereas circulating orbits run over all $\omega \in (-\pi, \pi)$. The separatrices between families of librating and circulating orbits are indicated with red dashed lines. Depending on the values of Γ and Θ , one could have either only circulating orbits (typically the case for large Θ , and always true for $-1/5 < \Gamma \leq 0$), only librating orbits (a rare case realized for $\Gamma = 1/5$, see Figure 3.5), or a mix of both (for low enough Θ in most dynamical regimes).

3.2.2 Fixed points and orbit families

We start by exploring characteristics of the fixed points of the system, around which phase-space orbits librate. As these points are extrema of H_1^* in (ω, j) space, they must be solutions of $d\omega/dt = dj/dt = 0$. With a small amount of algebra, we find from equations (3.10)-(3.11) two possible formal solutions for the non-trivial fixed points in our phase-space¹, namely $(\omega, j) = (\pm\pi/2, j_{\text{f}})$ where

$$j_{\text{f}}(\Theta, \Gamma) = \left(\frac{\Theta}{\Lambda}\right)^{1/4}, \quad \Lambda(\Gamma) = \frac{5\Gamma + 1}{10\Gamma}. \quad (3.12)$$

¹Fixed points also exist for all ω along the lines $j = 1$ and $j = \sqrt{\Theta}$, but these are trivial in the sense that they can never be reached by orbits which do not start on those lines. However, they are still important because they bound the phase-space and are often the locations of maximum/minimum H_1^* , see §3.2.5.

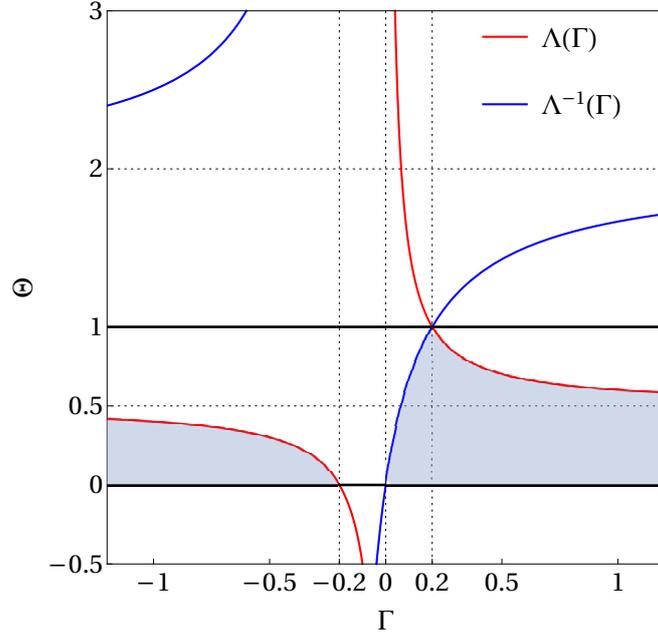


Figure 3.1: Ranges of Θ for which fixed points and librating orbits can exist (shaded regions), shown as a function of Γ . Red and blue curves show Λ (defined by equation (3.12)) and Λ^{-1} respectively. Shaded regions correspond to the constraints (3.13) and $0 \leq \Theta \leq 1$.

The value of ω corresponding to fixed points agrees with the phase portraits discussed in §3.2.1.

For a given Γ , fixed points can exist in the (ω, j) phase-space as long as $j_f(\Theta, \Gamma)$ satisfies the condition (3.9). Solving the inequality $\sqrt{\Theta} < j_f < 1$ gives the following constraint on Θ for the existence of fixed points:

$$\Theta < \min(\Lambda, \Lambda^{-1}). \quad (3.13)$$

Depending on the value of Γ , this constraint may or may not have meaningful solutions. The functions $\Lambda(\Gamma)$ and $\Lambda^{-1}(\Gamma)$ are plotted in Figure 3.1. We can see from this plot that there are four distinct Γ regimes, given by the ranges (3.3)-(3.6). We will return to Figure 3.1 when discussing the existence of fixed points in §§3.3-3.6.

3.2.3 Does a given orbit librate or circulate?

Next we work out whether a given phase-space orbit with specified H_1^* , Θ , Γ librates or circulates. We do this by considering the behaviour at $\omega = 0$: from the morphology of the phase portraits (Figures 3.4, 3.5, 3.6 & 3.7), it is clear that if constant $H_1^* = H_1^*(\Theta, \Gamma, \omega, j)$ has a physical solution for j at $\omega = 0$ then the orbit circulates; if not, it has to librate about one of the fixed points.

We find from equation (3.8) a formal solution

$$j^2(\omega = 0) = \frac{5/3 - 2\Gamma\Theta - 5\Gamma - H_1^*/3}{1 - 5\Gamma}. \quad (3.14)$$

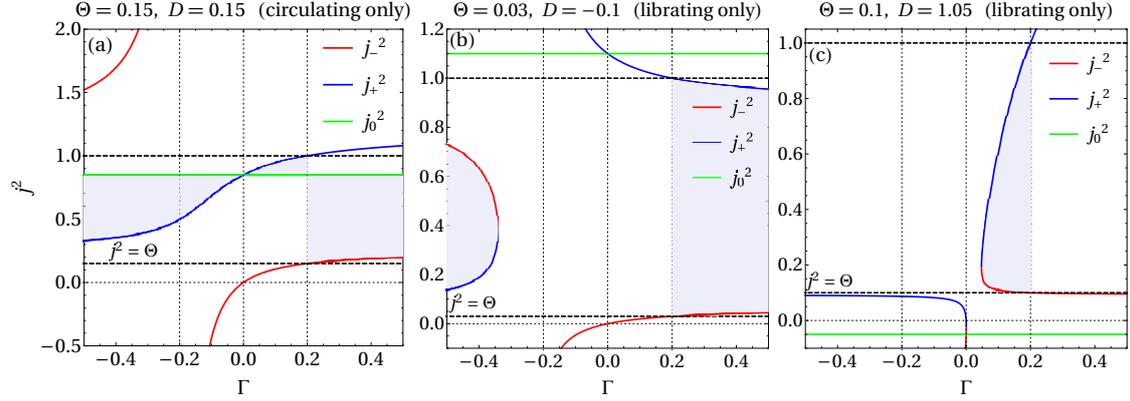


Figure 3.2: Plot showing j_0^2 and j_{\pm}^2 defined by equations (3.17) and (3.18) as a function of Γ for fixed (D, Θ) . Vertical dotted lines indicate $\Gamma = \pm 1/5, 0$. Thick horizontal dashed lines correspond to $j^2 = \Theta$ and $j^2 = 1$. Note that j_{\pm}^2, j_0^2 can take values outside the allowed $\Theta < j^2 < 1$ range and can even be negative, because they are simply formal solutions (see the text after equation (3.20)). In each Γ range, either two or none of j_{\pm}^2, j_0^2 are physically relevant — physical solutions j^2 must always lie in the shaded regions. Panel (a) corresponds to circulating orbits (which exist for all $\Gamma \neq 1/5$), while panels (b) and (c) correspond to librating orbits. From these plots one can read off the relative amplitudes of j_{\pm}^2, j_0^2 (and hence the values of $j_{\min/\max}$ and Δ) in each Γ regime.

The trajectory is circulating whenever the condition (3.9) is satisfied for $j(\omega = 0)$. If $j(\omega = 0)$ does not obey this inequality, then it does not represent a physically meaningful solution. As a result, the orbit must librate around one of the fixed points and never reach $\omega = 0$.

Let us define the quantity

$$D \equiv 1 - j^2(\omega = 0) = \frac{H_1^*/3 - 2/3 + 2\Gamma\Theta}{1 - 5\Gamma}. \quad (3.15)$$

It represents $e^2(\omega = 0)$ and is an integral of motion since it depends on H_1^* and Θ . It will prove useful to eliminate Θ and H_1^* in the above expression in favour of e, i, ω . We find after some algebra

$$D = e^2 \left(1 + \frac{10\Gamma}{1 - 5\Gamma} \sin^2 i \sin^2 \omega \right). \quad (3.16)$$

Equations (3.9) and (3.15) imply that for $0 < D < 1 - \Theta$ the trajectory is circulating, whereas for $D < 0$ or $D > 1 - \Theta$ it librates around a fixed point. If both families exist then the separatrix between them corresponds to either $D = 0$ or $D = 1 - \Theta$ (see the upcoming Figure 3.3).

3.2.4 Maximum and minimum eccentricities

We now find the minimum and maximum eccentricities that a binary reaches as it evolves along its phase-space trajectory. We do this by finding the extrema of j , which we call $j_{\min/\max}$.

From the phase portraits in Figures 3.4, 3.5, 3.6 & 3.7 it is clear that librating orbits, whenever they exist, have both their minimum and maximum eccentricities at $\omega = \pm\pi/2$.

On the other hand, for circulating orbits the maximum and minimum eccentricities can be at either $\omega = 0$ or $\omega = \pm\pi/2$ depending on the value of Γ ; see Figures 3.4 and 3.5. To find $j_{\min/\max}$, we therefore separately plug $\omega = 0, \pi/2$ into H_1^* (equation (3.8)) and solve for j .

The solution for $\omega = 0$ is simply the square root of equation (3.14), and we denote it j_0 :

$$j_0 \equiv j(\omega = 0) = \sqrt{1 - D}. \quad (3.17)$$

For $\omega = \pi/2$ there are two solutions, which we call j_{\pm} :

$$j_{\pm} \equiv j(\omega = \pi/2) = \sqrt{\frac{\Sigma \pm \sqrt{\Sigma^2 - 10\Gamma\Theta(1 + 5\Gamma)}}{1 + 5\Gamma}}, \quad (3.18)$$

where

$$\Sigma = \frac{1 + 5\Gamma}{2} + 5\Gamma\Theta + \left(\frac{5\Gamma - 1}{2}\right) D \quad (3.19)$$

$$= \frac{1 + 5\Gamma}{2}(1 - e^2) + 5\Gamma(\cos^2 i + e^2 \sin^2 i \cos^2 \omega). \quad (3.20)$$

We would like to stress here that although we use the notation j_{\pm}^2, j_0^2 , these quantities are nothing more than possible solutions to the equations $dj/dt = d\omega/dt = 0$, and should not therefore be interpreted as always positive. Indeed, depending on the Γ regime either two or none of j_{\pm}^2, j_0^2 will lie in the allowed physical range $(\Theta, 1)$ (equation (3.9)) — the remaining one or three will lie outside this range and can even be negative, but are physically irrelevant.

This is demonstrated in Figure 3.2, which shows j_{\pm}^2, j_0^2 as functions of Γ for various points in (D, Θ) space. Depending on the dynamical regime (determined by the value of Γ) each of these j_{\pm}, j_0 solutions, if they exist, can correspond to either the minimum or maximum of j , as we describe in further detail in §§3.3-3.6.

3.2.5 Range of parameter values

We would first like to determine the range of values that the integral of motion H_1^* can take. To do this we need to find the extrema of our Hamiltonian H_1^* in the (ω, e) (or equivalently (ω, j)) phase-space. It is clear from the phase portraits (Figures 3.4, 3.5, 3.6 & 3.7) that extrema of H_1^* can only occur in three distinct locations: fixed points² ($\omega = \pi/2, j = j_f$), the line $j = 1$, and the line $j = \sqrt{\Theta}$. Evaluating H_1^* in these locations, we find

$$H_1^*(j = 1) = 2(1 - 3\Gamma\Theta), \quad (3.21)$$

$$H_1^*(j = \Theta) = (5 - 3\Theta)(1 - 3\Gamma), \quad (3.22)$$

$$H_1^*(\omega = \pi/2, j = j_f) = \begin{cases} H_-, & \Gamma > 0, \\ H_+, & \Gamma \leq -1/5, \end{cases} \quad (3.23)$$

²Evaluating at $\omega = -\pi/2$ gives the same answers as evaluating at $\omega = \pi/2$. Fixed points exist in both locations, but to avoid confusion with other upcoming \pm signs we prefer just to consider $\omega = \pi/2$ from here on.

where³

$$H_{\pm} = 5(1 + 3\Gamma) + 24\Gamma\Theta \pm 6\sqrt{10\Gamma\Theta(1 + 5\Gamma)}. \quad (3.24)$$

(No fixed points exist in the range $-1/5 < \Gamma \leq 0$). One must then investigate each Γ regime independently to work out which of the above corresponds to a maximum or minimum. We will not pursue the details here but we state the results for each Γ regime in §§3.3-3.6, and summarise them in Table 3.2.

Rather than H_1^* , Θ , it is often convenient to take D and Θ to be our primary integrals of motion, so that for fixed Γ each phase-space trajectory of a binary corresponds to a point in the (D, Θ) plane (see the upcoming Figure 3.3). It is obvious that Θ can always run between 0 and 1 for circulating orbits, and is bounded by equation (3.13) for librating orbits. We would like to know which values D can take for a given Θ, Γ . From equation (3.15) we see that extrema of D are also extrema of H_1^* , so we must evaluate D at the same three locations as H_1^* , see equations (3.21)-(3.23). We find

$$D(j = 1) = 0, \quad (3.25)$$

$$D(j = \Theta) = 1 - \Theta, \quad (3.26)$$

$$D(\omega = \pi/2, j = j_f) = \begin{cases} D_-, & \Gamma > 0, \\ D_+, & \Gamma \leq -1/5, \end{cases} \quad (3.27)$$

where

$$D_{\pm} = \frac{1}{(1 - 5\Gamma)} \left[1 + 5\Gamma + 10\Gamma\Theta \pm 2\sqrt{10\Gamma\Theta(1 + 5\Gamma)} \right]. \quad (3.28)$$

In each Γ regime and for each type of orbit the maximum and minimum D will correspond to some combination of D_{\pm} , $D = 0$ and $D = 1 - \Theta$. These limits give rise to distinctive morphologies of the physically allowed regions in the (D, Θ) plane — see Figure 3.3. They are summarised in Table 3.1 and discussed in §§3.3-3.6.

3.2.6 Timescales of eccentricity oscillations

We can also derive a general expression for the timescale of secular eccentricity oscillations, for any binary and for any given Γ .

Since H_1^* is a conserved quantity we can rearrange (3.8) to get ω explicitly in terms of j :

$$\cos 2\omega = \frac{(j^2 - 3\Gamma\Theta)(5 - 3j^2) - j^2 H_1^*}{15\Gamma(j^2 - \Theta)(1 - j^2)}. \quad (3.29)$$

³Note that the functions H_{\pm} , unlike j_{\pm} , are *not* two independent solutions. The \pm sign is due merely to an algebraic peculiarity that arises when evaluating the single-valued function H_1^* at $(\pi/2, j_f)$, depending on the Γ range. For a given Γ , only one of H_{\pm} is correct. The same consideration holds for D_{\pm} in the next paragraph.

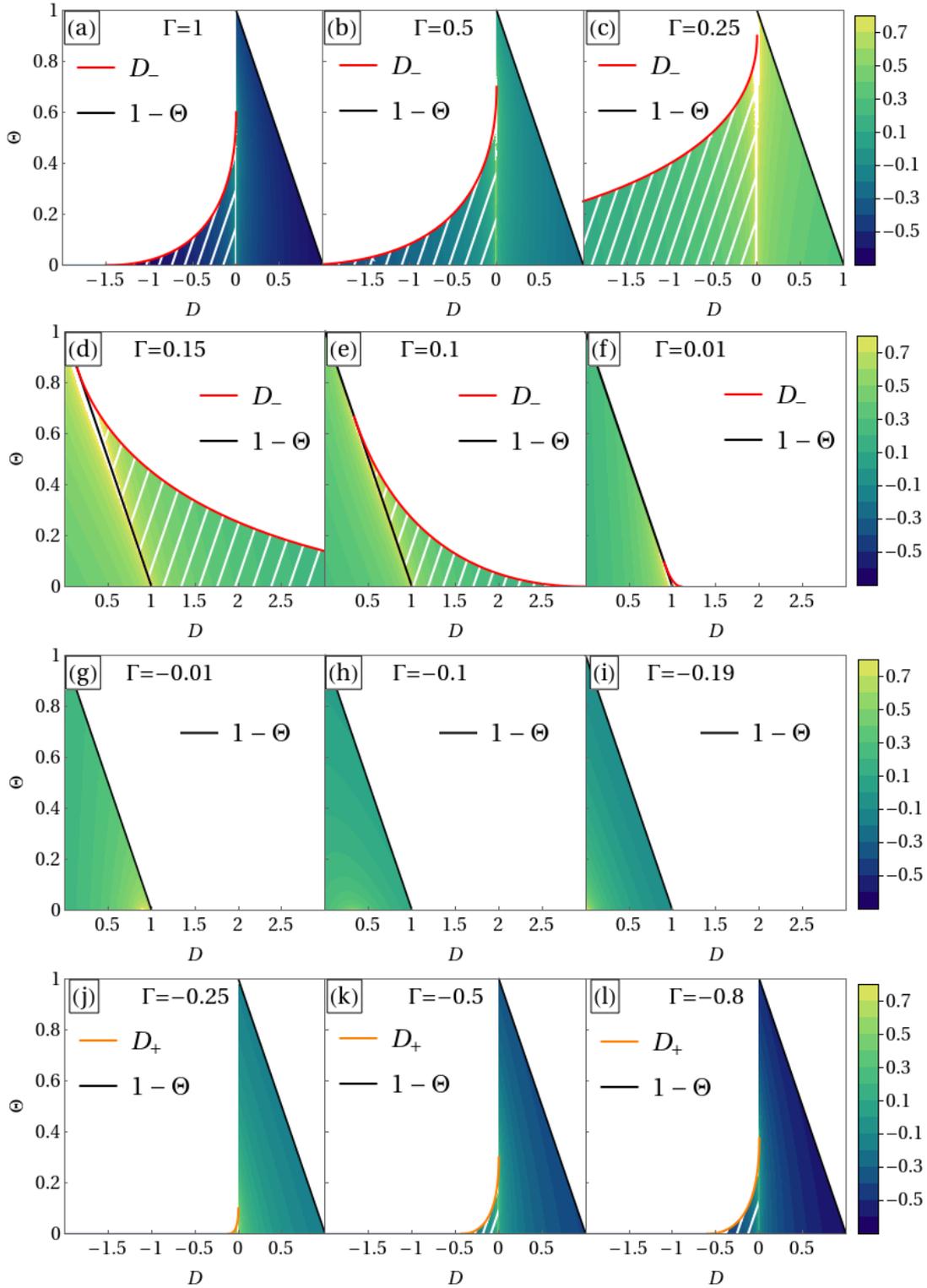


Figure 3.3: Contour plots of $\log_{10}(t_{\text{sec}}/t_1)$, where t_{sec} is the period of secular eccentricity oscillations (equation (3.33)), in the (D, Θ) plane for various values of Γ . Each row corresponds to one of the dynamical regimes (3.3)-(3.6). In each panel the circulating orbits fill the triangle $0 < D < 1 - \Theta$, while regions of librating orbits are indicated with white hashing (one of their boundaries is given by $D_{\pm}(\Theta)$ defined by equation (3.28)). The secular period diverges at the separatrix between the regions of librating and circulating orbits. It also diverges everywhere in (D, Θ) space in the special cases of $\Gamma = \pm 1/5$. Bifurcations at $\Gamma = 1/5, 0, -1/5$ change the morphology of the (D, Θ) plane.

Plugging this into equation (3.11) allows us to eliminate ω from dj/dt , turning it into an equation for j only. Factorizing the result and multiplying both sides by j we arrive at a simple equation for the rate of change of j^2 :

$$\frac{dj^2}{dt} = \pm \frac{12C}{L} \sqrt{(25\Gamma^2 - 1)(j_0^2 - j^2)(j_+^2 - j^2)(j^2 - j_-^2)}. \quad (3.30)$$

The square root here is well defined because for $|\Gamma| > 1/5$ and $< 1/5$ the signs of bracketed terms change in such a way that the whole expression under the square root is positive⁴, as can be checked using the results collected in Table 3.1 (also see Figure 3.2).

The maximum and minimum j reached by a given phase-space orbit and satisfying the constraint (3.9) are denoted $j_{\min/\max}$. They correspond to two of the three possible roots j_{\pm}, j_0 depending on the orbit type and the value of Γ , as we will see in §§3.3-3.6. Regardless of their precise values, an entire oscillation runs from j_{\min} to j_{\max} and back again, so an entire secular oscillation takes

$$t_{\text{sec}} = 2 \int_{j_{\min}^2}^{j_{\max}^2} \left(\frac{dj^2}{dt} \right)^{-1} dj^2, \quad (3.31)$$

which is expressible in terms of complete elliptic integrals of the first kind⁵ $K(k) = \int_0^{\pi/2} d\alpha / \sqrt{1 - k^2 \sin^2 \alpha}$ (Gradshteyn & Ryzhik 2014). Defining⁶

$$\Delta \equiv \max[j_+^2, j_-^2, j_0^2] - \min[j_+^2, j_-^2, j_0^2], \quad (3.32)$$

we find, in general, that

$$t_{\text{sec}} = \frac{t_1}{\sqrt{|1 - 25\Gamma^2| \Delta}} K \left(\sqrt{\frac{j_{\max}^2 - j_{\min}^2}{\Delta}} \right), \quad (3.33)$$

where

$$t_1 \equiv \frac{8}{3A} \sqrt{\frac{G(m_1 + m_2)}{a^3}} \quad (3.34)$$

$$= 1.7 \text{ Gyr} \times \left(\frac{A^*}{0.5} \right)^{-1} \left(\frac{\mathcal{M}}{10^5 M_{\odot}} \right)^{-1} \left(\frac{b}{\text{pc}} \right)^3 \left(\frac{m_1 + m_2}{M_{\odot}} \right)^{1/2} \left(\frac{a}{10 \text{ AU}} \right)^{-3/2} \quad (3.35)$$

is the characteristic secular timescale. In the numerical estimate we assumed the binary orbits a spherical cluster with scale radius b and total mass \mathcal{M} , and defined $A^* \equiv A/(GM/b^3)$. (Recall that maps of A^* for different cluster potentials and binary orbits are presented in Chapter 2). Note that in the LK case, $A = G\mathcal{M}/[2a_g^3(1 - e_g^2)3/2]$, where

⁴For $\Gamma = \pm 1/5$ we have $dj^2/dt = 0$. This reflects the fact that secular evolution stalls and $t_{\text{sec}} \rightarrow \infty$ for $\Gamma = \pm 1/5$ — see equation (3.33) and Figure 3.3.

⁵See Figure 5.4.

⁶Note that, in general, $\min[j_+^2, j_-^2, j_0^2] \neq j_{\min}^2$. This is because $\min[j_+^2, j_-^2, j_0^2]$ can take any value (including negative values, see Figure 3.2), whereas j_{\min} is the physical minimum angular momentum reached by a given binary, which must lie between $\sqrt{\Theta}$ and 1. An analogous statement holds for j_{\max} .

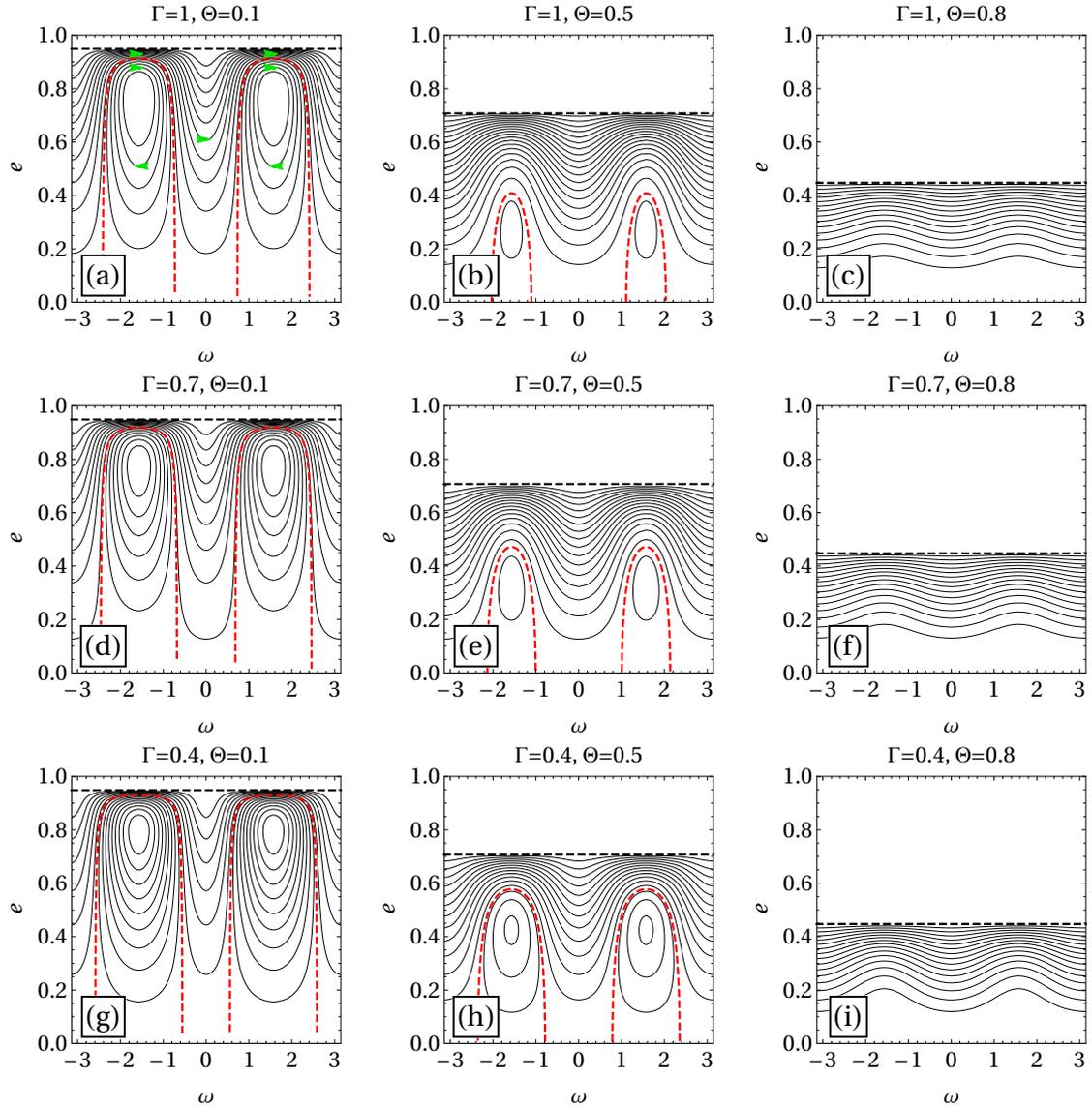


Figure 3.4: Contour plots of constant H_1^* in the (ω, e) phase-space for $\Gamma > 1/5$. Phase portraits are shown for $\Gamma = 1, 0.7, 0.4$ and $\Theta = 0.1, 0.5, 0.8$. Contours are spaced linearly from $H_{1,\text{min}}^*$ to $H_{1,\text{max}}^*$. The black dashed line shows the limiting eccentricity $e_{\text{lim}} = \sqrt{1 - \Theta}$ and the red dashed lines show separatrices between regions of librating and circulating orbits. Green arrows in panel (a) show the direction of phase-space trajectories.

a_g and e_g are respectively the semimajor axis and eccentricity of the outer orbit. Thus one can evaluate any numerical estimate like that above in the LK limit by putting $\Gamma = 1$, $A^* = 0.5$ and $b = a_g(1 - e_g^2)^{1/2}$.

In Chapter 2 we noted that $A \sim 4\pi^2 A^* T_\phi^{-2}$, where $T_\phi = 2\pi(GM/b^3)^{-1/2}$ is the characteristic azimuthal period of the outer orbit of the binary around the cluster. Introducing the period of the inner orbit of the binary $T_b = 2\pi[G(m_1 + m_2)/a^3]^{-1/2}$, we can then use equation (3.34) to estimate t_1 as

$$t_1 \sim \frac{4}{3\pi} \frac{T_\phi^2}{T_b}. \quad (3.36)$$

A similar estimate of the characteristic secular timescale — ratio of the square of the outer orbital period to the inner orbital period — is known to hold for the LK problem (Naoz 2016). Although equation (3.36) was derived for a spherical cluster potential, we also expect the same scaling to hold in (non-spherical) axisymmetric potentials.

The ratio t_{sec}/t_1 is plotted in the (D, Θ) plane for various fixed Γ in Figure 3.3. In each panel, circulating orbits fill the triangle $0 < D < 1 - \Theta$ while regions of librating orbits are shown with white hashing. In §3.2.2 we noted that the four distinct Γ regimes (equations (3.3)-(3.6)) would give rise to different phase-space behaviours. We see in Figure 3.3 that they also give rise to different morphologies of the allowed regions in the (D, Θ) plane, to be discussed in §§3.3-3.6.

The analytic derivation of eccentricity oscillation period (equation (3.33)) was previously done in the LK limit ($\Gamma = 1$) by Vashkov'yak (1999) and Kinoshita & Nakai (2007), and in the $\Gamma = 1/3$ limit by Brassler (2001). Our expression (3.33) generalizes their results for arbitrary external tidal potentials of the type explored in Chapter 2.

Much of §§3.3-3.6 will be focused on deriving the values of $j_{\text{min/max}}$, Δ , and the bounds on Θ , D and H_1^* appropriate to each of the distinct Γ regimes (3.3)-(3.6). For ease of reference all of the results are collected in Tables 3.1 and 3.2, which will be discussed in more detail in §3.9.

3.3 Secular dynamics in the case $\Gamma > 1/5$

In this section we focus exclusively on the case $\Gamma > 1/5$. In short, we find the dynamics in this regime to be qualitatively similar to (but quantitatively different from) the ‘test particle quadrupole’ LK problem. A similar investigation in the LK limit ($\Gamma = 1$) was carried out by Antognini (2015). This regime also covers the case of $\Gamma = 1/3$ relevant for Oort Cloud comets perturbed by the Galactic tide (Heisler & Tremaine 1986).

In Figure 3.4 we plot contours of the dimensionless doubly-averaged perturbing Hamiltonian H_1^* (equation (3.2)) in the (ω, e) plane, with $\Gamma = 1, 0.7, 0.4$ from top to bottom and $\Theta = 0.1, 0.5, 0.8$ from left to right.

Fixed points exist in the left and centre columns, but not in the rightmost column (large Θ) where there are only circulating orbits. Circulating orbits show prograde pericentre precession $\dot{\omega} > 0$, while librating orbits traverse clockwise loops in the (ω, e) plane. As we increase Θ (i.e. move from left to right), the maximum eccentricity reached by the average orbit sharply decreases.

We see that whenever a fixed point is present, the circulating orbits run ‘over the top’ of the librating orbits. As a result, *the eccentricity of a fixed point provides a lower bound on the maximum eccentricity reached by any orbit in the phase-space*⁷. Furthermore, the eccentricity of the fixed point increases slightly as we decrease Γ (move down the page) — see equation (3.12). Thus for Γ close to (but greater than) 0.2 and $\Theta \rightarrow 0$ more binaries may reach very high eccentricities than for $\Gamma \sim 1$.

We now proceed to explain these features quantitatively.

3.3.1 Fixed points and orbit families

Looking at Figure 3.1, we see that whenever $\Gamma > 1/5$, for fixed points to exist (shaded regions) Θ must be less than Λ . However, Λ has a minimum value of $1/2$ in this Γ range (namely as $\Gamma \rightarrow \infty$); hence fixed points *always* exist for $\Gamma > 1/5$ provided we choose Θ small enough, see equation (3.13). The precise requirement is

$$\Theta \in (0, \Lambda), \quad \Gamma > 1/5. \quad (3.37)$$

The range of Θ for which fixed points exist increases as Γ decreases.

This result allows us to understand the lack of librating orbits in panels (c), (f) and (i) in Figure 3.4: their Θ value is too large for the range (3.37). Physically, at large Θ the inclination of the binary with respect to the symmetry plane of the cluster is too low for its tidal field to efficiently torque the binary to high eccentricities. In the LK case $\Gamma = 1$ we recover the classic result that the critical Θ range for fixed points to exist is $\Theta \in (0, 3/5)$. For initially circular binaries the resulting minimum inclination is then $i_{\min} = \cos^{-1}(\sqrt{3/5}) \approx 39.2^\circ$ (see also §3.9.1).

We can convert (3.12) into an eccentricity via $e = \sqrt{1 - j^2}$:

$$e_f = \left[1 - \sqrt{\frac{\Theta}{\Lambda}} \right]^{1/2}. \quad (3.38)$$

⁷This is not true for general Γ — c.f. §3.4.

This helps us to understand why in Figure 3.4, the maximum eccentricity is largest for small Θ , and only weakly dependent on Γ : since each trajectory reaches a maximum eccentricity which is *at least* e_f , decreasing Θ will increase that maximum. And since Λ is a weak function of Γ in this range (taking values $\Lambda \in (1/2, 1)$ — see Figure 3.1), dependence on Γ is not very strong.

3.3.2 Range of parameter values

In §3.2.5 we mentioned that Θ is bounded by equation (3.13) for librating orbits and runs between 0 and 1 for circulating orbits. From equations (3.25)-(3.27) we know that in the regime $\Gamma > 1/5$, the extrema of D correspond to some combination of $D = 0$, $D = 1 - \Theta$ and D_- . It can be shown that D_- is negative all for Θ when $\Gamma > 1/5$, while $D = 1 - \Theta$ is obviously non-negative. Hence the bounds on D are:

$$D \in \begin{cases} (D_-, 0), & \Gamma > 1/5, \text{ librating orbits,} \\ (0, 1 - \Theta), & \Gamma > 1/5, \text{ circulating orbits.} \end{cases} \quad (3.39)$$

These ranges dictate the morphology of the (D, Θ) plane in the top row of Figure 3.3.

It is easy to verify that for $\Gamma > 1/5$ the minimum of H_1^* is attained at $j^2 = \Theta$ (i.e. along the black dashed lines of limiting eccentricity in Figure 3.4), so $H_{1,\min}^*$ is given by equation (3.22). As for $H_{1,\max}^*$, if fixed points exist for a given Γ, Θ , then the Hamiltonian is maximised at the fixed point and $H_{1,\max}^* = H_-$, see equation (3.23). If fixed points do not exist, then $H_{1,\max}^*$ is attained at $j^2 = 1$ (i.e. along the line of zero eccentricity), and hence is given by equation (3.21).

3.3.3 Maximum and minimum eccentricities

For $\Gamma > 1/5$, librating orbits, if they exist, will have minimum angular momentum $j_{\min} = j_-$ and maximum angular momentum $j_{\max} = j_+$, since these are the two solutions at $\omega = \pi/2$, and $j_+ > j_-$ (see Figure 3.4 and equation (3.18)).

Meanwhile, circulating orbits also reach minimum angular momentum at $\omega = \pi/2$ (Figure 3.4), so that their $j_{\min} = j_-$. At the same time, $j_{\max} = j_0$ for circulating orbits since the maximum value of their angular momentum (lowest eccentricity) is reached at $\omega = 0$.

Maximum and minimum eccentricities are then given by $e_{\max/\min} = (1 - j_{\min/\max}^2)^{1/2}$ for both types of phase-space trajectories.

3.3.4 Timescales of eccentricity oscillations

To find the timescale t_{sec} using equation (3.31) we need to know the values of $j_{\text{min/max}}$ (§3.3.3) and Δ for each orbit family. First of all, in the $\Gamma > 1/5$ regime it is clear (equation (3.18)) that we always have $j_+^2 > j_-^2 \geq 0$ (see Figure 3.2 for an illustration).

Librating orbits in this regime have $D < 0$ (see Figure 3.3), so we see from equation (3.14) that $j_0^2 > 1$. Since j_{\pm}^2 provide upper and lower bounds on the true angular momentum j^2 , it must be the case that $j_-^2 < j^2 < j_+^2 < j_0^2$, from which we read off that librating orbits have $\Delta = j_0^2 - j_-^2$.

Meanwhile for circulating orbits $D > 0$ and we know that j^2 is bounded from above by $j_0^2 < 1$, so we must have $j_-^2 < j^2 < j_0^2 < j_+^2$, see Figure 3.2a. Hence $\Delta = j_+^2 - j_-^2$ for circulating orbits.

Using these results we plot the ratio $\log_{10}(t_{\text{sec}}/t_1)$ in (D, Θ) space for $\Gamma = 1, 0.5, 0.25$ in the top row of Figure 3.3. The triangle $0 < D < 1 - \Theta$ contains the circulating orbits in each case; the orbits outside of this triangle librate. The bounds on Θ and D are given by equations (3.37) and (3.39) respectively. The timescale for oscillations is seen to depend primarily on the proximity to the separatrix between librating and circulating orbits at $D = 0$. Along the separatrix the timescale for secular oscillations diverges⁸. (Note, however, that the divergence only occurs in a very narrow region around $D = 0$ — see Figure 1 of Antognini (2015), as well as Appendix 3.B.4). As we decrease Γ from 1, the region containing librating orbits gets larger, though of course the triangle of circulating orbits is unchanged. The value t_{sec}/t_1 is amplified when we decrease Γ , so that the timescale for oscillations increases as we decrease Γ (at fixed A).

For any Γ in the approximate range $0.25 \leq \Gamma \lesssim 0.5$, and sufficiently far from the separatrix, t_1 provides a good estimate of t_{sec} . For $\Gamma \gtrsim 0.5$, large portions of the (D, Θ) space have secular timescales that are shorter than t_1 by a factor of a few. As $\Gamma \rightarrow 0.2$, the timescale diverges everywhere in (D, Θ) space (see equation (3.33)).

All of the results arrived at in this section will change when we leave the regime $\Gamma > 1/5$.

3.4 The case $0 < \Gamma \leq 1/5$

We now turn to the second regime, $0 < \Gamma \leq 1/5$, which is realised quite naturally for example by binaries orbiting close to the core of a spherical cluster (see Chapter 2).

We begin as in §3.3 by showing the phase portraits as one varies Γ and Θ . In Figure 3.5 we plot contours of H_1^* , with $\Gamma = 0.2, 0.1, 0.01$ from top to bottom and $\Theta = 0.1, 0.5, 0.8$

⁸To see this note that as $D \rightarrow 0$, we have $j_+ \rightarrow j_0 = 1$, and the function $K(k)$ diverges logarithmically as $k \rightarrow 1$.

from left to right. The green arrows in panel (d) show the sense in which orbits traverse their trajectories. We immediately note qualitative differences between the plots with $0 < \Gamma \leq 1/5$ (Figure 3.5) and those for $\Gamma > 1/5$ (Figure 3.4). For $\Gamma = 1/5 = 0.2$ only librating orbits exist, as we can see in the top row of Figure 3.5. In panels (d) and (e) we again have fixed points at $\omega = \pm\pi/2$ and librating orbits surrounding them, but note that the circulating orbits which exist for $\Gamma = 0.1, 0.01$ now have eccentricity *minima* at $\omega = \pm\pi/2$ and *maxima* at $\omega = 0$, which is the opposite of the $\Gamma > 1/5$ case. In the (ω, e) phase plane, circulating orbits now run ‘underneath’ the librating orbits, whereas for $\Gamma > 1/5$ they ran ‘over the top’. As a result, *fixed points no longer provide a lower bound on the maximum eccentricity*. The librating orbits still run clockwise but the circulating orbits now display retrograde precession, $\dot{\omega} < 0$, whereas in the $\Gamma > 1/5$ case we had $\dot{\omega} > 0$.

3.4.1 Fixed points and orbit families

Figure 3.1 shows that for $\Gamma = 1/5$ we have $\Lambda = \Lambda^{-1} = 1$. This means that all $\Theta \in (0, 1)$ accommodate librating trajectories, and nothing circulates. Moving to lower Γ in Figure 3.1 one finds $\Lambda^{-1}(\Gamma) < \Lambda(\Gamma)$, so that for fixed points and librating trajectories to exist Θ must now be less than Λ^{-1} (i.e. the shaded region is now bounded by the blue curve):

$$\Theta \in \left(0, \Lambda^{-1}\right), \quad 0 < \Gamma < 1/5. \quad (3.40)$$

The range of Θ for which fixed points exist diminishes as $\Gamma \rightarrow 0$.

The fact that circulating orbits have changed their sense of pericentre precession from prograde to retrograde is easily explained by calculating $d\omega/dt$ (equation (3.10)) at $\omega = 0$. The result is proportional to $(5\Gamma - 1)$, so that $\dot{\omega}$ is positive when $\Gamma > 1/5$ and negative when $\Gamma < 1/5$.

3.4.2 Range of parameter values

We again want to derive the bounds on the (D, Θ) plane and to find the extrema of H_1^* . We begin as in §3.2.5 by considering the limits on D .

First of all, according to its definition (3.17), D diverges when $\Gamma = 1/5$, which is in agreement with the absence of circulating orbits in the top row of Figure 3.5 (circulating orbits require $0 < D < 1 - \Theta$). For $0 < \Gamma < 1/5$ the quantity $10\Gamma/(1 - 5\Gamma)$ is positive; then it follows from equation (3.16) that the minimum value of D for any fixed Θ is zero (attained for $e = 0$). Hence, $D = 0$ is the lower bound. The other possible bounds on D are D_- and $1 - \Theta$ (see equations (3.25)-(3.27)); it turns out that $D_- \geq 1 - \Theta$ for all Θ in this Γ range, so we conclude that:

$$D \in \begin{cases} (1 - \Theta, D_-), & 0 < \Gamma < 1/5, \text{ librating orbits,} \\ (0, 1 - \Theta), & 0 < \Gamma < 1/5, \text{ circulating orbits.} \end{cases} \quad (3.41)$$

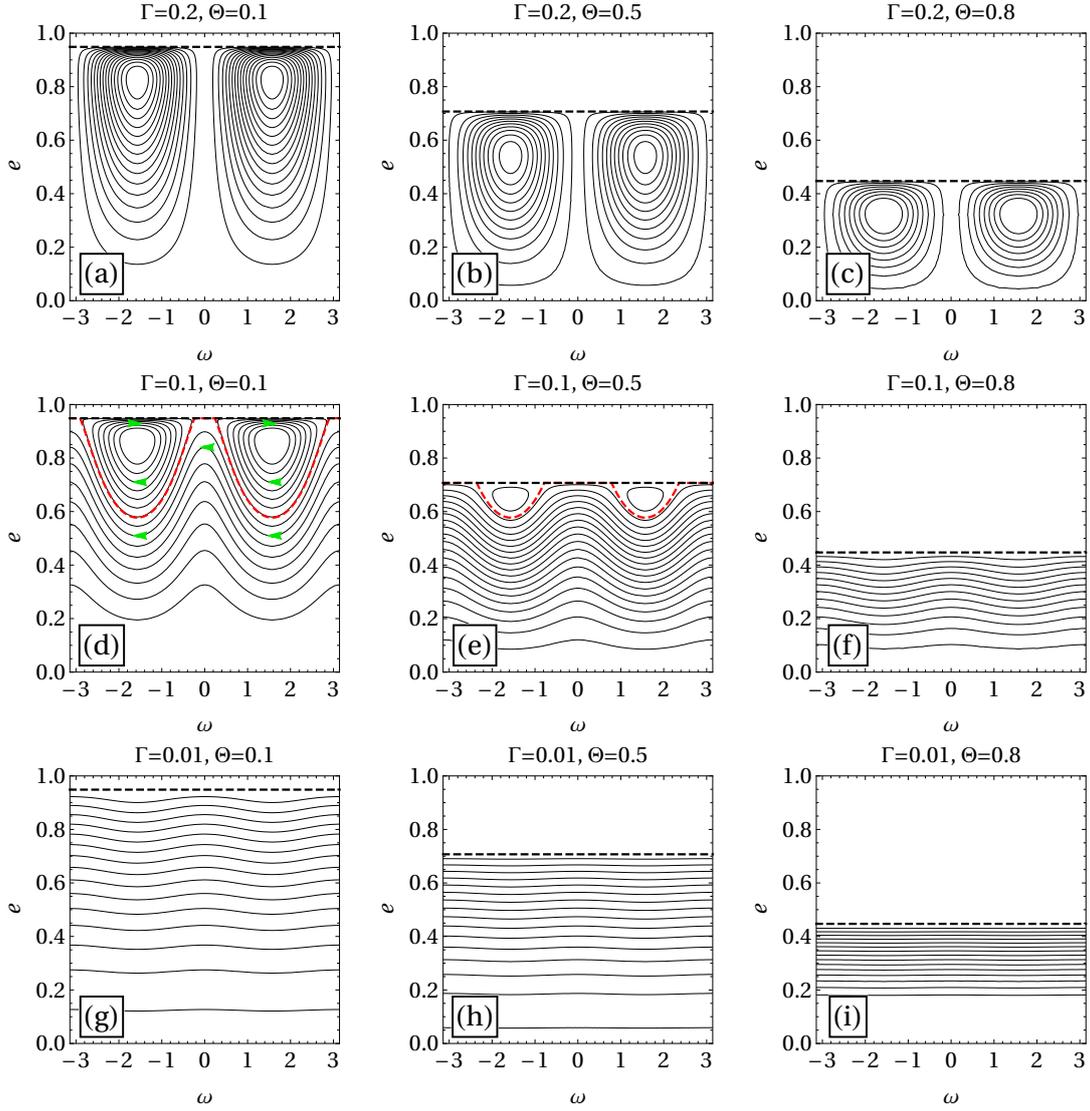


Figure 3.5: Contour plots of constant H_1^* in the (ω, e) plane as in Figure 3.4, but now for the regime $0 < \Gamma \leq 1/5$. Phase portraits are shown for $\Gamma = 0.2, 0.1, 0.01$. Note the different sign of precession of circulating orbits compared to Figure 3.4 as well as the change of morphology: circulating orbits (when they exist, i.e. for $\Gamma \neq 1/5$) now run *below* the islands of libration.

Looking at the timescale plots in the second row of Figure 3.3, we see that the (D, Θ) plane morphology has completely changed compared to $\Gamma > 1/5$ (top row).

This time the minimum of H_1^* is situated at $j^2 = 1$ (i.e. along the line of zero eccentricity in Figure 3.5), so $H_{1,\min}^*$ is given by equation (3.21). If fixed points exist for a given (Γ, Θ) , then $H_{1,\max}^* = H_-$ is found at the fixed point, see equation (3.23). Otherwise $H_{1,\max}^*$ is found on the line $j^2 = \Theta$ (the line of limiting eccentricity), and its value is given by equation (3.22).

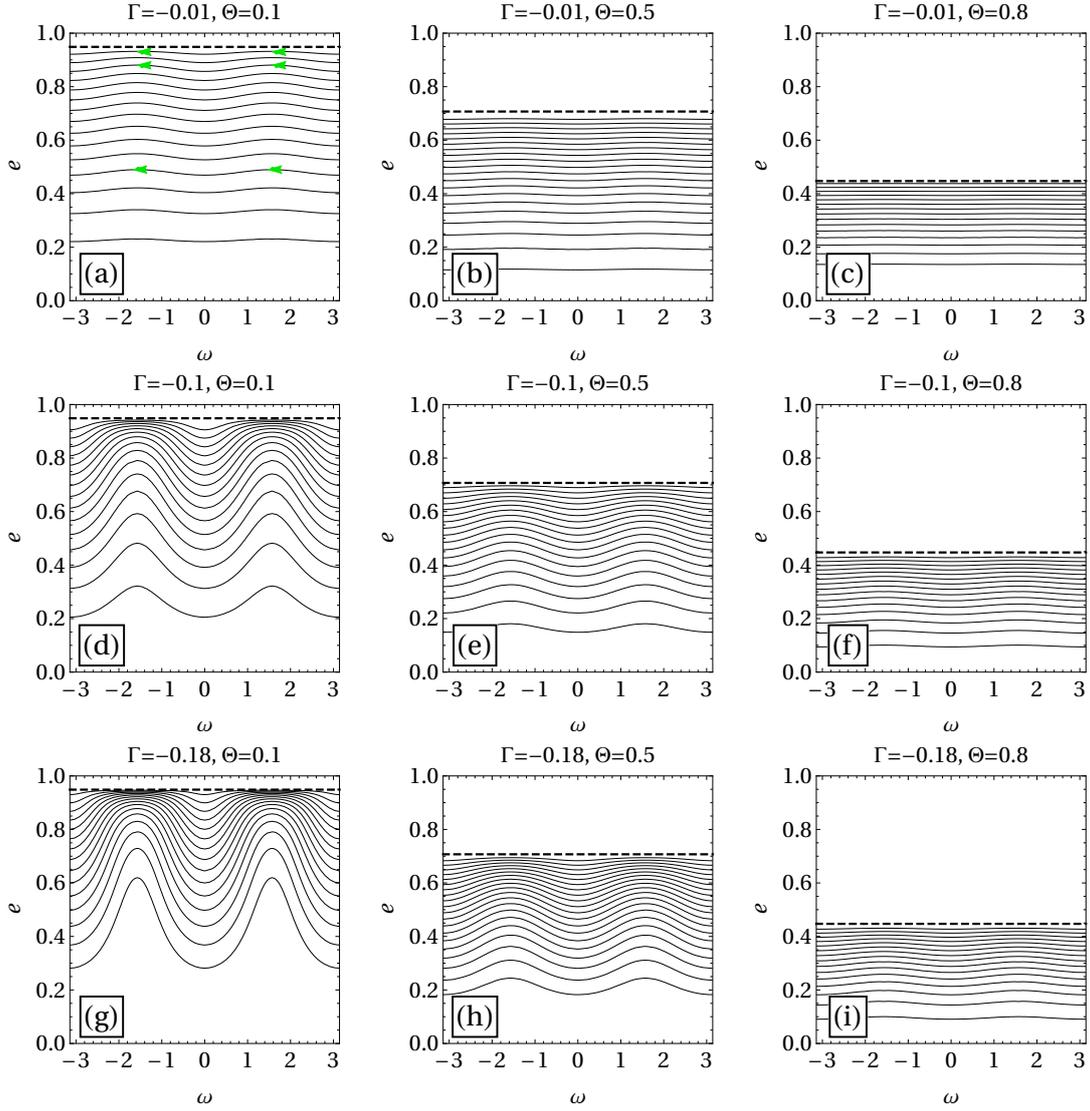


Figure 3.6: Contour plots of constant H_1^* in the (ω, e) plane as in Figure 3.4, but now for the regime $-1/5 < \Gamma \leq 0$. Phase portraits are shown for $\Gamma = -0.01, -0.1, -0.18$. Note the absence of fixed points and librating orbits for all Γ and Θ in this regime.

3.4.3 Maximum and minimum eccentricities

For $\Gamma > 1/5$ it was easy to determine for example that librating orbits have $j_-^2 < j^2 < j_+^2 < j_0^2$. From this we were able to instantly read off $j_{\min/\max} = j_{\pm}$ and $\Delta = j_0^2 - j_-^2$ (§3.3.4). When $\Gamma \leq 1/5$ things become more complicated. While it is possible to calculate the analogous results in each Γ regime algebraically, it is easier and more informative to look at Figure 3.2, in which we plot j_{\pm}^2, j_0^2 as a function of Γ for various fixed (D, Θ) .

Let us consider only Figure 3.2a to begin with, which is for $\Theta = 0.15, D = 0.15$ and therefore sits inside the triangle of circulating orbits ($0 < D < 1 - \Theta$) for any $\Gamma \neq 1/5$. The properties of circulating orbits in each Γ regime can be read off from Figure 3.2a. In the current regime $0 < \Gamma < 1/5$ we see that $j_-^2 < j_0^2 < j_+^2$. Moreover, the horizontal

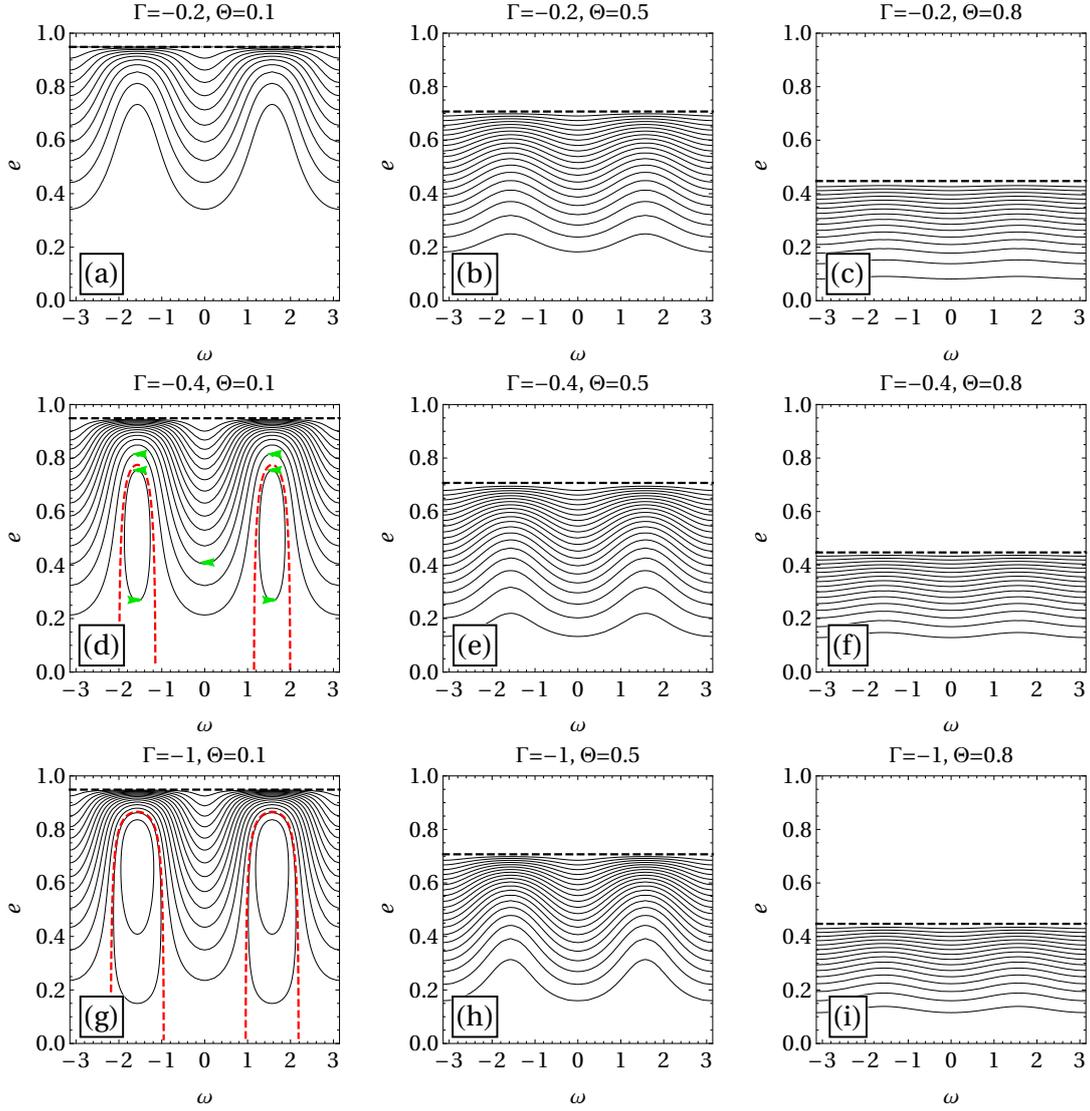


Figure 3.7: Contour plots of constant H_1^* in the (ω, e) plane as in Figure 3.4, but now for the regime $\Gamma \leq -1/5$. Phase portraits are shown for $\Gamma = -0.2, -0.4, -1$.

dotted lines represent $j^2 = \Theta$ and $j^2 = 1$, so physical solutions j^2 must lie between these two lines. Hence j^2 must be bounded by the green and blue lines (that is by j_0^2 and j_+^2). We therefore deduce that $j_{\min} = j_0$, $j_{\max} = j_+$, and $\Delta = j_+^2 - j_-^2$.

This exercise can be repeated with panels (b) and (c) to find the corresponding results for librating orbits. The only subtlety is that as one changes Γ , the region of the (D, Θ) plane corresponding to librating orbits changes shape and moves (this is most clearly seen by comparing different panels along the same row in Figure 3.3). The (D, Θ) points we have chosen for panels (b) and (c) in Figure 3.2 do not *quite* fall inside the region of librating orbits for some Γ because that region becomes too small, which is why for example the red and blue curves in panel (b) are not defined in a small range near $\Gamma = -0.2$. However these plots are good enough for anticipating the results of the

algebraic calculations. When $0 < \Gamma < 1/5$ the j^2 ranges for librating orbits can be read off from Figure 3.2c: for j^2 to lie between the upper and lower horizontal dotted lines we must have $j_0^2 < j_-^2 < j^2 < j_+^2$; as a result, $j_{\min} = j_-$, $j_{\max} = j_+$, and $\Delta = j_+^2 - j_0^2$.

3.4.4 Timescales of eccentricity oscillations

In the second row of Figure 3.3 we present contour plots of $\log_{10}(t_{\text{sec}}/t_1)$ in (D, Θ) space for $\Gamma = 0.15, 0.1, 0.01$. This time the bounds on Θ and D are given by equations (3.40) and (3.41) respectively. The triangle $0 < D < 1 - \Theta$ again contains the circulating orbits, but now the librating orbits have moved to the right of this triangle, and the separatrix corresponds to $D = 1 - \Theta$. Along the separatrix the timescale for secular oscillations again diverges. The timescale is also infinite everywhere in the (D, Θ) plane in the special case $\Gamma = 1/5$ (see equation (3.33)).

For $0 \leq \Gamma \lesssim 0.15$, we see that t_1 again provides a fair estimate of t_{sec} , although it is really a lower bound on t_{sec} in much of the (D, Θ) space, whereas for $\Gamma > 1/5$ it was typically an effective upper bound.

3.5 The case $-1/5 < \Gamma \leq 0$

The two regimes $-1/5 < \Gamma \leq 0$ and $\Gamma \leq -1/5$ cannot be realised by binaries orbiting spherical potentials. In fact, they typically require rather extreme orbits in strongly aspherical potentials. For this reason, here and in §3.6 we simply summarise the qualitative results for each regime — the details are given in Appendices 3.A and 3.B respectively.

The regime $-1/5 < \Gamma \leq 0$ is typically realised when the binary's outer orbit makes large excursions in the Z direction in an oblate potential, i.e. is highly inclined with respect to the potential's symmetry plane (see Chapter 2). Orbits in prolate potentials can also result in this range of Γ .

In Figure 3.6 we plot contours of constant H_1^* for $\Gamma = -0.01, -0.1, -0.18$ from top to bottom and $\Theta = 0.1, 0.5, 0.8$ from left to right. The phase portrait has undergone another bifurcation as we passed through $\Gamma = 0$. We see from Figure 3.6 that only circulating orbits exist in this Γ regime, all with retrograde precession ($\dot{\omega} < 0$). Fixed points and librating trajectories do not emerge at all in this regime, which is explained in Appendix 3.A along with some more technical details.

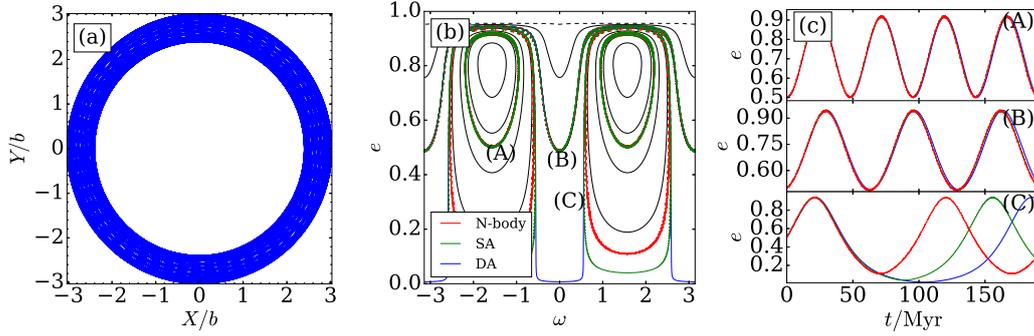


Figure 3.8: Comparison of the evolution of inner binary orbital elements computed via N-body, SA and DA integrations (see legend in panel (b)), over 100 azimuthal periods of \mathbf{R}_g . The binary ($m_1 = m_2 = 0.5M_\odot$) orbits a spherical isochrone cluster (3.42) with total mass $\mathcal{M} = 10^5 M_\odot$ and scale radius $b = 1\text{pc}$. The outer orbit (plotted in the (X, Y) plane in panel (a)) has initial conditions $(R, v_R, Z, v_Z, \phi, v_\phi) = (3.03b, 0, 0, 0, 0, 0.363\sqrt{GM/b})$, giving the theoretical values $A = 0.0206(GM/b^3)$ and $\Gamma = 0.4$. Panels (b) and (c) display the (ω, e) phase-space portrait and time dependence of eccentricity, respectively, for three binaries. Their initial orbital elements are $(a, e, i, \Omega, M) = (10^3\text{AU}, 0.5, 70^\circ, 17.188^\circ, 161.36^\circ)$, with initial ω taking the values (A) 91.67° , (B) 5.73° , (C) 34.14° . See §3.7.2 for discussion.

3.6 The case $\Gamma \leq -1/5$

Situations where $\Gamma \leq -1/5$ may arise, for example, from orbits that are highly inclined with respect to the symmetry plane of a strongly flattened potential. Cylindrical potentials, i.e. ones in which $\Phi(R, Z) = \Phi(R)$ (no Z -dependence, extremely prolate configurations), also fall into this regime as they always have $\Gamma = -1/3$.

In Figure 3.7 we plot contours of H_1^* with $\Gamma = -0.2, -0.4, 1$ from top to bottom and $\Theta = 0.1, 0.5, 0.8$ from left to right. A final bifurcation has occurred as we moved below $\Gamma = -1/5$, such that the phase portrait morphology now looks similar to the case $\Gamma > 1/5$, with fixed points at $\omega = \pm\pi/2$ and circulating orbits running ‘over the top’ of librating orbits in the (ω, e) plane. However, the direction of ω precession is the opposite to the $\Gamma > 1/5$ case: circulating orbits in the $\Gamma \leq -1/5$ regime have retrograde precession ($\dot{\omega} < 0$) while librating orbits run anticlockwise. Appendix 3.B provides additional details on this Γ regime.

This completes our detailed exploration of the dynamical regimes corresponding to the different ranges of Γ (equations (3.3)-(3.6)) considered in this Chapter.

3.7 Accuracy of the doubly-averaged approximation

In Chapter 2 we developed three successive levels of approximation for the evolution of the binary’s inner orbital elements.

First, we had the set of six time-dependent equations for the relative position and velocity of the binary components, subject to the gravitational force of each other and the smooth background potential of the cluster. No tidal approximation had been made at this stage. These six equations can only be solved by direct numerical (‘N-body’) integration.

Second, we had a set of four ‘singly-averaged’ (SA) equations, which were obtained by tidally expanding the six N-body equations, recasting them in Hamiltonian form, and averaging over the binary’s mean anomaly (i.e. over the ‘inner orbit’). The singly-averaged equations are still explicitly time-dependent through the external potential $\Phi(\mathbf{R}_g(t))$, where $\mathbf{R}_g(t)$ is the barycentric position of the binary (assumed to move as a test particle in the cluster potential).

Finally, we derived a system of two ‘doubly-averaged’ (DA) time-independent equations (3.10)-(3.11) resulting from the secular Hamiltonian (3.1), which was itself obtained by averaging the singly-averaged Hamiltonian over many ‘outer orbits’ of the binary around the cluster. Our time-averaging procedure relied on the assumption that $\mathbf{R}_g(t)$ densely fills an axisymmetric torus whose symmetry axis coincides with the symmetry (Z) axis of the potential (in the case of a spherical potential, the ‘torus’ is a two-dimensional annulus perpendicular to Z). More technically, we required that the time-averages of the derivatives $\Phi_{\alpha\beta}$ of the potential (where $\Phi_{xy} = \partial^2\Phi/\partial X\partial Y$, etc.) converge to constant values — in particular, $\overline{\Phi_{xy}} \equiv \overline{\Phi_{xz}} = \overline{\Phi_{yz}} = 0$. This condition is almost always satisfied for orbits in any axisymmetric potential after a sufficient number of outer orbital periods. Then, the time-dependent torque on the binary can be replaced with a converged time-independent torque that arises from an axisymmetric mass distribution as seen from the binary’s frame. In §§3.2-3.6 we examined the dynamics that arise from the doubly-averaged equations (3.10)-(3.11).

However, it is to be expected that the DA theory will break down under certain circumstances. The goal of this section is to explore the validity of the DA approximation for computing inner binary dynamics in the different Γ regimes covered in §§3.3-3.6. To do so, we present several examples comparing numerical integrations of the N-body⁹, SA and DA equations.

We will see that the secular approximation is good as long as the ratio of the secular timescale to the outer orbital period, t_{sec}/T_ϕ , is large enough. Given that $t_{\text{sec}} \sim T_\phi^2/T_b$ (see equations (3.33) and (3.36)), this is equivalent to the statement that T_ϕ/T_b be sufficiently large. Hence, alongside each case where the DA theory fails we present another example with identical initial conditions except for a smaller binary semi-major axis. The effect of this is to decrease the inner binary orbital period T_b , rendering the DA theory valid.

⁹Note the term ‘N-body’ is not meant to imply that we integrate an entire cluster of, say, 10^5 particles. Instead we integrate the exact two-body equations of motion in the presence of the time-dependent external potential calculated via numerical orbit integration of $\mathbf{R}_g(t)$.

3.7.1 Method

We use two particular forms of the background potential Φ in our examples. The first is the spherical isochrone potential

$$\Phi_{\text{iso}}(r) = -\frac{GM}{b + \sqrt{b^2 + r^2}}, \quad (3.42)$$

where $r = \sqrt{X^2 + Y^2 + Z^2}$ is the spherical radius. This is a model potential for a spherical star cluster with total mass \mathcal{M} and scale radius b (Binney & Tremaine 2008). Since the potential (3.42) is spherical we can always choose the plane of the binary's outer orbit \mathbf{R}_g to be the (X, Y) plane. The other potential we will use is the Miyamoto-Nagai potential (Miyamoto & Nagai 1975):

$$\Phi_{\text{MN}}(R, Z) = -\frac{GM}{\sqrt{R^2 + (b_\ell + \sqrt{Z^2 + b_h^2})^2}}, \quad (3.43)$$

where $R = \sqrt{X^2 + Y^2}$ is the usual cylindrical radius. Here M is the total mass of the model, b_ℓ is the scale length and b_h is the scale height. By changing the value of b_h/b_ℓ , the Miyamoto-Nagai potential interpolates between the Kuzmin potential of a razor thin disk ($b_h \ll b_\ell$) and the spherical Plummer potential ($b_h \gg b_\ell$) frequently used to model globular clusters (Binney & Tremaine 2008).

In either case the binary's outer orbit $\mathbf{R}_g(t)$ is stipulated via its initial conditions $(R, v_R, Z, v_Z, \phi, v_\phi)$, where $\phi = \tan^{-1}(Y/X)$ is the azimuthal angle in cylindrical coordinates, v_R is the velocity of the binary in the direction of increasing R , etc. We integrate the orbit $\mathbf{R}_g(t)$ in this potential numerically using `galpy` (Bovy (2015); see §2.F for details), and then feed the resulting time series into the SA and N-body equations¹⁰. The DA equations (3.10)-(3.11) are integrated using the theoretical values of A and Γ when they are available (i.e. in the spherical isochrone case); otherwise we use the numerical prescription outlined in §2.F and denote them by¹¹ $A_{\text{num}}, \Gamma_{\text{num}}$. In all numerical examples the binary has constituent masses $m_1 = m_2 = 0.5M_\odot$.

3.7.2 Accuracy of the doubly-averaged approximation for $\Gamma > 1/5$.

Here we give two examples where the N-body, SA and DA integrations are in very good agreement, both in the regime $\Gamma > 1/5$ (explored in §3.3). In the first the binary orbits the spherical isochrone potential, and in the second it orbits the Miyamoto-Nagai potential. The details of each example are given in the following paragraphs, and the results are shown in Figures 3.8 and 3.9 respectively.

¹⁰Note that in some examples where there is an extremely large separation between the secular timescale and inner orbital period it becomes prohibitively expensive to integrate the N-body equations of motion, so we just show the DA and SA results.

¹¹In spherical potentials, A and Γ can be calculated theoretically by stipulating the outer orbit's peri/apocentre $r_{p/a}$. See Chapter 2 for more information about calculating $A, A_{\text{num}}, \Gamma, \Gamma_{\text{num}}$ and possible small discrepancies between the theoretical and numerical values.

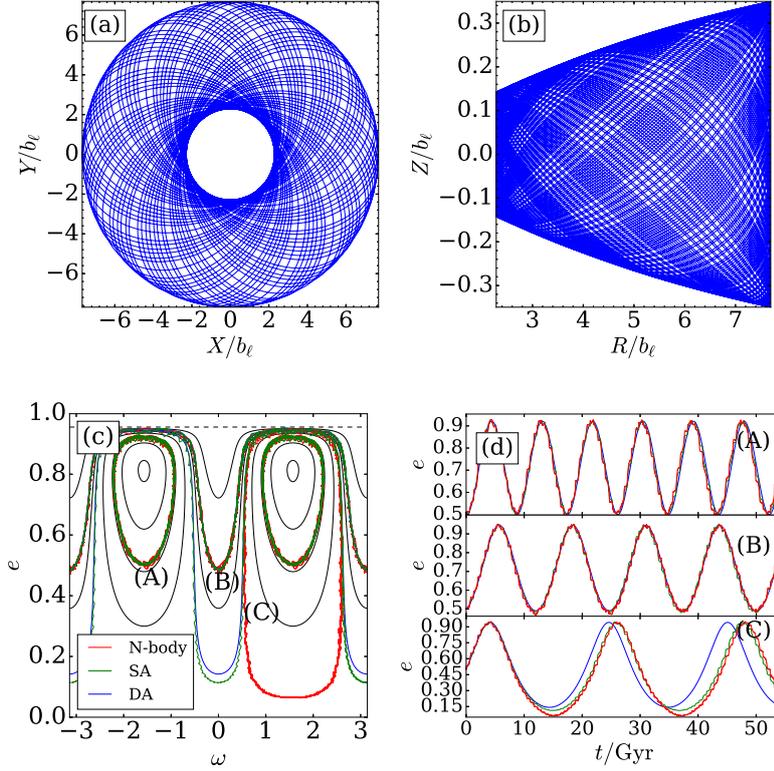


Figure 3.9: Comparison of N-body, SA and DA integrations in a non-spherical cluster. The binary ($m_1 = m_2 = 0.5M_\odot$) orbits a Miyamoto-Nagai potential (3.43) with total mass $\mathcal{M} = 10^{11}M_\odot$ and $b_\ell = b_h = 3.5\text{kpc}$. The outer orbit has initial conditions $(R, v_R, Z, v_Z, \phi, v_\phi) = (2.29b_\ell, 0, 0.143b_\ell, 0, 0, 0.667\sqrt{GM/b_\ell})$, its projections onto the (X, Y) (equatorial) and (R, Z) planes are shown in panels (a) and (b); numerically we find $A_{\text{num}} = 0.0149(GM/b_\ell^2)$ and $\Gamma_{\text{num}} = 0.370$. Panels (c) and (d) are similar to panels (b) and (c) of Figure 3.8. Initial orbital elements are $(a, e, i, \Omega, M) = (5 \times 10^4\text{AU}, 0.5, 70^\circ, 17.188^\circ, 161.36^\circ)$, with initial ω taking the values (A) 91.67° , (B) 5.73° , (C) 30.65° . See §3.7.2 for discussion.

[**Figure 3.8: Isochrone potential, $\Gamma = 0.4$.]** We consider a binary orbiting a spherical isochrone cluster (3.42) with scale radius $b = 1\text{pc}$ and total mass $M = 10^5M_\odot$. The initial conditions for the outer orbit \mathbf{R}_g are as follows:

$$(R, v_R, Z, v_Z, \phi, v_\phi) = (3.03b, 0, 0, 0, 0, 0.363\sqrt{GM/b}). \quad (3.44)$$

It is easy to show that this choice of initial conditions is equivalent to a choice of peri/apocentre $(r_p/b, r_a/b) = (2.40, 3.03)$. The theoretical A, Γ values that result are $A = 0.0206(GM/b^3)$ and $\Gamma = 0.4$. The outer orbit is shown in Figure 3.8a in blue. In panels (b) and (c) we show the (ω, e) phase space evolution and the $e(t)$ timeseries for three characteristic trajectories (A)-(C). The black lines in panel (b) just show the level curves of the underlying doubly-averaged Hamiltonian, like in Figures 3.4-3.7. All three panels in Figure 3.8 show $100T_\phi$ worth of data.

The (rather soft) binary has initial orbital elements

$$(a, e, i, \Omega, M) = (10^3\text{AU}, 0.5, 70^\circ, 17.188^\circ, 161.36^\circ), \quad (3.45)$$

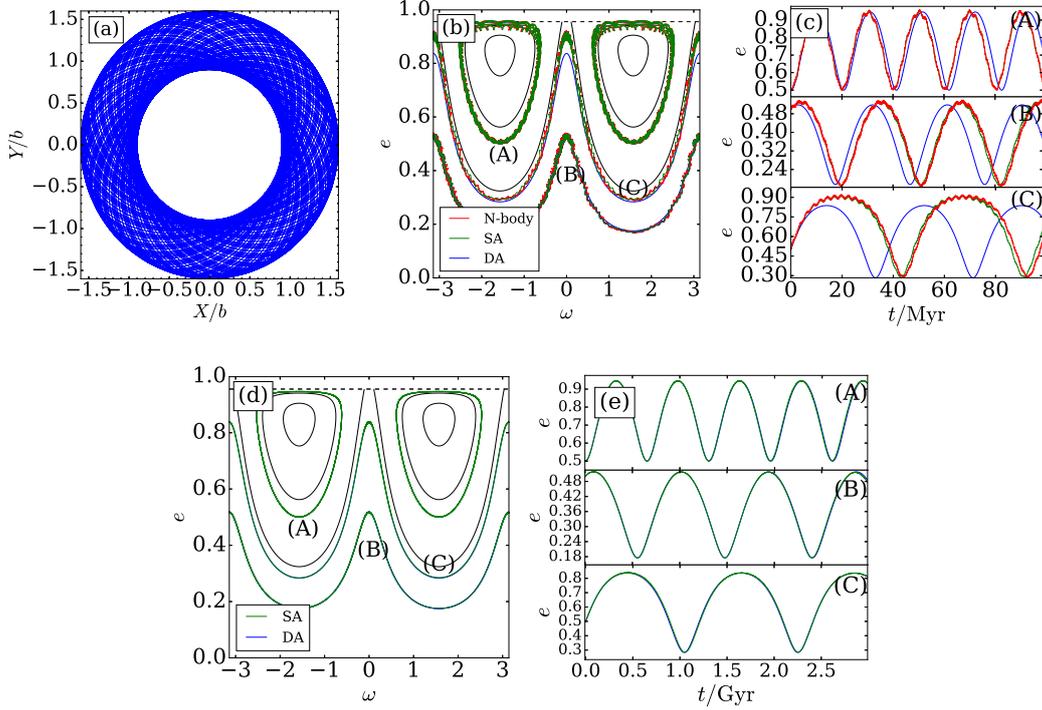


Figure 3.10: Same as Figure 3.8, including the form of the cluster potential, but now focusing on the regime $0 < \Gamma \leq 1/5$. The binary’s outer orbit (shown for $100T_\phi$ in panel (a)) has initial conditions $(R, v_R, Z, v_Z, \phi, v_\phi) = (1.6b, 0, 0, 0, 0, 0.27\sqrt{GM/b})$, resulting in $A = 0.124(GM/b^3)$ and $\Gamma = 0.162$. In panels (b) and (c) the initial binary semi-major axis is $a = 10^3 \text{AU}$ and we integrate the equations of motion for $100T_\phi$, while in panels (d) and (e) it is $a = 100 \text{AU}$ and we integrate for $3000T_\phi$. In each case, the other initial orbital elements are $(e, i, \Omega, M) = (0.5, 70^\circ, 17.188^\circ, 161.36^\circ)$, with initial ω taking the values (A) 91.67° , (B) 5.73° , (C) 34.14° . See §3.7.3 for discussion.

with three different initial values of ω , namely (A) 91.67° (librating phase-space orbit), (B) 5.73° (circulating orbit), and (C) 34.14° (an orbit very close to the separatrix). From Figures 3.8b,c we see that the agreement is extremely good between N-body, SA and DA calculations for phase-space trajectories (A) and (B), even though their secular timescales are longer than T_ϕ by just a factor of 20 – 30. Although initially the agreement in trajectory (C) is also excellent, there is a divergence between N-body, SA and DA calculations when we reach low eccentricity because (C) was chosen to be so close to the separatrix (where the secular timescale formally is infinite — see Figure 3.3). The DA result circulates while the others librate.

[**Figure 3.9: Miyamoto-Nagai potential, $\Gamma_{\text{num}} = 0.370$.**] This time the binary orbits the Miyamoto-Nagai potential (3.43) with $b_\ell = b_h = 3.5 \text{kpc}$ and total mass $\mathcal{M} = 10^{11} M_\odot$. The initial conditions of the outer orbit are

$$(R, v_R, Z, v_Z, \phi, v_\phi) = (2.29b_\ell, 0, 0.143b_\ell, 0, 0, 0.667\sqrt{GM/b_\ell}). \quad (3.46)$$

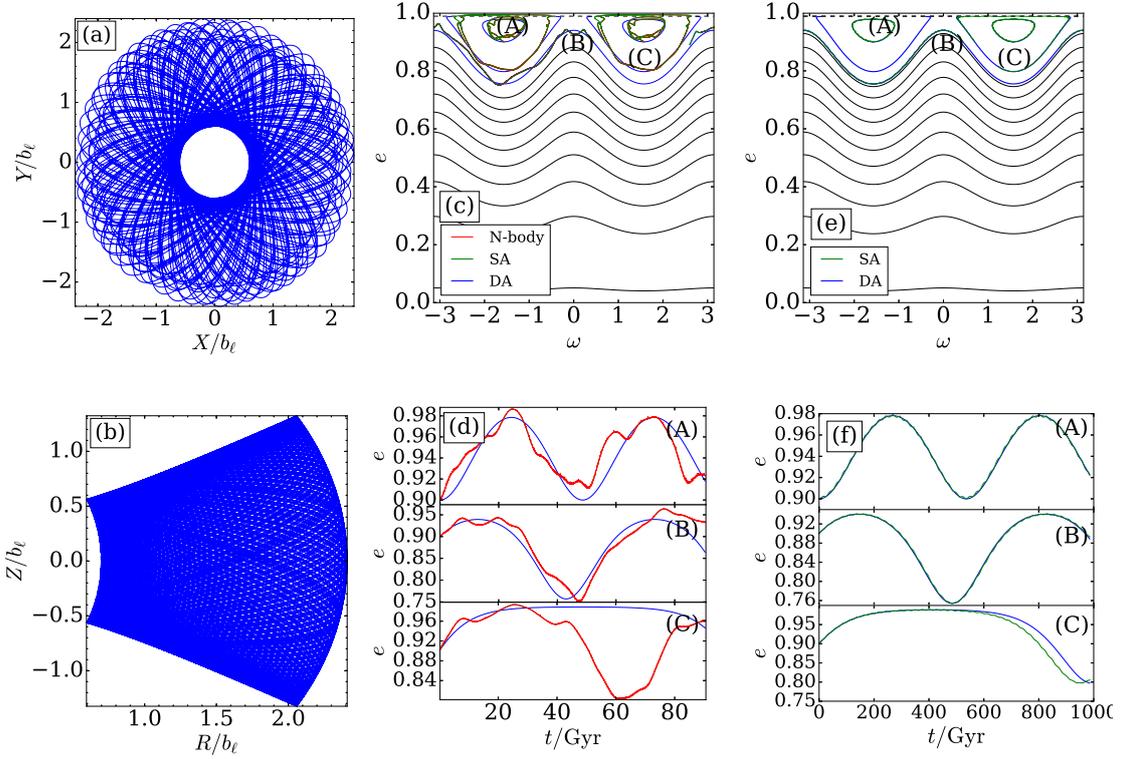


Figure 3.11: Validity of the doubly-averaged secular approximation for a substantially inclined outer orbit of the binary in a non-spherical cluster. The binary ($m_1 = m_2 = 0.5M_\odot$) orbits a Miyamoto-Nagai potential (3.43) with total mass $\mathcal{M} = 10^{11}M_\odot$ and $b_\ell = 3.5\text{kpc}$ and $b_h/b_\ell = 5$ (less flattened than in Figure 3.9). The outer orbit (shown for $100T_\phi$ in panels (a) and (b)) has initial conditions $(R, v_R, Z, v_Z, \phi, v_\phi) = (2.29b_\ell, 0.05\sqrt{GM/b_\ell}, 0.143b_\ell, 0.05\sqrt{GM/b_\ell}, 0, 0.05\sqrt{GM/b_\ell})$; numerically we find $A_{\text{num}} = 0.00826(GM/b_\ell^3)$ and $\Gamma_{\text{num}} = 0.045$. In panels (c) and (e) the initial binary semi-major axis is $a = 5 \times 10^4\text{AU}$ and we integrate the equations of motion for $100T_\phi$, while in panels (d) and (f) it is $a = 10^4\text{AU}$ and we integrate for $1000T_\phi$. Initial orbital elements are $(e, i, \Omega, M) = (0.9, 70^\circ, 17.188^\circ, 161.36^\circ)$, with initial ω taking the values (A) 90° , (B) 24.9° , (C) 40.1° . See §3.7.3 for discussion.

The projections of the outer orbit onto the (X, Y) and (R, Z) planes shown in Figures 3.9a,b. Again all panels show the first $100T_\phi$ of integration time. Since the outer orbit is not planar we do not have theoretical A, Γ values; numerically we find $A_{\text{num}} = 0.0149(GM/b_\ell^3)$ and $\Gamma_{\text{num}} = 0.370$ (not too different from $\Gamma = 1/3$ corresponding to binaries near the midplane of a thin disc).

The binary has initial orbital elements

$$(a, e, i, \Omega, M) = (5 \times 10^4 \text{ AU}, 0.5, 70^\circ, 17.188^\circ, 161.36^\circ), \quad (3.47)$$

and we consider three initial values of ω , namely (A) 91.67° (librating orbit), (B) 5.73° (circulating orbit), and (C) 30.65° (an orbit very close to the separatrix). It is again evident (Figures 3.9c,d) that the agreement is extremely good between N-body, SA and DA calculations for trajectories (A) and (B). The agreement in the eccentricity timeseries

for the separatrix trajectory (C) is also good over the first half-oscillation, but there is then once again a discrepancy in the phase portrait as to whether the orbit ought to librate or circulate. Since the semi-major axis used in this example is typical for Oort Cloud comets, we conclude that the DA approximation should work well for characterising the secular evolution of the long-period comets (Heisler & Tremaine 1986).

In both of these examples the DA theory is very accurate, except for describing phase-space trajectories that lie extremely close to the separatrix between librating and circulating orbits. Note that in each case, 100 outer orbital periods was enough time for at least two secular cycles to take place, so t_{sec}/T_ϕ was at most ~ 50 and usually smaller. We will see in the upcoming sections that in some circumstances a much greater timescale separation t_{sec}/T_ϕ is required for the secular approximation to be valid.

3.7.3 Accuracy of the doubly-averaged approximation for $0 < \Gamma \leq 1/5$.

We now consider the case $0 < \Gamma \leq 1/5$ studied in §3.4. We again provide one example in the spherical isochrone potential and one in the non-spherical Miyamoto-Nagai potential.

This time, in each instance we use two different initial semi-major axes to show how the secular approximation improves for smaller a (shorter T_b).

[Figure 3.10: Isochrone potential, $\Gamma = 0.162$.] The potential is exactly as in the $\Gamma = 0.4$ case discussed in §3.7.2 (Figure 3.8), but now we choose different initial conditions for the outer orbit, namely

$$(R, v_R, Z, v_Z, \phi, v_\phi) = (1.6b, 0, 0, 0, 0, 0.27\sqrt{GM/b}). \quad (3.48)$$

This corresponds to $(r_p/b, r_a/b) = (0.9, 1.6)$, giving the theoretical values $\Gamma = 0.162$ and $A = 0.124(GM/b^3)$. Figure 3.10a shows the outer orbit for the first $100T_\phi$ of integration time.

In Figures 3.10b,c the binary's initial orbital elements are also unchanged from those in Figure 8 (in particular, we still use $a = 10^3\text{AU}$). We again integrate for $100T_\phi$. We see that the agreement between N-body, SA and DA calculations is good for trajectory (A) and reasonable for (B). Trajectory (C), which is very near the separatrix, does not show very good agreement between the DA and other approximations. The reason that the DA theory is so much less accurate in this example than for $\Gamma = 0.4$ (Figure 3.8) — despite having similar ratios of t_{sec}/T_ϕ — is that this time the binary does not fill its annulus fast enough for the time-averaged potential coefficients $\overline{\Phi}_{\alpha\beta}$ to converge rapidly.

To show that the secular approximation can be improved, we rerun integrations of the same three trajectories with exactly the same initial conditions except we use a smaller

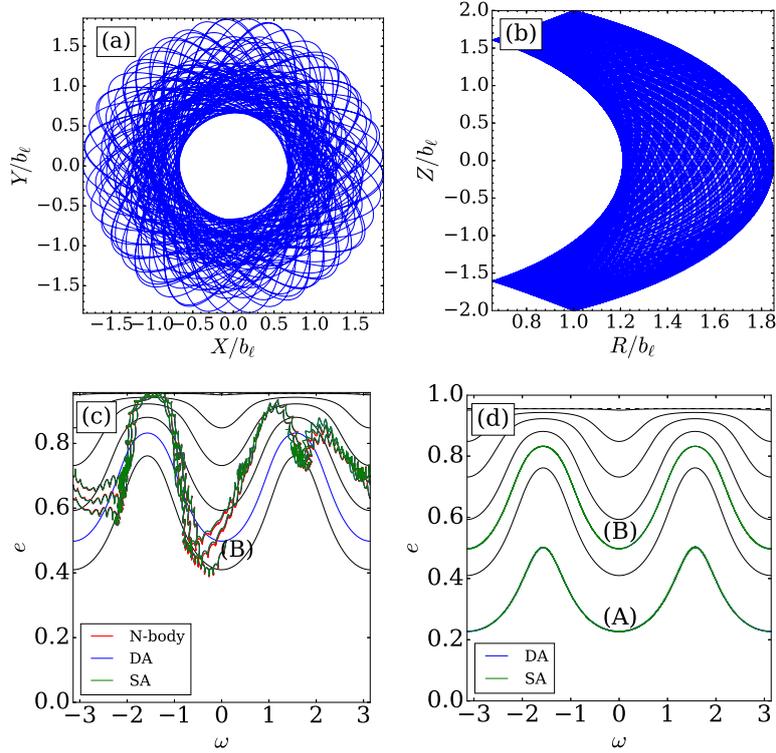


Figure 3.12: Same as Figure 3.9 except that the initial vertical coordinate for the outer orbit is now $Z = 2b_\ell$ (otherwise, the potential and initial conditions of the outer orbit are the same). This gives a negative Γ value, namely $\Gamma_{\text{num}} = -0.163$ and $A_{\text{num}} = 0.0392(GM/b_\ell^3)$. The outer orbit is shown for the first $100T_\phi$ in panels (a) and (b). In panel (c) we integrate the equations of motion for a single phase-space trajectory (‘B’), with $a = 5 \times 10^4 \text{AU}$, for $100T_\phi$. In panel (d) we show this trajectory and another trajectory (‘A’) for an $a = 5 \times 10^3 \text{AU}$ binary over $2000T_\phi$. In all cases the other inner orbital elements of trajectories (A) and (B) are exactly the same as in Figure 3.9, namely $(e, i, \Omega, M) = (0.5, 70^\circ, 17.188^\circ, 161.36^\circ)$, with initial ω taking the values (A) 91.67° , (B) 5.73° . See §3.7.4 for discussion.

semi-major axis, $a = 100 \text{AU}$ (lowering T_b by ~ 30 and, correspondingly, increasing t_{sec} by the same factor). We integrate for $3000T_\phi$. The results are shown in Figures 3.10d,e. The secular timescales are much longer now (from $t_{\text{sec}}/T_\phi \sim 650$ for trajectory (A), to ~ 1300 for (C)). As a result, the binary fills its annulus many times per secular period and the agreement between DA and SA integrations is almost perfect.

[Figure 3.11: Miyamoto-Nagai potential, $\Gamma_{\text{num}} = 0.045$.] The binary orbits the Miyamoto-Nagai potential (3.43) with $b_\ell = 3.5 \text{kpc}$ and $b_h/b_\ell = 5$, which is a less flattened potential than the $b_h/b_\ell = 1$ example from Figure 3.9. The total mass is again $M = 10^{11} M_\odot$. The initial conditions of the outer orbit are

$$(R, v_R, Z, v_Z, \phi, v_\phi) = (2.29b_\ell, 0.05\sqrt{GM/b_\ell}, 0.143b_\ell, 0.05\sqrt{GM/b_\ell}, 0, 0.05\sqrt{GM/b_\ell}). \quad (3.49)$$

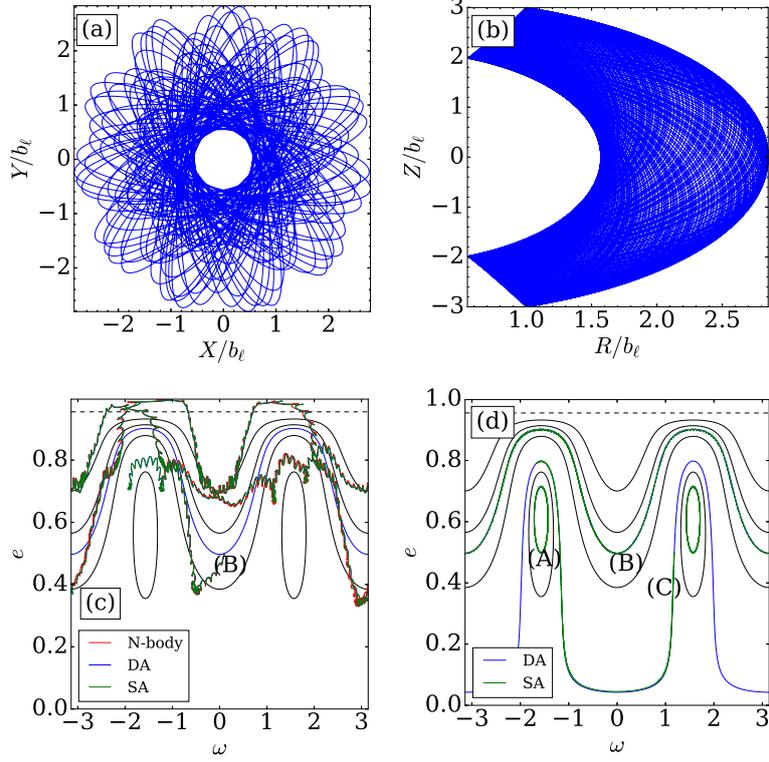


Figure 3.13: Same as Figure 3.9 except that the initial vertical coordinate for the outer orbit is now $Z = 3b_\ell$ (other initial conditions and the potential are the same). This gives a highly inclined outer orbit illustrated in panels (a) and (b) with $\Gamma_{\text{num}} = -0.384$ and $A_{\text{num}} = 0.0142(GM/b_\ell^3)$. In panel (c) we integrate the equations of motion for a single (circulating) phase-space trajectory, with $a = 5 \times 10^4 \text{AU}$, for $100T_\phi$. In panel (d) we show three trajectories for an $a = 5 \times 10^3 \text{AU}$ binary over $2000T_\phi$. In each case the other orbital elements are $(e, i, \Omega, M) = (0.5, 70^\circ, 17.188^\circ, 161.36^\circ)$, with initial ω taking the values (A) 91.67° (librating), (B) 5.73° (circulating), (C) 67.6° (close to the separatrix). See §3.7.4 for discussion.

From Figures 3.11a,b (which both show the first $100T_\phi$ of integration time) we see that \mathbf{R}_g makes large excursions in the Z direction: the binary’s barycentric orbit is about as thick vertically as it is radially. Numerically we find $A_{\text{num}} = 0.00826(GM/b_\ell^3)$ and $\Gamma_{\text{num}} = 0.045$.

In Figures 3.11c,d the binary has initial orbital elements

$$(a, e, i, \Omega, M) = (5 \times 10^4 \text{AU}, 0.9, 70^\circ, 17.188^\circ, 161.36^\circ), \quad (3.50)$$

and three initial values of ω , namely (A) 90° (librating orbit), (B) 24.9° (circulating orbit), and (C) 40.1° (separatrix). We integrate the equations of motion for $100T_\phi$. Although the N-body and SA integrations agree extremely well (the red and green lines in Figures 3.11c,d are almost indistinguishable), the agreement with the DA integration is only reasonable for trajectories (A) and (B) and is poor for trajectory (C). Moreover, the smooth periodicity of the DA solution has disappeared in the SA and N-body integrations, which show chaotic small-scale oscillations. This is unsurprising — despite the fact that

$t_{\text{sec}}/T_\phi \gtrsim 50$ in each case, the binary does not fill its torus densely enough over sufficiently few T_ϕ to render the axisymmetric secular approximation valid.

The DA theory fares much better in Figures 3.11e,f, in which we rerun the same three trajectories except with a smaller semi-major axis, $a = 10^4 \text{AU}$ (i.e. lowering T_b by ~ 10 and increasing t_{sec} by the same amount). We integrate for $1000T_\phi$, so that $t_{\text{sec}}/T_\phi \gtrsim 500$. Of course, the timescales involved here are much longer than the age of the universe and so are not relevant in practice, but this example demonstrates how the secular approximation becomes more accurate when the binary’s outer orbit has a better chance to fill its axisymmetric torus.

3.7.4 Accuracy of the doubly-averaged approximation in the cases $-1/5 < \Gamma \leq 0$ and $\Gamma \leq -1/5$

In this section we present one numerical example in each of the regimes $-1/5 < \Gamma \leq 0$ and $\Gamma \leq -1/5$ (explored in §3.5 and §3.6 respectively). In both cases we use a Miyamoto-Nagai potential with $b_h/b_\ell = 1$; each time we give an example in which the DA theory works very poorly, and one in which it works well.

[Figure 3.12: Miyamoto-Nagai potential, $\Gamma_{\text{num}} = -0.163$.] The potential and outer orbit initial conditions are exactly as in Figure 3.9, except that the initial vertical coordinate is $Z = 2b_\ell$. The resulting (vertically extended) outer orbit (shown in Figures 3.12a,b for the first $100T_\phi$ of integration time) results in a negative Γ value: we find $\Gamma_{\text{num}} = -0.163$ and $A_{\text{num}} = 0.0392(GM/b_\ell^3)$.

For clarity we only show a single phase-space orbit (‘B’) in Figure 3.12c (integrated for $100T_\phi$), which obviously circulates since there are no fixed points in this Γ regime (§3.5). The initial orbital elements for trajectory (B) are identical to those in Figure 3.9 — in particular, we again use $a = 5 \times 10^4 \text{AU}$. From the (ω, e) phase-space portrait (Figure 3.12c) we see that the N-body and SA integrations agree nicely, but the agreement with the DA theory is very poor. Comparing with Figure 3.9, we note that we have caused the DA theory to fail simply by changing a single parameter, the initial Z coordinate of the outer orbit. This is true because the outer orbit is now much more vertically extended and so takes many more outer orbital periods to fill its torus, delaying the convergence of the time averages of the potential derivatives $\Phi_{\alpha\beta}$ that enter the DA Hamiltonian.

Much better agreement is found in Figure 3.12d, in which we use a semi-major axis of $a = 5 \times 10^3 \text{AU}$ (increasing the secular timescale by a factor ~ 30) and integrate for $2000T_\phi$. Other than semi-major axis, trajectory (B) in this figure has the same initial conditions as trajectory (B) in Figure 3.12c. Trajectory (A) differs from (B) only in that it has initial $\omega = 91.67^\circ$.

[Figure 3.13: Miyamoto-Nagai potential, $\Gamma_{\text{num}} = -0.384$.] The potential and the initial conditions of the outer orbit are exactly the same as in Figure 3.12, except the initial Z coordinate is now even larger, $Z = 3b_\ell$, making the orbit very highly inclined. This thickens the outer orbit’s torus further (the first $100T_\phi$ are shown in Figures 3.13a,b) and results in $\Gamma_{\text{num}} = -0.384$ and $A_{\text{num}} = 0.0142(GM/b_\ell^3)$. From panel (a) it is clear that the outer orbit has not come close to filling its torus after $100T_\phi$.

Figure 3.13c shows a trajectory with exactly the same initial conditions as trajectory (B) in Figure 3.12c (i.e. $a = 5 \times 10^4$ AU), integrated for $100T_\phi$. The DA theory is seen to be very inaccurate here. For a better example, in Figure 3.13d we again lower the semi-major axis to $a = 5 \times 10^3$ AU and integrate for $2000T_\phi$. We see that this time (B) is captured almost perfectly by the DA theory, as are two other trajectories, namely (A) (identical initial conditions except $\omega = 91.67^\circ$) and (C) (with $\omega = 67.6^\circ$).

3.7.5 Discussion: validity of the doubly-averaged secular theory

The examples presented in §§3.7.2-3.7.4 illustrate two possible ways in which the DA theory can be in error:

(i) The secular approximation is only a good one if the timescale for evolution of the inner orbital elements is much longer than the time for the time-averages of $\Phi_{\alpha\beta}$ to converge (i.e. for the binary to ‘fill its torus’). Otherwise, the secular approximation can break down and the DA equations can fail completely to describe the evolution.

(ii) The torque experienced by the binary fluctuates on the timescale of its outer orbital period, leading to small fluctuations in the orbital elements on this timescale. Even when the DA equations provide a good description of the dynamics on average, they will always fail to resolve these short-timescale fluctuations (which are fully captured at the level of the SA approximation).

So far in this section we primarily explored the validity of the secular approximation in the sense of $\Phi_{\alpha\beta}$ convergence (effect (i)). However, in practice both effects (i) and (ii) are typically present in our calculations and distinguishing them is important.

The best illustration of the two effects operating simultaneously is provided by Figures 3.12c and 3.13c, in which the secular (DA) approximation completely fails to follow the dynamics quantitatively over long timescales. First, the overall shape of each trajectory (the ‘carrier signal’) computed with both the N-body and SA theory does not line up with the DA prediction. This mismatch is entirely the consequence of effect (i). Additionally, the N-body and SA curves also exhibit short-timescale ‘wiggles’ on

top of the smoother ‘carrier’ phase curve. These wiggles are caused by effect (ii), i.e. torque variations along the outer orbit.

Regardless of how good is the secular approximation by the measure of $\Phi_{\alpha\beta}$ convergence, there are always small short-timescale fluctuations due to effect (ii) that the DA theory cannot capture (e.g. see Figures 3.9c,d, 3.10b,c & 3.11c,d). In many applications of the (quadrupole) LK theory effect (i) is largely absent because there the timescale for a binary to effectively ‘fill its torus’ is simply equal to its outer orbital period. The primary deviations from the DA prediction in the LK case are then due to short-timescale fluctuations (effect (ii)), which has been studied widely in this setup (Ivanov et al. 2005; Katz & Dong 2012; Antonini & Perets 2012; Bode & Wegg 2014; Antonini, Murray, et al. 2014; Antognini et al. 2014; Luo et al. 2016; Grishin et al. 2018). While we largely pass over them from now on, the short-period SA fluctuations can become vitally important when eccentricities get very close to 1 (see LK references above). Their magnitude depends strongly on the shape of the outer orbit. Accounting for such fluctuations in a systematic way in the general Γ case is a central focus of Chapter 6.

Focusing now on the issue of the $\Phi_{\alpha\beta}$ convergence, the two examples presented in §3.7.2 (Figures 3.8 and 3.9 — both in the regime $\Gamma > 1/5$) showed very good agreement between DA theory and direct numerical integration, even for binaries whose secular evolution timescales were significantly shorter than $100T_\phi$ (for example, trajectory (A) in Figure 3.8 had $t_{\text{sec}}/T_\phi \sim 25$). However, all other examples required a much larger ratio of t_{sec}/T_ϕ for the DA theory to be rendered accurate (typically \sim a few $\times 100$). This is because the secular approximation is valid only when the timescale for secular evolution is much longer than the time for the binary to fill its torus densely. The number of outer orbital periods required to fill the torus densely depends strongly on the form of the potential and the choice of outer orbit.

The DA approximation is often most easily satisfied (i.e. it works for relatively small values of t_{sec}/T_ϕ) in spherical potentials, because then the ‘torus’ reduces to a two-dimensional annulus (e.g. Figure 3.8). Not only does this decrease the physical volume that must be visited by the outer orbit, but also the derivatives Φ_{xz}, Φ_{yz} automatically vanish and so pose no problem to the convergence. Circular outer orbits, and outer orbits that avoid any central core, tend to fill their annuli particularly efficiently (and typically correspond to $\Gamma > 1/5$; see Chapter 2).

However, many spherical cluster potentials are cored (such as the isochrone and Plummer models), and so binaries that spend significant time near the centre of these clusters (i.e. those with small r_p) experience an almost-harmonic potential. Since orbits in a harmonic potential are closed ellipses, the apsidal precession of such outer orbits can be very slow; this frequently leads to unfilled gaps being left in the annulus even after

$\sim 100T_\phi$ (see Figure 2.7c for an example). As a result, the secular approximation may require relatively large values of t_{sec}/T_ϕ to be valid, as was the case in Figure 3.10. In spherical potentials, small r_p tends to correspond to small (but always positive) Γ , so this issue will arise most often for binaries in the regime $0 < \Gamma \leq 1/5$.

In non-spherical potentials the situation is often worse simply because there is a third dimension of the torus for the outer orbit to fill. In addition, the derivatives Φ_{xz}, Φ_{yz} are no longer identically zero in general, so we must wait for them to converge, and this typically takes longer than for the other $\Phi_{\alpha\beta}$ (see §2.7.2). In these potentials the secular approximation is most easily satisfied by binaries on outer orbits that are coplanar or nearly coplanar — then the torus is small in volume, and, if the potential is strongly flattened, the vertical oscillations tend to be very rapid, so the torus is filled efficiently (a good example is Figure 3.9). Orbits confined near the midplane of a strongly flattened potential have $\Gamma \approx 1/3$ (§2.6.3).

On the other hand, when the outer orbit has a large vertical extent, filling a torus takes many more outer orbital periods and hence very large values of t_{sec}/T_ϕ are required for the secular theory to be accurate (e.g. Figures 3.12 & 3.13). This in turn implies that for $\Gamma < 0$, the DA theory may be of limited practical use in certain cases (when t_{sec}/T_ϕ is not large enough) because achieving negative Γ typically requires outer orbits that make large excursions in the Z direction.

3.8 Effect of short-range forces on the cluster-tide driven evolution

So far our secular theory considered only the gravitational tidal effect of a stellar cluster on binary evolution. However in a realistic astrophysical situation there could be other, short-range forces which must be taken into account, particularly at high eccentricity when the pericentre distance becomes small (see §3.9.2). Depending on the type of binary (i.e the masses and sizes of its components) and its semi-major axis these could include (i) prograde precession of the argument of pericentre ω due to general relativity (GR), (ii) precession due to the oblateness or tidal distortions of the binary components, (iii) loss of energy and angular momentum due to gravitational wave emission, (iv) tidal dissipation within the components of the binary, etc. The first two effects are conservative in that they do not change the energy of the system and preserve the binary semi-major axis, while the latter two lead to energy losses and tend to shrink the binary orbit. In this section we will briefly consider (i), namely GR pericentre precession, to gauge its effect on the secular dynamics — a full, detailed study of the effect of GR precession is given in Chapter 4.

We can include GR precession in our doubly-averaged theory by adding the following term to the Hamiltonian (3.1) (Fabrycky & Tremaine 2007):

$$\langle H_{\text{GR}} \rangle_M = -\frac{3G^2(m_1 + m_2)^2}{c^2 a^2 \sqrt{1 - e^2}} = -\frac{3G^4(m_1 + m_2)^4}{c^2 L^3 J}. \quad (3.51)$$

The angle brackets remind us that this term is derived by averaging over the binary's mean anomaly. The Hamiltonian $\langle H_{\text{GR}} \rangle_M$ is independent of the longitude of the ascending node Ω , so the z -component of angular momentum J_z is conserved; hence, $\Theta = (1 - e^2) \cos^2 i$ remains an integral of motion. Another integral of motion is the full perturbation energy $\overline{\langle H_1 \rangle}_M + \langle H_{\text{GR}} \rangle_M$ — the sum of the cluster tide and GR Hamiltonians, equations (3.1) and (3.51).

We put (3.51) into dimensionless form by dividing by $C = Aa^2/8$ (see equation (3.8)). Then we must compare the perturbation H_1^* due to the potential of the cluster to the corresponding dimensionless perturbation due to GR:

$$H_{\text{GR}}^* = -\frac{24G^2(m_1 + m_2)^2}{c^2 Aa^4 \sqrt{1 - e^2}} = -\frac{24G^6(m_1 + m_2)^6}{c^2 AL^8 j} = -\frac{\epsilon_{\text{GR}}}{\sqrt{1 - e^2}} = -\frac{\epsilon_{\text{GR}}}{j}. \quad (3.52)$$

where the relative strength of GR precession compared to the cluster tide is measured by the (not necessarily small) parameter

$$\epsilon_{\text{GR}} \equiv \frac{24G^2(m_1 + m_2)^2}{c^2 Aa^4} \quad (3.53)$$

$$= 0.258 \times \left(\frac{A^*}{0.5}\right)^{-1} \left(\frac{\mathcal{M}}{10^5 M_\odot}\right)^{-1} \left(\frac{b}{\text{pc}}\right)^3 \left(\frac{m_1 + m_2}{M_\odot}\right)^2 \left(\frac{a}{20 \text{ AU}}\right)^{-4}. \quad (3.54)$$

In the numerical estimate we have again assumed that the binary is orbiting a spherical cluster with scale radius b and total mass \mathcal{M} . The parameter ϵ_{GR} represents, up to a constant factor, the ratio of the GR apsidal precession rate for a *circular* binary to the binary precession rate due to the cluster tide.

The prograde pericentre precession rate induced by GR is

$$\dot{\omega}_{\text{GR}} \equiv \frac{C}{L} \frac{\partial H_{\text{GR}}^*}{\partial j} = \frac{3G^{3/2}(m_1 + m_2)^{3/2}}{a^{5/2} c^2 (1 - e^2)}. \quad (3.55)$$

With this we can evaluate the ratio of $\dot{\omega}_{\text{GR}}$ to the precession rate $\dot{\omega}_1$ due to the background cluster tide alone (equation (3.10)):

$$\begin{aligned} \frac{\dot{\omega}_{\text{GR}}}{\dot{\omega}_1} &= \frac{\epsilon_{\text{GR}} j}{6} \left(5\Gamma\Theta - j^4 - 5\Gamma \cos 2\omega(\Theta - j^4)\right)^{-1} \\ &= \frac{\epsilon_{\text{GR}} j}{6\Theta} \left(10\Gamma \sin^2 \omega + \frac{j^4}{\Theta}(5\Gamma - 1)\right)^{-1}. \end{aligned} \quad (3.56)$$

We expect GR effects to be most important at high eccentricity, when $j^2 \sim \Theta \ll 1$. Also, in most cases of interest j_{\min} occurs at $\omega = \pi/2$. Plugging these relations into (3.56) we can evaluate

$$\frac{\dot{\omega}_{\text{GR}}}{\dot{\omega}_1} \approx \frac{\epsilon_{\text{GR}} j}{60\Gamma \Theta}. \quad (3.57)$$

We can ignore GR precession only when $|\dot{\omega}_{\text{GR}}/\dot{\omega}_1| \ll 1$, which requires rather small $\epsilon_{\text{GR}} \ll 60\Gamma\Theta^{1/2}(j/\Theta^{1/2})^{-1}$.

3.9 Discussion

The main result of this Chapter is the unveiling of a variety of new dynamical regimes that characterise the orbital evolution of a binary system subject to an external gravitational tidal field (‘cluster’). While the results of §3.2 are completely general, we found that we need to investigate dynamics in four separate regimes, corresponding to certain ranges of the parameter Γ characterising the external tidal field of the cluster and the binary orbit in it. For $\Gamma > 1/5$ (§3.3), the results were found to be qualitatively very similar to those previously derived in the test particle quadrupole Lidov-Kozai problem, which is recovered exactly by taking $\Gamma = 1$ (Vashkov’yak 1999; Kinoshita & Nakai 2007; Antognini 2015).

However, when leaving the regime $\Gamma > 1/5$ several qualitative differences emerge. The condition for the existence of fixed points changes, as do locations of minimum and maximum eccentricities in the (ω, e) phase-space and the morphology of the (D, Θ) plane; even the very existence of the fixed points and orbits librating around them changes with Γ . The first bifurcation in the qualitative dynamics happens at $\Gamma = 1/5$ but there are others at $\Gamma = 0$ and $\Gamma = -1/5$, so we separately treated the regimes $0 < \Gamma \leq 1/5$ (described in §3.4), $-1/5 < \Gamma \leq 0$ (in §3.5), and $\Gamma \leq -1/5$ (in §3.6).

In Tables 3.1 & 3.2 we collect the results of §§3.3-3.6. In Table 3.1 we provide the locations and values of minimum/maximum j , the values of Δ which enter the secular timescale (3.33), and the allowed ranges of the constants of motion D and Θ , for each family of phase-space orbit and in each Γ regime. Table 3.2 collects the locations and values of the extrema of the dimensionless Hamiltonian H_1^* , depending on the Γ regime and whether or not fixed points exist.

In this Chapter we never explicitly considered the possibility of $\Gamma > 1$ — all our examples were given for $\Gamma \leq 1$. Situations in which Γ exceeds unity are possible. However we found in Chapter 2 that this regime is realised only for rather extreme binary orbits inside the cluster, e.g. close to polar, which justifies our overall neglect of the $\Gamma > 1$ possibility. Also, we should note that none of the results obtained in the $\Gamma > 1/5$ regime (§3.3) explicitly assumed $\Gamma \leq 1$; they also hold when $\Gamma > 1$. The only substantial

difference in this case would be that A becomes negative for $\Gamma > 1$. As a result, maxima and minima of the (dimensional) perturbing Hamiltonian would swap their locations in the (ω, e) phase-space, and the phase-space trajectories would be traversed in the opposite direction compared to the $1/5 < \Gamma \leq 1$ case.

For our doubly-averaged theory to properly characterise binary orbital evolution certain conditions should be met. We already saw in §3.7 that the description based on the doubly-averaged Hamiltonian (3.1) may fail when the secular timescale t_{sec} is not much longer than the period of the outer orbit of the binary T_ϕ . In such cases one should resort to the SA framework (although even this can break down in extreme cases — see §6.5.2). Other phenomena that may affect the orbital evolution of binaries in clusters (on top of the smooth cluster tide-driven secular evolution) are discussed in §3.9.3.

3.9.1 Critical inclination for the existence of fixed points

A classic result of LK theory is the value of the critical initial inclination $i_0 = i_c$, above which fixed points appear in the (ω, e) phase-space. This critical angle marks the onset of large eccentricity oscillations, and a qualitative departure from classical Laplace-Lagrange dynamics. It provides a constraint on which binary orientations can lead to large eccentricity excursions.

Assuming an initial binary eccentricity of zero, we can calculate i_c for general Γ using the conservation of $\Theta = \cos^2 i_0$. The upper bounds on Θ for fixed points to exist ($= 3/5$ in the test particle quadrupole LK case) are given by equations (3.37), (3.40) and (3.65); it then easily follows that the existence of fixed points requires $i_0 > i_c$ where

$$i_c = \begin{cases} \cos^{-1}(\Lambda^{1/2}), & |\Gamma| > 1/5, \\ \cos^{-1}(\Lambda^{-1/2}), & 0 < \Gamma \leq 1/5, \end{cases} \quad (3.58)$$

with $\Lambda(\Gamma)$ given in equation (3.12). There are no fixed points for any initial inclination if $-1/5 < \Gamma \leq 0$ — see §3.5. If the initial eccentricity of the binary e_0 is non-zero, the argument of the \cos^{-1} needs to be additionally divided by $\sqrt{1 - e_0^2}$, further lowering i_c .

We plot i_c as a function of Γ in Figure 3.14. In the LK limit $\Gamma = 1$ we recover the classic result $i_c = \cos^{-1} \sqrt{3/5} \approx 39.2^\circ$. As we decrease Γ from 1, fixed points exist for ever smaller initial inclinations, until i_c reaches zero at $\Gamma = 1/5$; note however that the secular timescale diverges as $|\Gamma| \rightarrow 1/5$. Asymptotes at $\Gamma = -1/5$ and $\Gamma = 0$ ensure that fixed points never exist between those values. As $\Gamma \rightarrow -\infty$ we find $\Lambda \rightarrow 1/2$ and so $i_c \rightarrow 45^\circ$.

Table 3.1: Summary of (ω, j) locations and values of minima and maxima of j , values of Δ , and Θ and D ranges, for different types of orbit in each Γ regime. These determine minimum and maximum eccentricities as well as timescales of secular eccentricity oscillations via equation (3.33).

| Type of orbit | Regime | (ω, j_{\min}) | (ω, j_{\max}) | Δ | $\Theta \in$ | $D \in$ |
|------------------------------------|-------------------------------|----------------------|----------------------|-----------------|-----------------------------------|---------------------|
| Librating, $\Gamma > 1/5$ | $j_-^2 < j^2 < j_+^2 < j_0^2$ | $(\pm\pi/2, j_-)$ | $(\pm\pi/2, j_+)$ | $j_0^2 - j_-^2$ | $(0, \frac{1+5\Gamma}{10\Gamma})$ | $(D_-, 0)$ |
| Circulating, $\Gamma > 1/5$ | $j_-^2 < j^2 < j_0^2 < j_+^2$ | $(\pm\pi/2, j_-)$ | $(0, j_0)$ | $j_+^2 - j_-^2$ | $(0, 1)$ | $(0, 1 - \Theta)$ |
| Librating, $0 < \Gamma \leq 1/5$ | $j_0^2 < j^2 < j_+^2 < j_-^2$ | $(\pm\pi/2, j_-)$ | $(\pm\pi/2, j_+)$ | $j_+^2 - j_0^2$ | $(0, \frac{10\Gamma}{1+5\Gamma})$ | $(1 - \Theta, D_-)$ |
| Circulating, $0 < \Gamma \leq 1/5$ | $j_-^2 < j^2 < j_0^2 < j_+^2$ | $(0, j_0)$ | $(\pm\pi/2, j_+)$ | $j_+^2 - j_-^2$ | $(0, 1)$ | $(0, 1 - \Theta)$ |
| $-1/5 < \Gamma \leq 0$ | $j_-^2 < j_+^2 < j^2 < j_0^2$ | $(\pm\pi/2, j_+)$ | $(0, j_0)$ | $j_0^2 - j_-^2$ | $(0, 1)$ | $(0, 1 - \Theta)$ |
| Librating, $\Gamma \leq -1/5$ | $j_+^2 < j^2 < j_-^2 < j_0^2$ | $(\pm\pi/2, j_+)$ | $(\pm\pi/2, j_-)$ | $j_0^2 - j_+^2$ | $(0, \frac{1+5\Gamma}{10\Gamma})$ | $(D_+, 0)$ |
| Circulating, $\Gamma \leq -1/5$ | $j_+^2 < j^2 < j_0^2 < j_-^2$ | $(\pm\pi/2, j_+)$ | $(0, j_0)$ | $j_-^2 - j_+^2$ | $(0, 1)$ | $(0, 1 - \Theta)$ |

Table 3.2: Summary of (ω, j) locations and values of minima and maxima of H_1^* , depending on Γ and whether fixed points (FPs) exist.

| Category | H_1^* min location | H_1^* min | H_1^* max location | $H_{1,\max}^*$ |
|--|----------------------|------------------------------|----------------------|------------------------------|
| $\Gamma > 1/5$, FPs exist | $j = \sqrt{\Theta}$ | $(5 - 3\Theta)(1 - 3\Gamma)$ | $(\pm\pi/2, j_f)$ | H_- |
| $\Gamma > 1/5$, FPs do not exist | $j = \sqrt{\Theta}$ | $(5 - 3\Theta)(1 - 3\Gamma)$ | $j = 1$ | $2(1 - 3\Theta\Gamma)$ |
| $0 < \Gamma \leq 1/5$, FPs exist | $j = 1$ | $2(1 - 3\Theta\Gamma)$ | $(\pm\pi/2, j_f)$ | H_- |
| $0 < \Gamma \leq 1/5$, FPs do not exist | $j = 1$ | $2(1 - 3\Theta\Gamma)$ | $j = \sqrt{\Theta}$ | $(5 - 3\Theta)(1 - 3\Gamma)$ |
| $-1/5 < \Gamma \leq 0$ | $j = 1$ | $2(1 - 3\Theta\Gamma)$ | $j = \sqrt{\Theta}$ | $(5 - 3\Theta)(1 - 3\Gamma)$ |
| $\Gamma \leq -1/5$, FPs exist | $(\pm\pi/2, j_f)$ | H_+ | $j = \sqrt{\Theta}$ | $(5 - 3\Theta)(1 - 3\Gamma)$ |
| $\Gamma \leq -1/5$, FPs do not exist | $j = 1$ | $2(1 - 3\Theta\Gamma)$ | $j = \sqrt{\Theta}$ | $(5 - 3\Theta)(1 - 3\Gamma)$ |

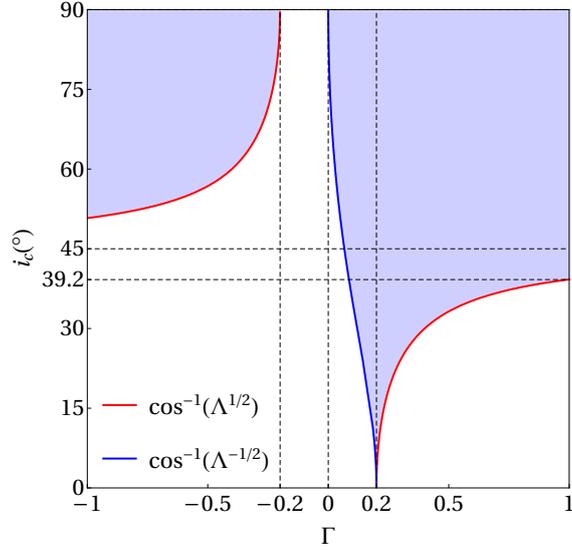


Figure 3.14: Critical inclination i_c assuming zero initial eccentricity (equation (3.58)), as a function of Γ . For initial inclinations greater than i_c (shaded regions), fixed points exist in the (ω, e) phase-space. There are no fixed points in the range $-1/5 < \Gamma \leq 0$. The classic LK result $i_c = 39.2^\circ$ is recovered when $\Gamma = 1$.

3.9.2 High eccentricity behaviour

One is often interested in the high-eccentricity behaviour of binaries undergoing LK-like cycles, because it is at small pericentre distances that exotic effects like GR precession, mass transfer, gravitational wave emission and tidal circularisation become important. To explore these possibilities we consider orbits that are capable of reaching $e \rightarrow 1$ or $j \rightarrow 0$, which necessarily requires

$$\Theta \ll 1, \quad (3.59)$$

(see definitions 3.7) and study their behavior as they evolve through highest eccentricity. We will focus on orbits that start at eccentricities that are not too close to unity, and ignore the effects of GR precession.

We base our discussion on equation (3.30). One of the roots (j_\pm, j_0) corresponds to the smallest value of j satisfying the constraint (3.9), which we have called j_{\min} . Normally, when (3.59) is true, the other two roots are much larger in magnitude. With this in mind equation (3.30) can be approximately integrated in the vicinity of this root j_{\min} as

$$j(t) = \sqrt{1 - e^2(t)} \approx j_{\min} \left[1 + \left(\frac{t}{t_{\min}} \right)^2 \right]^{1/2}, \quad (3.60)$$

where we defined

$$t_{\min} = \frac{L}{6C\sqrt{25\Gamma^2 - 1}} \frac{j_{\min}}{j_1 j_2} = \frac{t_1 \Theta^{1/2}}{2j_1 j_2 \sqrt{25\Gamma^2 - 1}} \frac{j_{\min}}{\Theta^{1/2}}, \quad (3.61)$$

(see equation (3.34)) and j_1 and j_2 are the other two roots of equation (3.30), i.e not j_{\min} (which we normalized by its smallest possible value $\Theta^{1/2}$); $|j_1|, |j_2| \gg j_{\min}$. Time t is counted from the point of reaching highest eccentricity, i.e. $j(t=0) = j_{\min}$, while t_{\min} is the characteristic evolution timescale in the vicinity of j_{\min} — the time it takes for j to change from j_{\min} (at $t=0$) to $2^{1/2}j_{\min}$. Note also that $j(t) \propto t$ when $j \gtrsim j_{\min}$.

The fact that the time spent near maximum eccentricity is of order $t_{\min} \propto j_{\min} = \sqrt{1 - e_{\max}^2}$ is a familiar result from the Lidov-Kozai case (e.g. Miller & Hamilton 2002). It has been used to characterise the timescale for gravitational-wave induced mergers of binaries driven to high eccentricity through the LK effect (e.g. Thompson 2011; Antonini & Perets 2012; Bode & Wegg 2014; Liu & Lai 2018; Grishin et al. 2018; Randall & Xianyu 2018). The more general result (3.60)-(3.61) is applicable for an arbitrary axisymmetric perturbation (not just that of a point mass companion). This formula has to be modified when one includes GR precession (Chapter 4).

Finally, phase space morphology (as determined by the value of Γ) is a crucial factor in determining how many phase space orbits are able to reach $e \rightarrow 1$. For example, provided (3.59) is satisfied, all orbits in the regime $\Gamma > 1/5$ will reach very high eccentricities, whereas this ceases to be true for $0 < \Gamma < 1/5$ (compare Figures 3.4a,d,g with Figures 3.5d,g). This effect is important when calculating merger rates of compact object binaries in stellar clusters (7).

3.9.3 Stellar scattering and other non-ideal effects

Throughout this thesis we assume that the gravitational field of the cluster can be adequately approximated as time-independent. This is of course not true in general. For example, globular clusters in the Milky Way can be shocked and tidally stripped as they move through the Galactic disk. Also, stellar clusters inevitably undergo some secular evolution on \gtrsim Gyr timescales which eventually leads to core collapse. Both of these effects would directly modify the mean field potential Φ and could therefore alter both the outer and inner orbit behavior dramatically.

Additionally, we have assumed that the cluster's potential is perfectly smooth. However, one must remember that the cluster's true potential is in fact the sum of the potentials of the many individual stars comprising it. As a result the true potential felt by the binary is both granular in space and stochastic in time. In practice, these effects can be explored by considering the impact of individual stellar passages in the vicinity of a binary on its orbital elements. The issue of binaries undergoing flyby encounters has been studied widely. Heggie & Rasio (1996) first considered the case of 'secular encounters', where the scattering event takes much longer than the orbital period of the inner binary (Hamers 2018a). This regime is appropriate if one is studying perturbations to the orbits of relatively tight

systems (hard binaries), such as millisecond pulsars or hot Jupiters. On the other hand, Collins & Sari (2008) (see also Collins & Sari 2010) considered the opposite regime in which the timescale for the flyby interaction is much shorter than the inner binary period, so that the encounter can be treated in the impulse approximation. This is the correct description when studying the dynamics of the Oort Cloud comets in the Galaxy or very soft binaries in clusters. Finally, when the approach distance and velocity of the external perturber are comparable to the semi-major axis and the orbital speeds of the binary components, the binary changes its orbital elements in a dramatic fashion on a short (non-secular) timescale, with a high chance of being disrupted (Heggie 1975; Goodman & Hut 1993). This would completely reset the course of the smooth secular evolution of the binary orbit explored in this Chapter. Thus, the prescription needed for estimating the effects of stellar scattering depends on the physical problem one wishes to address.

We defer a careful study of the coupling between the effects of stochastic stellar encounters and the smooth cluster tide-driven evolution of binaries to future work (see §8.2). Here we simply estimate the characteristic time between close encounters of a binary with field stars. Assuming that all perturbers have mass m and can be drawn from an homogeneous, isotropic Maxwellian distribution with number density n and velocity dispersion σ , we can estimate the typical time elapsed before the binary experiences an encounter with impact parameter $q_{\text{enc}} = a/2$ as (Binney & Tremaine 2008):

$$t_{\text{enc}} = \frac{4}{\pi n \sigma a^2} \left(1 + \frac{4G(m_1 + m_2 + m)}{3\sigma^2 a} \right)^{-1} \\ \approx 5 \text{ Gyr} \times \frac{1}{1 + \xi_{\text{GF}}} \left(\frac{n}{10^4 \text{ pc}^{-3}} \right)^{-1} \left(\frac{\sigma}{10 \text{ kms}^{-1}} \right)^{-1} \left(\frac{a}{10 \text{ AU}} \right)^{-2}, \quad (3.62)$$

where

$$\xi_{\text{GF}} \equiv \frac{4G(m_1 + m_2 + m)}{3\sigma^2 a} \\ = 1.2 \times \left(\frac{m_1 + m_2 + m}{M_{\odot}} \right) \left(\frac{\sigma}{10 \text{ kms}^{-1}} \right)^{-2} \left(\frac{a}{10 \text{ AU}} \right)^{-1}, \quad (3.63)$$

is a measure of gravitational focusing, and we have used typical values of n and σ for a globular cluster (although as the binary moves through the cluster, the velocity dispersion and number density of field stars it experiences may change dramatically; this is especially true in the cores of nuclear clusters, which can easily have $n > 10^5 \text{ pc}^{-3}$ and $\sigma > 100 \text{ km s}^{-1}$). One can see that depending on cluster mass, number density, binary semi-major axis, etc., t_{enc} can be larger or smaller than the secular timescale due to cluster tides (equations (3.33) and (3.34)). Moreover, weaker (secular) encounters (Heggie & Rasio 1996) which cause slow random walk of the binary orbital elements would occur more frequently.

Thus, it is usually very important to take into account the effect of stellar flybys. However, we do not believe that this diminishes the astrophysical relevance of cluster

tides, for several reasons. First, tidal effects can be important even in the outskirts of clusters where n is low and stellar encounters are rare. In fact, tides can drive compact object binaries to merge out to cluster-centric distances of several parsecs¹² (Chapter 7). Second, in massive centrally cusped clusters (such as nuclear clusters with or without a central massive black hole), secular timescales due to tides can be as short as $\sim 10^5$ yr, potentially leading to interesting effects before close encounters occur. Third, while dense stellar environments can lead to frequent disruption of binaries they can also result in efficient binary formation, i.e. they can act as a *source* of new binaries that can then undergo tidal evolution.

3.9.4 Relation to previous work

The secular dynamics of binaries presented in this Chapter have been investigated thoroughly by other authors in the LK ($\Gamma = 1$) limit (the ‘test particle quadrupole’ LK problem). In particular, Vashkov’yak (1999) and Kinoshita & Nakai (2007) derived analytically the maximum and minimum eccentricities and the timescale of LK oscillations. Antognini (2015) rederived the same results and provided an approximate fitting formula for the timescale.

A study of the phase-space portrait of binaries perturbed by the Galactic tide — a problem investigated by Heisler & Tremaine (1986) and many others, see Chapter 2 — has been performed by Brassler (2001). Keeping only the $\partial^2\Phi/\partial Z^2$ contribution in the tidal expansion of the potential (equivalent to $\Gamma = 1/3$), they derived the fixed points, secular timescale, and criteria for circulation and libration in (ω, e) space.

Petrovich & Antonini (2017) considered an extension to the LK problem in which a binary orbits a supermassive black hole (SMBH), and its (outer) orbit is perturbed by a non-spherical nuclear cluster potential (the inner orbit is assumed to be unperturbed by the cluster). Unlike our study, Petrovich & Antonini (2017) only looked at the effect of the cluster potential on the outer orbit of the binary and completely ignored the direct effect of the cluster potential on the secular dynamics of the inner orbital elements. Relevant for this Chapter, part of their paper involves an investigation of the (ω, e) phase portrait of the inner binary in the quadrupole approximation, assuming (a) the outer orbit is almost circular and (b) the cluster potential is only weakly flattened. However, our doubly-averaged formalism does not cover this part of Petrovich & Antonini (2017)’s paper, because in this particular limit the outer orbit’s nodal precession timescale is long compared to the secular evolution time, so the perturbing potential cannot be considered axisymmetric (the situation here is similar to that described in §3.7.4).

¹²That is, of course, provided that they do not sink to the centre of the cluster via dynamical friction on a much shorter timescale.

3.10 Summary

In this Chapter we considered the secular dynamics of binaries arising from the general doubly-averaged tidal Hamiltonian derived in Chapter 2. Our study focused on exploring the phase portraits describing the evolution of binaries perturbed by the tidal field of a host cluster. We unravelled a number of new dynamical regimes previously not accounted for in studies of binary evolution problem, and provided their full classification. Our results can be briefly summarized as follows.

- We find that that under a wide range of initial conditions, a generic axisymmetric potential (including spherical potentials) can generate a sufficient tidal torque on a binary to allow it to perform large-amplitude secular eccentricity oscillations reminiscent of the LK mechanism.
- The morphology of the binary evolution in the phase-space of its orbital elements (e.g. ω and e) is uniquely set by the value of a single dimensionless parameter Γ , which encodes all information about the shape of the cluster potential and the binary orbit within it. We mapped out different dynamical behaviours of the binary as a function of Γ .
- Although the dynamics are qualitatively similar to the LK mechanism for $\Gamma > 1/5$, there are bifurcations in the phase-space portrait when $\Gamma = \pm 1/5$ and 0 such that the dynamics become drastically different from LK case. We provide detailed descriptions of the binary evolution in each of the corresponding dynamical regimes.
- We numerically verify our theoretical predictions and find that they work well when the timescale for secular evolution is much longer than the time for the binary's outer orbit to fill an axisymmetric torus inside the cluster. Such circumstances may be rare when $\Gamma < 0$, because this regime typically requires strongly non-coplanar outer orbits that may take large number (several hundred) of orbital periods to fill a torus.
- While the LK mechanism is efficient at driving high eccentricity oscillations, it requires the presence of a long-term distant companion to a binary. In contrast, every binary in a cluster feels the cluster potential, just as every comet feels the Galactic tide. As a result, the cluster tide-driven effect considered in this Chapter, while possibly weaker than in the standard LK scenario, should be more ubiquitous in nature since it is available to any binary bound to an axisymmetric host system.

Appendices

3.A Detailed characteristics of the $-1/5 < \Gamma \leq 0$ regime.

Here we provide more details about the properties of the dynamical regime $-1/5 < \Gamma \leq 0$, see §3.5.

3.A.1 Fixed points and orbit families

When $\Gamma = 0$ the dimensionless Hamiltonian is simply $H_1^* = 2 + 3e^2$, so the phase portrait consists of straight horizontal lines: all orbits circulate with $e(t) = e(0)$.

When $-1/5 < \Gamma < 0$, we use the constraint (3.13) to explore the possibility of fixed points. Since both Λ and Λ^{-1} are negative (see Figure 3.1), while $\Theta > 0$, we conclude there are no fixed points in this Γ regime. As a result, all orbits circulate, in agreement with Figure 3.6.

3.A.2 Range of parameter values

In the absence of fixed points in this Γ regime, the only possible bounds on D are $D = 0$ and $D = 1 - \Theta$. Hence the (D, Θ) plane consists simply of a triangle of circulating orbits (see the third row of Figure 3.3):

$$D \in (0, 1 - \Theta), \quad -1/5 < \Gamma \leq 0. \quad (3.64)$$

It is easy to show that the Hamiltonian is maximised at $j^2 = \Theta$ (i.e. at the upper limit on eccentricity in Figure 3.6), so $H_{1,\max}^*$ obeys equation (3.22). Similarly it is minimised along the line of zero eccentricity ($j^2 = 1$), so $H_{1,\min}^*$ is given by equation (3.21).

3.A.3 Maximum and minimum eccentricities

Figure 3.6 shows that circulating orbits' maximum and minimum eccentricities are back at $\omega = \pm\pi/2$ and $\omega = 0$ respectively, as they were in the $\Gamma > 1/5$ case (§3.3).

To understand the ordering of j_{\pm}^2, j_0^2 we only need to consider panel (a) of Figure 3.2, because we have only circulating orbits in this Γ regime. For $-1/5 < \Gamma \leq 0$ the ordering of j_0^2 and j_+^2 has flipped compared to $0 < \Gamma \leq 1/5$, while j_-^2 still lies outside of the physical region (i.e. it is not bounded by the horizontal dotted lines $j^2 = \Theta, 1$). Hence we must have $j_-^2 < j_+^2 < j^2 < j_0^2$, so that $j_{\min} = j_+$, $j_{\max} = j_0$, and $\Delta = j_0^2 - j_-^2$.

3.A.4 Timescales of eccentricity oscillations

The timescale $\log_{10}(t_{\text{sec}}/t_1)$ is plotted in the third row of Figure 3.3 for $\Gamma = -0.01, -0.1, -0.18$. We have only a triangle of circulating orbits, and their secular timescale is rather well approximated by t_1 .

3.B Detailed characteristics of the $\Gamma \leq -1/5$ regime.

Here we provide more details about the properties of the dynamical regime $\Gamma \leq -1/5$, see §3.6.

3.B.1 Fixed points and orbit families

When $\Gamma = -1/5$ there are still no librating orbits, because $\Lambda(-1/5) = 0$. However, librating orbits emerge as we decrease Γ further. In terms of the constraint on Θ for fixed points and librating orbits to exist, the regime $\Gamma < -1/5$ mirrors the first ($\Gamma > 1/5$) regime in that we again require $\Theta < \Lambda$ (in Figure 3.1, the shaded region is bounded by the red curve):

$$\Theta \in (0, \Lambda), \quad \Gamma < -1/5. \quad (3.65)$$

3.B.2 Range of parameter values

It is perhaps unsurprising from the morphology of the phase portraits (compare Figures 3.4 and 3.7) that the (D, Θ) plane for $\Gamma \leq -1/5$ (bottom row of Figure 3.3) looks similar to the $\Gamma > 1/5$ case (top row of Figure 3.3). However, in this case the librating orbits are bounded to the left by D_+ (see equation (3.27)):

$$D \in \begin{cases} (D_+, 0), & \Gamma < 1/5, \text{ librating orbits,} \\ (0, 1 - \Theta), & \Gamma < 1/5, \text{ circulating orbits.} \end{cases} \quad (3.66)$$

As for the extrema of H_1^* , the only change from the case $-1/5 < \Gamma \leq 0$ is that we now have fixed points available. If they exist then the Hamiltonian is minimised at the fixed point and so $H_{1,\min}^*$ obeys equation (3.23); if not, it is minimised at the zero eccentricity line $j^2 = 1$ (equation (3.21)). The maximum $H_{1,\max}^*$ is always found at $j = \sqrt{\Theta}$ and is therefore given by equation (3.22).

3.B.3 Maximum and minimum eccentricities

For circulating orbits we may again inspect Figure 3.2a. When $\Gamma \leq -1/5$ the physical solutions j^2 run from j_+^2 to j_0^2 as in the $-1/5 < \Gamma < 0$ case, but $j_-^2 > 1$ is suddenly larger than the others, so $j_+^2 < j^2 < j_0^2 < j_-^2$. Thus $j_{\min} = j_+$, $j_{\max} = j_0$ and $\Delta = j_-^2 - j_+^2$.

For librating orbits we read off from Figure 3.2b that $j_+^2 < j^2 < j_-^2 < j_0^2$ (with $j_0 > 1$), giving $j_{\min} = j_+$, $j_{\max} = j_-$ and $\Delta = j_0^2 - j_+^2$.

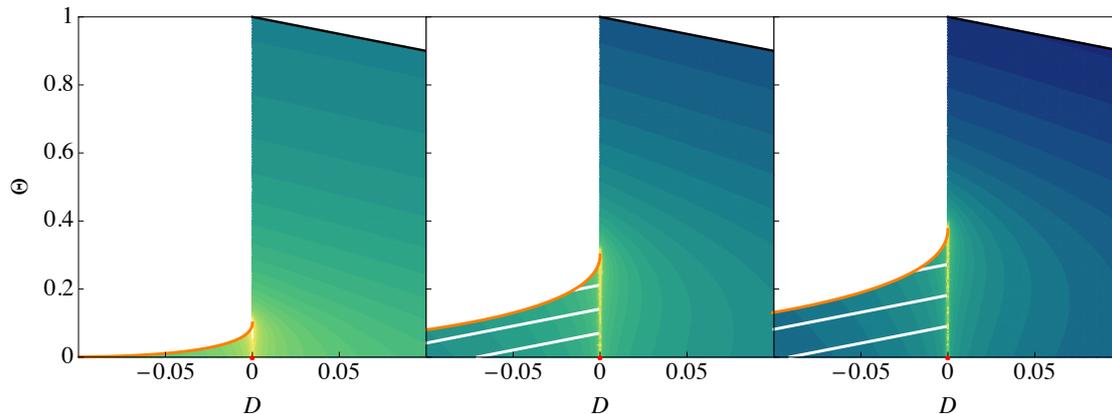


Figure 3.15: Exactly as in the bottom three panels of Figure 3.3, except we show only the region $D \in (-0.1, 0.1)$.

3.B.4 Timescales of eccentricity oscillations

We plot $\log_{10}(t_{\text{sec}}/t_1)$ for $\Gamma = -0.25, -0.5, -0.8$ in the bottom row of Figure 3.3. Bounds on Θ and D are given by equations (3.65) and (3.66) respectively. The separatrix lies along $D = 0$.

Along the separatrix the timescale for secular oscillations once again diverges. To illustrate the nature of the divergence further we present Figure 3.15, in which we plot exactly the same three panels as in the bottom row of Figure 3.3, and with the same colour scale, except we zoom in on the region $D \in (-0.1, 0.1)$. We see that one has to have D extremely close to zero for t_{sec} to be amplified by even an order of magnitude.

The timescale also diverges everywhere in (D, Θ) space for $\Gamma = -1/5$, see equation (3.33). However, as Γ is lowered, one can see that t_{sec} becomes substantially smaller than t_1 , just as in the case of $\Gamma \rightarrow 1$ considered in §3.3.4.

The work presented in this Chapter has been published
in *Monthly Notices of the Royal Astronomical Society* as
Hamilton & Rafikov (2021).

4

The effect of general relativistic precession

Contents

| | | |
|-------------------|--|------------|
| 4.1 | Introduction | 111 |
| 4.2 | Dynamical framework | 113 |
| 4.3 | Phase space behaviour | 115 |
| 4.3.1 | Phase space behaviour in the case $\Gamma > 1/5$ | 118 |
| 4.3.2 | Phase space behaviour in the case $0 < \Gamma \leq 1/5$ | 119 |
| 4.3.3 | Fixed points | 121 |
| 4.3.4 | Determination of the maximum eccentricity of a given orbit | 124 |
| 4.4 | High eccentricity behaviour | 128 |
| 4.4.1 | Phase space behaviour for $\Theta \ll 1$, $\Gamma > 0$ | 128 |
| 4.4.2 | High eccentricity behaviour for $\epsilon_{\text{GR}} = 0$ | 130 |
| 4.4.3 | Modifications brought about by finite ϵ_{GR} | 130 |
| 4.4.4 | High- e behaviour in the weak-to-moderate GR limit | 131 |
| 4.4.5 | High- e behaviour in the strong GR limit | 138 |
| 4.4.6 | Evolution of $\omega(t)$ and $\Omega(t)$ as $e \rightarrow 1$ | 138 |
| 4.5 | Discussion | 139 |
| 4.5.1 | Summary of ϵ_{GR} regimes and their physical interpretation | 140 |
| 4.5.2 | Relation to LK studies | 143 |
| 4.5.3 | Approximations and limitations | 144 |
| 4.6 | Summary | 145 |
| Appendices | | 146 |
| 4.A | Mathematical details of phase space behaviour for $\Gamma > 0$ | 146 |
| 4.A.1 | Fixed points at $\omega = \pm\pi/2$ | 146 |
| 4.A.2 | Fixed points at $\omega = 0$ | 147 |
| 4.A.3 | Does a given orbit librate or circulate? | 148 |
| 4.B | High-eccentricity behaviour for orbits whose eccentricity maxima are found at $\omega = 0$ | 149 |
| 4.C | Analytic solution for orbital elements at high eccentricity | 150 |
| 4.C.1 | Analytic solution in the LK limit | 152 |
| 4.C.2 | Simplified analytic solution in the limit $\epsilon_{\text{GR}} = 0$ | 152 |
| 4.C.3 | Validity of the analytic solution | 156 |
| 4.D | Phase space behaviour and maximum eccentricity in $\Gamma \leq 0$ regimes | 159 |

| | | |
|-------|---|-----|
| 4.D.1 | Phase space behaviour in the case $-1/5 < \Gamma \leq 0$ | 160 |
| 4.D.2 | Phase space behaviour in the case $\Gamma \leq -1/5$ | 163 |
| 4.D.3 | Orbit families and maximum eccentricity for $\Gamma \leq 0$ regimes . . | 164 |

4.1 Introduction

At the very beginning of Chapter 1 we mentioned that the mathematical solution of the two-body problem was obtained for the first time by Newton in 1687. Of course, this solution is only ‘correct’ if one considers purely Newtonian gravity, and ignores general relativistic corrections. In 1915, Einstein updated Newton’s solution by showing that the lowest order correction to the binary’s elliptical orbit (in the small parameter $G(m_1 + m_2)/ac^2$, with m_i the constituent masses and a the binary semimajor axis) was simply an extra prograde apsidal precession at a rate

$$\dot{\omega}_{\text{GR}} = \frac{\dot{\omega}_{\text{GR}}|_{e=0}}{1 - e^2} \quad \text{with} \quad \dot{\omega}_{\text{GR}}|_{e=0} = \frac{3[G(m_1 + m_2)]^{3/2}}{a^{5/2}c^2}, \quad (4.1)$$

where e is the orbital eccentricity, and $\dot{\omega}_{\text{GR}}|_{e=0}$ is the GR precession rate for a circular orbit. Einstein’s solution is now known as the first post-Newtonian (1PN) approximation to the two-body problem.

A century on from Einstein’s discovery, the LIGO/Virgo Collaboration has detected, and continues to detect, dozens of merging compact object (black hole or neutron star) binaries (The LIGO Scientific Collaboration et al. 2019; The LIGO Scientific Collaboration et al. 2020). As reviewed in §1.1.1 these discoveries warrant an astrophysical explanation, which is complicated by the fact that the timescale for an isolated compact object binary to merge via gravitational wave (GW) emission is typically much longer than the age of the Universe. Thus nature must have a way of forcing these relativistic binaries to small separations. The cluster tide-driven eccentricity excitation revealed in Chapter 3 is a candidate mechanism that may help achieve this (a scenario which is further explored in Chapters 5 and 7). LK coupling of a compact object binary with a tertiary perturber is a special case of this theory that has also been widely promulgated as a possible solution to this problem (e.g. Naoz 2016).

However, when investigating such merger channels it is almost always necessary to account for the effect of (1PN) GR precession of the binary’s pericentre angle. This is because GW emission primarily occurs during close pericentre passages when the binary is highly eccentric, and this is precisely the regime in which GR precession is most important (equation (4.1)). For similar reasons it is often necessary to include GR precession (as well as other short-range precession effects, such as those arising from rotational or tidal bulges, see e.g. Liu, Muñoz, et al. (2015) and Muñoz et al. (2016)) in studies of LK secular

evolution, which rely on tidal dissipation (inside one or both binary components) to shrink the binary orbit (Fabrycky & Tremaine 2007; Antonini, Chatterjee, et al. 2016). Tidal dissipation is strongest when $e \rightarrow 1$, meaning that GR precession is important as well.

GR precession is routinely accounted for in LK population synthesis calculations (e.g. Antonini, Murray, et al. 2014; Rodriguez, Morscher, et al. 2015; Liu, Muñoz, et al. 2015; Liu & Lai 2018; Hamers, Bar-Or, et al. 2018a; Samsing et al. 2019). Its effect is understood as increasing a binary’s prograde apsidal precession rate as $e \rightarrow 1$, preventing the perturber from coherently torquing the binary and effectively stopping the reduction of the binary’s angular momentum. A result of this so-called ‘relativistic quenching’ effect is a reduction in the maximum eccentricity e_{\max} that the binary can reach, even if the initial inclination between binary and perturber orbits is favourable (Fabrycky & Tremaine 2007). Some authors have derived approximations to this maximum eccentricity in the limit where GR precession can be treated as a small perturbation to the LK evolution (Miller & Hamilton 2002; Blaes et al. 2002; Wen 2003; Veras & Ford 2010; Liu, Muñoz, et al. 2015; Anderson, Lai, et al. 2017; Grishin et al. 2018). Also, Iwasa & Seto (2016) looked at the modification of the phase space portrait of the LK problem in the presence of GR precession, although their study was far from exhaustive. However, so far nobody has studied carefully the impact of the GR precession for binaries perturbed by general tidal potentials (Brasser et al. 2006; Bub & Petrovich 2020) where we expect similar considerations to apply.

The main purpose of this Chapter is to explore systematically the effect of GR precession on the underlying phase space dynamics and eccentricity evolution of a tidally perturbed binary. We will focus exclusively upon the doubly-averaged (DA) dynamics of binaries perturbed by quadrupole-order tidal potentials. We will also make the test-particle approximation, i.e. assume that the binary’s outer orbital motion relative to its perturber contains much more angular momentum than its internal Keplerian orbit (Naoz 2016). The quadrupolar and test-particle approximations are very good ones for the applications we have in mind (such as compact object binaries perturbed by globular cluster tides), but they can be relaxed, see §4.5.3.

In §4.2 we write down the doubly-averaged perturbing Hamiltonian and establish the notation that we will use for the rest of the paper. In particular we introduce the key parameter ϵ_{GR} which measures the strength of GR precession relative to tidal torques. In §4.3 we explore the phase space behaviour as ϵ_{GR} is varied; the quantitative results that we quote in this section are derived in Appendix 4.A. In §4.4 we investigate very high eccentricity behaviour in the presence of weak or moderate GR precession. In particular we explore how finite ϵ_{GR} modifies both the maximum eccentricity reached and the timescale of high eccentricity episodes. In §4.5 we discuss our results in the light of the existing literature, and comment on the limitations of our study. We summarise in §4.6. Lastly,

in Appendix 4.C we provide for the first time an explicit, analytical solution to the DA equations of motion for all orbital elements in the high eccentricity limit. We also check the accuracy of this solution against direct numerical integration of the DA equations.

4.2 Dynamical framework

As already discussed in §3.8, when GR precession is included the dynamical evolution of the binary's inner orbital elements is governed by the secular ‘doubly-averaged’ perturbing Hamiltonian¹:

$$H = CH^* \equiv C(H_1^* + H_{\text{GR}}^*), \quad \text{where } C \equiv Aa^2/8. \quad (4.2)$$

Here H_1^* and H_{GR}^* are the dimensionless Hamiltonians accounting for quadrupole-order cluster tides and GR pericentre precession, respectively:

$$H_1^* = (2 + 3e^2)(1 - 3\Gamma \cos^2 i) - 15\Gamma e^2 \sin^2 i \cos 2\omega, \quad (4.3)$$

$$H_{\text{GR}}^* = -\epsilon_{\text{GR}}(1 - e^2)^{-1/2}. \quad (4.4)$$

The relative strength of GR precession is measured in equation (4.4) by the crucial parameter

$$\epsilon_{\text{GR}} \equiv \frac{24G^2(m_1 + m_2)^2}{c^2 A a^4} \sim \frac{n_{\text{K}}^2}{A} \left(\frac{v}{c}\right)^2 \sim \dot{\omega}_{\text{GR}}|_{e=0} t_{\text{sec}} \quad (4.5)$$

$$= 0.258 \times \left(\frac{A^*}{0.5}\right)^{-1} \left(\frac{\mathcal{M}}{10^5 M_{\odot}}\right)^{-1} \left(\frac{b}{\text{pc}}\right)^3 \left(\frac{m_1 + m_2}{M_{\odot}}\right)^2 \left(\frac{a}{20 \text{ AU}}\right)^{-4}. \quad (4.6)$$

Here $n_{\text{K}} = \sqrt{G(m_1 + m_2)/a^3}$ and $v \sim \sqrt{G(m_1 + m_2)/a}$ are the Keplerian mean motion and typical orbital speed of the inner orbit of the binary, respectively, while t_{sec} is the timescale of secular eccentricity oscillations in the non-GR limit (equation (3.33)). As usual, in the numerical estimate (4.6) we have assumed that the binary is orbiting a spherical cluster with scale radius b and total mass \mathcal{M} . We reiterate that one can always evaluate any numerical result in the LK limit by setting $\Gamma = 1$, $A^* = 0.5$ and $b = a_{\text{g}}(1 - e_{\text{g}}^2)^{1/2}$. Thus, for illustration, we might consider a black hole binary with $m_1 = m_2 = 30M_{\odot}$, on a circular outer orbit with radius $a_{\text{g}} = 0.1\text{pc}$ around the supermassive black hole at the centre of our Galaxy, which has mass $\mathcal{M} = 4 \times 10^6 M_{\odot}$. In this case $\epsilon_{\text{GR}} \approx 0.02 \times (a/20\text{AU})^{-4}$.

Converting from orbital elements to Delaunay variables we can choose to rewrite the dimensionless Hamiltonians (4.3)-(4.4) in the form

$$H_1^* = \left[(j^2 - 3\Gamma\Theta)(5 - 3j^2) - 15\Gamma(j^2 - \Theta)(1 - j^2) \cos 2\omega \right] j^{-2}, \quad (4.7)$$

$$H_{\text{GR}}^* = -\epsilon_{\text{GR}} j^{-1}. \quad (4.8)$$

¹For simplicity we have replaced the notation $\overline{\langle H_1 \rangle}_M$, $\langle H_{\text{GR}} \rangle_M$ from Chapters 2 and 3 with H_1, H_{GR} .

Since both Hamiltonians are independent of Ω , the dimensionless quantity Θ is an integral of motion. The total dimensionless Hamiltonian $H^* = H_1^* + H_{\text{GR}}^*$ can be taken as the other integral of motion. The equations of motion fully describing the evolution of the dimensionless variables ω , j are

$$\begin{aligned} \frac{d\omega}{dt} &= \frac{C}{L} \frac{\partial H^*}{\partial j} = \frac{C}{L} \frac{\partial}{\partial j} (H_1^* + H_{\text{GR}}^*) \\ &= \frac{6C}{L} \frac{[5\Gamma\Theta - j^4 + 5\Gamma(j^4 - \Theta) \cos 2\omega + \epsilon_{\text{GR}}j/6]}{j^3}, \end{aligned} \quad (4.9)$$

$$\begin{aligned} \frac{dj}{dt} &= -\frac{C}{L} \frac{\partial H^*}{\partial \omega} = -\frac{C}{L} \frac{\partial H_1^*}{\partial \omega} \\ &= -\frac{30\Gamma C}{L} \frac{(j^2 - \Theta)(1 - j^2)}{j^2} \sin 2\omega. \end{aligned} \quad (4.10)$$

Since ω, j are decoupled from Ω , the evolution of the nodal angle Ω can be explored separately using the equation of motion

$$\begin{aligned} \frac{d\Omega}{dt} &= C \frac{\partial H^*}{\partial J_z} = C \frac{\partial H_1^*}{\partial J_z} \\ &= -\frac{6C\Gamma}{L} \Theta^{1/2} \frac{5 - 3j^2 - 5 \cos 2\omega(1 - j^2)}{j^2}. \end{aligned} \quad (4.11)$$

Obviously, the equation of motion for J_z is trivial, $dJ_z/dt = -\partial H/\partial \Omega = 0$.

Given that $H^*(\omega, j)$ is a constant we can use equations (4.7)-(4.8) to eliminate ω from equation (4.10). Following a derivation analogous to that of equation (3.30), and without making any approximations, we find

$$\begin{aligned} \frac{dj}{dt} &= \pm \frac{6C}{Lj^2} \left\{ (25\Gamma^2 - 1) \left[(j_+^2 - j^2)(j^2 - j_-^2) - \frac{\epsilon_{\text{GR}}}{3(1 + 5\Gamma)}j \right] \right. \\ &\quad \left. \times \left[j^2(j_0^2 - j^2) + \frac{\epsilon_{\text{GR}}}{3(5\Gamma - 1)}j \right] \right\}^{1/2}, \end{aligned} \quad (4.12)$$

where

$$j_{\pm}^2 \equiv \frac{\Sigma \pm \sqrt{\Sigma^2 - 10\Gamma\Theta(1 + 5\Gamma)}}{1 + 5\Gamma}, \quad (4.13)$$

$$j_0^2 \equiv 1 - D, \quad (4.14)$$

with

$$\Sigma \equiv \frac{1 + 5\Gamma}{2} + 5\Gamma\Theta + \left(\frac{5\Gamma - 1}{2} \right) D, \quad (4.15)$$

$$\begin{aligned} D &\equiv \frac{H^*/3 - 2/3 + 2\Gamma\Theta}{1 - 5\Gamma} \\ &= e^2 \left(1 + \frac{10\Gamma}{1 - 5\Gamma} \sin^2 i \sin^2 \omega \right) - \frac{\epsilon_{\text{GR}}}{3(1 - 5\Gamma)\sqrt{1 - e^2}}. \end{aligned} \quad (4.16)$$

Note that the definitions of j_{\pm}^2 , j_0^2 , Σ , and D are equivalent to those given in Chapter 3 (equations (3.18), (3.17), (3.20) and (3.16) respectively) except that we have replaced H_1^* in equation (4.16) by $H^* = H_1^* + H_{\text{GR}}^*$, i.e. we have used the value of the Hamiltonian that includes GR precession. Therefore in the limit $\epsilon_{\text{GR}} = 0$, equation (4.12) reduces to equation (3.30). Note also that, just like in Chapter 3, j_{\pm}^2 , j_0^2 are not necessarily positive. We will use equation (4.12) extensively when we study high-eccentricity behaviour in §4.4.

Some shorthand notation will be necessary as we proceed. In particular, several different values of e and j will come with distinct subscripts. We provide a summary of our notation in Table 4.1. A summary of key results derived in this Chapter can also be found in Table 4.2.

We saw in Chapter 2 that to get negative Γ values typically requires a highly inclined outer orbit in a sufficiently non-spherical potential. Since our applications are mostly concerned with spherical or near-spherical potentials such as those of globular clusters, for which negative Γ values are very rare, we concentrate on the $\Gamma > 0$ regimes in the main body of the paper. Discussion of the negative- Γ regimes can be found in Appendix 4.D.

4.3 Phase space behaviour

To gain a qualitative understanding of the dynamics driven by the Hamiltonian equations (4.9)-(4.10), one can fix the values of Γ , Θ , ϵ_{GR} and then plot (ω, e) phase space portraits, i.e. contours of constant H^* in the (ω, e) plane.

Let us briefly recap what we found for $\epsilon_{\text{GR}} = 0$ in Chapter 3. In this case the phase portraits are simply contours of constant H_1^* — see Figures 3.4-3.7. Then, one finds that two distinct phase space orbit families are possible for $\Gamma > 0$: circulating orbits, which run over all $\omega \in (-\pi, \pi)$, and librating orbits, which loop around fixed points located at $(\omega = \pm\pi/2, e = e_f)$, where $e_f \equiv (1 - j_f^2)^{1/2}$ and

$$j_f = \left(\frac{10\Gamma\Theta}{1 + 5\Gamma} \right)^{1/4}, \quad (4.17)$$

see equation (3.12). The precise requirement for fixed points to exist is (§3.2.2):

$$\Theta < \min(\Lambda, \Lambda^{-1}), \quad \text{where} \quad \Lambda(\Gamma) \equiv \frac{5\Gamma + 1}{10\Gamma}. \quad (4.18)$$

For $\Gamma > 0$, fixed points always exist for sufficiently small Θ when $\epsilon_{\text{GR}} = 0$. Neither this nor the existence of fixed points exclusively at $\omega = \pi/2$ are generally true for $\epsilon_{\text{GR}} \neq 0$, as we will see below. In §4.4 we will be interested exclusively in situations where some fraction of binaries can reach eccentricities very close to unity (i.e. $1 - e \ll 1$). Given that $\Theta = (1 - e^2) \cos^2 i$ is conserved, a necessary condition for this is that $\Theta \ll 1$. Finally, a

Table 4.1: Key to different variables and their defining equations.

| Symbol | Description | Def. |
|----------------------------------|---|----------------|
| ϵ_{GR} | Parameter determining strength of GR precession. | (4.5) |
| $\epsilon_{\pi/2}$ | Critical value of ϵ_{GR} dictating the behaviour of fixed points at $\omega = \pm\pi/2$. | (4.19) |
| ϵ_{strong} | Critical value of ϵ_{GR} for the onset of ‘strong’ GR precession, $\epsilon_{\text{strong}} = 3(1 + 5\Gamma)$. | (4.25) |
| ϵ_{weak} | Critical value of ϵ_{GR} defining the upper bound of the ‘weak’ GR regime. | (4.44) |
| e, j | Binary eccentricity, dimensionless angular momentum $j = \sqrt{1 - e^2}$. | (3.7) |
| e_f, j_f | e, j values of fixed points at $\omega = \pm\pi/2$ when $\epsilon_{\text{GR}} = 0$. | (4.17) |
| $e_{f,\pi/2}, j_{f,\pi/2}$ | e, j values of fixed points at $\omega = \pm\pi/2$ when $\epsilon_{\text{GR}} \neq 0$. | (4.63) |
| $e_{f,0}, j_{f,0}$ | e, j values of fixed points at $\omega = 0$, possible only when $\epsilon_{\text{GR}} \neq 0$. | (4.27) |
| $e_{\text{max}}, j_{\text{min}}$ | Maximum e , minimum j values. | |
| e_{lim} | Upper limit on possible values of eccentricity, $e_{\text{lim}} = \sqrt{1 - \Theta}$. | (3.9) |
| e_0, i_0, ω_0 | Initial values of e, i, ω . | |
| j_{\pm}, j_0 | Important functions of $e_0, i_0, \omega_0, \Gamma, \epsilon_{\text{GR}}$ (note j_0 is <i>not</i> the initial j value). | (4.13), (4.14) |

Table 4.2: Key to references for different asymptotic GR regimes and corresponding high eccentricity results.

| | Very weak GR | Weak GR | Moderate GR | Strong GR |
|---|---|---|---|---|
| | $\epsilon_{\text{GR}} \ll \epsilon_{\pi/2}$ | $\epsilon_{\text{GR}} \ll \epsilon_{\text{weak}}$ | $\epsilon_{\text{GR}} \ll \epsilon_{\text{strong}}$ | $\epsilon_{\text{GR}} \gg \epsilon_{\text{strong}}$ |
| j_{min} [if found at $\omega = \pm\pi/2$] | (4.47) | (4.47) | (4.53) | N/A |
| j_{min} [if found at $\omega = 0$] | (4.76) if (4.75) true, else §§4.A.3,4.B | §§4.A.3,4.B | §§4.A.3,4.B | N/A |
| $j(t)$ near $e \rightarrow 1$ | (4.35) | (4.49) | (4.55) | (4.60) |
| $j_{f,\pi/2}$ | (4.20) | (4.21) | (4.21) | none |
| $j_{f,0}$ | (4.27) | (4.27) | (4.27) | none |

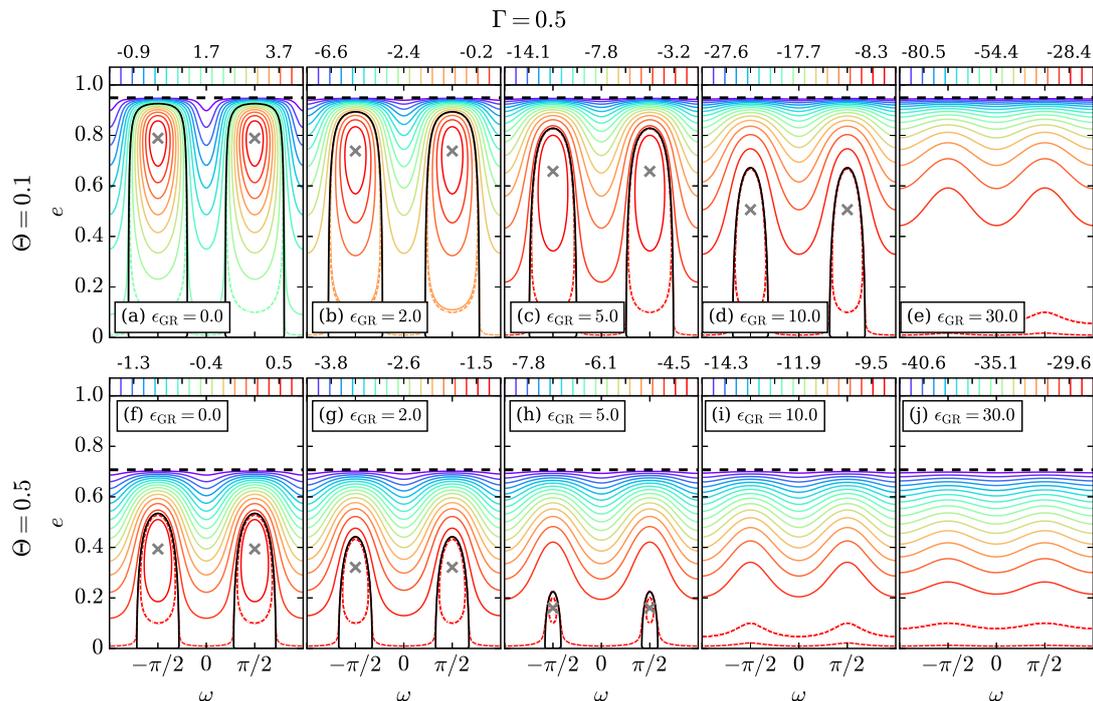


Figure 4.1: Contour plots of constant Hamiltonian $H^* \equiv H_1^* + H_{\text{GR}}^*$ in the (ω, e) plane for $\Gamma = 0.5$. In the top (bottom) row we fix $\Theta = 0.1$ ($\Theta = 0.5$). We increase ϵ_{GR} from left to right, using the values $\epsilon_{\text{GR}} = 0, 2, 5, 10, 30$ indicated in each panel. Contours are spaced linearly from the minimum (blue) to maximum (red) value of H^* — see the colour bar at the top of each panel. We have also added by hand dashed contours passing through $(\omega, e) = (0, 0.01)$ and $(\omega, e) = (\pm\pi/2, 0.1)$ in each panel. Dashed black horizontal lines indicate the limiting possible eccentricity $e_{\text{lim}} = \sqrt{1 - \Theta}$, while fixed points are shown with grey crosses should they exist. Black separatrices illustrate the boundary between families of librating and circulating phase space trajectories.

key result of Chapter 3 (again with $\epsilon_{\text{GR}} = 0$) was that for $\Gamma > 1/5$, whenever fixed points exist, e_f provides a *lower bound* on e_{max} . Equation (4.17) implies that e_f is close to unity whenever $\Theta \ll 1$; high eccentricity excitation is then ubiquitous. On the other hand, for $0 < \Gamma \leq 1/5$ the fixed points no longer provide a lower bound on circulating trajectories' e_{max} and so high- e excitation is much rarer.

In the rest of this section we explore how the phase space behaviour uncovered in Chapter 3 (i.e. for $\epsilon_{\text{GR}} = 0$) is modified in the case of finite ϵ_{GR} , which we do separately for $\Gamma > 1/5$ (§4.3.1) and for $0 < \Gamma \leq 1/5$ (§4.3.2). In §4.3.3 we describe some properties of the fixed points that arise in the phase portraits. In §4.3.4 we show how to calculate the maximum eccentricity of a given binary. Details of the mathematical results that we quote throughout §§4.3.1-4.3.4 are given in Appendix 4.A. Note that the $\Gamma \leq 0$ regimes are treated in Appendix 4.D.

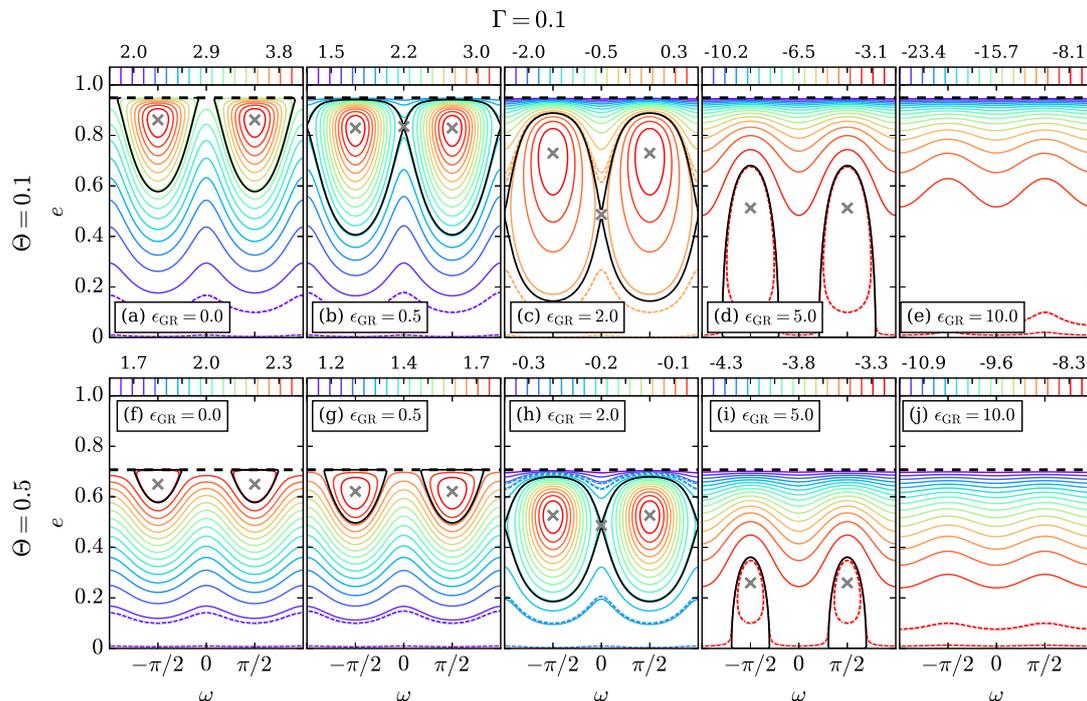


Figure 4.2: Same as Figure 4.1 except for $\Gamma = 0.1$, and we have taken different values of ϵ_{GR} to better demonstrate the new phase space behaviour. Note that dashed contours above the saddle points have the same H^* value as the low- e dashed contours. See text for details.

4.3.1 Phase space behaviour in the case $\Gamma > 1/5$

Figure 4.1 shows phase portraits for the case $\Gamma = 0.5 > 1/5$. In the top (bottom) row we set $\Theta = 0.1$ (0.5). From left to right we vary ϵ_{GR} taking $\epsilon_{\text{GR}} = 0, 2, 5, 10, 30$. In each panel a black horizontal dashed line shows the limiting possible eccentricity $e_{\text{lim}} = \sqrt{1 - \Theta}$ (equation (3.9)). Contours are spaced linearly from the minimum (blue) to maximum (red) value of H^* indicated by the colour bar at the top of each panel. Since the linearly sampled contours become too widely separated at low eccentricity, to illustrate the low- e behaviour we have added dashed contours passing through $(\omega, e) = (0, 0.01)$ and $(\omega, e) = (\pm\pi/2, 0.1)$ in each panel. Just like in Chapter 3, trajectories are split into librating and circulating families. We plot the separatrices between these families with solid black lines. Fixed points are denoted with grey crosses.

In panels (a) and (f) we encounter the usual $\epsilon_{\text{GR}} = 0$ behaviour familiar from the LK problem: (I) there are fixed points² at $\omega = \pm\pi/2$, each of which is surrounded by a region of librating orbits, (II) these librating islands are connected to $e = 0$ line, (III) the family of circulating orbits runs ‘over the top’ of the librating regions, and (IV) all phase space trajectories reach maximum eccentricity at $\omega = \pm\pi/2$.

²We have deliberately chosen Θ values such that the fixed points do exist for $\epsilon_{\text{GR}} = 0$, i.e. satisfying (4.18).

Inspecting the other panels, we see that the effect of increasing ϵ_{GR} from zero is simply to push the fixed points at $\omega = \pm\pi/2$ to lower eccentricity. As a result, large amplitude eccentricity oscillations along a given secular trajectory are noticeably quenched as ϵ_{GR} is increased, and the region of librating orbits is diminished in both area and vertical extent. Eventually the eccentricity of the fixed points reaches zero and so they vanish altogether, leaving only circulating orbits (panels (e), (i) and (j)).

The phase space evolution for non-zero ϵ_{GR} exhibited in Figure 4.1 is characteristic of all systems with $\Gamma > 1/5$, including the LK case $\Gamma = 1$, which has already been discussed to some degree by Iwasa & Seto (2016) — see §4.5.2. Of course, the precise characteristics, such as the eccentricity of the fixed points, the critical ϵ_{GR} for fixed points to vanish, etc., do depend on the value of Γ , as we detail in §4.3.3.

4.3.2 Phase space behaviour in the case $0 < \Gamma \leq 1/5$

For $0 < \Gamma \leq 1/5$ (a regime typical of binaries orbiting the inner regions of a cored cluster), a similar but slightly more complex picture emerges. In Figure 4.2 we show phase portraits similar to Figure 4.1 except that we now take $\Gamma = 0.1 < 1/5$, and pick some new values of ϵ_{GR} to better demonstrate the modified phase space behaviour. The strength of GR still increases from left to right. As in Figure 4.1 we have added in dashed contours that take the values $H^*(\omega = 0, e = 0.01)$ and $H^*(\omega = \pm\pi/2, e = 0.1)$.

Starting with the non-GR case $\epsilon_{\text{GR}} = 0$, we immediately notice a qualitative difference between the phase space morphologies for $0 < \Gamma \leq 1/5$ (Figures 4.2a,f) and $\Gamma > 1/5$ (Figures 4.1a,f), discussed at length in Chapter 3. Although there are again fixed points at $\omega = \pm\pi/2$, the librating islands that surround them are now connected to $e = e_{\text{lim}}$ (and not to $e = 0$, like in the $\Gamma > 1/5$ case). As a result, circulating orbits run ‘underneath’ librating orbits (rather than ‘over the top’ as for $\Gamma > 1/5$) and the maximum eccentricity of circulating orbits is found at $\omega = 0$ (rather than at $\omega = \pm\pi/2$). Crucially, unlike for $\Gamma > 1/5$, a binary that starts at low eccentricity does not necessarily reach a high eccentricity even if there are fixed points located near $e = 1$. This fact is responsible for the dearth of cluster-tide driven mergers in cored clusters, which host many binaries with $0 < \Gamma \leq 1/5$ (Chapter 7).

As we increase ϵ_{GR} from zero, the fixed points again get pushed to lower eccentricity (panels (b) and (g)). However, the effect of this for $\Gamma = 0.1$ is to initially *increase*, rather than decrease, the fraction of the phase space area that is encompassed by the librating islands. Additionally, as the fixed points get pushed to lower eccentricity, a new family of high-eccentricity circulating orbits emerges once ϵ_{GR} exceeds a threshold value which we determine in §4.3.3. These phase space trajectories run ‘over the top’ of the fixed points and have their eccentricity maxima at $\omega = \pm\pi/2$ (panels (b), (c) and (h)). The

qualitative change from $\epsilon_{\text{GR}} = 0$ is reflected in the fact that the librating island is now truly an *island*, disconnected from both $e = 0$ and $e = e_{\text{lim}}$. This is different from the case $\Gamma > 1/5$, in which the lower portion of the librating regions *always* stretches down to $e = 0$ until ϵ_{GR} becomes so large that fixed points cease to exist.

Physically these new features might have been anticipated. First of all, in Chapter 3 we saw that for $0 < \Gamma \leq 1/5$, the cluster-driven ω evolution of circulating trajectories is always retrograde, contrary to the $\Gamma > 1/5$ case in which it is prograde. Since GR always promotes prograde precession, its effect in this Γ regime is initially to slow down the overall precession rate $\dot{\omega}$, allowing for a more coherent torque compared to the case of $\epsilon_{\text{GR}} = 0$. This leads to the appearance of new librating solutions. This is first true for binaries that previously lay just below the separatrix in the (ω, e) phase space, as these circulate the slowest (recall that the secular period diverges on the separatrix itself), giving $\dot{\omega} \sim 0$ for a relatively small value of ϵ_{GR} . On the other hand, at the highest binary eccentricities (near $e = e_{\text{lim}}$) GR may dominate the dynamics, causing the binary's pericentre angle ω to precess rapidly, leading to the appearance of new high- e circulating solutions.

Simultaneously with the new high- e family of orbits, two saddle points (i.e. fixed points that are not local extrema of H^*) emerge at $\omega = 0, \pi$. Passing through them are separatrices that isolate the distinct phase space orbital families in panels (b), (c) and (h). This is an entirely new phase space feature that is not found in LK theory, as it is only possible for $\Gamma \leq 1/5$ and only when GR is present, as we show in §4.3.3. The ‘two-eyed’ phase space structure of panels (b), (c) and (h) has therefore not been uncovered before. Note that for a system exhibiting this structure, a circulating trajectory ‘above’ the saddle point can have the same H^* value as a circulating trajectory ‘below’ the saddle point. In other words a single value of the Hamiltonian can correspond to two entirely different phase space trajectories. This can be seen in Figures 4.2c,h where the dashed contours circulating *above* the librating islands appear because they have the same H^* values as the manually added dashed low- e contours passing through $(\omega, e) = (\pm\pi/2, 0.1)$ and $(0, 0.01)$.

As ϵ_{GR} is increased further, the eccentricity of the saddle points diminishes, similar to the fixed points at $\omega = \pi/2$. It is interesting to note that these various types of fixed points move at different ‘speeds’ down the phase portrait as ϵ_{GR} grows. In particular, panels (d) and (i) of Figure 4.2 demonstrate that for $0 < \Gamma \leq 1/5$ there is a range of ϵ_{GR} values where the saddle point at $\omega = 0$ has gone below $e = 0$ and so no longer exists, but the $\omega = \pm\pi/2$ fixed points still do exist.

Even these remaining fixed points get pushed to (and past) $e = 0$ as ϵ_{GR} is increased ever further, leaving the entire phase space filled with circulating trajectories that have their eccentricity maxima at $\omega = \pm\pi/2$, just as for $\Gamma > 1/5$ (Figure 4.2e,j). The amplitude of eccentricity oscillations decreases correspondingly until cluster tides are completely negligible and only GR apsidal precession remains.

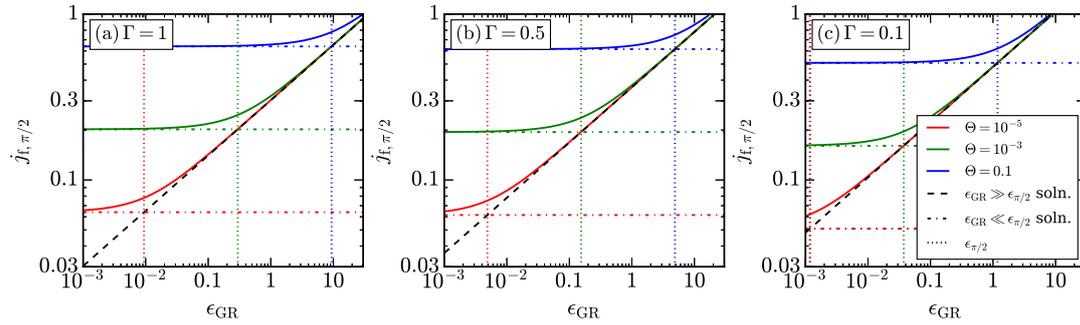


Figure 4.3: Plots of $j_{f, \pi/2}$ — i.e. the values of j for fixed points at $\omega = \pi/2$ — for several values of Γ and Θ (indicated on panels using labels and colours). Solid lines show the exact solution $j_{f, \pi/2}(\Gamma, \Theta, \epsilon_{GR})$ found by solving the quartic equation (4.63). Dot-dashed and dashed lines indicate the asymptotic solutions (4.20) and (4.21) respectively, while vertical dotted lines indicate $\epsilon_{GR} = \epsilon_{\pi/2}$ (see equation 4.19) where the two asymptotic solutions match.

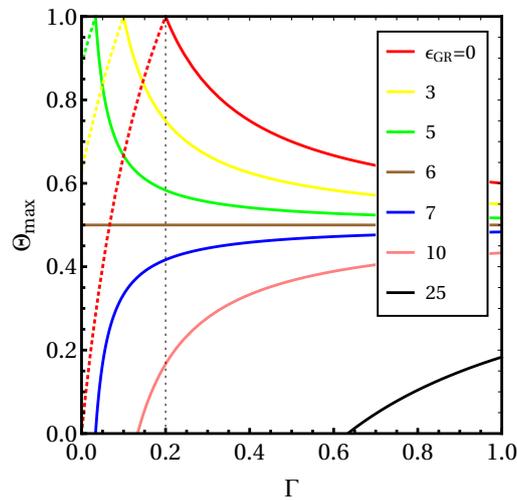


Figure 4.4: Plots of Θ_{\max} , the maximum value of Θ for which fixed points could exist at $\omega = \pm\pi/2$, defined by equation (4.23), as a function of Γ for various values of ϵ_{GR} . Solid lines show $\Theta_{\max} = \Theta_1$ while dotted lines show $\Theta_{\max} = \Theta_2$. The vertical dotted line corresponds to $\Gamma = 1/5$. This figure is discussed in more detail after equation (4.67).

4.3.3 Fixed points

We now proceed to understand mathematically the nature of the various fixed points that we found in the phase portraits in §§4.3.1-4.3.2. By setting $dj/dt = 0$ in equation (4.10), we see that all possible non-trivial fixed points³ are located on (i) the lines $\omega = \pm\pi/2$, as in Chapter 3, and/or (ii) the lines $\omega = 0, \pm\pi$, consistent with Figures 4.1 and 4.2. Finding the j values of the fixed points requires plugging these ω values into $d\omega/dt = 0$, given by equation (4.9), and solving the resulting algebraic equation for j . We do this next for each of the fixed points.

³i.e. not corresponding to $j^2 = \Theta$ or $j^2 = 1$.

Fixed points at $\omega = \pm\pi/2$

In §4.A.1, we show how to calculate the j value of the fixed points at $\omega = \pm\pi/2$, which we call $j_{f,\pi/2}$, for arbitrary ϵ_{GR} and for any $\Gamma > 0$, i.e. for both types of phase portraits shown in Figures 4.1, 4.2. The values of $j_{f,\pi/2}$ are found as solutions to the quartic polynomial (4.63), and we illustrate their behaviour in Figure 4.3 for several values of Γ and Θ . One can see that $j_{f,\pi/2}$ always *increases* with ϵ_{GR} (see (4.66)), explaining why in Figures 4.1,4.2 the fixed points at $\omega = \pm\pi/2$ always get pushed to lower e as ϵ_{GR} is gradually increased from zero.

While the explicit expressions for $j_{f,\pi/2}$ are too complicated to be shown here, we can gain important insights by considering two limiting cases, namely when ϵ_{GR} is much smaller/larger than a particular critical value:

$$\epsilon_{\pi/2} \equiv 6(10\Gamma\Theta)^{3/4}(1+5\Gamma)^{1/4}. \quad (4.19)$$

(In the top and bottom rows of Figure 4.1, $\epsilon_{\pi/2}$ takes values around 4.9 and 16.3 respectively). In the limit $\epsilon_{\text{GR}} \ll \epsilon_{\pi/2}$, which we will call ‘very weak GR’ regime, the ϵ_{GR} term in (4.63) is small and we find to lowest order in $\epsilon_{\text{GR}}/\epsilon_{\pi/2}$

$$j_{f,\pi/2} \approx j_f \left(1 + \frac{\epsilon_{\text{GR}}}{4\epsilon_{\pi/2}} \right). \quad (4.20)$$

In the opposite limit $\epsilon_{\text{GR}} \gg \epsilon_{\pi/2}$ the right hand side in (4.63) becomes small and we find to lowest order in $\epsilon_{\pi/2}/\epsilon_{\text{GR}}$

$$j_{f,\pi/2} \approx \left[\frac{\epsilon_{\text{GR}}}{6(1+5\Gamma)} \right]^{1/3} \left[1 + \frac{1}{3} \left(\frac{\epsilon_{\pi/2}}{\epsilon_{\text{GR}}} \right)^{4/3} \right]. \quad (4.21)$$

Figure 4.3 shows that these asymptotic solutions match the actual $j_{f,\pi/2}$ behaviour in the appropriate limits very well.

In §4.A.1 we show also that for fixed points at $(\omega, j) = (\pm\pi/2, j_{f,\pi/2})$ to exist for a given $\Gamma > 0$, the quantities Θ and ϵ_{GR} must obey the inequalities

$$6\Theta^{1/2}[(1+5\Gamma)\Theta - 10\Gamma] < \epsilon_{\text{GR}} < 6[1+5\Gamma - 10\Gamma\Theta], \quad (4.22)$$

and

$$\Theta < \Theta_{\text{max}} \equiv \begin{cases} \Theta_1, & \Gamma > 1/5, \\ \min[\Theta_1, \Theta_2], & 0 < \Gamma \leq 1/5. \end{cases} \quad (4.23)$$

Here Θ_2 is the smallest positive real solution to equation (4.67), while

$$\Theta_1 \equiv \frac{1+5\Gamma}{10\Gamma} \left(1 - \frac{\epsilon_{\text{GR}}}{2\epsilon_{\text{strong}}} \right) = \frac{1+5\Gamma - \epsilon_{\text{GR}}/6}{10\Gamma}, \quad (4.24)$$

and we have defined

$$\epsilon_{\text{strong}} \equiv 3(1 + 5\Gamma), \quad (4.25)$$

a quantity that will appear repeatedly throughout this Chapter, as well as throughout Chapter 5.

In Figure 4.4 we plot Θ_{max} (equation (4.23)) as a function of Γ for various values of ϵ_{GR} (c.f. Figure 3.1). It is easy to check that the combinations of Γ , Θ and ϵ_{GR} that give rise to $\omega = \pm\pi/2$ fixed points in Figure 4.1 do obey the inequalities (4.22)-(4.23).

Of course, in the limit $\epsilon_{\text{GR}} \rightarrow 0$ equations (4.23)-(4.25) reduce to the non-GR constraint (4.18). Finally we note that for sufficiently small Θ , the conditions (4.22), (4.23) reduce simply to the requirement that

$$\epsilon_{\text{GR}} < 2\epsilon_{\text{strong}} \quad (\text{for } \Theta \ll 1). \quad (4.26)$$

In other words, if (4.26) is not satisfied then there are no fixed points even for initially orthogonal inner and outer orbits ($i_0 = 90^\circ$). ϵ_{GR} beyond a critical value of $6(1 + 5\Gamma)$ the right hand of (4.21) is necessarily > 1 , so that fixed points no longer exist even for $\Theta = 0$, having disappeared through $e = 0$. This reflects what we see in Figures 4.1d,e and 4.2d,e (see also §4.4.1).

Fixed points at $\omega = 0, \pi$

Fixed points at $\omega = 0, \pi$ are unique to the $0 < \Gamma \leq 1/5$ regime (for $\Gamma > 0$). They are always saddle points, and we explore their properties mathematically in §4.A.2. From now on, for brevity we will simply refer to them as being located at $\omega = 0$ rather than $\omega = 0, \pm\pi$, because phase space locations separated in ω by multiples of π are equivalent (see equation (4.7)).

As we demonstrate in §4.A.2, these saddle points are always located at $(\omega, j) = (0, j_{f,0})$ where

$$j_{f,0} \equiv \left[\frac{\epsilon_{\text{GR}}}{6(1 - 5\Gamma)} \right]^{1/3}. \quad (4.27)$$

Note that $j_{f,0}$ is independent of Θ , which can be reflected in Figure 4.2c,h. Also, the constraint (3.9) implies that fixed points exist at $(\omega, j) = (0, j_{f,0})$ if and only if

$$\Theta^{3/2} < \frac{\epsilon_{\text{GR}}}{6(1 - 5\Gamma)} < 1. \quad (4.28)$$

Obviously ϵ_{GR} must be finite for the inequality (4.28) to hold even for very small Θ — hence fixed points at $\omega = 0$ do not exist for $\epsilon_{\text{GR}} = 0$, which is why they were not found in Chapter 3 and do not exist in Figures 4.1a,f or 4.2a,f.

In addition we learn from (4.28) that there are never any fixed points at $\omega = 0$ for $\Gamma > 1/5$ regardless of ϵ_{GR} , which explains the phase space structure in Figure 4.1. In particular this implies that for the LK problem ($\Gamma = 1$) the only possible fixed point locations are the standard ones at $\omega = \pm\pi/2$, regardless of the value of ϵ_{GR} . For positive Γ , fixed points at $\omega = 0$ can be realised only for $\Gamma \leq 1/5$ and we see from (4.27)-(4.28) that when ϵ_{GR} exceeds the threshold value $6(1 - 5\Gamma)\Theta^{3/2}$ (corresponding to $\epsilon_{\text{GR}} = 0.095$ and 1.06 in the top and bottom rows of Figure 4.2, respectively), a fixed point appears at the limiting eccentricity $e_{\text{lim}} = \sqrt{1 - \Theta}$. Increasing ϵ_{GR} at fixed Γ always acts to increase $j_{f,0}$, i.e. to decrease the eccentricity $e_{f,0} \equiv (1 - j_{f,0}^2)^{1/2}$ of this particular fixed point. As we increase ϵ_{GR} to the threshold value $\epsilon_{\text{GR}} = 6(1 - 5\Gamma)$ (which is independent of Θ and corresponds to $\epsilon_{\text{GR}} = 3$ in Figure 4.2), the saddle point vanishes through $e = 0$, leaving only the fixed points at $\omega = \pm\pi/2$.

Beyond that threshold, as mentioned in §4.3.3, there is a range of ϵ_{GR} values for which the saddle point at $\omega = 0$ is no longer present, but the $\omega = \pm\pi/2$ fixed points still do exist. Combining the constraints (4.22), (4.23) and (4.28) we see that for $\Theta \ll 1$ this range is given approximately by

$$6(1 - 5\Gamma) < \epsilon_{\text{GR}} < 6(1 + 5\Gamma). \quad (4.29)$$

The lower limit here is exact, while the upper limit is correct to zeroth order in Θ . Within this range the qualitative behaviour resembles the $\Gamma > 1/5$ behaviour we saw in Figure 4.1; in particular, the maximum eccentricity of all orbits is found at $\omega = \pm\pi/2$. The range (4.29) is important because it allows for eccentricity excitation of initially near-circular binaries, which is not possible in the $0 < \Gamma \leq 1/5$ regime when $\epsilon_{\text{GR}} = 0$ (see §4.3.4).

4.3.4 Determination of the maximum eccentricity of a given orbit

Our next goal is to calculate the maximum eccentricity e_{max} reached by a binary given the initial conditions $(\omega_0, e_0, \Theta, \Gamma, \epsilon_{\text{GR}})$. In particular, we wish to know if a binary will reach $e_{\text{max}} \rightarrow 1$, since this is the regime in which dissipative effects (e.g. GW emission) can become important.

For $\Gamma > 1/5$, a binary's maximum eccentricity is always found at $\omega = \pi/2$ regardless of whether its phase space orbit librates or circulates (Figure 4.2). Plugging $\omega = \pi/2$ into $H^*(\omega, j)$ gives us a depressed quartic equation:

$$j^4 + \left(\frac{H^* - 24\Gamma\Theta - 5 - 15\Gamma}{3(1 + 5\Gamma)} \right) j^2 + \frac{\epsilon_{\text{GR}}}{3(1 + 5\Gamma)} j + \frac{10\Gamma\Theta}{1 + 5\Gamma} = 0. \quad (4.30)$$

We call real roots of equation (4.30) $j(\omega = \pi/2)$. In the limit $\epsilon_{\text{GR}} = 0$, equation (4.30) reduces to a quadratic for $j^2(\omega = \pi/2)$ and we recover the non-GR solution (3.18). For $\epsilon_{\text{GR}} \neq 0$ the real roots of (4.30) can still be written down analytically

but they are too complicated to be worth presenting here. The minimum angular momentum j_{\min} (corresponding to the maximum eccentricity $e_{\max} \equiv \sqrt{1 - j_{\min}^2}$) will then be given by the smallest physical root $j(\omega = \pi/2)$, i.e. the smallest root of (4.30) that satisfies $\sqrt{\Theta} < j(\omega = \pi/2) < 1$.

The situation is slightly more complex for $0 < \Gamma \leq 1/5$. In this case we must first work out whether an orbit librates or circulates (and if it circulates, to which circulating family it belongs, since it can be above or below the saddle point, as in Figures 4.2b,c,h). To do so we use the procedure given in §4.A.3 to calculate $j(\omega = 0)$, which is the solution to the depressed cubic equation (4.69) that results from plugging $\omega = 0$ into $H^*(\omega, j)$. If the orbit circulates ‘below’ the librating regions and the saddle point then we have $j_{\min} = j(\omega = 0)$. Otherwise j_{\min} is found at $\omega = \pm\pi/2$ and we proceed as for $\Gamma > 1/5$ by solving equation (4.30).

Maximum eccentricity achieved by initially near-circular binaries

We can gain further insight and connect to the results of previous LK studies by considering the simplified case of initially near-circular binaries, $e_0 \approx 0$. Evaluating the integrals of motion H^* and Θ with the initial condition $e_0 = 0$ we find

$$H^* = 2(1 - 3\Gamma \cos^2 i_0) - \epsilon_{\text{GR}}, \quad \Theta = \cos^2 i_0. \quad (4.31)$$

Note the lack of ω_0 dependence in these constants.

Now, for $\Gamma > 1/5$ eccentricity is always maximised at $\omega = \pi/2$, so can be found by solving (4.30). Plugging (4.31) into (4.30) we find that j_{\min} is the solution to the equation

$$0 = (j - 1) \left[j^3 + j^2 - \frac{(10\Gamma \cos^2 i_0 + \epsilon_{\text{GR}}/3)}{1 + 5\Gamma} j - \frac{10\Gamma \cos^2 i_0}{1 + 5\Gamma} \right]. \quad (4.32)$$

In the LK limit of $\Gamma = 1$, equation (4.32) is equivalent to e.g. equation (34) of Fabrycky & Tremaine (2007) or equation (50) of Liu, Muñoz, et al. (2015)⁴. Note that $j_{\min} = 1$ (i.e. $e_{\max} = 0$) is a solution to this equation. It is the correct solution in the special case of a perfectly initially circular orbit, $e_0 \equiv 0$, which necessarily remains circular forever. This is because perfectly circular binaries feel no net torque from the external tide, which can be seen by plugging $j = 1$ into equation (4.10).

Meanwhile, an orbit that has e_0 infinitesimally larger than zero can have j_{\min} corresponding to a non-trivial solution of (4.32). This will be the case if and only if the $\omega = \pm\pi/2$ fixed points have not yet disappeared below $e = 0$ (panels (a)-(d) and (f)-(h) of Figure 4.1). Because of the constraint (4.22), a necessary (and for $i_0 \rightarrow 90^\circ$, sufficient) requirement for this is $\epsilon_{\text{GR}} < 6(1 + 5\Gamma)$. In that case the fixed points bound

⁴Note that there is a typo in Liu, Muñoz, et al. (2015)’s equation (50) — the factor of 3/5 on the right hand side should be 5/3.

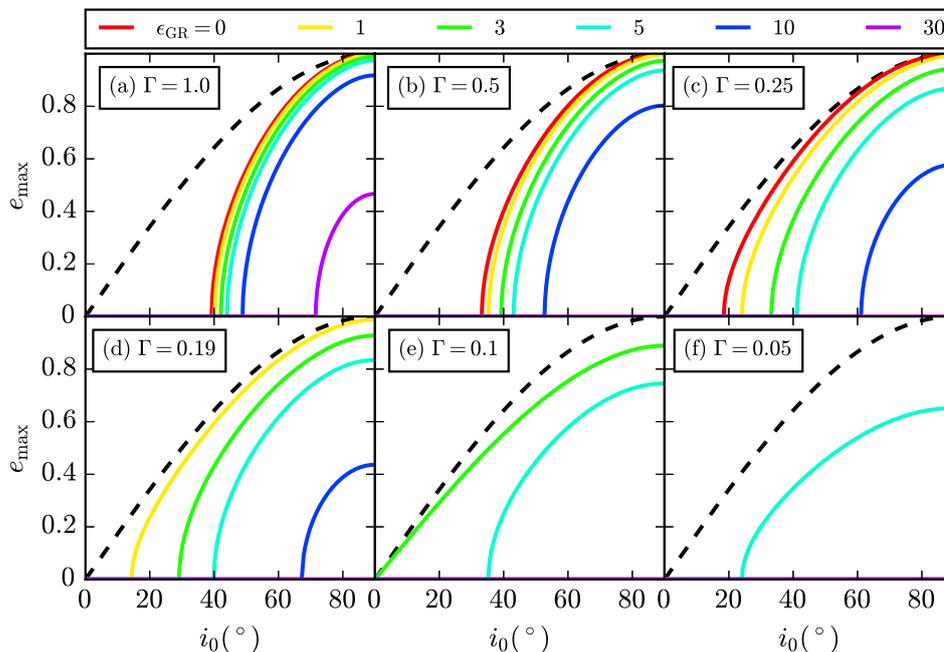


Figure 4.5: Maximum eccentricity e_{\max} as a function of i_0 for initially near-circular binaries. Panels (a)-(c) are for $\Gamma > 1/5$ while panels (d)-(f) correspond to $0 < \Gamma \leq 1/5$. In each panel, different coloured lines represent the different values of ϵ_{GR} (see legend). A dashed black line corresponds to $e_{\max} = e_{\text{lim}} = \sin i_0$. Note that for initially circular orbits to reach a non-zero e_{\max} we require fixed points to exist in the phase portrait at $\omega = \pm\pi/2$ but not at $\omega = 0$; for $i_0 \approx 90^\circ$ this corresponds to $6(1 - 5\Gamma) < \epsilon_{\text{GR}} < 6(1 + 5\Gamma)$ — see equation (4.29). Note also that for $\Gamma < 1/5$, eccentricity excitation of near-circular binaries may be possible regardless of inclination, even when $i_0 = 0^\circ$, as for $\epsilon_{\text{GR}} = 3$ in panel (e).

the maximum eccentricity from below, so $e_{\max} > e_{f,\pi/2} \equiv (1 - j_{f,\pi/2}^2)^{1/2}$. On the other hand, if ϵ_{GR} is large enough that the fixed points *have* disappeared through $e = 0$ then we simply have $e_{\max} = 0$ (see panels (e), (i), (j) of Figure 4.1).

Next we turn to the regime $0 < \Gamma \leq 1/5$. By consulting Figure 4.2 one can see that a finite eccentricity is only achieved if ϵ_{GR} is sufficiently large that the saddle point at $\omega = 0$ has passed ‘down’ the (ω, e) phase space and disappeared through $e = 0$, but also sufficiently small that the $\omega = \pm\pi/2$ fixed points still exist (as in Figure 4.2d,i). A necessary requirement for this (which is again sufficient in the case $i_0 \rightarrow 90^\circ$) is that (4.29) be true. Then e is maximised at $\omega = \pm\pi/2$ and j_{\min} is a non-trivial solution to equation (4.32). On the other hand, if (4.29) is not satisfied then a binary that starts at $e_0 \approx 0$ never increases its eccentricity⁵ even for $i_0 = 90^\circ$.

Overall then, we see that for near-circular binaries to reach finite e_{\max} we require fixed points to exist in the phase portrait at $\omega = \pm\pi/2$ but not at $\omega = 0$, and this necessarily requires ϵ_{GR} to satisfy (4.29).

⁵There is another solution at $\omega = 0$ given by equation (4.93) which is unphysical for $\Gamma > 0$ but will become important for $\Gamma \leq 0$ — see §4.D.3.

In Figure 4.5 we plot e_{\max} as a function of i_0 for initially near-circular binaries. Panels (a)-(c) are for $\Gamma > 1/5$ (c.f. Figure 3 of Fabrycky & Tremaine (2007) and Figure 6 of Liu, Muñoz, et al. (2015)) while panels (d)-(f) correspond to $0 < \Gamma \leq 1/5$. In each panel, different coloured solid lines represent the different values of ϵ_{GR} , while a dashed black line indicates the limiting eccentricity $e_{\text{lim}} = \sqrt{1 - \Theta} = \sin i_0$ (the highest possible e for an initially near-circular binary, corresponding to $j = \cos i_0$). We see that for $\Gamma > 1/5$, the effect of increasing ϵ_{GR} at a fixed i_0 (and therefore a fixed Θ) is always to decrease e_{\max} . This is what we would expect by comparing the top and bottom rows of Figure 4.1. Moreover, if we consider the most favourable orbital inclination $i_0 = 90^\circ$ then we can easily derive the exact solution to (4.32). We find that either $j_{\min} = 1$ (so $e_{\max} = 0$), or that

$$j_{\min} = \frac{1}{2} \left[\left(1 + \frac{4\epsilon_{\text{GR}}}{\epsilon_{\text{strong}}} \right)^{1/2} - 1 \right], \quad (4.33)$$

with ϵ_{strong} defined in (4.25); in the LK limit this result reduces to equation (35) of Fabrycky & Tremaine (2007). Expanding the solution (4.33) for $\epsilon_{\text{GR}}/\epsilon_{\text{strong}} \ll 1$ we find

$$e_{\max} \approx 1 - \frac{1}{2} \left(\frac{\epsilon_{\text{GR}}}{\epsilon_{\text{strong}}} \right)^2. \quad (4.34)$$

Thus we expect $e_{\max} \rightarrow 1$ for these favourably inclined binaries when GR is negligible, but also that e_{\max} will deviate from 1 considerably when ϵ_{GR} starts approaching ϵ_{strong} , which is what we see in Figure 4.5a,b,c. Obviously this means that the smaller is Γ , the smaller ϵ_{GR} needs to be to suppress the very highest eccentricities. Finally we note that there is no magenta curve — corresponding to $\epsilon_{\text{GR}} = 30$ — in either panel (b) or panel (c). This is because for these Γ values the constraint (4.29) is violated for $\epsilon_{\text{GR}} = 30$, so the only possible solution to (4.32) is $e_{\max} = 0$.

Now consider the regime $0 < \Gamma \leq 1/5$ exhibited in panels (d)-(f). The reader will notice the diminishing number of curves in these panels. Indeed, there is not even a red curve corresponding to $\epsilon_{\text{GR}} = 0$. This again is a consequence of the fact that for $\epsilon_{\text{GR}} = 0$, equation (4.29) cannot be satisfied, so that initially circular orbits achieve no eccentricity excitation ($e_{\max} = 0$).

A related phenomenon is that in panel (e), the green ($\epsilon_{\text{GR}} = 3$) curve asymptotes to the black dashed line $e = e_{\text{lim}}$ as $i_0 \rightarrow 0^\circ$. This is also as expected: since $\Gamma = 0.1$, equation (4.29) tells us $\epsilon_{\text{GR}} = 3$ is precisely the lower bound on GR strength above which initially near-circular binaries can reach non-zero eccentricities at all i_0 , since at this value of ϵ_{GR} the saddle point crosses $e = 0$ — see Figure 4.6i,j,k.

Note that this $0 < \Gamma \leq 1/5$ behaviour is completely different from that found for near-circular binaries in the $\Gamma > 1/5$ regime (and therefore to the known LK results). For $\Gamma > 1/5$, taking $i_0 \approx 0^\circ$ inevitably leads to $e_{\max} \approx 0$ — in other words there

is no eccentricity excitation for initially coplanar ($i_0 = 0$) orbits, regardless of ϵ_{GR} . Moreover, for $\Gamma > 1/5$ even if a binary can reach a finite maximum eccentricity for $\epsilon_{\text{GR}} = 0$, increasing ϵ_{GR} always decreases this maximum eccentricity. On the contrary, for $0 < \Gamma \leq 1/5$ reaching a finite e_{max} may be possible even for initially almost coplanar orbits, and a finite ϵ_{GR} is actually *necessary* to trigger the eccentricity excitation starting from a circular orbit. Despite this, comparison of the top and bottom rows of Figure 4.5 reinforces the idea that the $0 < \Gamma \leq 1/5$ regime admits far fewer high-eccentricity solutions than $\Gamma > 1/5$ as ϵ_{GR} is varied.

4.4 High eccentricity behaviour

Our next goal is to understand the impact of GR precession on the time dependence of the binary orbital elements in the important limit of very high eccentricity, $e \rightarrow 1$. This limit is relevant in a variety of astrophysical contexts. For example, the dramatic reduction of the binary pericentre distance that occurs when e approaches unity can trigger short-range effects such as tidal dissipation (leading to hot Jupiter formation), GW emission (leading to compact object mergers), and so on. Thus we wish to understand in detail how GR precession affects not only the maximum eccentricity e_{max} , but also the behaviour of $e(t)$ and other orbital elements in the vicinity of e_{max} .

In §4.3.4 we explained how to find e_{max} for arbitrary $\Gamma > 0$, initial conditions (e_0, i_0, ω_0) , and value of ϵ_{GR} . Here we will examine the solutions quantitatively in the high eccentricity limit, and explore the time spent near highest eccentricity. To this end we will make extensive use of equation (4.12), which tells us dj/dt as a function of j . It is important to note that the solutions for extrema of j at $\omega = 0$ and $\omega = \pm\pi/2$ are all contained within (4.12). Indeed, setting the first square bracket inside the square root in (4.12) to zero gives the depressed quartic equation (4.30) whose roots correspond to extrema of j at $\omega = \pm\pi/2$, i.e. what we have so far called $j(\omega = \pm\pi/2)$. Setting the other square bracket to zero gives the depressed cubic (4.69)-(4.70) which determines the roots at $\omega = 0$, i.e. what we called $j(\omega = 0)$.

In this section we will focus on situations in which e_{max} is achieved at $\omega = \pm\pi/2$, since this is the most common prerequisite for $e \rightarrow 1$ (§§4.3.1-4.3.2). The rare cases in which e approaches unity at $\omega = 0$ are covered in Appendix 4.B.

4.4.1 Phase space behaviour for $\Theta \ll 1$, $\Gamma > 0$

We are interested in binaries that start with initial eccentricity e_0 not close to unity, and that are capable of reaching extremely high eccentricities $e_{\text{max}} \rightarrow 1$, i.e. $j_{\text{min}} \rightarrow 0$. For this to be possible a necessary condition is that $\Theta \ll 1$, owing to the constraint (3.9). Hence it is important to understand the regime $\Theta \ll 1$ in detail.

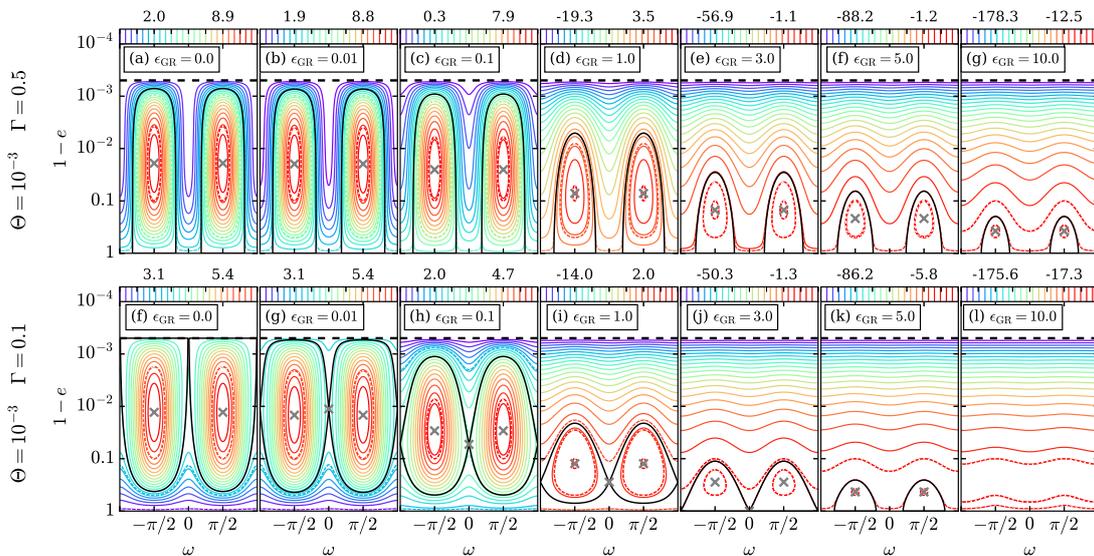


Figure 4.6: As in Figures 4.1 and 4.2, except we have (I) fixed $\Theta = 10^{-3}$ and used two values of Γ (namely 0.5 and 0.1 for the top and bottom row respectively), (II) plotted $1 - e$ on the vertical axis using an inverted logarithmic scale (so that e still increases vertically), (III) added by hand additional dashed contours with the value $H^*(\omega = \pm\pi/2, e = 0.9)$, and (IV) used some new values of ϵ_{GR} .

In Figure 4.6 we show phase portraits for $\Gamma = 0.5$ (top row) and $\Gamma = 0.1$ (bottom row), this time fixing $\Theta = 10^{-3}$ in both cases, and adding in extra dashed contours⁶ with the value $H^*(\omega = \pm\pi/2, e = 0.9)$. Note that on the vertical axis we now plot $1 - e$ on a logarithmic scale, with eccentricity still increasing vertically as in Figures 4.1, 4.2. This allows us to see in detail how trajectories separate from $e \approx e_{\text{lim}}$ as we increase ϵ_{GR} .

In these plots, Θ is sufficiently small that to a very good approximation the requirement for fixed points at $\omega = \pm\pi/2$ to exist is just $\epsilon_{\text{GR}} < 2\epsilon_{\text{strong}}$ (equation (4.26)). This critical value is surpassed in panel (l), since in that case $\epsilon_{\text{GR}} = 10$ while $2\epsilon_{\text{strong}} = 9$, which is why all fixed points have disappeared. Meanwhile the criterion for a saddle point to exist at $\omega = 0$ (equation (4.28)) for $\Gamma = 0.1$ and $\Theta = 10^{-3}$ is approximately $10^{-5} < \epsilon_{\text{GR}} < 3$. This is consistent with what we see in panels (f)-(l) — note in particular the transitional point $\epsilon_{\text{GR}} = 3$ in panel (j).

Comparing the top and bottom rows of Figure 4.6, one observes a striking difference between behaviour in the $\Gamma > 1/5$ and $0 < \Gamma \leq 1/5$ dynamical regimes. For $\Gamma = 0.5 > 1/5$, an initially near-circular binary can be driven to very high eccentricity ($\gtrsim 0.99$) even for $\epsilon_{\text{GR}} = 1.0$ (panel (d)). Conversely, for $\Gamma = 0.1 < 1/5$ the phase space structure simply does not allow such behaviour (panels (f)-(i)). More precisely, for $0 < \Gamma \leq 1/5$, the eccentricity of the saddle point (4.27) acts as a hard boundary on the maximum eccentricity of low- e orbits, and most of them do not get close even to that value. Even

⁶In addition to the dashed contours already included in Figures 4.1 and 4.2.

when ϵ_{GR} is increased so that a new family of circulating orbits appears, and the librating region is significantly enlarged, the system admits very few solutions that start at low- e and achieve high- e . It is therefore unsurprising that one finds fewer cluster-tide driven compact object mergers from systems such as globular clusters that have a relatively high fraction of binaries in the $0 < \Gamma \leq 1/5$ regime (Chapter 7).

4.4.2 High eccentricity behaviour for $\epsilon_{\text{GR}} = 0$

Before embarking on a full study of high eccentricity evolution for arbitrary ϵ_{GR} , we first consider the case $\epsilon_{\text{GR}} = 0$. In that case the non-zero roots of the polynomial on the right hand side of (4.12) are j_{\pm}, j_0 , one of which will correspond to the minimum angular momentum j_{min} . Then we can integrate (4.12) with $\epsilon_{\text{GR}} = 0$ to find $t(j)$; the resulting expression involves an incomplete elliptical integral of the first kind (see §3.2.6 for the general Γ case, and Vashkov'yak (1999) and Kinoshita & Nakai (2007) in the LK case of $\Gamma = 1$). Next, assuming that $j^2 \ll 1$, we can expand this elliptical integral to find⁷ (see §3.9.2):

$$j(t) = j_{\text{min}} \sqrt{1 + \left(\frac{t}{t_{\text{min}}}\right)^2}, \quad \text{where} \quad t_{\text{min}} \equiv \frac{j_{\text{min}}}{j_1 j_2} \tau, \quad (4.35)$$

j_1, j_2 are the two roots *not* corresponding to j_{min} , and τ is a characteristic secular timescale which is independent of e_0, i_0, ω_0 :

$$\tau \equiv \frac{L}{6C\sqrt{|25\Gamma^2 - 1|}}. \quad (4.36)$$

Using the definitions of C and L one can show that τ is, up to constant factors, the same as t_{sec} defined after equation (4.6). Note we have taken the origin of the time coordinate to coincide with $j = j_{\text{min}}$. Clearly t_{min} is the characteristic evolution timescale in the vicinity of j_{min} , i.e. the time it takes for j to change from j_{min} to $\sqrt{2}j_{\text{min}}$.

Note that the solution (4.35) is quadratic in t for $t \lesssim t_{\text{min}}$ and linear when $t \gtrsim t_{\text{min}}$, as long as j remains $\ll 1$. It provides a better approximation to $j(t)$ over a wider interval of time near the peak eccentricity than the purely quadratic approximation adopted by Randall & Xianyu (2018), in their calculation of the GW energy emitted by a binary undergoing LK oscillations (§4.C.3; see also §5.A).

4.4.3 Modifications brought about by finite ϵ_{GR}

Before we proceed to examine the $j(t)$ behaviour, it is important to realise that including a finite ϵ_{GR} affects the right hand side of (4.12), and therefore the value of j_{min} , in two distinct ways. First, there is the obvious explicit dependence on ϵ_{GR} that appears twice in equation (4.12). Second, there is also an implicit dependence on ϵ_{GR} in (4.12)

⁷Note that one can get the same result simply by expanding the right hand side of (4.12) for $j \ll 1$.

through the values of j_{\pm} and j_0 (see equations (4.13), (4.14)). We will now discuss this implicit dependence, and then use the results to understand $j(t)$ behaviour in different asymptotic ϵ_{GR} regimes.

In the limit $\Theta \ll 1$, and assuming that e_0 is not too close to 1 and Γ is not too close to $1/5$, equations (4.15), (4.16) tell us that

$$\Sigma \approx (\epsilon_{\text{strong}} + \epsilon_{\text{GR}})/6 + \mathcal{O}(e_0^2). \quad (4.37)$$

Equation (4.37) implies that Σ , and hence j_{\pm}^2 , will be modified significantly by GR only if $\epsilon_{\text{GR}} \gtrsim \epsilon_{\text{strong}}$, in agreement with what we saw in Figures 4.1, 4.2 & 4.6. In this case, a perturbative approach around the non-GR solution will fail. We therefore say that any binary with $\epsilon_{\text{GR}} \gtrsim \epsilon_{\text{strong}}$ exists in the regime of ‘strong GR’, which we explore in §4.4.5. Conversely, if ϵ_{GR} is in what we will call the ‘weak-to-moderate GR’ regime:

$$\epsilon_{\text{GR}} \ll \epsilon_{\text{strong}}, \quad (4.38)$$

then $\Sigma \sim 1$, and so for fixed $\Theta \ll 1$ the values of j_{\pm} will stay close to their non-GR values as we increase ϵ_{GR} , namely (see equation (4.13)):

$$j_+^2 \approx \frac{2\Sigma}{1+5\Gamma} \sim 1, \quad \text{and} \quad j_-^2 \approx \frac{5\Gamma\Theta}{\Sigma} \sim \Theta \ll 1. \quad (4.39)$$

In other words the relative perturbations to j_{\pm} induced by GR can normally be neglected. Note that the weak-to-moderate GR regime (4.38) already encompasses the very weak GR regime introduced in §4.3.3. In §4.4.4 we will further delineate distinct ‘weak GR’ and ‘moderate GR’ regimes.

Also, using equations (4.14), (4.16) it is easy to show that the *absolute* change to j_0 incurred by including GR will be small ($\ll 1$) whenever

$$\epsilon_{\text{GR}} \ll 3|1 - 5\Gamma|\sqrt{1 - e_0^2}. \quad (4.40)$$

Note that for Γ not close to $1/5$ and e_0 not close to unity, the condition (4.40) is automatically guaranteed by the weak-to-moderate GR condition (4.38). In that case the relative perturbation to j_0 due to GR precession can be neglected (if $j_0 \sim 1$).

4.4.4 High- e behaviour in the weak-to-moderate GR limit

In the non-GR limit ($\epsilon_{\text{GR}} = 0$), for $\Gamma > 0$ the vast majority of phase space trajectories that are capable of reaching very high eccentricities reach them at⁸ $\omega = \pm\pi/2$. As shown in Chapter 3, for $\Gamma > 0$ the corresponding minimum angular momentum for

⁸The exception is for circulating orbits with $0 < \Gamma < 1/5$ that lie very close to the separatrix. These rare orbits are discussed in Appendix 4.B.

these orbits is always $j_{\min} = j_- \ll 1$. The maximum j is equal to j_0 if the orbit circulates and j_+ if it librates.

We now want to see what happens to (4.12) for finite $\epsilon_{\text{GR}} \ll \epsilon_{\text{strong}}$. From the discussion in §4.4.3 we expect that we may neglect j^2 compared to j_+^2 , j_0^2 in this regime. As a result we can write

$$\frac{dj}{dt} \approx \pm \frac{6C}{Lj^{3/2}} \sqrt{(25\Gamma^2 - 1)j_+^2 j_0^2 [j^2 - j_-^2 - \gamma j j_-] [j + \sigma j_-]}, \quad (4.41)$$

where we defined the following dimensionless numbers:

$$\gamma \equiv \frac{\epsilon_{\text{GR}}}{3(1 + 5\Gamma)j_+^2 j_-} = \frac{2\epsilon_{\text{GR}}}{\epsilon_{\text{weak}}}, \quad (4.42)$$

$$\sigma \equiv \frac{\epsilon_{\text{GR}}}{3(5\Gamma - 1)j_0^2 j_-} = \frac{2\epsilon_{\text{GR}}}{\epsilon_{\text{weak}}} \times \frac{5\Gamma + 1}{5\Gamma - 1} \frac{j_+^2}{j_0^2}, \quad (4.43)$$

with

$$\epsilon_{\text{weak}} \equiv 6(1 + 5\Gamma)j_+^2 j_- \approx (720\Gamma\Sigma)^{1/2} \Theta^{1/2}. \quad (4.44)$$

To get the second equality in (4.44) we used the approximation (4.39). Both ϵ_{weak} and γ are manifestly positive in the weak-to-moderate GR regime given $\Gamma > 0$. Except in pathological cases, σ is also positive for the regimes we are interested in here⁹.

To find the minimum j at $\omega = \pm\pi/2$ we require the right hand side of (4.41) to equal zero, which, as we mentioned earlier, means that the first square bracket inside the square root must vanish. This gives a quadratic equation for j_{\min} , the only meaningful (positive) solution to which is

$$\begin{aligned} j_{\min} &= \frac{\gamma j_-}{2} \left[1 + \sqrt{1 + 4\gamma^{-2}} \right] \\ &= \frac{1}{2j_+^2 \epsilon_{\text{strong}}} \left[\epsilon_{\text{GR}} + \sqrt{\epsilon_{\text{GR}}^2 + \epsilon_{\text{weak}}^2} \right]. \end{aligned} \quad (4.45)$$

Equations (4.41) and (4.45) work as long as $\Theta \ll 1$ and ϵ_{GR} is in the weak-to-moderate GR regime, i.e. satisfies (4.38) and (4.40).

Equation (4.45) has been used by several authors in the LK limit of $\Gamma = 1$ — see §4.5.2. Importantly, it allows us to write down a solution for the maximum eccentricity reached by initially near-circular binaries in the weak-to-moderate GR regime. Indeed, let us put $\Theta = \cos^2 i_0$ and assume $\Theta \ll 1$ (i.e. $i_0 \approx 90^\circ$) so that the binary is capable of reaching very high eccentricity. Then $j_+ \approx 1$ and from (4.45) we find

$$j_{\min} \approx \frac{1}{2} \left[\frac{\epsilon_{\text{GR}}}{3(1 + 5\Gamma)} + \sqrt{\left[\frac{\epsilon_{\text{GR}}}{3(1 + 5\Gamma)} \right]^2 + \frac{40\Gamma \cos^2 i_0}{1 + 5\Gamma}} \right]. \quad (4.46)$$

⁹This is true because $(5\Gamma - 1)j_0^2$ is positive in the $\epsilon_{\text{GR}} = 0$ limit for all the cases we care about, namely any orbit with $\Gamma > 1/5$ and librating orbits with $0 < \Gamma \leq 1/5$. The inclusion of GR subtracts from $(5\Gamma - 1)j_0^2$ by an amount $\epsilon_{\text{GR}}/(3\sqrt{1 - e_0^2})$. For $j_0 \sim 1$ and $e_0^2 \ll 1$, this modification will not make $(5\Gamma - 1)j_0^2$ negative as long as (4.40) is satisfied.

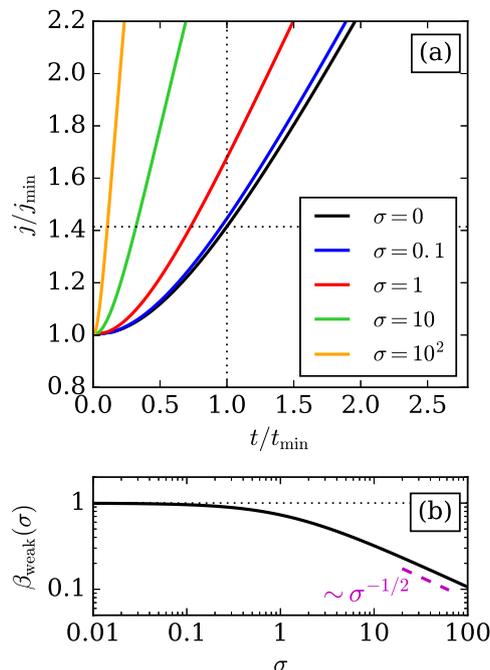


Figure 4.7: High- e behaviour in the weak GR regime (§4.4.4). Panel (a) shows the solution (4.49) for $j(t)$ near the eccentricity peak for various values of σ (equation (4.43)). The horizontal dotted line shows $j/j_{\min} = \sqrt{2}$ and the vertical dotted line shows $t = t_{\min}$. Panel (b) shows $\beta_{\text{weak}}(\sigma)$, which is the time (in units of t_{\min}) over which the binary’s j/j_{\min} changes from 1 to $\sqrt{2}$, defined by setting $j/j_{\min} = \sqrt{2}$ in the right hand side of (4.49). Note that both axes are on a logarithmic scale in this panel. A dashed magenta line shows the scaling $\beta_{\text{weak}} \propto \sigma^{-1/2}$ for $\sigma \gg 1$.

Note that this result is consistent with what we found in §4.3.4, where we assumed near-circularity from the outset and made no (explicit) assumptions about j_{\min} or ϵ_{GR} other than (4.29). For instance: (I) we can alternatively derive (4.46) by solving equation (4.32) in the limit $j \ll 1$; (II) if we take $i_0 = 90^\circ$ in (4.46) then we get exactly the same result as if we expand (4.33) for $\epsilon_{\text{GR}} \ll \epsilon_{\text{strong}}$, namely equation (4.34). Moreover, in the LK limit $\Gamma = 1$ we recover from (4.46) a well-known result, identical to¹⁰ e.g. equation (8) of Miller & Hamilton (2002) and equation (52) of Liu, Muñoz, et al. (2015).

It is now instructive to investigate separately the high- e behaviour in the asymptotic regimes of *weak* and *moderate* GR precession (still assuming eccentricity is maximised at $\omega = \pm\pi/2$).

Weak GR, $\epsilon_{\text{GR}} \ll \epsilon_{\text{weak}}$

In the asymptotic regime of *weak GR*, defined by $\epsilon_{\text{GR}} \ll \epsilon_{\text{weak}}$, the solution (4.45) becomes approximately

$$j_{\min} \approx j_- \left(1 + \frac{\gamma}{2}\right) = j_- \left(1 + \frac{\epsilon_{\text{GR}}}{\epsilon_{\text{weak}}}\right). \quad (4.47)$$

In other words, GR causes only a slight perturbation of j_{\min} away from the non-GR value of j_- at the relative level $\epsilon_{\text{GR}}/\epsilon_{\text{weak}} \ll 1$.

To determine the time dependence of $j(t)$ in the vicinity of j_{\min} , we make use of the weak GR assumption to drop the γ term in the first square bracket in (4.41). The result is

$$\frac{dj}{dt} \approx \pm \frac{6C}{Lj^{3/2}} \sqrt{(25\Gamma^2 - 1)j_+^2 j_0^2 [j^2 - j_-^2] [j + \sigma j_-]}. \quad (4.48)$$

Integration of (4.48) gives an implicit solution for $j(t)$ in the form

$$\frac{t}{t_{\min}} = \int_1^{j/j_{\min}} \frac{dx x^{3/2}}{\sqrt{(x^2 - 1)(x + \sigma)}}, \quad (4.49)$$

where t_{\min} is defined in equation (4.35). In Figure 4.7a we plot the implicit solution for j/j_{\min} as a function of t/t_{\min} for various values of σ .

We can gain insight into σ in the weak GR regime by using the fact that in this regime, $\epsilon_{\text{GR}} \ll j_{\min} \approx j_- \sim \Theta^{1/2}$. Evaluating (4.16) at $\omega = \pm\pi/2$, $j = j_{\min}$ and using these scalings we find

$$D \approx 1 + \frac{10\Gamma}{1 - 5\Gamma} \left(1 - \frac{\Theta}{j_{\min}^2}\right)^{-1}. \quad (4.50)$$

Plugging this into (4.14) and the resulting expression into (4.43) gives

$$\sigma \approx \frac{\epsilon_{\text{GR}}}{30\Gamma j_{\min}} \left(1 - \frac{\Theta}{j_{\min}^2}\right)^{-1} = \frac{\epsilon_{\text{GR}}\chi}{30\Gamma j_{\min}}, \quad (4.51)$$

where $\chi \geq 1$ is defined in equation (4.82). For typical values of $\chi \sim 1$, since $\epsilon_{\text{GR}} \ll j_{\min}$ we expect $\sigma \ll 1$. However, when χ greatly exceeds unity, $\sigma \gtrsim 1$ or even $\sigma \gg 1$ is also possible¹¹.

In the case $\sigma \ll 1$, the term σj_- in the final square bracket in (4.48) can also be dropped compared to j . Then equation (4.48) takes the same functional form as its non-GR analogue; integrating, we get a solution $j(t)$ in precisely the form (4.35) with¹² $j_{\min} \rightarrow j_-$ and $j_1, j_2 \rightarrow j_+, j_0$. This is reflected in Figure 4.7a, in which the black

¹⁰Note that our definition of ϵ_{GR} differs from what Miller & Hamilton (2002) call θ_{PN} and what Liu, Muñoz, et al. (2015) call ϵ_{GR} . Our ϵ_{GR} is defined for any outer orbit in any axisymmetric potential, whereas their parameters are defined only in the Keplerian (LK) limit. In this limit, $\epsilon_{\text{GR}} = 6\theta_{\text{PN}} = 16\epsilon_{\text{GR}}$.

¹¹Note that contrary to what a naive interpretation of (4.43) might suggest, the condition for $\sigma \gg 1$ is not that $\Gamma \rightarrow 1/5$.

¹²Note that j_{\pm}, j_0 depend on ϵ_{GR} through (4.13)-(4.14) only weakly, at the relative level $\mathcal{O}(\epsilon_{\text{GR}}/\epsilon_{\text{weak}})$.

line ($\sigma = 0$) is exactly the non-GR result from (4.35), and as expected $j = \sqrt{2}j_{\min}$ coincides with $t = t_{\min}$ in that case.

However the assumption $\sigma \ll 1$ may not always be valid. Figure 4.7a shows that as we increase σ the behaviour of $j(t)$ becomes more sharply peaked around j_{\min} (when time is measured in units of t_{\min}), although against this trend one must remember that to change σ is to change one or more of ϵ_{GR} , Γ , j_0 and j_- , any of which will modify t_{\min} . We are particularly interested in the value of $t_{\min}^{\text{weak}}(\sigma)$, which is the time it takes for j to go from j_{\min} to $\sqrt{2}j_{\min}$ in the weak GR regime, to compare with the solution (4.35). By setting $j/j_{\min} = \sqrt{2}$ on the right hand side of (4.49) and $t = t_{\min}^{\text{weak}}$ on the left, we find

$$t_{\min}^{\text{weak}}(\sigma) = \beta_{\text{weak}}(\sigma)t_{\min}, \quad (4.52)$$

where $\beta_{\text{weak}}(\sigma)$ is plotted as a function of σ in Figure 4.7b. As expected $\beta_{\text{weak}} \rightarrow 1$ for $\sigma \rightarrow 0$, i.e. in the limit of negligible GR precession. For finite GR, typical values of β_{weak} are ~ 1 except for very large $\sigma \gtrsim 10$. For $\sigma \gg 1$ we see that β_{weak} falls off like $\sim \sigma^{-1/2}$.

In Figures 4.11, 4.12, 4.14 and 4.15 we compare the weak GR solution for $j(t)$, namely equation (4.49), to direct numerical integration of the DA equations of motion (4.9), (4.10), for binaries in different dynamical regimes. Full details are given in §4.C.3; here we only note that the values of the key quantities Γ , ϵ_{GR} , ϵ_{weak} , σ , etc. are shown at the top of each figure. In every example, panel (a) shows $\log_{10}(1 - e)$ behaviour in the vicinity of peak eccentricity, while panel (b) shows the same thing zoomed out over a much longer time interval¹³. The weak GR solution for $j(t)$ (equation (4.49)) is plotted in panels (a) and (b) with a dashed green line, while the numerical solution is shown with a solid blue line. We see that for $\epsilon_{\text{GR}} \ll \epsilon_{\text{weak}}$ (Figures 4.11, 4.12) this weak GR solution works very well, but that substantial errors begin to set in when ϵ_{GR} approaches ϵ_{weak} (Figures 4.14, 4.15). Finally, in each of these plots we also show with red dashed lines an ‘analytic’ solution, equation (4.78), which coincides with (4.35) provided $j_{\min}^4/\Theta \ll 1$. As we have already stated, in the weak GR regime $j(t)$ takes the form (4.35) provided that $\sigma \ll 1$, so it is unsurprising that in the plot with very small σ (Figure 4.11) this analytic solution (equation 4.78, denoted with red dashed lines) overlaps with the weak GR solution (equation 4.49, shown with green dashed lines).

Moderate GR, $\epsilon_{\text{weak}} \ll \epsilon_{\text{GR}} \ll \epsilon_{\text{strong}}$

Perhaps more interesting is the asymptotic regime of *moderate* GR, defined as $\epsilon_{\text{weak}} \ll \epsilon_{\text{GR}} \ll \epsilon_{\text{strong}}$. In this regime one finds from (4.45) that

$$j_{\min} \approx \frac{\epsilon_{\text{GR}}}{3(1 + 5\Gamma)j_+^2} = \gamma j_- = \frac{2\epsilon_{\text{GR}}}{\epsilon_{\text{weak}}} j_- \gg j_-, \quad (4.53)$$

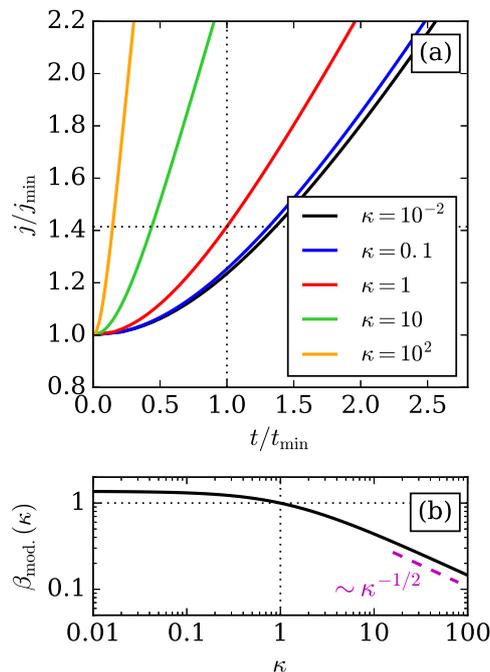


Figure 4.8: Similar to Figure 4.7 except now for the moderate GR regime (§4.4.4). In panel (a) the solution for $j(t)$ is defined implicitly by equation (4.55) and we plot it for various values of κ (equation (4.56)). In panel (b) we show $\beta_{\text{mod}}(\kappa)$, which is the time (in units of t_{\min}) over which j changes from j_{\min} to $\sqrt{2}j_{\min}$ in this regime. Dotted lines show $\kappa = 1$ and $\beta_{\text{mod}} = 1$, while a dashed magenta line shows the scaling $\beta_{\text{mod}} \propto \kappa^{-1/2}$ for $\kappa \gg 1$.

i.e. a significant perturbation of j_{\min} away from j_- , resulting in a significantly reduced maximum eccentricity e_{\max} .

To determine the time dependence of $j(t)$ around j_{\min} we neglect j_-^2 compared to j^2 in the first square bracket in (4.41) and find

$$\frac{dj}{dt} \approx \pm \frac{6C}{Lj} \sqrt{(25\Gamma^2 - 1)j_+^2 j_0^2 [j - \gamma j_-] [j + \sigma j_-]}. \quad (4.54)$$

Integration of (4.54) gives an implicit solution for $j(t)$ in the form

$$\frac{t}{t_{\min}} = \int_1^{j/j_{\min}} \frac{x dx}{\sqrt{(x-1)(x+\kappa)}}, \quad (4.55)$$

where

$$\kappa \equiv \frac{\sigma}{\gamma} = \frac{j_+^2}{j_0^2} \frac{5\Gamma + 1}{5\Gamma - 1}. \quad (4.56)$$

Note that t_{\min} in (4.55) is still defined by equation (4.35) but taking j_{\min} equal to its GR-modified value, namely γj_- . In Figure 4.8a we plot this implicit solution for various values of κ . Note also that $\kappa = 1$ (red line) gives precisely the solution j/j_{\min} in the form

¹³Note however that on the horizontal axis we plot time in units of t'_{\min} (equation (4.79)) rather than t_{\min} , as explained in Appendix 4.C.

(4.35), and so unsurprisingly $j/j_{\min} = \sqrt{2}$ coincides with $t/t_{\min} = 1$ in that case. As we increase κ we see that the time spent near the minimum j decreases (when measured in units of t_{\min} , which itself also depends on κ).

We can get a better feel for the quantity κ in the moderate GR regime using the fact that in this regime, $\epsilon_{\text{GR}} \sim j_{\min} \gg \Theta^{1/2}$. Then from (4.15), (4.16) we get

$$\Sigma \approx \frac{\epsilon_{\text{GR}}}{6j_{\min}}, \quad \text{and} \quad D \approx 1 + \frac{10\Gamma}{1-5\Gamma} \left(1 - \frac{\epsilon_{\text{GR}}}{30\Gamma j_{\min}}\right)^{-1}. \quad (4.57)$$

Plugging these results into (4.13) and (4.14) and inserting the resulting expressions into (4.56), we find

$$\kappa \approx \left(\frac{30\Gamma j_{\min}}{\epsilon_{\text{GR}}} - 1\right)^{-1}. \quad (4.58)$$

Since $j_{\min} \sim \epsilon_{\text{GR}}$ we typically expect the first term in the bracket to be $\gg 1$, resulting in $\kappa \ll 1$. However, as we will see in Appendix 4.C.3, much larger values of κ are also possible. Moreover, in writing down the scaling $\epsilon_{\text{GR}} \sim j_{\min} \gg \Theta^{1/2}$ we have implicitly assumed that Θ is fixed while ϵ_{GR} is increased in equation (4.45). This is not how things work when we include GW emission and track the evolution of ϵ_{GR} , Θ for a shrinking binary — see Chapter 5. In that case both Θ and ϵ_{GR} increase as the semimajor axis a decreases, while $\Theta^{1/2}/j_{\min} = |\cos i_{\min}|$ remains very nearly constant. Thus the results of this subsection will need to be modified there (for instance we will find that $\kappa < 0$ is common).

Analogous to §4.4.4, by setting $j/j_{\min} = \sqrt{2}$ on the right hand side of (4.55) and $t = t_{\min}^{\text{mod}}$ on the left, we find that the time for j to increase from j_{\min} to $\sqrt{2}j_{\min}$ in the moderate GR regime is

$$t_{\min}^{\text{mod}}(\kappa) = \beta_{\text{mod}}(\kappa)t_{\min}, \quad (4.59)$$

where $\beta_{\text{mod}}(\kappa)$ is plotted as a function of κ in Figure 4.8b. Clearly when $\kappa \sim 1$ (which is true for Γ not too close to $1/5$) we have $\beta \sim 1$ and so $t_{\min}^{\text{GR}} \sim t_{\min}$. But for $\kappa \gg 1$ the time spent in the high eccentricity state is somewhat reduced, with a scaling $t_{\min}^{\text{GR}} \propto \kappa^{-1/2}t_{\min}$. However for this to be a significant effect requires rather extreme values of $\kappa \gg 1$.

Finally, in panels (a) and (b) of Figures 4.13 and 4.16 we compare the moderate GR solution (4.55), shown with dashed cyan lines, to direct numerical integration of the DA equations of motion, shown in solid blue. In both cases the moderate GR solution provides an excellent fit to the numerical result despite ϵ_{GR} only being slightly larger than ϵ_{weak} .

4.4.5 High- e behaviour in the strong GR limit

The final asymptotic case to consider is that of strong GR, $\epsilon_{\text{GR}} \gg \epsilon_{\text{strong}}$. This regime is important for understanding the later stages of evolution of shrinking compact object binaries. Indeed, GW emission eventually brings any merging binary to a small enough semimajor axis to put it in this regime.

In this limit GR precession is the dominant effect, exceeding the secular effects of the external tide — see e.g. panels (e) and (j) of Figures 4.1 and 4.2. Thus we anticipate that at high eccentricity the lowest order solution will be one of constant eccentricity and uniform prograde precession:

$$j(t) = j(0), \quad \omega = \omega(0) + \dot{\omega}_{\text{GR}}(0)t, \quad (4.60)$$

where $\dot{\omega}_{\text{GR}}(0) = (C/L)\epsilon_{\text{GR}}/j^2(0)$ — see equation (4.1). High eccentricity can therefore only be achieved if $j(0) \ll 1$ to start with. This is actually a highly relevant scenario in practice because it is at very high eccentricity that GW emission, and hence the shrinkage of a and the growth of ϵ_{GR} , is concentrated. Binaries periodically torqued to very high eccentricity by cluster tides eventually become trapped in a highly eccentric orbit as they enter the strong GR regime (Chapter 5). Their phase space trajectories are then well described by (4.60).

Interestingly, the minimum angular momentum $j_{\text{min}} = j(0)$ predicted by the solution (4.60) can still be described by the expression (4.45) in the limit $\epsilon_{\text{GR}} \gg \epsilon_{\text{strong}}$. To see this we note that for $\epsilon_{\text{GR}} \gg \epsilon_{\text{strong}}$ equation (4.45) gives $j_{\text{min}} \approx \epsilon_{\text{GR}}/(j_+^2 \epsilon_{\text{strong}})$. Then we take the expression (4.39) for j_+^2 , and substitute into it the value of Σ we get by taking $\epsilon_{\text{GR}} \gg 1$ in (4.15)-(4.16), namely $\Sigma \approx \epsilon_{\text{GR}}(1 - e_0^2)^{-1/2}/6$. Putting these pieces together we find $j_{\text{min}} \approx \sqrt{1 - e_0^2} = j(0)$. Thus the solution (4.45) interpolates smoothly between the different asymptotic GR regimes.

4.4.6 Evolution of $\omega(t)$ and $\Omega(t)$ as $e \rightarrow 1$

So far we focused on understanding the behaviour of $j(t)$. Once this is determined one can understand the evolution of other orbital elements as well. In particular, since the Hamiltonian (4.2) is conserved, we can use equations (4.7)-(4.8) to express $\cos 2\omega$ entirely in terms of $j(t)$ and conserved quantities, leading to an explicit analytical expression for $\omega(t)$. Finally one can plug this $\cos 2\omega(t)$ and $j(t)$ into the equation of motion (4.11) for $\Omega(t)$ and integrate the result. Together with $J_z(t) = J_z(0) = \text{const.}$, this constitutes a complete solution to the DA, test-particle quadrupole problem with GR precession.

Unfortunately this proposed solution for $\omega(t)$ and $\Omega(t)$ is very messy for arbitrary values of j . Luckily, in the high eccentricity regime $j \ll 1$, one can make substantial

progress by (i) making the additional (and often well-justified) assumption given by equation (4.77) and (ii) adopting the ansatz¹⁴ (4.35) for $j(t)$. Then, as we show in Appendix 4.C, one can derive relatively simple explicit analytical solutions for $\omega(t)$ and $\Omega(t)$. These solutions work very well as long as equation (4.35) is a good approximation to the $j(t)$ behaviour near the peak eccentricity, as we verify numerically in §4.C.3. To our knowledge, an explicit high- e solution of this form accounting for GR precession has not been derived before even for the LK problem. It can be used for instance in order to explore the short-timescale (i.e. non-DA) effects near peak eccentricity, which are important for accurate calculation of the LK-driven merger rate (Grishin et al. 2018).

4.5 Discussion

In this Chapter we have studied the impact of 1PN GR precession on secular evolution of binaries perturbed by cluster tides. A single dimensionless number Γ effectively encompasses all information about the particular tidal potential and the binary’s outer orbit within that potential. Meanwhile the relative strength of GR precession compared to external tides is characterised by the dimensionless number ϵ_{GR} (equation (4.5)).

In the main body of the Chapter we only discussed the systems with $\Gamma > 0$. Although the resulting dynamics are significantly complicated by bifurcations that occur at $\Gamma = \pm 1/5, 0$, for $\Gamma > 1/5$ our qualitative results are intuitive, falling in line with those gleaned from previous LK ($\Gamma = 1$) studies that accounted for GR precession. However, for $0 < \Gamma \leq 1/5$ we uncovered a completely new pattern of secular evolution, which we characterised in detail. Secular dynamics of binaries with negative Γ (possible for binaries on highly inclined outer orbits in strongly non-spherical potentials, see Chapter 2) and non-zero ϵ_{GR} is covered in Appendix 4.D. Like in Chapter 3 we find that the $\Gamma \leq 0$ regime splits into two further regimes, namely $-1/5 < \Gamma \leq 0$ and $\Gamma \leq -1/5$. It turns out that in both of these regimes the resulting phase space structures and maximum eccentricity behaviour are considerably more complex and counter-intuitive than for $\Gamma > 0$.

Furthermore, we have explored the evolution of binary orbital elements in the limit of very high eccentricity (§4.4). This investigation revealed a number of distinct dynamical regimes that are classified according to the value of ϵ_{GR} . In §4.5.1 we summarise and systematise these regimes based on their physical characteristics. In §4.5.2 we compare our study to the existing LK literature, and in §4.5.3 we discuss its limitations.

¹⁴Strictly speaking this ansatz is valid only for $\sigma \ll 1$, but is often a good approximation in the vicinity of j_{min} even for $\sigma \gg 1$ — see §4.C.3.

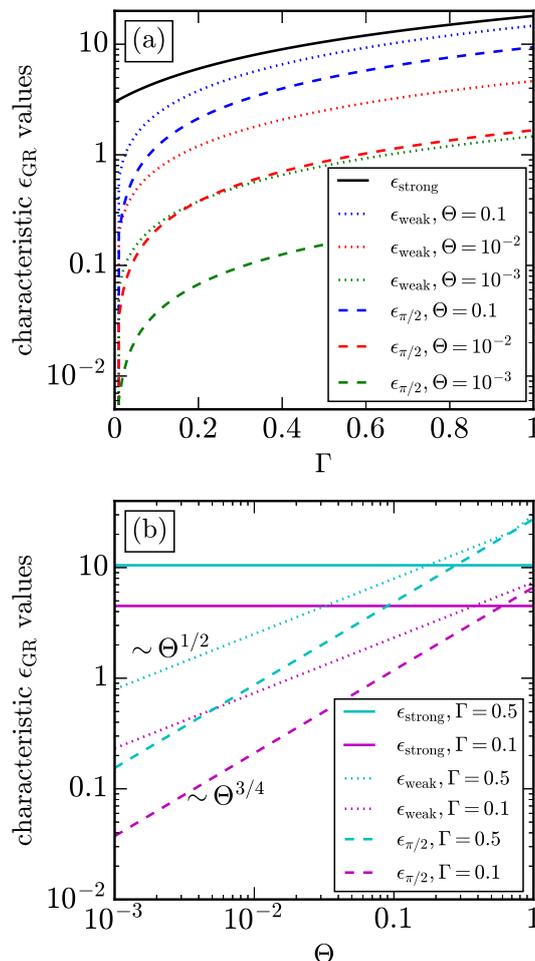


Figure 4.9: (a) Plot showing several characteristic values of the GR strength ϵ_{GR} as functions of $\Gamma > 0$: ϵ_{strong} (black solid line), ϵ_{weak} (dotted lines) and $\epsilon_{\pi/2}$ (dashed lines). The values of ϵ_{weak} and $\epsilon_{\pi/2}$ depend on Θ so we show them for three Θ values, namely 0.1 (blue), 10^{-2} (red) and 10^{-3} (green). (b) Same, but now as a function of Θ , for $\Gamma = 0.5$ (cyan) and $\Gamma = 0.1$ (magenta).

4.5.1 Summary of ϵ_{GR} regimes and their physical interpretation

In this Chapter we introduced three characteristic values of ϵ_{GR} , namely $\epsilon_{\pi/2}$ (equation (4.19)), ϵ_{weak} (equation (4.44)), and ϵ_{strong} (equation (4.25)). The first two scale with Θ in such a way that in the high- e limit, when $\Theta \ll 1$, one finds

$$\epsilon_{\pi/2} \lesssim \epsilon_{\text{weak}} \lesssim \epsilon_{\text{strong}}. \quad (4.61)$$

This hierarchy is illustrated in Figure 4.9, in which we show how¹⁵ $\epsilon_{\pi/2}$, ϵ_{weak} , ϵ_{strong} depend on both $\Gamma > 0$ and Θ . We see that decreasing Θ widens the gap between ϵ_{strong} and ϵ_{weak} . This is as expected because small Θ tends to promote high- e , and since $\dot{\omega}_{\text{GR}} \propto \epsilon_{\text{GR}}/(1 - e^2)$, the higher is e , the smaller is the critical value of ϵ_{GR} at which GR

¹⁵to calculate ϵ_{weak} for this figure we set $e_0 = \epsilon_{\text{GR}} = 0$, so that j_{\pm}^2 is given by (4.13) with $\Sigma = (1 + 5\Gamma + 10\Gamma\Theta)/2$.

effects become important. Note however that even for $\Theta = 10^{-3}$ the difference between ϵ_{strong} and ϵ_{weak} does not exceed $\sim 10^2$. Since $\epsilon_{\text{GR}} \propto a^{-4}$ (equation (4.5)), a relatively small change in semimajor axis a can easily shift ϵ_{GR} from one asymptotic regime to another.

These characteristic values of ϵ_{GR} allow us naturally to delineate four important regimes of secular dynamics:

- very weak GR: $\epsilon_{\text{GR}} \lesssim \epsilon_{\pi/2}$,
- weak GR: $\epsilon_{\pi/2} \lesssim \epsilon_{\text{GR}} \lesssim \epsilon_{\text{weak}}$,
- moderate GR: $\epsilon_{\text{weak}} \lesssim \epsilon_{\text{GR}} \lesssim \epsilon_{\text{strong}}$,
- strong GR: $\epsilon_{\text{GR}} \gtrsim \epsilon_{\text{strong}}$.

Based on our findings in §§4.3-4.4 we now provide a description of the basic features of each regime.

Very weak GR In this limit (below the dashed curves in Figure 4.9) GR precession has essentially no effect on the dynamics for $\Gamma > 1/5$. More precisely, GR is too weak to affect either the locations of the fixed points at $\omega = \pm\pi/2$, which are given by equation (4.20), or the maximum eccentricity reached by the binary at the same ω in the course of its tide-driven secular evolution, given by equation (4.47). Thus all $\Gamma > 1/5$ results of Chapter 3, which were derived for $\epsilon_{\text{GR}} = 0$, are valid¹⁶. However, for $0 < \Gamma \leq 1/5$ an important modification arises if $6(1 - 5\Gamma)\Theta^{3/2} < \epsilon_{\text{GR}} \lesssim \epsilon_{\pi/2}$, which is that saddle points appear at $\omega = 0, \pi$. These saddles do not exist for $\epsilon_{\text{GR}} = 0$ (see equation (4.28) and §4.3.3), but they do change the maximum eccentricity reached by the binary — see Appendix 4.B.

Weak GR In this regime (between the dashed and dotted curves in Figure 4.9) GR precession starts to modify the j locations of the fixed points at $\omega = \pm\pi/2$, which are now given by equation (4.21). At the same time, GR precession does not appreciably change j_{min} (or equivalently e_{max}), which stays close to its $\epsilon_{\text{GR}} = 0$ value j_- (see equation (4.47)). If $\sigma \gg 1$ then GR also modifies the time spent in the high eccentricity state (equation (4.52)).

Moderate GR In this regime (between the dotted and solid curves in Figure 4.9) GR precession modifies not only the locations of the fixed points but also the values of j_{min} (and hence of e_{max}), now given in equation (4.53). GR also modifies the time spent in the high- e state (equation (4.59)). In other words, in this regime GR precession presents an efficient barrier suppressing the maximum eccentricity reached by the binary in the course of its secular evolution.

¹⁶Note that this is not necessarily true in the SA approximation, in which even a very weak amount of GR can totally alter the phase space dynamics if the binary is capable of reaching extremely high eccentricity — see Chapter 6.

Strong GR In this limit (above the solid curves in Figure 4.9) GR precession dominates the binary dynamics at all times; the quantities j_{\pm} are significantly affected by GR precession (equation (4.37)) and all fixed points in the phase portrait disappear (equations (4.26), (4.28)). Cluster tides drive only very small eccentricity oscillations on top of uniform GR precession, so that e is roughly constant — see equation (4.60).

In Table 4.2 we summarise the main features of the asymptotic ϵ_{GR} regimes that we have found in this and previous sections.

We may use this regime separation to shed light on the physical meaning of the characteristic ϵ_{GR} values introduced in this work. To do so, we first note that the GR precession rate (4.1) can be written as $\dot{\omega}_{\text{GR}}(j) = \dot{\omega}_{\text{GR}}|_{e=0} j^{-2} \sim \epsilon_{\text{GR}} t_{\text{sec}}^{-1} j^{-2}$ (see the definition (4.5)). Next, consider some arbitrary cluster tide-driven process occurring on a characteristic timescale t_{ch} . GR precession will affect this process if ϵ_{GR} is such that

$$\dot{\omega}_{\text{GR}}(j)t_{\text{ch}} \sim 1, \quad \text{i.e.} \quad \epsilon_{\text{GR}} \frac{t_{\text{ch}}}{t_{\text{sec}}} j^{-2} \sim 1. \quad (4.62)$$

If ϵ_{GR} satisfies (4.62), or exceeds that value, then GR breaks the coherence of the tidal torque over the timescale $\sim t_{\text{ch}}$, and so GR precession substantially interferes with the secular evolution. We now demonstrate how this simple physical argument leads one to the critical values ϵ_{strong} , ϵ_{weak} and $\epsilon_{\pi/2}$.

First, in the strong GR regime we expect GR precession to dominate binary evolution at all times, even for near-circular orbits. Setting $t_{\text{ch}} \sim t_{\text{sec}}$ and $j \sim 1$ we obtain $\epsilon_{\text{GR}} \sim 1$, which is consistent with the definition (4.25) of ϵ_{strong} up to a numerical coefficient.

Second, in the moderate GR regime, we anticipate that GR precession will present an effective barrier that stops the decrease of j if t_{ch} is the characteristic timescale of secular evolution near the eccentricity peak. In §4.4 we find quite generally this timescale to be $t_{\text{ch}} \sim t_{\text{min}} \sim j_{\text{min}} t_{\text{sec}}$ — see e.g. equations (4.35) and (4.59). Plugging this into the condition (4.62) and evaluating $\dot{\omega}_{\text{GR}}$ at j_{min} we immediately find that $j_{\text{min}} \sim \epsilon_{\text{GR}}$, in agreement with equation (4.53). When the GR barrier first emerges at the transition between weak and moderate regimes, j_{min} is still well approximated by the $\epsilon_{\text{GR}} = 0$ solution $j_- \sim \Theta^{1/2}$ (see equation (4.39)). As a result, the ϵ_{GR} value corresponding to this transition is $\sim \Theta^{1/2}$, in agreement with the definition (4.44) of ϵ_{weak} .

Third, we expect fixed points in the phase portrait at $\omega = \pm\pi/2$ to be substantially displaced by GR precession when $\dot{\omega}_{\text{GR}}(j_f)$ becomes comparable to the characteristic secular frequency $\dot{\omega}$ of libration around a fixed point. Since we are interested in the displacement of j_f by an amount $\sim j_f$, we take this $\dot{\omega}$ from the $H^* = \text{const.}$ contour centred on the $\omega = \pm\pi/2$ fixed point and with vertical extent $\sim j_f$. Plugging $j \sim j_f$ into the equation (4.9) and using the expression (4.17) for j_f we find $\dot{\omega} \sim \Theta^{1/4} t_{\text{sec}}^{-1}$, so that in

this case $t_{\text{ch}} \sim \Theta^{-1/4} t_{\text{sec}}$. Substituting this into the condition (4.62) and again setting $j \sim j_{\text{f}} \sim \Theta^{1/4}$ we find $\epsilon_{\text{GR}} \sim \Theta^{3/4}$ for the transition between the weak and very weak GR regimes. This agrees with the definition of $\epsilon_{\pi/2}$ in equation (4.19).

Note that while these considerations allow us to understand the scalings of characteristic ϵ_{GR} values with Θ , one still needs the full analysis presented in §§4.3-4.4 to obtain the numerical coefficients, which are actually quite important. Indeed, equations (4.19), (4.44), and (4.25) feature constant numerical factors which can substantially exceed unity, especially for the LK case of $\Gamma = 1$.

4.5.2 Relation to LK studies

Many authors who studied the LK mechanism and its applications have included 1PN GR precession in their calculations. The maximum eccentricity of an initially near-circular binary undergoing LK oscillations (i.e. the $\Gamma = 1$ limit of §4.3.4) was derived by Miller & Hamilton (2002), Blaes et al. (2002), Wen (2003), Fabrycky & Tremaine (2007), and Liu, Muñoz, et al. (2015). Of these, Fabrycky & Tremaine (2007) and Liu, Muñoz, et al. (2015) also produced plots very similar to Figure 4.5 that show how increasing ϵ_{GR} decreases the maximum eccentricity achieved by initially near-circular binaries. Various authors have derived equations identical to, or very similar to, the quartic (4.30) and the weak-to-moderate maximum eccentricity solution (4.45) in the LK limit — see for instance equation (A7) of Blaes et al. (2002), equation (8) of Wen (2003), equation (A6) of Veras & Ford (2010), and equations (64)-(65) of Grishin et al. (2018). Of course, because these studies only work with $\Gamma = 1$, the rather non-intuitive behaviour for $\Gamma < 1/5$ revealed in §4.3.4 and Appendix 4.D.3 has not been unveiled before. Moreover, to our knowledge no previous study has presented a clear classification of the different ϵ_{GR} regimes (which we do in §4.5.1), even in LK theory.

The quantitative results in the aforementioned papers have been employed in many practical calculations. Typically one simply adds the term (4.1) to the singly- or doubly-averaged equations of motion along with any other short range forces or higher PN effects. In population synthesis calculations of compact object mergers (Antonini & Perets 2012; Antonini, Murray, et al. 2014; Silsbee & Tremaine 2017; Liu & Lai 2018) one often puts a sensible lower limit on the semimajor axis distribution below which GR is so strong that sufficient eccentricity excitation is impossible. As explained in §2.1 of Rodriguez & Antonini (2018) there are at least two ways to decide when GR dominates. One method is to take j_{min} corresponding to the pericentre distance that needs to be reached according to the problem at hand, and then equate $\dot{\omega}_{\text{GR}}(j_{\text{min}})$ with the precession rate due to the tidal perturbations (see their equation (29)), which is what we did in §3.8. As discussed in §4.5.1, this method would set a rough upper limit of $\epsilon_{\text{GR}} \lesssim j_{\text{min}}$; see equation (4.53)

for a more accurate expression. A second method is to demand that $\omega = \pm\pi/2$ fixed points do exist in the phase portrait (Fabrycky & Tremaine 2007) allowing for substantial eccentricity excitation to occur starting from the near-circular orbits, which is equivalent to $\epsilon_{\text{GR}} \lesssim \epsilon_{\text{strong}}$. However, this is not a very stringent requirement, and does not guarantee that the majority of systems with such ϵ_{GR} would reach the required j_{min} — many of them will be stopped by the GR barrier at eccentricities much lower than needed. The former method of setting an upper limit on ϵ_{GR} is typically more stringent and allows more efficient selection of systems for Monte Carlo population synthesis (see Figure 4.9).

With regard to phase space structure, the only study we know of that resembles our §4.3 is that by Iwasa & Seto (2016). They considered a hierarchical triple consisting of a star on an orbit around a supermassive black hole (SMBH), with another massive black hole also orbiting the SMBH on a much larger, circular orbit and acting as the perturber of the star-SMBH ‘binary’. Their §C provides a brief explanation of the phase space behaviour as a parameter they call γ , which is equivalent to our $\epsilon_{\text{GR}}/3$, is varied. Since the LK problem has $\Gamma = 1 > 1/5$, their Figure 2 is qualitatively the same as our Figure 4.1.

4.5.3 Approximations and limitations

To derive the Hamiltonian (4.3) we truncated the perturbing tidal potential at the quadrupole level. This is justified if the semimajor axis of the binary is much smaller than the typical outer orbital radius. Next order corrections to the perturbing potential — so called octupole terms — are routinely accounted for in LK studies (Naoz, Kocsis, et al. 2013; Will 2017). In §2.E we provided the octupole correction to (4.3) for arbitrary Γ . When octupole-order effects are important, the maximum eccentricity can actually be *increased* by GR precession (Ford, Kozinsky, et al. 2000; Naoz, Kocsis, et al. 2013; Antonini, Murray, et al. 2014). However for the applications we have in mind, e.g. a compact object binary of $a \sim 10\text{AU}$ orbiting a stellar cluster at $\sim 1\text{pc}$, octupole corrections are typically negligible.

We also employed the test particle approximation, which is valid if the outer orbit contains much more angular momentum than the inner orbit. One can relax the test particle approximation: in particular, this is often necessary for weakly-hierarchical triples. Anderson, Lai, et al. (2017) made a detailed study of the ‘inclination window’ that allows fixed points to exist in the (quadrupole) LK phase space for different ϵ_{GR} , as one varies the ratio of inner to outer orbital angular momenta. We recover their results in the test particle limit valid for our applications.

Finally, several of the results derived at very high eccentricity (§4.4) are rather delicate when Γ is close to $\pm 1/5$ or when the binary’s phase space trajectory is close to a separatrix. These are not major caveats; for instance, in a given stellar cluster potential only a small fraction of binaries will have Γ values close enough to $1/5$ to be affected (Chapter 2).

4.6 Summary

In this Chapter we completed our investigation of doubly-averaged (test-particle quadrupole) cluster tide-driven binary dynamics in the presence of 1PN general relativistic pericentre precession. Throughout, we parameterised the strength of GR precession relative to tides using the dimensionless number ϵ_{GR} (equation (4.5)). We can summarise our results as follows:

- We investigated the effect of non-zero ϵ_{GR} on phase space morphology. For values of ϵ_{GR} much less than a critical value ϵ_{strong} , bifurcations in the dynamics happen at $\Gamma = \pm 1/5, 0$, so that we must consider four Γ regimes separately. We found that for $\Gamma \leq 1/5$ a non-zero ϵ_{GR} can lead to entirely new phase space morphologies, including (previously undiscovered) fixed points located at $\omega = 0, \pm\pi$.
- We presented general recipes for computing the locations of fixed points in the phase portrait, for determining whether a given phase space trajectory librates or circulates, and for finding its maximum eccentricity, for arbitrary ϵ_{GR} .
- We considered how the maximum eccentricity reached by an initially circular binary is affected by GR precession. For $\Gamma > 1/5$ the intuitive picture holds that a larger ϵ_{GR} leads to a lower maximum eccentricity, but this is not always the case for $\Gamma \leq 1/5$.
- We delineated four distinct regimes of secular evolution with GR precession depending on the value of ϵ_{GR} — ‘strong GR’, ‘moderate GR’, ‘weak GR’, and ‘very weak GR’ — and provided physical justification for transitions between them.
- We also studied secular evolution with GR precession in the limit of very high eccentricity. We determined the GR-induced modifications to the minimum angular momentum j_{min} achieved by the binary and the time dependence of $j(t)$ near the eccentricity peak, which can be rather non-trivial.
- We also provided an approximate analytic description for the evolution of other orbital elements — pericentre and nodal angles — near the eccentricity peak, accounting for the GR precession.

In Chapter 5 we will use these results to understand the long-term evolution of compact object binaries whose inner orbits decay through GW emission, ultimately leading to their mergers and the production of LIGO/Virgo GW sources. Furthermore, these results will inform our study of the effect of short-timescale fluctuations (‘singly-averaged effects’) on binaries undergoing cluster tide-driven secular evolution (Chapter 6), as well as our population synthesis calculation of merger rates (Chapter 7).

Appendices

4.A Mathematical details of phase space behaviour for $\Gamma > 0$

In this Appendix we provide some mathematical details for the results quoted in §4.3.

4.A.1 Fixed points at $\omega = \pm\pi/2$

Putting $\omega = \pm\pi/2$ and $d\omega/dt = 0$ into equation (4.9) gives us the following quartic equation for j values of the fixed points, which we will call $j_{f,\pi/2}$:

$$j_{f,\pi/2} \left(j_{f,\pi/2}^3 - \frac{\epsilon_{\text{GR}}}{6(1+5\Gamma)} \right) = \frac{10\Gamma\Theta}{1+5\Gamma}. \quad (4.63)$$

For any $\Gamma > 0$ the right hand side of (4.63) is obviously positive. Thus, for there to be a positive (not necessarily physical) solution to equation (4.30) a necessary but insufficient requirement is that

$$j_{f,\pi/2} \geq \left[\frac{\epsilon_{\text{GR}}}{6(1+5\Gamma)} \right]^{1/3}, \quad (4.64)$$

which, since $j < 1$, in turn means that ϵ_{GR} must necessarily be $\leq 6(1+5\Gamma)$. By differentiating (4.63) it is then easy to show that

$$\left(\frac{\partial j_{f,\pi/2}}{\partial \Theta} \right)_{\epsilon_{\text{GR}}} = \frac{10\Gamma}{1+5\Gamma} \left(4j_{f,\pi/2}^3 - \frac{\epsilon_{\text{GR}}}{6(1+5\Gamma)} \right)^{-1} > 0, \quad (4.65)$$

$$\left(\frac{\partial j_{f,\pi/2}}{\partial \epsilon_{\text{GR}}} \right)_{\Theta} = \frac{j_{f,\pi/2}}{6(1+5\Gamma)} \left(4j_{f,\pi/2}^3 - \frac{\epsilon_{\text{GR}}}{6(1+5\Gamma)} \right)^{-1} > 0. \quad (4.66)$$

In other words, for $\Gamma > 0$ the fixed points at $(\omega, j) = (\pm\pi/2, j_{f,\pi/2})$ always get pushed to lower eccentricity when we increase ϵ_{GR} or Θ (see Figures 4.1, 4.2 & 4.6).

The criteria for these $\omega = \pm\pi/2$ fixed points to exist can be found by demanding that the condition (3.9) is obeyed, i.e. that $\sqrt{\Theta} < j_{f,\pi/2} < 1$. Let us begin by fixing Γ and Θ ; then, owing to the monotonic behaviour of $j_{f,\pi/2}(\epsilon_{\text{GR}})$ (equation (4.66)), we simply look for the ϵ_{GR} values that correspond to $j_{f,\pi/2} = \sqrt{\Theta}$ and $j_{f,\pi/2} = 1$. Doing so, we arrive straightforwardly at the condition (4.22) on ϵ_{GR} .

Next we wish to instead fix Γ and ϵ_{GR} and look for the resulting condition on Θ that allows the fixed points to exist. To begin with, we look for the critical Θ values for which $j_{f,\pi/2} = 1$ and $j_{f,\pi/2} = \sqrt{\Theta}$. The former is Θ_1 , the expression for which is given in equation (4.24), and the latter is Θ_2 which is determined implicitly through the equation

$$\Theta_2^{1/2} \left(\Theta_2 - \frac{10\Gamma}{1+5\Gamma} \right) = \frac{\epsilon_{\text{GR}}}{6(1+5\Gamma)}. \quad (4.67)$$

For $\Gamma > 0$ this equation can have meaningful ($0 \leq \Theta_2 \leq 1$) solutions only if $10\Gamma/(1+5\Gamma) \leq 1$, i.e. if $\Gamma \leq 1/5$. For $\Gamma > 1/5$ the value of Θ_2 has no physical significance. Next, to

determine the proper constraint on Θ we begin by setting $\Theta = 0$ in (4.63) — we see that the fixed point exists and has value $j_{f,\pi/2} = (\epsilon_{\text{GR}}/[6(1+5\Gamma)])^{1/3}$, so that $j_{f,\pi/2} > \sqrt{\Theta}$ at $\Theta = 0$. As we increase Θ , there are two possibilities. The first is that $j_{f,\pi/2}$ increases steeply enough that it reaches unity (at $\Theta = \Theta_1$) before it intersects $\sqrt{\Theta}$. In this case the constraint on Θ for fixed points to exist is $\Theta < \Theta_1$. The second scenario is that $j_{f,\pi/2}$ intersects $\sqrt{\Theta}$ (at Θ_2) before it reaches unity. Then for the inequality $\sqrt{\Theta} < j_{f,\pi/2}$ to be satisfied one needs $\Theta < \Theta_2$. Of course, since Θ_2 is only physically meaningful for $\Gamma \leq 1/5$, the second scenario can only occur in that Γ regime. This reasoning leads us to the constraint (4.23) for fixed points to exist at $\omega = \pm\pi/2$. In the limit $\epsilon_{\text{GR}} \rightarrow 0$ this constraint reduces to the non-GR constraint (4.18).

In summary, for $\Gamma > 0$, fixed points at $\omega = \pm\pi/2$ exist if the constraints on both Θ and ϵ_{GR} are satisfied simultaneously; thus they exist in the sub-volume of $(\Gamma, \Theta, \epsilon_{\text{GR}})$ space bounded by the inequalities (4.22), (4.23). We can use this information to understand Figure 4.4 in more detail. Recall that in this figure we plotted $\Theta_{\text{max}}(\Gamma)$, namely the maximum value of Θ for which fixed points exist at $\omega = \pm\pi/2$ for a given ϵ_{GR} (equation (4.23)). We now seek to understand separately the behaviour for $\Gamma > 1/5$ and $0 < \Gamma \leq 1/5$.

For $\Gamma > 1/5$, the lines in Figure 4.4 correspond to $\Theta_{\text{max}} = \Theta_1$ (equation (4.24)). Then $\partial\Theta_1/\partial\Gamma = \Gamma^{-2}(\epsilon_{\text{GR}} - 6)/60$, so that Θ_1 increases (decreases) monotonically with Γ for $\epsilon_{\text{GR}} > 6$ ($\epsilon_{\text{GR}} < 6$). For the special value $\epsilon_{\text{GR}} = 6$ we have $\Theta_1 = 1/2 = \text{const.}$, hence the straight horizontal brown line in Figure 4.4.

On the other hand, for $0 < \Gamma \leq 1/5$ we have $\Theta_{\text{max}} = \min[\Theta_1, \Theta_2]$. By equating $\Theta_1 = \Theta_2$ in equations (4.24), (4.67) it is straightforward to show that Θ_2 becomes smaller than Θ_1 when Γ is reduced below a critical value $\Gamma_{\text{crit}} = (6 - \epsilon_{\text{GR}})/30$, and that this happens at $\Theta_1 = \Theta_2 = 1$. This is reflected in Figure 4.4 — as we decrease Γ starting from $1/5$, the red ($\epsilon_{\text{GR}} = 0$), yellow ($\epsilon_{\text{GR}} = 3$) and green ($\epsilon_{\text{GR}} = 5$) lines transition from solid (Θ_1) to dotted (Θ_2) at the points $(1/5, 1)$, $(1/15, 1)$ and $(1/30, 1)$ respectively. For $\epsilon_{\text{GR}} > 6$ (blue, pink and black lines in Figure 4.4) we have $\Gamma_{\text{crit}} < 0$, so this transition never occurs for positive Γ . Finally, for the special value $\epsilon_{\text{GR}} = 6$ we have from equation (4.67) that $\Theta_2 = (1 - 5\Gamma)^{-1} > 1$, which is obviously greater than $\Theta_1 = 1/2$, so $\Theta_{\text{max}} = \Theta_1$. Hence the brown horizontal line in Figure 4.4 extends all the way to $\Gamma \rightarrow 0$.

4.A.2 Fixed points at $\omega = 0$

For $\Gamma \leq 1/5$ we found (e.g. Figure 4.2) that hitherto undiscovered fixed points could arise at $\omega = 0$. To find the eccentricity of these fixed points we plug $\omega = 0$ into $d\omega/dt = 0$ using equation (4.9). The result is a cubic equation for j with no quadratic or linear terms. The solution is $j = j_{f,0}$ with $j_{f,0}$ given in equation (4.27), which is physically meaningful only

for $0 < \Gamma \leq 1/5$ (i.e. fixed points at $\omega = 0$ do not exist for $\Gamma > 1/5$). The determinant of the Hessian matrix of $H^*(\omega, j)$ evaluated at the point $(0, j_{f,0})$ is equal to

$$-180\epsilon_{\text{GR}}\Gamma \times \frac{(j_{f,0}^2 - \Theta)(1 - j_{f,0}^2)}{j_{f,0}^5}. \quad (4.68)$$

Clearly for $0 < \Gamma \leq 1/5$ the determinant (4.68) is negative whenever the fixed point $j_{f,0}$ exists, so $(\omega, e) = (0, e_{f,0})$ is necessarily a saddle point in the phase portrait, consistent with Figures 4.2b,c,h.

4.A.3 Does a given orbit librate or circulate?

Here we show how to determine whether a phase space trajectory is librating or circulating, given $\Gamma, \Theta, \epsilon_{\text{GR}}$ and the initial phase space coordinates (ω_0, e_0) . For $\Gamma > 0$, librating orbits cannot cross $\omega = 0$ and so any trajectory that passes through $\omega = 0$ must be circulating¹⁷. Therefore we can figure out whether an orbit librates or circulates by determining whether it crosses $\omega = 0$. Plugging $\omega = 0$ into $H^*(\omega, j)$ gives us a depressed cubic polynomial:

$$j^3 - j_0^2 j + q = 0, \quad (4.69)$$

where

$$q \equiv \frac{\epsilon_{\text{GR}}}{3(1 - 5\Gamma)}, \quad (4.70)$$

and we used the definition (4.14), (4.16) of j_0^2 , which need not be positive. We call the real roots of this polynomial $j(\omega = 0)$. In the limit $\epsilon_{\text{GR}} = 0$ we have $q = 0$ and so we find $j(\omega = 0) = j_0$, recovering the expression for $j(\omega = 0)$ from equation (3.17). For $\epsilon_{\text{GR}} \neq 0$, the nature of the roots of (4.69) depends on the sign of the discriminant

$$\Delta \equiv 4j_0^6 - 27q^2. \quad (4.71)$$

We can evaluate Δ given $(\omega_0, e_0, \Theta, \Gamma, \epsilon_{\text{GR}})$. There are then a few different cases to consider:

- If $\Delta < 0$, equation (4.69) has one real root, which may or may not be physical. If $j_0^2 > 0$ then this root can be written as

$$j(\omega = 0) = -2\frac{|q|}{q}\sqrt{\frac{j_0^2}{3}} \cosh\left(\frac{1}{3}\operatorname{arccosh}\left[\frac{3|q|}{2j_0^2}\sqrt{\frac{3}{j_0^2}}\right]\right), \quad (4.72)$$

whereas for $j_0^2 < 0$ it is given by

$$j(\omega = 0) = -2\sqrt{\frac{-j_0^2}{3}} \sinh\left(\frac{1}{3}\operatorname{arcsinh}\left[\frac{-3q}{2j_0^2}\sqrt{\frac{-3}{j_0^2}}\right]\right). \quad (4.73)$$

Once $j(\omega = 0)$ has been determined, the orbit circulates if $\sqrt{\Theta} < j(\omega = 0) < 1$, and librates otherwise.

¹⁷This general statement does not hold for $\Gamma \leq 0$ — see Appendix 4.D.

- If $\Delta > 0$ (which necessarily requires $j_0^2 > 0$) there are three distinct real roots, and they can be expressed as

$$j(\omega = 0) = 2\sqrt{\frac{j_0^2}{3}} \cos\left(\frac{1}{3} \cos^{-1}\left[\frac{-3q}{2j_0^2} \sqrt{\frac{3}{j_0^2}}\right] - \frac{2\pi k}{3}\right), \quad (4.74)$$

for $k = 0, 1, 2$. From the theory of polynomial equations we also know that the product of the three real roots of (4.69) is $-q = \epsilon_{\text{GR}}/[3(5\Gamma - 1)]$ and their sum is 0. For $\Gamma > 1/5$ this implies that two roots (namely $k = 1, 2$) must be negative and one ($k = 0$) positive. Thus the orbit circulates if the $k = 0$ solution lies in $(\sqrt{\Theta}, 1)$, and librates otherwise.

For $0 < \Gamma \leq 1/5$, one root ($k = 2$) must be negative and the other two ($k = 0, 1$) positive. If either or both of the two positive roots lies in $(\sqrt{\Theta}, 1)$ then the orbit circulates. If neither of them do then it librates. The case of both positive roots lying in $(\sqrt{\Theta}, 1)$ corresponds to two coexisting families of circulating orbits that share values of H^* , one above $e_{f,0}$ and one below, as in Figure 4.2b,c,h. To determine the family of circulating orbits to which the trajectory belongs we compare its initial eccentricity e_0 with that of the saddle point $e_{f,0}$. If $e_0 > e_{f,0}$ then the orbit circulates in the family ‘above’ the saddle point, and vice versa.

4.B High-eccentricity behaviour for orbits whose eccentricity maxima are found at $\omega = 0$

When GR is switched off, the only binaries whose eccentricity is maximised at $\omega = 0$ are those on circulating phase space trajectories in the regime $0 < \Gamma \leq 1/5$ (e.g. Figure 4.6f; see Chapter 3 for a thorough discussion). The minimum j in this case is $j_{\min} = j_0$ (Chapter 3), which is given in equation (4.14). For this to correspond to very high eccentricity one needs $D \approx 1$, and the orbit must sit very close to the separatrix between librating and circulating orbits, which can be hard to achieve in practice.

Nevertheless, suppose $j_0 \sim \Theta^{1/2} \ll 1$ for $\epsilon_{\text{GR}} = 0$; then for e_{\max} not to be changed radically when we do include GR, a necessary but insufficient condition is (4.40). Finding the minimum j at $\omega = 0$ requires that we set the final square bracket in (4.12) to zero, which is the same as solving the depressed cubic (4.69). In §4.A.3 we explained how to determine the appropriate explicit solution to (4.69) for arbitrary initial conditions. In particular, we note that the $\epsilon_{\text{GR}} = 0$ solution $j_{\min} = j_0$ corresponds exactly to equation (4.74) with $k = 0$. However, the general solutions for $\epsilon_{\text{GR}} \neq 0$ are not very enlightening. We can make some analytical progress if we further assume that

$$\epsilon_{\text{GR}} \ll |1 - 5\Gamma|j_0^3. \quad (4.75)$$

If (4.75) is true, then the first order solution for finite ϵ_{GR} is

$$j_{\min} \approx j_0 \left[1 - \frac{\epsilon_{\text{GR}}}{6(1-5\Gamma)j_0^3} \right] = j_0 \left[1 - \left(\frac{j_{f,0}}{j_0} \right)^3 \right]. \quad (4.76)$$

In other words, since $j_0^2 \sim \Theta \ll 1$ by construction, j_{\min} starts to substantially deviate from j_0 when the saddle points appear at $\omega = 0$ — see equation (4.28) and §4.3.3. Note that the condition (4.75) is very stringent and requires that the binary be deep in the very weak GR regime (§4.5.1), so (4.76) may not be useful in practice.

4.C Analytic solution for orbital elements at high eccentricity

In this Appendix we present an analytic solution to the DA equations of motion for all orbital elements in the limit of high eccentricity, assuming $\Gamma > 0$. To do this we will make the following four assumptions:

- (I) $j = j_{\min} \ll 1$ is realised at $\omega = \pm\pi/2$,
- (II) Weak or very weak GR, i.e. $\epsilon_{\text{GR}} \ll \epsilon_{\text{weak}}$,
- (III) $\sigma \ll 1$ (equation (4.43)),
- (IV) $j^4/\Theta \ll 1$.

Assumptions (I)-(III) are familiar from §4.4.4 (and of course if we set $\epsilon_{\text{GR}} = 0$ then (II) and (III) are satisfied automatically). Taken together, assumptions (I)-(III) imply, in particular, that $j(t)$ takes the form (4.35) with $j_{\min} = j_-$, $j_1 = j_+$, $j_2 = j_0$, which can be seen by expanding the weak GR equation (4.48) for $\sigma \ll 1$.

However, assumption (IV) is new. It is equivalent to the requirement that

$$\frac{j}{\cos^2 i_{\min}} \ll 1, \quad \text{where} \quad \cos^2 i_{\min} \equiv \Theta/(1 - e_{\max}^2) \quad (4.77)$$

is the cosine of the binary's minimum inclination. Assumption (IV) is nearly always satisfied at high eccentricities since we normally have¹⁸ $j^2 \sim \Theta$. The additional assumption (IV) allows us to take the solution for $j(t)$ from (4.35) that we got using assumptions (I)-(III) and simplify the expression for t_{\min} . The result is:

$$j(t) = j_{\min} \sqrt{1 + \left(\frac{t}{t'_{\min}} \right)^2}, \quad (4.78)$$

¹⁸Indeed, equation (4.17), which is valid in the very weak GR regime, tells us that $j_f^4 \sim \Theta \ll 1$ and we know that j_f then provides an upper bound on j_{\min} for $\Gamma > 1/5$.

where¹⁹

$$t'_{\min} \equiv \frac{j_{\min}^3}{60\Gamma\sqrt{\Theta(j_{\min}^2 - \Theta)}} \frac{L}{C}. \quad (4.79)$$

We note that t'_{\min} diverges as $j_{\min} \rightarrow \Theta$, that is as $e_{\max} \rightarrow e_{\lim}$. This is as expected from e.g. Figure 4.6a, since trajectories that approach e_{\lim} become ever ‘flatter’ in the vicinity of $\omega = \pm\pi/2$, i.e. less and less sharply peaked around their eccentricity maxima, so the fraction of a secular period they spend in the vicinity of e_{\max} increases.

Next we obtain the solution for $\omega(t)$. First, using the conservation of H^* (equation (4.2)) and assumptions (I) and (IV) we easily get an expression for $\cos^2 \omega(j)$ without stipulating any particular form of $j(t)$:

$$\cos^2 \omega = \frac{\Theta}{j^2 - \Theta} \left[\frac{j^2}{j_{\min}^2} - 1 + \frac{\epsilon_{\text{GR}} j_{\min}}{30\Gamma\Theta} \left(\frac{j^2}{j_{\min}^2} - \frac{j}{j_{\min}} \right) \right]. \quad (4.80)$$

Now plugging in the particular form (4.78) for $j(t)$ we find the following explicit solution for²⁰ $\omega(t)$:

$$\omega(t) = \frac{\pi}{2} + \text{sgn}(t) \cos^{-1} \left(\sqrt{\frac{1 + (t/t'_{\min})^2 - \epsilon_{\text{GR}} P(t)}{1 + (t/t'_{\min})^2 \chi}} \right), \quad (4.81)$$

where

$$\chi \equiv \frac{j_{\min}^2}{j_{\min}^2 - \Theta} = \left(1 - \frac{\Theta}{j_{\min}^2} \right)^{-1} = \frac{1}{\sin^2 i_{\min}}, \quad (4.82)$$

and

$$P(t) \equiv \frac{\chi}{30\Gamma j_{\min}} \left[1 + (t/t'_{\min})^2 - \sqrt{1 + (t/t'_{\min})^2} \right]. \quad (4.83)$$

is a dimensionless function of time. In Figure 4.10a we show how χ varies as a function of $j_{\min}/\Theta^{1/2} \geq 1$. We see that $\chi > 1$ always and that typical values of χ are \sim a few.

Finally we can get the solution for $\Omega(t)$ by using assumption (I) in equation (4.11), plugging in the solutions (4.78) and (4.81) for $j(t)$ and $\omega(t)$ respectively, and integrating in time. The result is

$$\begin{aligned} \Omega(t) = \Omega(0) + \text{sgn}(j_z) \left\{ -\tan^{-1} \left(\sqrt{\chi} \frac{t}{t'_{\min}} \right) \right. \\ + \frac{\epsilon_{\text{GR}}}{30\Gamma} \frac{\chi}{j_{\min}} \left[\tan^{-1} \left(\sqrt{\chi} \frac{t}{t'_{\min}} \right) \right. \\ \left. \left. - \sqrt{\frac{\chi}{\chi - 1}} \tan^{-1} \left(\sqrt{\chi - 1} \frac{t/t'_{\min}}{\sqrt{1 + (t/t'_{\min})^2}} \right) \right] \right\}, \quad (4.84) \end{aligned}$$

¹⁹To see this we take the explicit expressions for $j_1 = j_+$ and $j_2 = j_0$ from (4.13)-(4.16) and simplify them using assumption (I). Plugging the simplified expressions into (4.35) and expanding the result using assumption (IV) we recover equation (4.79).

²⁰We have included the $\text{sgn}(t)$ factor in (4.89) because for $\Gamma > 0$ the pericentre angle ω must increase towards $\pi/2$ as j decreases to j_{\min} (at $t = 0$), and continue to increase as j increases away from j_{\min} .

where we introduced $j_z \equiv J_z/L = j \cos i$.

In equation (4.84) the value of $\Omega(0)$ is an arbitrary constant to be prescribed. Otherwise, equations (4.78)-(4.79), (4.81)-(4.84), and the equation $j_z(t) = j_z(0)$ provide a complete, explicit description of the DA dynamics in the high- e limit whenever assumptions (I)-(IV) are satisfied.

We note that ω, Ω make finite ‘swings’ across the maximum eccentricity peak. Indeed, equation (4.81) tells us that ω takes asymptotic values

$$\omega(t \rightarrow \pm\infty) = \frac{\pi}{2} \pm \cos^{-1} \left(\sqrt{\frac{1}{\chi} - \frac{\epsilon_{\text{GR}}}{30\Gamma j_{\text{min}}}} \right), \quad (4.85)$$

giving a total swing of magnitude $|\Delta\omega| = 2 \cos^{-1} \sqrt{[\chi^{-1} - \epsilon_{\text{GR}}/(30\Gamma j_{\text{min}})]}$. Clearly the larger ϵ_{GR} , the bigger is this swing²¹, which makes sense since GR promotes fast apsidal precession. Similarly from (4.84) we find:

$$\Omega(t \rightarrow \pm\infty) = \Omega(0) \mp \text{sgn}(j_z) \left\{ \frac{\pi}{2} - \frac{\epsilon_{\text{GR}}\chi}{30\Gamma j_{\text{min}}} \left[\frac{\pi}{2} - \sqrt{\frac{\chi}{\chi-1}} \tan^{-1}(\sqrt{\chi-1}) \right] \right\}. \quad (4.86)$$

Thus, the size of the swing in Ω across the eccentricity peak $|\Delta\Omega| = \pi - \mathcal{O}(\epsilon_{\text{GR}})$ is *reduced* by GR effects.

4.C.1 Analytic solution in the LK limit

To apply the analytic solution to the LK case of hierarchical triples, let the tertiary perturber have mass \mathcal{M} and the outer orbit have semimajor axis a_g and eccentricity e_g . We may then use the results of §2.B. In particular, the value of A in this case is (Chapter 2, Appendix B):

$$A_{\text{LK}} = \frac{GM}{2a_g^3(1-e_g^2)^{3/2}}, \quad (4.87)$$

and obviously $\Gamma = 1$. We then evaluate ϵ_{GR} using equations (4.5) and take t'_{min} equal to

$$t'_{\text{min,LK}} \equiv \frac{4a_g^3(1-e_g^2)^{3/2}\sqrt{m_1+m_2}}{15G^{3/2}\mathcal{M}a^{3/2}} \frac{\chi^{3/2}}{\sqrt{\chi-1}}. \quad (4.88)$$

4.C.2 Simplified analytic solution in the limit $\epsilon_{\text{GR}} = 0$

One can get a simplified version of the analytic solution if one takes the non-GR limit. For $\epsilon_{\text{GR}} = 0$ the solution for $j(t)$ takes the same form (4.78), while (4.81) and (4.84) simplify to

$$\omega(t) = \frac{\pi}{2} + \text{sgn}(t) \cos^{-1} \left(\sqrt{\frac{1 + (t/t'_{\text{min}})^2}{1 + (t/t'_{\text{min}})^2\chi}} \right), \quad (4.89)$$

$$\Omega(t) = \Omega(0) - \text{sgn}(j_z) \tan^{-1} \left(\frac{\sqrt{\chi}t}{t'_{\text{min}}} \right). \quad (4.90)$$

²¹Of course this value becomes ill-defined when $\epsilon_{\text{GR}} > 30\Gamma j_{\text{min}}/\chi \sim \epsilon_{\text{weak}}/\chi$. For typical values of $\chi \sim 1$ this is never an issue in the weak GR regime $\epsilon_{\text{GR}} \ll \epsilon_{\text{weak}}$.

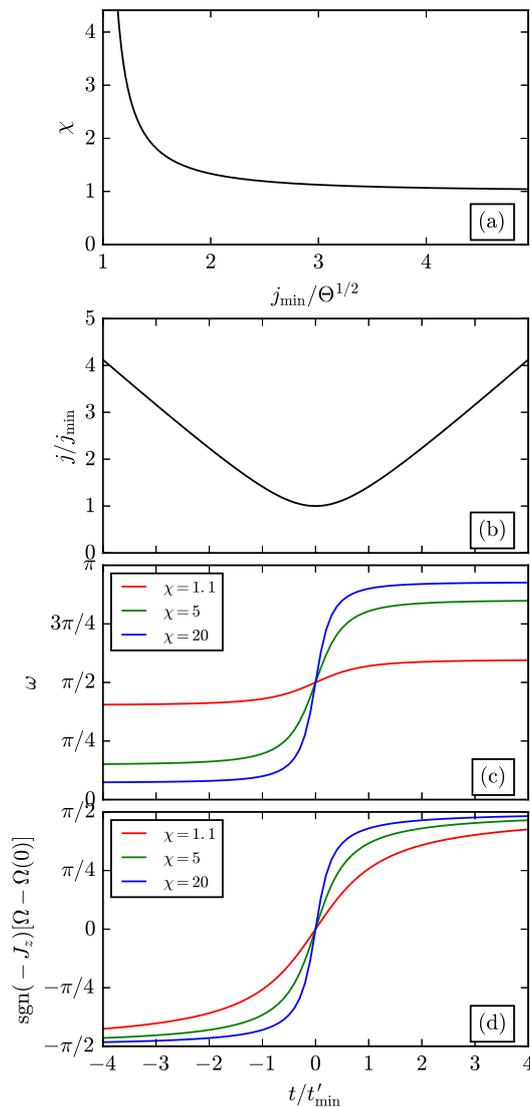


Figure 4.10: Analytic solution to the DA equations of motion without GR at high eccentricity, for binaries that achieve maximum eccentricity at $\omega = \pm\pi/2$. The solution depends on the parameter χ (equation (4.82)) which is plotted as a function of $j_{\min}/\Theta^{1/2}$ in panel (a). Panels (b)-(d) show the evolution of $j(t)$, $\omega(t)$ and $\Omega(t)$ respectively, where we have taken the maximum eccentricity to coincide with $t = 0$.

In Figure 4.10 we show the characteristic behaviour of this non-GR solution. In panel (b) we plot j/j_{\min} as a function of t/t'_{\min} : obviously $j(t)$ is quadratic in t for $t \lesssim t'_{\min}$ and linear for $t \gtrsim t'_{\min}$. Panels (c) and (d) demonstrate how the solutions for $\omega(t)$ and $\Omega(t)$ look for various χ . We note that both angles evolve very rapidly during the interval $-1 \lesssim t/t'_{\min} \lesssim 1$ and rather slowly otherwise, particularly for $\chi \gg 1$. We see also that the behaviour of ω is quite strongly dependent on χ ; it completes a swing $|\Delta\omega| = 2\cos^{-1}(\chi^{-1/2})$ as t runs from $-\infty$ to $+\infty$. The evolution of Ω depends somewhat less strongly on χ , and its asymptotic value is independent of χ , so that that the total swing in Ω across the eccentricity peak is always $|\Delta\Omega| = \pi$. Of course these swings in

ω and Ω are not completely correct because we expect our analytic formulae to break down once e differs significantly from unity (§4.C.3).

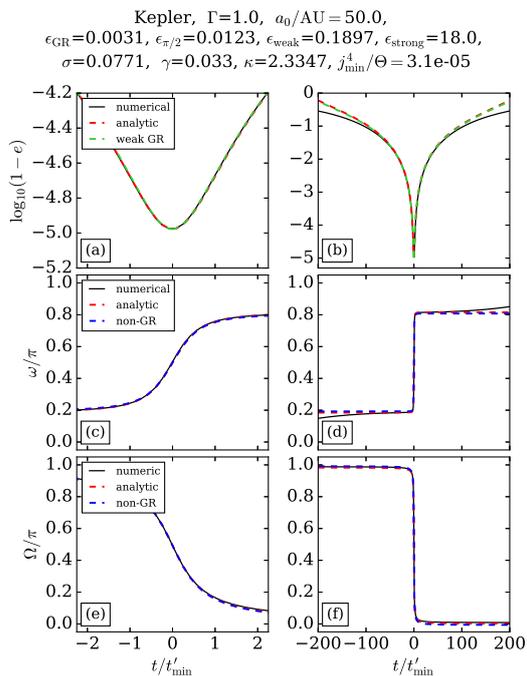


Figure 4.11: Comparing analytic results at high eccentricity with exact numerical integration (see §4.C.3 for details). In this example $\epsilon_{\text{GR}} < \epsilon_{\pi/2}$ so the binary is in the very weak GR regime. Note also that $\sigma \ll 1$.

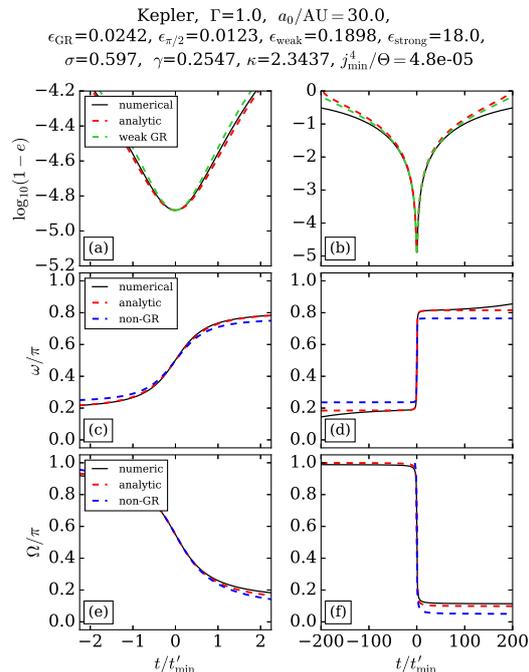


Figure 4.12: As in Figure 4.11 except we take $a_0 = 30\text{AU}$, so the binary is in the weak GR regime, $\epsilon_{\pi/2} < \epsilon_{\text{GR}} < \epsilon_{\text{weak}}$, and the value of σ is now approaching unity.

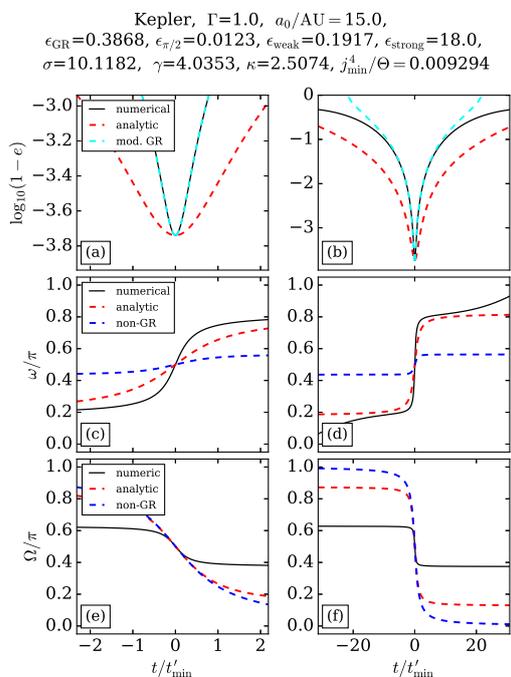


Figure 4.13: As in Figure 4.11 except we take $a_0 = 15\text{AU}$, so the binary is in the moderate GR regime, $\epsilon_{\text{weak}} < \epsilon_{\text{GR}} < \epsilon_{\text{strong}}$. Note that σ is significantly larger than unity in this case.

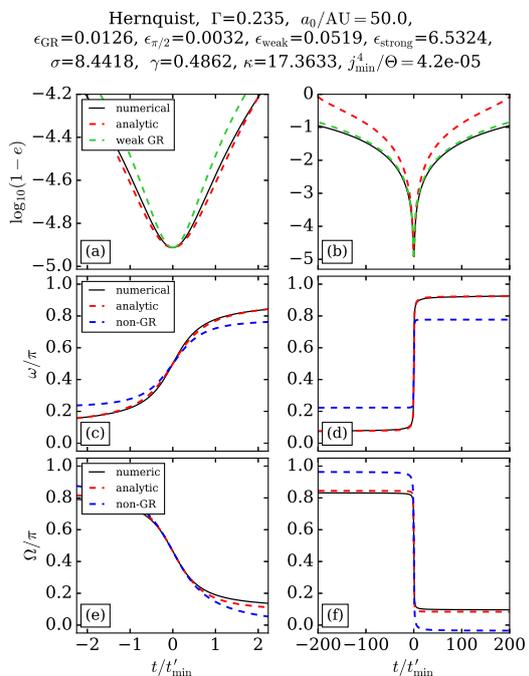


Figure 4.14: As in Figure 4.11 except we replace the Kepler potential with the Hernquist potential with scale radius 1pc . This results in $\Gamma = 0.235$ so $\sigma \gg 1$ even though the binary is in the weak GR regime.

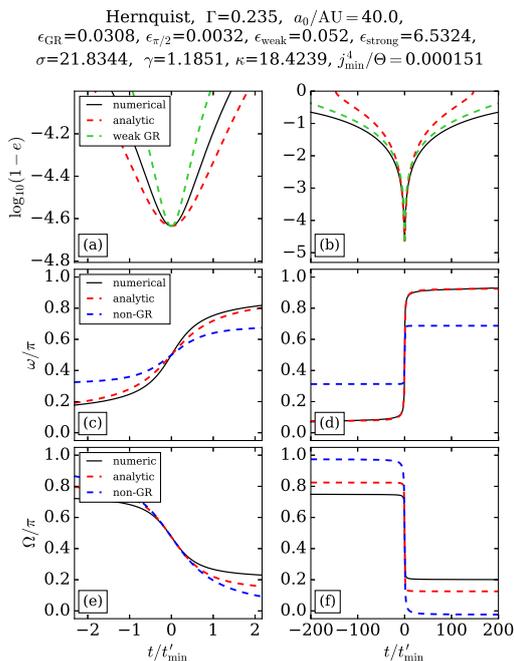


Figure 4.15: As in Figure 4.14 except we take $a_0 = 40\text{AU}$. Again the binary is in the weak GR regime, $\epsilon_{\pi/2} < \epsilon_{\text{GR}} < \epsilon_{\text{weak}}$, but is approaching the moderate GR regime.

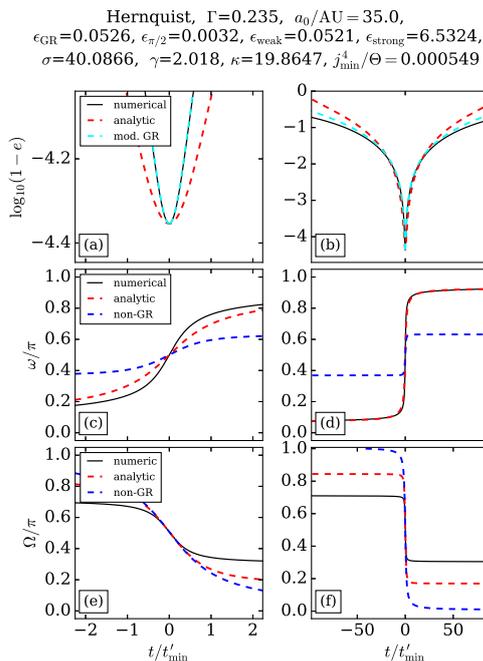


Figure 4.16: As in Figure 4.14 except we take $a_0 = 35\text{AU}$. In this case the binary is (just) in the moderate GR regime $\epsilon_{\text{weak}} < \epsilon_{\text{GR}} < \epsilon_{\text{strong}}$.

4.C.3 Validity of the analytic solution

In this section we test the accuracy of the analytical solution (4.78), (4.81), (4.84) and the simplified solution (4.89), (4.90) derived in the non-GR limit, against direct numerical integration of the DA equations of motion (4.9), (4.10), (4.11), in different dynamical regimes.

Three examples with $\Gamma = 1$

First we consider some examples in the LK case of $\Gamma = 1$. Precisely, we consider a binary with component masses $m_1 = m_2 = 1M_{\odot}$ orbiting a point mass $\mathcal{M} = 10^5 M_{\odot}$. For the outer orbit we choose a pericentre distance $r_p = 0.4\text{pc}$ and an apocentre distance $r_a = 0.6\text{pc}$. The outer orbit is then an ellipse with semimajor axis $a_g = 0.5\text{pc}$ and eccentricity $e_g = 0.2$. For the inner binary orbit we take the initial conditions $e_0 = 0.5$, $i_0 = 89.75^\circ$ (so that $\Theta = 1.4 \times 10^{-5}$), $\omega_0 = 0^\circ$. When we integrate the equations of motion we will shift the time coordinate so that maximum eccentricity is achieved at $t = 0$; in each example we choose a value of Ω_0 so that $\Omega(0) \approx \pi/2$. All that remains is to specify the initial semimajor axis a_0 .

In Figure 4.11 we take $a_0 = 50\text{AU}$. Then in each panel we plot the result of the direct numerical integration with a black line and we show the analytic solution (4.78), (4.81), (4.84) with a dashed red line. Panel (a) shows the evolution of $\log_{10}(1 - e)$ as

a function of time t over a short time interval $-2.2 \leq t/t'_{\min} \leq 2.2$ centred on the eccentricity peak. Panel (b) shows the same solution zoomed out over a much longer time interval $-200 \leq t/t'_{\min} \leq 200$. Analogously, panels (c) and (d) show the numerical and analytical solutions for the apsidal angle $\omega(t)$ over these same time intervals, while panels (e) and (f) show the evolution of the nodal angle $\Omega(t)$. Finally, in panels (c)-(f) we plot blue dashed lines which correspond to the simple non-GR form of the analytic solution, namely equations (4.89), (4.90), though to evaluate it we still use the GR-modified value of j_{\min} . (There are also green dashed lines in panels (a), (b) — see §4.4.4). At the top of the figure we show the values of various key quantities that allow us to check the validity of the assumptions (I)-(IV).

Overall, in Figure 4.11 the analytic solution provides an excellent fit to the exact numerical integration. Errors are only noticeable once e falls below ~ 0.9 (panels (b) and (d)). This good agreement reflects the fact that $\sigma, j_{\min}^4/\Theta \ll 1$ and $\epsilon_{\text{GR}} \ll \epsilon_{\text{weak}}$, meaning that all assumptions (I)-(IV) are fulfilled. Moreover, we see that the full analytic solution (red dashed lines) and non-GR solution (blue dashed lines) overlap almost exactly in panels (c)-(f). This is unsurprising because the binary actually sits in the *very weak* GR regime $\epsilon_{\text{GR}} < \epsilon_{\pi/2}$, meaning GR effects are negligible (§4.5.1).

In Figure 4.12 we use all the same system parameters as in Figure 4.11 except we set $a_0 = 30\text{AU}$. This increases ϵ_{GR} and puts the binary in the *weak* GR regime $\epsilon_{\pi/2} < \epsilon_{\text{GR}} < \epsilon_{\text{weak}}$. The fact that ϵ_{GR} is no longer smaller than $\epsilon_{\pi/2}$ is responsible for the disagreement between the analytic and non-GR solutions in panels (c)-(f). Nevertheless, the analytic solution still matches the numerical one very well for $e \gtrsim 0.9$, although not quite as well as in Figure 4.11, owing to the fact that σ is now comparable to unity (breaking assumption (III)).

Next, in Figure 4.13 we again run the same experiment but this time with $a_0 = 15\text{AU}$. This puts the binary in the *moderate* GR regime, $\epsilon_{\text{weak}} < \epsilon_{\text{GR}} < \epsilon_{\text{strong}}$, which violates assumption (II). Additionally we have $\sigma \gg 1$, violating assumption (III). We see solutions (4.78), (4.81), (4.84) largely fail to capture the high-eccentricity behaviour even over a very short timescale. At the same time, we note that the moderate GR solution (4.55) captures the $j(t)$ behaviour extremely well in this case.

Three examples with $\Gamma = 0.235$

Next we consider some examples with a different value of Γ . To achieve this we replace the Kepler potential with a Hernquist potential $\Phi(r) = -GM(b+r)^{-1}$, where the total mass $\mathcal{M} = 10^5 M_{\odot}$ and the scale radius $b = 1\text{pc}$. (The outer orbit still has $r_p = 0.4\text{pc}$ and $r_a = 0.6\text{pc}$, but will now fill a 2D annulus rather than forming a closed ellipse —

see Chapter 2). As a result we find $\Gamma = 0.235$. Also in this case both σ and κ attain large values, putting our analytical solutions to a demanding test.

In Figure 4.14 we integrate exactly the same system as in Figure 4.11 except for this replacement of the potential — in particular, we again take $a_0 = 50\text{AU}$. We see that this puts the binary in the weak GR regime $\epsilon_{\pi/2} < \epsilon_{\text{GR}} < \epsilon_{\text{weak}}$ (as in Figure 4.12), but that σ is much larger than unity (unlike in Figure 4.12). One consequence of this is that the analytic approximation to $\log_{10}(1 - e)$ fails rather early on, with significant errors by the time e falls below 0.99 (Figure 4.14b). Despite this, the analytic approximations to $\omega(t)$ and $\Omega(t)$ are still excellent (panels (c)-(f)). This is because ω and Ω are sensitive only to the eccentricity behaviour at the very peak — they change very rapidly over the interval $-2.2 < t/t'_{\text{min}} < 2.2$ (panels (c) and (e)), but are almost constant the rest of the time. Thus, as long as $j(t)$ is captured well near the very peak eccentricity, as it is in panel (a), the analytic solutions for ω , Ω work well despite assumption (II) being broken.

In Figure 4.15 we investigate the same system except with a_0 reduced to 40AU. The binary is still in the weak GR regime but only just so, violating assumption (II). It also has $\sigma \gg 1$ like it did in Figure 4.14, violating assumption (III). We see that the analytic fit to $\log_{10}(1 - e)$ is quite poor even at the very peak (panel (a)). Interestingly though, the evolution of ω (panel (d)) is reproduced rather accurately, highlighting how sensitive $\omega(t)$ is to the value of peak eccentricity j_{min} (see equation (4.85)), and how insensitive it is to anything else. However, the evolution of $\Omega(t)$ is not reproduced very well. The same conclusions hold for Figure 4.16, in which we have reduced the semimajor axis further to $a_0 = 35\text{AU}$, putting the binary squarely in the moderate GR regime (so that both assumptions (II) and (III) are broken).

Conclusions

While assumptions (I) and (IV) are almost always good provided we consider binaries that reach very high eccentricity ($1 - e \ll 0.1$), assumptions (II) and (III) are liable to fail in some regimes.

We have seen that for $\log_{10}(1 - e)$ to be accurately reproduced by the analytical solution (4.78) for $e \gtrsim 0.9$, all four assumptions (I)-(IV) must be valid.

However, the analytic solution (4.84) for Ω can be very accurate even for $\sigma \gg 1$ (violating assumption (III)) provided the behaviour of $j(t)$ in the close vicinity of j_{min} is reproduced reasonably well.

What is more, the solution (4.81) for $\omega(t)$, and the swing $\Delta\omega$ in particular, can be very accurate even if the system is in the moderate GR regime, invalidating both assumptions (III) and (IV). This is because ω is extremely sensitive to the behaviour of j around absolute peak eccentricity and largely insensitive to j otherwise.

Lastly, if all assumptions (I)-(IV) are valid and we additionally have $\epsilon_{\text{GR}} \lesssim \epsilon_{\pi/2}$, then one can employ the simpler non-GR form of the solution for ω, Ω (equations (4.89), (4.90)).

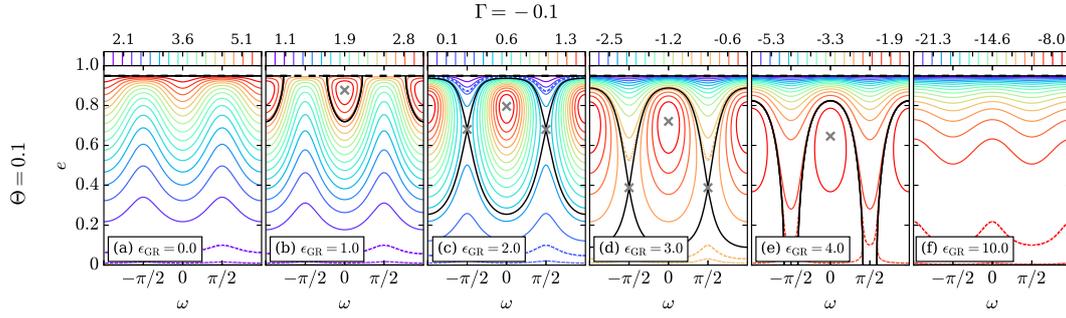


Figure 4.17: As in the top row of Figure 4.2 except we take $\Gamma = -0.1$ and use some new ϵ_{GR} values.

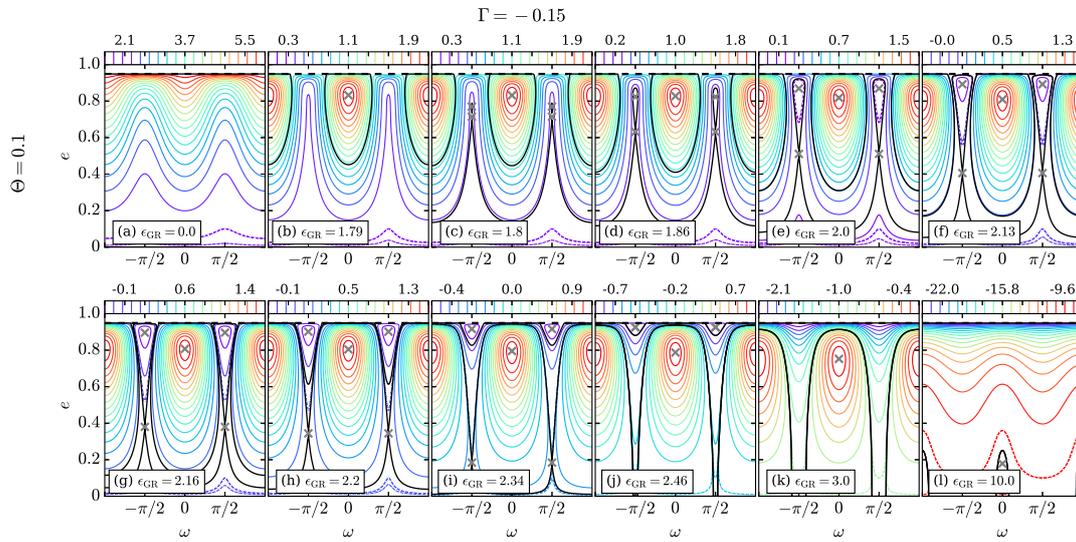


Figure 4.18: As in Figure 4.17 except for $\Gamma = -0.15$, and for twelve values of ϵ_{GR} .

4.D Phase space behaviour and maximum eccentricity in $\Gamma \leq 0$ regimes

In this Appendix we discuss the dynamical behaviour that arises in negative Γ regimes. This behaviour can be significantly more complicated than for positive Γ . In what follows we offer an overview of the phase space dynamics for $-1/5 < \Gamma \leq 0$ (in §4.D.1) and

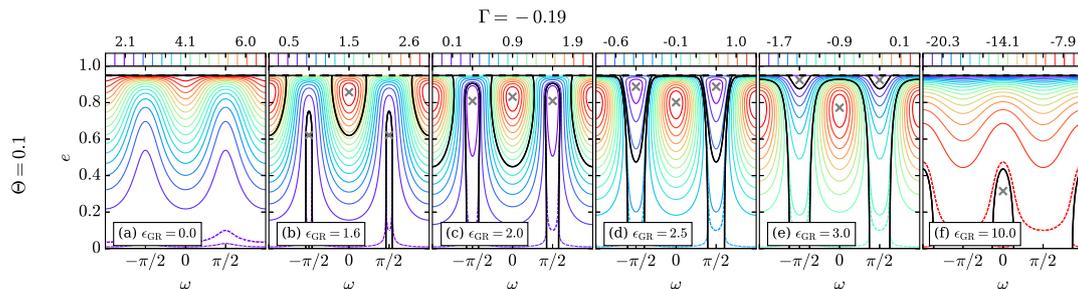


Figure 4.19: As in Figure 4.17 except for $\Gamma = -0.19$ and some different ϵ_{GR} values.

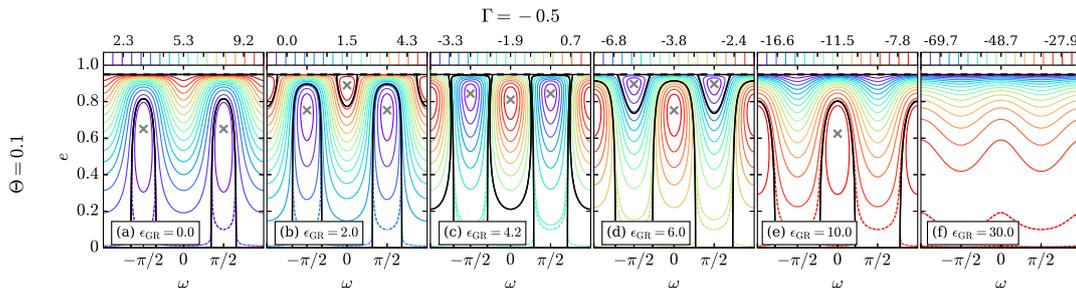


Figure 4.20: As in Figure 4.17 except for $\Gamma = -0.5$, and some new ϵ_{GR} values.

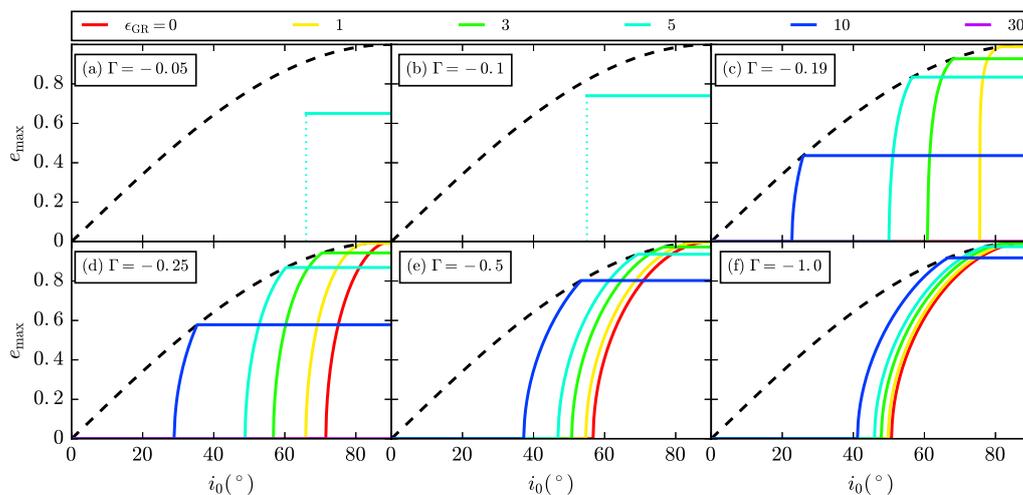


Figure 4.21: As in Figure 4.5 except for $-1/5 < \Gamma \leq 0$ (panels (a)-(c)) and $\Gamma \leq -1/5$ (panels (d)-(f)). The vertical dotted lines in panels (a) and (b) show the critical inclination $i_0 = \cos^{-1} \sqrt{\Theta_1(\Gamma, \epsilon_{\text{GR}})}$ for $\epsilon_{\text{GR}} = 5$ — see §4.D.3.

$\Gamma \leq -1/5$ (in §4.D.2). Lastly we consider the eccentricity maxima of binaries with negative Γ (§4.D.3), focusing mainly on initially near-circular orbits.

4.D.1 Phase space behaviour in the case $-1/5 < \Gamma \leq 0$

Unlike for $\Gamma > 0$, the dynamical behaviour in the regime $-1/5 < \Gamma \leq 0$ cannot be understood using only one value of Γ as an example. Thus, we consider three values. In Figures 4.17, 4.18 and 4.19 we plot contours of constant H^* in the (ω, e) phase space for $\Theta = 0.1$, taking $\Gamma = -0.1$, $\Gamma = -0.15$ and $\Gamma = -0.19$ respectively. The manually-added dashed contours are the same as in Figure 4.2. We now discuss these three figures in turn.

First we discuss Figure 4.17 ($\Gamma = -0.1$). From Chapter 3 we know that when $\epsilon_{\text{GR}} = 0$, fixed points never exist in the phase space for $-1/5 < \Gamma \leq 0$. Thus all phase space trajectories circulate and their maximum eccentricity is found at $\omega = \pm\pi/2$, as in panel (a). Now we consider finite ϵ_{GR} . In panel (b), namely for $\epsilon_{\text{GR}} = 1.0$, we see that fixed points have appeared at $\omega = 0, \pm\pi$, which we will refer to simply as $\omega = 0$ from now on. These

fixed points are not saddle points like they were for $0 < \Gamma \leq 1/5$ (Figure 4.2); instead they are maxima of H^* and host a region of librating orbits that is connected to e_{lim} .

As we increase ϵ_{GR} further we see that these fixed points move down the page to lower eccentricity, and their associated librating islands become larger in area. At some threshold value of ϵ_{GR} the librating islands become disconnected from the line $e = e_{\text{lim}}$, coinciding with the appearance of new saddle points at $\omega = \pm\pi/2, e = e_{\text{lim}}$. As we increase ϵ_{GR} beyond this threshold the fixed point at $\omega = 0$ continues to move down the page (panels (c) and (d)), as do the saddle points at $\omega = \pm\pi/2$, and a new family of high- e circulating orbits runs over the top of the librating islands, reminiscent of what we saw for $0 < \Gamma \leq 1/5$ in Figure 4.2. Partitioning the different librating islands and circulating regions in panels (c) and (d) are separatrices that cross at the saddle points. Continuing to increase ϵ_{GR} forces both kinds of fixed point to move to lower eccentricities. The saddle points move fastest and disappear first; in panel (e), the fixed point at $\omega = 0$ remains but the saddle points at $\omega = \pm\pi/2$ have disappeared through $e = 0$. Increasing ϵ_{GR} even further still, the $\omega = 0$ fixed point reaches $e = 0$ and then disappears. This leaves a phase space filled with circulating trajectories (panel (f)), which is similar to the $\epsilon_{\text{GR}} = 0$ case shown in panel (a) except that the maximum eccentricities are now found at $\omega = 0$ rather than $\omega = \pm\pi/2$, and the locations of the maxima and minima of H^* are reversed (see the colourbars).

Moving on to Figure 4.18 ($\Gamma = -0.15$), we find a completely different picture of rather impressive dynamical diversity. In this figure we have to use twelve panels to fully illustrate the complex phase space behaviour. To begin with, panels (a) and (b) in Figure 4.18 have the same morphology as Figures 4.17a,b. However, panel (c) is very different from Figure 4.17c. This time, at some threshold value of ϵ_{GR} a *pair* of fixed points emerges from a single point at $\omega = \pi/2, e = e_{\text{f},\pi/2}$, and the same thing happens at $\omega = -\pi/2$. An increase in ϵ_{GR} nudges these fixed points apart in their eccentricity values (panel (d)): one of them moves up the page and the other moves down. In each pair, the fixed point with higher e is a minimum of H^* and hosts a region of librating orbits. The fixed point with lower e is a saddle point, and sits on the separatrix that surrounds the upper point's librating region. In addition we still have the usual fixed point and accompanying librating island at $\omega = 0$. As a result, we now find two families of circulating trajectories. One runs close to $e = 0$ under the separatrices passing through the saddle points. The other runs above these separatrices, but below the separatrices surrounding the librating islands centred on $\omega = 0, \pm\pi$. This second type of circulating trajectory reaches high eccentricity by running above the upper fixed points at $\omega = \pm\pi/2$. Quite remarkably, these circulating trajectories also exhibit *non-monotonic* behaviour of $\omega(t)$, i.e. $\dot{\omega}$ is > 0 at some times and < 0 at others, despite the trajectory being a circulating one.

Increasing ϵ_{GR} further, the upper fixed point (minimum) at $\omega = \pm\pi/2$ continues to move up the page, while the lower fixed point (saddle) moves down (panels (e)-(g)). Meanwhile the $\omega = 0$ fixed points also move down the page, albeit much more slowly. Eventually the librating region surrounding the upper fixed point at $\omega = \pm\pi/2$ becomes connected to $e = e_{\text{lim}}$. Simultaneously, the saddle point at $\omega = \pm\pi/2$ and its associated separatrices merge with the separatrices surrounding the $\omega = 0$ librating regions (see the transition from panel (g) to panel (h)). Accompanying this transition is the change in the nature of the second family of finite eccentricity circulating orbits described above — they now run above (below) the saddle points at $\omega = 0, \pm\pi$ ($\omega = \pm\pi/2$). As ϵ_{GR} continues to increase the pair of fixed points at $\omega = \pm\pi/2$ continue to move apart in eccentricity, until eventually the lower one disappears at $e = 0$ (panel (j)) followed by the upper one at $e = e_{\text{lim}}$ (panel (k)). In panels (k) and (i) we retain only the fixed points at $\omega = 0$, with qualitatively the same overall phase space behaviour as in Figure 4.17e. The $\omega = 0$ fixed points also disappear once ϵ_{GR} becomes sufficiently large.

Figure 4.19 ($\Gamma = -0.19$) shows yet again a different qualitative behaviour. Like in Figures 4.17, 4.18, fixed points emerge at $\omega = 0$ followed by additional fixed points at $\omega = \pm\pi/2$, $e = e_{f,\pi/2}$ (panel(b)). However, this time the $\omega = \pm\pi/2$ fixed points do not come in pairs like they did in Figure 4.18. Instead they are minima of H^* and are surrounded by a librating island that stretches to $e = 0$ (though $e_{f,\pi/2} \neq 0$ for any ϵ_{GR}). These fixed points move up the page as we increase ϵ_{GR} (panel(c)) until they become connected to $e = e_{\text{lim}}$ (panel (d)). At this stage circulating trajectories exhibit a transition similar to that in Figure 4.18. Thereafter we have qualitatively the same behaviour as in Figure 4.18j.

As these three examples demonstrate, the qualitative dynamical behaviour in the regime $-1/5 < \Gamma \leq 0$ is highly complex. It is also very difficult to analyse mathematically. The simplest place to start is with the fixed points at $\omega = 0$, $j = j_{f,0}$. The formulae describing these fixed points can be carried over from §4.3.2 and Appendix 4.A: the value of $j_{f,0}$ is still determined by equation (4.27) and the fixed points exist provided equation (4.28) is true. The key difference for negative Γ compared to positive Γ is that the determinant of the Hessian matrix of H^* evaluated at the fixed points, namely the expression (4.68), is now manifestly positive rather than negative. Thus the fixed points at $\omega = 0$, $j = j_{f,0}$ are now true extrema (more precisely, maxima) of H^* and host a librating island, which is reflected in Figures 4.17-4.19.

Understanding the fixed points at $\omega = \pm\pi/2$, $j = j_{f,\pi/2}$ is much harder. Just like for $\Gamma > 0$, to find $j_{f,\pi/2}$ we must solve the depressed quartic (4.63). From this equation we can once again derive a necessary but insufficient condition for fixed points to exist

at $\omega = \pm\pi/2$; however since $10\Gamma/(1+5\Gamma) < 0$ in this Γ regime, rather than the upper bound $\epsilon_{\text{GR}} < 6(1+5\Gamma)$ that we found for $\Gamma > 0$ (§4.3.1) we instead get a lower bound, $\epsilon_{\text{GR}} > 6(1+5\Gamma)\Theta^{3/2}$. Unfortunately it is not easy to write down analogues of the sufficient conditions (4.22)-(4.23)²². Indeed, as we saw in Figure 4.18, for $-1/5 < \Gamma \leq 0$ fixed points can arise in pairs at $\omega = \pi/2$ (with another, separate pair at $\omega = -\pi/2$), corresponding to there being two physical solutions to the quartic (4.63).

Finally, even the nature of the $\omega = \pm\pi/2$ fixed points is a non-trivial issue. The determinant of the Hessian of $H^*(\omega, j)$ evaluated at $(\pm\pi/2, j_{\text{f},\pi/2})$ is given by

$$[3(1+5\Gamma)j_{\text{f},\pi/2}^4 + 10\Gamma\Theta] \frac{360\Gamma(j_{\text{f},\pi/2}^2 - \Theta)(1 - j_{\text{f},\pi/2}^2)}{j_{\text{f},\pi/2}^6}, \quad (4.91)$$

where we eliminated ϵ_{GR} using equation (4.63). For negative Γ , the sign of (4.91) depends on the sign of the first bracket. If $[3(1+5\Gamma)j_{\text{f},\pi/2}^4 + 10\Gamma\Theta] > 0$ then the fixed point at $\omega = \pm\pi/2$ is a saddle point; otherwise it is a true extremum (in fact a minimum). This puts an implicit constraint on ϵ_{GR} (since $j_{\text{f},\pi/2}$ depends on ϵ_{GR}) when determining the nature of the fixed points. That constraint is responsible for the fact that even for a fixed $\Theta = 0.1$, the $\omega = \pm\pi/2$ fixed points are saddle points in Figure 4.17, minima in Figure 4.19, and both are present in Figure 4.18.

4.D.2 Phase space behaviour in the case $\Gamma \leq -1/5$

We now turn to the final Γ regime, $\Gamma \leq -1/5$, which luckily is not as complicated as $0 < \Gamma \leq 1/5$. We need only illustrate it with a single example, namely Figure 4.20, which is for $\Theta = 0.1$ and $\Gamma = -0.5$.

For $\epsilon_{\text{GR}} = 0$ (panel (a)) the phase portrait looks almost identical to those typical of $\Gamma > 1/5$ (e.g. Figure 4.1a). However, as we noted in Chapter 3, despite their similarities the dynamical regimes $\Gamma > 1/5$ and $\Gamma \leq -1/5$ are significantly different. In particular, the phase space trajectories in each regime are traversed in opposite directions (to see this, compare the arrows in Figure 3.4a to those in Figure 3.7d). One consequence of this is that for $\Gamma \leq -1/5$, increasing ϵ_{GR} always pushes the fixed points at $\omega = \pi/2$ *up* the page to higher eccentricity — see panels (b) and (c) of Figure 4.20. This behaviour is easy to reconstruct mathematically. Since $10\Gamma/(1+5\Gamma) > 0$ in this Γ regime, equation (4.63) tells us that for fixed points at $\omega = \pm\pi/2$ to exist necessarily requires $j_{\text{f},\pi/2}^3 > \epsilon_{\text{GR}}/[6(1+5\Gamma)]$. Then it is easy to show (c.f. equations (4.65)-(4.66)) that

$$\left(\frac{\partial j_{\text{f},\pi/2}}{\partial \Theta}\right)_{\epsilon_{\text{GR}}} > 0, \quad \text{and} \quad \left(\frac{\partial j_{\text{f},\pi/2}}{\partial \epsilon_{\text{GR}}}\right)_{\Theta} < 0. \quad (4.92)$$

²²The difficulty arises because the signs of $\partial j_{\text{f},\pi/2}/\partial \Theta$ and $\partial j_{\text{f},\pi/2}/\partial \epsilon_{\text{GR}}$ (expressions for which are given in (4.65)-(4.66)) are not fixed in this Γ regime, so we cannot look for e.g. the bounding values of ϵ_{GR} that give $j = \sqrt{\Theta}, 1$.

In other words, increasing Θ decreases the eccentricity of the fixed points at $\omega = \pm\pi/2$ should they exist (same as $\Gamma > 0$), but increasing ϵ_{GR} increases their eccentricity (opposite to $\Gamma > 0$). The condition on ϵ_{GR} for the existence of these fixed points is the same as (4.22) but reversing the inequalities, i.e. replacing each ‘<’ with ‘>’. The condition on Θ is the same as that given for $0 < \Gamma \leq 1/5$ in equation (4.23). Additionally, the fixed points at $\omega = \pm\pi/2$ are always true extrema (minima) of H^* in this Γ regime since the expression (4.91) is always positive.

Meanwhile, the fixed points at $\omega = 0$ follow exactly the same rules as for $-1/5 < \Gamma \leq 0$, appearing at $\omega = 0, e = e_{\text{lim}}$ when ϵ_{GR} reaches the critical value $\epsilon_{\text{GR}} = 6(1 - 5\Gamma)\Theta^{3/2}$ and then working their way down the page towards $e = 0$ as ϵ_{GR} is increased, disappearing for $\epsilon_{\text{GR}} > 6(1 - 5\Gamma)$ (equation (4.28)). The only difference is that these fixed points are maxima of H^* , not saddle points, which follows from the fact that the quantity (4.68) is positive for $\Gamma < 0$.

4.D.3 Orbit families and maximum eccentricity for $\Gamma \leq 0$ regimes

Owing to the highly complex phase space morphology, working out a trajectory’s orbital family analytically is often a very tedious job for negative Γ values. The same goes for finding a binary’s maximum eccentricity: in practice it is best simply to take a brute-force approach by solving the cubic and quartic equations (4.69), (4.30) numerically to get all seven possible roots, and then declaring j_{min} to be the real root closest to but smaller than the initial j value. We will not pursue any further technical details here.

Maximum eccentricity for initially near-circular binaries

With this brute-force approach it is straightforward to calculate e_{max} for a given i_0, Γ and ϵ_{GR} for initially near-circular binaries when $\Gamma \leq 0$ (c.f. §4.3.4). In Figure 4.21 we show $e_{\text{max}}(i_0)$ for various ϵ_{GR} values. In each panel we use a different negative value of Γ (c.f. Figure 4.5).

In the top row of Figure 4.21 (panels (a)-(c)) we explore the regime $-1/5 < \Gamma \leq 0$. To understand panels (a) and (b) it is worth looking back at Figures 4.17 and 4.18 and asking what we expect of the behaviour of initially near-circular orbits. We expect from Figures 4.17a,b,c,d and 4.18a-i that for low enough ϵ_{GR} the maximum eccentricity will be zero. This immediately explains, for instance, why there is no red line (corresponding to $\epsilon_{\text{GR}} = 0$) in panels (a) and (b) of Figure 4.21. However, when ϵ_{GR} takes a value such that (I) the fixed point exists at $\omega = 0$ and (II) the librating region that this fixed point hosts is connected to $e = 0$, then the eccentricity of initially circular orbits is maximised at $\omega = 0$ and is nonzero (Figure 4.17e and Figure 4.18j,k,l).

We know that (I) is true if and only if ϵ_{GR} satisfies (4.28). We also know that for (II) to be true the fixed points at $\omega = \pm\pi/2$ must have disappeared below $e = 0$. By examining equations (4.63) and (4.66) in the limit of $j_{f,\pi/2} \approx 1$ and small Θ , we find that (II) becomes true when ϵ_{GR} is increased beyond the threshold value $6(1 + 5\Gamma)$. Putting these constraints together and using the fact that $-1/5 < \Gamma \leq 0$, we find that in the limit $\Theta \rightarrow 0$ a necessary condition for both (I) and (II) to be true is $6(1 + 5\Gamma) < \epsilon_{\text{GR}} < 6(1 - 5\Gamma)$, which is the same as (4.29) if we replace ‘<’ with ‘>’. For $\Gamma = -0.05$ this gives $4.5 < \epsilon_{\text{GR}} < 7.5$, which is why there is only a cyan line in Figure 4.21a. Similarly, for $\Gamma = -0.1$ we get $3 < \epsilon_{\text{GR}} < 9$, hence the solo cyan line in Figure 4.21b.

This necessary constraint on ϵ_{GR} was derived for $\Theta \rightarrow 0$, i.e. $i_0 \rightarrow 90^\circ$. To find the necessary constraint on i_0 for (I) and (II) to be true, we need a constraint on Θ (equivalent to $\cos^2 i_0$ for $e_0 \approx 0$). By considering equations (4.28), (4.63) and (4.65) for $j_{f,\pi/2} \approx 1$ and Γ not too close to²³ $-1/5$, we can show that (I) and (II) are true provided $\Theta < \Theta_1(\Gamma, \epsilon_{\text{GR}})$. So initially near-circular binaries whose Γ values produce phase portraits like in Figures 4.17, 4.18 can achieve a finite e_{max} only if they have i_0 greater than the critical value $\cos^{-1} \sqrt{\Theta_1(\Gamma, \epsilon_{\text{GR}})}$. For panels (a) and (b) of Figure 4.21 these values are $i_0 = 66^\circ$ and $i_0 = 55^\circ$ respectively, which we show with vertical dotted lines. Plugging H^*, Θ from (4.31) into the depressed cubic (4.69), we find that the corresponding minimum j value is:

$$j_{\text{min}} = \frac{1}{2} \left(-1 + \sqrt{1 + \frac{4\epsilon_{\text{GR}}}{3(1 - 5\Gamma)}} \right), \quad (4.93)$$

which is independent of i_0 . In panels (a) and (b) of Figure 4.21, the straight horizontal cyan lines for $\epsilon_{\text{GR}} = 5$ correspond to the solution (4.93).

Panels (c)-(f) of Figure 4.21 all share a similar morphology, so we will consider them together. In each panel, for a fixed ϵ_{GR} a finite e_{max} arises at some critical value of i_0 , increases as a function of i_0 until it reaches $e_{\text{max}} = e_{\text{lim}}$, and then is constant for all larger values of i_0 up to 90° . Note that on the non-constant parts of these curves we have essentially the opposite of the intuitive $\Gamma > 1/5$ result: for a fixed initial inclination, a larger ϵ_{GR} leads to a *larger* e_{max} . The behaviour we see in these panels is consistent with what we expect from the $\Gamma = -0.19$ example given in Figure 4.19 and the $\Gamma \leq -1/5$ example we studied in Figure 4.20. In those figures, the fixed points that emerge at $\omega = \pm\pi/2, e = e_{f,\pi/2}$ host regions of librating orbits that are connected to $e = 0$. In each case, the maximum eccentricity of initially circular orbits is determined by the eccentricity of the separatrix at the point $\omega = \pm\pi/2$. As we increase ϵ_{GR} , the value of

²³Values of Γ close to $-1/5$ are more complicated, essentially because the sign of the right hand side of (4.65) is liable to change in this regime even for $j_{f,\pi/2} \approx 1$. This is the case in particular for $\Gamma = -0.19$, which is why the behaviour in Figure 4.21c is different from the other $-1/5 < \Gamma \leq 0$ examples in Figure 4.21a,b.

$e_{f,\pi/2}$ is increased, pushing the separatrix to higher e , and so the maximum eccentricity of initially circular orbits grows. Eventually, however, $e_{f,\pi/2}$ is increased so much that the separatrix reaches $e = e_{\text{lim}}$ (dashed black line) — see the transition between Figure 4.20c and 4.20d. At the same time, the librating islands that are hosted by fixed points at $\omega = 0$ become connected to $e = 0$. After that the maximum eccentricity is given by (4.93) and is independent of i_0 — hence the straight horizontal lines in Figure 4.21c-f. The main qualitative difference between panel (c) and panels (d)-(f) is that panel (d) exhibits no red line, i.e. no solution for $\epsilon_{\text{GR}} = 0$. This is because in the regime $-1/5 < \Gamma \leq 0$ a finite ϵ_{GR} is always required for any fixed points to exist (Figures 4.17-4.19).

Finally we may briefly compare Figure 4.21 with Figure 4.5. Consider what happens if we fix ϵ_{GR} and increase i_0 from zero. In Figure 4.21 ($\Gamma \leq 0$), the larger is ϵ_{GR} , the lower i_0 is required to achieve a non-zero e_{max} (provided ϵ_{GR} is not so large that no eccentricity excitation is possible). On the contrary, in Figure 4.5 ($\Gamma > 0$) the most favourable situation for eccentricity excitation is always to have ϵ_{GR} as small as possible: the larger ϵ_{GR} , the larger i_0 is required to get a non-zero maximum eccentricity. Though the two regimes differ in this respect, they are similar in that for binaries with $i_0 \approx 90^\circ$ a larger ϵ_{GR} always leads to a smaller maximum eccentricity (again provided ϵ_{GR} is such that eccentricity excitation is possible).

5

The effect of gravitational wave emission

Contents

| | | |
|-------------------|---|------------|
| 5.1 | Introduction | 168 |
| 5.1.1 | Motivation | 168 |
| 5.1.2 | Outlook for this Chapter | 169 |
| 5.2 | High eccentricity results without GW emission | 170 |
| 5.2.1 | Minimum and maximum j in the weak-to-moderate GR regime | 171 |
| 5.2.2 | The strong GR regime | 174 |
| 5.3 | Gravitational wave emission | 174 |
| 5.3.1 | Conservation laws | 175 |
| 5.3.2 | Evolution of a shrinking binary through phase space | 177 |
| 5.3.3 | The secular timescale, t_{sec} | 185 |
| 5.3.4 | The decay in semimajor axis over one secular cycle, Δa | 187 |
| 5.3.5 | The time evolution of semimajor axis, $a(t)$ | 190 |
| 5.4 | Numerical examples | 193 |
| 5.4.1 | Example 1: $\Gamma = 0.42 > 1/5$. An initially librating orbit in the weak GR regime | 193 |
| 5.4.2 | Example 2: $\Gamma = 0.42 > 1/5$. A binary initially in the moderate GR regime | 197 |
| 5.4.3 | Example 3: $\Gamma = 0.176 < 1/5$. | 201 |
| 5.5 | Discussion | 203 |
| 5.5.1 | Merger timescale | 204 |
| 5.5.2 | Relation to studies of LK-driven mergers | 206 |
| 5.6 | Summary | 207 |
| Appendices | | 209 |
| 5.A | Relation to Randall & Xianyu (2018) | 209 |
| 5.A.1 | Calculation of Δa | 209 |
| 5.A.2 | Decrease in t_{sec} with time | 211 |

5.1 Introduction

5.1.1 Motivation

By the midpoint of their third Observing Run in October 2020, the LIGO/Virgo Collaboration had detected a total of 50 compact object binary mergers (The LIGO Scientific Collaboration et al. 2019; The LIGO Scientific Collaboration et al. 2020). These 50 events imply a merger rate in the local Universe of $23.9^{+14.9}_{-8.6} \text{Gpc}^{-3} \text{yr}^{-1}$ for black hole-black hole (BH-BH) binaries and $320^{+490}_{-240} \text{Gpc}^{-3} \text{yr}^{-1}$ for neutron star-neutron star (NS-NS) binaries. However, despite these impressive experimental achievements, there is still ambiguity on the theoretical side about which mechanisms drive the mergers of compact object binaries. As discussed at length in Chapter 1, one much-studied candidate is the Lidov-Kozai (LK) mechanism (Lidov 1962; Kozai 1962; Naoz 2016), whereby a binary is torqued by a bound tertiary perturber. LK theory tells us that under the right circumstances, the binary’s eccentricity e can be driven periodically to large values. This greatly reduces the pericentre distance $p = a(1 - e)$, potentially leading to more rapid mergers (see e.g. Blaes et al. 2002; Wen 2003; Antonini & Perets 2012; Antognini et al. 2014; Silsbee & Tremaine 2017 and references therein).

However, this LK-driven merger channel obviously requires the existence of a rather special scenario, in which a compact object binary is accompanied by a third body on a much larger, though still bound, orbit. On the other hand, every compact object binary that resides in a stellar cluster feels that cluster’s potential. As we have seen in Chapters 2-4, the cluster potential provides a tidal torque on the binary just as a tertiary perturber would. One can therefore consider the cluster itself to be a ubiquitous ‘third body’, and explore its capacity to induce ‘cluster tide-driven’ mergers. Of course, to do this quantitatively means we must include GW emission in our equations of motion, which we have not done so far. In the present Chapter we will introduce GW emission into our calculations and attempt to understand the physics of cluster tide-driven compact object mergers.

At this stage the skeptical reader might ask: what is to be gained from a Chapter on cluster tide-driven mergers that is not already known from LK merger theory? Aren’t we going to just end up with the usual formulae, but with a few new factors of Γ floating around? Isn’t the physics essentially understood already? And indeed, the most basic tenets of LK-driven mergers are certainly well understood (Miller & Hamilton 2002; Blaes et al. 2002; Wen 2003; Thompson 2011; Liu & Lai 2017). GW emission is very strongly concentrated at the binary’s peak eccentricity $e \rightarrow e_{\text{max}} \approx 1$, and because of this one can use the value of e_{max} to characterise the amount of GW energy emitted over one secular cycle. These losses accumulate over multiple cycles until the binary is so small that it

‘decouples’ from the tertiary perturber and undergoes a GW-dominated inspiral following the Peters (1964) formulae. However, *very little detailed theory has been developed beyond this simple picture, even in the LK case*. Instead, most of the aforementioned studies opt either for direct numerical integration of the binary equations of motion, or they aim at a parameterisation of the total merger timescale as a function of¹ e_{\max} . On the other hand, LK-driven (and, by extension, cluster tide-driven) merger dynamics are in fact much richer than this overview would suggest, exhibiting non-trivial time evolution of the binary’s semimajor axis, maximum and minimum eccentricity, phase space location, etc. To take just one example, in LK literature one is frequently confronted with the counter-intuitive result that as a binary’s semimajor axis a shrinks, the timescale t_{sec} for the next secular eccentricity oscillation *decreases* (Randall & Xianyu 2018). This trend is rarely mentioned and until now has not been fully explained. The purpose of the present Chapter is to explain such dynamical characteristics.

5.1.2 Outlook for this Chapter

All binaries emit gravitational waves, so given enough time every binary will ultimately merge. We concentrate on those binaries whose merger timescale is significantly shortened by secular eccentricity excitation — the so-called ‘cluster tide-driven mergers’ (which includes LK-driven mergers as a special case). Following Randall & Xianyu (2018) we can separate these binaries further into ‘fast mergers’ and ‘slow mergers’. Fast mergers are those that occur after only one (or at most a few) secular eccentricity cycles. By contrast, slow mergers occur after many secular eccentricity cycles. Inevitably they involve a gradual transition of the binary from the weak-to-moderate GR regime to the strong GR regime. In fact, one can consider slow mergers to be a somewhat more general case, since one can imagine any fast merger merely as the final few secular cycles of a slow merger. Thus, in this Chapter we will focus on understanding the physics of slow mergers.

Let us therefore consider a binary with initial eccentricity not close to unity, and suppose that unless excited to high eccentricity it will not merge within a Hubble time. For the required eccentricity excitation to be possible, we know from Chapter 4 that the binary must begin its life in the weak-to-moderate GR regime. Supposing this is the case, and that the binary does indeed achieve high values of e periodically, then its semimajor axis a will be decreased by some amount Δa during each high-eccentricity episode because of GW emission (whereas away from $e \approx 1$, GW emission will be completely negligible, so

¹Or a version of e_{\max} that is modified by non-ideal effects, e.g. the breakdown of the test particle approximation (i.e. allowing the binary’s inner orbit to have comparable angular momentum to the tertiary’s outer orbit), the DA approximation (allowing for short-timescale fluctuations in the tidal torque) and the quadrupole approximation (including octupole or higher terms in the perturbing potential) — see Antonini, Murray, et al. 2014; Anderson, Lai, et al. 2017; Liu & Lai 2018; Grishin et al. 2018 and references therein.

we can treat a as constant). For a slow merger we know by assumption that each individual decrease is small $|\Delta a| \ll a$, though of course Δa will itself depend on the value of a . Next, the time between these high-eccentricity episodes is $t_{\text{sec}}(a)$. Therefore on timescales longer than a few secular periods we can approximate the slow decay of $a(t)$ using

$$\frac{da}{dt} \approx \frac{\Delta a(a)}{t_{\text{sec}}(a)} \implies t - t_i \approx \int_{a_i}^{a(t)} da' \frac{t_{\text{sec}}(a')}{\Delta a(a')}, \quad (5.1)$$

where $a_i \equiv a(t_i)$ and t_i is some earlier reference time. Equation (5.1) is an implicit equation for $a(t)$ in the weak-to-moderate GR regime. Clearly to perform the integral on the right hand side one needs to find expressions for $t_{\text{sec}}(a)$ and $\Delta a(a)$ — one of the primary goals of this Chapter will be to compute these two quantities. It will turn out that the qualitatively different behaviours exhibited by $\Delta a(a)$, $t_{\text{sec}}(a)$, and ultimately $a(t)$, will map onto different GR regimes and phase space features explored in Chapter 4. Eventually a becomes small enough that the binary reaches the strong GR regime and gets ‘trapped’ at high eccentricity. At this stage (5.1) breaks down and we must use a different prescription to follow $a(t)$ accurately all the way to merger.

The rest of this Chapter is organised as follows. In §5.2 we gather a few key results from Chapters 3-4 which are valid when GW emission is switched off. We do this for arbitrary $\Gamma > 0$, which covers the great majority of realistic cases, including LK theory. In §5.3 we introduce the effect of GW emission, and show how two new approximate conservation laws — that of the minimum pericentre distance and minimum inclination reached by the binary — allow us to write all important quantities (j_{min} , t_{sec} , Δa , etc.) as functions of semimajor axis a . This in turn allows us understand qualitatively the long-term decay of semimajor axis as a function of time, $a(t)$, in three key asymptotic regimes. In §5.4 we illustrate our results via numerical examples, in which we integrate directly the DA equations of motion including both GR precession and GW emission. In §5.5 we discuss our results, including their implications for LK-driven mergers. We summarise in §5.6. Finally, in Appendix 5.A we compare our calculations with those of Randall & Xianyu (2018).

5.2 High eccentricity results without GW emission

In this section we gather some results from Chapters 3-4 concerning cluster tide-driven secular dynamics *without* gravitational wave emission (but including GR precession). Though there is nothing strictly new here, it will be useful to have these results gathered in one place and written in a form that makes their meaning transparent. In particular, we focus on binaries that reach high eccentricity and write down how various important quantities depend on four dimensionless constants (Γ , Θ , ϵ_{GR} , j_{min}). This will form the basis for the new results we will derive from §5.3 onwards with GW emission included (though of course, at that stage Θ , ϵ_{GR} , j_{min} will no longer be constant).

We will assume throughout this Chapter that $\Gamma > 0$, and that the binary's maximum eccentricity e_{\max} is achieved at $\omega = \pm\pi/2$. As we have seen in previous Chapters these are not restrictive conditions, but they allow us to avoid discussing certain pathological cases. It immediately follows from these assumptions that j_{\min} satisfies (4.45). Furthermore, without GWs the entire eccentricity evolution is dictated by equation (4.12). It will therefore be important that we are able to derive simple expressions for j_{\pm} , j_0 etc. in the high-eccentricity limit. By evaluating (4.15), (4.16) at $\omega = \pm\pi/2$, $e \approx 1$ we find the approximate results

$$\Sigma \approx \frac{\epsilon_{\text{GR}}}{6j_{\min}} + \frac{5\Gamma\Theta}{j_{\min}^2}, \quad (5.2)$$

$$D \approx 1 + \frac{10\Gamma}{1-5\Gamma} \left(1 - \frac{\Theta}{j_{\min}^2} - \frac{\epsilon_{\text{GR}}}{30\Gamma j_{\min}} \right). \quad (5.3)$$

It then follows from (4.39), (4.14) that

$$j_+^2 \approx \frac{10\Gamma}{1+5\Gamma} \left(\frac{\Theta}{j_{\min}^2} + \frac{\epsilon_{\text{GR}}}{30\Gamma j_{\min}} \right), \quad (5.4)$$

$$j_-^2 \approx j_{\min}^2 \left(1 + \frac{\epsilon_{\text{GR}} j_{\min}}{30\Gamma\Theta} \right)^{-1}, \quad (5.5)$$

$$j_0^2 \approx \frac{10\Gamma}{5\Gamma-1} \left(1 - \frac{\Theta}{j_{\min}^2} - \frac{\epsilon_{\text{GR}}}{30\Gamma j_{\min}} \right). \quad (5.6)$$

5.2.1 Minimum and maximum j in the weak-to-moderate GR regime

Minimum j

At this stage we have not said anything about Θ , nor about the size of ϵ_{GR} (i.e. we have not specified to the weak/moderate/strong GR regime). However, we know from Chapter 4 that for a binary to start at $e \sim 0$ and to end up at $e \rightarrow 1$, it must initially reside in the weak-to-moderate GR regime and have $\Theta \ll 1$. Supposing this is true, for all orbits that maximise their eccentricity at $\omega = \pm\pi/2$ the solution j_{\min} is given by equation (4.45). We used this result to argue in §4.4 that for weak GR,

$$j_{\min} \sim \Theta^{1/2} \gg \epsilon_{\text{GR}}, \quad (\text{weak GR}). \quad (5.7)$$

For moderate GR, we argued in §4.4 that $j_{\min} \sim \epsilon_{\text{GR}} \gg \Theta^{1/2}$. However, as we mentioned there, this scaling implicitly assumed that Θ was kept fixed while ϵ_{GR} was increased, and this is no longer true when we consider shrinking a . In fact it will turn out that Θ/j_{\min}^2 is essentially constant as the binary shrinks (§5.3.1). Thus for the moderate GR regime we find in this case

$$j_{\min} \sim \epsilon_{\text{GR}} \sim \Theta^{1/2}, \quad (\text{moderate GR}). \quad (5.8)$$

It follows from equations (5.4), (5.5) and (5.7)-(5.8) that in the weak-to-moderate regime, provided $\Gamma \sim 1$, we always have $j_+^2 \sim 1$ and $j_-^2 \lesssim j_{\min}^2 \sim \Theta \ll 1$.

Maximum j

So, for all the phase space orbits in which we are interested, j_{\min} is given by the same formula (4.45). However it will turn out that it is extremely important to be able to distinguish between qualitatively different orbits, and in particular to know their maximum angular momentum j_{\max} (corresponding to minimum eccentricity e_{\min}), so we will devote some effort to this now.

The maximum j can either be found at $\omega = \pm\pi/2$ (if the phase space trajectory librates) or at $\omega = 0$ (if it circulates). In the librating case we find²

$$j_{\max} \approx j_+, \quad (\text{librating orbits}). \quad (5.10)$$

Finding an approximate expression for j_{\max} for circulating orbits is more complex, because there are two qualitatively different types of circulating orbit to consider, and it will be important that we differentiate them clearly. The first type of circulating trajectory, which we call ‘Type 1’, corresponds to $j_{\max} \sim 1$ or $e_{\min} \sim 0$, i.e. the binary undergoes an order-unity oscillation in eccentricity during each secular cycle. This is the classic type of circulating orbit undergone by, for instance, a binary with $\Gamma > 1/5$ in the weak GR regime starting out with small eccentricity at $\omega \approx 0$ — see e.g. 5.6g for illustration. The second type of circulating solution, which we call ‘Type 2’, corresponds to $j_{\max} \ll 1$ or $e_{\min} \sim 1$, so that the ‘oscillation’ in eccentricity is actually rather small despite e_{\max} being very large. These Type 2 orbits are important because every binary passes through this stage while in the moderate GR regime during a slow merger, as a precursor to the strong GR regime³ — see e.g. Figure 5.6h. Indeed, it will turn out that Type 2 circulating orbits in the moderate GR regime have very similar $a(t)$ decay behaviour to those in the strong GR regime, whereas Type 1 orbits show a qualitatively different behaviour — see §5.3.5.

To make this distinction between Type 1 and Type 2 circulating orbits quantitative, recall that in either case j_{\max} is a solution to the cubic (4.69), which we write here as

$$j_{\max}(j_{\max}^2 - j_0^2) = -q \equiv \frac{\epsilon_{\text{GR}}}{3(5\Gamma - 1)}. \quad (5.11)$$

²To see this, recall that for librating orbits j_{\max} is a solution to the quartic found by setting the first square bracket in (4.12) to zero. We can simplify this quartic by noting that, since librating orbits loop around fixed points at $\omega = \pm\pi/2$, they necessarily have $j_{\max}^2 > j_{f,\pi/2}^2$. We know from Figure 4.3 that $j_{f,\pi/2}^2 \gg \Theta$, and from equations (5.7)-(5.8) that $\Theta \gtrsim j_-^2$, so we can ignore j_- in the quartic and write

$$j_{\max}^3 - j_+^2 j_{\max} + \epsilon_{\text{GR}}/\epsilon_{\text{strong}} \approx 0. \quad (5.9)$$

Since in the weak-to-moderate GR regime we have $j_+^2 \sim 1$ and $\epsilon_{\text{GR}} \ll \epsilon_{\text{strong}}$, the solution is obviously $j_{\max} \approx j_+$.

³For $0 < \Gamma \leq 1/5$, we know from Chapter 4 that Type 2 circulating trajectories are immediately formed once ϵ_{GR} exceeds $6(1 - 5\Gamma)\Theta^{3/2}$, which is $\ll \epsilon_{\text{weak}}$ (equation (4.28); see also e.g. Figure 4.6). Thus, one does not necessarily need to be in the moderate GR regime to have Type 2 circulating orbits. However most of our focus in this Chapter will be on Type 2 circulating orbits that exist in the moderate GR regime, which occur for all Γ .

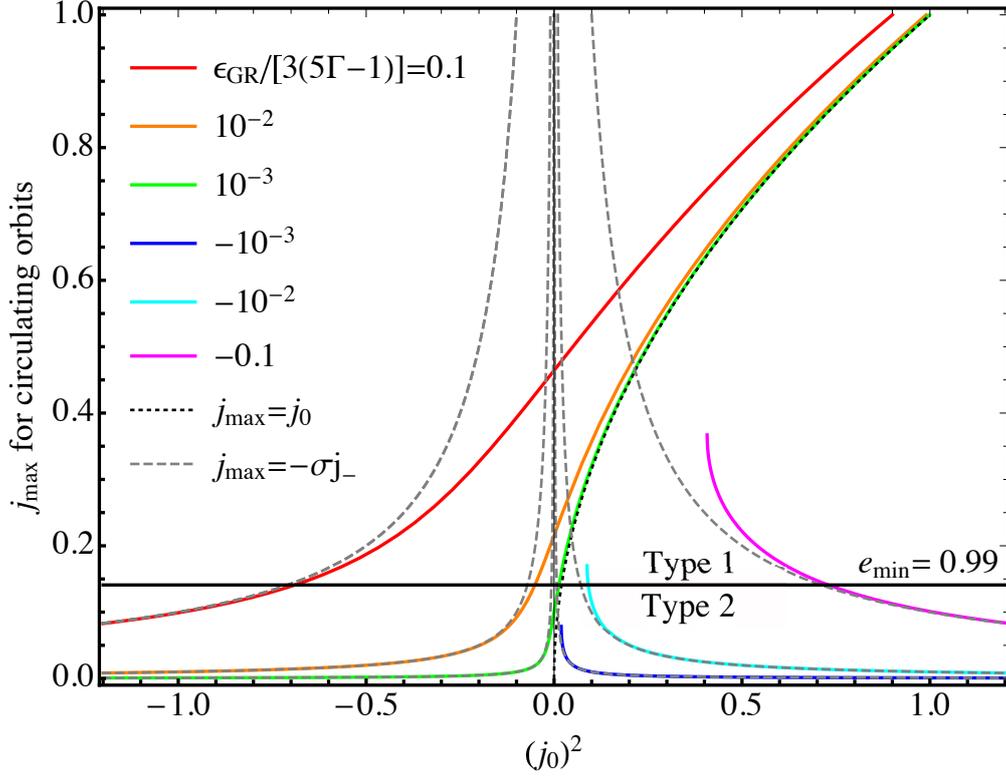


Figure 5.1: Solid coloured lines show j_{\max} (the physical solutions to equation (5.11)) for circulating orbits as a function of j_0^2 , for different values of $\epsilon_{\text{GR}}/[3(5\Gamma - 1)]$, shown with different colours. (We know from Chapters 3-4 that j_0^2 can be negative). A horizontal black line at $j_{\max} = 0.141$ ($e_{\min} = 0.99$) separates ‘Type 1’ and ‘Type 2’ circulating solutions. Grey dashed lines show the solution $j_{\max} = -\sigma j_-$, valid for Type 2 circulating orbits (see equation (5.13)). The black dotted line shows $j_{\max} = j_0$, valid for Type 1 circulating orbits (equation (5.12)).

One can solve this cubic analytically using the machinery of §4.A.3, but for simplicity here we will just plot the answer. Figure 5.1 shows j_{\max} as a function of j_0^2 (which we know can be positive or negative, see Chapters 3-4) for different values of $-q \equiv \epsilon_{\text{GR}}/[3(5\Gamma - 1)]$ (different coloured solid lines). We have (somewhat arbitrarily) chosen to split the figure into a Type 1 region ($j_{\max} > 0.141$, i.e. $e_{\min} < 0.99$) and a Type 2 region ($j_{\max} < 0.141$, i.e. $e_{\min} > 0.99$). Thus for nearly all positive values of j_0^2 the solid red, orange and green curves correspond to Type 1 circulating orbits; they are well approximated by ignoring the ϵ_{GR} term in (5.11) and putting

$$j_{\max} \approx j_0 \sim 1, \quad (\text{Type 1 circulating orbits}). \quad (5.12)$$

In Figure 5.1 we also plot $j_{\max} = j_0$ with a dotted black curve, and we see that the Type 1 curves asymptote towards this solution for sufficiently large $j_0^2 > 0$. Meanwhile the solid curves in Figure 5.1 with $j_{\max} \lesssim 0.141$ correspond to Type 2 circulating orbits. Assuming for these orbits that $j_{\max} \ll |j_0|$, we find that (5.11) has the approximate solution

$$j_{\max} \approx \frac{\epsilon_{\text{GR}}}{3(5\Gamma - 1)(-j_0^2)} = -\sigma j_- \ll 1, \quad (\text{Type 2 circulating orbits}). \quad (5.13)$$

In Figure 5.1 we plot this solution with different grey dashed curves for the different values of $\epsilon_{\text{GR}}/[3(5\Gamma - 1)]$ as for the solid lines. We again see that the solid Type 2 curves asymptote towards the solution (5.13) for sufficiently large $|j_0^2|$.

Of course by specifying these two asymptotic branches we have ignored one possible regime, namely that of Type 2 circulating orbits with small $|j_0|$, i.e. $|j_0| \lesssim j_{\text{max}} \ll 1$. However as Figure 5.1 shows, such solutions only exist for a narrow range of j_0^2 values centred around zero; and as we will see later, this regime is typically ‘short-lived’ in the sense that a shrinking binary passes through it rather quickly on the way to merger. In what follows we will ignore this case, so that Type 2 circulating orbits always have $|j_0| \gg j_{\text{max}}$ (which obviously implies $|j_0| \gg j_{\text{min}}$).

For negative values of $\epsilon_{\text{GR}}/[3(5\Gamma - 1)]$, i.e. for $\Gamma < 1/5$, we see that no solution exists below some (positive) value of j_0^2 . This is consistent with what we found in §4.B, where we saw that the existence of such solutions (which for $\epsilon_{\text{GR}} = 0$ just correspond to $j_{\text{max}} = j_0$) requires the binary to reside in the very weak GR regime. Indeed, a sufficiently positive j_0^2 is always required to render negative the left hand side of equation (5.11).

5.2.2 The strong GR regime

Finally, we recall from §4.4.5 that in the asymptotic limit of strong GR, cluster tides are negligible and the lowest order solution consists of pure precession at a rate

$$\dot{\omega}_{\text{GR}} = \frac{3[G(m_1 + m_2)]^{3/2}}{c^2 a^{5/2} j^2}, \quad (5.14)$$

In this limit there are no secular oscillations and hence there is no ‘secular timescale’ to speak of.

5.3 Gravitational wave emission

We now wish to include GW emission in our calculations. Our primary goal in this section is to compute $\Delta a(a)$ and $t_{\text{sec}}(a)$ for the different types of orbit and different dynamical regimes described above, and to use these to understand the slow evolution of $a(t)$.

To 2.5th post-Newtonian order, GW emission causes the binary’s semimajor axis and eccentricity to evolve according to the Peters (1964) equations⁴:

$$\left(\frac{da}{dt}\right)_{\text{GW}} = -\frac{64G^3 m_1 m_2 (m_1 + m_2)}{5c^5} \times \frac{1}{a^3(1 - e^2)^{7/2}} \left(1 + \frac{73}{24}e^2 + \frac{37}{96}e^4\right), \quad (5.15)$$

$$\left(\frac{de}{dt}\right)_{\text{GW}} = -\frac{304G^3 m_1 m_2 (m_1 + m_2)}{15c^5} \times \frac{1}{a^4(1 - e^2)^{5/2}} \left(1 + \frac{121}{304}e^2\right). \quad (5.16)$$

⁴Of course, in principle one can carry the post-Newtonian expansion to higher order (e.g. Pati & Will 2002). The higher-order terms are important for e.g. LIGO/Virgo templates of inspiralling binary waveforms. However, since this is only important at very late times when compared to the long secular evolution that we are considering here, we always truncate at 2.5pN.

To include this effect in our formalism we therefore add the following terms to our equations of motion for L , J and J_z :

$$\left(\frac{dL}{dt}\right)_{\text{GW}} = \frac{1}{2} \sqrt{\frac{G(m_1 + m_2)}{a}} \left(\frac{da}{dt}\right)_{\text{GW}}, \quad (5.17)$$

$$\left(\frac{dJ}{dt}\right)_{\text{GW}} = \sqrt{1 - e^2} \left(\frac{dL}{dt}\right)_{\text{GW}} - \frac{eL}{\sqrt{1 - e^2}} \left(\frac{de}{dt}\right)_{\text{GW}}, \quad (5.18)$$

$$\left(\frac{dJ_z}{dt}\right)_{\text{GW}} = \frac{J_z}{J} \left(\frac{dJ}{dt}\right)_{\text{GW}}, \quad (5.19)$$

Note that at this order, GW emission does not directly affect ω or Ω . It is clear from equations (5.15) and (5.16) that for a fixed semimajor axis a , GW losses are very strongly concentrated around peak eccentricity $e \rightarrow 1$, as we already anticipated in §5.1.2. In the weak-to-moderate GR regime we expect these GW contributions to the equations of motion to be completely negligible except in the vicinity of $e \approx 1$.

5.3.1 Conservation laws

When GW emission was switched off, the DA dynamics respected three exact conservation laws. The first was the conservation of a , which resulted from the ‘adiabatic’ assumption that the timescale of variation of the weak cluster perturbation (i.e. the outer orbital timescale) was much longer than that of the binary’s inner orbit. The second was the conservation of H^* , which followed from the fact that cluster tides were sufficiently weak that we could average over the outer orbital period and thus treat the perturbation as time-independent. The third was conservation of the z component of the binary’s angular momentum, or equivalently Θ , which followed from the axisymmetry of the time-averaged potential.

Now that we are including GW emission, the binary’s binding energy and inner orbital angular momentum will be dissipated according to equations (5.15)-(5.19), and so none of a , H^* or Θ will be strictly conserved any longer. On the other hand, they can be treated as roughly conserved on a timescale $\lesssim t_{\text{sec}}$, since in the weak-to-moderate GR regime GW emission is ineffective except for a short burst around $e \approx e_{\text{max}}$. Additionally, in their place two new (approximate) conservation laws will arise which are valid on much longer timescales — indeed they hold almost all the way to merger, and this will ease our calculations greatly. These are the conservation of the minimum pericentre distance p_{min} , and the conservation of the minimum inclination i_{min} .

Conservation of p_{\min}

Consider the rate of change of the pericentre distance with respect to semimajor axis:

$$p \equiv a(1 - e) \quad \implies \quad \frac{dp}{da} = 1 - e - a \frac{de}{da}. \quad (5.20)$$

Dividing (5.16) by (5.15) and taking $e \rightarrow 1$, it is easy to show that⁵ $dp/da \propto (1 - e)^2 \rightarrow 0$. In other words, the pericentre distance does not change across the eccentricity peak. Thus we arrive at the conservation of the *minimum pericentre distance*:

$$p_{\min} \equiv a(1 - e_{\max}) \approx \frac{1}{2} a j_{\min}^2 = \text{const.} \quad (5.21)$$

This quantity is approximately constant from one secular peak to the next, even though a and e_{\max} themselves change (see e.g. Figure 5.4.1d for a numerical example). Physically, this is because dissipation of energy in the form of GWs is concentrated around the peak of each secular eccentricity cycle, when $p \approx p_{\min}$. Since the effective ‘lever arm’ p is then very small, the binary dissipates angular momentum rather inefficiently compared to energy⁶ $\propto [a(1 - e^2)]^{1/2} \approx (2p)^{1/2}$, the angular momentum at the peak $J_{\min} \propto (2p_{\min})^{1/2}$ does not change across the peak (see Wen 2003, §3.1). And since the system is periodic, for small $|\Delta a| \ll a$ the binary returns to essentially the same value of J_{\min} (and hence the same p_{\min}) at the peak of the following secular cycle.

Equation (5.21) implies a simple scaling for j_{\min} :

$$j_{\min}(a) \approx \left(\frac{2p_{\min}}{a} \right)^{1/2}, \quad \text{where} \quad p_{\min} = a(t_i) \times (1 - e(t_i)), \quad (5.22)$$

and t_i is some reference time. Equation (5.22) is a central result of this Chapter and will allow significant simplification of our analytical formulae as we proceed.

We note that a similar argument to this one also applies when the binary gets trapped at high eccentricity in the strong GR regime. In that case there are no more secular oscillations (the binary having decoupled from cluster tides), but GW energy is still being dissipated efficiently while angular momentum is not (for more details see Wen (2003) and Antognini et al. (2014)). As a result p (rather than just p_{\min}) is constant, so $j \propto a^{-1/2}$. We will use this scaling to solve for $a(t)$ in the strong GR regime in §5.3.5.

⁵On the other hand, if one naively calculates dp/dt and takes $e \rightarrow 1$, then one finds that dp/dt diverges. However, it is also the case that da/dt diverges for $e \rightarrow 1$ (as, for that matter, does de/dt). The key point is that although dp/dt becomes very large at high eccentricity, that high eccentricity phase is over extremely quickly, so that p suffers almost no change across it. Put differently, dp/dt and da/dt both diverge but $dp/da = (dp/dt)/(da/dt)$ does not.

⁶Somewhat more precisely, the very small periastron distance means that the GW emission acts effectively like an instantaneous force on the binary components, directed radially with respect to their barycentre at the time of closest passage (pulling them both ‘in’). This radial nudge will change the inner orbital energy, but obviously cannot change the angular momentum.

Conservation of i_{\min}

All the key quantities in §5.2 were written in terms of Γ , ϵ_{GR} , Θ and j_{\min} . For a given outer orbit, Γ is constant. And from equations (4.5) and (5.22) we know how ϵ_{GR} and j_{\min} scale with semimajor axis a . All that is left is to work out how Θ depends on a .

Recalling $\Theta \equiv J_z^2/L^2$ and using the chain rule it is easy to show that

$$\frac{d\Theta}{da} = -\frac{\Theta}{a} \left(1 - \frac{2a}{J} \frac{dJ}{da} \right). \quad (5.23)$$

We know from equation (5.21) that at high eccentricity, $p \approx J^2/[2G(m_1 + m_2)]$. It follows that for $e \rightarrow 1$,

$$\frac{2a}{J} \frac{dJ}{da} \approx \frac{1}{1-e} \frac{dp}{da} \rightarrow 0. \quad (5.24)$$

Thus $d\Theta/da \approx -\Theta/a$, which has the solution

$$\frac{\Theta(t)}{\Theta_i} = \frac{a_i}{a(t)}, \quad (5.25)$$

where $\Theta_i \equiv \Theta(t_i)$. Indeed, the fact that $\Theta \propto 1/a$ could have been guessed a priori on physical grounds, since $\Theta = J_z^2/[G(m_1 + m_2)a]$ and GWs do not appreciably change the binary angular momentum.

Thus, Θa is constant. Combining this with the definition $\Theta \equiv j_{\min}^2 \cos^2 i_{\min}$ and the conservation of p_{\min} (equation (5.21)), one concludes that

$$i_{\min} = \text{const.} \quad (5.26)$$

In other words, the binary reaches the same minimum inclination at the peak of each secular eccentricity cycle, regardless of the decay in a (see e.g. Figure 5.4.1c for a numerical example).

For the remainder of this Chapter we will take p_{\min} and i_{\min} as our two primary, a -independent constants of motion in the weak-to-moderate GR regime. We can write Θ in terms only of a and these conserved quantities as

$$\Theta(a) \approx \frac{2p_{\min}}{a} \cos^2 i_{\min}. \quad (5.27)$$

5.3.2 Evolution of a shrinking binary through phase space

In the previous subsection we have seen how j_{\min} and Θ depend on a — see equations (5.22) and (5.27). We can now use these results to understand how a binary moves through phase space as its semimajor axis shrinks.

To begin, we substitute (5.22), (5.27) and (4.5) into equations (5.4)-(5.6) to get j_{\pm}^2 , j_0^2 as explicit functions of semimajor axis:

$$j_+^2 \approx \frac{10\Gamma}{1+5\Gamma} \left[\cos^2 i_{\min} + \left(\frac{d}{a}\right)^{7/2} \right], \quad (5.28)$$

$$j_-^2 \approx \frac{2p_{\min}}{a} \left[1 + \frac{1}{\cos^2 i_{\min}} \left(\frac{d}{a}\right)^{7/2} \right]^{-1}, \quad (5.29)$$

$$j_0^2 \approx \frac{10\Gamma}{5\Gamma-1} \left[\sin^2 i_{\min} - \left(\frac{d}{a}\right)^{7/2} \right], \quad (5.30)$$

where we defined the lengthscale

$$\begin{aligned} d &\equiv \left(\frac{4G^2(m_1+m_2)^2}{5c^2 A \Gamma (2p_{\min})^{1/2}} \right)^{2/7} \\ &\approx 13 \text{ AU} \times \Gamma^{-2/7} \left(\frac{A^*}{0.5} \right)^{-2/7} \left(\frac{\mathcal{M}}{10^6 M_{\odot}} \right)^{-2/7} \left(\frac{b}{\text{pc}} \right)^{6/7} \left(\frac{m_1+m_2}{M_{\odot}} \right)^{4/7} \left(\frac{p_{\min}}{10^{-2} \text{ AU}} \right)^{-1/7}, \end{aligned} \quad (5.31)$$

$$(5.32)$$

which is independent of a . As usual, in the numerical estimate (5.32) we have assumed a spherical cluster of mass \mathcal{M} and scale radius b . (We reiterate that to evaluate any numerical result in the LK case one simply sets $\Gamma = 1$, $A^* = 0.5$ and $b = a_g(1 - e_g^2)^{1/2}$, where a_g and e_g are respectively the semimajor axis and eccentricity of the outer orbit). We can also calculate the important dimensionless quantities γ , σ and κ as functions of a , as follows. First, by combining equations (4.44), (5.28) and (5.29) it is straightforward to show that

$$\gamma(a) \equiv \frac{2\epsilon_{\text{GR}}}{\epsilon_{\text{weak}}} \approx \frac{1}{\sqrt{\zeta(\zeta+1)}}, \quad \text{where} \quad \zeta \equiv (a/d)^{7/2} \cos^2 i_{\min}. \quad (5.33)$$

Second, plugging (4.5), (5.29) and (5.30) into (4.43) we get:

$$\sigma(a) \approx \left[\left(\frac{d}{a}\right)^{7/2} \frac{1}{\cos^2 i_{\min}} + 1 \right]^{1/2} \left[\left(\frac{a}{d}\right)^{7/2} \sin^2 i_{\min} - 1 \right]^{-1}. \quad (5.34)$$

Third, we can take the ratio of (5.34) and (5.33) to get κ (equation (4.56)):

$$\kappa(a) \approx \left[\left(\frac{a}{d}\right)^{7/2} \cos^2 i_{\min} + 1 \right] \left[\left(\frac{a}{d}\right)^{7/2} \sin^2 i_{\min} - 1 \right]^{-1}. \quad (5.35)$$

It will also be helpful to define some critical values of the semimajor axis a . A binary that starts its life in the weak GR regime will inevitably move into the moderate GR regime at some point as its semimajor axis shrinks. Thus the first critical value is a_{weak} , which we define to be the semimajor axis corresponding to $\epsilon_{\text{GR}} = \epsilon_{\text{weak}}$, i.e. to the transition between the weak and moderate GR regimes. Using (5.33) this is

$$a_{\text{weak}} \equiv d \times \left(\frac{\sqrt{2}-1}{2 \cos^2 i_{\min}} \right)^{2/7} \approx 0.63 (\cos i_{\min})^{-4/7} d. \quad (5.36)$$

Note that, since $\cos i_{\min} \lesssim 1$, we will nearly always have $a_{\text{weak}} \sim d$, allowing us to simplify the preceding expressions for j_{\pm}^2, γ , etc. in the particular asymptotic regimes of weak ($(a/d)^{7/2} \gg 1$) or moderate ($(a/d)^{7/2} \ll 1$) GR. The second critical semimajor axis is a_{sep} , which we define to be the value corresponding to $j_+^2 = j_0^2 = 1$. Physically, $a = a_{\text{sep}}$ corresponds to a separatrix between librating ($a > a_{\text{sep}}$) and Type 1 circulating ($a < a_{\text{sep}}$) orbits. This is relevant because a binary initially on a librating phase space orbit will transition to a circulating orbit once a drops below this threshold value, as we will see. We find from equations (5.28), (5.30) that

$$a_{\text{sep}} \equiv \left(\frac{1 + 5\Gamma}{10\Gamma} - \cos^2 i_{\min} \right)^{-2/7} d. \quad (5.37)$$

Of course, a_{sep} only has physical meaning if $\cos^2 i_{\min} < (1 + 5\Gamma)/10\Gamma$. This is because for $\cos^2 i_{\min} > (1 + 5\Gamma)/10\Gamma$, the binary is already on a circulating orbit even for $a \gg d$, so it never crosses a separatrix on its way to $a \rightarrow 0$. We also note that typically $a_{\text{sep}} \gtrsim d$. The third critical semimajor axis is that corresponding to $j^2 = 0$, which we note also corresponds to a divergence in σ and κ ; this is a_{div} , given by:

$$a_{\text{div}} \equiv (\sin i_{\min})^{-4/7} d. \quad (5.38)$$

Lastly, for completeness let us define a_{strong} , which corresponds to $\epsilon_{\text{GR}} = \epsilon_{\text{strong}} \equiv 3(1 + 5\Gamma)$, i.e. it demarcates the inevitable transition from moderate to strong GR regimes. Using (4.5) we get

$$\begin{aligned} a_{\text{strong}} &\equiv \left(\frac{8G^2(m_1 + m_2)^2}{c^2 A(1 + 5\Gamma)} \right)^{1/4}, \quad (5.39) \\ &\approx 6.4\text{AU} \times \left(\frac{1 + 5\Gamma}{6} \right)^{-1/4} \left(\frac{A^*}{0.5} \right)^{-1/4} \left(\frac{\mathcal{M}}{10^6 M_{\odot}} \right)^{-1/4} \left(\frac{b}{\text{pc}} \right)^{3/4} \left(\frac{m}{1.4 M_{\odot}} \right)^{1/2}. \end{aligned} \quad (5.40)$$

In summary, in this subsection we have defined four critical semimajor axis values $a_{\text{weak}}, a_{\text{sep}}, a_{\text{div}}, a_{\text{strong}}$. The weak GR regime corresponds to $a > a_{\text{weak}}$, while the moderate GR regime corresponds to $a_{\text{strong}} < a < a_{\text{weak}}$. We emphasise that we have purposely written e.g. $a > a_{\text{weak}}$ rather than $a \gg a_{\text{weak}}$ here: it will turn out that different dynamical regimes are not very well separated in semimajor axes (indeed, $a_{\text{weak}}, a_{\text{sep}}$ and a_{div} are all clustered around $\sim d$), despite being well separated in ϵ_{GR} .

Phase space evolution for $\Gamma > 1/5$

We now use these results to understand more precisely how binaries move through phase space as their semimajor axis shrinks. These considerations will help us develop approximate expressions for $t_{\text{sec}}, \Delta a$ and $a(t)$ in different regimes — see §§5.3.3-5.3.5. We begin with the regime $\Gamma > 1/5$.

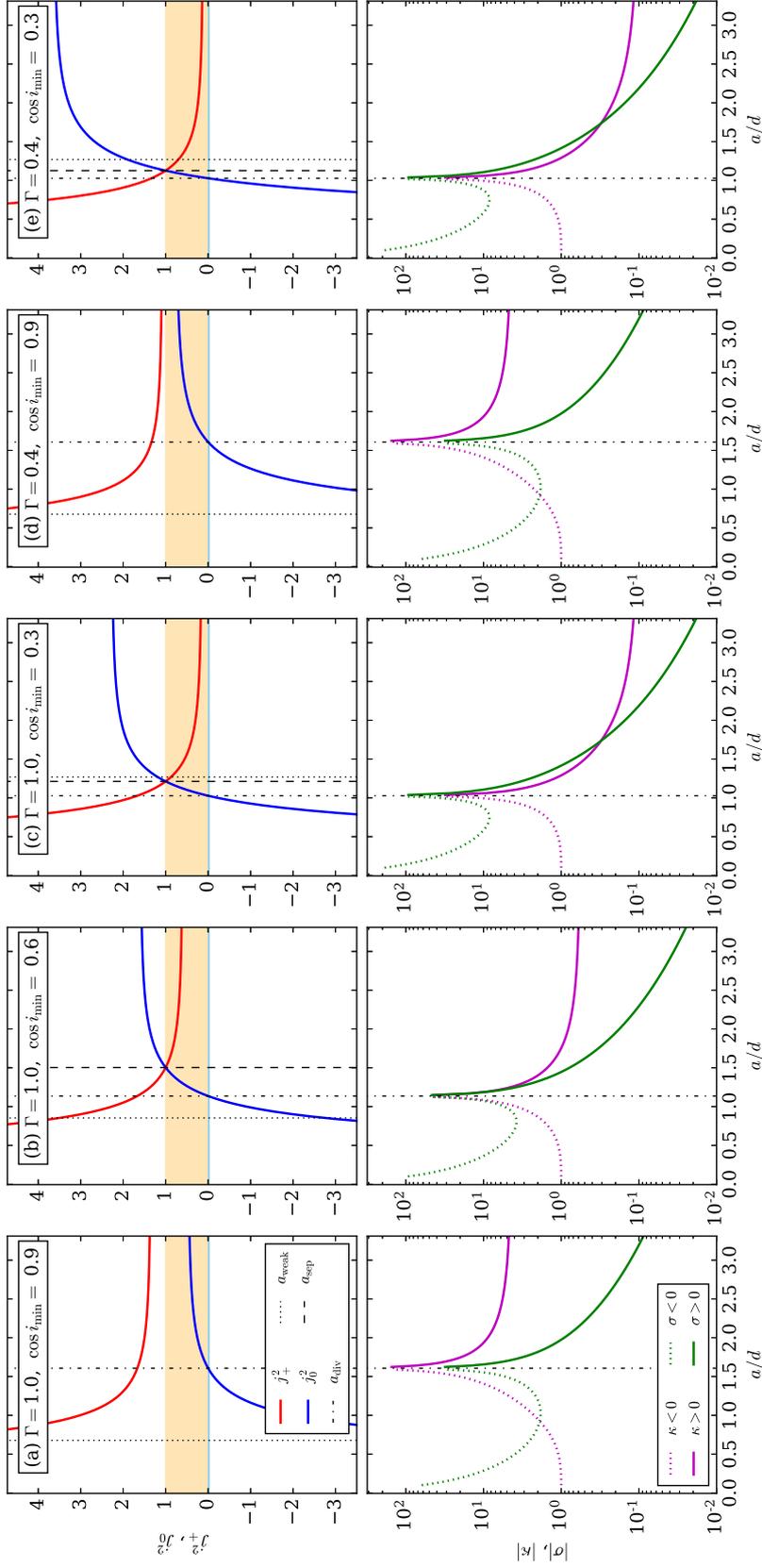


Figure 5.2: Plots of the key quantities j_+^2 , j_0^2 , σ and κ as functions of a/d according to equations (5.28), (5.30), (5.34) and (5.35) respectively, for different values of $\Gamma > 1/5$ and $\cos i_{\min}$. In the lower panels, for which the vertical axis is on a logarithmic scale, we show negative values of σ, κ with dotted curves and positive values with solid curves. In each panel we also show a_{weak} (vertical dotted line), a_{sep} (vertical dashed line) and a_{div} (vertical dot-dashed line), defined in equations (5.36), (5.37) and (5.38) respectively. Finally in the upper row we show with blue shading the (very thin) region $|j^2| < (0.141)^2 \approx 0.02$, within which the split into ‘Type 1’ and ‘Type 2’ circulating orbits is invalid (§5.2.1), and with pale orange shading the region $0 < j^2 < 1$.

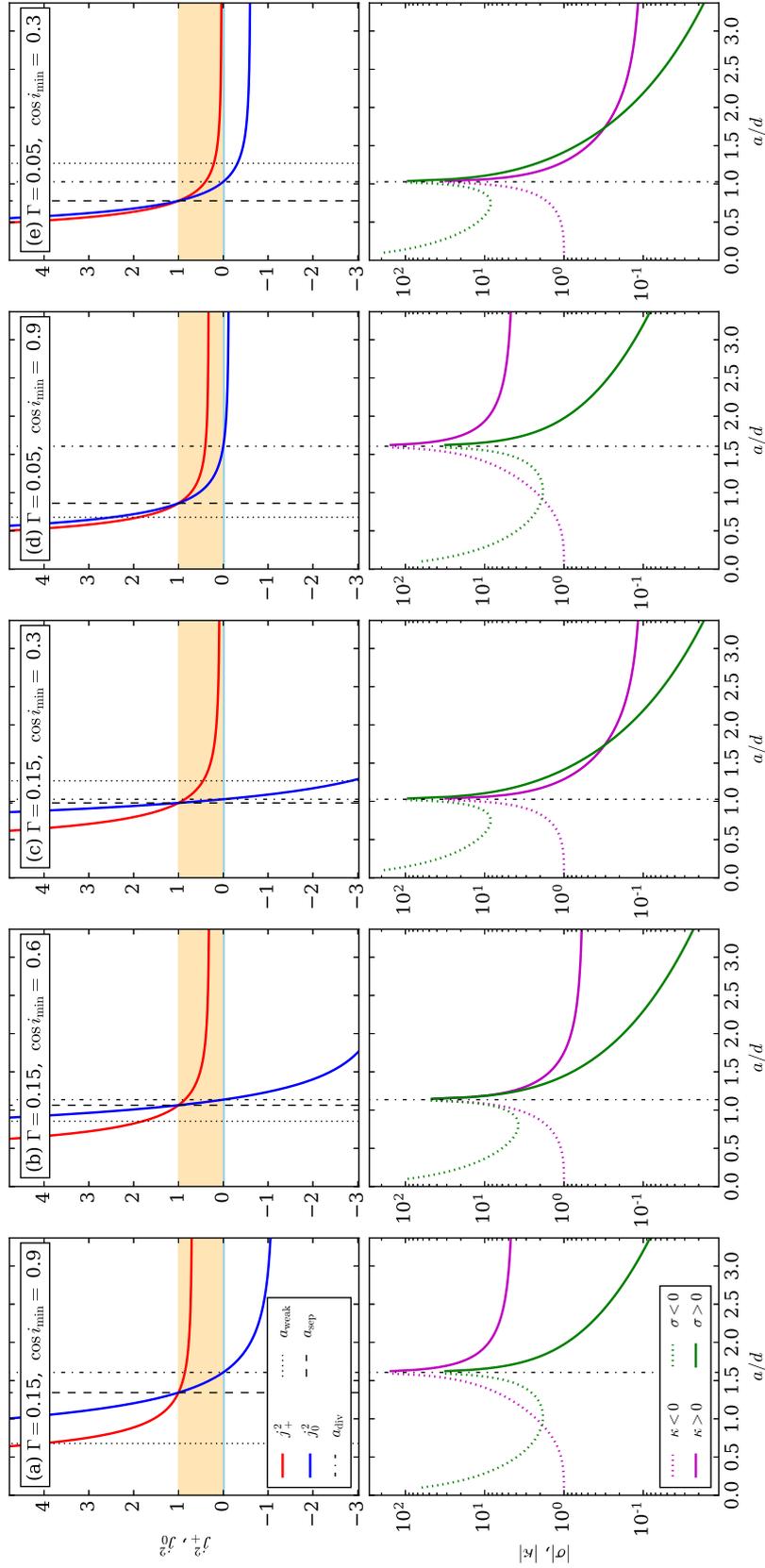


Figure 5.3: As in Figure 5.2 but for the regime $0 < \Gamma \leq 1/5$.

In Figure 5.2 we plot j_+^2 , j_0^2 , $|\sigma|$ and $|\kappa|$ as functions of a/d for various fixed values of Γ and $\cos i_{\min}$, according to equations (5.28), (5.30), (5.34) and (5.35) respectively. We also show the critical values a_{weak} (dotted vertical line), a_{sep} (dashed vertical line) and a_{div} (dot-dashed vertical line), defined in equations (5.36)-(5.38). Additionally, in the upper panels we show with blue shading the region $|j^2| < (0.141)^2$ and with orange shading the region $0 < j^2 < 1$. In particular, by looking at the runs of j_+^2 and j_0^2 and whether they lie in this orange region, we will be able to infer the value of j_{\max} and hence infer what type of phase space orbit the binary is on.

Without loss of generality, for each example (a)-(e) we can consider a binary that starts at the extreme right of each panel, i.e. with $a \gg a_{\text{weak}}$ (the weak GR regime), and follow it as a decreases. First we focus on examples (a)-(c), which are for $\Gamma = 1$ (the LK limit).

In panel (a) we see that for very large a , we have $j_+^2 > 1$ while $j_0^2 \sim 0.5$. This corresponds to a binary on a Type 1 circulating orbit in the weak GR regime, with $j_{\max} \approx j_0$ (equation (5.12)). Indeed in this case, since $\cos^2 i_{\min} > (1 + 5\Gamma)/10\Gamma = 0.6$, we have $j_+^2 > 1$ for all a (equation (5.28)), so the red curve never enters the orange region and the binary remains on a circulating orbit always. However, once a becomes smaller than a_{div} we quickly get j_0^2 values that are strongly negative, so the binary has transitioned to a Type 2 circulating orbit (Figure 5.1) with $j_{\max} \approx -\sigma j_- \ll 1$. It will remain on such an orbit until it gets trapped at high eccentricity in the strong GR regime around $a \sim a_{\text{strong}}$ (not shown here).

Example (b), in which we replace $\cos i_{\min} = 0.9$ with $\cos i_{\min} = 0.6$, shows different behaviour. In this cases the binary ‘begins’ at large a with $0 < j_+^2 < 1$ and $j_0^2 > 1$; this means that it is on a librating orbit in the weak GR regime, with $j_{\max} \approx j_+$ (equation (5.10)). The reason for the difference compared to panel (a) is that we now have $\cos^2 i_{\min} \leq (1 + 5\Gamma)/10\Gamma$, meaning librating orbits are possible. Of course as a is decreased j_+^2 is always increased, while j_0^2 is decreased, and when $a = a_{\text{sep}}$ the two cross over, $j_0^2 = j_+^2 = 1$. At this point the binary switches to a Type 1 circulating orbit with $j_{\max} \approx j_0$. In this case $a_{\text{sep}} < a_{\text{weak}}$, so that the separatrix crossing occurs while the binary is still nominally in the weak GR regime. Thereafter the binary follows the same evolutionary track as in panel (a), moving into the moderate GR regime and onto a Type 2 circulating orbit.

Example (c) shows very similar behaviour to example (b), except that the smaller value of $\cos i_{\min}$ means that the three values a_{weak} , a_{sep} and a_{div} are now even more closely clustered together around $a/d \approx 1$ (note also that a_{sep} is now very slightly smaller than a_{weak}). Because of this clustering, example (c) is perhaps the ‘cleanest’ of the three examples shown so far: for a significantly larger than d the binary is clearly on a librating orbit in the weak GR regime, whereas for a significantly smaller than d it is clearly on a

Type 2 circulating orbit in the moderate GR regime. In practice the transitions between these two various regimes are not always so well demarcated.

At this stage it is worth noting how different quantities scale with a in each regime. From all examples (a)-(c) we see that in the weak GR regime ($a > a_{\text{weak}}$) we nearly always have $|j_+^2|, |j_0^2| \gg 0.1$, and both of these quantities scale very weakly with a . In the moderate GR regime ($a < a_{\text{weak}}$) the scaling of j_+^2 and j_0^2 with a is much stronger, as we would expect from equations (5.28), (5.30). Moreover, as anticipated in §5.2.1, in every case it is clear that $|j_0|^2$ lies in the blue shaded region only for a very narrow range of semimajor axes surrounding a_{div} , and so we were justified in ignoring the small j_0 regime when discussing Type 2 circulating orbits in §5.2.1. Turning to the bottom panels, we see that $|\sigma|$ and $|\kappa|$ both vary over several orders of magnitude as a is decreased. However, it is noteworthy that for a far away from a_{div} , the value of $|\kappa|$ is usually $\mathcal{O}(1)$ and scales weakly with a .

Finally we turn to examples (d) and (e), which are for the same $\cos i_{\text{min}}$ value as examples (a) and (c) respectively except with $\Gamma = 0.4$. The physical interpretation of these examples is identical to those of (a) and (c), demonstrating a broad uniformity of evolution for all binaries in the $\Gamma > 1/5$ regime⁷.

Phase space evolution for $0 < \Gamma \leq 1/5$

In Figure 5.3 we plot the same quantities as in Figure 5.2, except this time we focus on the regime $0 < \Gamma \leq 1/5$, replacing $\Gamma = 1, 0.4$ with $\Gamma = 0.15, 0.05$. We see that a rather different phase space evolution emerges.

First we consider panel (a), which is for $\Gamma = 0.15$ and $\cos i_{\text{min}} = 0.9$. In this case, for large $a \gg d$ we have $j_+^2 \lesssim 1$ and $j_0^2 \approx -1$. This means that in the asymptotic weak GR regime the binary is on a librating orbit, with $j_{\text{max}} \approx j_+$. This is unsurprising since, as we saw in Chapter 3 with GR switched off, for $0 < \Gamma \leq 1/5$, circulating orbits very rarely reach high eccentricity. However, once a decreases below a_{div} in this plot, we see that j_0^2 becomes positive (though still smaller than j_+^2). Soon a reaches a_{sep} , below which both j_+^2 and j_0^2 are greater than unity: the binary has transitioned onto a Type 2 circulating orbit (Figure 5.1). We note that all of this happens well before the binary reaches the moderate GR regime. This also is not so surprising because we know that a family of high-eccentricity circulating orbits (i.e. Type 2) naturally arises in the $0 < \Gamma \leq 1/5$ regime as soon as ϵ_{GR} exceeds $6(1 - 5\Gamma)\Theta^{3/2} \ll \epsilon_{\text{weak}}$ (Chapter 4). The binary stays on its Type 2 circulating orbit as a shrinks into the moderate GR regime $a < a_{\text{weak}}$ and onwards to the strong GR regime.

A very similar story holds in panels (b)-(e). The only important difference is that as we decrease Γ or $\cos i_{\text{min}}$, or both, the value of $j_{\text{max}} \approx j_+$ for asymptotically weak GR

⁷As might be expected from Chapters 3 and 4 this broad-brush picture can break down very close to $\Gamma = 1/5$, but we ignore this complication here.

($a \gg d$) decreases. This means that librating orbits with high $e_{\max} \rightarrow 1$ in the very weak GR regime do not reach low⁸ e_{\min} (recall that we have assumed $j_{\min} \ll 1$ in deriving our expression for j_+^2). Said differently, binaries that initially have $e \sim 0$ do not tend to reach $e \rightarrow 1$, so this is typically not the type of situation in which we are interested.

Finally we mention that in all examples shown in Figure 5.3, for a sufficiently far from a_{div} we again have $|\kappa| \sim \mathcal{O}(1)$ or smaller, and κ varies only weakly with a .



The phase space evolution described in this section is quite complex: the various dynamical regimes are not typically well separated in semimajor axis ($a_{\text{weak}} \sim a_{\text{sep}} \sim a_{\text{div}} \sim d$), and the transitions between different regimes do not always happen in the same order. However, to gain a qualitative understanding, in what follows we will focus on three basic asymptotic regimes.

1. The asymptotic regime of weak GR, $a \gg a_{\text{weak}}$, in which the orbit can be librating or (Type 1) circulating but we always have $j_{\min} \approx j_- \sim \cos i_{\min}$ and $|j_+^2|, |j_0^2| \sim 1$. Moreover, in this weak GR regime both $|j_+^2|$ and $|j_0^2|$ scale weakly with a , as does κ .
2. The moderate GR regime in which crucially *the binary is assumed to be on a Type 2 circulating orbit* with $j_{\min} \approx \gamma j_- \gg j_-$, $j_{\max} = -\sigma j_-$, and we assume that a is sufficiently small that $|j_+^2|, |j_0^2| \sim (a/d)^{7/2} \gg 1$, and $|\kappa|$ is roughly constant and not large compared to unity.
3. The third is the strong GR regime, $a < a_{\text{strong}}$, in which the binary gets trapped at high eccentricity, and its semimajor axis decays while $p = a(1 - e)$ is kept roughly constant.

Again, we emphasise that in general the separation between these three specific regimes is far from clean, and these are not the only three possibilities; but they will allow us to make analytical progress, and will give us a qualitative understanding of the behaviour of t_{sec} , Δa and $a(t)$ throughout slow mergers, which is what we turn to next.

⁸This is essentially because the fixed points at $\omega = \pm\pi/2$, $j = j_f = (10\Gamma\Theta/(1 + 5\Gamma))^{1/4}$, sit at too high an eccentricity.

5.3.3 The secular timescale, t_{sec}

One crucial quantity in any study of tide-driven mergers is the period of secular eccentricity oscillations, t_{sec} , since this gives the time elapsed between each burst of GW emission. The exact expression for t_{sec} is

$$t_{\text{sec}} = 2 \int_{j_{\text{min}}}^{j_{\text{max}}} \left(\frac{dj}{dt} \right)^{-1} dj, \quad (5.41)$$

with dj/dt given in (4.12). The right hand side of (5.41) of course depends on semimajor axis a . Here we want to derive a tractable approximation to $t_{\text{sec}}(a)$ which is valid in the first two regimes mentioned above, namely for weak GR (librating or circulating) and moderate GR (Type 2 circulating). In the third asymptotic regime, that of strong GR, there are no secular oscillations.

Weak GR

In the case of weak GR we know that the explicit ϵ_{GR} terms in (4.12) only affect the very high eccentricity behaviour. We also know from §4.4.4 that the effect of such terms is simply to reduce the time spent in the high-eccentricity state (see equation (4.52)), which is already a fraction $j_{\text{min}} \ll 1$ of the total secular period. Put differently, we can split (5.41) into $\int_{j_{\text{min}}}^{j_{\text{max}}} = \int_{j_{\text{min}}}^{kj_{\text{min}}} + \int_{kj_{\text{min}}}^{j_{\text{max}}}$; then this argument suggests there exists $k \sim$ a few such that the explicit ϵ_{GR} terms in (4.12) only contribute to the first integral, and that the value of this first integral is negligible compared to the second one. As a result, a good approximation to the secular period for most orbits in the weak-to-moderate GR regime is found by applying the results of Chapter 3. Defining $\Delta \equiv \max[j_+^2, j_-^2, j_0^2] - \min[j_+^2, j_-^2, j_0^2]$, and assuming $k^2 j_{\text{min}}^2 \ll j_{\text{max}}^2$, we have from equations (3.33)-(3.34) that

$$t_{\text{sec}} \approx \frac{8}{3A} \sqrt{\frac{G(m_1 + m_2)}{|25\Gamma^2 - 1|}} \times \frac{1}{a^{3/2}\sqrt{\Delta}} K\left(\frac{j_{\text{max}}}{\sqrt{\Delta}}\right). \quad (5.42)$$

where $K(x) \equiv \int_0^{\pi/2} d\alpha / \sqrt{1 - x^2 \sin^2 \alpha}$. This result is clearly insensitive to the precise value of k . We emphasise however that GR is still implicitly present in equation (5.42) because it affects the values of j_{\pm} , j_0 , Δ , j_{max} that must be plugged into the right hand side. Finally, Figure 5.4 shows that although $K(x)$ diverges for $x \rightarrow 1$ (which corresponds to the binary approaching a separatrix in the (ω, e) phase space), far away from $x = 1$ the function $K(x)$ scales rather weakly with x .

Now, to evaluate (5.42) explicitly we need to know j_{max} and Δ , which in turn requires knowledge of the Γ regime, whether the orbit librates or circulates, etc. — see Table 3.1. We will not go through each case individually, but simply note that for weak GR we almost always have either $j_{\text{max}} \approx j_+$ and $\sqrt{\Delta} \approx |j_0|$, or $j_{\text{max}} \approx j_0$ and $\sqrt{\Delta} \approx |j_+|$; and assuming

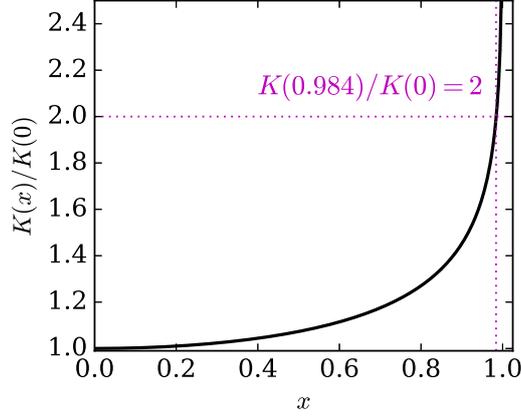


Figure 5.4: Value of the elliptical integral $K(x)$ normalised by $K(0) \equiv \pi/2$ as a function of x . The scaling is weak except for $x \gtrsim 0.99$.

$a \gg d$, both j_+, j_0 scale weakly with a (Figures 5.2-5.3). As a result, for most of the semimajor axis range of the weak GR regime we expect from equation (5.42) a rough scaling

$$t_{\text{sec}} \propto a^{-3/2}. \quad (5.43)$$

Physically this scaling arises because the binary's inner orbital period is proportional to $a^{3/2}$. It has been noted by many authors when estimating a LK-driven merger timescale (e.g. Wen 2003; Thompson 2011; Randall & Xianyu 2018).

Moderate GR, Type 2 circulation

The validity of equation (5.42) breaks down when we consider Type 2 circulating orbits, because in this case we cannot claim that the binary spends the great majority of its time at $j \sim 1$. Indeed, we know from §5.2.1 that for Type 2 circulating orbits in the moderate GR regime we have $j_{\text{max}} \approx -\sigma j_- = -\epsilon_{\text{GR}}/[3(5\Gamma - 1)j_0^2]$, so that $j^3 - j_0^2 j$ can never be assumed to dominate over the final term in the second square bracket in (4.12).

To get a formula for the secular timescale in this regime we assume that $j \ll 1$ is always small compared to $|j_0|$ (§5.2.1). Also, moderate GR means that $j_{\text{min}} \approx \gamma j_- \gg j_-$ and thus we may ignore j_-^2 compared to j^2 . With these assumptions equation (4.12) becomes

$$\frac{dj}{dt} \approx \pm \frac{6C}{Lj} \sqrt{|25\Gamma^2 - 1| |j_+ j_0|} \sqrt{(j - j_{\text{min}})(j_{\text{max}} - j)}, \quad (5.44)$$

where $j_{\text{min}} \approx \gamma j_-$ and $j_{\text{max}} = -\sigma j_-$. Plugging this into equation (5.41) and performing the integral we get the secular timescale for Type 2 circulating orbits in the moderate GR regime:

$$\begin{aligned} t_{\text{sec}} &\approx \frac{L\pi(j_{\text{min}} + j_{\text{max}})}{6C\sqrt{|25\Gamma^2 - 1| |j_+ j_0|}} \\ &\approx \frac{8}{3A} \sqrt{\frac{G(m_1 + m_2)}{|25\Gamma^2 - 1|}} \times \frac{\pi}{2a^{3/2}} \frac{(1 - \kappa)j_{\text{min}}}{|j_+ j_0|}. \end{aligned} \quad (5.45)$$

where in the second line we used $j_{\min} + j_{\max} \approx \gamma j_- - \sigma j_- = (1 - \kappa)j_{\min}$ (see equation (5.13)). Assuming also that the terms in $(a/d)^{7/2}$ dominate in the expressions for j_{\pm}^2 and j_0^2 (equations (5.28)-(5.30)), we find

$$t_{\text{sec}} \approx \frac{8}{3A} \sqrt{\frac{G(m_1 + m_2)}{|25\Gamma^2 - 1|}} \times \frac{\pi \sqrt{25\Gamma^2 - 1} (1 - \kappa) (2p_{\min})^{1/2} a^{3/2}}{20\Gamma d^{7/2}}. \quad (5.46)$$

Further assuming that a is sufficiently small compared to a_{div} for κ to vary only weakly with a (see Figures 5.2-5.3), we have in this regime rough scaling

$$t_{\text{sec}} \propto a^{3/2}. \quad (5.47)$$

According to (5.47), *the secular timescale decreases as the semimajor axis shrinks*, rather than increasing like one would naively expect (and as was predicted for the weak GR regime in equation (5.43)). This is because in the Type 2 circulating regime a smaller a leads to a larger $|j_0^2|$, hence a smaller j_{\max} (Figure 5.1). In turn, a smaller j_{\max} means that the binary spends more time at ‘high’ eccentricities. The cluster tide-driven secular evolution is faster at high e than at⁹ $e \sim 0$, which is how we end up with a shorter secular timescale. We will further explicate these ideas in §5.4 and §5.A.

Note also that in this case the secular timescale (equation (5.46)) is significantly shorter than in the weak GR case (equation (5.42)); the ratio of the former to the latter is of the order $(2p_{\min})^{1/2} a^3 / d^{7/2} \sim (a/d)^{7/2} j_{\min} \ll 1$.

5.3.4 The decay in semimajor axis over one secular cycle, Δa

Next we derive expressions for the decay in semimajor axis over one secular cycle, Δa , in terms of a , in the same asymptotic regimes as in §5.3.3. This expression can then be plugged into the right hand side of (5.1) along with t_{sec} (§5.3.3) to calculate $a(t)$.

To begin, we integrate equation (5.15) over one secular cycle, approximating a as constant to lowest order (which is valid since $|\Delta a| \ll a$ by assumption for a slow merger). The result is¹⁰

$$\Delta a \approx -\frac{\lambda_0}{a^3} \int_{\text{sec. cycle}} \frac{dt}{(1 - e^2)^{7/2}} \left(1 + \frac{73}{24} e^2 + \frac{37}{96} e^4 \right), \quad (5.48)$$

where $\lambda_0 \equiv (64/5)G^3 c^{-5} m_1 m_2 (m_1 + m_2)$ is independent of a . Assuming the binary reaches very high maximum eccentricity $e_{\max} \rightarrow 1$ we can approximate this as

$$\Delta a \approx -\frac{2\lambda_1}{a^3} \int_{j_{\min}}^{j_{\max}} \frac{dj}{j^7} \left(\frac{dj}{dt} \right)^{-1}, \quad (5.49)$$

⁹This is true because even though the torque on a binary with, say, $e = 0.1$ is comparable to that on a binary with $e = 0.9$, the angular momentum of the latter is significantly smaller, so the relative change in angular momentum occurs over a much shorter timescale.

¹⁰Note that equation (5.48) is essentially identical to the first line in equation (55) of Randall & Xianyu (2018) — see §5.A.

where $\lambda_1 \equiv (1 + 73/24 + 37/96)\lambda_0 = (170/3)G^3c^{-5}m_1m_2(m_1 + m_2)$. Of course, in general dj/dt — given in equation (4.12) — is so complicated that even this approximate integral is intractable. However, noting the very strong j^{-7} dependence in (5.49) we expect the integral to be dominated by the contributions from very high eccentricity, i.e. $j \ll j_+, |j_0|$. In this limit we can approximate dj/dt using equation (4.41); moreover, since we know that the minimum j_{\min} is a zero of the first square bracket in (4.41) we can write it as

$$\frac{dj}{dt} \approx \pm \frac{3Aa^{3/2}}{4\sqrt{G(m_1 + m_2)}j^{3/2}} \sqrt{(25\Gamma^2 - 1)j_+^2(-j_0^2)(j - j_{\min})(j + |j_\alpha|)(j_\sigma - j)}, \quad (5.50)$$

where $j_\alpha \equiv \gamma j_- [1 - \sqrt{1 + 4\gamma^{-2}}]/2 < 0$ is the other root of the first square bracket in (4.41), and $j_\sigma \equiv -\sigma j_-$. Using the results of §5.2 one can straightforwardly check that the sign of the quantity inside the square root is positive. For instance, for $\Gamma > 1/5$ we recall that Type 2 circulating orbits have $j_0^2 < 0$ and $j_{\max} = -\sigma j_- = j_\sigma$, while Type 1 circulating orbits have $j_0^2 > 0$, $j_\sigma < 0$ and $j_{\max} \sim 1$.

We now take (5.50) and plug it into (5.49). Defining

$$x_{\max} \equiv j_{\max}/j_{\min}, \quad x_\alpha \equiv j_\alpha/j_{\min}, \quad x_\sigma \equiv j_\sigma/j_{\min}, \quad (5.51)$$

and using (5.22), the result is

$$\Delta a \approx -\lambda_2 \times \frac{\xi(x_{\max}, x_\alpha, x_\sigma)}{a^{3/2}|j_+ j_0|}, \quad (5.52)$$

where $\lambda_2 \equiv 1360G^{7/2}m_1m_2(m_1 + m_2)^{3/2}/[9c^5A(2p_{\min})^3\sqrt{|25\Gamma^2 - 1|}]$ is independent of a , and

$$\xi(x_{\max}, x_\alpha, x_\sigma) \equiv \int_1^{x_{\max}} \frac{dx}{x^{11/2}\sqrt{(x-1)(x+|x_\alpha|)|x_\sigma-x|}}. \quad (5.53)$$

To simplify these expressions further we need to specify the asymptotic regime of interest.

Weak GR

For weak GR we have $j_{\min} \ll j_{\max}$ so that $x_{\max} \gg 1$. We also have $j_\sigma < 0$, so that $|x_\sigma - x| = x + |x_\sigma|$. In this case the integral in (5.53) is completely dominated by the contribution from $x \approx 1$, and so we may take the upper limit of the integral to $x_{\max} \rightarrow \infty$ with impunity. Finally, since $j_{\min} \approx j_-$ and $\gamma \ll 1$ in this limit we can simply replace x_σ with σ and $x_\alpha \rightarrow 1$. An excellent approximation to the resulting integral (accurate to within a few percent over several decades of $|\sigma|$) is then

$$\xi \approx \frac{8}{15} \frac{1}{\sqrt{1+|\sigma|}}. \quad (5.54)$$

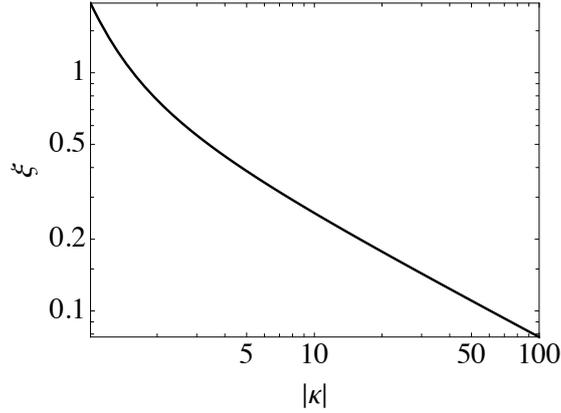


Figure 5.5: Plot of the function ξ as a function of $|\kappa|$ according to (5.57), which is valid for Type 2 circulating orbits in the moderate GR regime.

From Figures 5.2 and 5.3 we know that for weak GR, $\sigma \lesssim 1$; thus ξ will also be $\mathcal{O}(1)$ and scale weakly with a in the weak GR regime. Since j_+ and j_0 also scale weakly with a in this regime, we find from (5.52) that

$$\Delta a \propto a^{-3/2}. \quad (5.55)$$

Physically this just reflects the fact that the time spent in the high eccentricity state is proportional to the secular timescale, and that in the weak GR regime this timescale is proportional to $a^{-3/2}$ (equation (5.43)). Indeed, one might have guessed the result (5.55) a priori by noting from (5.15) that for $j \approx j_{\min}$ we have $(da/dt)_{\text{GW}} \propto a^{-3} j_{\min}^{-7}$. Since the time spent at high eccentricity is $t_{\min} \sim j_{\min} t_{\text{sec}}$ (§4.4.4), we get $\Delta a \sim (da/dt)_{\text{GW}} \times t_{\min} \propto a^{-3} j_{\min}^{-6} t_{\text{sec}}$. Plugging in (5.22) for j_{\min} we simply get $\Delta a \propto t_{\text{sec}} \propto a^{-3/2}$.

Moderate GR, Type 2 circulation

Alternatively, for Type 2 circulating orbits in the moderate GR regime, assuming the $(d/a)^{7/2}$ terms dominate in equations (5.28), (5.30) we get from (5.52)

$$\Delta a \approx -\frac{\lambda_2 \xi}{d^{7/2}} \frac{\sqrt{|25\Gamma^2 - 1|}}{10\Gamma} \times a^2. \quad (5.56)$$

Moreover, in this regime it is easy to show that $|x_\alpha| \approx \gamma^{-2} \ll 1$, so we can ignore $|x_\alpha|$ compared to x in (5.53). Also in this case $x_\sigma \approx (-\sigma j_-)/(\gamma j_-) = -\sigma/\gamma \equiv -\kappa$ (equation (4.56)) so that

$$\xi \approx \int_1^{|\kappa|} \frac{dx}{x^6 \sqrt{(x-1)(|\kappa|-1)}}. \quad (5.57)$$

In Figure 5.5 we plot ξ as a function of $|\kappa|$ according to (5.57); in particular we see that $\xi \sim 1$ except for very large $|\kappa| \gg 1$. Away from $a \approx a_{\text{div}}$ we know κ is almost never large

compared to unity, and scales weakly with a (Figures 5.2-5.3), so we can treat ξ as an order-unity constant for a rough analysis. As a result (5.56) predicts a scaling

$$\Delta a \propto a^2. \quad (5.58)$$

We see from (5.58) that, unlike for weak GR (equation (5.55)), here the individual decrements in semimajor axis Δa get *smaller* as the semimajor axis a shrinks.

One can understand the result (5.58) qualitatively as follows. Like for weak GR, at very high e we again have roughly $(da/dt)_{\text{GW}} \propto a^{-3} j_{\text{min}}^{-7}$, and again using (5.22) this is $\propto a^{1/2}$. This time, since the binary spends a large fraction of its secular period in the vicinity of j_{min} , we get a rough estimate of Δa by multiplying $(da/dt)_{\text{GW}}$ not by t_{min} , but by t_{sec} . Using the scaling (5.47) we get $\Delta a \propto a^{1/2} \times a^{3/2} \propto a^2$. Loosely speaking, the factor of ξ in equation (5.56) corrects for the fraction of time that the binary actually spends in the vicinity of j_{min} during each secular cycle.

5.3.5 The time evolution of semimajor axis, $a(t)$

The above results for Δa and t_{sec} allow us to write down implicit formulae for $a(t)$ valid in the aforementioned asymptotic regimes, which can then be approximately solved to get $a(t)$ explicitly. We can also solve for $a(t)$ explicitly in the strong GR regime.

Weak GR

In the weak GR regime, we take t_{sec} from (5.42) and Δa from (5.52) and (5.54); plugging these into (5.1) we find

$$t - t_i \approx - \frac{3c^5 (2p_{\text{min}})^3}{170G^3 m_1 m_2 (m_1 + m_2)} \int_{a_i}^{a(t)} da \frac{|j_+ j_0|}{\xi \sqrt{\Delta}} K \left(\frac{j_{\text{max}}}{\sqrt{\Delta}} \right). \quad (5.59)$$

Unfortunately one cannot go further than this without explicit expressions for j_{max} , Δ etc., which of course requires knowledge of the Γ regime and of whether the orbit librates or circulates. On the other hand, as we have argued above everything in the integrand will typically be $\mathcal{O}(1)$ and vary slowly with a in the weak GR regime. Thus, approximating $Q_1 \equiv \frac{|j_+ j_0|}{\xi \sqrt{\Delta}} K \left(\frac{j_{\text{max}}}{\sqrt{\Delta}} \right) \sim 1$ as a constant, we find a roughly linear decay for $a(t)$:

$$a(t) \approx a_i \left(1 - \frac{t - t_i}{\tau_1} \right), \quad (5.60)$$

where

$$\tau_1 \approx \frac{3c^5 (2p_{\text{min}})^3}{170G^3 m_1 m_2 (m_1 + m_2)} \times Q_1 a_i \quad (5.61)$$

$$= 4.1 \text{ Gyr} \times \left(\frac{p_{\text{min}}}{10^{-3} \text{ AU}} \right)^3 \left(\frac{m}{1.4 M_{\odot}} \right)^{-3} \left(\frac{Q_1}{1} \right) \left(\frac{a_i}{20 \text{ AU}} \right). \quad (5.62)$$

We stress that (5.60)-(5.61) is only a very rough prescription since we have ignored any scaling of Q_1 with a ; nevertheless, a linear decay law for $a(t)$ will turn out to be a good approximation in the weak GR regime (§5.4). In addition, the result (5.60)-(5.61) allows us to understand the origin of a frequently-used formula for the LK-driven merger timescale — see the discussion in §5.5.1.

Moderate GR, Type 2 circulation

For Type 2 circulating orbits in the moderate GR regime, we take t_{sec} from (5.45), and plug this and (5.52) into (5.1) to find

$$t - t_i \approx -\frac{3c^5(2p_{\min})^3}{170G^3m_1m_2(m_1+m_2)} \times \frac{\pi}{2} \int_{a_i}^{a(t)} da \frac{j_{\min} + j_{\max}}{\xi}, \quad (5.63)$$

with ξ given in (5.57). Note that we did *not* take Δa from (5.56): instead we used the more general equation (5.52), which allows us to take advantage of the cancellation of the factors $a^{3/2}|j_+j_0|$ *without* having to assume the dominance of the $(d/a)^{7/2}$ terms in j_+^2, j_0^2 . (Similarly, ξ as given in equation (5.57) does not rely on this assumption). Again without further approximation we can write $j_{\min} + j_{\max} \approx j_{\min}(1 - \kappa)$. If we finally assume that $Q_2 \equiv \pi(1 - \kappa)/2\xi$ is $\sim \mathcal{O}(1)$ and roughly constant, then plugging equation (5.22) for j_{\min} into (5.63) and integrating we get a non-linear decay law for $a(t)$:

$$a(t) \approx a_i \left(1 - \frac{t - t_i}{\tau_2}\right)^2, \quad (5.64)$$

where

$$\tau_2 \approx \frac{3c^5(2p_{\min})^3}{170G^3m_1m_2(m_1+m_2)} \times \pi Q_2 \sqrt{2p_{\min}a_i} \quad (5.65)$$

$$= 180 \text{ Myr} \times \left(\frac{p_{\min}}{0.01 \text{ AU}}\right)^{7/2} \left(\frac{m}{1.4M_{\odot}}\right)^{-3} \left(\frac{Q_2}{1}\right) \left(\frac{a_i}{20 \text{ AU}}\right)^{1/2}. \quad (5.66)$$

We emphasise once again that, perhaps even more so than in the weak regime, the solution (5.64)-(5.65) is only a very approximate one. Indeed in practice, it is often the case that the asymptotic regime considered here is barely established before the binary enters the strong GR regime (see §5.4). Nevertheless, it is good enough for a qualitative understanding and comparison to the weak GR result. Thus, caveats aside, we see that not only is the decay as predicted by (5.64) quadratic rather than linear as it was for weak GR, but the characteristic timescale of the decay is much shorter (compare (5.65) to (5.61)):

$$\tau_2 \sim \left(\frac{2p_{\min}}{a_i}\right)^{1/2} \tau_1 \sim j_{\min}(t_i)\tau_1 \ll \tau_1. \quad (5.67)$$

The fact that the decay of semimajor axis is much more rapid in this regime compared to the weak GR regime will be verified numerically in §5.4.

Strong GR

Of course, once the strong GR regime is reached equation (5.1) is no longer valid, because the binary has decoupled from cluster tides and so no longer undergoes secular eccentricity oscillations (§5.2.2). Instead the evolution of a , e is dictated purely by equations (5.15), (5.16). Supposing the transition to the strong GR regime happens at some reference time t_i , we know from §5.3.1 that for $t > t_i$ the binary conserves its value of $p = a(1 - e) = p_{\min}$, meaning $j = (2p_{\min})^{1/2}a^{-1/2}$. Thus from equation (5.15) we have

$$\frac{da}{dt} \approx -\frac{\lambda_1}{(2p_{\min})^{7/2}}a^{1/2}. \quad (5.68)$$

Integrating this equation we find the following explicit solution for $a(t)$ in the strong GR regime:

$$a(t) \approx a_i \left(1 - \frac{t - t_i}{\tau_{\text{strong}}}\right)^2, \quad (5.69)$$

where

$$\tau_{\text{strong}} \approx \frac{6c^5(2p_{\min})^{7/2}a_i^{1/2}}{170G^3m_1m_2(m_1 + m_2)} = \frac{2}{\pi Q_2} \times \tau_2, \quad (5.70)$$

and τ_2 is given in equation (5.65). Thus the only difference between the evolution of $a(t)$ during Type 2 circulation in the moderate GR regime (equations (5.64), (5.65)) and that during the strong GR regime (equations (5.69), (5.70)) is that the characteristic decay rate is modified by an order-unity constant $2/(\pi Q_2)$.

Finally, we can get a better feeling for how long the strong GR regime lasts before the binary merges if we assume that the transition into the strong GR regime occurs precisely at $a = a_{\text{strong}}$ (equation (5.40)). Setting $a_i = a_{\text{strong}}$ and $t_i = t_{\text{strong}}$ in (5.69)-(5.70) we get a merger at time $t = t_{\text{strong}} + \tau_{\text{strong}}$, where the time spent in the strong GR regime is

$$\begin{aligned} \tau_{\text{strong}} \approx & 96 \text{ Myr} \times \left(\frac{1 + 5\Gamma}{6}\right)^{-1/8} \left(\frac{A^*}{0.5}\right)^{-1/8} \left(\frac{\mathcal{M}}{10^6 M_\odot}\right)^{-1/8} \left(\frac{b}{\text{pc}}\right)^{3/8} \\ & \times \left(\frac{m}{1.4 M_\odot}\right)^{-11/4} \left(\frac{p_{\min}}{10^{-2} \text{ AU}}\right)^{7/2}. \end{aligned} \quad (5.71)$$

Of course this will not be exactly correct, because (i) the transition into the asymptotic strong GR regime (with completely negligible cluster tides) will not happen precisely at $a = a_{\text{strong}}$; and (ii) since $j \propto a^{-1/2}$, the assumption of very high eccentricity used to derive our results will eventually break down when a is sufficiently small, so that the constancy of p_{\min} , i_{\min} will ultimately fail¹¹. Nevertheless, for the mergers we have in mind we will typically have $\tau_{\text{strong}} \ll t_{\text{strong}}$ so these details do not make a substantial difference to the overall merger timescale.

¹¹See §5.4.2 for a detailed numerical example.

5.4 Numerical examples

In this section we will verify the approximate theoretical results derived in §5.3. We do this by direct numerical integration of the DA equations of motion, including both GR precession and GW emission, for various binaries that undergo slow mergers. First we give two Examples with $\Gamma > 1/5$ (§§5.4.1-5.4.2), the first of which exhibits all the hallmark behaviour of a slow merger beginning in the weak GR regime, and the second of which allows us to focus on the late-stage (moderate and strong GR) evolution. We then provide one further Example, this time for a binary with $0 < \Gamma \leq 1/5$ (§5.4.3). Note that we also provide one additional numerical example in the LK limit in Appendix 5.A, when comparing our work with that of Randall & Xianyu (2018).

To be clear, we note that we have run many more numerical experiments of slow mergers than those shown here. We have chosen to present here the minimal number of examples that still capture qualitatively all the possible interesting evolutionary scenarios. (There are of course non-interesting cases, such as binaries that are so tightly bound that the cluster essentially plays no role in their evolution, but we neglect to include them here).

5.4.1 Example 1: $\Gamma = 0.42 > 1/5$. An initially librating orbit in the weak GR regime

In Figure 5.6 we show the result of integrating the DA equations of motion for a binary with constituent masses $m_1 = m_2 = 10M_\odot$, orbiting inside a $\mathcal{M} = 10^6 M_\odot$ spherical Hernquist cluster with $\Gamma = 0.42 > 1/5$. The other initial conditions ($b, r_p/b, r_a/b, a_0, e_0, i_0, \omega_0$) are given in the text at the top of the figure. Note that these parameters have been chosen for pedagogical reasons, as the resulting dynamical evolution exhibits a clean separation of asymptotic behaviours, allowing us to illustrate clearly the different analytic results derived in §5 all in a single representative example. However, in reality they may prove an unrealistic choice, in particular because the initial semimajor axis is large ($a_0 = 250\text{AU}$). In a real stellar cluster one would expect such a wide binary to experience at least one hard encounter within the runtime of the integration (~ 7 Gyr) — see §3.9.3. We will ignore such caveats here, focusing only on illustrating the physics of the ‘idealised’ secular problem considered in this Chapter.

We will begin by describing how Figure 5.6 is organised, before moving on to a discussion of its physical content. In panels (a)-(d) we plot the evolution of binary’s semimajor axis a , eccentricity e , inclination i and pericentre distance p respectively as functions of time with black lines, alongside various critical values shown with coloured lines. Meanwhile panel (e) shows the corresponding run of $\epsilon_{\text{GR}}(t)$ as well as the two critical GR strengths ϵ_{weak} and ϵ_{strong} , defined in equations (4.44) and (4.25) respectively. The reader will also notice three colour-shaded vertical stripes in each panel (a)-(e): one

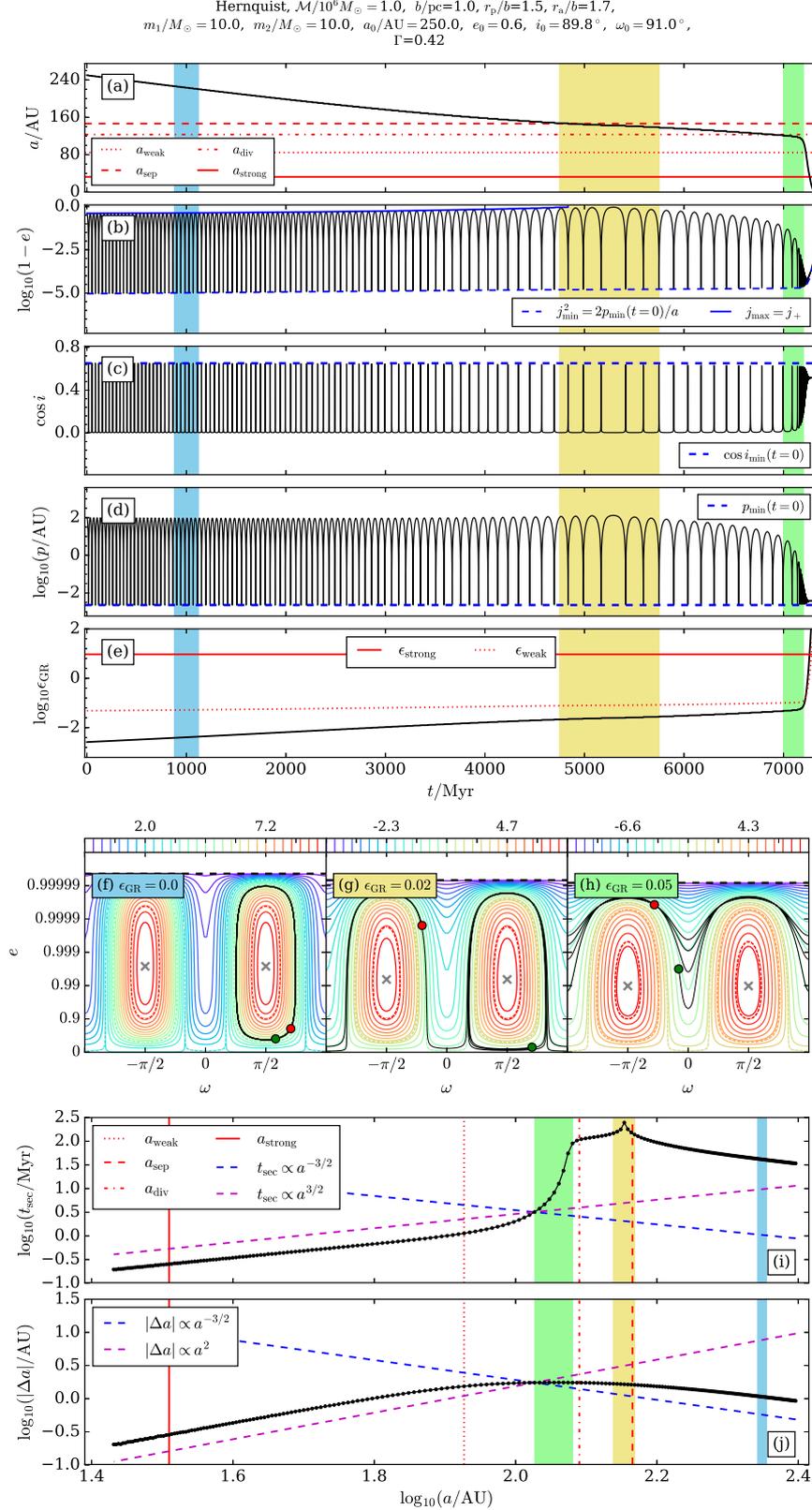


Figure 5.6: *Example 1.* The binary ($m_1 = m_2 = 10M_\odot$, $a_0 = 250\text{AU}$) orbits a $\mathcal{M} = 10^6 M_\odot$ Hernquist cluster. Other parameters are given above panel (a); the resulting Γ value is 0.42. Panels (a)-(e) show the time evolution of a , e , i , p and ϵ_{GR} respectively. Panels (f), (g) and (h) show the phase space evolution during the time intervals shaded in blue, yellow and green respectively in panels (a)-(e). Finally, panels (i) and (j) show with black dots the values of t_{sec} and $|\Delta a|$ as functions of a . See §5.4.1 for a detailed discussion.

blue, one yellow and one green. These three shaded stripes define three representative time intervals, and in panels (f)-(h) we plot the trajectory of the binary in the (ω, e) phase space (familiar from Chapter 4) during those respective time intervals¹². The phase space trajectory in each panel (f)-(h) is shown in black: it starts at the green dot and ends at the red dot. Finally, in panels (i) and (j) we show with black dots the secular timescale t_{sec} and the decay in semimajor axis $|\Delta a|$, respectively, as functions of semimajor axis a . The secular timescale was computed by finding the time elapsed between adjacent eccentricity maxima in panel (b), while Δa was computed by calculating the semimajor axis before (a_{bef}) and after (a_{aft}) each peak, and defining $\Delta a(a_{\text{bef}}) \equiv a_{\text{aft}} - a_{\text{bef}}$. Note that both axes in panels (i) and (j) are logarithmic. The binary evolves from right to left in these two panels, and to aid the reader's comprehension we have again added coloured vertical stripes, this time covering the semimajor axis range that corresponds to a given time interval stripe in panel (a). There are also various critical values and scalings shown with coloured lines in panels (i) and (j), which we will address in a moment.

Now that we are familiar with the structure of Figure 5.6 as a whole, we can discuss the binary's dynamical evolution. Let us first focus on the initial ~ 3000 Myr. From panel (a) we see that at $t = 0$ the binary has $a = 250$ AU, and that for the first ~ 3000 Myr, $a(t)$ exceeds significantly each of the four critical values a_{weak} , a_{sep} , a_{div} , a_{strong} , which were defined in §5.3.2 and which we show with horizontal red lines (see legend). It follows that during this time, the binary resides in the weak GR regime: and indeed, we see from panel (e) that ϵ_{GR} is initially far smaller than ϵ_{weak} (red horizontal dotted line). Moreover, panel (b) shows that the binary undergoes the expected secular eccentricity oscillations (initially on a timescale of ~ 30 Myr), and reaches a very high maximum eccentricity of $1 - e_{\text{max}} \approx 10^{-5}$. Concomitantly there are secular oscillations in inclination i (panel (c)) and pericentre distance p (panel (d)), though as predicted in §5.3.1 the values of $\cos i_{\text{min}}$ and p_{min} reached at the peak of each secular eccentricity cycle are very nearly conserved (see the dashed blue horizontal lines in these panels). Similarly, from panel (b) we see that the maximum eccentricity of the binary is well described by $j_{\text{min}}^2 = 2p_{\text{min}}/a$ (equation (5.22)), while its minimum is well described by $j_{\text{max}} = j_+$. The latter fact implies that the binary is initially on a librating phase space orbit (equation (5.10)), and this is confirmed by panel (f), in which we show the phase space evolution during the time interval denoted by the blue shaded stripe. The high eccentricity behaviour shown in panel (b) of course leads to bursts of GW emission, which results in a decay of the semimajor axis: returning to panel (a) we see that $a(t)$ decays roughly linearly while the binary is in the weak GR regime, which is exactly what we predicted in §5.3.5 (equation (5.60)).

¹²To draw the contours of constant H^* in these panels we took $\Gamma = 0.42$ and took the values of Θ and ϵ_{GR} at the midpoint of the corresponding coloured stripe. Note also that to keep the plots clean we refrain from showing explicit separatrices.

So, we have a binary on a librating orbit in the weak GR regime, whose semimajor axis is slowly decaying with time. As we know from §5.3.2 there are two key things that happen next to such a binary: one is that it enters the moderate GR regime, and the other is that its (ω, e) phase space trajectory crosses the separatrix and becomes circulating (ultimately a Type 2 circulating orbit). We also know from Figure 5.2 that these two occurrences can happen in any order. In this particular case the binary crosses the separatrix first: we see from panel (a) that a intercepts a_{sep} around $t = 4800$ Myr, and from panel (b) that around this time the minimum eccentricity gets very close to zero and then starts to increase and ceases to be well described by $j_{\text{max}} = j_+$. This inference is confirmed in panel (g), in which we see explicitly the evolution from libration to circulation that occurs during the time interval denoted by the yellow shaded stripe. Another interesting observation is that around $t = 5200$ Myr the behaviour of the secular timescale undergoes a qualitative change, as is obvious from inspection of panels (b)-(d): until around 5200 Myr we had t_{sec} increasing with time, whereas afterwards it decreases with time. This is consistent with what we predicted in §5.3.3, and we will say more about it momentarily.

Then, at around $t = 7100$ Myr, the binary’s dynamical evolution changes dramatically: the semimajor axis a approaches a_{weak} and its decay is accelerated, and the secular timescale becomes very short. Furthermore we see that $j_{\text{min}}^2 = 2p_{\text{min}}/a$ is still a good approximation for the maximum eccentricity and, though we do not show it here, the minimum eccentricity at this stage is fairly well-described by $j_{\text{max}} = -\sigma j_-$ (we defer a more careful, ‘zoomed in’ discussion of the late stages of a slow merger to Example 2 — see §5.4.2). These characteristics are the hallmark of Type 2 circulating orbits in the moderate GR regime. To confirm this, we look at panel (h), which shows the phase space evolution during the green striped time interval. We see clearly that the binary is indeed now stuck on a high-eccentricity circulating trajectory, whereafter it soon enters the strong GR regime and then merges.

To conclude our discussion of Example 1 we consider panels (i) and (j), which demonstrate precisely the asymptotic behaviours expected from §5.3.3 and §5.3.4 respectively. The critical semimajor axes shown with red vertical lines in these panels are the same as those shown with horizontal red lines in panel (a). Now, for large $a > a_{\text{sep}}$ we know that the binary is on a librating orbit in the weak GR regime far from the separatrix, and so it is unsurprising to find from panel (i) that $t_{\text{sec}} \propto a^{-3/2}$ (equation (5.43)) and from panel (j) that $|\Delta a| \propto a^{-3/2}$ (equation (5.55)), as we illustrate with blue dashed lines. Next, as a decreases towards a_{sep} the secular timescale briefly undergoes a sharp increase — which again is unsurprising as we know from Chapter 3 that t_{sec} diverges on separatrices — before decreasing again. Similarly, as a decreases towards a_{weak} the value of Δa reaches

a maximum¹³, after which it decreases monotonically with time while a shrinks. Lastly, once a has become significantly smaller than a_{weak} and a_{sep} , we know that the binary has transitioned to the moderate GR regime and is on a Type 2 circulating orbit. The purple dashed lines in panels (i) and (j) confirm the predicted scalings for this regime, namely $t_{\text{sec}} \propto a^{3/2}$ (equation (5.47)) and $|\Delta a| \propto a^2$ (equation (5.58)).

5.4.2 Example 2: $\Gamma = 0.42 > 1/5$. A binary initially in the moderate GR regime

In Figure 5.7 we show the evolution of a lower mass binary ($m_1 = m_2 = 1.4M_{\odot}$), but on the same outer orbit around the same cluster as in Example 1; thus we again have $\Gamma = 0.42$. We choose different initial conditions for the inner orbit, in particular $a_0 = 49$ AU. We see immediately from panel (a) that initially $a < a_{\text{weak}}$, putting the binary just inside the moderate GR regime (see also panel (e)). On the other hand, initially $a > a_{\text{sep}}$ so the phase space trajectory librates. This is fundamentally different to Example 1 since in that case, by the time the binary reached the moderate GR regime, its phase space orbit was already circulating. Thus with Example 2 we will not only be able to focus on the ‘late-time’ behaviour of slow mergers, i.e. their evolution through the moderate and strong GR regimes (as promised in §5.4.1), but also to see a phase space transition from librating to circulating within the moderate GR regime.

Let us note here that the decay of $a(t)$ shown in panel (a) is qualitatively very similar to that in Example 1 (Figure 5.6a). Indeed, this turns out to be the case for all slow mergers: their semimajor axes decay approximately linearly to begin with, then more rapidly (roughly following the quadratic law of equation (5.64)) once they get well into the Type 2 circulating regime, and then like equation (5.69) in the strong GR regime. What we want to focus on here is the behaviour of e , p and i in these latter stages of moderate and strong GR.

From panel (b) we see that while the binary is on a librating phase space orbit ($t \lesssim 4500$ Myr) its maximum eccentricity is well described by $j_{\text{max}} = j_+$ and its secular period increases with time. Once it enters the circulating region its secular period begins to decrease with time. Of course this was also the case in Example 1, though in that case this transition occurred in the weak GR regime. We note however from panel (a) that the semimajor axis decay at this stage, and indeed up until $t \sim 10000$ Myr, remains a gentle, linear one, by virtue of the fact that the binary is on a Type 1 circulating orbit

¹³We should note that the peak in t_{sec} is not centred precisely on a_{sep} , nor is the maximum of $|\Delta a|$ centred precisely on a_{weak} . This is partly because the expressions (5.36)-(5.38) are only approximate, derived in a particular high eccentricity limit and assuming exact conservation of p_{min} , i_{min} . Also, and perhaps more importantly, the transition between weak and moderate GR behaviours around a_{weak} is typically rather blurred, so in truth $a = a_{\text{weak}}$ rarely signifies a precise boundary between these different dynamical regimes, as we see in all numerical examples in this section.

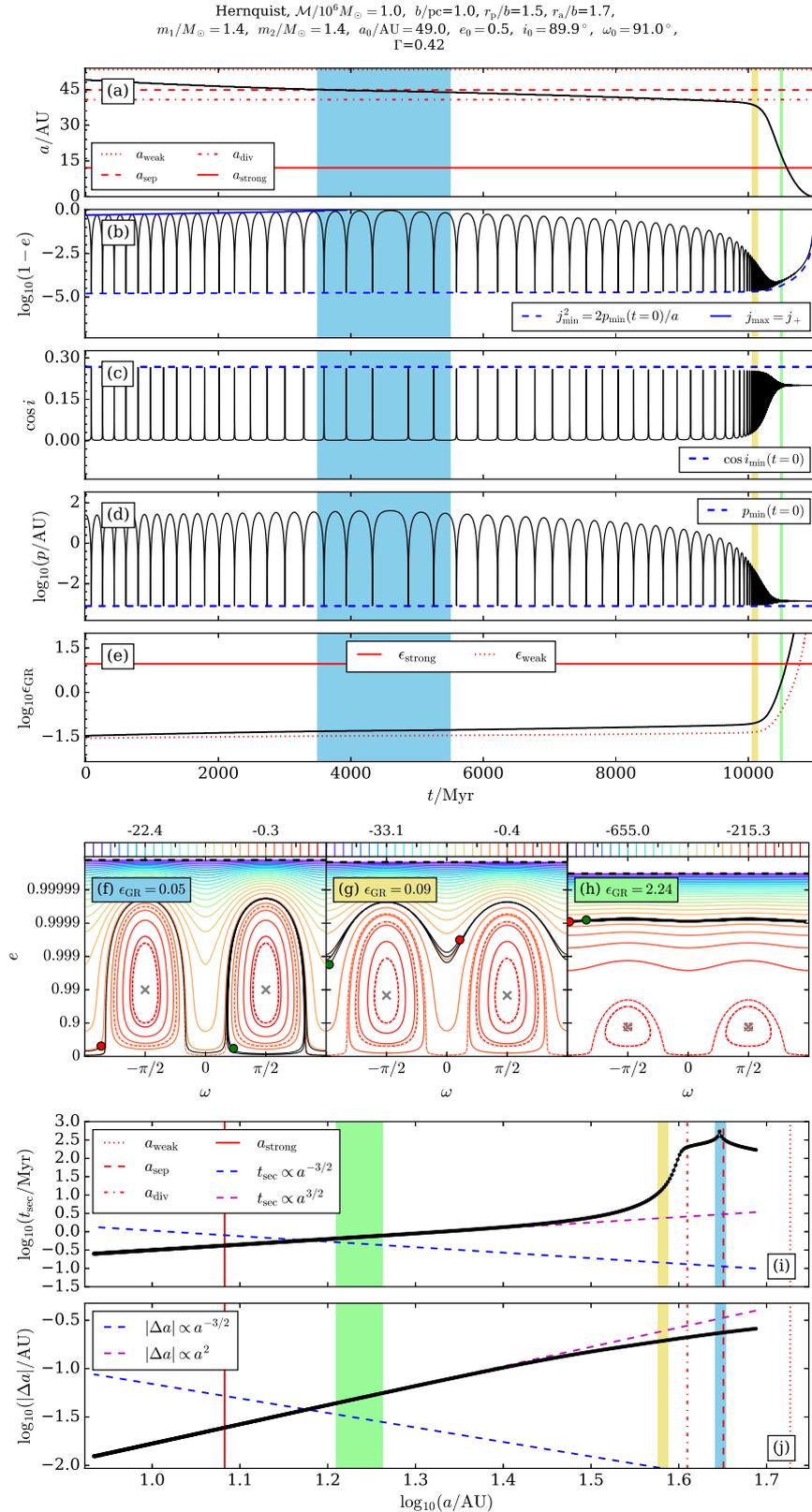


Figure 5.7: *Example 2.* The potential and outer orbit are the same as in Example 1 (Figure 5.6), so that again $\Gamma = 0.42$, but the binary constituent masses and inner orbit initial conditions are different. In this case the binary begins in the moderate GR regime on a librating phase space trajectory.

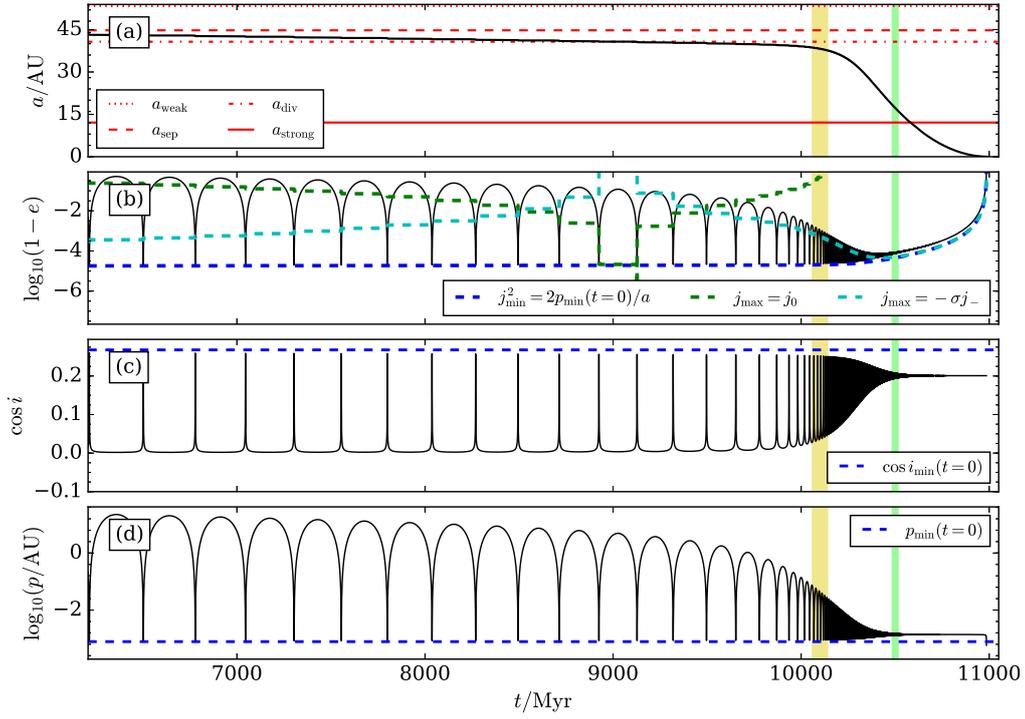


Figure 5.8: Zoomed-in version of panels (a)-(d) from Figure 5.7, focusing on $t > 6100$ Myr.

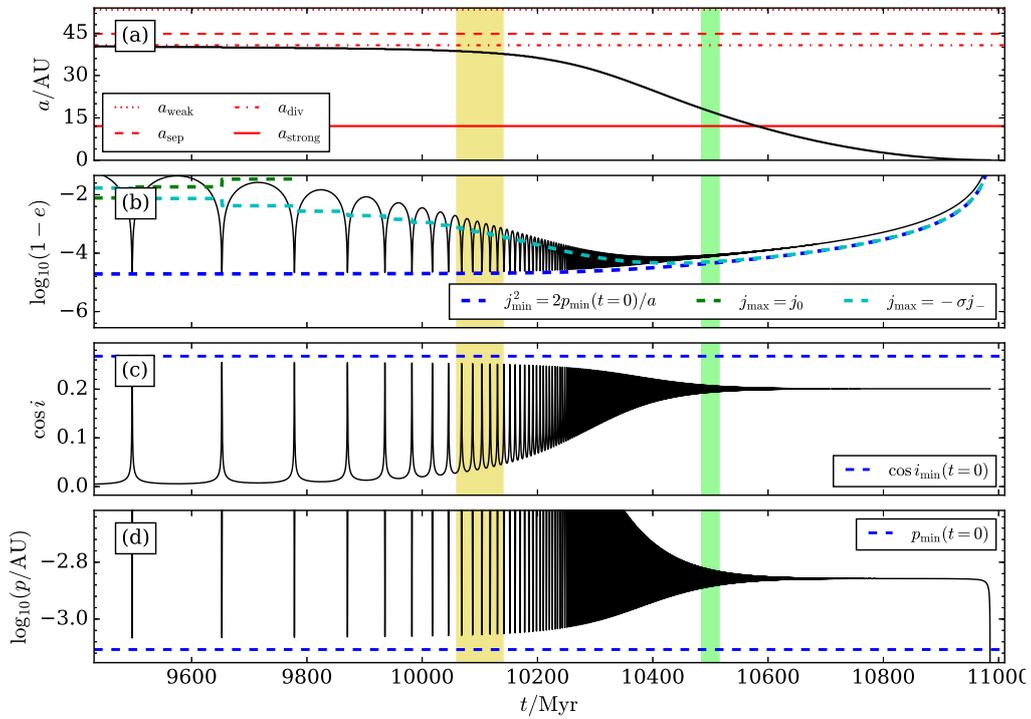


Figure 5.9: Further zoomed-in version of panels (a)-(d) from Figure 5.7, this time showing from $t = 9400$ Myr to merger.

(e.g. panel (f)) despite belonging to the moderate GR regime. The fast, nonlinear decay in semimajor axis does not occur until $t \gtrsim 10000$ Myr, because it is not until then that the circulating orbit can be well-described as Type 2 (panel (g)).

In a moment we will zoom in further on this very late-stage evolution and look at panels (a)-(d) more closely. For now, to conclude the discussion of Figure 5.7 we consider panels (i) and (j). These show that the binary exhibits the expected asymptotic behaviour for $t_{\text{sec}}(a)$ and $|\Delta a(a)|$ at small semimajor axes (i.e. with a significantly smaller than a_{weak} , a_{sep}) — see the purple dashed lines. This is unsurprising since we know from panel (g) that the binary has reached Type 2 circulation by this stage. The expected $t_{\text{sec}} \propto a^{-3/2}$ scaling in panel (i) is also manifest for a larger than a_{sep} , i.e. at the earliest times when the binary is on a librating phase space orbit. On the other hand, the $|\Delta a| \propto a^{-3/2}$ behaviour at large a is never realised, simply because the binary does not begin its life in the weak GR regime.

Let us now turn to Figure 5.8, in which we merely zoom in on panels (a)-(d) of Figure 5.7, focusing on $t > 6100$ Myr. In panel (b) we no longer show the $j_{\text{max}} = j_+$ solid blue line, since we know that for this time range the binary is certainly on a circulating orbit. However, we have added a green dashed line that shows the minimum eccentricity that would be obtained if the binary was on a Type 1 circulating orbit, i.e. with $j_{\text{max}} = j_0$ (equation (5.12)). We have also added a dashed cyan line showing the Type 2 solution $j_{\text{max}} = -\sigma j_-$ (equation (5.13)). We see that until around 8000 Myr the evolution is best described as a Type 1 circulating orbit with $j_{\text{max}} = j_0$. There is then a transitional stage around $t \approx 9000$ Myr wherein neither Type 1 nor Type 2 is a good description (this corresponds to j_0^2 approaching and then crossing zero from above in Figures 5.1 and 5.2). After $t \approx 9400$ Myr the evolution is quite well-described as a Type 2 circulating orbit, $j_{\text{max}} = -\sigma j_-$. On the other hand there is small systematic error in this prediction, which we will explain momentarily.

We finish our discussion of Example 2 by zooming in on this very latest stage of the evolution, $t > 9400$ Myr, which we plot in Figure 5.9. It is clear from this figure that at these late times the conservation of p_{min} (panel (d)) begins to fail, evolving from its value on the blue dashed line ($p_{\text{min}}(t=0) \approx 10^{-3.11}$) towards a slightly larger value. As discussed in §3.3 of Wen (2003), this happens because the role of GR precession at peak eccentricity is gradually becoming more pronounced, blocking the binary from reaching quite the same small p_{min} value from one eccentricity peak to the next¹⁴. The magnitude of the oscillations in p also diminish with time until, after around $t = 10600$ Myr, the binary reaches the strong GR regime and there oscillations are quenched. As predicted in §5.3.1 the value of p itself then remains effectively constant almost all the way to merger,

¹⁴There is a different way to think about this in the context of §5.3.1: as e_{max} is diminished from one secular cycle to the next, the approximation $dp/da \approx 0$ — which was derived in the limit $e \rightarrow 1$ — gets gradually worse.

taking a value¹⁵ $p_{\text{strong}} \approx 10^{-2.86} \approx 1.8p_{\text{min}}(t = 0)$. We see from panel (b) that the resulting underestimate of p_{min} at these late times leads to a slight overestimate of both the maximum and minimum eccentricities (the blue and cyan dashed lines each sit slightly too low in this panel) — hence the systematic error in the prediction $j_{\text{max}} = -\sigma j_-$ which was computed using the p_{min} value from $t = 0$. Finally, we note that the conservation of i_{min} also fails in these latter stages of the evolution (panel (c)), with $\cos i_{\text{min}}$ undergoing a decrease from $\cos i_{\text{min}}(t = 0) = 0.265$ to roughly $\cos i_{\text{strong}} \approx 0.20$. This change — which occurs for the same reason as that in p_{min} , and is also discussed in §3.3 of Wen (2003) — does not make a significant difference to our analysis because the $\cos^2 i_{\text{min}}$, $\sin^2 i_{\text{min}}$ terms in e.g. equations (5.28)-(5.30) are already dominated by the $(d/a)^{7/2}$ terms by this stage.

5.4.3 Example 3: $\Gamma = 0.176 < 1/5$.

In Figure 5.10 we provide one more Example, this time of a $m_1 = m_2 = 10M_{\odot}$ binary orbiting a $\mathcal{M} = 10^7 M_{\odot}$ Hernquist cluster. We take new initial parameters for the inner orbit, as well as a much smaller outer orbit $(r_p/b, r_a/b) = (0.1, 0.5)$, resulting in $\Gamma = 0.176 < 1/5$.

Clearly the binary begins its life squarely in the weak GR regime ($a > a_{\text{weak}}$), and is initially on a librating phase space orbit ($a > a_{\text{sep}}$). It moves into the circulating regime ($a < a_{\text{sep}}$) at around $t = 600$ Myr and then into the moderate regime ($a < a_{\text{weak}}$) at around 720 Myr, and by around 860 Myr it has merged. By now it will come as no surprise that the semimajor axis evolution is qualitatively very similar to that in Example 1 (for $\Gamma > 1/5$), with a roughly linear decrease in $a(t)$ in the weak GR regime and a steeper, nonlinear decay thereafter. Also unsurprisingly, the secular period increases with time while the phase space orbit librates, and decreases with time once it begins to circulate. What *is* notable in this case however, is that unlike in Examples 1 and 2 the phase space transition is very sharp. Indeed, at the beginning of the yellow shaded time interval the binary is on a librating orbit (starting at the green dot in panel (g)); then at around 600 Myr it suddenly joins the family of circulating orbits; and by the end of the yellow interval it is on a very high eccentricity circulating orbit (ending at the red dot). In other words, at some point around the middle of the yellow time interval the binary effectively ‘jumps’ from libration to Type 2 circulation. The reason for this jump is that in this Γ regime, the minimum eccentricity is forced to be larger than that of the saddle point at $\omega = 0$, namely $e_{f,0}$ — see equation (4.27). (Also, it is perhaps not

¹⁵Wen (2003) estimated that $p_{\text{strong}}/p_{\text{min}}$ should lie in the approximate range (1, 3) — see her equation (31) and the surrounding discussion. In fact, the following simple physical argument suggests the value ought to be ≈ 2 (Ford & Rasio 2006). Since GW emission is very poor at dissipating angular momentum, $J \equiv \sqrt{\mu a(1 - e^2)} = \sqrt{\mu p(1 + e)}$ is roughly constant during this phase. When $p = p_{\text{min}}$ we have $e \approx 1$, whereas upon circualisation we have $p = p_{\text{strong}}$ and $e \sim 0$; thus $2p_{\text{min}} \approx p_{\text{strong}}$, i.e. $p_{\text{strong}}/p_{\text{min}} \approx 2$. For all numerical examples presented in this Chapter, $p_{\text{strong}}/p_{\text{min}} \in (1.4, 1.8)$.

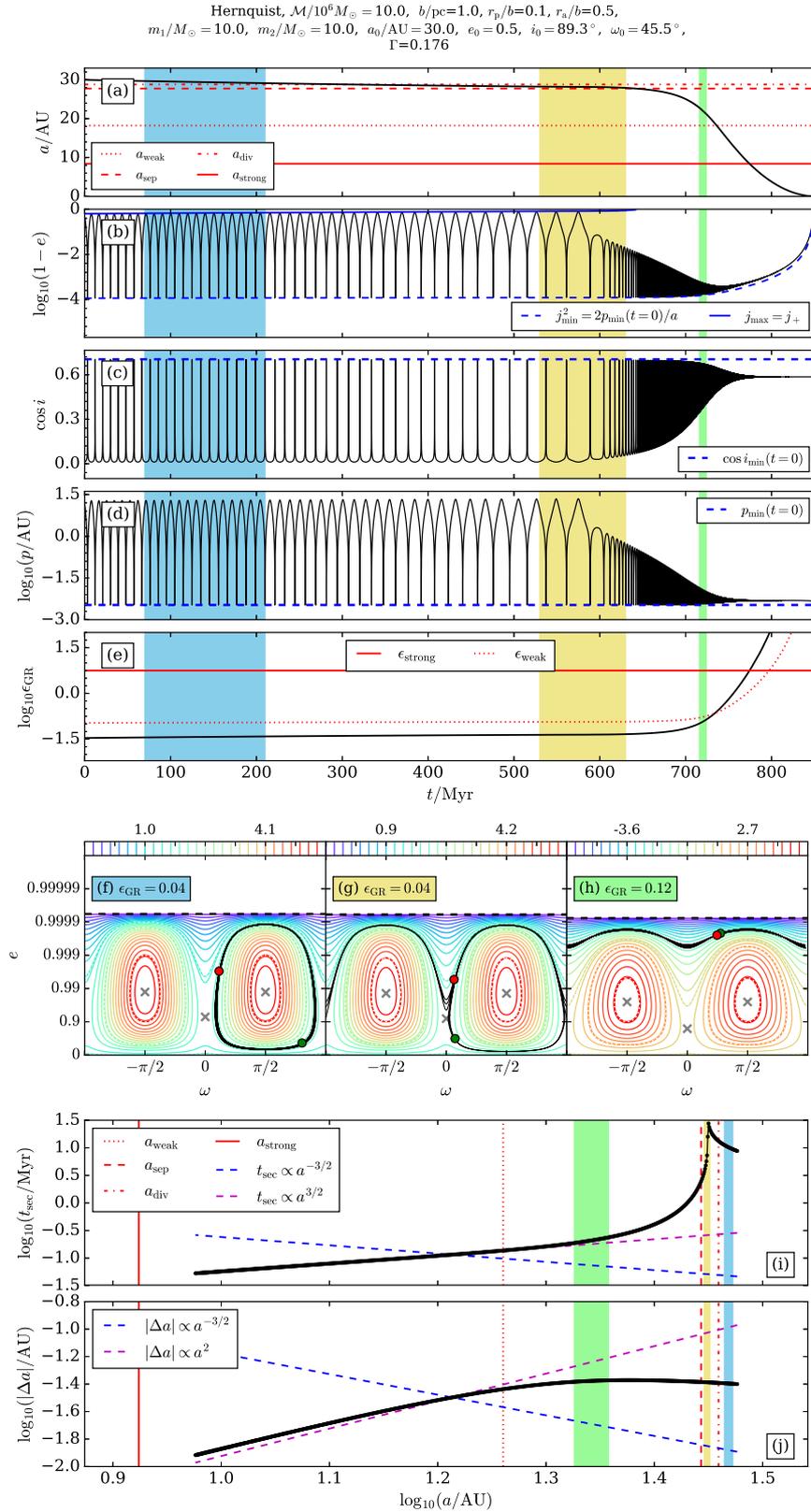


Figure 5.10: Example 3. Slow merger in the $0 < \Gamma \leq 1/5$ regime. The binary orbits an $M = 10^7 M_\odot$ Hernquist cluster. The outer orbit is much smaller than in Examples 1 and 2, giving $\Gamma = 0.176 < 1/5$.

surprising that Type 2 circulation is reached quickly given the lack of Type 1 solutions for $\Gamma < 1/5$ shown in Figure 5.1). One result of this sharp transition is that the change in minimum eccentricity behaviour (panel (b)) is very abrupt around 600 Myr, rather than smooth as it was for $\Gamma > 1/5$. Thereafter the evolution matches the usual Type 2 behaviour as seen for $\Gamma > 1/5$, followed by a merger.

Finally, at the extremes of panels (i) and (j) we see the expected asymptotic behaviour more or less taking shape. However the proper scalings are never fully developed either at large a , simply because at $t = 0$ the binary is too close to the separatrix and to a_{div} for j_+ , j_0 etc. to be considered near-constant.

5.5 Discussion

In this Chapter, we have extended our theory of secular dynamics of binaries in stellar clusters by accounting for the effect of GW emission. We have demonstrated that cluster tides are capable of driving binaries to very high eccentricity, where they can emit GW bursts, shrink in semimajor axis, and ultimately merge. Our results also encompass — and in several aspects extend — the theory of LK-driven binary mergers, which is of course recovered exactly in the limit $\Gamma = 1$.

We have focused throughout this Chapter on understanding the physics of ‘slow mergers’, i.e. those mergers that require many secular periods, but would not have occurred within a Hubble time had the cluster tidal effect not been present. As in Chapters 2-4 we have worked in the DA, test-particle quadrupolar limit, ignoring octupole effects, short timescale fluctuations, stellar flybys and the like. Yet even in this relatively simple setting we have seen that the evolution of a binary from ‘birth’ to merger can be rather complex and is, in general, analytically intractable. Key to making analytical progress was our identification of three key asymptotic regimes: weak GR, Type 2 circulation with moderate GR, and strong GR. We emphasise that the results derived in these regimes are only approximations (often very rough ones), and that the three regimes themselves by no means exhaust the possibilities of slow-merging binary dynamics. Indeed, in practice the boundaries between these asymptotic regimes are blurry and poorly separated, especially in a -space. Nevertheless, they have been sufficient for our purpose, which was to gain analytical — and consequently, physical — insight into an important class of problems that have traditionally been outsourced to a computer.

To conclude this Chapter, we first discuss in §5.5.1 the implications of our results for the calculation of the total merger timescale (which will be important for our population synthesis calculations of merger rates in Chapter 7). Finally in §5.5.2 we discuss our work more broadly in the context of previous studies of LK-driven mergers.

5.5.1 Merger timescale

Secular dynamics of binaries including GW emission is a problem that has been considered many times in the LK context of hierarchical triple systems. Many LK studies that include GW emission are focused upon resulting observable merger rate, i.e. the number of binaries that merge per cubic Gpc per year in the local universe. To compute such a rate — as we will do for cluster tide-driven compact object mergers in Chapter 7 — one needs to know the time it takes for a given binary to merge as a function of its initial conditions. There are effectively two ways to approach this problem. One can either integrate the equations of motion (DA, SA or N-body) directly and read off the merger time from the simulation, or one can seek an approximate (semi-)analytic formula that parameterises the merger time in terms of those initial conditions. The latter approach is obviously much faster when one is dealing with millions or billions of binary initial conditions in a Monte-Carlo population synthesis. In principle one can check the accuracy of such an analytic formula using direct numerical integration for a small number of cases.

For slow mergers, a rather general formula for the merger time t_m is found by setting $t_i = 0$, $t = t_m$ and $a(t_m) = 0$ in equation (5.1):

$$t_m \approx \int_{a(0)}^0 da' \frac{t_{\text{sec}}(a')}{\Delta a(a')}. \quad (5.72)$$

(see equation (57) of Randall & Xianyu (2018)). Of course, as it stands (5.72) is an entirely impractical formula given the complexity of the general analytic expressions for $t_{\text{sec}}(a)$ and $\Delta a(a)$ that must then be integrated over. Instead, the merger time formula usually used in compact object merger calculations in the LK literature is $t_m \approx T_m$ where

$$\begin{aligned} T_m &\equiv T_m^{\text{iso}}(a(0), e_{\text{max}}(0)) \times (1 - e_{\text{max}}^2(0))^{-1/2} \\ &= \frac{3c^5 a(0)^4}{85G^3(m_1 + m_2)m_1 m_2} (1 - e_{\text{max}}(0))^3, \end{aligned} \quad (5.73)$$

and

$$T_m^{\text{iso}}(a, e) = \frac{3c^5 a^4}{85G^3(m_1 + m_2)m_1 m_2} (1 - e)^{7/2}, \quad (5.74)$$

is simply the merger time of an isolated binary with initial semimajor axis a and very high initial eccentricity $e \approx 1$ (Peters 1964). The formula (5.73) is typically justified via the following heuristic argument (Miller & Hamilton 2002; Thompson 2011; Liu & Lai 2018; Randall & Xianyu 2018). First, one assumes that the GW emission is negligible except around $e \approx e_{\text{max}}$, and so the total amount of time that needs to be spent at $e \approx e_{\text{max}}$ before the binary merges is $\approx T_m^{\text{iso}}(e_{\text{max}})$. But the amount of time that the binary actually spends in the vicinity of e_{max} in each secular cycle is $\approx j_{\text{min}} t_{\text{sec}} \equiv (1 - e_{\text{max}}^2)^{1/2} t_{\text{sec}}$; thus the number of secular cycles required until the time spent around e_{max} accumulates to

$T_m^{\text{iso}}(e_{\text{max}})$ is $T_m^{\text{iso}}(e_{\text{max}})/[(1 - e_{\text{max}}^2)^{1/2}t_{\text{sec}}]$. To get the total merger time we multiply this by t_{sec} . Finally, evaluating everything at $t = 0$ we get the formula (5.73).

Of course, this heuristic derivation can be criticised on several levels. First, it makes no distinction between the values of e_{max} , t_{sec} at $t = 0$ and their values at later times, even though we know (§5.3) that both of these quantities vary with a . Second, it does not accurately treat the behaviour of e around e_{max} , instead assuming that e is precisely equal to e_{max} within a discrete time window which lasts for $(1 - e_{\text{max}}^2)^{1/2}t_{\text{sec}}$, and that GW emission is negligible outside that window. We also know from Chapter 4 that the time spent in the vicinity of high eccentricity is not always well approximated by $(1 - e_{\text{max}}^2)^{1/2}t_{\text{sec}}$, especially for large values of σ and/or κ .

Despite these shortcomings, the formula (5.73) actually works well in practice (to within a factor of order unity) when compared to direct numerical integration of the (DA, test particle quadrupole) equations of motion for triple systems (Thompson 2011; Liu & Lai 2018; Randall & Xianyu 2018). To see why this might be the case, we now show that one can actually derive the formula (5.73) in a less hand-waving fashion using the results of this Chapter. First we must assume that the majority of a slow merger is spent in the weak GR regime, and that by ignoring the time spent in the moderate and strong GR regimes we do not impart any major error. Then we expect $a(t)$ to undergo a linear decay following equation (5.60), and thus to merge after time $t_m \approx \tau_1$ (equation (5.61)). (Of course this is precisely equivalent to just evaluating the right hand side of (5.72) in the weak GR limit and assuming j_+ , j_0 are constants — see §5.3.5). Since $p_{\text{min}}(t) \equiv a(t) \times (1 - e_{\text{max}}(t))$ is conserved throughout a slow merger, we can then substitute in equation (5.61) the expression

$$(2p_{\text{min}})^3 = 8(1 - e_{\text{max}}(0))^3 a(0)^3 \approx (1 - e_{\text{max}}^2(0))^3 a(0)^3, \quad (5.75)$$

where in the second equality we assumed $e_{\text{max}}(0) \approx 1$. Plugging (5.75) into (5.61) and comparing the result to (5.73), we find that in this approximation the merger occurs at time

$$t_m \approx T_m \times \frac{Q_1}{2}. \quad (5.76)$$

Thus provided $Q_1 \sim 1$, we recover the standard estimate of the merger timescale (5.73) to within a factor of order unity.

If anything, one might expect that T_m will be an overestimate of the ‘true’ merger time (even if one calculates this ‘true’ time by integrating the DA quadrupolar equations, i.e. ignoring SA effects, octupolar terms, and so on). That is because, as we saw in §5.3.3 and §5.4, the decay of $a(t)$ speeds up substantially once the binary reaches its Type 2 circulating phase in the moderate GR regime (see equation (5.64)). Thus, approximating the entire decay using the weak GR equation (5.60) may seem overly

conservative. It is therefore surprising to note Figure 8 of Thompson (2011) and Figure 5 of Randall & Xianyu (2018), both of which suggest that T_m typically *underestimates* the true (DA) merger time by a factor ~ 2 for compact object binaries in hierarchical triple systems. In future work it might be interesting to understand more deeply the reason for this trend. It may also be profitable to try to use the results of this Chapter to calibrate a merger timescale formula that is more accurate than (5.73) — even a result that was typically in error at the level of only a few tens of percent would be a significant improvement. On the other hand, for practical calculations such a formula may be of limited interest, since the true merger time can be greatly shortened when one includes sub-secular (e.g. ‘singly-averaged’) effects, octupolar terms, and so on — see e.g. Antonini, Murray, et al. (2014) and Grishin et al. (2018).

5.5.2 Relation to studies of LK-driven mergers

As mentioned in §5.5.1, most LK studies ‘solve’ the problem of GW-assisted mergers either by direct numerical integration or by stating and then evaluating the merger time formula (5.73) after calculating e_{\max} from simple theory. There does not exist much in the literature that lies in between these extremes, in which an attempt is made to understand in detail the physics of each stage of the merger or to derive analytic results in specific asymptotic regimes as we have done here. Nevertheless, many of the key individual ideas covered in this Chapter have been considered by other authors, as we now describe.

A central result of §5.3 was the approximate conservation of p_{\min} and i_{\min} during slow mergers: this was ultimately what allowed us to express various important quantities (j_{\min} , t_{sec} , etc.) as functions of a . These conservation laws (as well as their breakdown in the latter stages of a slow merger) seem to have first been described in the LK limit by Wen (2003). The behaviour of p during slow mergers has subsequently been appreciated as an important diagnostic of different regimes; for instance, Antonini has followed the p evolution in order to distinguish between ‘LK dominated’ and ‘GW dominated’ regimes (Antonini & Perets 2012; Antonini, Murray, et al. 2014; Antonini, Toonen, et al. 2017). Some basic scalings of j_{\min} , t_{sec} , etc. with a were also written down by e.g. Miller & Hamilton (2002), Wen (2003), and Thompson (2011), although none of these authors venture beyond the weak GR regime in their analytical efforts, and so did not derive the peculiar results in the moderate GR regime that we have found here (particularly for Type 2 circulating orbits). What is more, nobody has progressed beyond simple scaling relations to write down explicit formulae like we did in §5.3.

One other important achievement of the present Chapter has been to understand the interplay between the time-evolution of key dynamical quantities like (a, e) , and the underlying phase space structure. The fact that a binary initially on a librating phase

space trajectory necessarily transitions into the circulating regime as it shrinks was first mentioned by Blaes et al. (2002) (although they did not note the accompanying qualitative change in t_{sec} behaviour). Of course, since LK theory corresponds to $\Gamma = 1 > 1/5$, no previous authors have noticed the new behaviour that arises in the $0 < \Gamma \leq 1/5$ regime.

The only LK study we know of to have written down a formula for the decay in semimajor axis Δa over one secular cycle is Randall & Xianyu (2018) — see their equation (55). These authors also wrote down an expression (their equation (57)) that is essentially the same as our equation (5.1), pertaining to the slow evolution of a . In addition, Randall & Xianyu (2018) also seem to be the only authors who have previously mentioned the fact that t_{sec} can occasionally decrease as a shrinks, even though every author who has integrated the equations of motion numerically will have encountered this phenomenon. In Appendix 5.A we compare in detail our results with those of Randall & Xianyu (2018). As we show there, Randall & Xianyu (2018)’s analytical theory implicitly assumed weak GR. To our knowledge no authors have derived explicit results in the moderate GR regime until now.

5.6 Summary

In this Chapter we studied the (2.5pN) GW-driven orbital decay and subsequent merger of binary systems which are torqued to high eccentricity by cluster tides on secular timescales. We worked in the DA, quadrupole approximation and included the effect of (1pN) GR precession in our calculations. Our results may be summarised as follows.

- Fundamentally, cluster tides are indeed capable of torquing binaries to sufficiently high eccentricity that they emit bursts of GWs and ultimately merge. Cluster tide-driven eccentricity excitation is therefore a viable mechanism for producing LIGO/Virgo mergers, similar to LK-driven mergers that have been widely explored in the past. Indeed, in the DA, test-particle quadrupole limit, LK mergers are simply a special case of the tide-driven mergers considered here.
- For slow mergers (those that take place over many secular periods) there are two approximate conservation laws that hold as the semimajor axis a decays, namely conservation of the minimum pericentre distance $p_{\text{min}} = a(1 - e_{\text{max}})$ and conservation of the minimum inclination reached i_{min} . The evolution of a decaying binary through phase space can be understood in terms of these conserved quantities.
- We uncovered three basic asymptotic regimes in which analytic progress was possible provided the binary’s phase space trajectory was not close to any separatrices. These were weak GR, moderate GR (wherein the binary undergoes ‘Type 2’ circulation

in phase space), and strong GR. These different regimes each exhibit their own characteristic behaviour of secular timescale $t_{\text{sec}}(a)$, decay in semimajor axis per cycle $\Delta a(a)$, and consequently time evolution of semimajor axis $a(t)$, as we confirmed numerically.

- We derived a formula for the merger timescale that has been much used in LK theory, and gave a more detailed justification for it than those that have been offered previously.

The insights from this Chapter will inform future studies of LK-driven and cluster tide-driven binary mergers. In addition, some of the ideas developed here may provide qualitative insight into other problems where secular forcing, apsidal precession and short-range dissipation compete over the dynamics of a two-body system. One such problem might be the LK-driven formation of short period binaries and hot Jupiters in triple systems, wherein the dissipation is not due to GW emission but instead due to internal fluid tidal friction in the star(s) and/or planet (Fabrycky & Tremaine 2007; Anderson, Storch, et al. 2016; Vick et al. 2019).

Appendices

5.A Relation to Randall & Xianyu (2018)

We have mentioned several times the work of Randall & Xianyu (2018), whose paper largely inspired the present Chapter. These authors are among the few who have attempted to gain a more analytical understanding of LK-driven slow mergers (and indeed it is from their paper that we have taken the terminology ‘slow merger’). In particular, to our knowledge Randall & Xianyu (2018) were the first authors to (i) calculate Δa explicitly, and (ii) mention explicitly the decrease in t_{sec} as the binary shrinks and offer an explanation thereof. On the other hand, we feel that both (i) and (ii) as presented in Randall & Xianyu (2018) could be improved. Here we explain how our calculations differ from those of Randall & Xianyu (2018).

5.A.1 Calculation of Δa

Randall & Xianyu (2018) begin their calculation of Δa by writing down their equation (55), the first two lines of which are essentially identical to our equation (5.48) when we evaluate the final bracket at $e = e_{\text{max}}$. One is then faced with the computation of an integral, $\Delta a \propto \int dt(1 - e^2(t))^{-7/2}$, over one complete secular cycle. To compute this integral in §5.3.4 we changed variables from $t \rightarrow j \in (j_{\text{min}}, j_{\text{max}})$ and hence wrote down equation (5.49). On the other hand, Randall & Xianyu (2018) chose to compute the integral by first approximating $e(t)$ as a quadratic in time (see their equation (53)). In particular, using our notation and letting the maximum eccentricity occur at $t = 0$ without loss of generality, their equation (52) reads

$$e(t) = e_{\text{max}} + \frac{1}{2} \left(\frac{d^2 e}{dt^2} \right)_{t=0} t^2. \quad (5.77)$$

Randall & Xianyu (2018) then plugged this into $\int dt(1 - e^2(t))^{-7/2}$ and integrated over $t \in (-\infty, \infty)$ to get Δa . The result is their second equation (55), which in our notation and evaluating at $e_{\text{max}} \approx 1$ reads

$$\Delta a_{\text{RX18}} \approx -\frac{544G^3 m_1 m_2 (m_1 + m_2)}{9c^5 a^3 j_{\text{min}}^6} \times \left| \frac{d^2 e}{dt^2} \right|_{t=0}^{-1/2}. \quad (5.78)$$

Randall & Xianyu (2018) then evaluate $\ddot{e}|_{t=0}$ using their equation (53).

There are a few issues with this calculation, however. First, the assumption that $e(t)$ is quadratic for small t is equivalent to the assumption that $j(t)$ is quadratic for small t . We know from §4.4.4 (see also §4.C) that for this quadratic ansatz to be a good one, it must be the case that $\epsilon_{\text{GR}} \ll \epsilon_{\text{weak}}$ (i.e. weak GR) and $\sigma \ll 1$. This is

certainly true for many binaries of interest, but it is not always true, and in particular it is not true for the example given in Figure 3 of Randall & Xianyu (2018) which begins in the moderate GR regime (see our Figure 5.11). Second, the equation that Randall & Xianyu (2018) quote for $\ddot{e}|_{t=0}$ — namely their equation (53) — is not a good one in general. To see this we compute it directly by differentiating $e = (1 - j^2)^{1/2}$ twice, using equations (4.9)-(4.10) and demanding that at $t = 0$, $j = j_{\min}$, $dj/dt = 0$ and $\omega = \pm\pi/2$. Without any approximations we find

$$\left(\frac{d^2e}{dt^2}\right)_{t=0} = -\frac{60\Gamma C}{L} \frac{(j_{\min}^2 - \Theta)e_{\max}}{j_{\min}} \left(\frac{d\omega}{dt}\right)_{t=0}. \quad (5.79)$$

For this to coincide with equation (53) of Randall & Xianyu (2018) in the LK limit, one must have $(j_{\min}^2 - \Theta) \approx j_{\min}^2 = (1 - e_{\max}^2)$, which is only true if¹⁶

$$j_{\min}^2 \gg \Theta, \quad \text{i.e.} \quad \cos^2 i_{\min} \ll 1. \quad (5.80)$$

As we have seen in §5.4 this is certainly not true in general. In fact, we know from §5.2.1 that if a binary starts in the weak GR regime it has $j_{\min}^2 \sim \Theta$ and this relation holds all the way into the moderate GR regime and beyond. The condition (5.80) *does* happen to be true in the specific numerical example shown in Randall & Xianyu (2018)'s Figure 3, but that is because their example is rather a peculiar one and begins in the moderate GR regime, which contradicts the weak GR assumption they have implicitly made when approximating $e(t)$ as a quadratic.

Nevertheless, let us follow the Randall & Xianyu (2018) method and compute Δa via equations (5.78) and (5.79), evaluating $d\omega/dt$ at maximum eccentricity using (4.9), and compare the result to our equation (5.52). (We will *not* make the assumption (5.80)). Using $j_{\min}^2 = 2p_{\min}/a$ and after some algebra we arrive at

$$\frac{\Delta a_{\text{RX18}}}{\Delta a} = \frac{8}{15} \frac{\sqrt{|25\Gamma^2 - 1|} |j_+ j_0|}{10\Gamma \xi} \frac{j_{\min}^2}{\sqrt{j_{\min}^2 - \Theta}} \left(10\Gamma\Theta - (1 + 5\Gamma)j_{\min}^4 + \frac{\epsilon_{\text{GR}} j_{\min}}{6}\right)^{-1/2}. \quad (5.81)$$

We can make sense of (5.81) by evaluating the right hand side in the weak and moderate GR regimes.

In the weak GR regime we have $\epsilon_{\text{GR}} j_{\min} \ll \Theta$ (equation (5.7)). If we also assume $j_{\min}^4 \ll \Theta$ (see 4.C) and ignore the a -dependent terms in (5.28), (5.30), we get

$$\frac{\Delta a_{\text{RX18}}}{\Delta a} \approx \frac{8}{15\xi}, \quad (5.82)$$

¹⁶It is easy to show that the condition (5.80) is also required to make equation (54) of Randall & Xianyu (2018) agree with our (4.9) at maximum eccentricity.

with ξ given in (5.54). Note that for $\sigma = 0$ we get exactly $\Delta a_{\text{RX18}}/\Delta a = 1$. Thus we recover precisely the formula that Randall & Xianyu (2018) derived when we make the approximations that they (implicitly) did, namely weak GR and $\sigma \ll 1$.

In the moderate GR regime we assume that the ϵ_{GR} term dominates the final bracket in (5.81), and that the a -dependent terms dominate equations (5.28), (5.30). With these assumptions we get

$$\begin{aligned} \frac{\Delta a_{\text{RX18}}}{\Delta a} &= \frac{8\sqrt{60\Gamma}}{15} \left(\frac{d}{a}\right)^{7/2} \frac{1}{\xi} \frac{j_{\text{min}}^2}{\sqrt{j_{\text{min}}^2 - \Theta}} \sqrt{\frac{6}{\epsilon_{\text{GR}} j_{\text{min}}}} \\ &\sim \frac{\sqrt{10\Gamma}}{\xi} \left(\frac{d}{a}\right)^{7/2} \frac{j_{\text{min}}}{\epsilon_{\text{GR}}}, \end{aligned} \quad (5.83)$$

with ξ given in equation (5.57) (and plotted in Figure 5.5). All three fractions in (5.83) are $\mathcal{O}(1)$ or (often significantly) larger. Thus we typically have $\Delta a_{\text{RX18}}/\Delta a \gg 1$, meaning that the method of Randall & Xianyu (2018) drastically *overestimates* the value of Δa in this regime.

5.A.2 Decrease in t_{sec} with time

As we mentioned in §5.1, the decrease in t_{sec} with time during a slow merger was already pointed out in the LK ($\Gamma = 1$) limit by Randall & Xianyu (2018) in their §3.1, when discussing their Figure 3. In Figure 5.11 we reproduce exactly this numerical example from Randall & Xianyu (2018), using the same layout of panels as we did in §5.4. We see that the binary initially sits in the moderate GR regime on a circulating orbit, and that the secular timescale does indeed decrease as the binary shrinks, before it finally merges at around $t = 7000$ yr.

When interpreting the scaling of $t_{\text{sec}}(a)$ physically, Randall & Xianyu (2018) noted that smaller a (larger ϵ_{GR}) promotes faster apsidal precession, which is obviously true. However, they then claimed that this faster precession directly leads to a shorter secular period, as well as the corresponding increase in maximum eccentricity with time and decrease in minimum eccentricity with time. This interpretation is not quite right (and also does not explain why in the librating regime t_{sec} increases with shrinking a). In reality, in the weak-to-moderate regime, GR precession is unimportant except during an extremely high eccentricity episode, and typically that extreme eccentricity episode is over very quickly. In other words, for most phase space orbits the second (GR) term in equation (56) of Randall & Xianyu (2018) is completely negligible during the majority of the evolution, so barely affects t_{sec} . What GR precession *does* do, when coupled with GW emission, is to alter the phase space morphology, and to periodically nudge the binary onto a new phase space trajectory every time it reaches high eccentricity (note how closely

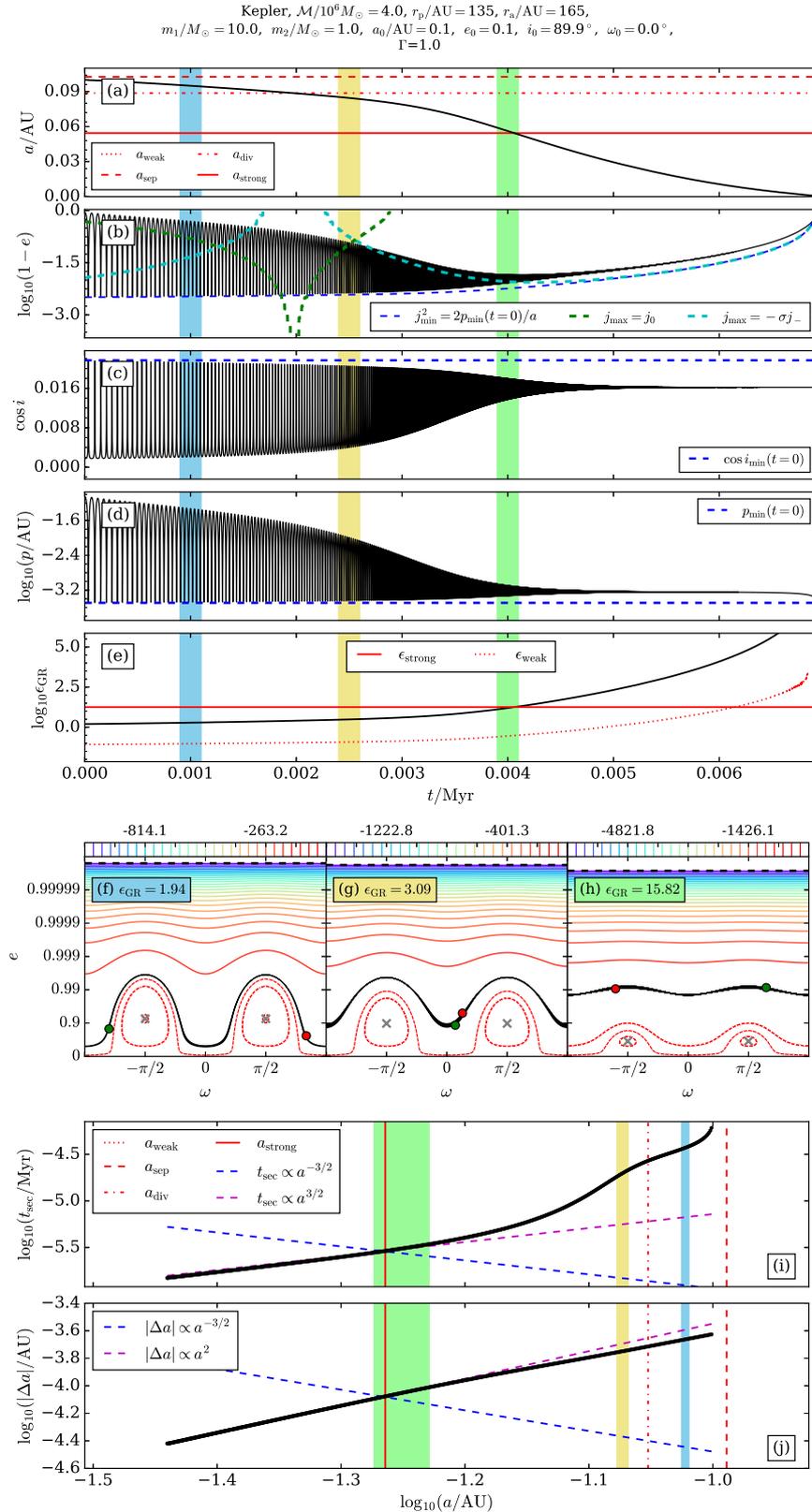


Figure 5.11: Reproducing Figure 3 of Randall & Xianyu (2018). In this case a binary with $m_1 = m_2 = 10M_\odot$ orbits a supermassive black hole (i.e. Kepler potential) of mass $\mathcal{M} = 4 \times 10^6 M_\odot$. The outer orbit has semimajor axis $a_g = (r_a + r_p)/2 = 150$ AU and eccentricity $e_g = (r_a - r_p)/(r_a + r_p) = 0.1$. See §5.A for discussion.

the contours of H^* are bunched at these high eccentricities). As a is decreased and ϵ_{GR} is increased, the binary gets pushed further away from the separatrix between librating and circulating orbits, towards the Type 2 circulating region where (5.46) applies. As long as this process continues the binary gets pushed to higher minimum eccentricity (smaller and smaller j_{max}), even though its e_{max} is getting smaller. On average the binary spends more and more time at ‘high’ (say $e \gtrsim 0.9$) eccentricities where cluster tide-driven secular evolution is fast. We emphasise that this last statement is true regardless of GR precession: indeed, the binary typically does not care about GR precession directly when, say, $e = 0.9$. In fact, whereas Randall & Xianyu (2018) attributed the evolution of t_{sec} and $e_{\text{min/max}}$ to fast GR-aided ω precession, both of these phenomena are present even in the weak GR regime where apsidal precession is nearly always negligible — see Figure 5.6.

6

Short-timescale fluctuations

Contents

| | | |
|------------|---|------------|
| 6.1 | Introduction | 214 |
| 6.2 | Dynamical framework | 217 |
| 6.3 | Short-timescale fluctuations and high eccentricity behaviour | 218 |
| 6.3.1 | Two examples in the Hernquist potential | 218 |
| 6.3.2 | Two examples in the Plummer potential | 222 |
| 6.3.3 | Analysis of fluctuating behaviour | 224 |
| 6.3.4 | Characteristic amplitude of δj fluctuations | 228 |
| 6.4 | The effect of GR precession | 231 |
| 6.4.1 | Two examples in the Hernquist potential | 231 |
| 6.4.2 | Two examples in the Plummer potential | 233 |
| 6.4.3 | An example in the Lidov-Kozai limit | 234 |
| 6.4.4 | Physical interpretation and quantitative analysis | 237 |
| 6.5 | Discussion | 242 |
| 6.5.1 | Astrophysical relevance of RPSD | 242 |
| 6.5.2 | Breakdown of the SA approximation | 243 |
| 6.5.3 | Relation to LK literature | 245 |
| 6.6 | Summary | 246 |
| | Appendices | 247 |
| 6.A | Singly-averaged equations of motion | 247 |
| 6.B | Fluctuating Hamiltonian | 249 |
| 6.C | Phase dependence of fluctuating behaviour | 250 |

6.1 Introduction

As discussed at length at the beginning of this thesis, decades of effort have gone into understanding in detail the secular evolution of tidally perturbed binaries. Much recent research in this field has been motivated by the discovery of exotic astrophysical systems — black hole mergers, blue stragglers, hot Jupiters, and so on — which were likely formed

through high eccentricity migration. In each case, the body that perturbs the ‘binary’ might simply be a tertiary point mass, it might be the Galactic tide, or it might be the mean field potential of a stellar cluster in which the binary orbits. Regardless of the particular system under consideration, the same two key questions almost always arise: (i) under what circumstances can a binary reach extremely high eccentricity on an astrophysically relevant timescale? and (ii) how does the binary behave when it reaches such extreme eccentricities?

Question (i) is straightforward to answer using secular theories, the simplest of which involve truncating the perturbing potential at quadrupole order, taking the ‘test particle’ approximation and then ‘double-averaging’ (DA). In Chapters 2-4 we developed the most comprehensive such theory to date, capable of describing the secular evolution of any binary perturbed by any fixed axisymmetric potential (in the test particle, quadruple limit). In this theory, the binary’s maximum eccentricity e_{\max} can be calculated (semi-)analytically as a function of initial conditions — see e.g. §4.3.4. As a result one can easily determine the region of parameter space that leads to extremely high e_{\max} . Crucially, one always finds that e_{\max} is limited by the initial relative inclination i_0 between the inner and outer orbits:

$$e_{\max} \leq e_{\text{lim}} \equiv (1 - \Theta)^{1/2}, \quad \text{where} \quad \Theta \equiv (1 - e_0^2) \cos^2 i_0, \quad (6.1)$$

and e_0 is the initial eccentricity. Hence, as we have seen repeatedly thus far, a necessary (but not always sufficient) part of the answer to question (i) is that $\Theta \ll 1$.

However, DA theories often do not provide an accurate answer to question (ii). That is because DA theory ignores a component of the torque that fluctuates on the timescale of the outer orbit, and washes out to zero upon averaging over that timescale. This becomes problematic at extremely high eccentricity, when the relative changes in the binary’s (very small) angular momentum due to this fluctuating torque can become $\mathcal{O}(1)$ — as a result, the DA theory fails to capture the dynamics in detail. A more accurate description is provided by the singly-averaged (SA) theory, i.e. the theory governed by the SA Hamiltonian that we wrote down in §2.3.1. As its name suggests, the SA theory only involves one average, namely over the binary’s inner Keplerian orbit, and hence captures fully the fluctuations in the orbital elements on the outer orbital timescale. In particular, these short-timescale fluctuations (sometimes called ‘SA fluctuations’) can increase a binary’s maximum eccentricity beyond e_{\max} . Because of this they can be of great significance when predicting LK-driven merger rates of black hole (BH) or neutron star (NS) binaries, blue straggler formation rates, white dwarf collision rates, and so on (e.g. Antonini, Murray, et al. 2014; Grishin et al. 2018).

Faced with this assessment, one might decide simply to abandon DA theory altogether and only work with the SA equations of motion. Alternatively one might choose to

forego all averaging, and instead to integrate the ‘N-body’ equations of motion directly (like in §3.7). There are three main objections to these approaches. First, numerical integration of the SA or N-body equations is prohibitively expensive if one wants to evolve millions of binary initial conditions, as we shall in Chapter 7. Second, SA and N-body approaches necessarily demand more initial data, inflating the parameter space. Third, whatever one gains through brute-force computation, one also often sacrifices in terms of analytical and physical insight. Instead, our approach will be to understand the SA problem at high eccentricity in an approximate analytical fashion, guided by the DA theory and by numerical integrations where appropriate.

Short-timescale fluctuations¹ have been a major focus of LK studies (e.g. Ivanov et al. 2005; Katz & Dong 2012; Antonini & Perets 2012; Bode & Wegg 2014; Antonini, Murray, et al. 2014; Antognini et al. 2014; Luo et al. 2016; Grishin et al. 2018; Lei et al. 2018; Lei 2019). However, they have not yet been explored in the more general case of binaries perturbed by arbitrary axisymmetric potentials, e.g. that of a host globular cluster, nuclear cluster or galaxy in which the binary orbits. The primary aim of this Chapter is to extend the well-known LK results to these more general systems, and to expand upon them. Finally, as we know from §4.C, even in the DA approximation the quantities j , ω , Ω can exhibit $\mathcal{O}(1)$ fractional changes on timescales much shorter than t_{sec} at high eccentricity. We will see that if this short timescale is comparable to T_ϕ , and GR precession is switched on, then we get a new effect which we dub ‘relativistic phase space diffusion’ (RPSD). In RPSD the coincidence of very high eccentricity secular behaviour, short-timescale fluctuations and GR precession conspire to shift the binary to a qualitatively new phase space orbit. This happens even in the LK limit, but has not been shown before.

This Chapter is structured as follows. In §6.2 we very briefly recap some central results from earlier Chapters and establish some notation. In §6.3 we provide several numerical examples that illustrate the phenomenology of short-timescale fluctuations when GR is *not* included, particularly with regard to high eccentricity behaviour. We then proceed to explain the observed behaviour quantitatively, and derive an approximate expression for the magnitude of angular momentum fluctuations at high e . In §6.4 we switch on GR precession and give several numerical examples of systems exhibiting phase space diffusion. We then analyse this phenomenon more quantitatively and offer a physical explanation for it. In §6.5 we consider the astrophysical importance of the new effect we have uncovered, and discuss our results in the context of the existing LK literature. We summarise in §6.6.

¹Throughout this Chapter, we use the term ‘short-timescale fluctuations’ in the sense defined in the previous paragraphs — i.e. those fluctuations that arise in SA theory when compared with DA theory. We do not consider other fluctuations that might occur on short timescales, e.g. flyby encounters with passing stars.

6.2 Dynamical framework

Let us briefly recap some results from earlier Chapters that we will need here. The perturbing Hamiltonian that encodes the effect of quadrupole-order cluster tides upon a binary was derived in Chapter 2 and is given in equation (2.9). When we average (2.9) over the shortest timescale in the problem, i.e. over the inner orbit's mean anomaly M , the resulting 'singly-averaged' (SA) Hamiltonian is²

$$H_{1,\text{SA}} = \frac{1}{2} \sum_{\alpha\beta} \Phi_{\alpha\beta} \langle r_\alpha r_\beta \rangle_M. \quad (6.2)$$

The averages $\langle r_\alpha r_\beta \rangle_M$ are given explicitly in terms of orbital elements in §2.A. Since we have eliminated the angle M the conjugate action $L = \sqrt{\mu a}$ is conserved, and so the binary's semi-major axis a is constant. The singly-averaged Hamiltonian $H_{1,\text{SA}}$ is a function of the variables J, J_z, ω, Ω and the time t through the time-dependent coefficients $\Phi_{\alpha\beta}(\mathbf{R}_g(t))$. The equations of motion that result from differentiation of (6.2) are called the SA equations, and are given explicitly in Appendix 6.A (equations (6.33)-(6.36)). When we further average the Hamiltonian (6.2) over the outer orbital motion $\mathbf{R}_g(t)$ (i.e. over the orbital torus or annulus — see Chapter 2), the resulting doubly-averaged (DA) perturbing Hamiltonian is

$$\begin{aligned} H_{1,\text{DA}} &= \frac{1}{2} \sum_{\alpha\beta} \bar{\Phi}_{\alpha\beta} \langle r_\alpha r_\beta \rangle_M = \frac{1}{2} \bar{\Phi}_{xx} \langle x^2 + y^2 \rangle_M + \frac{1}{2} \bar{\Phi}_{zz} \langle z^2 \rangle_M \\ &= C \times \frac{1}{L^2 J^2} \left[(J^2 - 3\Gamma J_z^2)(5L^2 - 3J^2) - 15\Gamma(J^2 - J_z^2)(L^2 - J^2) \cos 2\omega \right]. \end{aligned} \quad (6.3)$$

The DA equations of motion arising from (6.3) are given in equations (4.9)-(4.11). Whether we use SA or DA theory, if we wish to include GR precession then we must add to our Hamiltonian a term

$$H_{\text{GR}} = -\frac{CL\epsilon_{\text{GR}}}{J^2}. \quad (6.4)$$

This obviously affects the equation of motion for $d\omega/dt$ by adding an extra term $CL\epsilon_{\text{GR}}/J^2$ (equation (4.1)). We will ignore GW emission throughout this Chapter.

Following this, we know from Chapter 4 that in the DA approximation, i.e. under the dynamics prescribed by the total DA Hamiltonian $H_{\text{DA}} \equiv H_{1,\text{DA}} + H_{\text{GR}}$, there are two independent integrals of motion. For the purposes of this Chapter, it will be most useful to take these to be $j_z \equiv J_z/L$ and D , defined as (see equation (4.16)):

$$D \equiv e^2 \left(1 + \frac{10\Gamma}{1-5\Gamma} \sin^2 i \sin^2 \omega \right) - \frac{\epsilon_{\text{GR}}}{3(1-5\Gamma)\sqrt{1-e^2}}. \quad (6.5)$$

²In Chapters 2 and 3 we referred to $H_{1,\text{SA}}$ as $\langle H_1 \rangle_M$. We will stick to the $H_{1,\text{SA}}$ notation in what follows. An analagous statement holds for the upcoming $H_{1,\text{DA}}$ and H_{GR} .

In the SA approximation, i.e. under the dynamics prescribed by the total SA Hamiltonian $H_{\text{SA}} \equiv H_{1,\text{SA}} + H_{\text{GR}}$, the quantities $j_z(t)$, $D(t)$ are not precise integrals of motion: instead they are time-dependent (the latter is found by evaluating the right hand side of (6.5) using the SA orbital elements). However, they can typically be treated as constants when averaged over several outer orbital periods, which means they are still useful for characterising the ‘underlying’ secular behaviour. Indeed, a binary that is described by SA dynamics with near-constant j_z and D will simply fluctuate around the underlying DA solution corresponding to one of the level curves in the characteristic (ω, e) phase space — see Chapter 3. This is what allows us to consider short-timescale fluctuations as a perturbation on top of a dominant secular effect (Ivanov et al. 2005; Luo et al. 2016). On the other hand, in §6.4 we will see that for non-zero ϵ_{GR} and very high eccentricities, this perturbative assumption can break down — even the time-averaged value of D can change dramatically and abruptly, reflective of phase space diffusion.

6.3 Short-timescale fluctuations and high eccentricity behaviour

In this section we provide several numerical examples that demonstrate the phenomenology of short-timescale fluctuations, and discuss them qualitatively, first in the Hernquist potential (§6.3.1) and then in the Plummer potential (§6.3.2). Crucially, for simplicity and in order to cleanly separate certain physical effects, *we do not include GR precession in any of these examples* (GR precession will be added in §6.4). In §6.3.3 we provide a quantitative analysis of the behaviour we have uncovered. Finally in §6.3.4 we derive an approximate expression for the magnitude of angular momentum fluctuations at highest eccentricity.

We note that in all numerical examples presented in this Chapter, we checked that the numerical method for integrating the SA equations had converged. In particular we integrated the outer orbit using a timestep Δt which was small enough that further shortening of the timestep did not affect the SA results (typically $T_\phi/\Delta t \sim 10^3$).

6.3.1 Two examples in the Hernquist potential

(A) Fiducial example in the Hernquist potential, $i_0 = 90.3^\circ$

In Figure 6.1A we give an example of a binary that undergoes significant short-timescale fluctuations at high eccentricity. This figure is very rich in information and exhibits several interesting features that we wish to explore throughout the Chapter. We will also see several other figures with this or similar structure. It is therefore worth describing the structure of Figure 6.1A in detail.

At the very top of the figure (top line of text) we provide the values of 6 input parameters $(\Phi, \mathcal{M}, b, r_p, r_a, \phi_0)$ that define the perturbing potential as well as the

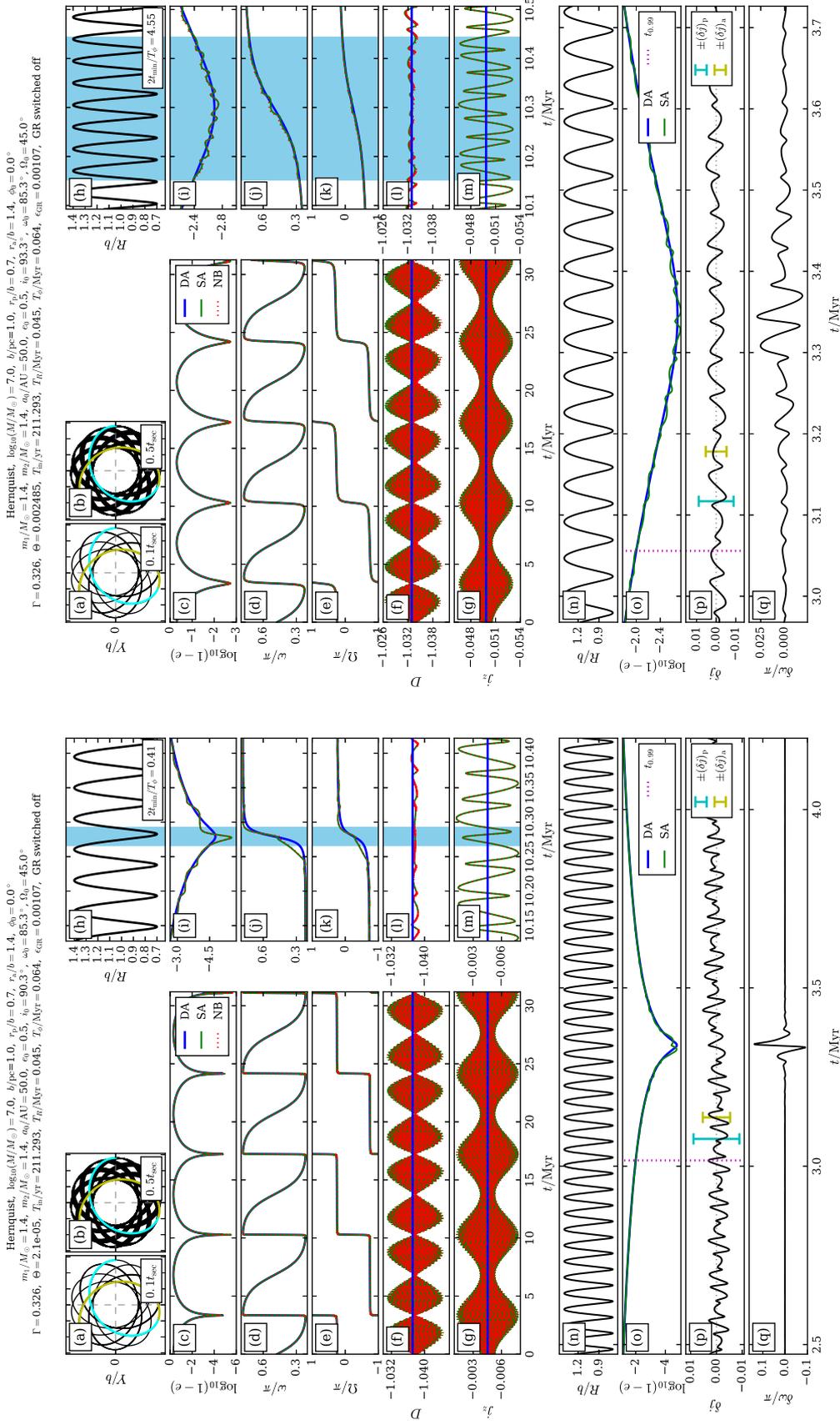


Figure 6.1: (A) Example of a binary that undergoes significant short-timescale fluctuations at very high eccentricity. The details of the plot are explained fully in §6.3.1. (B) We change i_0 to 93.3° . As a result ϵ_{max} is decreased and $2t_{\text{min}}$ becomes significantly larger than T_{ϕ} .

outer orbit’s initial conditions. In this case we are considering a binary in a Hernquist potential (see equation (2.52)) with total mass $\mathcal{M} = 10^7 M_\odot$ and scale radius $b = 1\text{pc}$. Since the potential is spherical, the outer orbit can be specified with three numbers: its pericentre $r_p = 0.7b$, its apocentre $r_a = 1.4b$, and its initial azimuthal coordinate $\phi_0 = 0^\circ$, where ϕ is the outer orbit’s azimuthal angle relative to the X axis (see Chapter 2). Unless otherwise stated, in this Chapter we always initiate the outer orbit at $t = 0$ from $(R, \phi) = (r_a, \phi_0)$ with $\dot{\phi} > 0$.

In the second line of text we list 7 input parameters $(m_1, m_2, a_0, e_0, i_0, \omega_0, \Omega_0)$ that concern the binary’s inner orbit; the subscript ‘0’ denotes initial values. In this example we are considering a NS-NS binary ($m_1 = m_2 = 1.4M_\odot$) with initial semimajor axis $a_0 = 50\text{AU}$. Note also that the initial inclination i_0 is chosen close to 90° , which is necessary to achieve very large eccentricities ($\Theta \ll 1$).

In the third and final line of text at the top of the figure, we list 6 important quantities that follow from the choices of 13 input parameters above: Γ, Θ , the inner orbital period $T_{\text{in}} = 2\pi\sqrt{\mu/a^3}$, the outer orbit’s radial period T_R , its azimuthal period T_ϕ , and the strength of GR precession ϵ_{GR} . In this instance we have a Γ value of 0.326 and $\Theta = 2.1 \times 10^{-5}$, which allows e_{max} to become extremely high. Lastly, we emphasise that we have artificially switched off GR precession in this example. Thus, even though we provide the value of ϵ_{GR} for the binary in question, in practice this quantity is set to zero in the equations of motion and in evaluating D . This choice is also indicated in the third line of text. GR precession will be incorporated in §6.4, allowing for direct comparison with the results of this section.

Now we move on to the figure proper. In panels (a) and (b) we display the trajectory of the outer orbit through the (X, Y) plane, integrated using GALPY (Bovy 2015). In both panels we show the trajectory from $t = 0$ to $t = T_R$ (cyan line), and from $t = T_R$ to $t = 2T_R$ (yellow line). In black we show the entire trajectory traced up to time $t = 0.1t_{\text{sec}}$ (panel (a)) and $t = 0.5t_{\text{sec}}$ (panel (b)), where t_{sec} is the period of secular oscillations, computed using equation (3.33).

In total we integrated the outer orbit $\mathbf{R}_g(t)$ until $t = 4.5t_{\text{sec}}$. We then fed the resulting $\Phi_{\alpha\beta}(\mathbf{R}_g(t))$ timeseries into the SA equations of motion (6.33)-(6.36) and integrated them numerically. In panels (c), (d) and (e) we compare the numerical integrations of the SA equations of motion for e, ω, Ω (green curves) against the prediction of DA theory (blue curves). We also show the results of direct ‘N-body’ integration³ (red dotted curves). In panel (c) we see that the binary reaches extremely high eccentricity, with the DA result $1 - e_{\text{DA}}$ reaching values as low as $\sim 10^{-5}$. In this panel we already see that the maximum

³The ‘N-body’ integration involves directly integrating the exact binary equations of motion in the presence of the full, smooth time-dependent field $\Phi(\mathbf{R}_g(t))$ using REBOUND (Rein & Liu 2012).

eccentricity reached in the SA approximation can be rather different from the DA value, and changes from one eccentricity peak to the next. Also striking from panels (d) and (e) are the step-like jumps in ω and Ω that occur near maximum eccentricity in both SA and DA integrations — note that these are just what we expect from our investigation in §4.C.

In panels (f) and (g) we show the evolution of the quantities D (equation (6.5)) and j_z (equation (3.7)). We emphasise again that although the definition (6.5) of D involves ϵ_{GR} (and we report a finite ϵ_{GR} value in the third line at the top of the figure), we do not include GR at all in this example, i.e. we set $\epsilon_{\text{GR}} \equiv 0$ for the purposes of calculating D .

In the DA approximation, D and j_z are integrals of motion — hence the blue DA result is simply a straight horizontal line. We see that the SA result more or less oscillates around the constant DA value in both cases, with an envelope that has period t_{sec} for j_z and $t_{\text{sec}}/2$ for D . In panel (f) there is a small offset between D_{DA} and the mean value of D_{SA} , which is due to an initial phase offset of the outer orbit (Luo et al. 2016; Grishin et al. 2018). We also notice a characteristic behaviour which is that fluctuations in D are minimised around the eccentricity peak, while fluctuations in j_z are maximised there.

In the right hand column, in panels (i)-(m) we simply reproduce panels (c)-(g), except we zoom in on the sharp eccentricity peak at around 10.3 Myr. At the top of this column we have panel (h), which shows the outer orbital radius $R(t)$ during this high-eccentricity episode. In addition, in each panel (h)-(m) we shade in light blue the region

$$|t - t(j_{\text{min}})| < t_{\text{min}}. \quad (6.6)$$

Here $t(j_{\text{min}})$ is the time corresponding to the minimum of j_{DA} (i.e. peak DA eccentricity), namely when $j_{\text{DA}} = j_{\text{min}}$, and t_{min} is the time taken for j_{DA} to change from j_{min} to $\sqrt{2}j_{\text{min}}$ — see equation (4.35). In panel (h) we also indicate the value of the ratio $2t_{\text{min}}/T_\phi$, which will turn out to be very important when we switch on GR in §6.4. In this particular case we see that $2t_{\text{min}}/T_\phi = 0.41$, so that most of the interesting (very highest eccentricity) behaviour happens on a timescale shorter than an outer azimuthal period.

From panels (h)-(m) we see that the N-body and SA integrations agree very well even at extreme eccentricities, giving us confidence that the SA approximation is a good one⁴. However, the SA prediction differs markedly from the DA prediction at very high eccentricity. In particular, from panel (i) we see that $1 - e_{\text{DA}}$ becomes $\approx 10^{-4.6}$ at its minimum, while $1 - e_{\text{SA}}$ reaches a significantly smaller value still, $\approx 10^{-5.2}$. Panels (j) and (k) reveal that the large jumps in ω and Ω both happen on a timescale $\sim 2t_{\text{min}}$. In panels (l) and (m) we show how the integrals of motion j_z and D fluctuate around the maximum eccentricity.

⁴The SA approximation can itself break down at extremely high eccentricity — see §6.5.2 for discussion.

In panels (n)-(q) we again show the timeseries of R and $1 - e$ around the second eccentricity peak (although over a wider timespan), as well as the differences

$$\delta j(t) \equiv j_{\text{SA}}(t) - j_{\text{DA}}(t), \quad \delta \omega(t) \equiv \omega_{\text{SA}}(t) - \omega_{\text{DA}}(t), \quad (6.7)$$

between the results of SA and DA integration. The vertical dotted magenta line in panels (o) and (p) corresponds to $t = t_{0.99}$, which is the time when e_{DA} first reaches 0.99. From panel (p) we see that δj fluctuates in a complex but near-periodic manner, with period $\sim 5T_R$. The fluctuations themselves are not perfectly centred around zero; before the eccentricity peak, $\langle \delta j \rangle$ is slightly negative, whereas afterwards it is slightly positive. The blue and cyan bars in panel (p) correspond to simple approximations to the amplitude of δj at peak eccentricity — see §6.3.4. Meanwhile, from panel (q) we see that the fluctuation $\delta \omega$ is negligible until the very highest eccentricities are reached, where there is a sharp pulse before it decays to zero again. This pulse is approximately antisymmetric in time around $t = t(j_{\text{min}})$. The pulse episode lasts for $\sim 2T_R$.

(B) Changing the initial inclination to $i_0 = 93.3^\circ$

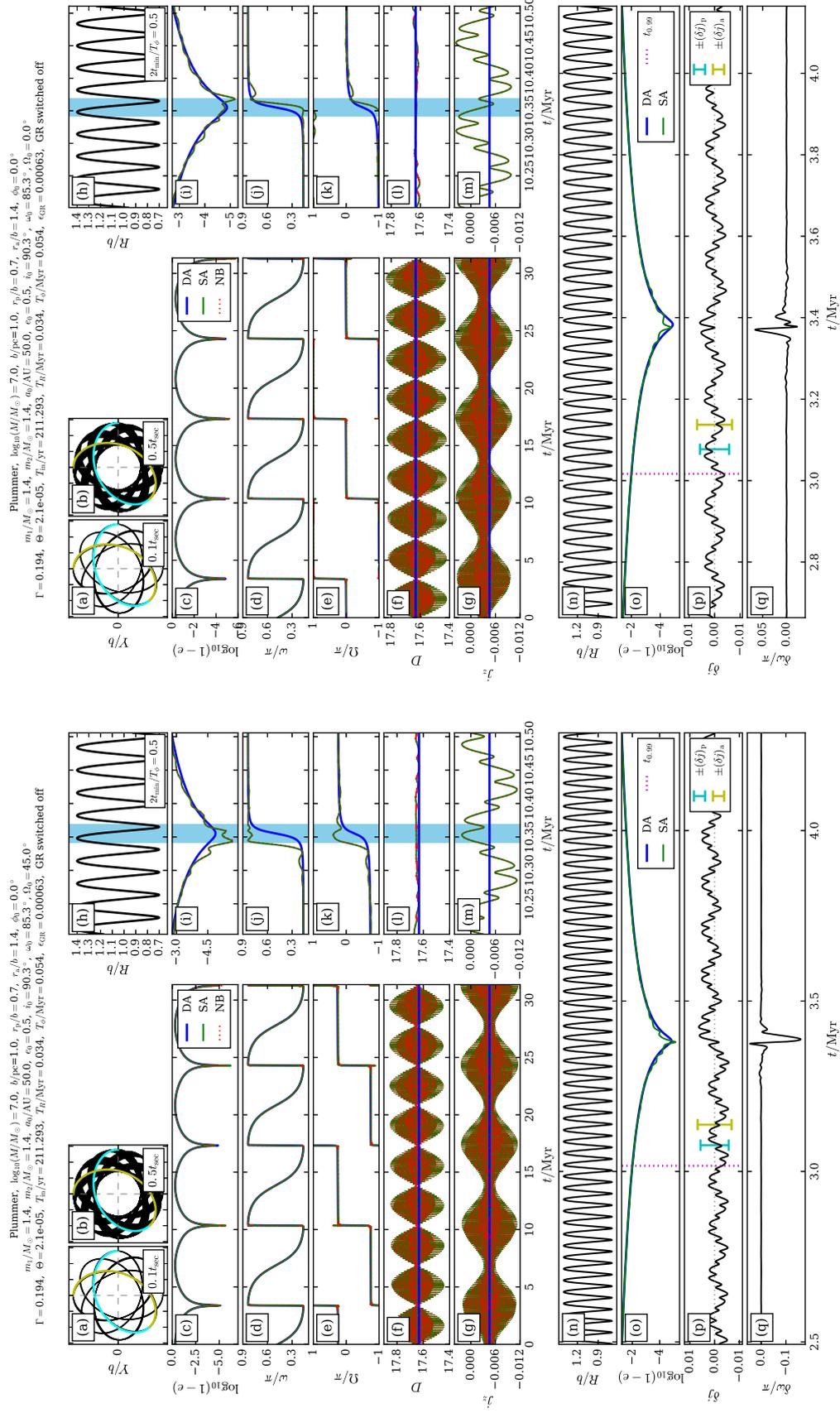
In Figure 6.1B we run the same calculation as in Figure 6.1A, except we take $i_0 = 93.3^\circ$ rather than 90.3° . The main effect of this choice is to reduce the maximum eccentricity significantly, so that $1 - e_{\text{max}} \approx 10^{-2.7}$. As a result, DA evolution near maximum eccentricity is slower than in Figure 6.1A, while the outer orbit is unchanged; hence we find $2t_{\text{min}}/T_\phi = 4.55$ in this case. The qualitative fluctuating behaviour of D , j_z and δj is quite similar between the two figures, although in Figure 6.1B many more fluctuations fit into the ‘blue stripe’ surrounding maximum eccentricity. The fluctuations $\delta \omega$ are very different: the brief ‘pulse’ that lasted for only $\sim 2T_R$ in Figure 6.1Aq has been replaced with a much broader signal with a significantly smaller amplitude.

6.3.2 Two examples in the Plummer potential

The phenomenology reported above is rather characteristic of binaries orbiting in cusped potentials, and will be analysed more quantitatively in §6.4.4. First, however, we perform the same calculation except this time with a cored potential, namely the Plummer sphere.

(A) Fiducial example in the Plummer potential

In Figure 6.2A we provide an example of a binary exhibiting short-timescale fluctuations in a cored potential. All input parameters are the same as in Figure 6.1A except we change the potential from Hernquist to Plummer (2.51). The change of potential means that we now have $\Gamma = 0.194 < 1/5$, which leads to a rather large D_{DA} value of around 17.6 (equation (6.5)). Once again, SA and N-body results agree very nicely even at extreme eccentricities.



Noteworthy in this case is the morphology of the fluctuations of $j_{z,\text{SA}}$ near highest eccentricity, around the DA value $j_{z,\text{DA}} = -\sqrt{\Theta} = -0.0045$ (panel (g)). In this case, small fluctuations of amplitude on the timescale $\sim T_R$ are superimposed upon a larger ‘carrier signal’ oscillation, which has amplitude ~ 0.006 and its own period $\sim 4T_R$. A similar morphology is exhibited by the δj timeseries (panel (p)). Again the fluctuations $\delta\omega$ (panel (q)) are negligible until the very highest eccentricities are reached, where there is a sharp, negative pulse of maximum amplitude $\sim 0.2\pi$, that lasts for $\sim 2T_R$ in total before decaying back to zero.

(B) Shifting the initial nodal angle Ω_0 by 45°

We give one final example in Figure 6.2B. The initial conditions this time are exactly the same as in Figure 6.2A, except we make an initial phase shift by putting $\Omega_0 = 0^\circ$ rather than $\Omega_0 = 45^\circ$. This choice of phase feeds into the SA equations of motion (6.33)-(6.36), but does not affect the DA equations. In the right hand column we zoom in on the same secular eccentricity peak as in Figure 6.2A. We see that the SA results at very high e are rather different between the two figures. In particular, e_{SA} reaches a different maximum value, and does so at a different time. Also, deviations of ω , Ω from the DA result are not so severe in Figure 6.2B as in Figure 6.2A. Overall though, the fluctuating behaviour in Figures 6.2A and 6.2B is qualitatively similar despite the applied phase shift. In fact this is a rather generic result, holding for various choices of initial conditions, potentials, etc. — phase shifts produce a quantitatively different, but qualitatively similar, fluctuating behaviour at high eccentricity when GR precession is switched off. However, phase dependence can become much more important when GR precession is included, as we explore in §6.C.

6.3.3 Analysis of fluctuating behaviour

We now wish to explain more quantitatively the behaviour that we observed in the previous two subsections. Precisely, we aim to understand the characteristic behaviours of δj and $\delta\omega$ around the eccentricity peak and to understand the envelopes of D and j_z fluctuations over secular timescales. That is what we will accomplish in this subsection. For more practical purposes we also wish to derive and test an approximate expression for the amplitude of fluctuations δj around the peak eccentricity — that will be done in §6.3.4.

Notation and approximation

To achieve these aims we must first introduce a clean, precise notation to describe fluctuations. Let us define the vector $\mathbf{w} \equiv [\omega, J, \Omega, J_z]$. Then the ‘SA solution’

$$\mathbf{w}_{\text{SA}}(t) \equiv [\omega_{\text{SA}}(t), J_{\text{SA}}(t), \Omega_{\text{SA}}(t), J_{z,\text{SA}}(t)], \quad (6.8)$$

is found by self-consistently integrating the SA equations (6.33)-(6.36), which are the Hamilton equations resulting from $H_{\text{SA}}(\omega_{\text{SA}}, J_{\text{SA}}, \Omega_{\text{SA}}, J_{z,\text{SA}}, t) \equiv H_{\text{SA}}(\mathbf{w}_{\text{SA}}, t)$. Meanwhile the ‘DA solution’

$$\mathbf{w}_{\text{DA}}(t) \equiv [\omega_{\text{DA}}(t), J_{\text{DA}}(t), \Omega_{\text{DA}}(t), J_{z,\text{DA}}], \quad (6.9)$$

is found by self-consistently integrating the DA equations of motion (e.g. (4.9)-(4.11)), which are the Hamilton equations for $H_{\text{DA}}(\omega_{\text{DA}}, J_{\text{DA}}, J_{z,\text{DA}}) \equiv H_{\text{DA}}(\mathbf{w}_{\text{DA}})$. Consistent with equation (6.7) we formally define:

$$\delta\mathbf{w}(t) \equiv \mathbf{w}_{\text{SA}}(t) - \mathbf{w}_{\text{DA}}(t). \quad (6.10)$$

Next, we will also find it useful to define a ‘fluctuating Hamiltonian’:

$$\Delta H(\mathbf{w}, t) \equiv H_{\text{SA}}(\mathbf{w}, t) - H_{\text{DA}}(\mathbf{w}), \quad (6.11)$$

which is written explicitly in Appendix 6.B — see equation 6.37. Using the fluctuating Hamiltonian we can *define* four new quantities

$$\Delta\mathbf{w}(\mathbf{w}, t) \equiv [\Delta\omega(\mathbf{w}, t), \Delta J(\mathbf{w}, t), \Delta\Omega(\mathbf{w}, t), \Delta J_z(\mathbf{w}, t)] \quad (6.12)$$

as the solution to the equations of motion

$$\frac{d\Delta\mathbf{w}(\mathbf{w}, t)}{dt} \equiv \left[\frac{\partial\Delta H(\mathbf{w}, t)}{\partial J}, -\frac{\partial\Delta H(\mathbf{w}, t)}{\partial\omega}, \frac{\partial\Delta H(\mathbf{w}, t)}{\partial J_z}, -\frac{\partial\Delta H(\mathbf{w}, t)}{\partial\Omega} \right]. \quad (6.13)$$

As an example, the partial derivative $\partial\Delta H/\partial\omega$ is given explicitly for spherical potentials in equation (6.42).

Now, we must bear in mind that in general,

$$\delta\mathbf{w}(t) \neq \Delta\mathbf{w}(\mathbf{w}_{\text{SA}}, t) \neq \Delta\mathbf{w}(\mathbf{w}_{\text{DA}}, t). \quad (6.14)$$

In other words, we cannot simply feed a numerical result $\mathbf{w}(t) = \mathbf{w}_{\text{SA}}(t)$ into equation (6.13) and expect to reproduce the ‘SA minus DA’ solution (6.10) by integrating forwards in time; for example:

$$\delta\omega(t) \equiv \omega_{\text{SA}}(t) - \omega_{\text{DA}}(t) \neq \Delta\omega(\mathbf{w}_{\text{SA}}, t) \equiv \int_0^t dt' \frac{\partial\Delta H(\mathbf{w}_{\text{SA}}(t'), t')}{\partial J_{\text{SA}}}. \quad (6.15)$$

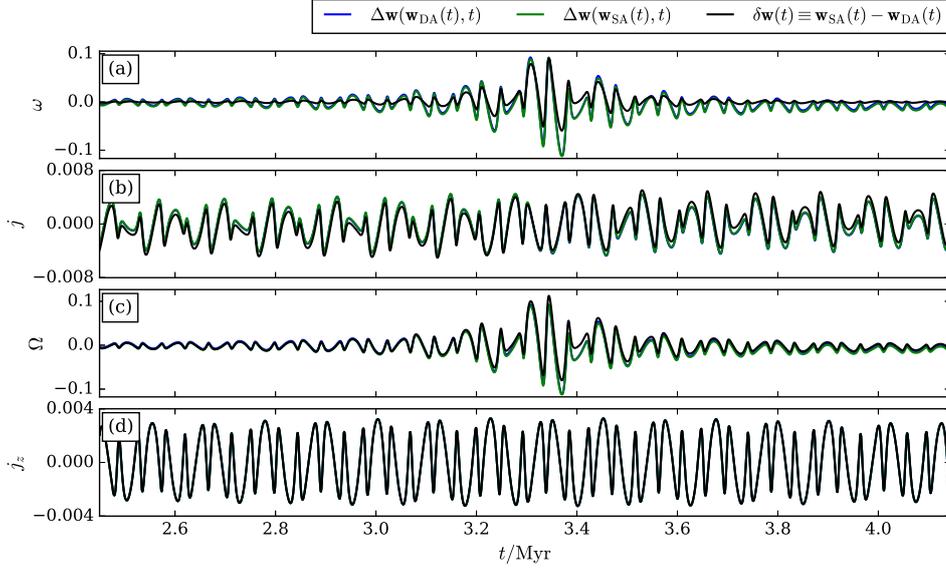


Figure 6.3: Justifying the approximation $\delta\mathbf{w}(t) \approx \Delta\mathbf{w}(\mathbf{w}_{\text{SA}}, t) \approx \Delta\mathbf{w}(\mathbf{w}_{\text{DA}}, t)$. The data used here is taken from the first secular period of the evolution shown in Figure 6.1B.

Similarly, $\delta J \neq \Delta J$ and $\delta\Omega \neq \Delta\Omega$. The exception is J_z , for which there is no DA evolution, so that $\delta J_z(t) = \Delta J_z(\mathbf{w}_{\text{SA}}, t)$. Figure 6.3 demonstrates this explicitly. The data in this figure is taken from the first secular period of evolution of the system shown in Figure 6.1B. In blue we show $\Delta\mathbf{w}(\mathbf{w}_{\text{DA}}(t), t)$, i.e. the result of feeding the DA solution to equations (6.13) and integrating forwards in time. Analogously, in green we plot $\Delta\mathbf{w}(\mathbf{w}_{\text{SA}}(t), t)$. Finally, overplotted in black we show $\delta\mathbf{w}(t) = \mathbf{w}_{\text{SA}}(t) - \mathbf{w}_{\text{DA}}(t)$. Panels (a)-(d) correspond to fluctuations in ω, j, Ω, j_z respectively. We see in each panel that the blue and green curves agree nicely, so that we may approximate $\Delta\mathbf{w}(\mathbf{w}_{\text{DA}}(t), t) \approx \Delta\mathbf{w}(\mathbf{w}_{\text{SA}}(t), t)$. Furthermore, panel (d) shows that for j_z fluctuations, all three curves agree very precisely, as expected. On the other hand, in panels (a)-(c) there is a visible offset between the blue/green curves and the black curves. This is particularly true for fluctuations in ω (panel (a)). However, the differences are small enough — particularly for fluctuations in j (panel (b)), which are what we care about most — to justify the approximation

$$\delta\mathbf{w}(t) \approx \Delta\mathbf{w}(\mathbf{w}_{\text{SA}}(t), t) \approx \Delta\mathbf{w}(\mathbf{w}_{\text{DA}}(t), t). \quad (6.16)$$

This approximate equality in (6.16) is useful because we have explicit expressions for $d\Delta\mathbf{w}/dt$, via equations (6.13) and the Hamiltonian (6.37).

Scaling of fluctuations at high eccentricity in spherical potentials

Having established the approximation (6.16), we can use the equations of motion (6.13) to gain a better understanding of the behaviour of fluctuating quantities in Figures 6.1-6.2. To

do this we take derivatives of (6.41) (which is valid for spherical potentials only) and then take the high eccentricity limit $L^2 \gg J^2 \rightarrow J_z^2$. As a result we find the following scalings:

$$\frac{d}{dt}\delta j \propto J^0, \quad \frac{d}{dt}\delta j_z \propto J^0, \quad \frac{d}{dt}\delta\omega \propto J^{-1}, \quad \frac{d}{dt}\delta\Omega \propto J^{-1}. \quad (6.17)$$

Thus, we expect the fluctuations δj , δj_z to be independent of j as $e \rightarrow 1$, i.e. as $j^2 \rightarrow j_z^2 \rightarrow 0$. In other words, we do not expect any sharp peak in δj or δj_z as we approach maximum eccentricity. On the contrary we do expect a spike in $\delta\omega$ and $\delta\Omega$ as we approach the highest eccentricities. Such behaviour is precisely what we found in Figure 6.3 and in panels (m), (p) and (q) of Figures 6.1-6.2.

Envelope of D , j_z fluctuations

When discussing the numerical examples above we saw that the envelopes of fluctuations in $j_{z,SA}$ and D_{SA} had periodic amplitudes, with periods of t_{sec} and $t_{\text{sec}}/2$ respectively. We now explain each of these behaviours in turn.

First, we consider the envelope of j_z fluctuations. It is clear from the numerical examples that the amplitude of this envelope is minimised at the midpoint of each secular cycle, as ω/π gradually moves through ~ 0.5 , and maximised when ω/π gets close to its extrema. Evaluating the torque formula $dj_{z,SA}/dt$ at these different values of ω using equation (6.43), we find that the envelope of j_z fluctuations reflects nothing more than the amplitude of this fluctuating torque⁵.

Second, we address the fluctuations in D_{SA} . In this case the amplitude of the envelope exhibits minima at times corresponding to both $j_{DA} = j_{\text{max}}$ and $j_{DA} = j_{\text{min}}$, and maxima in-between. To see why, we differentiate (6.5):

$$\frac{dD_{SA}}{dt} = -\frac{2Dj}{1-j^2} \frac{dj}{dt} + \frac{10\Gamma e^2}{1-5\Gamma} \left(\sin 2i \sin^2 \omega \frac{di}{dt} + \sin 2\omega \sin^2 i \frac{d\omega}{dt} \right), \quad (6.18)$$

where, strictly, all quantities should have the ‘SA’ suffix. Consider the first term on the right hand side of (6.18). For $j \rightarrow j_{\text{min}} \ll 1$ it is proportional to j (see equation (6.17)), and so is negligible. For $j \rightarrow j_{\text{max}} \sim 1$ it is finite but small, because the $(1-j^2)^{-1}$ factor cancels on the $(J^2 - L^2)$ prefactor in the SA torque formula (equation (6.42)). Now take the second term on the right hand side of (6.18). Near minimum eccentricity, e^2 is small, and $i \approx i_{\text{max}} \approx 90^\circ$, so both $\sin 2i$ and $\sin^2 i$ are very small. Near maximum eccentricity ω approaches $\pi/2$, so $\sin^2 \omega$ and $\sin 2\omega$ are both minimised; also the evolution of ω slows down dramatically, so that $d\omega/dt$ is small and negative⁶. These arguments no longer hold away from the extrema of j , so that the D_{SA} envelope is minimised at those extrema and maximised in-between.

⁵There is also the rapid ‘swing’ between extreme values of ω that occurs around maximum eccentricity, during which ω/π again passes through 0.5. However this swing happens too quickly for the envelope of j_z fluctuations to be significantly affected.

⁶Except, again, for the extremely rapid but brief swing in ω at high eccentricity — see the previous footnote.

6.3.4 Characteristic amplitude of δj fluctuations

Perhaps the most important consequence of short-timescale fluctuations is that they enhance the value of e_{\max} when e gets very large, which can lead to e.g. more rapid compact object binary mergers (Grishin et al. 2018). With this in mind, we wish to estimate $(\delta j)_{\max}$, which we define to be the absolute value of the maximum fluctuation δj in the vicinity of maximum eccentricity. Unfortunately, it is immediately obvious from Figures 6.1-6.2 that a binary can have a different value of $(\delta j)_{\max}$ from one secular eccentricity peak to the next. Nevertheless, our goal in this section will be to estimate the characteristic size of such fluctuations and then demonstrate numerically that our estimate is a reasonable one.

For simplicity we will assume that Φ is spherically symmetric. Then to evaluate the torque at high eccentricity we can use (6.42), which by (6.13) and (6.16) is a good approximation to $-\mathrm{d}\delta j/\mathrm{d}t$ if we evaluate it using DA quantities. The maximum eccentricity as predicted by the DA theory is $e_{\mathrm{DA}} = e_{\max} \approx 1$, and it always occurs either at $\omega_{\mathrm{DA}} = \pm\pi/2$ or at $\omega_{\mathrm{DA}} = 0$. Let the corresponding minimum inclination be i_{\min} . Evaluating (6.42) at these (assumed fixed) DA values, we find

$$\left. \frac{\mathrm{d}\delta J}{\mathrm{d}t} \right|_{\omega=\pm\pi/2} = \frac{5}{4} a^2 \cos i_{\min} \times 2f_-(R) \sin[2(\phi - \Omega)], \quad (6.19)$$

or the same thing with an additional minus sign if evaluating at $\omega = 0$. Note that the function $f_-(R)$, defined in equation (6.38), depends on the instantaneous value of the outer orbital radius $R(t)$. Finally, one can check that for a Keplerian potential $\Phi = -GM/R$ we recover from (6.19) equation (B4) of Ivanov et al. (2005).

Next, we use the fact that a and $\cos i_{\min}$ are constants and we assume that Ω is stationary on the timescale T_ϕ . Placing the maximum DA eccentricity at $t = 0$ without loss of generality, we set $\Omega = \Omega(0)$ (although as with stationarity of ω and e , this can be a poor assumption, as we know from §4.C). Then the only time dependence in equation (6.19) comes from $R(t)$ and $\phi(t)$. Furthermore, $f_-(R) < 0$ for all R in sensible cluster potentials⁷. As a result, the sign of the torque at highest eccentricity (equation (6.19)) is dictated entirely by the instantaneous value of the phase angle $2(\phi - \Omega)$. The fluctuation

⁷To see this, suppose the cluster has density profile $\rho(r)$. From Poisson's equation $\nabla^2\Phi = 4\pi G\rho$ it is straightforward to show that

$$\frac{\partial^2\Phi}{\partial R^2} - \frac{1}{R} \frac{\partial\Phi}{\partial R} = R \frac{\partial}{\partial R} \left(\frac{GM(R)}{R^3} \right), \quad (6.20)$$

where $M(R) = \int_0^R 4\pi r^2 dr \rho(r)$ is the mass enclosed inside a sphere of radius R . For any any model in which density is a non-increasing function of radius, the expression (6.20) is negative for all R .

$(\delta j)_{\max}$ is therefore accumulated over a quarter period in azimuth, say from $\phi(t_1) - \Omega = 0$ to $\phi(t_2) - \Omega = \pi/2$, after which the torque changes sign. Integrating (6.19) over time we find

$$(\delta j)_{\max} = \frac{5}{4} \left[\frac{a^3}{G(m_1 + m_2)} \right]^{1/2} \cos i_{\min} F(r_p, r_a) \quad (6.21)$$

$$= 10^{-4} \times \left(\frac{\sqrt{\Theta}/j_{\min}}{1} \right) \left(\frac{m_1 + m_2}{M_{\odot}} \right)^{-1/2} \left(\frac{a}{10\text{AU}} \right)^{3/2} \\ \times \left(\frac{F^*}{0.8} \right) \left(\frac{M}{10^5 M_{\odot}} \right)^{1/2} \left(\frac{b}{\text{pc}} \right)^{-3/2}, \quad (6.22)$$

where all the details of the potential and outer orbit have been absorbed by the function

$$F(r_p, r_a) = \int_{t_1}^{t_2} dt \left| \frac{\partial^2 \Phi}{\partial R^2} - \frac{1}{R} \frac{\partial \Phi}{\partial R} \right| \sin[2(\phi - \Omega(0))], \quad (6.23)$$

and in the numerical estimate (6.22) we defined the dimensionless number

$$F^* \equiv (GM/b^3)^{-1/2} F. \quad (6.24)$$

The problem with equations (6.21) and (6.23) as they stand is that we do not know precisely which quarter-period in ϕ will provide the dominant fluctuation, because this would require knowledge of $R(t)$ and $\phi(t)$. Thus, we cannot evaluate (6.23) directly for an arbitrary outer orbit — and even if we could, we would not expect the resulting $(\delta j)_{\max}$ to be exactly correct because of the non-stationarity of ω, j, Ω .

Circular approximation

The simplest (and practically speaking, only) way to proceed is to estimate F by imagining that the binary is on a circular outer orbit with radius R . Then $F = F_{\text{circ}}(R)$ where

$$F_{\text{circ}}(R) = \left| \frac{\partial^2 \Phi}{\partial R^2} - \frac{1}{R} \frac{\partial \Phi}{\partial R} \right| \bigg/ \sqrt{\frac{1}{R} \frac{\partial \Phi}{\partial R}}. \quad (6.25)$$

In Figure 6.4 we plot the dimensionless number $F_{\text{circ}}^* \equiv (GM/b^3)^{-1/2} F_{\text{circ}}$ as a function of R/b for circular outer orbits in various spherically symmetric cluster potentials. For reference we also plot F_{circ}^* for the Kepler potential $\Phi = -GM/r$ (the scale radius b is arbitrary in that case). We see that in the cored (Plummer and isochrone) models F_{circ}^* has a maximum value of order unity which is realised when $R \sim b$, and that it falls sharply to zero towards the centre of the cluster. For centrally cusped potentials (Hernquist and NFW) we again have $F_{\text{circ}}^* \sim 1$ at intermediate radii $R \sim b$, but F_{circ}^* diverges towards the centre as $\sim R^{-1/2}$, typically reaching $F_{\text{circ}}^* \sim 10$ at the smallest sensible radii. At very large radii $R \gg b$, the isochrone, Plummer and Hernquist potentials tend toward Keplerian behaviour, $F_{\text{circ}}^* \sim R^{-3/2}$. (The logarithm in the NFW potential (2.53) means it never quite becomes Keplerian at these radii).

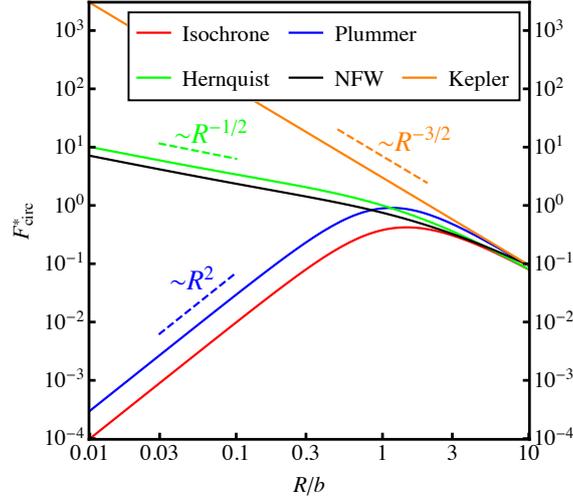


Figure 6.4: Plot of the dimensionless function $F_{\text{circ}}^* \equiv F_{\text{circ}}(R)/(GM/b^3)^{1/2}$, where F is defined in equation (6.25), for the four spherical potentials (2.50)-(2.53) as well as the Kepler potential $\Phi = -GM/r$.

From Figure 6.4 we learn that (i) the magnitude of short-timescale angular momentum fluctuations is roughly independent of potential type for $R \gtrsim b$, (ii) short-timescale fluctuations are significantly larger in cusped potentials than in cored potentials for $R < b$, and (iii) very large values of F^* can be reached at small radii in the Kepler potential. We also learn that in cusped potentials, the dominant short-timescale fluctuation in the angular momentum at very high eccentricity is likely to coincide with the outer orbit's pericentre passage (despite the fact that binaries spend more time at apocentre than pericentre). In other words, for binaries orbiting cusped potentials we should evaluate F_{circ} (equation (6.25)) at $R = r_p$. In cored potentials this is no longer true because of the turnover in F_{circ}^* at $R \sim b$. Then, for example, for orbits with $r_a \lesssim b$ the dominant j fluctuations arise around apocentre passage. However, outer orbits in cored potentials with $r_a \lesssim b$ tend to have $\Gamma < 1/5$ (Figure 2.5), so they tend not to reach such high eccentricities anyway⁸, and besides, the values of F_{circ}^* never exceeds ~ 1 regardless of R for these potentials. Hence, for an order of magnitude estimate we may choose simply to evaluate $(\delta j)_{\text{max}}$ using equation (6.25) with $R = r_p$, regardless of the potential or outer orbit.

In panel (p) of Figures 6.1-6.2 we show as ‘error bars’ the values of $\pm(\delta j)_p$ (cyan) and $\pm(\delta j)_a$ (yellow), which are calculated by evaluating $\pm(\delta j)_{\text{max}}$ (equation (6.21)) using the circular approximation (6.25) at $R = r_p$ and $R = r_a$ respectively. We see that the circular approximation gives a decent rough estimate of the amplitude of fluctuations δj . In the LK case, the result (6.21) with the circular approximation (6.25) was first used by Ivanov et al. (2005).

⁸Of course the $\Gamma < 1/5$ numerical examples shown in this section *do* reach very high e , but that is because we have purposely chosen a rather special set of initial conditions in order to make this so.

One important caveat here is that while the $\delta j(t)$ behaviour is often rather regular up to $e_{\text{DA}} \lesssim 0.99$, it often becomes very *irregular* in the immediate vicinity of the eccentricity peak, as can be seen in Figures 6.1-6.2. This is because of the rapid evolution of ω and Ω when $j \approx j_{\text{min}}$ (see the light blue shaded bands in panels (j) and (k) of each of those Figures) which introduces a significant phase dependence into the detailed fluctuation behaviour. Of course, since (6.21) was derived by assuming stationary ω, Ω, j , it necessarily fails to capture this irregular behaviour.

6.4 The effect of GR precession

In the previous section we gained a lot of insight into how fluctuations in the cluster tidal torque affect high eccentricity behaviour, but if we are to investigate many real systems — e.g. binary black hole mergers — then it is vital that we also account for GR precession. As we will see in this section, including GR can change the picture significantly. To begin with, in §§6.4.1-§6.4.2 we rerun the numerical calculations from §§6.3.1-6.3.2 except this time with GR precession switched on, and simply describe the altered phenomenology. In §6.4.3 we compare the GR and non-GR calculations in the special case of the LK problem. The main new result that arises in each of these subsections is the diffusion of the binary integral of motion D in the SA approximation. In §6.4.4 we offer a physical explanation for this new phenomenon and attempt to analyse it quantitatively.

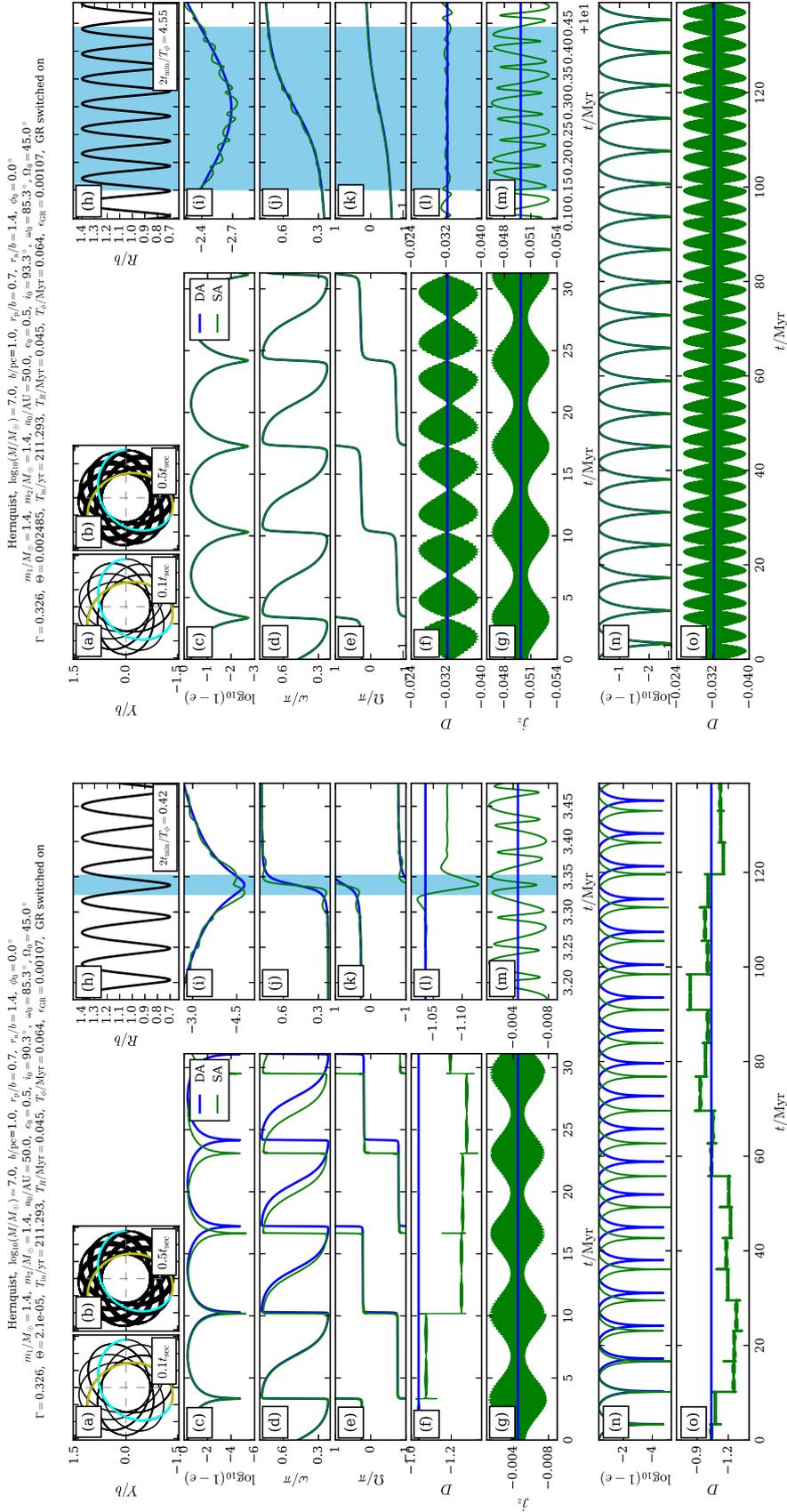
6.4.1 Two examples in the Hernquist potential

(A) Fiducial Hernquist example including GR precession

In Figure 6.5A we rerun the calculation from Figure 6.1A, except we switch on the GR precession term (with strength $\epsilon_{\text{GR}} = 0.00107$) in the SA and DA equations of motion. We now discuss Figure 6.5A in some detail.

The structure of panels (a)-(m) is identical to those of Figure 6.1A, except that we have dispensed with N-body integration since it is prohibitively computationally expensive, and that in panels (h)-(m) we have chosen to zoom in on the first secular eccentricity peak rather than the second. Comparing panels (a)-(m) with those of Figure 6.1A we immediately notice several qualitative differences. Whereas in Figure 6.1A the DA and SA predictions for $\log_{10}(1-e)$ agreed almost perfectly except at extremely high eccentricity, now in Figure 6.5A (GR precession on) they disagree manifestly after the first eccentricity maximum. Moreover, while the period of secular oscillations is fixed in the DA case, the SA secular period changes from one eccentricity peak to the next. By the time of the fourth eccentricity peak the DA and SA curves in panels (c)-(e) are completely out of sync.

Furthermore, from panel (f) we see that D_{SA} no longer fluctuates around D_{DA} indefinitely like it did in Figure 6.1Af, but rather exhibits discrete jumps during very



(A) Fiducial Hernquist example including GR precession.

(B) Changing the initial inclination to $i_0 = 93.3^\circ$.

Figure 6.5: (A) As in Figure 6.1A, except we switch on GR precession. This causes the time-averaged value of D_{SA} to diffuse away from D_{DA} , with discrete jumps occurring during episodes of extremely high eccentricity. We call this relativistic phase space diffusion ('RPSD'). (B) As in Figure 6.5A except we take $i_0 = 93.3^\circ$ — in other words, the same as Figure 6.1B except we switch on GR precession. There is no RPSD in this case.

high eccentricity episodes. In panel (l) we zoom in on the D_{SA} behaviour around the first eccentricity peak. We see that D_{SA} jumps from ≈ -1.05 to ≈ -1.1 and that this jump happens essentially within the ‘blue stripe’ (equation (6.6)), i.e. on the timescale $\sim 2t_{\text{min}}$. Meanwhile, we observe in panels (g) and (m) that $j_{z,\text{SA}}$ does not perform any discrete jumps at high-eccentricity, but rather exhibits a very similar behaviour to the non-GR case (Figure 6.1Ag,m). Jumping to a new D value while keeping j_z fixed means a new value of H^* , i.e. the binary now traces a different contour in the DA (ω, e) phase space (Chapters 3-4). Since this behaviour depends crucially on the presence of finite GR precession we choose to call it ‘relativistic phase space diffusion’ (RPSD). Finally, each phase space contour has its own secular period. Hence it is unsurprising that a jump in D_{SA} leads to a modified SA secular period compared to the fixed DA period⁹.

To investigate the RPSD behaviour further, we next ran the integration for a much longer time¹⁰, $20t_{\text{sec}}$. In panels (n) and (o) of Figure 6.5A we plot $1 - e$ and D respectively as functions of time over this entire duration. The same picture holds in that D_{SA} is roughly static between eccentricity peaks, but often makes a discernible jump during a peak. These jumps seem to have no preferred sign.

(B) Changing the initial inclination to 93.3°

The diffusion of D_{SA} at eccentricity peaks is *not* present in Figure 6.5B, in which we keep all conditions exactly the same as Figure 6.5A except that we change i_0 from 90.3° to 93.3° . (In other words, we rerun the same calculation as in Figure 6.1B except with GR switched on). Note that this case has $2t_{\text{min}}/T_\phi = 4.55 \gtrsim 1$.

6.4.2 Two examples in the Plummer potential

(A) Fiducial Plummer example including GR precession

In Figure 6.6A we give another example of RPSD, this time in the Plummer potential. This example is identical to that in Figure 6.2A except that we switched on GR precession, and zoomed in on the first eccentricity peak in the right column rather than the second. A major jump in D_{SA} can be seen in panel (o), around 39 Myr, which entirely shifts the SA evolution away from the original DA prediction.

⁹For instance, in Figure 6.5A, making D more negative while keeping j_z — and therefore Θ — fixed moves the binary further to the left of the separatrix in (D, Θ) phase space (see Figure 3.3), which accounts for the decrease in the period of the subsequent secular oscillation.

¹⁰Note that t_{sec} is not the exact DA secular period when we include GR precession, as it is calculated assuming no GR (equation (3.33)), but for $\epsilon_{\text{GR}} \ll 3(1 + 5\Gamma)$ it is a very good approximation (Chapter 4).

(B) Shifting the initial nodal angle Ω_0 by 45°

We give another example in Figure 6.6B, which uses the same initial conditions as Figure 6.2B except with GR switched on. In other words, this example is identical to Figure 6.6A except for an initial orbital phase change, taking $\Omega_0 = 0^\circ$. The result is a completely different SA evolution, with no discernible RPSD for the first three eccentricity peaks followed by a large kick at the fourth, around 24Myr (panel (a)). We see here that the effect of an initial orbital phase offset is far more dramatic with GR precession than without: said differently, RPSD dramatically amplifies small differences in the initial conditions, often producing entirely divergent secular behaviour from one case to the next. This phase dependence is illustrated further in §6.C.

6.4.3 An example in the Lidov-Kozai limit

The main phenomenological result we have found in this section is that when GR precession is switched on and a binary reaches very high eccentricity, short-timescale fluctuations may cause it to be ‘kicked’ to a qualitatively new phase space trajectory. It is important to note that this behaviour is not exclusive to the non-Keplerian potentials that we have investigated so far, but is present in the Lidov-Kozai problem as well (though to our knowledge, nobody has mentioned it explicitly). To demonstrate this we provide Figure 6.7.

(A) LK example without GR precession

In Figure 6.7A we run a calculation with exactly the same initial condition as in Figure 6.1A, except that we change the potential to the Keplerian one, $\Phi = -GM/\sqrt{R^2 + Z^2}$. In other words, we are now investigating the classic test particle quadrupole Lidov-Kozai problem, relevant to e.g. a NS-NS binary orbiting a SMBH (e.g. Antonini & Perets 2012; Hamers 2018a; Bub & Petrovich 2020), except that we have GR switched off. In panels (a) and (b) we simply see the outer orbital ellipse, which has semimajor axis $a_g = (r_a + r_p)/2 = 1.05b$ and eccentricity $e_g = (r_a - r_p)/(r_a + r_p) = 0.33$. Panels (c)-(m) show quite similar behaviour to that of Figure 6.1A (although note that this time we have chosen to zoom in on the first eccentricity peak in the right hand panels, rather than the second). Since $2t_{\min}/T_\phi = 0.15$, the highest eccentricity episode lasts for significantly less than one outer orbital period. Despite this, the SA result tracks the secular (DA) result indefinitely, and the N-body integration again confirms that the SA approximation is a good one.

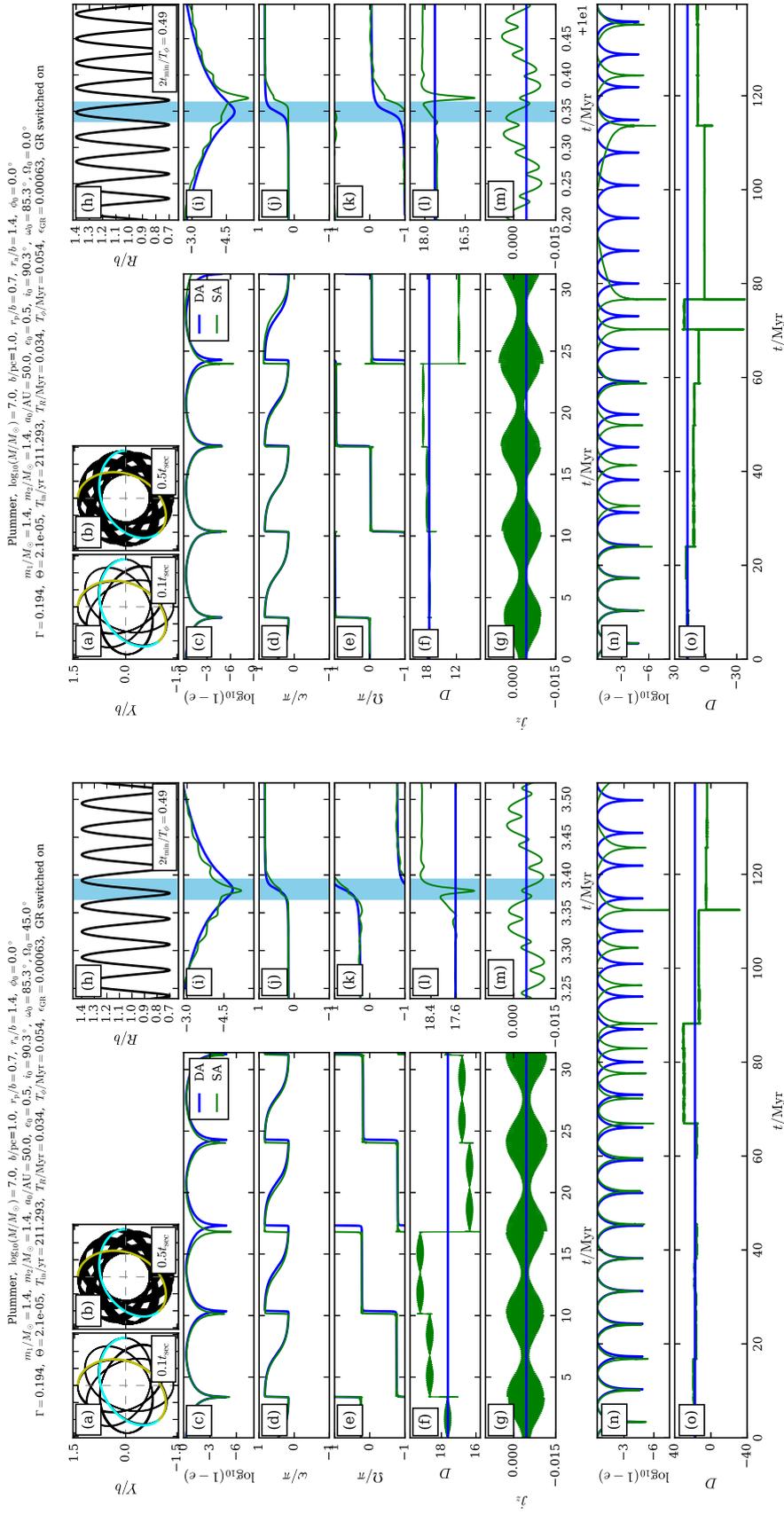


Figure 6.6: (A) As in Figure 6.5A except we use the Plummer potential — in other words, the same as Figure 6.2A except we switch on GR precession. (B) We take $\Omega_0 = 0^\circ$. The strong phase dependence of RPSD means that the SA secular behaviour completely diverges from that of Figure 6.6A..

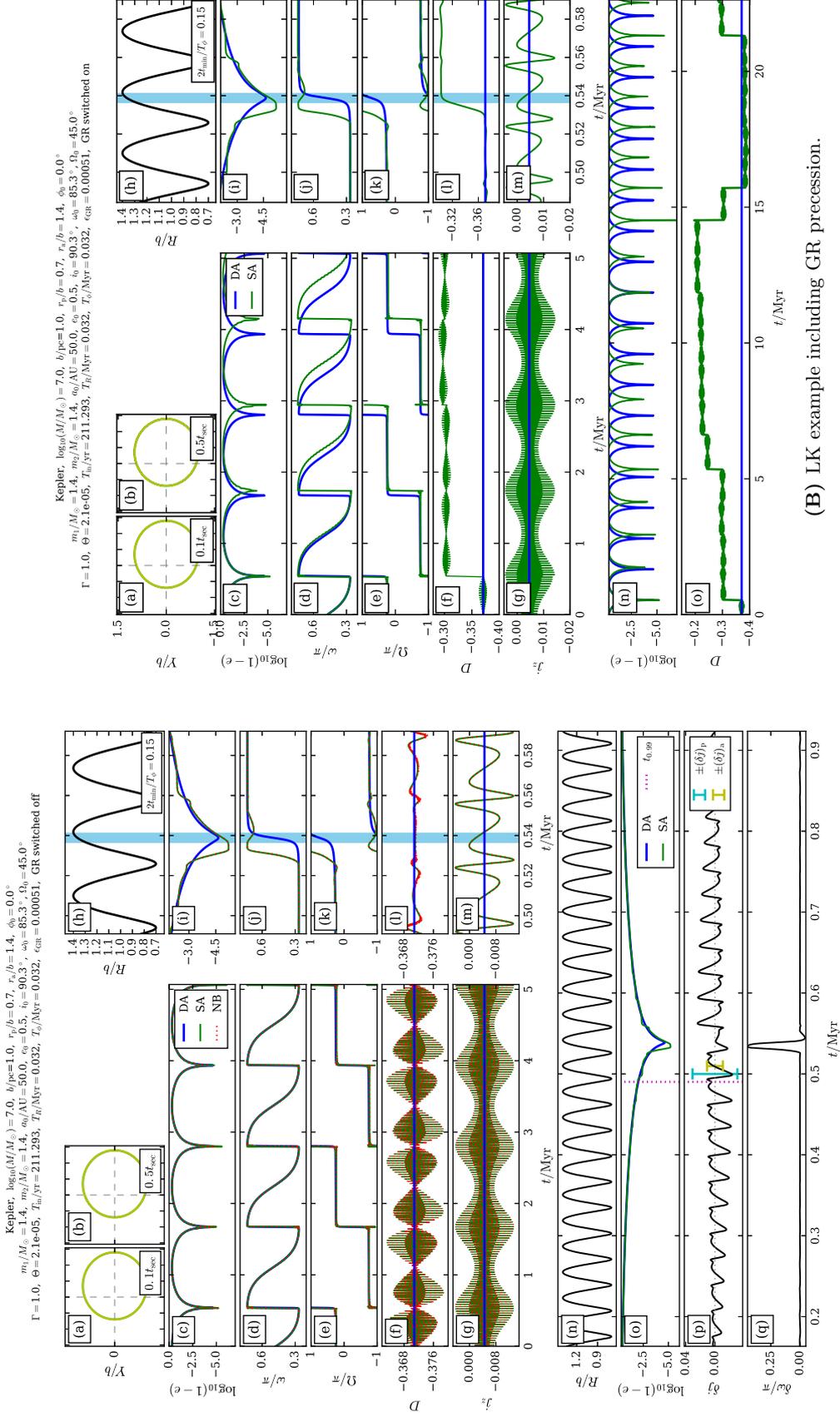


Figure 6.7: (A) Initial conditions as in Figure 6.1A, except we use the Kepler potential. GR precession is switched off here. (B) The same system with GR precession switched on.

(B) LK example including GR precession

In Figure 6.7B we perform the same calculation with GR precession switched on. Again, this causes significant and repeated diffusion of D_{SA} around the secular eccentricity peaks, meaning the binary ultimately follows a completely different trajectory to the one we would have predicted had we only used DA theory. This confirms that RPSD is present as a phenomenon in LK theory *at the test-particle quadrupole level* as long as 1pN GR precession is included, despite not having been discussed before in the literature.

6.4.4 Physical interpretation and quantitative analysis

By examining the phenomenology of several numerical experiments, we have uncovered a new effect, namely the ‘jump’ behaviour of D_{SA} at high eccentricity, which we have termed RPSD. However, we still lack a quantitative description of when and why it occurs. The aim of this section is to provide some physical understanding and to attempt a rough quantitative analysis of RPSD.

Physical interpretation of RPSD, and the necessary conditions for its occurrence

Why does RPSD happen? We take note of three pieces of evidence which will help us answer this question.

First, we have also seen empirically that when GR precession is switched off, there is no RPSD. Second, we have seen in Figures 6.5-6.7 that RPSD can occur when $2t_{\text{min}}/T_\phi \lesssim 1$ but never occurs when $2t_{\text{min}}/T_\phi \gtrsim 1$ (and we confirmed this in several additional numerical experiments not shown here). Third, we notice that there is no diffusion in the time-averaged value of $j_{z,\text{SA}}$ to accompany the diffusion of D_{SA} (recall that j_z and D are both integrals of motion in the DA approximation). Taking these different strands of evidence together, the physical mechanism behind RPSD soon becomes clear, as we now explain.

Whether we consider SA or DA dynamics, at eccentricities far from unity, significant changes in the orbital elements occur only on secular timescales, i.e. timescales much longer than T_ϕ . Of course D (equation (6.5)) is a function of these orbital elements. Thus at eccentricities far from unity, $D_{\text{SA}}(t)$ invariably exhibits relatively small and rapid (timescale $\sim T_\phi$) oscillations around the constant value D_{DA} . However, at the highest eccentricities $e \rightarrow 1$, significant changes in orbital elements can occur on much shorter timescales. This is true even in DA theory: indeed, in §4.C we saw several numerical examples of binaries whose ω_{DA} value is turned through $\sim 90^\circ$ or more on the timescale $2t_{\text{min}}$. Now, when $2t_{\text{min}}/T_\phi \gtrsim 1$ this is still slow from the point of view of the outer orbit, since the relevant timescale is still much longer than any orbital period. But in the regime $2t_{\text{min}}/T_\phi \lesssim 1$, the DA theory essentially predicts its own demise: $\mathcal{O}(1)$ relative changes in the orbital elements occur on the timescale of the outer orbit, contrary to the fundamental

assumption of the secular approximation (Chapter 2). In that case, it is possible that the fluctuations of SA theory may no longer be considered small, and would no longer oscillate rapidly around the DA orbital elements while those DA elements change significantly.

For an example, consider e.g. Figure 6.5Aj, for which $2t_{\min}/T_\phi = 0.42$. In this case ω_{DA} undergoes a large ‘swing’ in the time range shaded in blue, which is centred on peak DA eccentricity. This time range is so short that ω_{SA} does not have a chance to perform any full oscillation during it. By contrast, consider Figure 6.5Bj, for which $2t_{\min}/T_\phi = 4.55$. In this instance ω_{DA} undergoes a swing of very similar magnitude but on a much longer timescale, allowing ω_{SA} to perform many small oscillations while the swing is in progress (i.e. within the blue range).

How does this lack of timescale separation between t_{\min} and T_ϕ affect $j_{z,\text{SA}}$ and D_{SA} ? Well, when GR is switched off, they are not affected very significantly. This is a consequence of the fact that without GR, $\text{d}j_{z,\text{SA}}/\text{d}t$ and $\text{d}D_{\text{SA}}/\text{d}t$ scale rather weakly with j at very high eccentricity¹¹, and so do not respond vigorously to the rapid changes in the various orbital elements — instead they just oscillate around the constants $j_{z,\text{DA}}$ and D_{DA} indefinitely. However, when we *do* include GR precession, $\text{d}D_{\text{SA}}/\text{d}t$ gets an additional contribution: we must add to the right hand side of (6.18) the time derivative of the last term in equation (6.5):

$$\left(\frac{\text{d}D_{\text{SA}}}{\text{d}t}\right)_{\text{GR}} = \frac{\epsilon_{\text{GR}}}{3(1-5\Gamma)} \frac{1}{j^2} \frac{\text{d}j}{\text{d}t}, \quad (6.26)$$

where strictly we should evaluate all quantities at their SA values. Now, we know that $\text{d}j_{\text{SA}}/\text{d}t$ does not vary with j when we approach very high eccentricity (the doubly-averaged part of j is stationary at maximum DA eccentricity by definition, while the fluctuating part δj satisfies equation (6.17)). As a result $\text{d}D_{\text{SA}}/\text{d}t \propto \epsilon_{\text{GR}} j^{-2}$ as $j \rightarrow 0$. If $\epsilon_{\text{GR}} j_{\min}^{-2}$ is sufficiently large, this can lead to a fluctuation in D around peak eccentricity that is far larger than the ones we saw when GR was switched off. In this case, the relative fluctuations in D_{SA} can be so large, and D_{SA} can have so little time to oscillate back and forth around its ‘parent’ D_{DA} value before the DA orbital elements change, that upon emerging from the high-eccentricity blue stripe it ‘settles’ on a new parent D_{DA} value (see e.g. Figure 6.5Al). Equivalently, the binary ‘jumps’ to a new (ω, e) phase space contour while at high e (still at fixed Γ , Θ and ϵ_{GR}), aided by the fact that contours at high e are bunched so closely together (e.g. Figure 4.6). On the other hand, $\text{d}J_z/\text{d}t \propto J^0$ is unaffected by GR precession, so these conclusions do not apply to $j_{z,\text{SA}}$.

Let us attempt to be somewhat more quantitative. Suppose we tried to calculate the ‘jump’ that D_{SA} sustains across the eccentricity peak. First, we know $\text{d}D_{\text{SA}}/\text{d}t$ is

¹¹We already argued this for D_{SA} in §6.3.3; as for $\text{d}j_{z,\text{SA}}/\text{d}t$, we simply need to note that $\text{d}J_z/\text{d}t \propto J^0$ as $J \rightarrow J_z \rightarrow 0$ (see equation (6.17)).

driven by the GR term (6.26) at high eccentricity, and that we can ignore the other contributions. If we assume that far from the eccentricity peaks the fluctuations in D_{SA} are small, then we can choose the time coordinate to coincide with¹² $j_{\text{DA}}(0) = j_{\text{min}}$ and calculate D^{jump} by integrating (6.26):

$$D^{\text{jump}} \approx \frac{\epsilon_{\text{GR}}}{3(1-5\Gamma)} \int_{-\infty}^{\infty} dt \frac{1}{j_{\text{SA}}^2(t)} \left(\frac{dj_{\text{SA}}}{dt} \right)_{\mathbf{w}_{\text{SA}}(t), \mathbf{R}_{\text{g}}(t)}, \quad (6.27)$$

where we are justified in taking the limits of the integral to $\pm\infty$ since for $|t| \gg t_{\text{min}}$ the integrand decays as $\sim t^{-2}$. The notation here is to remind us that SA torque dj_{SA}/dt (see equation (6.34)) is to be evaluated using the SA solution $\mathbf{w}_{\text{SA}}(t)$ and the instantaneous outer orbital position $\mathbf{R}_{\text{g}}(t)$. Note that we cannot simply write the integral (6.27) as $2 \int j_{\text{SA}}^{-2} dj_{\text{SA}}$ because in general there is no one-to-one correspondence between j_{SA} and t like there is between j_{DA} and t . Now, if SA fluctuations were identically zero then we could replace $\mathbf{w}_{\text{SA}} \rightarrow \mathbf{w}_{\text{DA}}$ and in particular $j_{\text{SA}} \rightarrow j_{\text{DA}}$, resulting in $D^{\text{jump}} = 0$, as expected. This simply reflects the anti-symmetry of the resulting integrand $j_{\text{DA}}^{-2} dj_{\text{DA}}/dt$ around $t = 0$. To get a non-zero D^{jump} we need to break this anti-symmetry. Let us use the definition (6.16) and the approximation (6.16); then we can write the above equation as

$$D^{\text{jump}} \approx \frac{\epsilon_{\text{GR}}}{3(1-5\Gamma)} \int_{-\infty}^{\infty} dt \frac{1}{[j_{\text{DA}}(t) + \Delta j(\mathbf{w}_{\text{DA}}(t), t)]^2} \left(\frac{dj_{\text{SA}}}{dt} \right)_{\mathbf{w}_{\text{DA}}(t), \mathbf{R}_{\text{g}}(t)} \quad (6.28)$$

If $2t_{\text{min}}/T_{\phi} \gg 1$ then the SA fluctuations are small¹³: $|\Delta j| \ll j_{\text{DA}}$ at all times. They are also rapid, so that dj_{SA}/dt oscillates many times while j_{DA} is in the vicinity of j_{min} . In this case the fluctuations naturally wash out to zero, giving a negligible D^{jump} — see Figure 6.6B. On the other hand, if $2t_{\text{min}}/T_{\phi} \lesssim 1$ then the integrand in (6.28) is no longer close to anti-symmetric in time. Moreover, the value of this integrand can be boosted significantly if $j_{\text{DA}} + \Delta j$ conspires to come very close to 0 around the eccentricity peak (though in practice this depends very sensitively on orbital phase information). Under these circumstances we get a value for D^{jump} that is significantly different from zero.

This analysis tells us that RPSD truly requires three key ingredients. (i) We require $\epsilon_{\text{GR}} > 0$, so that the term (6.26) is switched on. (ii) We need very high eccentricity DA behaviour leading to rapid changes in all orbital elements around j_{min} ; this requires $\epsilon_{\text{GR}} \ll \epsilon_{\text{strong}}$ (Chapter 4). (iii) We need the short-timescale fluctuations to be significant so that the the integral in (6.28) does not wash out to zero; this requires $2t_{\text{min}}/T_{\phi} \lesssim 1$,

¹²Again we have centred the DA eccentricity peak at $t = 0$ without loss of generality.

¹³To see this, we look at equation (6.21). Very roughly speaking, using $\cos i_{\text{min}} \sim j_{\text{min}}$ and $F \sim T_{\phi}^{-1}$, we can say that $\delta j \sim j_{\text{min}} T_{\text{in}}/T_{\phi}$ (though the true value is very strongly phase-dependent). Coupling this with $t_{\text{min}} \sim j_{\text{min}} t_{\text{sec}}$ and $t_{\text{sec}} \sim T_{\phi}^2/T_{\text{in}}$ (Chapter 3) we have $\delta j/j_{\text{min}} \sim j_{\text{min}} T_{\phi}/t_{\text{min}}$. Clearly for $2t_{\text{min}}/T_{\phi} \gg 1$ this quantity is very small.

which, using $t_{\min} \sim j_{\min} t_{\text{sec}}$, $t_{\text{sec}} \sim T_{\phi}^2/T_{\text{in}}$ (Chapter 3) and assuming weak GR so that $j_{\min} \sim \Theta^{1/2}$, may be rewritten as:

$$\Theta \lesssim \left(\frac{T_{\text{in}}}{T_{\phi}} \right)^2 \quad (6.29)$$

$$\lesssim 5 \times 10^{-5} \times \left(\frac{m_1 + m_2}{2.8M_{\odot}} \right)^{-1} \left(\frac{a}{50\text{AU}} \right)^3 \left(\frac{\mathcal{M}}{10^7 M_{\odot}} \right)^1 \left(\frac{R}{1\text{pc}} \right)^{-3}. \quad (6.30)$$

To get the second line here we took a binary on a circular outer orbit of radius R in a spherical cluster of mass \mathcal{M} , and put $T_{\phi}^2 \sim GM/R^3$. For binaries that do not start off at very high eccentricities we have $\Theta \approx \cos^2 i_0$, implying that RPSD operates in the inclination window:

$$|\cos i_0| \lesssim 0.007 \times \left(\frac{m_1 + m_2}{2.8M_{\odot}} \right)^{-1/2} \left(\frac{a}{50\text{AU}} \right)^{3/2} \left(\frac{\mathcal{M}}{10^7 M_{\odot}} \right)^{1/2} \left(\frac{R}{1\text{pc}} \right)^{-3/2}. \quad (6.31)$$

Throughout this Chapter we have considered clusters with mass $10^7 M_{\odot}$, binaries with the mass of a NS-NS binary, $m_1 = m_2 = 1.4M_{\odot}$, and outer orbits with semimajor axis $a_g = 1\text{pc}$. Putting $R \sim a_g$ with these numbers into equation (6.31) gives $|\cos i_0| < 0.007$, which corresponds approximately to $i_0 \in (89.6^\circ, 90.4^\circ)$. This is concomitant with what we found numerically; all examples that exhibited RPSD had $i_0 = 90.3^\circ$, while the example for which there was no RPSD (Figure 6.6B) had $i_0 = 93.3^\circ$. We discuss the astrophysical implications of the result (6.31) in §6.5.1.

Finally, let us clear up one potential misconception. It is *not* the case that rapid GR-induced advance of ω at high eccentricity is the culprit for pushing the binary to a new phase space contour, and hence for RPSD — at least not directly. Indeed, even in the absence of GR, the rapid ‘swing’ in ω_{DA} is always present at very high eccentricity, as we have seen in Figures 6.1, 6.2, 6.7A. Instead, it is the addition to dD_{SA}/dt of the term (6.26), which crucially diverges as $j \rightarrow 0$ and requires non-zero ϵ_{GR} , that drives the phase space diffusion. The susceptibility of D to small fluctuations at very high eccentricity is due to the fact that D is proportional to the DA perturbing Hamiltonian (equation (4.16)), and the level curves of this Hamiltonian are bunched very tightly together as $e \rightarrow 1$ (e.g. Figure 4.6). Overall, the remarkable and non-intuitive truth is that in the weak GR regime, GR has essentially no effect on the DA dynamics, and yet modifies the SA dynamics — and hence the real binary evolution — fundamentally.

Statistical analysis of phase space diffusion

We know from §6.C that the erratic high eccentricity behaviour when GR is switched on is highly phase dependent. Since predicting the behaviour of D_{SA} for a given set of initial conditions is very difficult, a natural next step is to investigate the statistics

of jumps in D_{SA} , by creating an ensemble D^{jump} values for the same binary set off with different orbital phases.

To do this, it is necessary to first define more precisely the quantity we are interested in studying. This immediately raises some technical issues: (i) the value of D_{SA} is never actually fixed, and (ii) a steady-state approximation of D_{SA} clearly breaks down as e_{SA} approaches unity. We therefore choose to study time-averages of D_{SA} before and after the eccentricity peak, taken over time intervals that stay sufficiently far from the peak for the averaged value to be meaningful¹⁴, i.e. for $\langle D_{\text{SA}} \rangle$ to correspond accurately to the ‘parent’ value D_{DA} . We are then interested in the quantity

$$D^{\text{jump}} \equiv \langle D_{\text{SA}} \rangle_{\text{after}} - \langle D_{\text{SA}} \rangle_{\text{before}}. \quad (6.32)$$

To probe this quantity, we considered a system with exactly the same setup and initial conditions of Figure 6.5A except for the value of Ω_0 . We ran $N = 300$ numerical integrations of the SA equations, using the initial phase angles $\Omega_0 = 0, 2\pi/N, \dots, 2\pi(N - 1)/N$. We stopped each integration at $t \approx t_{\text{sec}}$. By this time the binary had reached one eccentricity maximum and come out the other side, followed by enough time for the value of $\langle D_{\text{SA}} \rangle_{\text{after}}$ to roughly achieve steady state. Thus, D^{jump} was well-established according to equation (6.32). In Figure 6.8a we plot the resulting D^{jump} values against the minimum j_{SA} value achieved during the corresponding eccentricity peak. The dark red dots are for $\Omega_0 = 0$, and they progress towards yellow and then green as Ω_0 approaches 2π . In Figure 6.8b we show a histogram of the magnitude of those D^{jump} values. We see that the D^{jump} behaviour is very strongly phase dependent and sadly there is no simple dependence of D^{jump} on $j_{\text{SA},\text{min}}$. In fact, we plotted several such figures for several different sets of initial conditions, cluster potential, and so on, tried increasing N substantially, and experimented with different ways of binning the values of D^{jump} — unfortunately there was never any real order to be found within the chaos.

Perhaps this is due to the fact that, by keeping R_0 fixed, we have not sampled the full range of outer orbital phases? To test this idea, we next ran exactly the same calculation except this time for $N = 10^4$ pairs of initial radial and azimuthal phases, drawn randomly from the appropriately weighted distributions as described in §6.C. In Figure 6.9 we show the histogram of $\log_{10} |D^{\text{jump}}/D_{\text{DA}}|$ that results, where apart from the initial orbital phases we have used the same initial conditions as in Figures 6.5. We have also separated the $D^{\text{jump}}/D_{\text{DA}} \geq 0$ (red) and $D^{\text{jump}}/D_{\text{DA}} < 0$ (blue) jumps. Yet again, there do not seem to be any clear patterns to the histograms of D^{jump} values. We see that red and blue histograms do not even overlap. Testing various other initial

¹⁴In practice, it is sufficient to average D over a several outer orbital periods during a time range corresponding to $e_{\text{SA}} \notin (0.99, 1)$.

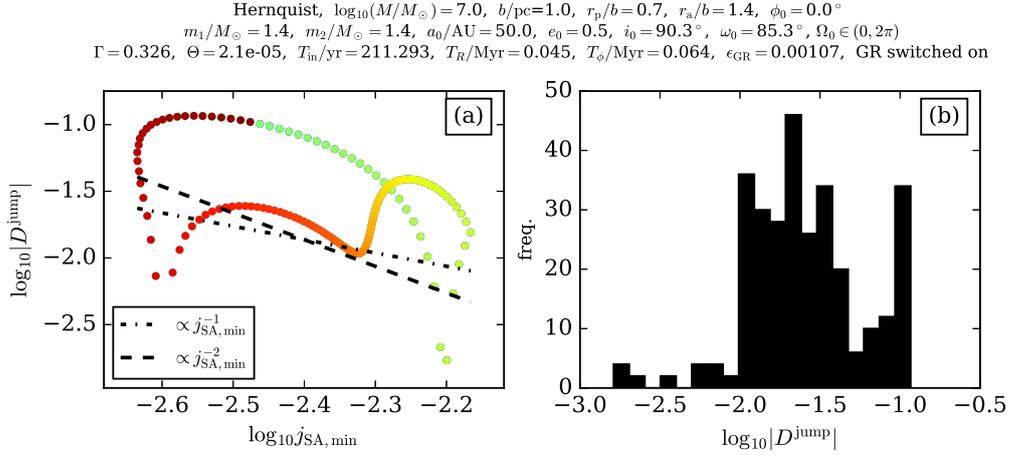


Figure 6.8: Rerun of the first secular period from Figure 6.5A, except for 300 different values of the initial Ω_0 . Panel (a) shows the magnitude of D^{jump} values from each run against the minimum j value achieved in the SA calculation. Panel (b) shows a histogram of the $|D^{\text{jump}}|$ values from panel (a).

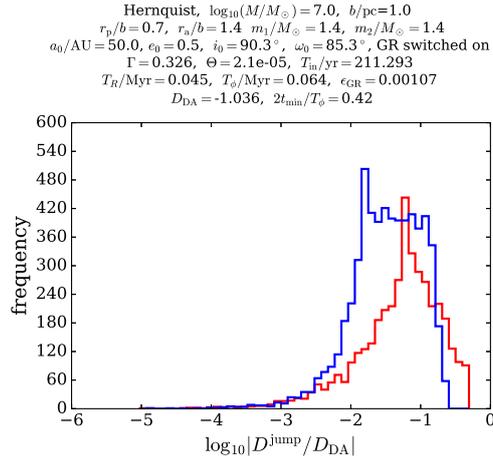


Figure 6.9: Histogram of $|D^{\text{jump}}/D_{\text{DA}}|$ values. Red is for positive $D^{\text{jump}}/D_{\text{DA}}$, blue for negative. We rerun the calculations from Figure 6.5A for $N = 10^4$ randomized values of the initial radial and azimuthal phase.

conditions, potentials, etc. (not shown here) reveals that there is no consistent symmetry between positive and negative jumps, nor does the distribution of jumps ever seem to converge towards any simple (e.g. lognormal) form. Overall, the statistics of D^{jump} values do not immediately reveal any striking insights.

6.5 Discussion

6.5.1 Astrophysical relevance of RPSD

Since we have uncovered a new effect in this Chapter the obvious question is: how relevant is it to astrophysical systems? With a view to considering compact object binary

mergers in Chapter 7, we will use that as a test case. Let us assume that our compact object binary satisfies $\epsilon_{\text{GR}} \ll \epsilon_{\text{weak}}$ and hence reaches very high eccentricity $j_{\text{min}} \sim \Theta^{1/2}$. Assume also that its initial eccentricity is not so large. Then the important necessary condition for significant RPSD to occur is equation (6.31). In Chapter 7, when considering binary neutron star mergers in globular and nuclear clusters, $10^7 M_{\odot}$ will be the upper limit on sensible cluster masses; $m_1 = m_2 = 1.4 M_{\odot}$ will be the lower limit on compact object mergers; and 50AU will be the upper limit on any sensible distribution of (still rather soft) NS-NS binaries. Since $\cos i_0$ is distributed uniformly $\in (0, 1)$ for an isotropic distribution of binaries, we conclude that RPSD will affect much less than 1% of our sample. Moreover, this fraction will be even smaller when we consider e.g. BH-BH binaries with $m_1 = m_2 = 30 M_{\odot}$. Thus, we do not expect RPSD to be important for the bulk populations of binary mergers that we consider in Chapter 7, and so we will ignore it there.

We speculate that RPSD may be important for certain exotic phenomena that involve even more extreme eccentricities, such as head-on collisions of white dwarfs in triple systems (Katz & Dong 2012). Note that RPSD occurs even for circular outer orbits and for equal mass binaries, i.e. in triples where the octupole contributions to the potential are very small, which are not usually considered promising for producing high merger rates. Further analysis of this possibility is left to future work.

6.5.2 Breakdown of the SA approximation

In this Chapter we have considered the SA dynamics of binaries driven by cluster tides, particularly at very high eccentricity. We have implicitly assumed that these equations are accurate. It is worth noting, however, that even the SA approximation can itself break down, in situations where any of (J, ω, J_z, Ω) evolve significantly on the timescale of the inner orbital period (Antonini, Murray, et al. 2014). Such a situation is shown in Figure 6.10. In this figure we consider a binary that reaches $1 - e \sim 10^{-4}$ in the Hernquist potential (we assume no GR precession in this example). We ran the integrations with very high resolution; in particular we used a very large number of timesteps for the outer orbit integration, $T_{\phi}/\Delta t = 5000$. Nevertheless, by the time of the 7th eccentricity peak, the N-body and SA results differ significantly¹⁵, with $1 - e$ differing at worst by ~ 0.5 dex. We see that ω and Ω change very rapidly near peak eccentricity, especially in the SA and N-body integrations. For instance, ω changes by $\sim \pi$ on a timescale of ~ 5000 yr. Given $T_{\text{in}} \approx 600$ yr, this corresponds to a rate $\dot{\omega} \sim 22^{\circ}$ per inner orbital period. Thus, given the $\cos 2\omega$ dependence of quadrupole-order terms in the SA equations of motion, it is not surprising that even the SA approximation breaks down here.

¹⁵Note that the binary component masses are equal in this example, $m_1 = m_2$, so any octupole terms in the tidal force ought to be zero, meaning any corrections are of hexadecapole or higher order. Nevertheless, we checked that the disagreement is really due to a breakdown in the orbit-averaging approximation, rather than these higher order terms, by running an N-body simulation with only the quadrupole potential included. The agreement with the full N-body simulation was excellent.

Hernquist, $\log_{10}(M/M_{\odot}) = 7.0$, $b = 1.0$ pc, $r_p/b = 0.3$, $r_a/b = 0.6$, $m_1 = 1.4 M_{\odot}$, $m_2 = 1.4 M_{\odot}$, $a_0 = 100.0$ AU, $e_0 = 0.8$, $i_0 = 91.3$, $\omega_0 = 85.3$, $\Omega_0 = 45.0$
 $\Gamma = 0.221$, $\Theta = 0.000185$, $\epsilon_{\text{GR}} = 1e-05$, $D = -12.945$, $T_{\text{in}} = 597.626$ yr, $T_R = 0.019$ Myr, $T_{\phi} = 0.03$ Myr, GR precession off

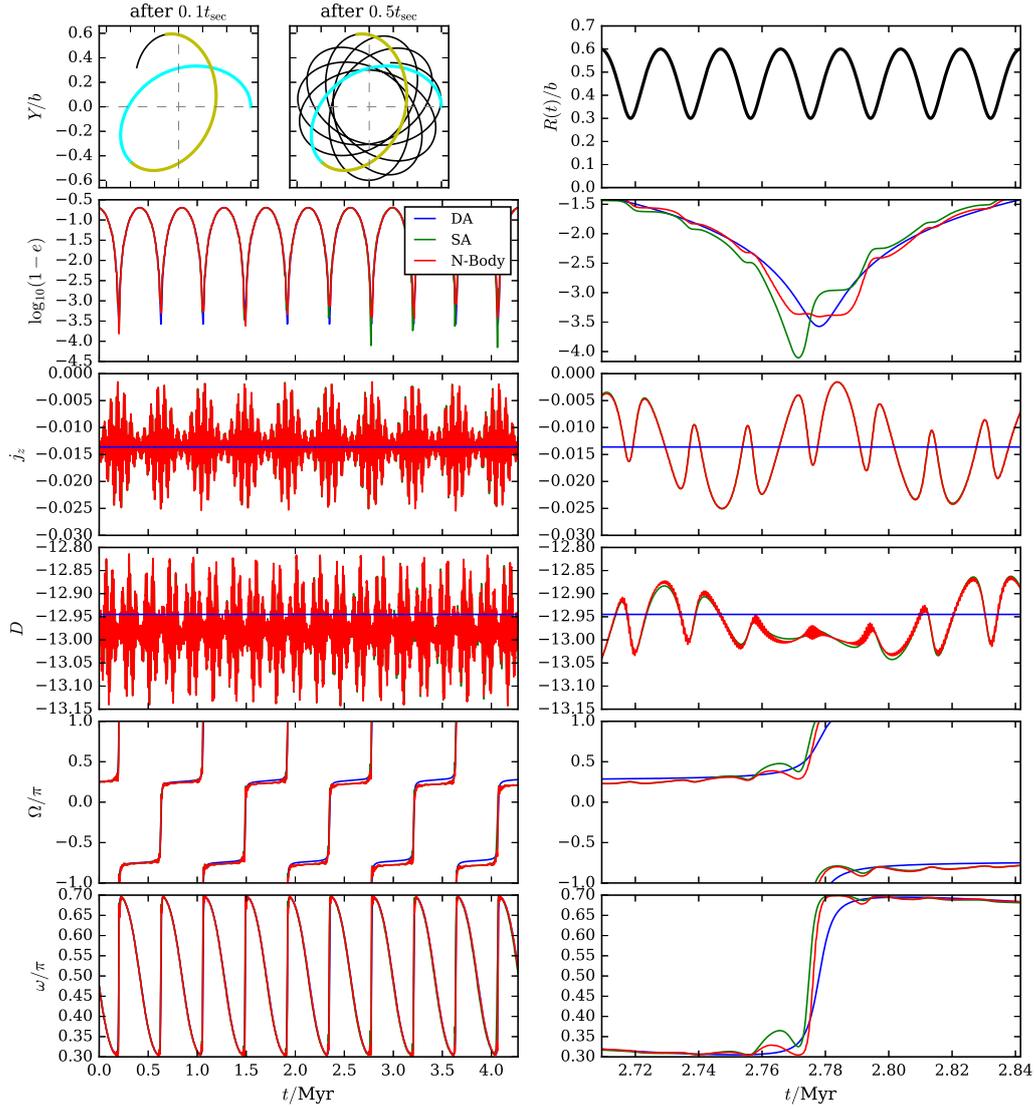


Figure 6.10: Example of SA and N-body integrations disagreeing at high eccentricity. In the right column we zoom in to the seventh eccentricity peak, where the two approaches for calculating $1 - e$ disagree by as much as ~ 0.5 dex. Here we took $T_{\phi}/\Delta t = 5000$ to be sure the results had converged.

6.5.3 Relation to LK literature

The question of how short-timescale fluctuations affect high eccentricity evolution in LK theory was first addressed by Ivanov et al. (2005), who derived the size of the angular momentum fluctuations experienced near maximum eccentricity by a binary undergoing LK oscillations. In their case the outer orbit was assumed circular, though a generalisation to eccentric outer orbits was derived by Haim & Katz (2018). Ivanov et al. (2005)'s result and other results very similar to it have since been used extensively for modelling hierarchical triples (e.g. Katz & Dong 2012; Bode & Wegg 2014; Antognini et al. 2014; Silsbee & Tremaine 2017; Grishin et al. 2018).

More recently, Luo et al. (2016) took a perturbative approach to the LK problem for arbitrary inner and outer eccentricities. Taking (implicitly) the limit $2t_{\min}/T_\phi \gg 1$, they showed that short-timescale fluctuations captured by the SA equations of motion can accumulate over many secular periods, resulting in secular evolution that does not resemble the original DA prediction. In other words, the time-averaged SA solution does not agree with the DA solution, but rather diverges from it¹⁶. Grishin et al. (2018) applied the formalism of Luo et al. (2016) to high eccentricity behaviour. Assuming a circular outer orbit they calculated the new maximum eccentricity arising from Luo et al. (2016)'s 'corrected' secular theory, as well as the magnitude of angular momentum fluctuations at highest eccentricity. Though we have not done so here, in the special case of circular outer orbits the results of Luo et al. (2016) and Grishin et al. (2018) could be trivially extended to arbitrary axisymmetric cluster potentials of the sort considered in this thesis. Unfortunately, for non-circular outer orbits we lack an analytic prescription for $\mathbf{R}_g(t)$ (Chapter 2), which renders the calculation intractable.

However, in the papers mentioned above (Ivanov et al. 2005; Luo et al. 2016; Grishin et al. 2018), GR precession was not directly included when calculating the fluctuating behaviour at high eccentricity. Those authors also all assumed the timescale separation $t_{\min}/T_\phi \gg 1$, allowing them to freeze (ω, J, Ω, J_z) on the timescale T_ϕ . Our work extends these results by including GR precession and investigating systems with $2t_{\min}/T_\phi \lesssim 1$. In this case, the RPSD effect that we have uncovered means that SA dynamics do not converge to the original DA prediction on average, just as was found by Luo et al. (2016) in the non-GR LK theory. However, unlike Luo et al. (2016)'s discovery, RPSD depends critically on the strength of GR precession and also happens essentially instantaneously (on a timescale $t_{\min} \lesssim T_\phi$) rather than accumulating over many secular periods.

Finally, we have considered only quadrupole terms in the tidal potential. The octupole terms are very small in most cases we consider, since a/R is very small. In fact, in all our

¹⁶Strictly, Luo et al. (2016) only considered the accumulated error due to the quadrupole-level SA dynamics. Their approach was generalised to arbitrary (octupole, hexadecapole, ...) order by Lei et al. (2018). The method was further extended to include fluctuations on the timescale $\sim T_{\text{in}}$ by Lei (2019).

numerical examples the octupole terms are identically zero because of the equal masses of the binary components, $m_1 = m_2$. In LK theory, octupole and higher-order terms are expected to become important when the outer orbit is significantly eccentric and the component masses are not equal, and can lead to e.g. spin flips and dynamical chaos (Li et al. 2015). We have shown that chaotic phase space behaviour can arise even at the pure quadrupole level provided 1PN GR precession is included.

6.6 Summary

In this Chapter we have investigated the role of short-timescale fluctuations of the orbital elements of a secularly evolving binary system, particularly with regard to high eccentricity behaviour. Our results can be summarised as follows.

- We investigated the pure test particle quadrupole problem without GR precession. Our findings here are a generalisation of those already found by others for the LK problem. In particular we derived a useful expression for the magnitude of angular momentum fluctuations at high eccentricity.
- We reran the same calculations with GR switched on, and observed somewhat similar behaviour except for a much stronger phase dependence. Related to this we uncovered a new effect, relativistic phase space diffusion (RPSD). RPSD kicks the binary to a new phase space contour, potentially leading to chaotic evolution and extreme eccentricities. A necessary, though insufficient, condition for RPSD to occur is that $\Theta \lesssim (T_{\text{in}}/T_{\phi})^2$.

Appendices

6.A Singly-averaged equations of motion

The full SA Hamiltonian is $H_{\text{SA}} = H_{1,\text{SA}} + H_{\text{GR}}$ where H_{GR} is given by (6.4) and $H_{1,\text{SA}}$ is given by (6.2). The corresponding SA equations of motion are

$$\begin{aligned}
\frac{d\omega}{dt} &= \frac{\partial H_{1,\text{SA}}}{\partial J} + \frac{\partial H_{\text{GR}}}{\partial J} \\
&= [L^2/(8J^4\mu^2)] \times \\
&\quad \left\{ (\Phi_{xx} + \Phi_{yy}) \left[-J^5(3 + 5\cos 2\omega) - 5JJ_z^2L^2(1 - \cos 2\omega) \right] \right. \\
&\quad + (\Phi_{xx} - \Phi_{yy}) \left[-J^5(3 + 5\cos 2\omega) \cos 2\Omega + 5JJ_z^2L^2(1 - \cos 2\omega) \cos 2\Omega \right. \\
&\quad \left. + 5J^2J_z(J^2 + L^2) \sin 2\Omega \sin 2\omega \right] \\
&\quad + \Phi_{zz} \left[-6J^5 + 10JJ_z^2L^2 + 10(J^5 - JJ_z^2L^2) \cos 2\omega \right] \\
&\quad + \Phi_{xy} \left[-10J^2J_z(J^2 + L^2) \cos 2\Omega \sin 2\omega - J^5 \sin 2\Omega(6 + 10\cos 2\omega) \right. \\
&\quad \left. + 10JJ_z^2L^2 \sin 2\Omega(1 - \cos 2\omega) \right] \\
&\quad + \Phi_{xz}(1 - J_z^2/J^2)^{-1/2} \left[J_z(6J^4 + 10J^2L^2 - 20J_z^2L^2) \sin \Omega \right. \\
&\quad + J(-20J^4 + 10J^2J_z^2 + 10J_z^2L^2) \sin 2\omega \cos \Omega \\
&\quad \left. + J_z(-10J^4 - 10J^2L^2 + 20J_z^2L^2) \cos 2\omega \sin \Omega \right] \\
&\quad + \Phi_{yz}(1 - J_z^2/J^2)^{-1/2} \left[-J_z(6J^4 + 10J^2L^2 - 20J_z^2L^2) \cos \Omega \right. \\
&\quad + J(-20J^4 + 10J^2J_z^2 + 10J_z^2L^2) \sin 2\omega \sin \Omega \\
&\quad \left. - J_z(-10J^4 - 10J^2L^2 + 20J_z^2L^2) \cos 2\omega \cos \Omega \right] \left. \right\} + \frac{CL\epsilon_{\text{GR}}}{J^2}, \tag{6.33}
\end{aligned}$$

$$\begin{aligned}
\frac{dJ}{dt} &= - \frac{\partial H_{1,\text{SA}}}{\partial \omega} \\
&= [5L^2/(2J^2\mu^2)](J^2 - L^2) \times \\
&\quad \left\{ -0.25(\Phi_{xx} + \Phi_{yy}) \left[(J^2 - J_z^2) \sin 2\omega \right] \right. \\
&\quad - 0.25(\Phi_{xx} - \Phi_{yy}) \left[(J^2 + J_z^2) \cos 2\Omega \sin 2\omega + 2JJ_z \cos 2\omega \sin 2\Omega \right] \\
&\quad + \Phi_{zz} \left[0.5(J^2 - J_z^2) \sin 2\omega \right] \\
&\quad + \Phi_{xy} \left[JJ_z \cos 2\omega \cos 2\Omega - 0.5(J^2 + J_z^2) \sin 2\omega \sin 2\Omega \right] \\
&\quad + \Phi_{xz} \left[J\sqrt{1 - J_z^2/J^2} (J \cos 2\omega \cos \Omega - J_z \sin 2\omega \sin \Omega) \right] \\
&\quad \left. + \Phi_{yz} \left[J\sqrt{1 - J_z^2/J^2} (J \cos 2\omega \sin \Omega + J_z \sin 2\omega \cos \Omega) \right] \right\}, \tag{6.34}
\end{aligned}$$

$$\begin{aligned}
\frac{d\Omega}{dt} &= \frac{\partial H_{1,SA}}{\partial J_z} \\
&= [L^2/(4J^3\mu^2)] \times \\
&\quad \left\{ (\Phi_{xx} + \Phi_{yy}) \left[0.5JJ_z(-3J^2 + 5L^2 + 5(J^2 - L^2)\cos 2\omega) \right] \right. \\
&\quad + (\Phi_{xx} - \Phi_{yy}) \left[-0.5JJ_z \cos 2\Omega(-3J^2 + 5L^2 + 5(J^2 - L^2)\cos 2\omega) \right. \\
&\quad \left. + 5J^2(J^2 - L^2)\sin 2\omega \sin 2\Omega \right] \\
&\quad - \Phi_{zz} \left[JJ_z(-3J^2 + 5L^2 + 5(J^2 - L^2)\cos 2\omega) \right] \\
&\quad - \Phi_{xy} \left[J[5J(J^2 - L^2)\cos 2\Omega \sin 2\omega + J_z(-3J^2 + 5L^2 + 5(J^2 - L^2)\cos 2\omega)\sin 2\Omega] \right] \\
&\quad + \Phi_{xz}(1 - J_z^2/J^2)^{-1/2} \left[5JJ_z(J^2 - L^2)\sin 2\omega \cos \Omega \right. \\
&\quad \left. + [3J^4 + 10J_z^2L^2 + J^2(-6J_z^2 - 5L^2) + (-5J^4 - 10J_z^2L^2 + J^2(10J_z^2 + 5L^2))\cos 2\omega] \sin \Omega \right] \\
&\quad + \Phi_{yz}(1 - J_z^2/J^2)^{-1/2} \left[5JJ_z(J^2 - L^2)\sin 2\omega \sin \Omega \right. \\
&\quad \left. - [3J^4 + 10J_z^2L^2 + J^2(-6J_z^2 - 5L^2) \right. \\
&\quad \left. + (-5J^4 - 10J_z^2L^2 + J^2(10J_z^2 + 5L^2))\cos 2\omega] \cos \Omega \right] \left. \right\}, \tag{6.35}
\end{aligned}$$

$$\begin{aligned}
\frac{dJ_z}{dt} &= - \frac{\partial H_{1,SA}}{\partial \Omega} \\
&= - [L^2/(4J^2\mu^2)] \times \\
&\quad \left\{ (\Phi_{xx} - \Phi_{yy}) \left[5JJ_z(J^2 - L^2)\cos 2\Omega \sin 2\omega + 0.5((J^2 - J_z^2)(3J^2 - 5L^2) \right. \right. \\
&\quad \left. \left. + 5(J^2 + J_z^2)(J^2 - L^2)\cos 2\omega) \sin 2\Omega \right] \right. \\
&\quad + \Phi_{xy} \left[((-J^2 + J_z^2)(3J^2 - 5L^2) - 5(J^2 + J_z^2)(J^2 - L^2)\cos 2\omega)\cos 2\Omega \right. \\
&\quad \left. + 10JJ_z(J^2 - L^2)\sin 2\omega \sin 2\Omega \right] \\
&\quad - \Phi_{xz} \left[J\sqrt{1 - J_z^2/J^2}(-5J(J^2 - L^2)\sin \Omega \sin 2\omega \right. \\
&\quad \left. + J_z(-3J^2 + 5L^2 + 5(J^2 - L^2)\cos 2\omega)\cos \Omega \right] \\
&\quad - \Phi_{yz} \left[J\sqrt{1 - J_z^2/J^2}(5J(J^2 - L^2)\cos \Omega \sin 2\omega \right. \\
&\quad \left. + J_z(-3J^2 + 5L^2 + 5(J^2 - L^2)\cos 2\omega)\sin \Omega \right] \left. \right\}. \tag{6.36}
\end{aligned}$$

It is straightforward to recover the DA equations from the SA equations (6.34)-(6.35) by replacing the time-dependent quantities $\Phi_{\alpha\beta}$ with their time-averages $\bar{\Phi}_{\alpha\beta}$, and using the identities $\bar{\Phi}_{xx} = \bar{\Phi}_{yy}$ and $\bar{\Phi}_{xy} = \bar{\Phi}_{xz} = \bar{\Phi}_{yz} = 0$.

6.B Fluctuating Hamiltonian

By subtracting the DA Hamiltonian from the SA Hamiltonian, assuming them to be functions of the same variables (J, ω, \dots) , and using $\Phi_{xx} = \bar{\Phi}_{yy}$ and $\bar{\Phi}_{xy} = \bar{\Phi}_{xz} = \bar{\Phi}_{yz} = 0$, we get an expression for the ‘fluctuating Hamiltonian’ (equation (6.11)):

$$\begin{aligned}
\Delta H &\equiv H_{\text{SA}}(J, \omega, \dots) - H_{\text{DA}}(J, \omega, \dots) \\
&= \frac{1}{2} \sum_{\alpha\beta} (\Phi_{\alpha\beta}(t) - \bar{\Phi}_{\alpha\beta}) \langle r_\alpha r_\beta \rangle_M \\
&= \frac{1}{2} \left\{ (\Phi_{xx} - \bar{\Phi}_{xx}) \langle x^2 \rangle_M + (\Phi_{yy} - \bar{\Phi}_{xx}) \langle y^2 \rangle_M + (\Phi_{zz} - \bar{\Phi}_{zz}) \langle z^2 \rangle_M \right. \\
&\quad \left. + \Phi_{xy} \langle xy \rangle_M + \Phi_{xz} \langle xz \rangle_M + \Phi_{yz} \langle yz \rangle_M \right\}. \tag{6.37}
\end{aligned}$$

(Note that the term involving ϵ_{GR} has disappeared, since it is the same in SA and DA theory). Equation (6.37) holds for binaries in arbitrary axisymmetric potentials.

We can simplify matters significantly if we restrict ourselves to spherical potentials $\Phi(r) = \Phi(\sqrt{R^2 + Z^2})$. Let us define

$$f_{\pm}(\mathbf{R}_g(t)) \equiv \frac{1}{2} \left[\left(\frac{\partial^2 \Phi}{\partial R^2} \right)_{\mathbf{R}_g} \pm \left(\frac{1}{R} \frac{\partial \Phi}{\partial R} \right)_{\mathbf{R}_g} \right], \tag{6.38}$$

and assume without loss of generality that \mathbf{R}_g is confined to $Z = 0$. Then it is easy to show (see equations (2.26)-(2.31)) that:

$$\Phi_{xx} = f_+ + f_- \cos 2\phi, \quad \Phi_{yy} = f_+ - f_- \cos 2\phi, \quad \Phi_{zz} = f_+ - f_-, \tag{6.39}$$

$$\Phi_{xy} = f_- \sin 2\phi, \quad \Phi_{xz} = \Phi_{yz} = 0. \tag{6.40}$$

(Note that we have dropped the ‘g’ subscript for ease of notation). If we also define $\Delta f_{\pm} \equiv f_{\pm} - \bar{f}_{\pm}$ where \bar{f}_{\pm} is the torus-averaged value of f_{\pm} then the fluctuating Hamiltonian can be written concisely as

$$\begin{aligned}
\Delta H &= \frac{1}{2} \left[\Delta f_+ \langle x^2 + y^2 + z^2 \rangle_M - \Delta f_- \langle z^2 \rangle_M + f_- \left(\langle x^2 - y^2 \rangle_M \cos 2\phi + 2 \langle xy \rangle_M \sin 2\phi \right) \right] \\
&= \frac{L^2}{8J^2\mu^2} \left\{ (3J^2 - 5L^2) \left[(\Delta f_- - 2\Delta f_+) J^2 - \Delta f_- J_z^2 + f_- (-J^2 + J_z^2) \cos[2(\Omega - \phi)] \right] \right. \\
&\quad \left. - 5(J^2 - L^2) \cos 2\omega \left[\Delta f_- (J^2 - J_z^2) + f_- (J^2 + J_z^2) \cos[2(\Omega - \phi)] \right] \right. \\
&\quad \left. + 10f_- J J_z (J^2 - L^2) \sin 2\omega \sin[2(\Omega - \phi)] \right\}. \tag{6.41}
\end{aligned}$$

As explained in §6.3.3, equations of motion can be derived from the fluctuating Hamiltonian $\Delta H(J, \omega, \dots)$ by taking its partial derivatives (equation (6.13)). In particular, we have

$$\begin{aligned} \frac{\partial \Delta H}{\partial \omega} = \frac{5}{4} \frac{L^2}{J^2 \mu^2} (J^2 - L^2) & \left\{ \sin 2\omega (J^2 - J_z^2) \Delta f_- + f_- \left((J^2 + J_z^2) \sin 2\omega \cos[2(\phi - \Omega)] \right. \right. \\ & \left. \left. - 2JJ_z \cos 2\omega \sin[2(\phi - \Omega)] \right) \right\}. \end{aligned} \quad (6.42)$$

Assuming this to be a good approximation to $-dj/dt$ at high e , as we do in §6.3.4, it constitutes a generalisation of equation (B4) of Ivanov et al. (2005). The result of Ivanov et al. (2005) is recovered if one assumes the outer orbit to be circular (so $\Delta f_{\pm} = 0$), the perturbing potential to be Keplerian, and evaluates (6.42) at $\omega = \pm\pi/2$ and $e \rightarrow 1$. We also have

$$\begin{aligned} \frac{\partial \Delta H}{\partial \Omega} = \frac{L^2}{4J^2 \mu^2} f_- & \left\{ 10JJ_z (J^2 - L^2) \sin 2\omega \cos[2(\phi - \Omega)] \right. \\ & \left. - ((J^2 - J_z^2)(3J^2 - 5L^2) + 5(J^2 + J_z^2)(J^2 - L^2) \cos 2\omega) \sin[2(\phi - \Omega)] \right\}, \end{aligned} \quad (6.43)$$

which coincides precisely with (minus) the right hand side of (6.36) if one assumes Φ to be spherical.

6.C Phase dependence of fluctuating behaviour

As mentioned several times in this Chapter, a very striking example of the difference between the GR and non-GR scenarios comes from examining the phase dependence of binary dynamical evolution. Let us now illustrate this further.

For simplicity let us consider binaries in spherical potentials only. For such potentials, to determine uniquely the SA evolution of a binary one must stipulate, on top of all the quantities needed for the DA equations, two key phases. These are the initial outer orbital radius R_0 and the initial difference $\phi_0 - \Omega_0$ between the outer azimuthal angle and the inner nodal angle. Without loss of generality we may fix $\phi_0 = 0$; then the two numbers we must stipulate are R_0, Ω_0 . Of course, if one wishes to study a representative sample of binaries, then R_0 must be drawn randomly from a distribution that correctly weights the time spent at each radius, i.e. $dN \propto dR_0/v_R$. Meanwhile Ω_0 can be drawn from a uniform distribution in $(0, 2\pi)$. We saw already in §6.3.2 that a simple offset of 45° between Ω_0 values — keeping $R_0 = r_a$ fixed — can lead to vastly different high eccentricity behaviour.

In Figure 6.11A we rerun the first secular period of SA evolution from Figure 6.1A (i.e. with GR switched off), except using many different values of the initial radial phase¹⁷,

¹⁷In practice we achieve this simply by performing the outer orbit integration first from $R_0 = r_a$, and then shifting the initial time of the SA integration by the appropriate amount in $(0, T_R)$.

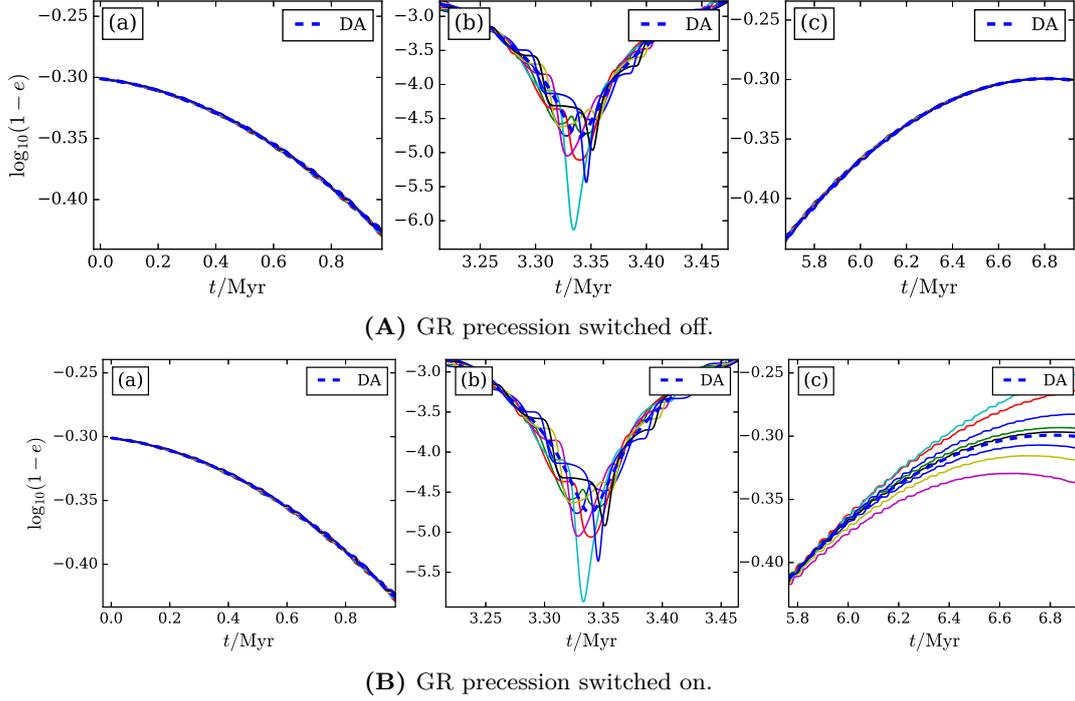


Figure 6.11: (A) We rerun the first secular period of the calculation from Figure 6.5A, except for various different values of the initial radial phase. Panels (a) and (c) show the beginning and end of the calculation, respectively, while panel (b) focuses on the eccentricity peak. The DA solution (independent of radial phase) is shown with a blue dashed line. (B) The same calculation with GR precession switched on.

keeping $\Omega_0 = 45^\circ$ fixed. We show $\log_{10}(1 - e)$ for each run (different coloured lines) at three stages of the evolution: (a) the very earliest stages near $t = 0$, (b) around the eccentricity peak, and (c) the latest stages near $t = t_{\text{sec}}$. The thick dashed blue line in each panel is the familiar DA solution from panel (n) of Figure 6.1A. While the trajectories are very close (at least in log space) in panels (a) and (c), we see that around the eccentricity peak in panel (b) the behaviour is strongly phase dependent. Some binaries do not even reach the DA maximum eccentricity $\log_{10}(1 - e_{\text{max}}) \approx -4.75$, while the binary represented by the cyan line peaks at an extremely high eccentricity, $\log_{10}(1 - e_{\text{max}}) < -6$.

Now we turn to Figure 6.11B, in which we perform exactly the same calculation except with GR switched on — in other words, we are rerunning panel (n) of Figure 6.5A for many different initial radial phases. The thick dashed blue line in each panel is the DA solution from Figure 6.5An. Comparing Figures 6.11Ab and 6.11Bb, we see that the maximum eccentricity reached by a given binary is slightly diminished by the inclusion of GR, as we would expect given that the binary resides in the weak GR regime. Regardless of this, by the time the eccentricity is returning to its minimum around 6.8 Myr (panel (c)) the various trajectories have diverged significantly. This did not happen when GR was switched off, and is a reflection of RPSD in action.

The strong phase dependence demonstrated here has not been mentioned in analogous LK calculations in the literature. Moreover, adding in the extra degree of freedom that we have not altered here — the azimuthal phase — would compound the phase dependence still further.

The work presented in this Chapter has been published in *The Astrophysical Journal Letters* as Hamilton & Rafikov (2019a).

7

Compact object binary mergers in stellar clusters

Contents

| | | |
|------------|--|------------|
| 7.1 | Introduction | 253 |
| 7.2 | Dynamical framework | 254 |
| 7.3 | Calculation of the merger fractions | 256 |
| 7.3.1 | Merger time T_m | 256 |
| 7.3.2 | Method | 257 |
| 7.3.3 | Merger fraction results | 259 |
| 7.4 | Merger rates | 261 |
| 7.4.1 | Merger rates from globular clusters | 262 |
| 7.4.2 | Merger rates from nuclear clusters | 263 |
| 7.5 | Discussion | 264 |
| 7.5.1 | Relation to existing literature | 265 |
| 7.5.2 | Further refinements | 266 |
| 7.5.3 | Summary | 267 |

7.1 Introduction

In the preceding Chapters we have shown that the smooth tidal potential of a host star cluster can drive wide binaries to perform LK-like secular eccentricity oscillations on timescales relevant for the production of LIGO/Virgo mergers. In particular, in Chapter 5 we added GW emission to our theory and demonstrated explicitly that this mechanism is capable of producing compact object binary mergers in less than a Hubble time. In the present Chapter we bring together results from all preceding Chapters, in order to explore the consequences of our mechanism for the merger rate of compact object binaries in the local Universe.

Throughout this short Chapter we make the simplifying assumption that binaries orbit spherical star clusters, and that the binary dynamical evolution is driven only by the smooth, time-independent cluster potential, truncated at quadrupole order in the tidal expansion. Thus we neglect the effects of flyby encounters, dynamical friction, evolution of the cluster, octupolar tides, and so on. Several of these are important effects that should be incorporated into future work; they are discussed further in §7.5 and §8.2. Thus, given the simplicity of the models used here, this Chapter should really be considered a ‘proof of concept’. It demonstrates that cluster tides can be important and should be taken seriously in merger calculations, but does not quantify precisely their impact relative to other key effects.

The Chapter is organised as follows. In §7.2 we recap some results for binaries in spherical clusters and outline the key assumptions of our model. Then in §7.3 we describe the approximate method we use to compute the binary merger fraction. We present the results of our merger fraction calculations in §7.3.3 and use them to calculate compact object merger rates in §7.4. We discuss our work and compare it to that of others in §7.5.

7.2 Dynamical framework

We consider a compact object binary with component masses m_1, m_2 orbiting in a fixed smooth background potential Φ of a spherically symmetric star cluster (globular or nuclear). Spherical symmetry implies that the binary’s ‘outer’ barycentric orbit is confined to a plane, and typically densely fills an axisymmetric annulus in this plane with inner and outer radii (r_p, r_a) . These two radii uniquely label an orbit in a given spherical potential $\Phi(r)$.

In our upcoming merger rate calculation, a key role will be played by the cluster tide-driven secular timescale. We know from Chapter 3 that this timescale is set by the parameter A ; roughly speaking, $t_{\text{sec}} \sim n_K/A$, where n_K is the Keplerian mean motion. The value of A is fully determined by stipulating the cluster potential Φ and the peri/apocentre (r_p, r_a) of the binary’s outer orbit. In Figure 7.1b we plot $A^*(R) \equiv A/(GM/b^3)$ assuming a circular outer orbit¹ of radius R in Plummer (cored) and Hernquist (cusped, with density $\rho \propto r^{-1}$ for $r \rightarrow 0$) potentials

$$\Phi_{\text{Plum}}(r) = -\frac{GM}{\sqrt{b_{\text{Plum}}^2 + r^2}}, \quad \Phi_{\text{Hern}}(r) = -\frac{GM}{b_{\text{Hern}} + r}, \quad (7.1)$$

where \mathcal{M} is the total mass of the cluster and $b_{\text{Plum}/\text{Hern}}$ are the corresponding scale radii. These are the two potentials that we will use throughout the Chapter; we will also fix $b_{\text{Plum}} = 1\text{pc}$ and $b_{\text{Hern}} = 0.544\text{pc}$ so that both potentials have the same half mass radius $r_h = 1.31\text{pc}$.

¹Note that the circular assumption is only used for illustration here; when we come to calculate the merger fraction we will not assume the outer orbit to be circular.

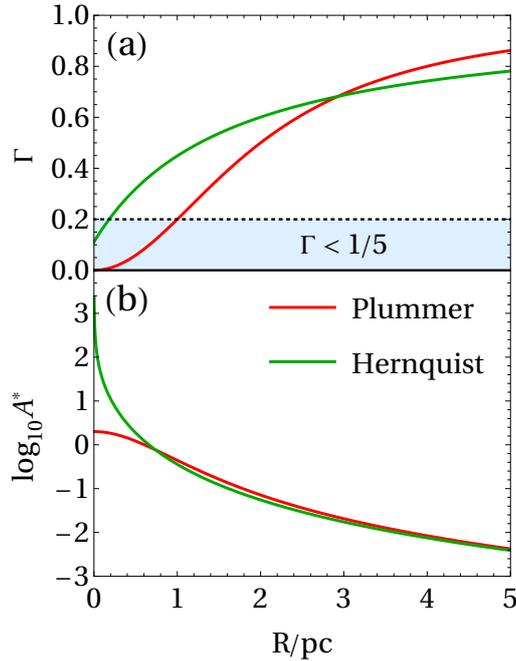


Figure 7.1: Plots of the important parameters Γ and A^* assuming the binary is on a circular outer orbit of radius R in Plummer (red) and Hernquist (green) potentials, each with scale radius chosen so that the half-mass radius is 1.31pc. For initial inclinations close to 90° , high eccentricity excitation is readily achieved when $\Gamma > 1/5$, but is much rarer when $\Gamma < 1/5$ (shaded region in panel (a)).

The second key parameter characterizing clusters and the binary’s outer orbit within them is Γ . For binaries in realistic spherical clusters we always have $0 < \Gamma \leq 1$; Figure 7.1a shows the profiles of $\Gamma(R)$ in the same Plummer and Hernquist potentials described above.

Compact object binaries are fairly simple in comparison with stellar binaries in that most short-range forces, such as tidal bulges, can be ignored entirely. However, one short-range effect that is almost always important in the present context is 1PN GR precession. Thus, to compute merger rates due to cluster tides we will rely heavily on the results of Chapter 4. In particular, we will compute a binary’s DA maximum eccentricity using the prescription given in §4.3.4, essentially by setting the right hand side of equation (4.12) to zero. Furthermore, we know from Chapter 4 that GR precession will quench the cluster tide-driven eccentricity oscillations if initially we have² $\epsilon_{\text{GR}} \gtrsim \epsilon_{\text{strong}} \equiv 3(1 + 5\Gamma)$. In practice for spherical clusters this means that that one should not expect high eccentricity oscillations to occur whenever $\epsilon_{\text{GR}} \gtrsim 10$. This requirement severely constrains the parameter space of initial conditions that can lead to cluster tide-driven mergers.

In Chapter 5 we coupled the DA equations of motion (including GR precession) to GW emission, and observed the resulting orbital decay — and ultimate merger — of

²We say ‘initially’ because of course, as a binary’s semimajor axis shrinks, its value of ϵ_{GR} increases indefinitely. It is the initial semimajor axis, and therefore the initial ϵ_{GR} , that we care about here.

compact object binaries. We saw that the associated $t_{\text{sec}}(a)$ and $a(t)$ behaviour was not universal; instead it depended on the binary's phase space orbit family (librating or circulating), the value of Γ (greater than or less than $1/5$), the relative strength of GR precession (weak, moderate or strong), and so on. We were able to predict the $t_{\text{sec}}(a)$ and $a(t)$ behaviour analytically, but only in certain asymptotic regimes. In future work one might attempt to stitch together these various asymptotic results and thereby derive an accurate semi-analytic formula for the total time taken for the binary to merge given its initial conditions. Here, however, we will skip over these details and use a much simpler formula that only relies on the results of Chapter 4. This formula (described in §7.3.1) is effectively a generalisation of the much-used result from the theory of LK-driven mergers that we discussed in §5.5.1.

Finally, from Chapter 6 we know that fluctuations in the tidal torque felt by the binary on the timescale of its outer orbital period (which are ignored by double-averaging) can increase its maximum eccentricity. Roughly (§6.3.4), one can think of these fluctuations as modifying the maximum eccentricity reached by the binary from e_{max} to $\tilde{e}_{\text{max}} = e_{\text{max}} + \delta e$, $\delta e > 0$. These short-timescale fluctuations will therefore enhance the merger rate, and we will take this effect into account in our approximate merger time formula. However, we will not take into account the other key finding of Chapter 6, namely phase space diffusion (RPSD). This is justified because, as we argued in §6.5.1, RPSD operates in such a narrow inclination window that it is likely to affect only a very small fraction of the binaries sampled here. Given that the estimates we must make in this Chapter already come with large error bars, omitting RPSD is insignificant in practice. Thus SA effects are taken into account purely in the form of fluctuations on top of the underlying smooth, and perfectly periodic, DA evolution.

7.3 Calculation of the merger fractions

The main goal of this Chapter is to compute the present day merger rate induced by cluster tides. Its calculation in §7.4 relies on knowledge of the time evolution of the *merger fraction* $f_{\text{m}}(t)$, which is found by taking a large ensemble of binaries and computing how many of them merge in a time $T_{\text{m}} < t$. Here we outline the details of the calculation, namely, our merger time prescription (§7.3.1), the method used (§7.3.2), and the results (§7.3.3).

7.3.1 Merger time T_{m}

As discussed in Chapter 1 and §5.5.1, a heuristic argument that allows us to approximate the total merger time is as follows. An isolated binary (in the absence of cluster tides)

with initial semi-major axis a_0 and eccentricity $e_0 \approx 1$ would merge due to GW emission in a time (Peters 1964):

$$T_m^{\text{iso}}(e_0) = \frac{3c^5 a_0^4}{85G^3(m_1 + m_2)m_1 m_2} (1 - e_0^2)^{7/2}. \quad (7.2)$$

However, the torque from the cluster potential causes the binary's eccentricity to vary in a cyclic fashion on a secular timescale t_{sec} , with $e \rightarrow 1$ under favorable circumstances. Because of the steep dependence of T_m^{iso} on $1 - e$, GW emission occurs in the form of discrete bursts around the sharp eccentricity maxima. Such high- e episodes last for about $\Delta t_{\text{max}} \approx t_{\text{sec}}(1 - e_{\text{max}}^2)^{1/2}$, where e_{max} is the maximum eccentricity obtained in the DA theory (see §4.4). This prolongs the time to merger (estimated using equation (7.2) at peak eccentricity) by a factor $\approx t_{\text{sec}}/\Delta t_{\text{max}} = (1 - e_{\text{max}}^2)^{-1/2}$ — see equation (7.3).

Moreover, as e passes through its peak value it also experiences short-timescale fluctuations. These variations periodically take e to its peak singly-averaged value \tilde{e}_{max} , which is higher than the DA value e_{max} . Again, because of the sharp dependence of GW emission on $1 - e$, GW losses mainly occur when $e \approx \tilde{e}_{\text{max}}$. For this reason, to approximately account for the singly-averaged effects we set the peak eccentricity determining the intensity of GW emission to \tilde{e}_{max} (rather than e_{max}) and obtain the following estimate of the merger time:

$$T_m \approx T_m^{\text{iso}}(\tilde{e}_{\text{max}}) \times (1 - e_{\text{max}}^2)^{-1/2} \quad (7.3)$$

$$\begin{aligned} &= \frac{3c^5 a_0^4}{85G^3(m_1 + m_2)m_1 m_2} \psi(e_{\text{max}}, \tilde{e}_{\text{max}}) \quad (7.4) \\ &= 1.0 \text{ Gyr} \left(\frac{m}{1.4M_\odot} \right)^{-3} \left(\frac{a_0}{10 \text{ AU}} \right)^4 \frac{\psi(e_{\text{max}}, \tilde{e}_{\text{max}})}{10^{-12}} \\ &= 0.5 \text{ Gyr} \left(\frac{m}{30M_\odot} \right)^{-3} \left(\frac{a_0}{30 \text{ AU}} \right)^4 \frac{\psi(e_{\text{max}}, \tilde{e}_{\text{max}})}{10^{-12}}, \end{aligned}$$

where $\psi(e_{\text{max}}, \tilde{e}_{\text{max}}) = (1 - \tilde{e}_{\text{max}}^2)^{7/2} (1 - e_{\text{max}}^2)^{-1/2}$. In the numerical estimates we used typical values for NS-NS and BH-BH binaries with $m_1 = m_2 = m$. Equation (7.4) is what we use in this Chapter for T_m (although see the end of §7.3.2).

7.3.2 Method

To compute the merger fraction $f_m(t)$, it is necessary that we are first able to calculate e_{max} and \tilde{e}_{max} for any binary. For a given cluster potential, both e_{max} and \tilde{e}_{max} are functions of the eight parameters that describe the inner (a, e, i, ω) and outer (r_p, r_a) orbits of the binary at $t = 0$ and the binary component masses, e.g.

$$e_{\text{max}} = e_{\text{max}}(r_p, r_a, a_0, e_0, i_0, \omega_0, m_1, m_2). \quad (7.5)$$

We obtain e_{\max} from our secular (DA) theory by following the method prescribed in Chapter 4, essentially by setting the right hand side of equation (4.12) to zero, and solving for $j_{\min} = \sqrt{1 - e_{\max}^2}$. We approximate $\tilde{e}_{\max} = \sqrt{1 - \tilde{j}_{\min}^2}$ by writing $\tilde{j}_{\min}^2 = j_{\min} - (\delta j)_{\text{p}}$, where $(\delta j)_{\text{p}}$ is the approximate fluctuation calculated in §6.3.4. (That calculation pretends that the outer orbit is circular with radius r_{p} , which is unrealistic but good enough for a rough estimate).

Then at each time t , for a given a_0, m_1, m_2 there exists a critical region in $(e_{\max}, \tilde{e}_{\max})$ space for which $T_{\text{m}} < t$ (equation (7.3)). All systems in the critical region can be considered ‘merged’ at time t . With a suitable Monte Carlo sampling of the eight parameters listed in (7.5) one can therefore compute the cumulative fraction $f_{\text{m}}(t)$ of systems that have merged as a function of time. To carry out the Monte Carlo procedure we draw a large number³ $N = 10^6$ of binaries with initial parameters randomly chosen from appropriate distributions described as follows.

Our compact object binaries come in three flavours: NS-NS, NS-BH and BH-BH. For the component masses m_1, m_2 we always use $1.4M_{\odot}$ (NS) and $30M_{\odot}$ (BH). We use three cluster masses: $\mathcal{M} = 10^5, 10^6, 10^7 M_{\odot}$. We consider two cluster potentials, the same as in Figure 7.1: the Plummer potential Φ_{Plum} to mimic cored potentials of globular clusters and the Hernquist potential Φ_{Hern} to approximate cusped nuclear clusters. Each of them is scaled to have half mass radius $r_{\text{h}} = 1.31\text{pc}$.

We randomly sample r_{p} and r_{a} (which characterize the binary’s outer orbit) from a self-consistent distribution function (DF) constructed as follows. We take the isotropic self-consistent DF $g(\tilde{\mathcal{E}}(r_{\text{p}}/b, r_{\text{a}}/b), b)$ that generates the underlying cluster potential with mass \mathcal{M} and scale radius b , where $\tilde{\mathcal{E}} \equiv \mathcal{E}/(G\mathcal{M}/b)$ and \mathcal{E} is the specific energy of an orbit in that potential. Thus, $g(\tilde{\mathcal{E}}, b_{\text{Plum}}) \propto b_{\text{Plum}}^{-3/2} (-\tilde{\mathcal{E}})^{7/2}$ for the Plummer potential, while for the Hernquist potential $g(\tilde{\mathcal{E}}, b_{\text{Hern}})$ is given by equation (4.51) of Binney & Tremaine (2008). We then draw the orbits of our binaries from a DF $\propto g(\tilde{\mathcal{E}}(r_{\text{p}}/b', r_{\text{a}}/b'), b')$, where the new scale radius b' is a parameter that we vary to account for the possibility of the massive compact object binaries being more centrally concentrated than the underlying stellar population (we leave the scale radius b of the cluster potential unchanged). We choose three values of b' such that the corresponding central *over-concentration* $c \equiv \rho(0, b')/\rho(0, b)$ — ratio of the central densities computed from the DFs $g(\tilde{\mathcal{E}}, b')$ and $g(\tilde{\mathcal{E}}, b)$ — is equal to 1, 10 and 100. Hence for $c = 1$ the binaries are essentially tracer particles drawn from the underlying stellar population, while for $c \gg 1$ they are much more centrally concentrated. In the Plummer case this requires $b'/b_{\text{Plum}} = 1, 10^{-1/3}$ and $10^{-2/3}$, while for the Hernquist sphere we must take $b'/b_{\text{Hern}} = 1, 10^{-1/2}$ and 10^{-1} . Variation of c helps to alleviate the observational uncertainty in the radial distribution of compact object binaries in clusters.

³We checked that a ‘higher resolution’ calculation which sampled $N = 10^7$ binaries gave essentially identical results.

We assume Opik’s law for the distribution of binary semi-major axes ($dN/da_0 \propto a_0^{-1}$), sampling it in the range $a_0 \in (a_{\min}, a_{\max})$. Here a_{\min} is the semi-major axis below which GR precession will suppress cluster tide-driven evolution; we estimate a_{\min} by solving equation (4.6) for a with $\epsilon_{\text{GR}} = 10$ and $A^* = 1.0$. We take $a_{\max} = 50 \text{ AU}, 100 \text{ AU}, 100 \text{ AU}$ for NS-NS, NS-BH and BH-BH binaries respectively, expecting that wider binaries would be quickly disrupted by stellar encounters. Initial binary eccentricities are drawn from a thermal distribution (uniform in e_0^2) in the range $e_0 \in (0.01, 0.995)$.

We assume random orientation of the binaries, implying that the initial pericentre angles ω_0 and initial cosines of inclination $\cos i_0$ are uniformly distributed in $(-\pi, \pi)$ and $(0, 1)$ respectively. However, the symmetry of the problem means that we may restrict the random sampling of ω_0 to the range $(0, \pi)$, allowing us to speed up the calculation. Moreover, only binaries with initial inclinations i_0 close to 90° are able to merge within a Hubble time, as follows from the conservation of $(1 - e^2)^{1/2} \cos i$ and the fact that very high eccentricities ($e_{\max} \rightarrow 1$) are required to enhance GW emission. Hence it is sufficient to sample $\cos i_0$ from a uniform distribution not in $(0, 1)$ but $(0, \Xi)$, where we took $\Xi = 0.05, 0.08, 0.1$ for NS-NS, NS-BH and BH-BH binaries, respectively⁴.

When calculating merger fractions f_m we account for the aforementioned truncation of the ranges of $a_0, e_0, \omega_0, \cos i_0$. In particular we assume that the overall population of binaries has a minimum semi-major axis 0.2 AU (whereas it is only sampled down to a_{\min}) while the maximum semi-major axis is still a_{\max} , and weight the number of merged binaries accordingly⁵. Similarly, in reality $\cos i_0 \in (0, 1)$, but binaries in $(\Xi, 1)$ never merge. The values of $f_m(t)$ we quote always reflect the fraction of the *total* population that has merged in time t , not just of the initial N sampled binaries.

Implicit in the derivation of the merger time T_m is the assumption that the binary undergoes at least one secular cycle by time t . However, equation (7.4) sometimes predicts merger times that are short compared to the secular timescale t_{sec} . In particular, if ever $(\delta j)_p > j_{\min}$ we interpret this as an ‘instantaneous’ merger (taking place in much less than one secular timescale). Since binaries must first *reach* their maximum eccentricity before they can actually merge, which on average takes $\approx t_{\text{sec}}/2$, we account for these ‘fast’ mergers by taking the actual merger time to be $\max(T_m, t_{\text{sec}}/2)$.

7.3.3 Merger fraction results

In Figure 7.2 we plot the cumulative merger fractions $f_m(t)$ for $t \in (1 \text{ Myr}, 12 \text{ Gyr})$, calculated using the method of §7.3.2. We consider NS-NS (left column), NS-BH (middle

⁴The Ξ values are calculated by putting $a_0 = a_{\min}$, $T_m = 12 \text{ Gyr}$ and $1 - e_{\max}^2 \sim \cos^2 i_0$ in equation (7.4) and solving for $\cos i_0$.

⁵Note that the lower limit we take for the semimajor axis distribution, 0.2 AU , is roughly the maximum semimajor axis for which an isolated, circular $m_1 = m_2 = 30M_\odot$ BH-BH binary will merge in a Hubble time in the absence of external perturbations. For NS-BH and NS-NS binaries that value is even smaller. Thus, by making this choice we are essentially excluding from our sample any binaries that would have merged on their own within a Hubble time, i.e. without any assistance from cluster tides.

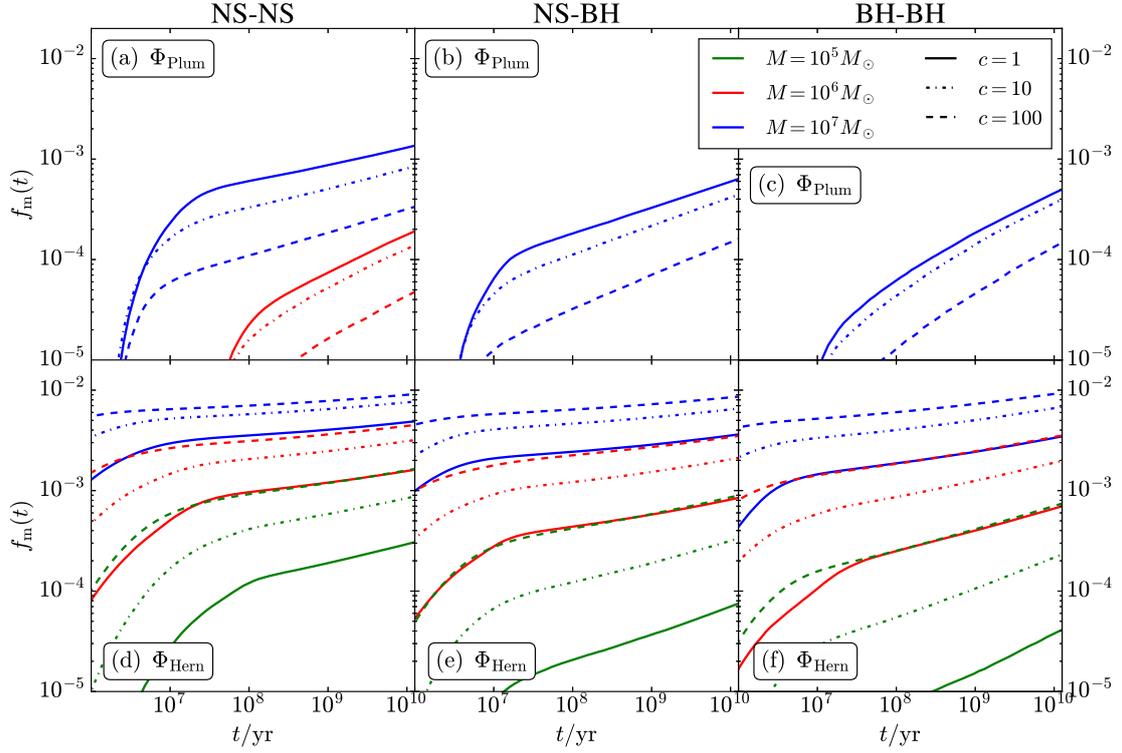


Figure 7.2: Cumulative merger fraction $f_m(t)$ over the domain $t \in (1 \text{ Myr}, 12 \text{ Gyr})$ for NS-NS, NS-BH and BH-BH binaries, each for cluster masses $\mathcal{M}/M_\odot = 10^5, 10^6, 10^7$ and binary central concentrations $c = 1, 10, 100$ in the Plummer and Hernquist potentials (see legend).

column), and BH-BH (right column) binaries, each for $\mathcal{M} = 10^5 M_\odot$ (green), $\mathcal{M} = 10^6 M_\odot$ (red) and $\mathcal{M} = 10^7 M_\odot$ (blue) clusters and concentrations $c = 1, 10, 100$ (solid, dot-dashed and dashed lines respectively), for the two potentials (7.1).

Cored (Plummer) models

Starting with the Plummer models (top row of Figure 7.2), we see that f_m is largest for the most massive clusters ($\mathcal{M} = 10^7 M_\odot$, blue lines) because the secular evolution is fastest in such clusters and therefore large eccentricity oscillations are less easily quenched by GR precession. For NS-NS binaries with central concentration $c = 1$, the final merger fraction is $f_m(12 \text{ Gyr}) \sim 10^{-3}$ in $\mathcal{M} = 10^7 M_\odot$ clusters. The corresponding result for NS-BH and BH-BH binaries is a factor of a few smaller because of the stronger GR precession barrier for these more massive systems. In $\mathcal{M} = 10^6 M_\odot$ clusters (red lines), we again find a non-negligible final NS-NS merger fraction, $f_m(12 \text{ Gyr}) \sim 10^{-4}$; however, we find no NS-BH and BH-BH mergers, because for those (heavy) binaries the cluster tides are no longer strong enough to beat the GR precession. For the same reason, f_m is negligible in cored (Plummer) $\mathcal{M} = 10^5 M_\odot$ clusters across all binary flavours.

In all three panels, increasing the central concentration c reduces the merger fraction because strongly centrally concentrated binaries in cored potentials fall into the $\Gamma < 1/5$

regime (see Figure 7.1) for which high eccentricity excitation is suppressed (see Chapters 3 and 4). Mass segregation of a population of heavy binaries would act to steadily increase $c(t)$ over the age of the cluster. In cored clusters this would lead to a lower merger fraction at late times compared to an unsegregated population.

Cusped (Hernquist) models

Cusped clusters represented by a Hernquist potential (bottom row of Figure 7.2) exhibit substantially higher f_m values than in the Plummer case. Indeed, even $10^5 M_\odot$ clusters (green curves) — which produced zero mergers in the Plummer potential — now have $f_m(12 \text{ Gyr})$ of at least a few $\times 10^{-5}$ and often as large as $\sim 10^{-3}$, depending on c and the binary type. Moreover, increasing c in these potentials *increases* f_m , which is the opposite trend to the Plummer case. As a result, mass segregation in cusped clusters would tend to additionally increase f_m at late times.

Both effects are due to the ubiquity of the $\Gamma > 1/5$ regime (promoting high e excitation) in the Hernquist potential, even near the cluster centre (Figure 7.1) — unlike in the Plummer case, there is little disadvantage to binaries being centrally concentrated. Moreover, secular evolution is fast near the centre of the Hernquist sphere ($t_{\text{sec}} \propto A^{-1}$ and the ‘tidal strength’ A diverges, see Figure 7.1), and short-timescale fluctuations there are strong. As a result, increasing c drives more binaries to merge within a Hubble time. Many binaries that orbit near the centres of cuspy clusters have $t_{\text{sec}} < 10^6 \text{ yr}$ — hence, several curves show nonzero $f_m(10^6 \text{ yr})$.

Also, f_m shows a weaker dependence on cluster mass \mathcal{M} than in the Plummer case. This is because of the large A values in the Hernquist case (see Fig. 7.1b), which act to suppress the effect of GR precession: equation (4.5) then yields $\epsilon_{\text{GR}} \rightarrow 0$, a limit in which e_{max} is independent of \mathcal{M} (equation (4.45)).

7.4 Merger rates

Our results on merger fractions $f_m(t)$ allow us to calculate the *specific merger rate* \mathcal{R} , which is the rate of compact object binary mergers of a given flavour per unit volume in the local universe, given the birth history of binaries of that type. The latter is described by the formation rate of such binaries per unit cluster mass $W(t)$, such that in the interval $(t, t + \delta t)$, a total of $W(t)\delta t$ systems are produced per unit cluster mass. The cumulative number of mergers from that binary type *per unit cluster mass* after time t is then

$$\mathcal{C}(t) \equiv \int_0^t dt' W(t') f_m(t - t'), \quad (7.6)$$

and the corresponding contribution to the specific merger *rate* at time t is $\mathcal{R} = \rho_{\text{cl}} d\mathcal{C}(t)/dt$, where ρ_{cl} is the cluster mass density in the local universe.

We consider two simple histories of compact object binary formation. The first takes the form of a burst, so that at $t = 0$ each cluster instantaneously forms a population of binaries. If X_{born} compact object binaries are born per unit cluster mass, then $W(t) = X_{\text{born}}\delta(t)$ so that $\mathcal{C}(t) = X_{\text{born}}f_{\text{m}}(t)$ and

$$\mathcal{R}(t) = X_{\text{born}}\rho_{\text{cl}}\frac{df_{\text{m}}(t)}{dt}. \quad (7.7)$$

The second model assumes a constant compact object binary formation rate $W(t) = Y_{\text{form}}$ per unit cluster mass. Then the cumulative merger number from that cluster is $\mathcal{C}(t) = Y_{\text{form}}\int_0^t dt' f_{\text{m}}(t - t') = Y_{\text{form}}\int_0^t dx f_{\text{m}}(x)$, resulting in the specific merger rate

$$\mathcal{R}(t) = Y_{\text{form}}\rho_{\text{cl}}f_{\text{m}}(t). \quad (7.8)$$

The results obtained for these two binary formation histories give an idea of the outcomes of more sophisticated models.

7.4.1 Merger rates from globular clusters

Globular clusters have cored density profiles, so to represent them we use f_{m} results for Plummer spheres (§7.3.3). Since globulars have a range of masses and f_{m} is a function of \mathcal{M} , appropriate averaging of the rates (7.7)-(7.8) over the cluster mass spectrum is needed. Following Rodriguez, Morscher, et al. (2015) we use a log-normal mass function for the number density of globulars (Harris et al. 2014):

$$\frac{dn_{\text{gc}}}{d\log_{10}(\mathcal{M}/M_{\odot})} = \frac{n_{\text{gc}}^{\text{tot}}}{\sqrt{2\pi}\sigma_{\mathcal{M}}} \times \exp\left[-\frac{(\log_{10}(\mathcal{M}/M_{\odot}) - \mu)^2}{2\sigma_{\mathcal{M}}^2}\right], \quad (7.9)$$

where $n_{\text{gc}}^{\text{tot}}$ is the total number density of globular clusters in the local universe integrated over \mathcal{M} , and $\sigma_{\mathcal{M}} = 0.52$, $\mu = 5.54$. The number density $n_{\text{gc}}^{\text{tot}}$ is an uncertain quantity (Portegies Zwart & McMillan 2000; Rodriguez, Morscher, et al. 2015; Rodriguez, Chatterjee, et al. 2016). In this Chapter, guided by existing estimates, we adopt $n_{\text{gc}}^{\text{tot}} = 3 \text{ Mpc}^{-3}$.

For simplicity, we do the averaging in an approximate fashion by splitting the cluster population into 3 mass bins $\mathcal{M}_i^{\text{min}} < \mathcal{M} < \mathcal{M}_i^{\text{max}}$, $i = 1, 2, 3$, where $\mathcal{M}_i^{\text{min}} = 5 \times 10^{3+i} M_{\odot}$ and $\mathcal{M}_i^{\text{max}} = 5 \times 10^{4+i} M_{\odot}$. The mass density in clusters in each mass bin is then $\rho_{\text{gc},i} = \int_{\mathcal{M}_i^{\text{min}}}^{\mathcal{M}_i^{\text{max}}} \mathcal{M} dn_{\text{gc}} = (3.9, 14.1, 3.3) \times 10^5 (n_{\text{gc}}^{\text{tot}}/3 \text{ Mpc}^{-3}) M_{\odot} \text{ Mpc}^{-3}$. We assign to each bin the value of f_{m} computed for Plummer models with $\mathcal{M} = \mathcal{M}_i = 10^{4+i} M_{\odot}$ (within the i -th bin). Then averaging of the merger rate over the distribution of \mathcal{M} amounts to replacing $\rho_{\text{cl}}f_{\text{m}}$ with

$$F_{\text{m}}(t) = \sum_{i=1}^3 \rho_{\text{gc},i} f_{\text{m}}(t; \mathcal{M}_i). \quad (7.10)$$

We now compute the present day rate \mathcal{R} for the two aforementioned binary birth histories.

Merger rates from globular clusters: a single burst of compact object binary formation

Globular clusters experience a large starburst at their formation. Compact objects get produced in supernova explosions shortly thereafter. If they remain bound and assemble into binaries on a timescale short compared to the Hubble time, then the single burst approximation (7.7) should characterize the current merger rate \mathcal{R} reasonably well.

Motivated by the calculations of Löckmann et al. (2010), in this Chapter we adopt $X_{\text{born}} = 10^{-3} M_{\odot}^{-1}$ for the specific birth rate of all compact binary species, similar to the value obtained in Rodriguez, Chatterjee, et al. (2016). We calculate the total merger rate using equation (7.7), averaging it over cluster mass via equation (7.10):

$$\begin{aligned} \mathcal{R} &= X_{\text{born}} \frac{dF_{\text{m}}}{dt} \\ &= 3 \times 10^{-3} \text{ Gpc}^{-3} \text{ yr}^{-1} \times \left(\frac{X_{\text{born}}}{10^{-3} M_{\odot}^{-1}} \right) \left(\frac{dF_{\text{m}}/dt|_{12 \text{ Gyr}}}{3 M_{\odot} \text{ Mpc}^{-3} \text{ Gyr}^{-1}} \right) \end{aligned} \quad (7.11)$$

where in the numerical estimate we assumed that the formation burst happened 12 Gyr ago, and took a value of dF_{m}/dt characteristic of Plummer models (§7.3.3).

Merger rates from globular clusters: a constant rate of compact object binary formation

An alternative birth history is the one in which the assembly of compact objects into binaries in globulars occurred at a steady (slow) rate Y_{form} over the last 12 Gyr. Here we adopt $Y_{\text{form}} = 10^{-4} M_{\odot}^{-1} \text{ Gyr}^{-1}$ so that upon integration over a Hubble time we reproduce roughly the specific compact binary occurrence rate X_{born} assumed in §7.4.1 (i.e. $Y_{\text{form}} \times 10 \text{ Gyr} = X_{\text{born}}$). Then from equation (7.8) the merger rate is

$$\begin{aligned} \mathcal{R} &= Y_{\text{form}} F_{\text{m}} \\ &= 0.3 \text{ Gpc}^{-3} \text{ yr}^{-1} \times \left(\frac{Y_{\text{form}}}{10^{-4} M_{\odot}^{-1} \text{ Gyr}^{-1}} \right) \left(\frac{F_{\text{m}}(12 \text{ Gyr})}{3 \times 10^3 M_{\odot} \text{ Mpc}^{-3}} \right) \end{aligned} \quad (7.12)$$

and again we took $F_{\text{m}}(12 \text{ Gyr})$ values characteristic of Plummer models (§7.3.3).

7.4.2 Merger rates from nuclear clusters

In the case of nuclear clusters we expect compact object binaries to be created at a relatively steady rate due to continuous star formation over long times (Figer et al. (2004); dynamical assembly due to 3-body processes is not as important here, although see Muno et al. (2005)). Thus, the constant formation rate assumption is more appropriate for nuclear clusters, and we again assume $Y_{\text{form}} = 10^{-4} M_{\odot}^{-1} \text{ Gyr}^{-1}$ for these systems.

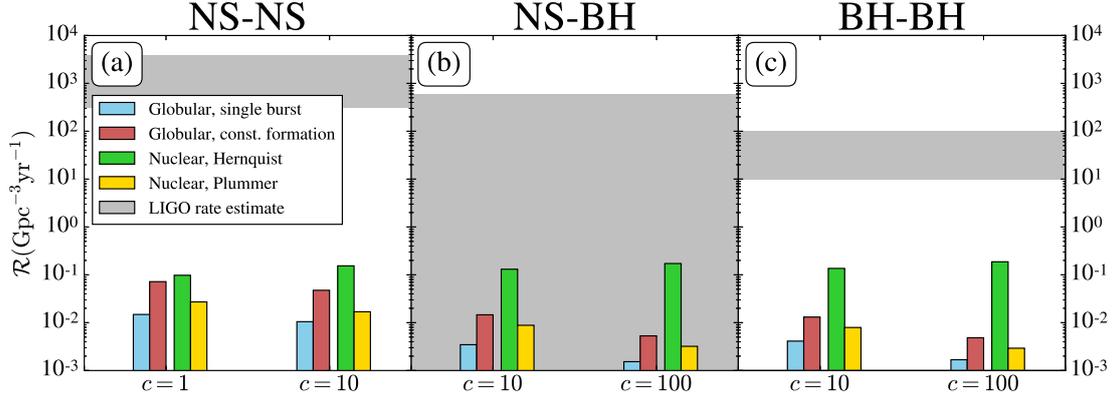


Figure 7.3: Merger rates \mathcal{R} of compact object binaries driven by the tidal fields of (spherical) globular and nuclear clusters. For each binary type we consider two values of the central concentration c . For globular clusters, modeled as cored (Plummer) systems, we look at two binary birth histories: single burst (blue) and constant formation rate (red). For nuclear clusters we calculate rates assuming either cored (Plummer, yellow) or cusped (Hernquist, green) profiles, considering only the constant binary formation history. Grey regions show the LIGO/Virgo merger rate estimates. See text for details.

For simplicity, we take all nuclear clusters to have mass $\mathcal{M}_{\text{nc}} = 10^7 M_{\odot}$ and assume $n_{\text{nc}} = 0.02 \text{ Mpc}^{-3}$ for their number density (Petrovich & Antonini 2017; Hamers, Bar-Or, et al. 2018a). Then $\rho_{\text{cl}} f_{\text{m}} = \mathcal{M}_{\text{nc}} n_{\text{nc}} f_{\text{m}}(\mathcal{M}_{\text{nc}})$ and the merger rate becomes (equation (7.8))

$$\mathcal{R} = Y_{\text{form}} \mathcal{M}_{\text{nc}} n_{\text{nc}} f_{\text{m}}(\mathcal{M}_{\text{nc}}) \quad (7.13)$$

$$= 0.2 \text{ Gpc}^{-3} \text{ yr}^{-1} \times \left(\frac{Y_{\text{form}}}{10^{-4} M_{\odot}^{-1} \text{ Gyr}^{-1}} \right) \left(\frac{n_{\text{nc}}}{0.02 \text{ Mpc}^{-3}} \right) \left(\frac{f_{\text{m}}(12 \text{ Gyr}; \mathcal{M}_{\text{nc}})}{10^{-2}} \right), \quad (7.14)$$

where for $f_{\text{m}}(12 \text{ Gyr}; \mathcal{M}_{\text{nc}})$ we adopted a value characteristic of cusped (Hernquist) models — see §7.3.3. Cored nuclear clusters have $f_{\text{m}}(12 \text{ Gyr}; \mathcal{M}_{\text{nc}})$ an order of magnitude lower, see §7.3.3.

7.5 Discussion

In Figure 7.3 we show present day compact binary merger rates due to cluster tides in globular and nuclear clusters. Rates for globulars use the results we obtained for Plummer models (for two birth histories, §§7.4.1-7.4.1), while for nuclear clusters we consider both Hernquist and Plummer models and a flat binary formation history (§7.4.2). For NS-NS binaries we consider only moderate concentrations $c = 1, 10$, while for (significantly heavier) NS-BH and BH-BH binaries we assumed a higher degree of central segregation, $c = 10, 100$. The grey regions in Figure 7.3 show the LIGO/Virgo rate estimates after O2 (The LIGO Scientific Collaboration et al. 2019): $110 - 3840 \text{ Gpc}^{-3} \text{ yr}^{-1}$ and $9.7 - 101 \text{ Gpc}^{-3} \text{ yr}^{-1}$ for NS-NS and BH-BH mergers in the local universe respectively, while the upper limit on the NS-BH merger rate is $610 \text{ Gpc}^{-3} \text{ yr}^{-1}$.

Focusing first on globular clusters, one can see that their merger rates fall short of providing a substantial contribution to the observed rates. We find $\mathcal{R} \sim 0.01 - 0.07 \text{ Gpc}^{-3} \text{ yr}^{-1}$ for NS-NS binaries and $\lesssim 0.02 \text{ Gpc}^{-3} \text{ yr}^{-1}$ for each of NS-BH and BH-BH binaries in globulars (Figure 7.3). The primary reason for fewer NS-BH and BH-BH mergers compared to NS-NS mergers is that the heavier binaries (i) suffer from stronger GR precession which cannot be overcome in a cored potential even at the cluster center, (ii) have higher central concentrations which brings them into the $\Gamma < 1/5$ regime, where high eccentricity excitation is suppressed (higher c always leads to lower \mathcal{R} in globulars). Also, a constant binary formation rate results in higher \mathcal{R} because many binaries merge soon after their birth: $f_m(t)$ curves rise substantially faster during the first $10^7 - 10^8$ yr, see Figure 7.2.

As for nuclear star clusters, if we assume a cusped density profile (Hernquist model) then $\mathcal{R} \sim 0.1 - 0.2 \text{ Gpc}^{-3} \text{ yr}^{-1}$ for NS-NS, NS-BH and BH-BH binaries. The NS-BH and BH-BH binaries merge *slightly* more often than NS-NS binaries because near the centre of cusped clusters the $\Gamma < 1/5$ regime is rare, and the tidal field is strong which helps to overcome GR precession. As a consequence, higher central concentration is advantageous (although not dramatically). However, in cored nuclear clusters the situation is more similar to that in globulars and \mathcal{R} drops appreciably with increasing c .

Overall, we see that NS-BH and BH-BH merger rates are very similar, assuming they are formed in equal numbers. Cusped nuclear clusters dominate the cluster tide-driven merger rate compared to globulars for all binary species. Whereas cluster tides acting alone are unlikely to produce many NS-NS mergers, they can still contribute at the level of several per cent to the observed NS-BH and BH-BH merger rates, given our assumptions.

7.5.1 Relation to existing literature

There are a number of existing estimates of compact object binary merger rates in globular and nuclear clusters (Antonini, Murray, et al. 2014; Stephan et al. 2016; Antonini, Chatterjee, et al. 2016; Fragione & Bromberg 2019). The studies which bear closest resemblance to our work consider binaries orbiting SMBHs at the centres of nuclear clusters and undergoing LK-driven evolution (Antonini & Perets 2012; Prodan et al. 2015; Petrovich & Antonini 2017; Hoang et al. 2018; Bub & Petrovich 2020; Arca Sedda 2020). However, apart from the last two, none of these studies accounted for the direct tidal torque on the inner orbit due to the cluster potential as we do here⁶. Additionally, in these studies the distribution of binary outer orbits is typically truncated at radii of $\lesssim 0.1 \text{ pc}$ from the cluster centre. We do not rely on the presence of a central black hole and still find mergers (out to much larger radii) by including a cluster potential.

⁶The work by Bub & Petrovich (2020) and Arca Sedda (2020) is discussed further in §8.2.

In nuclear clusters our BH-BH merger rate $\mathcal{R} \sim 0.1 - 0.2 \text{ Gpc}^{-3} \text{ yr}^{-1}$ is comparable to (but typically slightly smaller than) those of others, e.g. Antonini & Rasio (2016) ($\mathcal{R} \sim 1 \text{ Gpc}^{-3} \text{ yr}^{-1}$ from nuclear clusters without a SMBH), Petrovich & Antonini (2017) ($\mathcal{R} \sim 0.6 - 15 \text{ Gpc}^{-3} \text{ yr}^{-1}$ from non-spherical nuclear clusters with a SMBH, but they use higher Y_{form}). In globulars our BH-BH rate $\mathcal{R} \lesssim 0.02 \text{ Gpc}^{-3} \text{ yr}^{-1}$ is significantly smaller than those of e.g. Rodriguez, Chatterjee, et al. (2016) ($\mathcal{R} \sim 2 - 20 \text{ Gpc}^{-3} \text{ yr}^{-1}$ from hardening of dynamically formed binaries), see §7.4.1.

For NS-BH and NS-NS binaries in (cusped) nuclear clusters our rates, $\mathcal{R} \sim 0.1 - 0.2 \text{ Gpc}^{-3} \text{ yr}^{-1}$, are comparable to or greater than those of Petrovich & Antonini (2017) ($\mathcal{R} \sim 0.02 - 0.4 \text{ Gpc}^{-3} \text{ yr}^{-1}$ and $\mathcal{R} \lesssim 0.02 \text{ Gpc}^{-3} \text{ yr}^{-1}$ respectively). Our results are also comparable to those of Hamers, Bar-Or, et al. (2018a) who found a combined merger rate for all compact object binary flavours in nuclear clusters with SMBHs of $\mathcal{R} \sim 0.02 - 0.4 \text{ Gpc}^{-3} \text{ yr}^{-1}$.

Like most other dynamical merger channels, the rates produced by our mechanism fall short of those observed by LIGO by at least one order of magnitude.

7.5.2 Further refinements

Since the purpose of this Chapter is simply to demonstrate that cluster-tide driven mergers can contribute to the LIGO merger rate at a significant level, we defer a careful investigation of the parameter space to a more comprehensive future study (see §8.2). Apart from some technical simplifications used in this study (e.g. our approximation of T_{m} using equation (7.4), simple analytical estimate for the fluctuation δe , etc.), we have also deliberately omitted certain physical ingredients to focus on mergers arising due to secular effects alone.

Perhaps most crucially, we ignored the impact of flyby encounters on the binary's inner orbital elements (Heggie & Rasio 1996; Hamers 2018b). This is an important effect that can influence our results in non-trivial ways. Recently, Samsing et al. (2019) found that numerous distant flybys can systematically increase the number of binary mergers in stellar clusters (although they did not account for secular tide-driven evolution). Heisler & Tremaine (1986), in their study of Oort comet dynamics, found that stellar flybys contribute a significant portion of the torque at high eccentricity — in fact, the Oort comets exhibit a coupled behavior in which their orbital elements roughly follow a smooth, secular (Galactic tide-driven) trajectory on average, while simultaneously exhibiting a random walk in phase-space because of stochastic flyby encounters. We expect a similar behaviour to hold in our case, and will explore it in future work — see §8.2 for further discussion.

We also neglected time-dependence of the cluster properties, e.g. due to core collapse or disk shocking, and ignored the relaxation of the binary's outer orbit e.g. due to vector resonant relaxation (VanLandingham et al. 2016; Hamers, Bar-Or, et al. 2018a) or

dynamical friction (Arca Sedda 2020). In particular, mass segregation of heavy binaries would boost the central concentration c , which can increase or decrease merger fractions depending on the cluster potential and the level of concentration (§7.3). However, we note that our merger rates are often only mildly affected by variation of c (Figure 7.3). In addition we neglected the time-dependence of the number density of globular clusters in the local Universe. For instance, it is likely that our Galaxy used to contain 10-100 times as many globular clusters as it does now (Gnedin et al. 2014), and they have since been previously disrupted. Applying this argument to external galaxies might lead to a higher effective value for $n_{\text{gc}}^{\text{tot}}$, increasing the predicted merger rates, particularly at redshifts larger than 1 (Fragione & Kocsis 2018; Rodriguez & Loeb 2018).

Furthermore, to focus on the tidal effect of the smooth cluster mass distribution alone, in this Chapter we ignored the possibility of a central SMBH which could reside in nuclear clusters. Similarly, we assumed each cluster to be perfectly spherically symmetric, omitting the effects of possible oblateness on the outer orbit (Petrovich & Antonini 2017; Bub & Petrovich 2020). Again, to make our model realistic it is crucial to account for these effects and others (8.2).

7.5.3 Summary

In this Chapter we explored a new channel for producing compact object mergers in dense stellar clusters which relies on the secular evolution of binaries driven by the cluster's tidal gravitational field (a field which is unavoidably present in any merger model involving clusters). To summarise:

- We computed merger rates due to this mechanism by focusing on conditions in which the binary can be driven to such high eccentricity that GW emission becomes important, while fully accounting for the detrimental effect of GR precession.
- We showed that stellar systems with cored potentials (e.g. globular clusters) do not produce many mergers, owing to the inefficiency of high-eccentricity excitation in the cluster cores.
- Cusped nuclear clusters (even in the absence of a central SMBH) are significantly more effective and lead to observationally interesting merger rates.
- Our results are consistent with the (very poorly constrained) LIGO estimate for the NS-BH merger rate. Otherwise, our merger rates come closest to meeting current LIGO estimates for BH-BH binaries but still fall short by more than an order of magnitude.

On the other hand, we note that all current rate estimates — including ours — have (systematic) error bars of at least an order of magnitude. In a future study we shall explore the sensitivity of our results to variation of the underlying assumptions; in particular we will study the impact of the presence of a central SMBH on the merger rates in nuclear clusters (§8.2).

8

Conclusion

Contents

| | | |
|------------|--|------------|
| 8.1 | Summary | 269 |
| 8.2 | Future work | 271 |
| 8.2.1 | Flyby encounters | 272 |
| 8.2.2 | Asphericity of clusters, central supermassive black holes, and relaxation of the outer orbit | 274 |
| 8.2.3 | Population synthesis tools | 277 |

In this thesis we have concerned ourselves with the secular dynamics of binaries in stellar clusters. We said in §1.2 that the thesis would be focused on three questions. Below, we reproduce these questions and summarise the answers that we have arrived at throughout the preceding chapters (§8.1). Finally, to conclude the thesis we consider directions for future work (§8.2).

8.1 Summary

The first question we asked in §1.2 was:

- what happens to (idealised, point-mass, Newtonian) binary systems when they are perturbed weakly by an external body (e.g. tertiary companion, cluster, or galaxy) to which they are gravitationally bound?

This question can be of great importance when considering the origin of various exotic objects harboured by stellar clusters and galaxies. Chapters 2 and 3 were dedicated to providing the most general and comprehensive answer to this question to date.

In Chapter 2 we formulated the most general possible theory of tide-driven secular evolution of two bound point masses in arbitrary axisymmetric host systems (‘clusters’). We showed that the dynamical evolution of a given binary is governed by a secular, or ‘doubly-averaged’ (DA), Hamiltonian; here DA refers to time-averaging over both the inner Keplerian orbit of the binary and its outer orbit around the cluster. All information about the background cluster potential and the binary’s orbit within it is encapsulated in just two parameters, A and Γ , the first of which merely sets the secular timescale. The famous Lidov-Kozai (LK) problem is a special case of the theory corresponding to $\Gamma = 1$.

In Chapter 3 we explored in detail the dynamics arising from the DA Hamiltonian, for different values of Γ . For $\Gamma > 1/5$ we found that the phase-space structure and binary dynamics are qualitatively similar to LK theory, but that the behaviour changes dramatically outside of this regime because of bifurcations at $\Gamma = 1/5, 0, -1/5$. We calculated the important dynamical characteristics such as the secular evolution timescale, the maximum eccentricity, etc. for binaries in each Γ regime, and verified our theoretical results using numerical simulations.

Overall, these two Chapters provide the analytical bedrock upon which the remainder of the thesis is built.



The second question posed in §1.2 was:

- how is the resulting dynamical evolution impacted by non-Newtonian and non-secular effects, such as GR precession, GW emission and short-timescale fluctuations in the torque?

The three additional effects mentioned here — GR precession, GW emission and short-timescale fluctuations — were the focus of Chapters 4, 5 and 6 respectively. A recurrent theme in these three chapters was the uncovering of new results about LK theory that had not been appreciated before despite decades of work and hundreds of papers on the topic.

In Chapter 4 we generalised the DA secular theory to include GR pericentre precession of the binary’s inner orbit. We characterised the strength of GR using a dimensionless parameter ϵ_{GR} , and then studied the secular dynamics for arbitrary ϵ_{GR} , delineating several different dynamical regimes. In so doing we uncovered new and unintuitive phase space morphologies. We also derived several general results valid in the high eccentricity ($e \rightarrow 1$) limit, including an analytic solution to the DA equations of motion for all orbital elements. Despite its analytic simplicity such a solution had not been derived before, even in LK theory.

In Chapter 5, with one eye on an ultimate application to compact object merger physics, we added GW emission to our DA theory. Via a combination of analytical theory and numerical experiments we gained insight into the physics of cluster tide-driven decay of binary orbits (which ultimately led to ‘slow mergers’). While the results of this Chapter hold for arbitrary $\Gamma > 0$, they also shed important new light on LK-driven ($\Gamma = 1$) mergers, as many of the details had not been appreciated before even in this well-studied special case.

In Chapter 6 we finally broke the DA assumption and considered the impact of fluctuations in the torque on the timescale of the binary’s outer orbit. We derived an expression for the approximate magnitude of the resulting eccentricity fluctuations valid at $e \rightarrow 1$. We also chanced upon a new effect, relativistic phase space diffusion (RPSD), by which a very high-eccentricity binary can be ‘kicked’ from one ‘parent’ DA evolutionary track to another by sub-secular fluctuations. Once again this process occurs in classical three-body dynamics, but had not been pointed out before.



The third and final question posed in §1.2 read as follows:

- can tide-driven eccentricity excitation account for (some of) the compact object mergers currently being detected by LIGO/Virgo?

Of course, this question was already answered in the affirmative — at least qualitatively — in Chapter 5. To make that answer quantitative, in Chapter 7 we finally performed the calculation hinted at from the start of the thesis, namely the cluster-tide driven merger rate of compact object binaries. We computed merger rates for BH-BH, NS-BH and NS-NS binaries for various cluster potentials and binary concentrations, assuming all clusters to be spherical. We found that this dynamical channel can produce a significant number of mergers out to cluster-centric distances of several parsecs, and can contribute to the observed LIGO/Virgo rate at the level of several per cent.

8.2 Future work

The investigation undertaken in this thesis has unveiled some important — and previously neglected — aspects of binary dynamical evolution and compact object merger physics. However, it is far from whole story. There are many ways in which the theory presented here should be improved and extended (and already has been by some authors, as we mention below). Let us now comment on some of the most important and timely directions for future research.

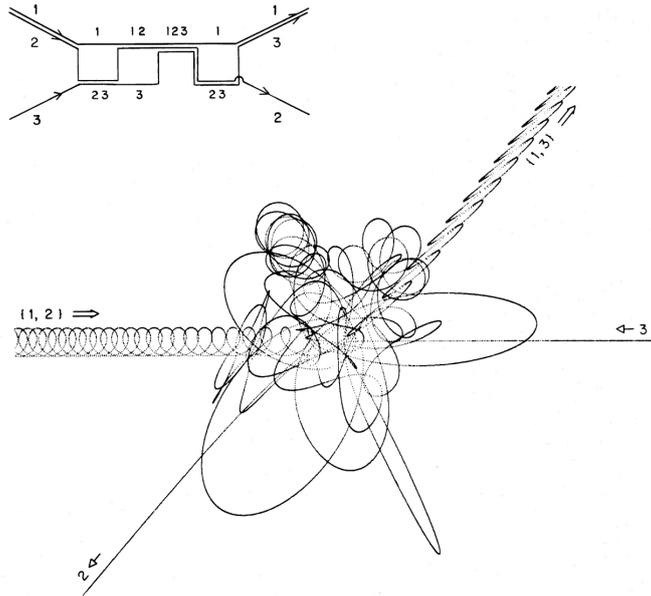


Figure 8.1: An example of a chaotic three-body interaction, taken from the classic paper by Hut & Bahcall (1983). A binary (stars 1 and 2) encounters a perturber (star 3) in a strong encounter with very small impact parameter. The schematic diagram on the top left illustrates the different configurations that the system goes through before it finally ‘ejects’ star 2.

8.2.1 Flyby encounters

Throughout this work we have assumed that the cluster potential that perturbs the binary is perfectly smooth — in other words, we have ignored the effect of any granularity in the potential, in particular due to passing stars. The most glaring omission from this thesis is therefore a proper accounting for the effect of stellar flybys on the binary evolution. These flybys are unavoidable in dense stellar systems, and can themselves modify sporadically a binary’s eccentricity, semimajor axis, etc., as we now discuss.

As already mentioned in §3.9.3, flyby encounters can be categorised into ‘strong’ and ‘weak’ encounters. Strong encounters have impact parameters (relative to the binary’s barycentre) that are comparable to the binary semimajor axis, a . The majority of these will fly past on nearly parabolic orbits and will ‘kick’ the binary impulsively (i.e. on a timescale much shorter than the inner orbital period), changing a and e as well as the other orbital elements (Heggie & Rasio 1996). However, some strong encounters, if they occur with a small enough impact parameter and relative velocity, can lead to a chaotic three-body interaction like the one illustrated in Figure 8.1. Of course, any such encounter would totally ‘reset’ the binary initial conditions and render our secular theory entirely inaccurate. In §3.9.3 we calculated the characteristic timescale that one has to wait for a given binary to undergo a strong encounter in a stellar cluster to be (equation (3.62)):

$$t_{\text{enc}} \approx 5 \text{ Gyr} \times \frac{1}{1 + \xi_{\text{GF}}} \left(\frac{n}{10^4 \text{ pc}^{-3}} \right)^{-1} \left(\frac{\sigma}{10 \text{ kms}^{-1}} \right)^{-1} \left(\frac{a}{10 \text{ AU}} \right)^{-2}, \quad (8.1)$$

where n is the cluster number density, σ is the typical speed of a passing star, and ξ_{GF} is related to the Safronov number which measures gravitational focusing. Equation (8.1) suggests that strong encounters should not be totally ignored in e.g. our merger rate calculation (Chapter 7), in which many binaries merged only after undergoing secular evolution for several Gyr (Figure 7.2). A more comprehensive calculation ought to account for such ‘resets’.

Perhaps more importantly, on the long secular timescales studied in this thesis one must take into consideration the effect of multiple ‘weak’ encounters on the binary evolution. Weak encounters are flybys with impact parameters $\gg a$. In this case, the flyby ‘duration’ (the time over which the perturber affects the binary significantly) is usually much longer than the binary’s inner orbital period, so these encounters are often termed ‘slow’ or ‘adiabatic’. Slow encounters do not change the binary’s semimajor axis but do cause a diffusion in the other orbital elements, including the eccentricity, as has been considered in detail by Heggie & Rasio (1996), Hamers (2018a), Hamers & Samsing (2019a), and Hamers & Samsing (2019b). Another way to think about weak encounters is to compare them to the RPSD phenomenon we found in Chapter 6. Recall that D and j_z are the primary integrals of motion for a binary undergoing secular evolution¹, and that RPSD led to diffusion of D , but only at high eccentricity and only when GR precession was included. On the contrary, a succession of weak flybys will lead to phase space diffusion in both D and j_z . It will occur at all binary eccentricities, and will not depend on the presence of GR; and rather than the violent jumps in D that we saw in the RPSD case, a succession of weak encounters will instead produce gentler kicks to D and j_z , resulting in something like a random walk in the (D, j_z) space.

Therefore, in the context of our work, it is clearly important to understand the ‘competition’ between the effect of numerous weak encounters upon a binary and the secular evolution of that binary driven by cluster tides. A small step in this direction was taken recently by Batygin et al. (2020) for a binary on a circular outer orbit in a Plummer cluster, but the general problem has yet to be studied in any detail. In future work, one ought to devise a general theory of binary dynamical evolution that would account for both of these effects simultaneously. More precisely, one could derive a Fokker-Planck-style kinetic theory for the phase space distribution function of an ensemble of binaries. This theory would allow for both secular forcing and diffusion due to stochastic kicks. It could be tested using a combination of precise few-body simulations and Monte-Carlo methods, for instance by integrating a BH-BH binary in a smooth cluster potential but including flybys from passing stars at random intervals (e.g. Hamers 2017).

¹If we ignore GW emission, that is.

8.2.2 Asphericity of clusters, central supermassive black holes, and relaxation of the outer orbit

The DA theory developed in this thesis is valid for any binary orbiting in any axisymmetric host system. This includes spherical systems as a special case but necessarily excludes triaxial systems. However, in our calculation of merger rates in Chapter 7, we assumed that all clusters were spherical. Of course in reality clusters are rarely spherical — globular clusters are typically flattened at the $\sim 10\%$ level (Mackey & Van Den Bergh 2005), while nuclear clusters are often triaxial (Neumayer et al. 2020). Allowing our clusters to be axisymmetric as opposed to spherical will drive nodal precession of the outer orbit, which could render the theory invalid if the outer orbit does not ‘fill its torus’ (Chapter 2) on a timescale much shorter than the secular period. Moreover, if the cluster is fully triaxial then there is a significant chance that the outer orbit will be chaotic (Binney & Tremaine 2008; Merritt 2013) and not amenable to any kind of outer orbit averaging. In either of these cases one must revert to the singly-averaged theory, which is valid for any potential, axisymmetric or otherwise. In future work it be good to drop the spherical assumption, for instance by making the cluster slightly flattened. One could then investigate systematically the effect that the chosen degree of flattening has on merger rates.

In Chapter 7 we also ignored the possibility that, when considering nuclear star clusters, there is very likely a central supermassive black hole (SMBH) present (Neumayer et al. 2020). This SMBH can drive LK oscillations in a nearby binary even in the absence of any cluster potential (Antonini & Perets 2012). Additionally, if we allow a sufficiently massive SMBH to exist at the centre of our spherical clusters, then we expect the resulting merger rates to be increased significantly, for three main reasons:

1. A central finding of this thesis is that for a given binary to reach extreme eccentricities $e \rightarrow 1$, and hence potentially merge, nearly always requires $\Gamma > 1/5$. As illustrated in Figure 8.2a the SMBH’s presence allows many more binaries to remain above the $\Gamma = 1/5$ barrier.
2. As illustrated in Figure 8.2b, the presence of a central SMBH means that most binaries will have higher A^* , leading to faster secular evolution and a smaller value of ϵ_{GR} (so GR precession is less likely to quench eccentricity oscillations).
3. As illustrated in Figure 8.2c, the central SMBH can boost the value of F_{circ}^* , which roughly sets the size of the angular momentum fluctuations at highest eccentricity (§6.3.4). Larger fluctuations reduce the merger time (Chapter 7).

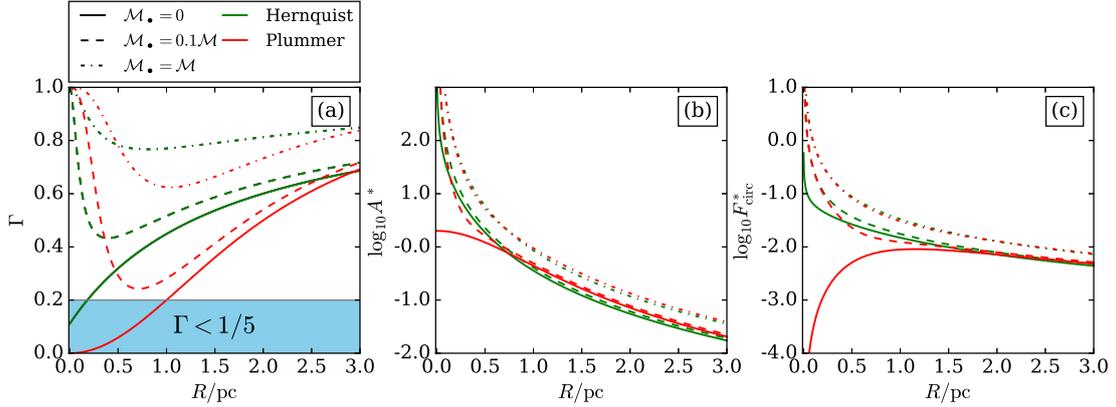


Figure 8.2: Plots of the important parameters (a) Γ , (b) $A^* \equiv A/(GM/b^3)$, and (c) $F_{\text{circ}}^* \equiv F_{\text{circ}}/(GM/b^3)^{1/2}$. We assume the binary is on a circular outer orbit of radius R in a spherical cluster with scale radius b and mass \mathcal{M} . Solid lines correspond to binaries in Plummer (red) and Hernquist (green) potentials, each scaled to have half-mass radius 1.31pc. We also show the results of adding a central point mass \mathcal{M}_{\bullet} (see legend). Note that the inclusion of the point mass can lift many binaries above the $\Gamma = 1/5$ barrier, below which high eccentricity excitation is very rare.

Thus, adding a central SMBH will likely lead to many more mergers, particularly in the more centrally concentrated models. Indeed, Arca Sedda (2020) has considered the problem of BH-BH binaries in spherical nuclear clusters with a central SMBH. He included the torquing effects of both the SMBH and the cluster potential upon the binary’s eccentricity evolution, although he assumed that the outer orbit was circular. He found that the BH-BH merger rate from this mechanism could account for perhaps $\sim 10\%$ of that measured by LIGO/Virgo, a significant boost compared to what we found in Chapter 7

In fact, a combination of cluster asphericity and a central SMBH can enhance merger rates even further. First of all, Petrovich & Antonini (2017) showed that in an aspherical nuclear cluster, if the outer orbit’s nodal precession timescale is comparable to the secular timescale for LK oscillations driven by a central SMBH, then the binary’s phase space evolution becomes chaotic. This can boost the merger rate significantly compared to the case of a spherical cluster with or without a SMBH. However, Petrovich & Antonini (2017) did not include the direct effect of the cluster potential on the *inner* orbit. Since then, Bub & Petrovich (2020) have combined the results of this thesis with those of Petrovich & Antonini (2017). They considered compact object binaries in triaxial nuclear clusters with and without a central point mass, and included the full effect of the cluster potential on both the binary’s inner and outer orbit. Under these conditions they found that the resulting BH-BH merger rate could be boosted by 1-2 orders of magnitude compared to what we found in Chapter 7.

Unfortunately, the studies by Bub & Petrovich (2020) and Arca Sedda (2020) were both severely limited in scope for computational reasons. The former authors approached the problem by simultaneously integrating the singly-averaged equations of motion of the inner

binary and the outer orbital motion in the combined SMBH-cluster potential. The latter used direct N-body simulations of the hierarchical triple (binary plus SMBH) system with extra terms added to the force to account for the cluster potential. When performing a population synthesis calculation, both approaches prove highly computationally expensive, and so neither study was able to cover a wide parameter space. Given our great ignorance of the masses and the density profiles of nuclear clusters, the masses of their central SMBHs, and of the semimajor axis and eccentricity distributions of compact object binaries in clusters, this is a significant drawback. In future it would be beneficial to develop a semi-analytical approach to this problem, which would not incur such a computational penalty, allowing us to investigate millions of binary initial conditions and study a wide parameter space.

Both asphericity and the presence of a SMBH modify the outer orbit, because they change the mean cluster potential. Additionally however, there can be modifications to the outer orbit that occur even though the mean potential of the system is fixed. For instance, vector resonant relaxation (Rauch & Tremaine 1996; Meiron & Kocsis 2018) can efficiently shuffle the orientation (i.e. direction of angular momentum unit vector) of orbits in spherical systems. This effect has been shown to increase LK-driven merger rates in galactic nuclei (Hamers, Bar-Or, et al. 2018a), but the analogous calculation has not yet been done for binaries in globular clusters². Also, heavy objects in galactic nuclei can experience a form of resonant dynamical friction that rapidly aligns them with a stellar disk (Szölgvény et al. 2021).

If the heavy object is a black hole binary then the binary might settle into the disk and consequently experience an enhanced rate of dynamical interactions with other stars, potentially hastening mergers. Circling back to something mentioned at the very start of this thesis (§1.1.1), if the heavy object is a single black hole then this dynamical friction mechanism might increase the prevalence of black holes found in AGN disks. The resulting high density of black holes can lead to the formation of black hole binaries; and those binaries can then be torqued by the central SMBH and/or cluster tides, or perturbed by other disk members. In summary, one wide-open avenue for future work is to consider the cluster tide-driven evolution of binaries whose outer orbits are slowly changing with time, even in fixed cluster potentials.

²Hamers, Bar-Or, et al. (2018a) showed that the merger rate of compact object binaries in the Galactic centre is enhanced most efficiently if the typical timescale for outer orbit reorientation by VRR is comparable to the LK secular timescale driven by the central SMBH. Applying the same argument to cluster tide-driven binary evolution, we must balance the VRR timescale $\sim \sqrt{N}T_\phi$ where N is the number of stars in the cluster (Rauch & Tremaine 1996; Meiron & Kocsis 2018) with the secular period $t_{\text{sec}} \sim T_\phi^2/T_b$. This implies $T_\phi/T_b \sim \sqrt{N}$, which is quite easily satisfied in a typical globular cluster. For instance, taking $N = 10^6$ and $T_\phi = 10^5 \text{yr}$, we find that VRR will be most effective for $T_b \sim 100 \text{yr}$, which is typical of, say, a $m_1 = m_2 = 30M_\odot$ BH-BH binary with $a \sim 40 \text{AU}$.

8.2.3 Population synthesis tools

Finally, there is the potential to bring all of the aforementioned effects together, and in so doing develop much more comprehensive numerical tools for population synthesis calculations. Precisely, one could combine the theoretical results of this thesis into a full semi-analytical model of compact object binary formation, evolution and mergers in stellar clusters, tracking populations of binaries from ‘birth’ to ‘death’ (i.e. merger or disruption) and including all important physical processes: cluster tides, flybys, GR precession, GW emission, compact object formation, etc. An ideal numerical toolkit would allow a user to input an ensemble of stars and compact objects with a given mass spectrum, binary fraction, etc., to focus on a particular environment (say, globular clusters), and would output predictions for observables like the merger rate, spin distribution of merger progenitors, and so on. Indeed, with small tweaks such a toolkit could be very useful not only for understanding compact object merger physics but also for attacking many related problems involving a ‘binary’ in a dense environment. For instance, it might be used to probe the formation of blue stragglers in globular clusters, or to understand the dynamics of comets around the Sun.

References

- Abbott, R. et al. (2020). “GWTC-2: Compact Binary Coalescences Observed by LIGO and Virgo During the First Half of the Third Observing Run”. In: *arXiv e-prints*, arXiv:2010.14527.
- Abbott et al., B. P. (2016). “Observation of gravitational waves from a binary black hole merger”. In: *Physical Review Letters* 116.6, p. 061102.
- Anderson, K. R., N. I. Storch, & D. Lai (2016). “Formation and stellar spin-orbit misalignment of hot Jupiters from Lidov–Kozai oscillations in stellar binaries”. In: *Monthly Notices of the Royal Astronomical Society* 456.4, pp. 3671–3701.
- Anderson, K. R., D. Lai, & N. I. Storch (2017). “Eccentricity and spin-orbit misalignment in short-period stellar binaries as a signpost of hidden tertiary companions”. In: *Monthly Notices of the Royal Astronomical Society* 467.3, pp. 3066–3082.
- Antognini, J. M. O. (2015). “Timescales of Kozai-Lidov oscillations at quadrupole and octupole order in the test particle limit”. In: *Monthly Notices of the Royal Astronomical Society* 452, pp. 3610–3619.
- Antognini, J. M. et al. (2014). “Rapid eccentricity oscillations and the mergers of compact objects in hierarchical triples”. In: *Monthly Notices of the Royal Astronomical Society* 439.1, pp. 1079–1091.
- Antonini, F. & F. A. Rasio (2016). “Merging Black Hole Binaries in Galactic Nuclei: Implications for Advanced-LIGO Detections”. In: *The Astrophysical Journal* 831, 187, p. 187.
- Antonini, F., S. Chatterjee, et al. (2016). “Black hole mergers and blue stragglers from hierarchical triples formed in globular clusters”. In: *The Astrophysical Journal* 816.2.
- Antonini, F., N. Murray, & S. Mikkola (2014). “Black hole triple dynamics: a breakdown of the orbit average approximation and implications for gravitational wave detections”. In: *The Astrophysical Journal* 781.1, p. 45.
- Antonini, F. & H. B. Perets (2012). “Secular Evolution of Compact Binaries near Massive Black Holes: Gravitational Wave Sources and Other Exotica”. In: *The Astrophysical Journal* 757.1, p. 27.
- Antonini, F., S. Toonen, & A. S. Hamers (2017). “Binary Black Hole Mergers from Field Triples: Properties, Rates, and the Impact of Stellar Evolution”. In: *The Astrophysical Journal* 841.2, p. 77.
- Arca Sedda, M. (2020). “Birth, Life, and Death of Black Hole Binaries around Supermassive Black Holes: Dynamical Evolution of Gravitational Wave Sources”. In: *The Astrophysical Journal* 891.1, 47, p. 47.
- Arnold, V. (1989). *Mathematical Methods of Classical Mechanics*. Graduate texts in mathematics. Springer.
- Bartos, I. et al. (2017). “Rapid and bright stellar-mass binary black hole mergers in active galactic nuclei”. In: *The Astrophysical Journal* 835.2, p. 165.
- Batygin, K. et al. (2020). “Dynamics of Planetary Systems within Star Clusters: Aspects of the Solar System’s Early Evolution”. In: *The Astronomical Journal* 159.3, p. 101.
- Belczynski, K. et al. (2016). “The first gravitational-wave source from the isolated evolution of two stars in the 40–100 solar mass range”. In: *Nature* 534.
- Bhattacharya, D. & E. P. J. van den Heuvel (1991). “Formation and evolution of binary and millisecond radio pulsars”. In: *Physics Reports* 203.1-2, pp. 1–124.
- Binney, J. & S. Tremaine (2008). *Galactic Dynamics: Second Edition*. Princeton University Press.
- Binney, J. & M. Merrifield (1998). *Galactic astronomy*. Vol. 9. Princeton University Press.
- Blaes, O., M. H. Lee, & A. Socrates (2002). “The Kozai Mechanism and the Evolution of Binary Supermassive Black Holes”. In: *The Astrophysical Journal* 578.2, p. 775.

- Bode, J. N. & C. Wegg (2014). “Production of EMRIs in supermassive black hole binaries”. In: *Monthly Notices of the Royal Astronomical Society* 438, pp. 573–589.
- Boffin, H. M., G. Carraro, G. Beccari, et al. (2015). *Ecology of Blue Straggler Stars*. Springer.
- Bovy, J. (2015). “galpy: A python Library for Galactic Dynamics”. In: *The Astrophysical Journals* 216, 29, p. 29.
- Brasser, R. (2001). “Some properties of a two-body system under the influence of the Galactic tidal field”. In: *Monthly Notices of the Royal Astronomical Society* 324.4, pp. 1109–1116.
- Brasser, R., M. Duncan, & H. Levison (2006). “Embedded star clusters and the formation of the Oort Cloud”. In: *Icarus* 184.1, pp. 59–82.
- Breiter, S., P. Dybczynski, & A. Elipe (1996). “The action of the Galactic disk on the Oort cloud comets. Qualitative study.” In: *Astronomy and Astrophysics* 315, pp. 618–624.
- Breiter, S., M. Fouchard, et al. (2007). “Two fast integrators for the Galactic tide effects in the Oort Cloud”. In: *Monthly Notices of the Royal Astronomical Society* 377.3, pp. 1151–1162.
- Bub, M. W. & C. Petrovich (2020). “Compact object mergers in the galactic center: evolution in triaxial clusters”. In: *The Astrophysical Journal* 894.1, p. 15.
- Byl, J. (1986). “The effect of the Galaxy on cometary orbits”. In: *Earth, Moon, and Planets* 36.3, pp. 263–273.
- Chang, P. (2009). “The effectiveness of the Kozai mechanism in the Galactic Centre”. In: *Monthly Notices of the Royal Astronomical Society* 393.1, pp. 224–228.
- Clark, G. (1975). “X-ray binaries in globular clusters”. In: *The Astrophysical Journal* 199, pp. L143–L145.
- Collins, B. F. & R. Sari (2008). “Lévy Flights of Binary Orbits due to Impulsive Encounters”. In: *The Astronomical Journal* 136.6, p. 2552.
- (2010). “A Unified Theory for the Effects of Stellar Perturbations and Galactic Tides on Oort Cloud Comets”. In: *The Astronomical Journal* 140.5, p. 1306.
- Correa-Otto, J., M. F. Calandra, & R. A. Gil-Hutton (2017). “A new insight into the Galactic potential: A simple secular model for the evolution of binary systems in the solar neighbourhood”. In: *A&A* 600, A59.
- de Mink, S. E. & I. Mandel (2016). “The chemically homogeneous evolutionary channel for binary black hole mergers: rates and properties of gravitational-wave events detectable by advanced LIGO”. In: *Monthly Notices of the Royal Astronomical Society* 460, pp. 3545–3553.
- Duquennoy, A. & M. Mayor (1991). “Multiplicity among solar-type stars in the solar neighbourhood. II - Distribution of the orbital elements in an unbiased sample”. In: *Astronomy and Astrophysics* 248, pp. 485–524.
- Evans, N. W. (1994). “The power-law galaxies”. In: *Monthly Notices of the Royal Astronomical Society* 267, pp. 333–360.
- Fabrycky, D. & S. Tremaine (2007). “Shrinking binary and planetary orbits by Kozai cycles with tidal friction”. In: *The Astrophysical Journal* 669.2, p. 1298.
- Farr, B., D. E. Holz, & W. M. Farr (2018). “Using spin to understand the formation of LIGO and Virgo’s black holes”. In: *The Astrophysical Journal Letters* 854.1, p. L9.
- Figer, D. F. et al. (2004). “An Extended Star Formation History for the Galactic Center from Hubble Space Telescope NICMOS Observations”. In: *The Astrophysical Journal* 601.1, pp. 319–339.
- Ford, E. B. & F. A. Rasio (2006). “On the relation between hot Jupiters and the Roche limit”. In: *The Astrophysical Journal Letters* 638.1, p. L45.
- Ford, E. B., B. Kozinsky, & F. A. Rasio (2000). “Secular Evolution of Hierarchical Triple Star Systems”. In: *The Astrophysical Journal* 535.1, p. 385.
- Fouchard, M. (2004). “New fast models of the Galactic tide”. In: *Monthly Notices of the Royal Astronomical Society* 349.1, pp. 347–356.
- Fouchard, M., C. Froeschlé, J. J. Matese, et al. (2005). “Comparison between different models of galactic tidal effects on cometary orbits”. In: *Celestial Mechanics and Dynamical Astronomy* 93.1-4, pp. 229–262.
- Fouchard, M., C. Froeschlé, G. Valsecchi, et al. (2006). “Long-term effects of the Galactic tide on cometary dynamics”. In: *Celestial Mechanics and Dynamical Astronomy* 95.1-4, pp. 299–326.

- Fragione, G. & O. Bromberg (2019). “Eccentric binary black hole mergers in globular clusters hosting intermediate-mass black holes”. In: *Monthly Notices of the Royal Astronomical Society* 488.3, pp. 4370–4377.
- Fragione, G. & B. Kocsis (2018). “Black hole mergers from an evolving population of globular clusters”. In: *Physical review letters* 121.16, p. 161103.
- (2019). “Black Hole Mergers from Quadruples”. In: *arXiv e-prints*, arXiv:1903.03112, arXiv:1903.03112.
- Gaidos, E. J. (1995). “Paleodynamics: Solar System formation and the early environment of the Sun”. In: *Icarus* 114.2, pp. 258–268.
- Generozov, A. et al. (2018). “An overabundance of black hole X-ray binaries in the Galactic Centre from tidal captures”. In: *Monthly Notices of the Royal Astronomical Society* 478.3, pp. 4030–4051.
- Giesers, B., S. Dreizler, et al. (2018). “A detached stellar-mass black hole candidate in the globular cluster NGC 3201”. In: *Monthly Notices of the Royal Astronomical Society* 475.1, pp. L15–L19.
- Giesers, B., S. Kamann, et al. (2019). “A stellar census in globular clusters with MUSE: Binaries in NGC 3201”. In: *Astronomy & Astrophysics* 632, A3.
- Gnedin, O. Y., J. P. Ostriker, & S. Tremaine (2014). “Co-evolution of galactic nuclei and globular cluster systems”. In: *The Astrophysical Journal* 785.1, p. 71.
- Goodman, J. & P. Hut (1993). “Binary-single-star scattering. V - Steady state binary distribution in a homogeneous static background of single stars”. In: *The Astrophysical Journal* 403, pp. 271–277.
- Gradshteyn, I. S. & I. M. Ryzhik (2014). *Table of integrals, series, and products*. Academic press.
- Grishin, E., H. B. Perets, & G. Fragione (2018). “Quasi-secular evolution of mildly hierarchical triple systems: analytics and applications for GW sources and hot Jupiters”. In: *Monthly Notices of the Royal Astronomical Society* 481.4, pp. 4907–4923.
- Guhathakurta, P. et al. (1998). “Globular cluster photometry with the Hubble Space Telescope. VII. Color gradients and blue stragglers in the central region of M30 from Wide Field Planetary Camera 2 observations”. In: *The Astronomical Journal* 116.4, p. 1757.
- Haas, J., L. Šubr, & D. Vokrouhlický (2011). “Secular theory of the orbital evolution of the young stellar disc in the Galactic Centre”. In: *Monthly Notices of the Royal Astronomical Society* 416, pp. 1023–1032.
- Haim, N. & B. Katz (2018). “Extreme close approaches in hierarchical triple systems with comparable masses”. In: *Monthly Notices of the Royal Astronomical Society* 479.3, pp. 3155–3166.
- Hamers, A. S. (2018a). “Secular dynamics of hierarchical multiple systems composed of nested binaries, with an arbitrary number of bodies and arbitrary hierarchical structure - II. External perturbations: flybys and supernovae”. In: *Monthly Notices of the Royal Astronomical Society* 476, pp. 4139–4161.
- (2018b). “Secular dynamics of hierarchical multiple systems composed of nested binaries, with an arbitrary number of bodies and arbitrary hierarchical structure - II. External perturbations: flybys and supernovae”. In: *Monthly Notices of the Royal Astronomical Society* 476, pp. 4139–4161.
- Hamers, A. S. & J. Samsing (2019a). “Analytic computation of the secular effects of encounters on a binary: features arising from second-order perturbation theory”. In: *Monthly Notices of the Royal Astronomical Society* 487.4, pp. 5630–5648.
- (2019b). “Analytic computation of the secular effects of encounters on a binary: third-order perturbation, octupole, and post-Newtonian terms; steady-state distribution”. In: *Monthly Notices of the Royal Astronomical Society* 488.4, pp. 5192–5209.
- Hamers, A. S. (2017). “On the formation of hot and warm Jupiters via secular high-eccentricity migration in stellar triples”. In: *Monthly Notices of the Royal Astronomical Society* 466.4, pp. 4107–4120.
- Hamers, A. S., B. Bar-Or, et al. (2018a). “The Impact of Vector Resonant Relaxation on the Evolution of Binaries near a Massive Black Hole: Implications for Gravitational-wave Sources”. In: *The Astrophysical Journal* 865.1, 2, p. 2.

- Hamers, A. S., B. Bar-Or, et al. (2018b). “The Impact of Vector Resonant Relaxation on the Evolution of Binaries near a Massive Black Hole: Implications for Gravitational-wave Sources”. In: *The Astrophysical Journal* 865.1, p. 2.
- Hamilton, C. & R. R. Rafikov (2021). “Secular dynamics of binaries in stellar clusters - III. doubly-averaged dynamics in the presence of general relativistic precession”. In: *Monthly Notices of the Royal Astronomical Society*. stab1284.
- (2019a). “Compact Object Binary Mergers Driven By Cluster Tides: A New Channel for LIGO/Virgo Gravitational-wave Events”. In: *The Astrophysical Journal Letters* 881.1, L13, p. L13.
- (2019b). “Secular dynamics of binaries in stellar clusters - I. General formulation and dependence on cluster potential”. In: *Monthly Notices of the Royal Astronomical Society* 488.4, pp. 5489–5511.
- (2019c). “Secular dynamics of binaries in stellar clusters - II. Dynamical evolution”. In: *Monthly Notices of the Royal Astronomical Society* 488.4, pp. 5512–5535.
- Harris, W. E. et al. (2014). “Globular cluster systems in brightest cluster galaxies: a near-universal luminosity function?” In: *The Astrophysical Journal* 797.2, p. 128.
- Heggie, D. C. (1975). “Binary evolution in stellar dynamics”. In: *Monthly Notices of the Royal Astronomical Society* 173, pp. 729–787.
- Heggie, D. & F. Rasio (1996). “The effect of encounters on the eccentricity of binaries in clusters”. In: *Monthly Notices of the Royal Astronomical Society* 282.3, pp. 1064–1084.
- Heisler, J. & S. Tremaine (1986). “The influence of the galactic tidal field on the Oort comet cloud”. In: *Icarus* 65, pp. 13–26.
- Higuchi, A. et al. (2007). “Orbital evolution of planetesimals due to the Galactic tide: formation of the Comet cloud”. In: *The Astronomical Journal* 134.4, p. 1693.
- Hills, J. & C. Day (1976). “Stellar collisions in globular clusters”. In: *Astrophysical Letters* 17, p. 87.
- Hoang, B.-M. et al. (2018). “Black Hole Mergers in Galactic Nuclei Induced by the Eccentric Kozai–Lidov Effect”. In: *The Astrophysical Journal* 856.2, p. 140.
- Hut, P. (1983). “Binary-single star scattering. II - Analytic approximations for high velocity”. In: *The Astrophysical Journal* 268, pp. 342–355.
- Hut, P. & J. N. Bahcall (1983). “Binary-single star scattering. I - Numerical experiments for equal masses”. In: *The Astrophysical Journal* 268, pp. 319–341.
- Iben, I. & M. Livio (1993). “Common envelopes in binary star evolution”. In: *Publications of the Astronomical Society of the Pacific* 105.694, p. 1373.
- Ito, T. & K. Ohtsuka (2019). “The Lidov-Kozai oscillation and Hugo von Zeipel”. In: *arXiv preprint arXiv:1911.03984*.
- Ivanov, P. B., A. G. Polnarev, & P. Saha (2005). “The tidal disruption rate in dense galactic cusps containing a supermassive binary black hole”. In: *Monthly Notices of the Royal Astronomical Society* 358, pp. 1361–1378.
- Iwasa, M. & N. Seto (2016). “Eccentricity boost of stars around shrinking massive black hole binaries”. In: *Phys. Rev. D* 93 (12), p. 124024.
- (2017). “Probabilistic eccentricity bifurcation for stars around shrinking massive black hole binaries”. In: *Monthly Notices of the Royal Astronomical Society* 472.2, pp. 1600–1617.
- Jiang, Y.-F. & S. Tremaine (2010). “The evolution of wide binary stars”. In: *Monthly Notices of the Royal Astronomical Society* 401, pp. 977–994.
- Kalogera, V. et al. (2007). “Formation of double compact objects”. In: *Publications of the Astronomical Society of Japan* 49, pp. 75–108.
- Katz, B. & S. Dong (2011). “Exponential growth of eccentricity in secular theory”. In: *arXiv preprint arXiv:1105.3953*.
- (2012). “The rate of WD-WD head-on collisions may be as high as the SNe Ia rate”. In: *arXiv preprint arXiv:1211.4584*.
- Katz, B., S. Dong, & R. Malhotra (2011). “Long-Term Cycling of Kozai-Lidov Cycles: Extreme Eccentricities and Inclinations Excited by a Distant Eccentric Perturber”. In: *Phys. Rev. Lett.* 107 (18), p. 181101.
- Katz, J. (1975). “Two kinds of stellar collapse”. In: *Nature* 253.5494, p. 698.

- Kinoshita, H. & H. Nakai (2007). “General solution of the Kozai mechanism”. In: *Celestial Mechanics and Dynamical Astronomy* 98, pp. 67–74.
- Knigge, C., N. Leigh, & A. Sills (2009). “A binary origin for blue stragglers in globular clusters”. In: *Nature* 457.7227, pp. 288–290.
- Kozai, Y. (1962). “Secular perturbations of asteroids with high inclination and eccentricity”. In: *The Astronomical Journal* 67, p. 591.
- Lei, H. (2019). “A semi-analytical model for secular dynamics of test particles in hierarchical triple systems”. In: *Monthly Notices of the Royal Astronomical Society* 490.4, pp. 4756–4769.
- Lei, H., C. Circi, & E. Ortore (2018). “Modified double-averaged Hamiltonian in hierarchical triple systems”. In: *Monthly Notices of the Royal Astronomical Society* 481.4, pp. 4602–4620.
- Leigh, N. W. C. et al. (2018). “On the rate of black hole binary mergers in galactic nuclei due to dynamical hardening”. In: *Monthly Notices of the Royal Astronomical Society* 474.4, pp. 5672–5683.
- Li, G. et al. (2014). “Eccentricity Growth and Orbit Flip in Near-coplanar Hierarchical Three-body Systems”. In: *The Astrophysical Journal* 785, 116, p. 116.
- Li, G. et al. (2015). “Implications of the eccentric Kozai-Lidov mechanism for stars surrounding supermassive black hole binaries”. In: *Monthly Notices of the Royal Astronomical Society* 451.2, pp. 1341–1349.
- Lidov, M. L. (1962). “The evolution of orbits of artificial satellites of planets under the action of gravitational perturbations of external bodies”. In: *Planetary Space Science* 9, pp. 719–759.
- Lithwick, Y. & S. Naoz (2011). “The eccentric Kozai mechanism for a test particle”. In: *The Astrophysical Journal* 742.2, p. 94.
- Liu, B. & D. Lai (2017). “Spin–Orbit Misalignment of Merging Black Hole Binaries with Tertiary Companions”. In: *The Astrophysical Journal Letters* 846.1, p. L11.
- (2018). “Enhanced black hole mergers in binary–binary interactions”. In: *Monthly Notices of the Royal Astronomical Society* 483.3, pp. 4060–4069.
- Liu, B., D. J. Muñoz, & D. Lai (2015). “Suppression of extreme orbital evolution in triple systems with short-range forces”. In: *Monthly Notices of the Royal Astronomical Society* 447.1, pp. 747–764.
- Löckmann, U., H. Baumgardt, & P. Kroupa (2008). “Origin of the S Stars in the Galactic Center”. In: *The Astrophysical Journal Letters* 683, p. L151.
- Löckmann, U., H. Baumgardt, & P. Kroupa (2010). “Constraining the initial mass function of stars in the Galactic Centre”. In: *Monthly Notices of the Royal Astronomical Society* 402.1, pp. 519–525.
- Luo, L., B. Katz, & S. Dong (2016). “Double-averaging can fail to characterize the long-term evolution of Lidov-Kozai Cycles and derivation of an analytical correction”. In: *Monthly Notices of the Royal Astronomical Society* 458, pp. 3060–3074.
- Mackey, A. & S. Van Den Bergh (2005). “The properties of Galactic globular cluster subsystems”. In: *Monthly Notices of the Royal Astronomical Society* 360.2, pp. 631–645.
- Manchester, R. N. et al. (2005). “The Australia telescope national facility pulsar catalogue”. In: *The Astronomical Journal* 129.4, p. 1993.
- Mandel, I. & S. E. de Mink (2016). “Merging binary black holes formed through chemically homogeneous evolution in short-period stellar binaries”. In: *Monthly Notices of the Royal Astronomical Society* 458.3, pp. 2634–2647.
- Mapelli, M., S. Sigurdsson, F. R. Ferraro, et al. (2006). “The radial distribution of blue straggler stars and the nature of their progenitors”. In: *Monthly Notices of the Royal Astronomical Society* 373, pp. 361–368.
- Mapelli, M., S. Sigurdsson, M. Colpi, et al. (2004). “The contribution of primordial binaries to the blue straggler population in 47 Tucanae”. In: *The Astrophysical Journal Letters* 605.1, p. L29.
- Martinez, M. A. et al. (2020). “Black Hole Mergers from Hierarchical Triples in Dense Star Clusters”. In: *The Astrophysical Journal* 903.1, p. 67.
- Matese, J. & D. Whitmire (1996). “Tidal imprint of distant galactic matter on the Oort comet cloud”. In: *The Astrophysical Journal Letters* 472.1, p. L41.
- Matese, J. J. & P. G. Whitman (1989). “The galactic disk tidal field and the nonrandom distribution of observed Oort cloud comets”. In: *Icarus* 82.2, pp. 389–401.

- Meiron, Y. & B. Kocsis (2018). “Diffusion and Mixing in Globular Clusters”. In: *The Astrophysical Journal* 855, 87, p. 87.
- Meiron, Y. & B. Kocsis (2019). “Resonant relaxation in globular clusters”. In: *The Astrophysical Journal* 878.2, p. 138.
- Merritt, D. (2013). *Dynamics and Evolution of Galactic Nuclei*. Princeton University Press.
- Mikkola, S. & P. Nurmi (2006). “Computing secular motion under slowly rotating quadratic perturbation”. In: *Monthly Notices of the Royal Astronomical Society* 371.1, pp. 421–423.
- Miller, M. C. & D. P. Hamilton (2002). “Four-Body Effects in Globular Cluster Black Hole Coalescence”. In: *The Astrophysical Journal* 576.2, p. 894.
- Miyamoto, M. & R. Nagai (1975). “Three-dimensional models for the distribution of mass in galaxies”. In: *Publications of the Astronomical Society of Japan* 17, pp. 533–543.
- Morbidelli, A. & H. F. Levison (2004). “Scenarios for the Origin of the Orbits of the Trans-Neptunian Objects 2000 CR₁₀₅ and 2003 VB₁₂ (Sedna)”. In: *The Astrophysical Journal* 128.5, pp. 2564–2576.
- Muno, M. P. et al. (2005). “An overabundance of transient x-ray binaries within 1 parsec of the Galactic Center”. In: *The Astrophysical Journal Letters* 622.2, p. L113.
- Muñoz, D. J., D. Lai, & B. Liu (2016). “The formation efficiency of close-in planets via Lidov-Kozai migration: analytic calculations”. In: *Monthly Notices of the Royal Astronomical Society* 460.1, pp. 1086–1093.
- Murray, C. D. & S. F. Dermott (1999). *Solar system dynamics*.
- Naoz, S., W. M. Farr, et al. (2011). “Hot Jupiters from secular planet–planet interactions”. In: *Nature* 473, p. 187.
- Naoz, S. (2016). “The eccentric Kozai-Lidov effect and its applications”. In: *Annual Review of Astronomy and Astrophysics* 54, pp. 441–489.
- Naoz, S., W. M. Farr, et al. (2013). “Secular dynamics in hierarchical three-body systems”. In: *Monthly Notices of the Royal Astronomical Society* 431.3, pp. 2155–2171.
- Naoz, S., B. Kocsis, et al. (2013). “Resonant Post-Newtonian Eccentricity Excitation in Hierarchical Three-body Systems”. In: *The Astrophysical Journal* 773.2, 187, p. 187.
- Neumayer, N., A. Seth, & T. Böker (2020). “Nuclear star clusters”. In: *The Astronomy and Astrophysics Review* 28.1, pp. 1–75.
- O’Leary, R. M. et al. (2006). “Binary Mergers and Growth of Black Holes in Dense Star Clusters”. In: *The Astrophysical Journal* 637, pp. 937–951.
- Paczynski, B. (1971). “Evolutionary processes in close binary systems”. In: *Annual Review of Astronomy and Astrophysics* 9.1, pp. 183–208.
- Pati, M. E. & C. M. Will (2002). “Post-Newtonian gravitational radiation and equations of motion via direct integration of the relaxed Einstein equations. II. Two-body equations of motion to second post-Newtonian order, and radiation reaction to 3.5 post-Newtonian order”. In: *Physical Review D* 65.10, p. 104008.
- Perets, H. B. & D. C. Fabrycky (2009). “On the triple origin of blue stragglers”. In: *The Astrophysical Journal* 697.2, p. 1048.
- Peters, P. C. (1964). “Gravitational Radiation and the Motion of Two Point Masses”. In: *Phys. Rev.* 136 (4B), B1224–B1232.
- Petrovich, C. (2015). “Steady-state Planet Migration by the Kozai-Lidov Mechanism in Stellar Binaries”. In: *The Astrophysical Journal* 799, 27, p. 27.
- Petrovich, C. & F. Antonini (2017). “Greatly Enhanced Merger Rates of Compact object Binaries in Non-spherical Nuclear Star Clusters”. In: *The Astrophysical Journal* 846, 146, p. 146.
- Phinney, E. S. & S. R. Kulkarni (1994). “Binary and Millisecond Pulsars”. In: *Annual Review of Astronomy and Astrophysics* 32, pp. 591–639.
- Pooley, D. et al. (2003). “Dynamical formation of close binary systems in globular clusters”. In: *The Astrophysical Journal Letters* 591.2, p. L131.
- Portegies Zwart, S. F. & S. L. W. McMillan (2000). “Black Hole Mergers in the Universe”. In: *The Astrophysical Journal Letters* 528, pp. L17–L20.
- Prodan, S., F. Antonini, & H. B. Perets (2015). “Secular Evolution of Binaries near Massive Black Holes: Formation of Compact Binaries, Merger/Collision Products and G2-like Objects”. In: *The Astrophysical Journal* 799.2, 118, p. 118.

- Randall, L. & Z.-Z. Xianyu (2018). “An Analytical Portrait of Binary Mergers in Hierarchical Triple Systems”. In: *The Astrophysical Journal* 864.2, p. 134.
- Rauch, K. P. & S. Tremaine (1996). “Resonant relaxation in stellar systems”. In: *New Astronomy* 1, pp. 149–170.
- Regev, O. (2006). *Chaos and Complexity in Astrophysics*. Cambridge University Press.
- Rein, H. & S. .-. Liu (2012). “REBOUND: an open-source multi-purpose N-body code for collisional dynamics”. In: *Astronomy and Astrophysics* 537, A128, A128.
- Rodriguez, C. L., S. Chatterjee, & F. A. Rasio (2016). “Binary black hole mergers from globular clusters: Masses, merger rates, and the impact of stellar evolution”. In: *Physical Review D* 93.8, p. 084029.
- Rodriguez, C. L. & A. Loeb (2018). “Redshift evolution of the black hole merger rate from globular clusters”. In: *The Astrophysical Journal Letters* 866.1, p. L5.
- Rodriguez, C. L., M. Zevin, et al. (2016). “Illuminating black hole binary formation channels with spins in advanced LIGO”. In: *The Astrophysical Journal Letters* 832.1, p. L2.
- Rodriguez, C. L. & F. Antonini (2018). “A Triple Origin for the Heavy and Low-spin Binary Black Holes Detected by LIGO/VIRGO”. In: *The Astrophysical Journal* 863.1, 7, p. 7.
- Rodriguez, C. L., M. Morscher, et al. (2015). “Binary Black Hole Mergers from Globular Clusters: Implications for Advanced LIGO”. In: *Phys. Rev. Lett.* 115 (5), p. 051101.
- Samsing, J., A. S. Hamers, & J. G. Tyles (2019). “Effect of distant encounters on black hole binaries in globular clusters: Systematic increase of in-cluster mergers in the LISA band”. In: *Publications of the Astronomical Society of Japan* 100.4, 043010, p. 043010.
- Sandage, A. (1953). “The color-magnitude diagram for the globular cluster M 3.” In: *The Astronomical Journal* 58, pp. 61–75.
- Silsbee, K. & S. Tremaine (2017). “Lidov-Kozai Cycles with Gravitational Radiation: Merging Black Holes in Isolated Triple Systems”. In: *The Astrophysical Journal* 836, 39, p. 39.
- Sridhar, S. & J. Touma (1999). “Stellar dynamics around black holes in galactic nuclei”. In: *Monthly Notices of the Royal Astronomical Society* 303.3, pp. 483–494.
- Stephan, A. P. et al. (2016). *Merging Binaries in the Galactic Center: The eccentric Kozai-Lidov mechanism with stellar evolution*.
- Stone, N. C., B. D. Metzger, & Z. Haiman (2017). “Assisted inspirals of stellar mass black holes embedded in AGN discs: solving the ‘final au problem’”. In: *Monthly Notices of the Royal Astronomical Society* 464.1, pp. 946–954.
- Šubr, L. ., J. Schovancová, & P. Kroupa (2009). “The warped young stellar disc in the Galactic centre”. In: *Astronomy and Astrophysics* 496, pp. 695–699.
- Szölgvény, Á., G. Máthé, & B. Kocsis (2021). “Resonant Dynamical Friction in Nuclear Star Clusters: Rapid Alignment of an Intermediate Mass Black Hole with a Stellar Disk”. In: *arXiv preprint arXiv:2103.14042*.
- Taam, R. E. & E. L. Sandquist (2000). “Common envelope evolution of massive binary stars”. In: *Annual Review of Astronomy and Astrophysics* 38.1, pp. 113–141.
- Tagawa, H. et al. (2021). “Eccentric Black Hole Mergers in Active Galactic Nuclei”. In: *The Astrophysical Journal Letters* 907.1, p. L20.
- The LIGO Scientific Collaboration, the Virgo Collaboration, & et al. (2019). “GWTC-1: a gravitational-wave transient catalog of compact binary mergers observed by LIGO and Virgo during the first and second observing runs”. In: *Physical Review X* 9.3, p. 031040.
- (2020). “Population Properties of Compact Objects from the Second LIGO-Virgo Gravitational-Wave Transient Catalog”. In: *arXiv e-prints*, arXiv:2010.14533.
- Thompson, T. A. (2011). “Accelerating Compact Object Mergers in Triple Systems with the Kozai Resonance: A Mechanism for “Prompt” Type Ia Supernovae, Gamma-Ray Bursts, and Other Exotica”. In: *The Astrophysical Journal* 741, 82, p. 82.
- Toonen, S., H. B. Perets, & A. S. Hamers (2018). “Rate of WD-WD head-on collisions in isolated triples is too low to explain standard type Ia supernovae”. In: *Astronomy and Astrophysics* 610.
- Tutukov, A. & L. Yungelson (1973). “Evolution of massive close binaries”. In: *Nauchnye Informatsii* 27, p. 70.

- VanLandingham, J. H. et al. (2016). “The Role of the Kozai–Lidov Mechanism in Black Hole Binary Mergers in Galactic Centers”. In: *The Astrophysical Journal* 828, 77, p. 77.
- Vashkov’yak, M. A. (1999). “Evolution of the orbits of distant satellites of Uranus”. In: *Astronomy Letters* 25, pp. 476–481.
- Veras, D. & N. W. Evans (2013a). “Exoplanets beyond the Solar neighbourhood: Galactic tidal perturbations”. In: *Monthly Notices of the Royal Astronomical Society* 430.1, pp. 403–415.
- (2013b). “Exoplanets beyond the Solar neighbourhood: Galactic tidal perturbations”. In: *Monthly Notices of the Royal Astronomical Society* 430.1, pp. 403–415.
- (2013c). “Planetary orbital equations in externally-perturbed systems: position and velocity-dependent forces”. In: *Celestial Mechanics and Dynamical Astronomy* 115.2, pp. 123–141.
- Veras, D. & E. B. Ford (2010). “Secular orbital dynamics of hierarchical two-planet systems”. In: *The Astrophysical Journal* 715.2, p. 803.
- Vick, M., D. Lai, & K. R. Anderson (2019). “Chaotic tides in migrating gas giants: forming hot and transient warm Jupiters via Lidov–Kozai migration”. In: *Monthly Notices of the Royal Astronomical Society* 484.4, pp. 5645–5668.
- Wen, L. (2003). “On the eccentricity distribution of coalescing black hole binaries driven by the Kozai mechanism in globular clusters”. In: *The Astrophysical Journal* 598.1, p. 419.
- Wiegert, P. & S. Tremaine (1999). “The evolution of long-period comets”. In: *Icarus* 137.1, pp. 84–121.
- Will, C. M. (2017). “Orbital flips in hierarchical triple systems: Relativistic effects and third-body effects to hexadecapole order”. In: *Phys. Rev. D* 96 (2), p. 023017.
- Yabushita, S. (1989). “On the orbital elements of nearly parabolic comets at previous returns as perturbed by the galactic tidal field”. In: *The Astronomical Journal* 97, pp. 262–264.
- Zevin, M. et al. (2019). “Eccentric Black Hole Mergers in Dense Star Clusters: The Role of Binary–Binary Encounters”. In: *The Astrophysical Journal* 871.1, p. 91.