Title	Efficient trajectory calculations for extreme mass-ratio inspirals using near-identity (averaging) transformations
Authors(s)	Lynch, Philip
Publication date	2022
Publication information	Lynch, Philip. “Efficient Trajectory Calculations for Extreme Mass-Ratio Inspirals Using near-Identity (Averaging) Transformations.” University College Dublin. School of Mathematics and Statistics, 2022.
Publisher	University College Dublin. School of Mathematics and Statistics
Item record/more information	http://hdl.handle.net/10197/13347

Efficient trajectory calculations for extreme mass-ratio inspirals using near-identity (averaging) transformations

by

Philip Lynch

UCD Student Number: 14376721

This thesis is submitted to University College Dublin in fulfilment of the requirements for the degree of Doctor of Philosophy in Mathematics.

UCD School of Mathematics and Statistics

Head of School: Dr. Edward Cox

Primary Supervisor: Dr Niels Warburton

Secondary Supervisor: Prof. Adrian Ottewill

Members of Research Studies Panel:

Chair: Dr Barry Wardell

Internal Examiner 1: Dr Christopher Kavanagh

Internal Examiner 2: Dr Miguel Bustamante

External Examiner: Prof. Éanna Flanagan

Contents

1 Introduction **1**
- 1.1 Gravitational Waves . . . 1
- 1.2 Extreme Mass Ratio Inspirals . . . 2
  - 1.2.1 Astrophysical Motivation . . . 2
  - 1.2.2 Science implications . . . 3
- 1.3 Approaches to the Relativistic Two Body Problem . . . 4
- 1.4 Gravitational Self-Force . . . 5
- 1.5 Modelling Extreme Mass Ratio Inspirals . . . 8
  - 1.5.1 Kludge Models . . . 8
  - 1.5.2 Adiabatic Models . . . 8
  - 1.5.3 Osculating Geodesics . . . 9
  - 1.5.4 Self-Forced Inspirals . . . 10
  - 1.5.5 Near-Identity Averaging Transformations . . . 11
  - 1.5.6 Orbital Resonances . . . 12
- 1.6 Organisation of this thesis . . . 12

2 Geodesic Motion **15**
- 2.1 Geodesic Motion around a Kerr black hole . . . 15
  - 2.1.1 Geodesic motion and orbital parametrization . . . 15
  - 2.1.2 Analytic solutions for radial and polar motion . . . 18
  - 2.1.3 Analytic solutions for Boyer-Lindquist time and azimuthal coordinates . . . 19
  - 2.1.4 Fundamental frequencies in other time parametrizations . . . 21
- 2.2 Classification of bound geodesic orbits . . . 22
  - 2.2.1 Circular and equatorial . . . 22
  - 2.2.2 Eccentric and equatorial . . . 23
  - 2.2.3 Spherical (circular and inclined) . . . 23
  - 2.2.4 Generic (eccentric and inclined) . . . 24
- 2.3 Resonant Orbits . . . 24

3 Osculating Geodesics **27**
- 3.1 The method of osculating geodesics . . . 27
- 3.2 Evolution equations of the orbital elements . . . 28
  - 3.2.1 Generic case . . . 28
  - 3.2.2 Spherical case . . . 30
- 3.3 Evolution of the orbital phases . . . 31
  - 3.3.1 Evolution of quasi-Keplerian angles . . . 31
  - 3.3.2 Evolution of action angles . . . 33
- 3.4 Equations of motion for self-forced generic Kerr inspirals . . . 34

4 Near Identity Transformations 37
4.1 A review of near-identity transformations for generic EMRI systems 37
4.1.1 Near Identity Transformation 37
4.1.2 Transformed Equations of Motion 38
4.1.3 Cancellation of oscillating terms at $\mathcal{O}(\epsilon)$ 38
4.1.4 Cancellation of oscillating terms at $\mathcal{O}(\epsilon^2)$ 39
4.1.5 Freedom in the averaged pieces 40
4.1.6 Evolution of extrinsic quantities 40
4.1.7 Summary of NIT Results 42
4.2 Near-resonant near identity transformations 43
4.2.1 Near Identity Transformation 43
4.2.2 Transformed Equations of Motion 43
4.2.3 Cancellation of oscillating terms at $\mathcal{O}(\epsilon)$ 44
4.2.4 Cancellation of oscillating terms at $\mathcal{O}(\epsilon^2)$ 45
4.2.5 Freedom in the averaged pieces 45
4.2.6 Evolution of extrinsic quantities 46
4.2.7 Summary of Near Resonant NIT 46
4.3 Averaging transformations for motion parametrized by Boyer-Lindquist coordinate time 47
4.4 Two-timescale expansion 50

5 Waveforms 53
5.1 Waveform Generation 53
5.1.1 Teukolsky Based Waveforms 53
5.1.2 Semi-Relativistic Quadrupole Waveforms 55
5.2 Waveform Analysis 57

6 Eccentric self-forced inspirals into a rotating black hole 61
6.1 Osculating Geodesics 61
6.2 Interpolated GSF model 62
6.3 Near Identity Transformations 63
6.4 Implementation 65
6.4.1 Offline Steps 65
6.4.2 Online Steps 66
6.5 Results 66
6.5.1 Consistency checks 67
6.5.2 Comparison between OG and NIT inspirals 67
6.5.3 Impact of adiabatic and post-adiabatic effects 75
6.5.4 Comparing inspirals driven using radiation gauge and Lorenz gauge self-force in Schwarzschild spacetime 77

7 Spherical self-forced inspirals into a rotating black hole 83
7.1 Spherical osculating geodesics 83
7.2 Interpolated models 84
7.2.1 Interpolated gravitational wave fluxes for spherical Kerr inspirals 84
7.2.2 Interpolated gravitational self-force model for spherical Kerr inspirals 85
7.3 Averaged equations of motion for spherical Kerr inspirals 87
7.4 Implementation 89
7.4.1 Offline Steps 89

7.4.2 Online Steps 90
7.5 Results 91
7.5.1 OG vs NIT and TTE inspirals 91
7.5.2 Using Higher Precision Fluxes 96
7.5.3 Impact of the self-force on spherical inspirals 97

8 Fast generic inspirals into a rotating black hole 99
8.1 Generic Kerr Osculating Geodesics 100
8.2 Generic Kerr GSF toy model 100
8.3 Non-Resonant NIT 103
8.3.1 Non-resonant NIT equations of motion 103
8.3.2 Implementation 104
8.3.3 Offline Steps 104
8.3.4 Online Steps 104
8.3.5 Results 105
8.4 Near-resonant NIT 106
8.4.1 Near -resonant NIT equations of motion 107
8.4.2 Implementation 108
8.4.3 Online Steps 108
8.4.4 Results 109
8.5 Resonance Transition 110
8.5.1 Transitioning through a single resonance 111
8.5.2 Transitioning through multiple resonances 114

9 Conclusion 119
9.1 Eccentric Self-forced inspirals into a rotating black hole 120
9.2 Spherical self-forced inspirals into a rotating black hole 120
9.3 Fast generic inspirals into a rotating black hole 121
9.4 Future Work 122

Appendices 123

A Self-force corrections to the periapsis advance around a spinning black hole 125

Bibliography 125

Abstract

Future space based gravitational wave detectors, such as the Laser Interferometer Space Antenna (LISA) will allow for the detection of previously undetectable gravitational wave sources. These include extreme mass ratio inspirals (EMRIs) which consist of a stellar mass compact object spiralling into a massive black hole (MBH) due to gravitational radiation reaction. These sources are of particular interest for their ability to accurately map the spacetime of the MBH, allowing for unprecedentedly accurate measurements of the MBH's mass and spin, and tests of general relativity in the strong field regime. In order to reach the science goals of the LISA mission, one requires waveform models that are (i) accurate to within a fraction of a radian, (ii) extensive in the source's parameter space and (iii) fast to compute, ideally in less than a second. This thesis focuses on the latter criteria by utilising techniques that will speed up inspiral trajectory calculations as well as extending prior models to include the MBH's spin.

To this end, we develop the first EMRI models that incorporate the spin of the MBH along with all effects of the gravitational self-force (GSF) to first order in the mass ratio. Our models are based on an action angle formulation of the method of osculating geodesics (OG) for generic inspirals in Kerr spacetime. For eccentric equatorial inspirals and spherical inspirals, the forcing terms are provided by an efficient pseudo-spectral interpolation of the first order GSF in the outgoing radiation gauge. For generic inspirals where sufficient GSF data is not available, we construct a toy model from the previous two models. However, the OG method is slow to evaluate due to the dependence of the equations of motion (EOM) on the orbital phases. Therefore, we apply a near-identity (averaging) transformation (NIT) to eliminate all dependence of EOM on the orbital phases while maintaining all secular effects to post-adiabatic order. This inspiral model can be evaluated in less than a second for any mass-ratio as we no longer have to resolve all $\sim 10^5$ orbit cycles of a typical EMRI. This work marks the first time this technique has been applied in Kerr spacetime for eccentric, spherical, and generic inspirals.

In the case of a non-rotating MBH, we compare eccentric inspirals evolved using GSF data computed in the Lorenz and radiation gauges. We find that the two gauges produce differing inspirals with a deviation of comparable magnitude to the conservative GSF correction. This emphasizes the need to include the (currently unknown) second order GSF for gauge independent, post-adiabatic waveforms.

For spherical orbits, we perform a second averaging transformation to parametrise the averaged EOM in terms of Boyer-Lindquist time instead of Mino time, which is much more convenient for LISA data analysis. We also implement a two-timescale expansion of the EOM and find that both approaches yield self-forced inspirals can be evolved to sub radian accuracy in less than a second. We further improve our spherical inspiral model by incorporating high precision gravitational wave flux calculations and find that without making this modification, the final waveform would be out of phase by as much as $10-10^4$ radians for typical LISA band EMRIs.

For generic inspirals, one can encounter transient orbital resonances where the standard NIT procedure breaks down. We use the standard NIT when far from these resonances and then we average all phases apart from the resonant phase when in their vicinity. This results in the fastest model to date which includes includes all resonant effects. Our preliminary results demonstrate that accurately modelling only the two lowest order resonances costs 10s of seconds for a typical EMRI, but the resulting waveforms are sufficiently accurate for LISA data science.

Statement of Original Authorship

I hereby certify that the submitted work is my own work, was completed while registered as a candidate for the degree of Doctor of Philosophy in Mathematics, and I have not obtained a degree elsewhere on the basis of the research presented in this submitted work.

Collaborations

The first five chapters serve as background to the final three chapters and were written by me with the help of feedback from Dr Warburton. Where results and insight are derived from other works, proper attribution has been given.

Chapter 6 is based on a published paper [1] which was written in collaboration with Dr Maarten van de Meent and Dr Niels Warburton who provided guidance throughout the project. Data for the first-order gravitational self force and code for the interpolated GSF model were provided by Dr van de Meent. All other work was undertaken by me. The text was written by me with some insightful additions and alterations by Dr Warburton, Dr van de Meent, and the reviewers at *Classical and Quantum Gravity*.

Chapter 7 is based on an unpublished manuscript written in collaboration with Dr Maarten van de Meent and Dr Niels Warburton who provided guidance throughout the project. Data for the gravitational wave fluxes and the first-order gravitational self force and as well as code for the interpolated models were provided by Dr van de Meent. All other work was undertaken by me. The the text was written by me with the help of some valuable feedback from Dr Warburton and Dr van de Meent.

Chapter 8 is based on an ongoing project in collaboration with Dr Maarten van de Meent, Dr Niels Warburton, and Dr Vojtech Witzany who have all provided guidance throughout the project. Dr Witzany produced a PN inspired GSF toy model which we ultimately chose not to use. All other work was undertaken by me. The text was written by me with the aid of some useful feedback from Dr Warburton.

Dedication

This thesis is dedicated to my late father, Michael Lynch. Through his constant love and support, he instilled in me the value of education and nurtured my curiosity about the world. He was so proud to see me begin my PhD and I know he would have been proud to see me finish it. Go raibh suaimhneas síoraí air.

Acknowledgements

While the last four years have been a turbulent time for us all, I have been privileged to spend that time pursuing work that I am dedicated and passionate about. Even still, this work would never have been completed without the friendship and support of many others, and so the success of this work is as much theirs as it is mine.

First and foremost, I would like to thank my supervisor, Dr Niels Warburton. Working alongside him has been a pleasure and he has been the best mentor one could possibly hope for. He has taught me importance of perseverance, work-life balance, and well designed cycle infrastructure. Without his hard work and insightful guidance, this thesis would never have been written.

I would also like to thank my collaborator Dr Maarten van de Meent. His contributions to this research and perceptive feedback have been vital for the success of this work. Furthermore, I would like to thank Dr Vojtech Witzany, whose collaboration on our most recent project has been invaluable.

I would like to thank all of the members of the UCD Relativity Group for fostering such a supportive research environment. In particular, I would like to thank my fellow PhD students Josh Mathews, Leanne Durkan, and Ben Leather. Between all the conversations, conferences, and companionship, I could not have asked for a better cohort to spend my PhD with. I would also like to thank all of the PhD students and postdocs to have passed through room B113. Through extended lunches and tactical coffee breaks, they have turned the often lonely and isolating process of research into a social experience that makes me enjoy coming into the office everyday.

I must also thank my friends who have stood by me throughout this time and constantly remind me to enjoy life outside of research. Last, but certainly not least, I would like to thank my family. Between loss, grief, and lock-downs, we have been through a lot in the last four years. I certainly would not have been able to overcome it all without them.

Chapter 1

Introduction

1.1 Gravitational Waves

On the 14th of September 2015, the the LIGO-VIRGO Collaboration (LVC) made the first ever detection of a gravitational wave and ushered in a new era of astronomy [2]. No longer would astronomers have to rely on electromagnetism or neutrinos to probe the universe, but now they could use the ripples in the fabric of spacetime itself. Albert Einstein's theory of General Relativity (GR) predicts that any object with mass is capable of producing such ripples[1], which cause intervals in space to stretch and compress as they pass by [3]. Due to the weakness of gravity compared to the other fundamental forces, only the most violent gravitational events are detectable by the time they reach Earth, such as those produced by the merging of celestial objects with masses of the same order of magnitude as our sun, like black holes and neutron stars.

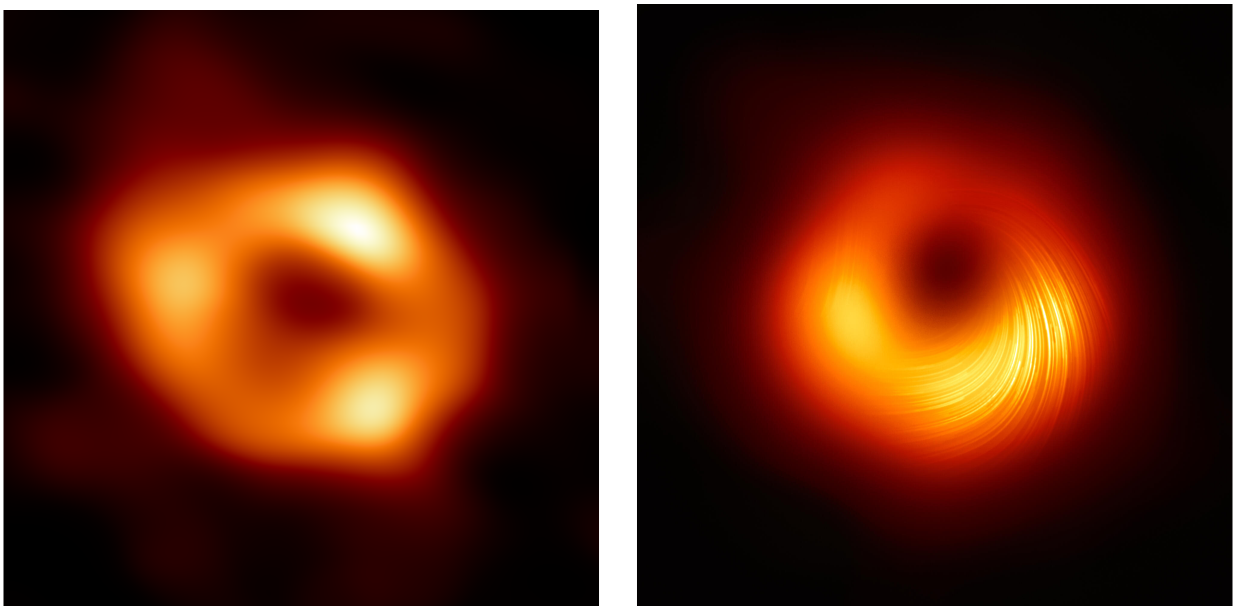
Now, seven years later, such observations are almost commonplace with the most recent observing runs, O3A and O3B, reporting a new candidate roughly once a week. With 90 confirmed detections catalogued thus far [4–6], we have obtained an unprecedented insight into the nature of compact binary coalescences (CBCs) [7] and have performed the most rigorous tests of GR in the strong field to date [8].

Current ground based detectors like LIGO, VIRGO and Kagra, are only sensitive to gravitational waves between 10Hz and 10kHz and future detectors like the Einstein Telescope [9] and Cosmic Explorer [10] may be able to push this lower bound down to 1-5Hz. However, detecting lower frequency gravitational waves with such detectors is difficult, largely due to the noise produced by the seismic activity of the Earth. This is the motivation for space based gravitational wave detectors such as the Japanese Aerospace Exploration Agency's (JAXA) DECIGO mission [11], the Chinese TianQin [12] and Teiji [13] missions and the European Space Agency's (ESA) Laser Interferometer Space Antenna (LISA) [14], all proposed to be launched in the 2030s. The former will be sensitive to gravitational waves between 0.1 Hz and 10 Hz, while the latter three will be sensitive to gravitational waves between 0.1 mHz and 1 Hz. This will allow us to detect gravitational waves from a variety of hitherto unseen sources such as the cosmological background, binary stars within the Milky Way and binaries with a total mass in the range of millions of solar masses [15].

Models of the gravitational waveforms from these sources will play a vital role in their detection, so long as the models can meet three important requirements. Firstly, they should be accurate enough so that systematic biases do not impact the scientific insights one draws form the data. Secondly, they should cover the wide variety of possible configurations for the source i.e. be extensive in the source's parameter space. Finally, data analysis techniques require a large number of waveforms to be computed,

---

[1] More specifically, the object must be accelerating in a way that is not spherically or rotationally symmetric. As such, the object's quadrupole moment and the second time derivative thereof must be non-zero for it to emit gravitational radiation.



(a) The shadow of a Sgr A*

(b) The shadow of M87.

Figure 1.1: Latest images from the Event Horizon telescope of the shadows of Sgr A* and M87. (Credit: EHT Collaboration)

either beforehand to produce a waveform template bank or on the fly, and so these models must be fast to evaluate. The goal of this work is to meet the third requirement, without sacrificing the other two, when modelling waveforms from one particular LISA source, extreme mass ratio inspirals.

1.2 Extreme Mass Ratio Inspirals

Extreme mass-ratio inspirals (EMRIs) are systems which consist of a stellar mass compact object of mass μ spiralling into a massive black hole of mass M. These systems are characterised by their extremely small mass ratio, typically between 10^{-4} and 10^{-7}. Unlike the signals detected by ground-based detectors, EMRIs will radiate in the LISA frequency band for up to hundreds of thousands of orbital cycles [16]. They are also expected to be eccentric and precessing, resulting in multi-year long waveforms with rich and intricate morphologies [17].

1.2.1 Astrophysical Motivation

To motivate this work, let us consider how likely it is to find such a system in nature. One requires a massive black hole (MBH) with a mass of order $10^5 - 10^8$ times the mass of the sun $(M_\odot)$ to act as the primary. Observational evidence suggests that massive dark compact objects can be found at the centre of most bright galaxies [18, 19]. This includes our own Milky Way, where the behaviour of stars at the centre of our galaxy indicates the presence of a dark compact object, Sgr A*, with a mass of $4.4 \times 10^6 M_\odot$ [20, 21]. Moreover, as shown in Fig. 1.1 the Event Horizon Telescope released images of the shadows of Sgr A* and the compact object at the heart of galaxy M87; both of which are consistent with the shadows cast by astrophysical black holes [22, 23]. As such, we have strong evidence for the existence of massive compact objects and have good reasons to accept the hypothesis that they can be described using the black hole solutions of GR [24].

One also requires a secondary with a mass of $\sim 1M_\odot - 100M_\odot$. While main sequence stars might meet this criteria, such objects will be torn apart by tidal forces long before reaching the last stable orbit [25]. However, compact objects such as white dwarfs, neutron stars and stellar black holes will remain intact throughout the entire inspiral. There is an abundance of direct observational evidence for white dwarfs and neutron stars, as well as indirect evidence for stellar mass black holes from x-ray binaries [26] and gravitational wave observations [4–6].

Since we have good reason to believe that the ingredients to form an EMRI exist, how might they be brought together? MBHs are often observed to be surrounded by a cluster of stars. Through multiple two body interactions, this population becomes segregated by mass, with heavier objects "sinking" towards the galactic centre faster than lighter ones. As such, compact objects should sink towards the galactic centre where they may be perturbed onto highly eccentric orbits around the MBH. If one of these orbits is in the "loss cone", its point of closest approach (or periapsis) will be close enough to the MBH such that the gravitational wave emission from the compact object passing by will release sufficient energy for it to be captured and form an EMRI. This is expected to be the primary formation channel for EMRIs meaning that we would expect most EMRIs to have significant residual eccentricity and be generically inclined with respect to the MBH's equatorial plane i.e. "generic" inspirals.

Another plausible scenario is a compact binary pair coming too close to the MBH and becoming tidally disrupted. One partner gets ejected from the galactic centre, while the other becomes captured and forms an EMRI. We already have evidence of this same mechanism occurring with binary stars in the Milky way, with one partner getting captured by Sgr A* while the other is ejected as a hypervelocity star [27]. EMRIs from this formation channel are likely to have circularized by the time they enter the LISA frequency band but still be generically inclined with respect to the MBH's equatorial plane and thus likely to form "spherical" inspirals.

A third scenario comes from a star forming in the accretion disk of the MBH and collapsing into a compact object at the end of its life. EMRIs formed this way will be aligned with the accretion disk and thus equatorial plane of the MBH and are not expected to have any significant eccentricity, thus forming equatorial, quasi-circular inspirals [15, 28, 29]. Between all these formation channels, we expect to observe anywhere between 1 - 1000 EMRIs over the course of the LISA mission, with the large uncertainty coming primarily from the estimated event rate for EMRI capture [16, 30].

1.2.2 Science implications

As we can see, the hypothesis that extreme mass ratios exist in the universe is in good standing, but mere existence is not a sufficient motivation for modelling these systems with high precision. EMRIs will undergo thousands of orbits while emitting gravitational waves in the LISA frequency band, precisely mapping the spacetime of the MBH as it does so [30]. This would allow us to either confirm or refute whether the exterior of astrophysical black holes are accurately described by the Kerr metric, test the no-hair conjecture and rule out some exotic compact objects that might mimic black holes like boson stars. Observation of EMRIs can also constrain the existence of effects only present in alternative theories of gravity such as extra gravitational wave polarizations, gravitational parity violations, or the breaking of Lorenz invariance [31, 32]. Moreover, EMRIs will allow for the most precise measurements of MBH masses and spins yet and such information is vital to understanding the nature of galactic centres and the evolution of galaxies [33]. Finally, the loudest and best localized EMRIs can be used as dark standard sirens to provide an independent measurement of the expansion rate of the universe [34].

The large number of gravitational wave cycles promises to provide some incredibly important science, but it is a double edged sword. As we will see, precisely modelling the relativistic two body problem over such a long timescale will prove to be an arduous task.

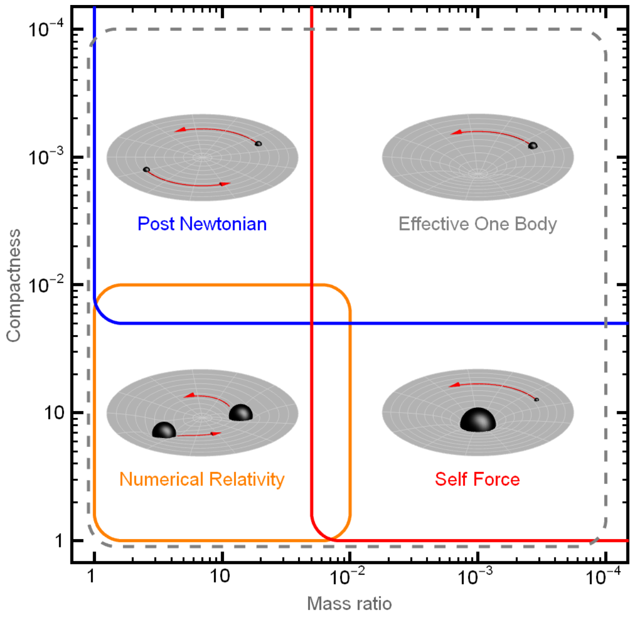


Figure 1.2: An illustration of the parameter space of the two body problem in GR with various approximation schemes overlaid onto their domains of validity. (Credit: By Maarten van de Meent - Own work, CC BY-SA 4.0, https://commons.wikimedia.org/w/index.php?curid=42495018)

1.3 Approaches to the Relativistic Two Body Problem

In Newtonian gravity, the two body problem can be easily solved and has exact, analytic solutions. This is not the case for two black holes interacting in General Relativity, as the non-linearities of the theory give rise to a much more dynamic system where one must solve not just for the positions of the two bodies, but also the entire past and future of the spacetime including the gravitational waves. Instead of the stable periodic orbits found in Newtonian gravity, the two black holes first slowly inspiral towards each other, then merge to form a single body which then rings down, settling back to a stable black hole solution, all while radiating gravitational waves. As a result, analytic solutions for these systems do not exist and so one must employ approximate methods. Each of these approaches are best suited to a different region of the parameter space, though there are regions of overlap which can be used to synergise multiple approaches. Before discussing in depth the approach that is best suited to EMRIs, we will first outline some of the other approaches most commonly practised today.

The oldest of these is post-Newtonian (PN) theory, arguably starting with Einstein's calculation of the perihelion precession of Mercury [35]. In this approach, one performs a perturbative expansion in powers of $(v/c)^2$, where v is velocity and c is the speed of light [36]. As such, this approximation works best for low velocities, which corresponds to large orbital separations i.e. the "weak field", but is ill-suited for modelling "strong field" dynamics such as the late inspiral, merger and ringdown phases. PN theory has had many successes, most famously the prediction of the change in the orbital period of binary pulsars like the Hulse-Taylor binary: PSR B1913+16 [37]. It is also invaluable for current gravitational wave searches through informing, benchmarking and hybridising with other approaches [38]. In spite of its age, PN theory is still a very active field of research, with state of the art calculations reaching beyond 4th PN order and incorporating techniques from effective field theory [39, 40].

Post-Minkowskian (PM) theory is a similar approach, where one expands in powers of the deviation of the metric from flat spacetime, often expressed as a power series in Newton's gravitational constant G [41]. This expansion is valid for large velocities, making it ideal for modelling hyperbolic encounters of two black holes [42]. However, like PN, it is ill-suited for the strong field dynamics one would expect from EMRIs, but it can synergise with and improve other approaches. This approach has become a very active subfield in recent years, due to the incorporation of techniques primarily used for calculating scattering amplitudes of particles in quantum field theory [43–45].

Instead of analytic expansions, one could decompose the spacetime into a 3+1 dimensional split and formulate the Einstein Field Equations into an initial value problem to be solved numerically on a super-computer. This is the approach taken by Numerical Relativity (NR). Initially, NR could only simulate single short lived two body problems like head on collisions [46, 47]. This all changed with a series of landmark papers in the mid 2000s, which allowed for long term simulations of quasi-circular black hole binaries [48–50]. Since then, catalogues of thousands of binaries have been computed, with additional physics like orbital eccentricity, spin precession and unequal masses now being incorporated into these simulations [51]. NR is best suited for capturing the strong field dynamics and can accurately simulate late inspiral, merger and ringdown. While they could in principle simulate the early inspiral as well, the computational cost of simulating long waveforms is unpractical as the number of orbital cycles before coalescence increases with the initial orbital separation of the simulated system. For reference, an NR waveform consisting a black hole binary with one twice as large as the other with moderate spins lasting only 10 orbital cycles takes on the order of 100,000 CPU hours to simulate [52]. For this reason, fast to evaluate surrogate models, which interpolate between NR simulations to cover more of the parameter space, are used for GW data analysis [53, 54]. Moreover, while recent work has tried to extend NR to large mass ratios [55–57], NR struggles with the discrepancy in spacial scales resulting from one body being much smaller than the other. For these two reasons NR is unsuitable for simulating EMRIs. Despite this, NR and is the gold standard for strong field dynamics and thus can inform analytic approaches, and continues to play a vital role GW data analysis.

The final approach takes inspiration from how one might simplify the Newtonian two problem by recasting it as an effective one body (EOB) system. The EOB approach in GR treats the secondary as following a geodesic in an effective spacetime characterised by a perturbed Hamiltonian [58]. The form of this Hamiltonian has free parameters which are informed by resumming PN and PM expansions. The resulting inspiral waveforms are then matched with analytic results for the ringdown, resulting in complete waveforms which cover inspiral, merger and ringdown phases [59]. NR is then used to fine tune the resuming to ensure greater accuracy [60]. In principle EOB can be used to generate full, accurate waveforms anywhere in the two-body problem parameter space in less than a second, and so has played a crucial role in LIGO-Virgo-Kagra (LVK) data analysis [61]. Even though it currently lacks the necessary information from the extreme mass ratio limit to accurately model EMRIs, recent work promises to rectify this in the near future [62, 63]. While not a self-consistent theory in its own right, it is arguably the most successful approach to date and its ability to incorporate new results as they become available will only improve it further.

As we see from Fig. 1.2, the approach best suited to EMRIs has yet to be discussed. Due to the central role this approach plays in this work, we will provide a more in depth discussion.

1.4 Gravitational Self-Force

The idea behind this approach is to first view the CO secondary as moving through the spacetime of the MBH primary. If the secondary had no gravitational field of its own, it would simply move along a time-like geodesic trajectory, just like a test body. Since this is not the case, the secondary will experience a force which will move it away from geodesic motion due to the presence of its own gravitational field. This is known as the “gravitational self-force” (GSF).

This force can be calculated perturbatively by expanding the metric of the binary around the metric

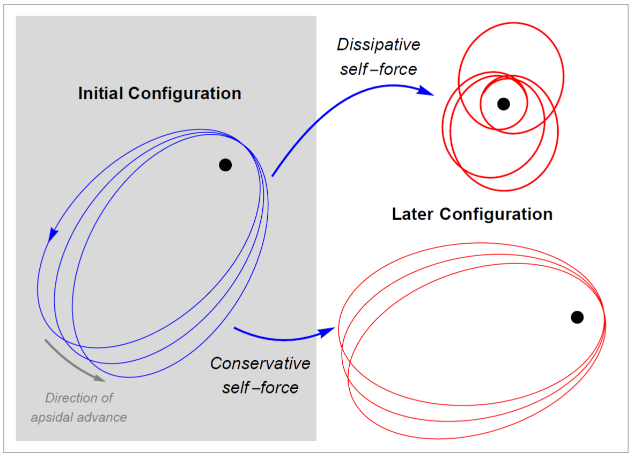


Figure 1.3: An illustration of the difference between the dissipative and conservative effects of the gravitational self force on the inspiral of a compact object around a Schwarzschild black hole. The dissipative force causes a secular decay of energy and angular momentum, causing the orbit to shrink and circularise. The conservative effects slow the rate of periapsis advance and causes small periodic changes to the energy and angular momentum. Image credit: [70]

of the primary in powers of the small mass ratio $\epsilon = \mu/M$, i.e., $g_{\mu\nu} = \bar{g}_{\mu\nu} + \epsilon h^{(1)}_{\mu\nu} + \epsilon^2 h^{(2)}_{\mu\nu} + \ldots$ where $\bar{g}_{\mu\nu}$ is the Kerr metric and the $h^{(n)}$ are the n-th order perturbations to the spacetime due to the presence of the secondary. Using matched asymptotic expansions, one can then derive the interaction between these metric perturbations and the motion of the secondary [64]. Unlike PN or PM expansions, there is no assumption about the velocity of the secondary or the strength of the gravitational interaction, meaning that this approach is valid for both weak and strong field dynamics. While this expansion works best for EMRIs where ϵ is extremely small, recent work comparing with NR has indicated that this approach can even work remarkably well, even for mass ratios of 1:10. [65, 66].

How far one needs to go in this expansion will depend on the desired accuracy requirement in the waveform phase. This was first formally derived using a two-timescale analysis in Ref. [67], where one defines a fast time t and a slow time $\mathcal{T} := \epsilon t$ which are then treated as independent variables. They found the orbital phase could be expanded in the following way

$$\varphi = \epsilon^{-1}\varphi^{\mathsf{Adb}} + \epsilon^{-1/2}\varphi^{\mathsf{Res}} + \varphi^{\mathsf{PA}} + \mathcal{O}(\epsilon). \tag{1.1}$$

A leading order adiabatic (Adb) waveform evolved for $\mathcal{O}(\epsilon^{-1})$ orbital cycles will accumulate a phase error $\mathcal{O}(1)$ radians. Moreover, for inclined and eccentric inspirals into a rotating black hole, it is possible for the inspiral to pass through an orbital resonance, and so one must accurately model the contribution to the orbital phase caused by such an encounter φ^{Res}, or be left with an additional phase error of $\mathcal{O}(\epsilon^{-1/2})$ [68]. A waveform model that includes these effects might be sufficient for detecting the presence of EMRIs in LISA data, but will not be accurate enough to achieve the LISA data science goals listed in Sec. 1.2.2. That will require waveforms which are accurate to roughly $1/\rho$ radians where ρ is the signal to noise ratio (SNR) [30, 67, 69]. As such, one must produce post-adiabatic (PA) waveforms whose phase error should be $\mathcal{O}(\epsilon)$.

To understand what we need for a post-adiabatic waveforms, we must first note that the effects of the self-force can be split into dissipative (time anti-symmetric) and conservative (time-symmetric) contributions [71]. (This is true for first order GSF but is unclear if this split can be done uniquely at second order as there are different definitions of conservative used for second-order GSF in the literature [72, 73]). The differences in conservative and dissipative effects are illustrated in Fig. 1.3.

The dissipative pieces of the self-force cause a loss in energy, angular momentum and Carter constant, causing the orbit to shrink until the secondary plunges into the primary. It generally causes the orbit to circularize, with the exception being just before the transition to plunge where the orbit gains eccentricity [74–77]. It also causes the orbital inclination with respect to the equatorial plane of the secondary to slightly increase over time [78–80]. The conservative pieces have more subtle effects on the inspiral, such as altering the rate of periapsis advance, changing the orbital frequencies and moving the location of the innermost stable circular orbit [81–87]. They can also cause the energy and angular momentum to change periodically, but such changes average to zero over a single geodesic orbit.

It has been found that computing post-adiabatic waveforms will require knowledge of both the dissipative and conservative pieces of the first-order self-force and the orbit average piece of the second-order self-force [67]. As such, one must expand the metric to at least second-order in the mass ratio, which has recently provided many important results for quasi-circular Schwarzschild inspirals [66, 88–90]. While the groundwork is being laid to extend second-order GSF to Kerr inspirals [91, 92], there are no results available as of yet. As such, we restrict ourselves only to first-order in the mass ratio results, even though this will result in less accurate inspirals.

At each instant, the self-force is a functional of the past history of the secondary which can make it challenging to compute. One approach is to couple the field equations and the equations of motion (EOM) and self-consistently solve both in a time-domain simulation. While this has been implemented for a toy model of a particle carrying a scalar charge orbiting a Schwarzschild black hole [93], numerical stability issues in the Lorenz gauge have so far stifled similar attempts for the gravitational case [94]. Moreover, this approach is computationally very expensive, making it impractical for generating large numbers of templates. However, it does promise waveforms against which more efficient schemes could be tested.

An alternative method is to compute the self-force for a body moving along fixed geodesics of the background spacetime and then use that force to move to another geodesic at a later time-step. The periodic nature of these geodesic orbits allows for calculations in frequency domain leading to many efficient calculations of the first-order self-force in both Schwarzschild [95–98] and Kerr [99, 100] spacetimes.

Solving the perturbation equations requires picking a gauge, and the resulting self-force is gauge dependent [101]. The self-force was first computed in the Lorenz gauge where the procedure for obtaining the regular metric perturbation part was best understood. Numerical calculations of the Lorenz gauge self-force have been made in both the frequency- [97, 98, 102] and time-domains [94–96] All these results have been for motion in Schwarzschild spacetime, with one exception in Kerr [103].

Calculating perturbations of the Kerr spacetime is hampered by the lack of separability of the linearized Einstein Field Equations on this background. This difficulty can be circumvented by using the Teukolsky formalism for describing perturbations to the Weyl scalars [104], which is fully separable in the frequency domain. From the Weyl scalars, the metric perturbation can be reconstructed in a radiation gauge [105–107]. There has also been recent progress understanding how to reconstruct the metric in the Lorenz gauge [108]. Regularization of the metric perturbation in radiation gauges is more subtle [109], but self-force calculations in the radiation gauge are now routine [99, 100, 110, 111].

As successful as this method has been, it is worth noting that the "geodesic self-force" approach is difficult to extend to second order in the mass ratio, as the source should now be on an inspiral trajectory instead of a geodesic. A more promising approach involves taking the two-timescale analysis of Ref. [67] and applying it to not only the equation of motion, but also to the Einstein Field Equations [112, 113]. This two-timescale framework is the current state of the art in GSF calculations and has been the only approach to produce second-order GSF effects [88, 89], post-adiabatic waveforms [66] and slow-time derivatives of the metric perturbation [90].

1.5 Modelling Extreme Mass Ratio Inspirals

Before describing how we have used interpolated GSF models to produce fast and accurate EMRI waveforms, we will first review the requirements for waveform models and the current attempts to meet these requirements. The majority of EMRIs will have a very low instantaneous signal-to-noise ratio (SNR), and so the data must be processed with matched filtering techniques which will allow for the build up of the SNR over time [114]. Such techniques require the development of theoretical waveform templates to compare against the data. To achieve LISA's science objectives, these templates need their phase to be accurate to within $\mathcal{O}(10^{-2})$ radians, even after hundreds of thousands of orbital cycles [69]. They also need to be extensive across the large parameter space of possible EMRI configurations. Moreover, since many template evaluations would be needed, they should also be fast to compute, ideally in less than a second.

1.5.1 Kludge Models

To meet this challenge, several so-called "kludge" models have been developed, which are both extensive and quick to compute [115–118]. These models come in two flavours; analytic and numeric.

Analytic models are the fastest to compute, taking only a matter of milliseconds to produce a waveform. The trajectory is modelled as an eccentric Keplerian orbit with the relativistic features of pericentre advance, Lense-Thiring precession and inspiral all approximated with an analytic formula. The waveform is then generated from the trajectory using the Peters-Mathews formula [119] for quadrupole radiation from two point masses on a Keplerian orbit. The "Analytic Kludge" (AK) was the first model to adopt this approach [116]. This has since been improved upon by the "Augmented Analytic Kludge" (AAK) which accurately maps to the Kerr geodesic orbital frequencies and uses updated evolution equations [118]. This results in significantly more accurate waveforms at almost no additional computational cost, and so AAK is still in use today [120].

The second flavour is the "Numeric Kludge" (NK) [116]. This approach assumes the secondary is on a geodesic in the Kerr background which then slowly loses energy, angular momentum, and Carter constant which is described by analytic PN approximations. One has to numerically solve for the evolution of the energy angular momentum and Carter constant from which one can then construct the trajectory. One then makes the semi-relativistic approximation of mapping solutions for the Boyer-Lindquist coordinates to flat-space coordinates and then using the flat-space quadrupole formula to generate the waveform from the inspiral trajectory. While this is more accurate than the AK or AAK, NK waveforms can take on the order of tens of seconds to compute.

While kludge models are fast and extensive in the parameter space, their use of non-relativistic approximations cause them to fall well short of the subradian accuracy requirement. Despite their shortcomings, these models may still have a role to play in detecting loud EMRI signals [121], and have proven to be invaluable for testing data analysis techniques for LISA through the mock data challenges [122–124].

1.5.2 Adiabatic Models

While fully relativistic, adiabatic waveforms would still not reach the subradian accuracy goal, one could use them to robustly search for EMRIs in LISA data [125, 126]. This would only require knowledge of the orbit averaged dissipative pieces of the first-order self-force. For energy and angular momentum these can be related, via balance laws, to the fluxes of GWs to infinity and down the event horizon. The same cannot be done for the rate of change of the Carter constant [126], but Ref. [127] found that this can be related to the average action of the dissipative self-force on a geodesic. Adiabatic inspirals are typically calculated using these calculations as they avoid having to regularize the metric perturbation [125,126,128,129]. The fluxes are calculated from solutions to the Regge-Wheeler equation in Schwarzschild [130] or the Teukolsky equation in Kerr [104], for perturbations sourced by a test mass on a geodesic orbit. Adiabatic inspirals have been calculated for quasi-circular [131] and eccentric [74] Schwarzschild inspirals, as well as quasi-circular [132, 133], eccentric [115], spherical [79, 80] and most recently generic [134] Kerr inspirals. However, only a subsection of the generic Kerr parameter space was explored, and so work remains to tile the rest of the parameter space with GW flux data. In particular no model has interpolated over the spin of the primary. Such inspirals are fast to compute, typically taking a few milliseconds, but they are not yet fully extensive in the parameter space.

As well as the fluxes, one also requires relativistic amplitudes from the Regge-Wheeler/ Teukolsky equations to compute fully relativistic waveforms. Constructing these waveforms requires summing over a large number of modes, which is very computationally expensive when this has to be repeated for every time-step in a multi-year long EMRI inspiral. The FastEMRIWaveforms (FEW) package overcomes this by employing neural networks and GPUs to speed up the waveform calculation, and can currently produce fully relativistic waveforms for eccentric Schwarzschild inspirals in a matter of milliseconds [120, 135]. While work needs to be done to extend this framework to Kerr inspirals, it is currently fastest and most accurate approach available for EMRI waveforms.

It is worth noting that an adiabatic model for EMRIs that is extensive in the parameter space has recently been made available [77]. This approach uses PN expansions of the solutions to the Teukolsky equation to recover the gravitational wave amplitudes, the GW fluxes to 5PN order and even dissipative resonant effects. While the model can in principle cover all of the generic Kerr parameter space, the current implementation works best in the weak field but struggles with very small separations or very large eccentricities, inclinations, or spins. However, it is a significant improvement over current kludge models and is well suited for LISA data analysis studies.

1.5.3 Osculating Geodesics

In order to go beyond adiabatic waveforms, one must understand how the trajectory is effected by the local GSF. One can do this using the geodesic equations with a forcing term on the right hand side, resulting in a set of coupled second-order ordinary differential equations (ODEs) for the evolution of the coordinates of the secondary. This can be recast into a series of first-order ODEs by using the long standing method of osculating orbital elements (or variation of constants) from Newtonian celestial mechanics. When applied to the curved spacetime around a black hole, this is known as the "method of osculating geodesics" [136]. In this approach, the inspiral is described as a smooth evolution through neighbouring geodesics that are instantaneously tangent to the true inspiral. Formally, one identifies a set of constants of motion which uniquely identify a geodesic, known as "orbital elements". These constants are then promoted to functions of time which are governed by a set coupled first-order ODEs that are derived from the "osculating conditions". These new equations of motion are then solved numerically to obtain the inspiral trajectory of the secondary. There are a number of equivalent formulations for these ODEs which have been derived for both Schwarzschild [136] and Kerr [113, 137] inspirals. In this work, we make use of a formulation based on action angles of the geodesic motion that was first sketched out in Ref. [137].

This method does assume that the orbit is evolving slowly or "adiabatically", and so this technique is only applicable for the inspiral part of the EMRI and breaks down when the secondary goes beyond the last stable orbit and plunges into the primary [138–142]. However, due to the EMRIs long time in the LISA band, the inspiral is expected to make up the vast majority of the accumulated SNR.

Importantly, the method does not assume that the force is small, and so it can be used for forces other than just the GSF. As such, this method has been used to model other effects on the binary such as gas drag [137], the presence of a third body acting as external perturber [143] and the spin-curvature force due to the spin of the secondary [144].

1.5.4 Self-Forced Inspirals

For this work, we are most interested in efficiently modelling the trajectory of an EMRI under the influence of the GSF. However, GSF calculations can take dozens of CPU hours to compute for a single geodesic orbit. As such, the prevailing approach has been to use precomputed GSF data across a section of the geodesic orbit parameter space to inform a GSF model that can be rapidly evaluated.

The first such model in this vein was computed for the first-order Lorenz gauge GSF for eccentric Schwarzschild inspirals [145]. The model decomposed each component of the force into a Fourier series and modelled the Fourier coefficients using a global fit to a power series in the orbital elements using least squares fitting. The resulting model was fast to compute, but used 1100 data points to achieve had a fractional accuracy of $< 10^{-3}$ across the parameter space and was only valid for eccentricities ≤ 0.2.

This model has since been surpassed by a model which also rapidly computes the first-order Lorenz gauge GSF for eccentric Schwarzschild inspirals, but can cover eccentricities ≤ 0.8 and much larger orbital separations [70]. Moreover, it was able to achieve a fractional accuracy of $\leq 10^{-8}$ for the adiabatic pieces of the force and $\leq 10^{-3}$ for the post-adiabatic pieces. This was achieved using local Chebyshev interpolation for subdomains of the parameter space and incorporating GW flux data from a high precision Regge-Wheeler code. However, it required 43875 data points from the flux code and 9602 data points from a Lorenz gauge self-force code.

While this is manageable for the two dimensional parameter space of eccentric Schwarzschild inspirals, generic Kerr inspirals have four parameters to tile over: orbital separation, eccentricity, inclination and spin of the primary. Using this same approach for generic Kerr inspirals would require $\mathcal{O}(10^9)$ flux data points and $\mathcal{O}(10^8)$ GSF data points respectively. This becomes even more intractable when one considers how much more computationally expensive Kerr GSF codes are compared to Schwarzschild codes.

For example, in this work we use the code of Ref. [99, 100, 111] to produce gravitational self force data for eccentric equatorial orbits and spherical orbits in Kerr spacetime. This code uses the Mano-Suzuki-Takasugi methods [146–149] to compute the perturbations to the Weyl scalars in the frequency domain. The metric is then reconstructed into an outgoing radiation gauge (including mass and angular momentum perturbations [150, 151] and gauge completion contributions [152]). The metric perturbation is then projected onto a basis of spherical harmonics before regularization is carried out using the mode-sum approach [109, 153].

Depending on the eccentricity or the inclination of the orbit, the code must compute the metric perturbation by summing over thousands to tens of thousands of Fourier harmonic modes. With the current Mathematica implementation, the self-force for a typical eccentric equatorial orbit, e.g., $a = 0.9M, p = 3.375, e = 0.5$ takes approximately 90 CPU hours to compute. As such, this code is much too slow to utilize the same interpolation techniques used for Schwarzschild inspirals and this has motivated us to find a more efficient interpolation method which will allow us to dramatically reduce the number of points needed.

For this work, we will produce global fits to the gravitational self force using the pseudo-spectral method of Chebyshev interpolation, instead of the many local fits which have been used previously [70]. This will allow us to produce the very first GSF models for eccentric and spherical Kerr inspirals, capable of achieving the $\leq 10^{-3}$ fractional accuracy using only $\mathcal{O}(10^2)$ GSF data points to cover a two dimensional slice of the parameter space. If this method were naively extended to the full generic Kerr parameter space, one would expect to need only $\mathcal{O}(10^4)$ data points, which is far more feasible.

Using any of these GSF models along with the osculating geodesic equations allows one to compute self-forced inspirals. In the quasi-circular case, these inspiral calculations are rapid, as the evolution of the orbital elements is decoupled from the evolution of the azimuthal phase due to the cylindrical symmetry of the Kerr spacetime. However, if the orbit is eccentric and/or inclined, then the evolution of the orbital elements is coupled with the evolution of the radial and/or polar phases. This results in the numerical solver taking small step sizes to resolve the orbital timescale physics in order to accurately calculate the long term secular evolution of the inspiral. This means that a single multi-year long EMRI inspiral can take minutes or hours to compute, which would make data analysis with such inspirals intractable.

1.5.5 Near-Identity Averaging Transformations

Following Ref. [154], we overcome this problem by applying a near-identity (averaging) transformation (NIT) [155] to the self-forced equations of motion. These transformed EOM have two important properties: (i) they no longer depend on the orbital phase, and (ii) they capture the long-term secular evolution of the original inspiral to the same order of approximation in the mass ratio as the original EOM. The first property means the transformed equation of motion can be numerically solved for any mass ratio in less than a second as the numerical integrator no longer needs to resolve the hundreds of thousands of orbital timescale oscillations.

Ref. [154] laid out how this approach can be used for a generic EMRI system and was then implemented for the case of eccentric self-forced inspirals in Schwarzschild [154]. This was further extended to highly eccentric Schwarzschild inspirals, and the relationship between the NIT orbital phases and the waveform "voices" was also established [156].

This work marks the first time this technique has been implemented for Kerr inspirals. Combined with an interpolated model of first-order gravitational self-force data, these averaged equations of motion allow us to efficiently compute equatorial and eccentric or spherical inspirals around a Kerr black hole which include all first order in the mass ratio effects. Since these inspirals are fast to compute, this approach can provide EMRI waveforms which could feasibly be used for practical data analysis when coupled to a fast waveform generation scheme, e.g., the FEW framework [135].

At this point, we emphasize that the inspirals we present do not reach the sub-radian accuracy required for EMRI data analysis as there are not yet any second-order self-force calculations in this domain. To stress the importance of the second-order contribution we examine the effects of driving the inspirals using first-order self-forces computed in two different gauges and demonstrate explicitly that without the inclusion of the second-order self-force the inspiral phase is not gauge invariant. Nonetheless, the framework we present can readily incorporate new self-force results as they become available.

Current implementations of the NIT EOM are not parametrised in terms of the Boyer-Lindquist time coordinate, which can be related to the time measured at the detector. Thus, time domain waveform calculations using these inspirals require an inconvenient, additional interpolation step. Following the procedure detailed in Ref. [113], we lay out how one can perform an additional averaging transformation to averaged EOM that are parametrized in terms of Boyer-Lindquist time.

A related framework for efficient inspiral calculations is the two-timescale expansion (TTE) [67,155]. This also allows one to average over the orbital timescale physics while accurately capturing the long term evolution of the system. This comes at the expense of doubling the number of ODEs one has to solve compared to the NIT EOM, but the resulting solutions are valid for all values of the mass ratio ϵ. It is important to note that the TTE EOM can be directly related to the NIT EOM [113,155], as the TTE is currently the framework employed for second-order GSF calculations [66,90,112,113]. Following the direction of Ref. [113], we perform a TTE on the NIT EOM and in order to investigate the practicalities of these calculations for self-forced inspirals for the first time.

1.5.6 Orbital Resonances

It is worth noting that for generic Kerr inspirals, near-identity transformations become ill-defined when the radial and polar frequencies are in a small integer ratio to one another, i.e. at an orbital resonance. While sustained resonances for EMRIs are highly unlikely [157], most EMRIs are expected to encounter at least one transient resonance while in the LISA band which will have a significant effect on resulting trajectory and waveform [158]. While passing through a resonance, the orbital elements evolve on a timescale that is slower than the orbital timescale but faster than the radiation reaction timescale. The gravitational wave flux can also be diminished or enhanced depending on the phase of the inspiral as it enters resonance [159]. By the time the resonance has past, the orbital elements have changed by a factor of $\mathcal{O}(\epsilon^{1/2})$, known as a resonance "jump". Not accounting for these jumps can result in a phase error $\mathcal{O}(\epsilon^{-1/2})$ as seen in Eq. (1.1) [68].

Orbital resonances were first explored using the method of osculating geodesics along with a PN inspired GSF model and comparing the OG inspirals to adiabatic inspirals [68]. This model was also applied to an astrophysically motivated EMRI population and it was found that a failure to model transient resonances could result in a loss of 4% of EMRI signals detected by LISA [160].

An effective resonance model has recently been developed, which modifies the numerical kludge with resonant jumps whose magnitudes are left as free parameters [161]. Using this as a phenomenological model, one can recover the size of the resonant jump from year long EMRI signal with SNR 20 with a relative precision of 10%. Comparing the recovered values with those predicted by GR would be a valuable test of fundamental physics, e.g. the Kerr hypothesis [24].

Recent work has also examined the impact of third body acting as an external tidal perturber, which have been found to have a negligible impact on the EMRI's trajectory except when the radial, polar and azimuthal frequencies are in a small integer ratio to each other i.e. a tidal resonance [143, 162]. Combining knowledge of the jumps induced by a tidal resonance with the effective resonance model has resulted in a fast and accurate model for tidal resonances [163]. This same model can be extended to model self-force transient resonances as soon as information for the first order generic Kerr GSF has been calculated across the parameter space.

In this work, we explore the problem of transient resonances using a generic Kerr toy model informed by the first-order self-force models for eccentric and spherical inspirals. We then modify the NIT to average out all phases except for the resonant phase in the vicinity of an orbital resonance. We demonstrate how one can transition from the non-resonant NIT to the near-resonance NIT and back again in order to efficiently evolve through multiple resonances for the first time.

1.6 Organisation of this thesis

The next four chapters serve as theoretical background for the rest of the thesis. Chapter 2 reviews geodesic motion in Kerr spacetime. Chapter 3 reviews the method of osculating geodesics and contains an original derivation of the action angle formulation which is then used throughout this work. Chapter 4 contains a review of near-identity averaging transformations, followed by a new derivation for a near-identity averaging transformation in the vicinity of an orbital resonance. We also review how one can perform additional averaging transformation such that the equations are parametrized in terms of the Boyer-Lindquist time coordinate and connections to the two-timescale expansion which were already covered in Ref. [113]. Chapter 5 contains a review of various waveform generation schemes for EMRIs and techniques for comparing waveforms.

The next three chapters describe practical implementations and results which are all original work. Chapter 6 focuses solely on eccentric, equatorial Kerr inspirals. We describe how we interpolated GSF data to produce the first eccentric Kerr GSF model and how this can be combined with the OG equations to produce the first self-forced inspirals in Kerr spacetime. We then outline the practical steps involved in our NIT implementation. We then present the results of this implementation including consistency checks, tests the speed and accuracy of the NIT inspirals, an exploration impact of adiabatic and post-adiabatic effects on the inspiral trajectory, and a comparison inspirals using self-force models calculated in two different gauges. The results in this chapter have already been published in Ref. [1].

Chapter 7 concentrates on spherical Kerr inspirals. We describe how we interpolated GSF data to produce the first spherical Kerr GSF model, and how this can be combined with the OG equations to produce the first spherical Kerr self-forced inspirals. In this chapter, we make use of an additional averaging transformation such that the resulting inspiral is parametrized in terms of Boyer-Lindquist time and relate these new averaged EOM to the TTE EOM. We then outline the practical steps involved in our NIT/ TTE implementation. Finally, we present the results of this implementation including tests the speed and accuracy of the NIT and TTE inspirals, an assessment of the impact from incorporating a high precision flux model on the inspiral trajectory and an exploration of the impact of the first-order self-force for spherical inspirals. The results in this chapter not yet been published but will be soon.

Chapter 8 deals with generic Kerr inspirals and orbital resonances. First we outline how we construct a generic Kerr GSF toy model from the eccentric and spherical GSF models. Then we implement the NIT far from resonances and test its speed and convergence. We then implement the near-resonant NIT in the vicinity of a low order resonance and perform the same speed and convergence tests. We follow this by outlining how one can transition from a non-resonant NIT to a near-resonant NIT and back again. We then test the accuracy and speed of this approach using two test inspirals; one that evolves through a single orbital resonance and one that evolves through two. This is gives us a promising proof of concept for efficiently evolving through multiple orbital resonances. However, there is still a considerable amount of work that needs to be done before the results in this chapter can be published.

We conclude with chapter 9, where we start by outlining the conclusions associated with chapters 6, 7 and 8. We then discuss the limitations of this work and how it might be extended in the future. Finally, since it does not fit within the narrative flow of the main text, we briefly discuss GSF corrections to the periapsis advance in Kerr spacetime in Appendix A.

Throughout this work we use the metric signature $(-+++)$ and work in geometrized units such that the gravitational constant and the speed of light are equal to unity i.e. $G = c = 1$.

Chapter 2

Geodesic Motion

In this work, we wish to describe the motion of a non-spinning compact object of mass μ moving in the Kerr spacetime under the influence of some arbitrary force. This will eventually be taken to be the self-force experienced by the secondary due to its interaction with its own metric perturbation. We denote the mass of the primary by M and parametrize its spin by $a = |J|/M$ where J is the angular momentum of the black hole. The Kerr metric can then be written in modified Boyer-Lindquist coordinates, $x^\alpha = \{t, r, z = \cos\theta, \phi\}$, as

$$\begin{aligned} ds^2 = &-\left(1 - \frac{2Mr}{\Sigma}\right) dt^2 + \frac{\Sigma}{\Delta} dr^2 + \frac{\Sigma}{1 - z^2} dz^2 \\ &+ \frac{1 - z^2}{\Sigma}(2a^2 r(1 - z^2) + \Sigma\varpi^2) d\phi^2 - \frac{4Mar(1 - z^2)}{\Sigma} dt d\phi \end{aligned} \tag{2.1}$$

where

$$\Delta(r) := r^2 + a^2 - 2Mr, \quad \Sigma(r, z) := r^2 + a^2 z^2, \quad \varpi(r) := \sqrt{r^2 + a^2}, \tag{2.2a-c}$$

If a force acts upon the secondary its motion can be described by the forced geodesic equation

$$u^\beta \nabla_\beta u^\alpha = a^\alpha \tag{2.3}$$

where $u^\alpha = dx^\alpha/d\tau$ is the secondary's four-velocity, ∇_β is the covariant derivative with respect to the Kerr metric, and a^α is the secondary's four-acceleration. We seek to recast Eq. (2.3) into a form useful for applying the near-identity (averaging) transformations.

Before considering the forced equation it is useful to first examine the geodesic limit, which will be the focus of this chapter. We will review geodesic motion and all of the analytic results that will be useful for future chapters. Then, we will classify the different types of geodesic orbits and discuss their properties. Finally, we will examine the case of resonant geodesic orbits.

2.1 Geodesic Motion around a Kerr black hole

2.1.1 Geodesic motion and orbital parametrization

In the absence of any perturbing force, the secondary will follow a geodesic, i.e.,

$$u^\beta \nabla_\beta u^\alpha = 0, \tag{2.4}$$

The symmetries of the Kerr spacetime allow for the identification of integrals of motion $\vec{\mathcal{P}} = \{\mathcal{E}, \mathcal{L}, \mathcal{K}\}$ given by

$$\mathcal{E} = -u_t, \quad \mathcal{L} = u_\phi, \quad \mathcal{K} = \mathcal{K}^{\alpha\beta} u_\alpha u_\beta, \tag{2.5a-c}$$

where $\mathcal{K}^{\alpha\beta}$ is the Killing tensor, $\mathcal{E}$ is the orbital energy per unit rest mass μ, $\mathcal{L}$ is the z-component of the angular momentum divided by μ and $\mathcal{K}$ is the Carter constant divided by μ^2 [164]. This definition of the Carter constant is related to another common definition of the Carter Constant, $\mathcal{Q}$, by

$$\mathcal{Q} = \mathcal{K} - (\mathcal{L} - a\mathcal{E})^2. \tag{2.6}$$

Using these integrals of motion, one can express the geodesic equations in first order form [165]:

$$\begin{aligned} \left(\frac{dr}{d\lambda}\right)^2 &= F^2 - \Delta\left(r^2 + \mathcal{K}\right) \\ &= (1 - \mathcal{E}^2)(r_1 - r)(r - r_2)(r - r_3)(r - r_4) := V_r \end{aligned} \tag{2.7a}$$

$$\begin{aligned} \left(\frac{dz}{d\lambda}\right)^2 &= \mathcal{Q} - z^2\left(\beta(1 - z^2) + \mathcal{L}^2 + \mathcal{Q}\right) \\ &= \left(z^2 - z_-^2\right)\left(\beta z^2 - z_+^2\right) := V_z \end{aligned} \tag{2.7b}$$

$$\frac{dt}{d\lambda} = \frac{\varpi^2}{\Delta}F - a^2\mathcal{E}(1 - z^2) + a\mathcal{L} := f_t^{(0)} \tag{2.7c}$$

$$\frac{d\phi}{d\lambda} = \frac{a}{\Delta}F + \frac{\mathcal{L}}{1 - z^2} - a\mathcal{E} := f_\phi^{(0)}. \tag{2.7d}$$

where $F = \mathcal{E}\varpi^2 - a\mathcal{L}$, $\beta = a^2(1 - \mathcal{E}^2)$, $r_1 > r_2 > r_3 > r_4$ are the roots of the radial potential V_r, $z_+ > z_-$ are the roots of the polar potential V_z, and λ is Mino(-Carter) time that decouples the radial and polar equations [128]. This time is related to the proper time of the particle, τ, by

$$d\tau = \Sigma d\lambda. \tag{2.8}$$

Rather than parametrize an orbit by the set $\vec{\mathcal{J}} = \{\mathcal{E}, \mathcal{L}, \mathcal{K}\}$ it is useful to instead use more geometric, quasi-Keplerian constants $\vec{P} = \{p, e, x\}$. Here p is the semi-latus rectum and e is the orbital eccentricity which are both related to the radial roots r_1 and r_2 via:

$$p = \frac{2r_1r_2}{M(r_1 + r_2)} \quad \text{and} \quad e = \frac{r_1 - r_2}{r_1 + r_2}. \tag{2.9a-b}$$

We also use x which is a measure of the orbital inclination given by

$$x = \cos\theta_{\text{inc}} = \sqrt{1 - z_-^2}. \tag{2.10}$$

The inclination angle θ_{inc} is related to θ_{min} (the minimum value of θ measured with respect to the primary's spin axis) by $\theta_{\text{inc}} = \pi/2 - \text{sgn}(\mathcal{L})\theta_{\text{min}}$. These angles and their relationship to each other are illustrated in Fig. 2.1. There are other common choices for inclination in the literature such as the inclination angle ι [78–80, 166] given by

$$\cos\iota = \frac{\mathcal{L}}{\sqrt{\mathcal{L}^2 + \mathcal{Q}}}. \tag{2.11}$$

However, we opt to use x as it smoothly parametrizes all inclinations between prograde equatorial motion

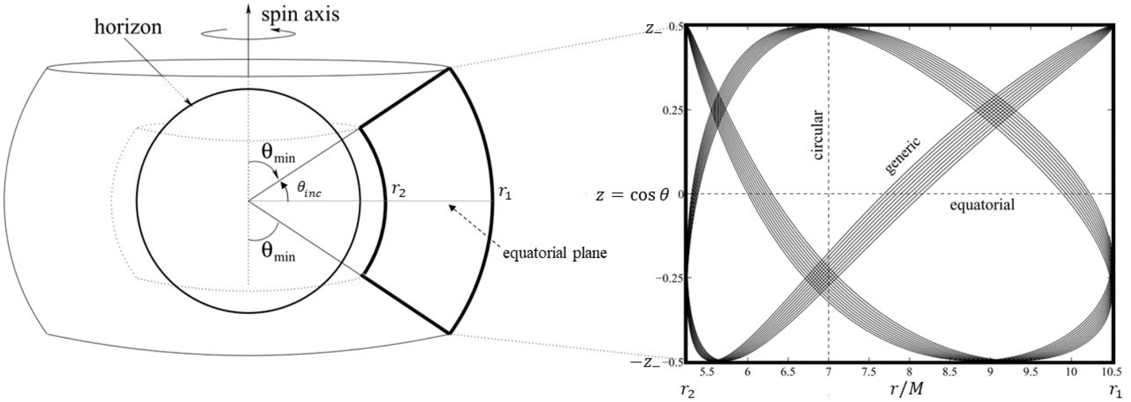


Figure 2.1: A illustration of the radial and polar motion present for generic Kerr orbits. The graph shows how the motion fills a torus which is bounded by the radial roots r_1 and r_2 and the value of the polar root z_-. This illustration was taken from [167] with minor alterations.

where $x = 1$ to retrograde equatorial motion where $x = -1$. Another advantage of using $\{p, e, x\}$ is that they are easily related to the radial and polar roots via

$$r_1 = \frac{pM}{1-e}, \quad r_2 = \frac{pM}{1+e}, \quad z_- = \sqrt{1-x^2}, \tag{2.12a-c}$$

As we can see in Fig. 2.1, the radial motion is bounded between a maximum of r_1 and a minimum of r_2, while the polar motion is bounded between z_- and $-z_-$.

One can also determine $\vec{\mathcal{J}}$ in terms of $\{r_1, r_2, z_-\}$ and thus in terms of $\{p, e, x\}$ [168]. First, we define the following terms:

$$d(r) := \Sigma(r)\Delta(r), \tag{2.13a}$$

$$f(r) := r^4 + a^2(r(r+2)z_-^2\Delta(r)), \tag{2.13b}$$

$$g(r) := 2ar, \quad \text{and} \tag{2.13c}$$

$$h(r) := r(r-2) + \frac{z_m^2}{x^2}\Delta(r). \tag{2.13d}$$

We then combine them to obtain to following quantities:

$$\alpha = d(r_1)h(r_2) - d(r_2)h(r_1) \tag{2.14a}$$

$$\zeta = d(r_1)g(r_2) - d(r_2)g(r_1) \tag{2.14b}$$

$$\rho = f(r_1)h(r_2) - f(r_2)h(r_1) \tag{2.14c}$$

$$\xi = f(r_1)g(r_2) - f(r_2)g(r_1) \tag{2.14d}$$

$$\upsilon = g(r_1)h(r_2) - g(r_2)h(r_1) \tag{2.14e}$$

From these, one can then express the integrals of motion $\vec{\mathcal{J}}$ as

$$\mathcal{E} = \sqrt{\frac{\alpha\rho + 2\zeta\upsilon - 2\mathsf{sgn}(x)\sqrt{\upsilon\left(\upsilon\zeta^2 + \rho\zeta\alpha - \xi\alpha^2\right)}}{\rho^4 + 4\xi\upsilon}}, \tag{2.15a}$$

$$\mathcal{L} = \frac{-\mathcal{E}g(r_1) + \mathsf{sgn}(x)\sqrt{d(r_1)h(r_1) + \mathcal{E}^2\left(g(r_1)^2 + f(r_1)h(r_1)\right)}}{h(r_1)}, \tag{2.15b}$$

$$\mathcal{K} = z_-^2\left(\beta + \frac{\mathcal{L}^2}{x^2}\right) + (\mathcal{L} - a\mathcal{E})^2. \tag{2.15c}$$

Finally, one can determine the remaining roots of the radial and polar potentials via

$$r_3 = \frac{M}{1-\mathcal{E}^2} - \frac{r_1+r_2}{2} + \sqrt{\left(\frac{r_1+r_2}{2} - \frac{M}{1-\mathcal{E}^2}\right)^2 - \frac{a^2\mathcal{Q}}{r_1 r_2(1-\mathcal{E}^2)}} \tag{2.16a}$$

$$r_4 = \frac{a^2\mathcal{Q}}{r_1 r_2 r_3(1-\mathcal{E}^2)}, \tag{2.16b}$$

$$z_+ = \sqrt{\beta + \frac{\mathcal{L}^2}{x^2}}. \tag{2.16c}$$

As such, given values for $\{p, e, x\}$, we can determine all of the other constants that appear in the geodesic equations of motion.

It is worth noting that not all values of $\{p, e, x\}$ correspond to bound geodesics and we denote the value of p at the last stable orbit (LSO) by $p_{\mathsf{LSO}}(a, e, x)$ [75, 169]. This is also known in the literature as the separatrix. When we come to discuss the special case of spherical orbits (i.e. $e = 0$), p_{LSO} is equivalent to the the radius of the inner most stable spherical orbit (ISSO) r_{ISSO}. We also have hyperbolic orbits when $e > 1$ [170, 171].

We note that it is also common in the literature [172, 173] to express the radial and polar motion in terms of quasi-Keplerian angles, ψ and χ, via

$$r(\psi) = \frac{pM}{1+e\cos\psi} \quad \text{and} \quad z(\chi) = z_- \cos\chi. \tag{2.17a-b}$$

With this parametrization, the evolution equations for ψ and χ depend on ψ and χ respectively i.e.,

$$\frac{d\psi}{d\lambda} = \frac{M}{1-e^2}\sqrt{(1-\mathcal{E}^2)\left((p-p_3) - e(p+p_3\cos\psi)\right)\left((p-p_4) + e(p-p_4\cos\psi)\right)} := \hat{\Upsilon}_r^{(0)}(\psi), \tag{2.18a}$$

$$\frac{d\chi}{d\lambda} = \sqrt{(z_+^2 - \beta z_-^2\cos^2\chi)} := \hat{\Upsilon}_z^{(0)}(\chi), \tag{2.18b}$$

where $p_3 = r_3(1-e)/M$ and $p_4 = r_4(1+e)/M$ [174].

2.1.2 Analytic solutions for radial and polar motion

In the action angle prescription of the geodesic motion, the orbital phases $\vec{q} = \{q_r, q_z\}$ are such that the geodesic equations can written in the form

$$\frac{dP_j}{d\lambda} = 0 \quad \text{and} \quad \frac{dq_i}{d\lambda} = \Upsilon_i^{(0)}(\vec{P}), \tag{2.19a-b}$$

where $\Upsilon_i(\vec{P})$ are the Mino time fundamental frequencies. For the radial and polar motion, these frequencies are given explicitly by

$$\Upsilon_r^{(0)} = \frac{\pi}{2K(k_r)}\sqrt{(1-\mathcal{E}^2)(r_1-r_3)(r_2-r_4)}, \quad \text{and} \quad \Upsilon_z^{(0)} = \frac{\pi z_+}{2K(k_z)}, \tag{2.20a-b}$$

where

$$k_r = \frac{(r_1-r_2)(r_3-r_4)}{(r_1-r_3)(r_2-r_4)}, \quad \text{and} \quad k_z = \beta\frac{z_-^2}{z_+^2}, \tag{2.21a-b}$$

K is complete elliptic integral of the first kind given by $K(m) = F(\pi/2|m)$ and F is the incomplete elliptic integral of the first kind given by

$$F(\phi|m) := \int_0^\phi \frac{d\theta}{\sqrt{1-m\sin^2\theta}}. \tag{2.22}$$

From geodesic equations of motion, the solutions for $\vec{P}(\lambda)$ and $\vec{q}_i(\lambda)$ are simply given by

$$P_j(\lambda) = P_j(0) \quad \text{and} \quad q_i = \Upsilon_i^{(0)}(\vec{P}(0))\lambda + q_i(0), \tag{2.23a-b}$$

where $\vec{P}(0)$ and $q_i(0)$ are the initial values at $\lambda = 0$. Using the expressions derived in Ref. [174] and then simplified in Ref. [175], one can express the radial and polar coordinates of a test particle under geodesic motion via

$$r(q_r) = \frac{r_3(r_1 - r_2)\mathrm{sn}^2\left(\frac{K(k_r)}{\pi}q_r|k_r\right) - r_2(r_1 - r_3)}{(r_1 - r_2)\mathrm{sn}^2\left(\frac{K(k_r)}{\pi}q_r|k_r\right) - (r_1 - r_3)}, \tag{2.24a}$$

$$z(q_z) = z_-\mathrm{sn}\left(K(k_z)\frac{2(q_z + \frac{\pi}{2})}{\pi}|k_z\right), \tag{2.24b}$$

where sn is Jacobi elliptic sine function given by $\mathrm{sn}(u|m) = \sin(\mathrm{am}(u|m))$ and the Jacobi amplitude am is the inverse function of the incomplete elliptic integral of the first kind F i.e. if $u = F(\phi|m)$, then $\phi = \mathrm{am}(u|m)$. Furthermore, we have adopted the convention that $q_r = 0$ corresponds to the periapsis of the radial motion (i.e. $r = r_2$), and $q_z = 0$ corresponds to the maximum value of the polar coordinate z (i.e. $z = z_-$).

2.1.3 Analytic solutions for Boyer-Lindquist time and azimuthal coordinates

One can also find analytic solutions for the Boyer-Lindquist time and azimuthal coordinates, however these expressions are a bit more involved [174]. First, we need to define the locations of the inner horizon r_- and the outer horizon r_+ of the Kerr black hole which are given by

$$r_\pm = M \pm \sqrt{M - a^2}. \tag{2.25}$$

This allows us to then define the following quantities

$$h_r := \frac{r_1 - r_2}{r_1 - r_3} \quad \text{and} \quad h_\pm := h_r\frac{r_3 - r_\pm}{r_2 - r_\pm}. \tag{2.26a-b}$$

We then also define the quantities ψ_r and ψ_z via

$$\psi_r := \text{am}\left(\frac{q_r K(k_r)}{\pi}\middle| k_r\right) \quad \text{and} \quad \psi_z := \text{am}\left(\frac{2\left(q_z + \frac{\pi}{2}\right)K(k_z)}{\pi}\middle| k_z\right). \tag{2.27a-b}$$

We note that the solutions to the geodesic equations (2.7c) and (2.7d) have the form

$$t(\lambda) = \Upsilon_t^{(0)}\lambda + t_r(q_r) + t_z(q_z), \tag{2.28a}$$

$$\phi(\lambda) = \Upsilon_\phi^{(0)}\lambda + \phi_r(q_r) + \phi_z(q_z), \tag{2.28b}$$

where $\Upsilon_k^{(0)}$ are the Mino-time fundamental frequencies associated with the time and azimuthal coordinates and t_r/ϕ_r and t_z/ϕ_z are oscillatory functions that depend solely on the radial and polar phases respectively. These expressions are given in terms of elliptic integrals, including the incomplete elliptic integral of the second kind

$$E(\phi|m) := \int_0^\phi \sqrt{1 - m\sin^2\theta}d\theta, \tag{2.29}$$

the complete elliptic integral of second kind $E(m) = E(\pi/2|m)$, the incomplete elliptic integral of the third kind

$$\Pi(n;\phi|m) \coloneqq \int_0^{\phi} \frac{d\theta}{\left(1 - n\sin^2\theta\right)\sqrt{1 - m\sin^2\theta}}, \tag{2.30}$$

and the complete elliptic integral of the third kind $\Pi(n|m) = \Pi(n;\pi/2|m)$. The radial oscillating terms are given by

$$\begin{aligned}
t_r(q_r) = &-\frac{\mathcal{E}}{\sqrt{\left(1-\mathcal{E}^2\right)\left(r_1-r_3\right)\left(r_2-r_4\right)}}\Bigg[4\left(r_2-r_3\right)\left(\frac{q_r}{\pi}\Pi\left(h_r|k_r\right) - \Pi\left(h_r;\psi_r|k_r\right)\right) \\
&- \frac{4\left(r_2-r_3\right)}{r_+ - r_-}\left(\frac{\left(r_+\left(4 - a\mathcal{L}/\mathcal{E}\right) - 2a^2\right)\left(\frac{q_r}{\pi}\Pi\left(h_+|k_r\right) - \Pi\left(h_+;\psi_r|k_r\right)\right)}{\left(r_2-r_+\right)\left(r_3-r_+\right)} - (+\leftrightarrow -)\right) \\
&+ \left(r_2-r_3\right)\left(r_1+r_2+r_3+r_4\right)\left(\frac{q_r}{\pi}\Pi\left(h_r|k_r\right) - \Pi\left(h_r;\psi_r|k_r\right)\right) \\
&+ \left(r_1-r_3\right)\left(r_2-r_4\right)\left(\frac{h_r\sin\psi_r\cos\psi_r\sqrt{1-k_r\sin^2\left(\psi_r\right)}}{1-h_r\sin^2\psi_r} + \frac{q_r}{\pi}E\left(k_r\right) - E\left(\psi_r|k_r\right)\right)\Bigg],
\end{aligned} \tag{2.31}$$

$$\begin{aligned}
\phi_r(q_r) = &\frac{2a\mathcal{E}}{\left(r_+-r_-\right)\sqrt{\left(1-\mathcal{E}^2\right)\left(r_1-r_3\right)\left(r_2-r_4\right)}} \\
&\times\left[\frac{\left(r_2-r_3\right)\left(2r_+ - a\mathcal{L}/\mathcal{E}\right)\left(\frac{q_r}{\pi}\Pi\left(h_+|k_r\right) - \Pi\left(h_+;\psi_r|k_r\right)\right)}{\left(r_2-r_+\right)\left(r_3-r_+\right)} - (+\leftrightarrow -)\right],
\end{aligned} \tag{2.32}$$

where we have adopted the shorthand $(+\leftrightarrow -)$ to indicate that one should repeat the previous term but with all positive symbols replaced with negative symbols (e.g., $r_+ \leftrightarrow r_-$ etc.). The oscillating polar terms are given by

$$t_z(q_z) = \frac{\mathcal{E}z_+}{1-\mathcal{E}^2}\left(\frac{2}{\pi}\left(q_z + \frac{\pi}{2}\right)E\left(k_z\right) - E\left(\psi_z|k_z\right)\right), \tag{2.33a}$$

$$\phi_z(q_z) = -\frac{\mathcal{L}}{z_+}\left(\frac{2}{\pi}\left(q_z + \frac{\pi}{2}\right)\Pi\left(z_-^2|k_z\right) - \Pi\left(z_-^2;\psi_z|k_z\right)\right). \tag{2.33b}$$

While the frequencies $\Upsilon_k^{(0)}$ have no dependence on the oscillating phases, they can be split up into contributions from the radial motion and from the polar motion as

$$\Upsilon_t^{(0)} = \Upsilon_{t,r}^{(0)} + \Upsilon_{t,z}^{(0)}, \tag{2.34}$$

$$\Upsilon_\phi^{(0)} = \Upsilon_{\phi,r}^{(0)} + \Upsilon_{\phi,z}^{(0)}. \tag{2.35}$$

$$\tag{2.36}$$

These have been found explicitly as

$$\begin{aligned}\Upsilon^{(0)}_{t,r} = &\frac{2\Upsilon^{(0)}_r}{\pi\sqrt{(1-\mathcal{E}^2)(r_1-r_3)(r_2-r_4)}}\Bigg[2\mathcal{E}M\left(r_3K\left(k_r\right)+\left(r_2-r_3\right)\Pi\left(h_r|k_r\right)\right)\\ &+\frac{\mathcal{E}}{2}\big\{\left(r_2-r_3\right)\left(r_1+r_2+r_3+r_4\right)\Pi\left(h_r|k_r\right)\\ &+\left(r_3\left(r_1+r_2+r_3\right)-r_1r_2\right)K\left(k_r\right)+\left(r_1-r_3\right)\left(r_2-r_4\right)E\left(k_r\right)\big\}\\ &+\frac{2M}{r_+-r_-}\left(\left(\frac{r_+\left(4\mathcal{E}M^2-a\mathcal{L}\right)-2a^2\mathcal{E}M}{r_3-r_+}\right)\left(K\left(k_r\right)-\frac{r_2-r_3}{r_2-r_+}\Pi\left(h_+|k_r\right)\right)-\left(+\leftrightarrow-\right)\right)\Bigg],\end{aligned} \tag{2.37}$$

$$\Upsilon^{(0)}_{t,z} = 4\mathcal{E}M^2+\frac{2\mathcal{E}\left(\beta^2+\mathcal{L}^2\right)\Upsilon^{(0)}_z\left(K\left(k_z\right)-E\left(k_z\right)\right)}{\pi(1-\mathcal{E}^2)x\sqrt{\beta x^2+\mathcal{L}^2}}, \tag{2.38}$$

$$\begin{aligned}\Upsilon^{(0)}_{\phi,r} = &\frac{2a\Upsilon^{(0)}_r}{\pi\left(r_+-r_-\right)\sqrt{(1-\mathcal{E}^2)\left(r_1-r_3\right)\left(r_2-r_4\right)}}\\ &\times\left[\frac{2\mathcal{E}Mr_+-a\mathcal{L}}{r_3-r_+}\left(K\left(k_r\right)-\frac{r_2-r_3}{r_2-r_+}\Pi\left(h_+|k_r\right)\right)-\left(+\leftrightarrow-\right)\right],\end{aligned} \tag{2.39}$$

$$\Upsilon^{(0)}_{\phi,z} = \frac{2\Upsilon^{(0)}_z}{\pi\sqrt{\frac{\beta}{\mathcal{L}^2}+\frac{1}{x^2}}}\Pi\left(z_-^2|k_z\right). \tag{2.40}$$

2.1.4 Fundamental frequencies in other time parametrizations

We note that the Mino-time fundamental frequencies $\{\Upsilon^{(0)}_t,\Upsilon^{(0)}_r,\Upsilon^{(0)}_z,\Upsilon^{(0)}_\phi\}$, are not the frequencies of the geodesic motion as observed by either an observer at infinity nor an observer co-moving with the secondary. However, both of these can be derived from the Mino-time frequencies. The frequencies observed at infinity are the Boyer-Lindquist fundamental frequencies and are given by

$$\Omega^{(0)}_r=\frac{\Upsilon^{(0)}_r}{\Upsilon^{(0)}_t},\quad \Omega^{(0)}_z=\frac{\Upsilon^{(0)}_z}{\Upsilon^{(0)}_t},\quad \Omega^{(0)}_\phi=\frac{\Upsilon^{(0)}_\phi}{\Upsilon^{(0)}_t}. \tag{2.41a-c}$$

One can also derive frequencies with respect to the secondary's proper time τ, but one first needs to first define the linear increase of proper time with Mino time $\Upsilon^{(0)}_\tau$ [175]. Much like the linear increase of Boyer-Lindquist time with Mino time Υ_t, this "frequency" has both a radial and a polar contribution

$$\Upsilon^{(0)}_\tau=\Upsilon^{(0)}_{\tau,r}+\Upsilon^{(0)}_{\tau,z}, \tag{2.42}$$

which are given by

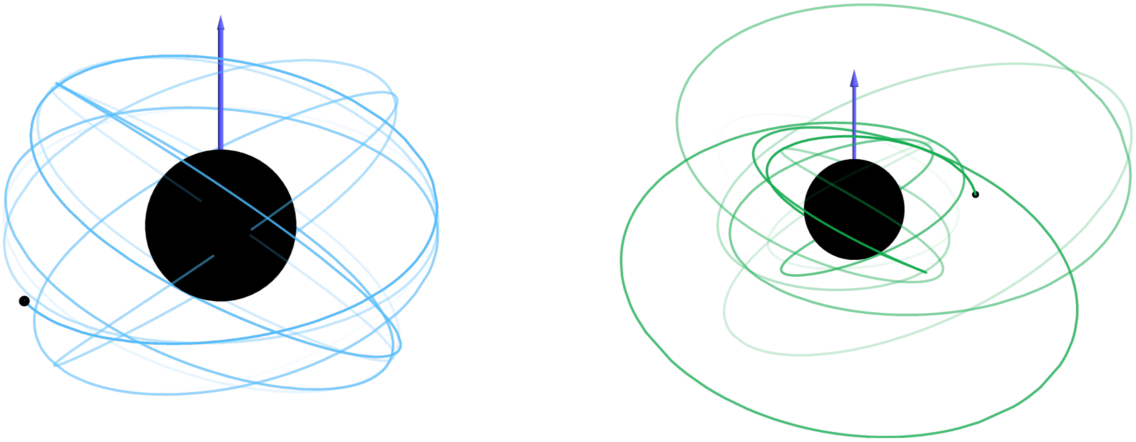
$$\begin{aligned}\Upsilon^{(0)}_{\tau,r} = &\frac{1}{2K\left(k_r\right)}\left[\left(r_2-r_3\right)\left(r_1+r_2+r_3+r_4\right)\Pi\left(\text{hr}\,|k_r\right)+\left(r_1-r_3\right)\left(r_2-r_4\right)E\left(k_r\right)\right]\\ &+\frac{r_1\left(r_3-r_2\right)+r_3\left(r_2+r_3\right)}{2}\end{aligned} \tag{2.43}$$

$$\Upsilon^{(0)}_{\tau,z} = \frac{z_p^2}{1-\mathcal{E}^2}\left(1-\frac{E\left(k_z\right)}{K\left(k_z\right)}\right). \tag{2.44}$$

(2.45)

(a) Circular, equatorial

(b) Eccentric, equatorial



(c) Spherical

(d) Generic

Figure 2.2: A collection of Kerr geodesic orbits, each with $a = 0.9$ and $p = 4$. The spherical and circular orbits have $e = 0$, while the eccentric and generic orbits have $e = 0.5$. The circular and eccentric equatorial orbits have $x = 1$, while the spherical and generic orbits have $x = 0.75$. The spherical and generic orbits are viewed from the front, but the equatorial orbits are viewed from above at an angle of 40° with respect to the equatorial plane.

With this in hand, one can straightforwardly derive the proper-time fundamental frequencies:

$$\omega_r^{(0)} = \frac{\Upsilon_r^{(0)}}{\Upsilon_\tau^{(0)}}, \quad \omega_z^{(0)} = \frac{\Upsilon_z^{(0)}}{\Upsilon_\tau^{(0)}}, \quad \omega_\phi^{(0)} = \frac{\Upsilon_\phi^{(0)}}{\Upsilon_\tau^{(0)}}. \tag{2.46a-c}$$

2.2 Classification of bound geodesic orbits

We will now briefly discuss the different cases of bound orbits around a Kerr black hole, that we make reference to throughout this work as illustrated in Fig. 2.2.

2.2.1 Circular and equatorial

The simplest orbits are circular orbits that lie in the equatorial plane and have $e = 0$ and $x = \pm 1$, as demonstrated in Fig. 2.2a. These orbits come in two flavours. Prograde orbits are where the secondary orbits in the same direction as the spin of primary and have $x = 1$. Retrograde are where the secondary orbits in the opposite direction as the spin of primary and have $x = -1$. Circular equatorial orbits also exist in the Schwarzschild limit when $a \to 0$, but since there is no spin on the primary, they cannot be classified as either prograde or retrograde.

The primary difference in each case is the location of the inner-most stable circular orbit (ISCO). For Schwarzschild orbits, the ISCO is located at $r_{\text{ISCO}} = 6M$. For prograde orbits around a fast spinning black hole, the ISCO asymptotically approaches the outer horizon of the primary i.e. $r_{\text{ISCO}} \to r_+ \to M$ as $a \to M$. For retrograde orbits, the ISCO approaches $9M$ as $a \to M$ [176].

We note due to the cylindrical symmetry of Kerr spacetime, the equations of motion have no dependence on the azimuthal phase ϕ. This becomes particularly useful when modelling inspirals as a continuous evolution through (quasi-)circular orbits. The slow secular evolution of the orbital radius is naturally decoupled from the rapidly oscillating azimuthal phase, and so equations of motion for quasi-circular inspirals can be rapidly solved numerically without the need to employ any sort of averaging scheme. As such, this work will not focus on these orbits, and will focus instead to the three remaining cases.

2.2.2 Eccentric and equatorial

Our first major focus of this work will be on eccentric and equatorial orbits with where $e \neq 0$ and $x = \pm 1$, like the one illustrated in Fig. 2.2b. As with the circular, equatorial orbits discussed above, eccentric, equatorial orbits can be classified as either prograde $(x = 1)$, retrograde $(x = -1)$, or neither in the Schwarzschild limit $(a = 0, x = \pm 1)$, which will effect the location of the last stable orbit (LSO). In Schwarzschild, the LSO is given by the simple formula $p_{\text{LSO}} = 6 + 2e$, but one must solve for the roots of a non-trivial quartic polynomial, which is given Ref. [177].

We note that the radius is no longer constant and evolves between a maximum value of r_1 and a minimum of r_2. We parametrize this evolution using a radial phase. Common choices for this phase are the quasi-Keplerian angle ψ as defined in Eq. (2.17) or the action angle q_r as defined in Eq. 2.19.

Unlike eccentric orbits in Newtonian gravity, eccentric orbits around a Kerr black hole are not simple ellipses. Instead the location of the periapsis will precess around the primary as there is a difference between the radial and azimuthal frequencies. For highly eccentric orbits close to the primary, there exist a class of "zoom-whirl" orbits just like the one Fig. 2.2b, where the secondary will "zoom" close to the primary near the innermost bound circular orbit (IBCO) where $\mathcal{E} = 1$, and complete one or more rapid rotations before "whirling" out to large orbital separations and starting the orbital cycle over again [178].

If one attempts to model an inspiral as an evolving series of eccentric orbits, the evolution equations for the semilatus rectum p and the eccentricity e will depend on the rapidly oscillating radial phase, unless one employs some sort of averaging scheme. Deriving and implementing such a scheme is the focus of Chapter 6.

2.2.3 Spherical (circular and inclined)

The second class of orbits that this work will focus on are circular and inclined orbits with where $e = 0$ but $-1 \leq x \leq 1$, such as the example presented in Fig. 2.2b. These are often referred to as "spherical" orbits, as highly inclined orbits almost fill out a sphere of constant Boyer-Lindquist radius centred on the primary.

Inclined orbits in Schwarzschild spacetime will stay in their plane of inclination due to spherical symmetry and so one can always rotate the coordinate system such that orbit lies in the equatorial plane. As such, spherical orbits are a phenomenon exclusive to Kerr spacetime.

The LSO for an inclined orbit in Schwarzschild is the same as in the circular case with $p_{\text{LSO}} = r_{\text{ISCO}}/M = 6$. However, for Kerr orbits the LSO is a root of 12^{th} degree polynomial in a and p and a 4^{th} polynomial in x^2 [169]. In the equatorial limit, the LSO recovers the results for the circular case. For extremal black holes with $a = M$, the LSO approaches $p_{\text{LSO}} \to 1$ as $x \to 1$, $p_{\text{LSO}} \to 9$ as $x \to -1$ and $p_{\text{LSO}} \to 1 + \sqrt{3} + \sqrt{3 + 2\sqrt{3}} \approxeq 5.275$ as $x \to 0$.

We note that the polar coordinate z is no longer constant and evolves between a maximum value of z_- to a minimum value of $-z_-$. This polar motion is parametrized using a polar phase with common choices being the quasi-Keplerian angle χ as defined in Eq. (2.17) or the action angle q_z as defined in Eq. 2.19.

If one attempts to model an inspiral as an evolving series of spherical orbits, the evolution equations for the semilatus rectum p and the inclination x will depend on the rapidly oscillating polar phase, unless one employs some sort of averaging scheme. Deriving and implementing such a scheme is the focus of Chapter. 7.

2.2.4 Generic (eccentric and inclined)

The final class of orbits we will focus on are eccentric and inclined orbits, or "generic" orbits where $0 < e < 1$ and $-1 < x < 1$. As one can see from Fig. 2.2d, these orbits have the richest structure out of all the possible orbits in Kerr spacetime. These orbits simultaneously have the radial motion of an eccentric orbit and the polar motion of a spherical orbit, and so we parametrize this motion using both a radial and a polar phase. Just like spherical orbits, generic orbits are a phenomenon exclusive to Kerr spacetime.

Similar to the spherical case, the LSO for generic Kerr orbits is a root of 12^{th} degree polynomial in a, p and e and a 4^{th} polynomial in x^2. The orbits can also exhibit the "zoom-whirl" behaviour, with the inner radius bounded by the innermost bound spherical orbit (IBSO) which is the smallest spherical orbit where $\mathcal{E} = 1$.

If one attempts to model an inspiral as an evolving series of generic orbits, the evolution equations for p, e and x will depend on the rapidly oscillating radial and polar phases, and we focus on deriving and implementing an average scheme for these orbits Chapter 8. However, there are a subset of generic orbits which pose an issue for any averaging procedure, which we will now discuss in more depth.

2.3 Resonant Orbits

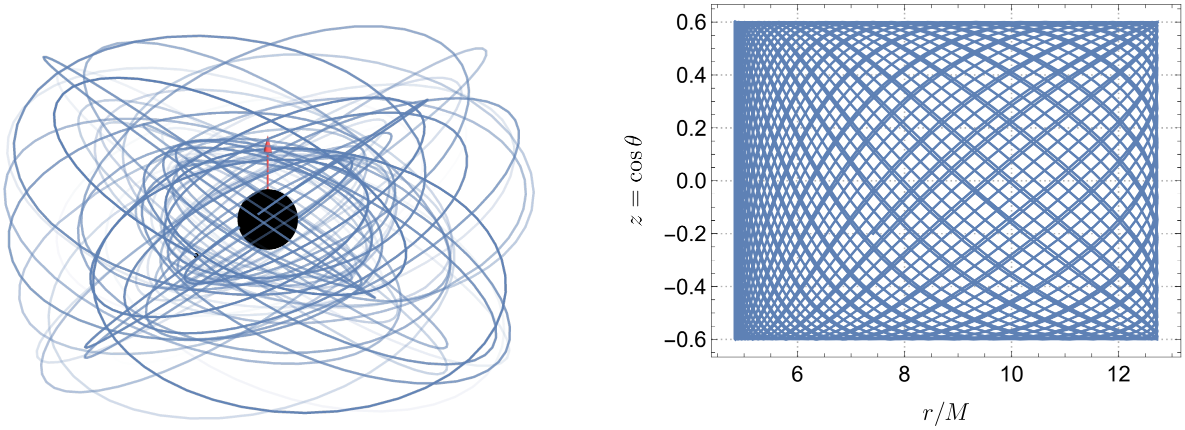
A resonant orbit occurs whenever the radial phase is related to the polar phase by a small number integer ratio i.e., $\vec{\kappa}_{\text{res}} \cdot \vec{\Upsilon}^{(0)} = \kappa_r \Upsilon_r^{(0)} + \kappa_z \Upsilon_z^{(0)} = 0$ for $\kappa_r, \kappa_z \in \mathbb{Z}$. We denote a specific orbital resonance using the fraction $\Upsilon_z^{(0)}/\Upsilon_r^{(0)} = |\kappa_z|/|\kappa_r|$.

As illustrated in Figs. 2.3a and 2.3b, if a generic orbit is allowed to evolve for infinitely many orbits, it will eventually fill the entirety of the (r, z) space bounded between r_1 and r_2, and z_- and $-z_-$. Such an orbit is said to be ergodic in the phase space. This allows us to equate the infinite Mino-time average for a geodesic with an integral over the 2-torus of the action angles q_r and q_z, i.e.,

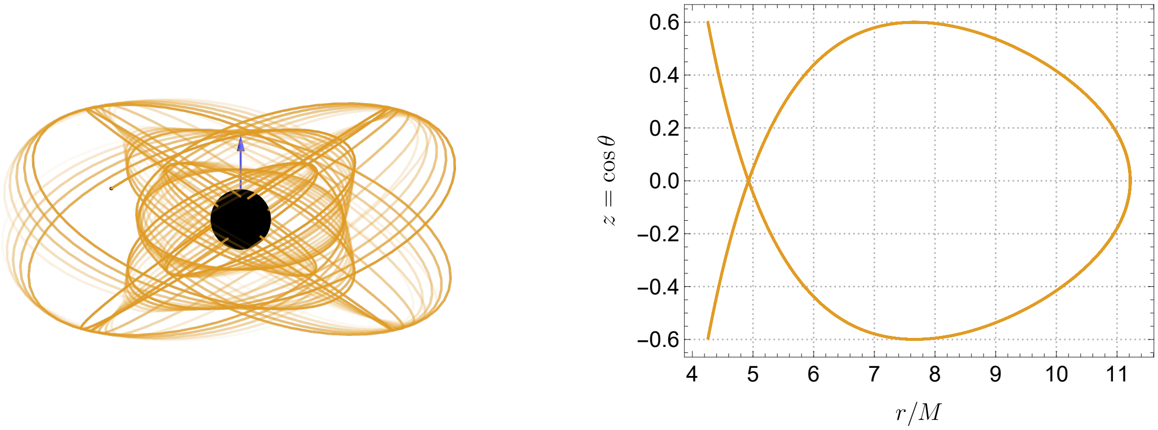
$$\langle A \rangle = \lim_{\lambda \to \infty} \frac{1}{2\lambda} \int_{-\lambda}^{\lambda} A(\lambda') d\lambda' = \frac{1}{(2\pi)^2} \int_0^{2\pi} \int_0^{2\pi} A(q_r, q_z) dq_r dq_z = A_{0,0}, \tag{2.47}$$

where $A_{0,0}$ is the zeroth Fourier coefficient.

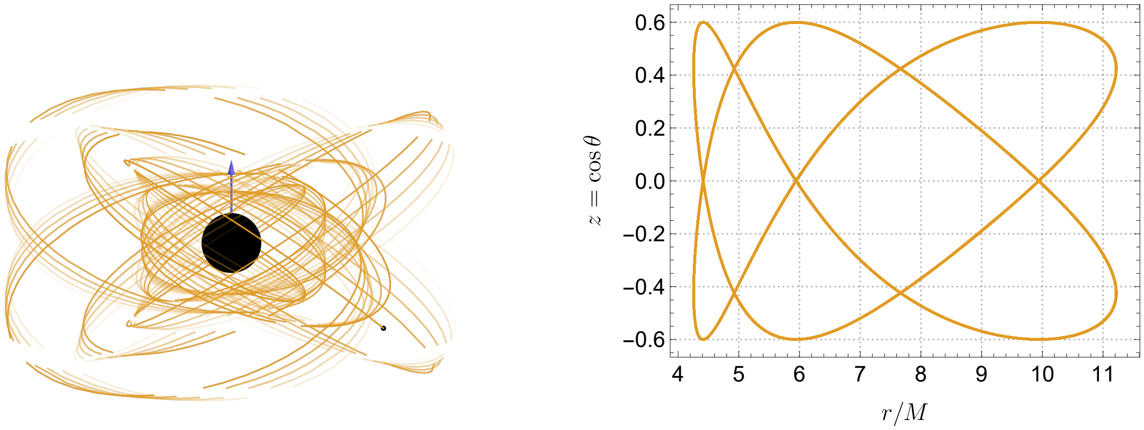
However, as seen in Figs. 2.3d and 2.3f, a resonant orbit does not fill the r, z space and instead repeatedly traces out the same trajectory in this space. Moreover, the dissimilarity between these figures demonstrates that the phase space trajectory is affected by the initial conditions for the phases i.e. $q_{r,0}$



(a) Generic orbit (b) Generic phase space



(c) Resonant orbit: $q_{r,0} = 0, q_{z,0} = 0$ (d) Resonant phase space: $q_{r,0} = 0, q_{z,0} = 0$



(e) Resonant orbit: $q_{r,0} = 0, q_{z,0} = \pi/4$ (f) Resonant phase space: $q_{r,0} = 0, q_{z,0} = \pi/4$

Figure 2.3: Parametric plots of r vs z for a generic orbit and resonant orbits which all have $a = 0.9M$, $e = 0.45$ and $x = 0.8$, with initial phases $q_{r,0} = q_{z,0} = 0$. The non-resonant orbit has $p = 7$ and the resonant orbits has $p = 6.171$ which is the location of the $2/3$ orbital resonance. The two resonant orbits differ only by their initial phases.

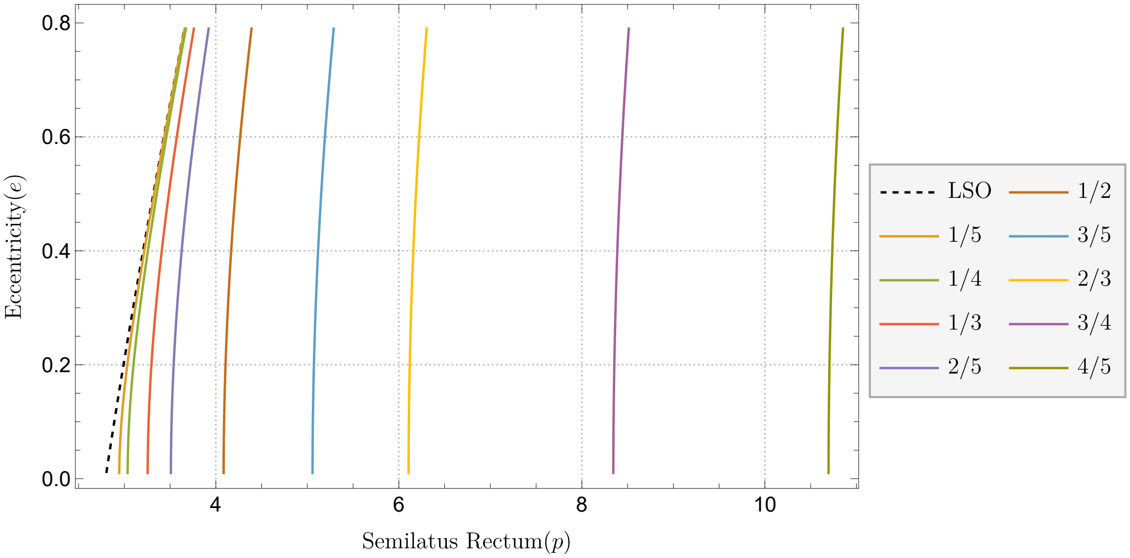


Figure 2.4: The locations of the resonant surfaces through (p, e) space for $a = 0.9M$ and $x = 0.8$. We also give the location of the last stable orbit (LSO) as a dashed black curve.

and $q_{z,0}$. Thus, one cannot equate the infinite Mino-time average for a resonant geodesic with the 2-torus average of the action angles and instead one gets [113]

$$\langle A \rangle_{\text{res}} = \lim_{\lambda \to \infty} \frac{1}{2\lambda} \int_{-\lambda}^{\lambda} A(\lambda')d\lambda' = \sum_{N \in \mathbb{Z}} A_{N\kappa_r, N\kappa_z} e^{iN(\kappa_r q_{r,0} + \kappa_z q_{z,0})}. \tag{2.48}$$

As such, any averaging procedure done in the presence of an orbital resonance will have to account for this new definition of orbit average.

Thankfully, sustained orbital resonances are very unlikely for EMRIs [157]. Moreover, the strength of these resonances is inversely proportional to the order of the resonance i.e. κ_r. Thus even though integers are dense in the real numbers, meaning there are an infinite number of potential orbital resonances, one only has to worry about resonances where $\kappa_r \lesssim 5$.

However, it is quite likely that an inspiral will pass through at least one of low order resonance while in the LISA band [160]. This is due to the resonances forming a 3D hyper-surface in the 4D generic Kerr orbit parameter space [179]. To get a flavour for where these resonances occur in the parameter space, Fig. 2.4 illustrates the location in (p, e) space in a 2D slice of the parameter space where we fixed $a = 0.9M$ and $x = 0.8$. As we can see, most low order resonances occur near the location of the last stable orbit (LSO). Since our inspiral models start to break down in this region anyway, these resonances are not the biggest concern. What is more concerning are the resonances that occur at a significant distance from the last stable orbit, as the effect of inaccurately modelling these resonances can accumulate over a large number of orbits. As such, the lowest order resonance of concern that we would expect most EMRIs to pass through is the $2/3$ resonance. Any averaging procedure employed to efficiently model EMRI trajectories will have to carefully account for the presence of these resonances in order to maintain subradian accuracy in the orbital phases.

Chapter 3

Osculating Geodesics

3.1 The method of osculating geodesics

We now wish to describe the forced motion of a body obeying Eq. (2.3). Writing this in terms of the Christoffel symbols and the worldline of the secondary gives us:

$$\frac{d^2x}{d\tau^2} + \Gamma^\alpha_{\beta\gamma}\frac{dx^\beta}{d\tau}\frac{dx^\gamma}{d\tau} = a^\alpha. \tag{3.1}$$

We first identify a set of orbital elements that uniquely identify a given geodesic, such as the integrals of motion $\vec{P}$ along with the initial values of the orbital phases of the geodesic orbit $\vec{q}_0$ and designate them as a set of "orbital elements" $\vec{I} = \{\vec{P}, \vec{q}_0\}$. For accelerated orbits, these orbital elements are promoted from constants to functions of Mino time. Note that now the orbital elements $\vec{q}_0(\lambda)$ are different quantities from the values of $\vec{q}$ evaluated at $\lambda = 0$, i.e., $\vec{q}(0)$. Our goal is to transform the forced geodesic equation into evolution equations for a set of orbital elements I^A. While there is choice in the selection of orbital elements, we require that the number of orbital elements is equal to the degrees of freedom on the orbit. For instance, eccentric and spherical orbits only have 3 degrees of freedom and so one only needs evolution equations for $(p, e, q_{r,0})$ or $(p, x, q_{z,0})$ respectively, while generic orbits have 5 and so one needs evolution equations for $(p, e, x, q_{r,0}, q_{z,0})$.

We assume that the worldline and four-velocity of the secondary at each instant can be described as the worldline and four-velocity of a test body on a tangent geodesic,

$$x^\alpha(\tau) = x^\alpha_G(I^A(\tau), \tau) \tag{3.2a}$$

$$\frac{dx^\alpha}{d\tau}(\tau) = \frac{\partial x^\alpha_G}{\partial\tau}(I^A(\tau), \tau), \tag{3.2b}$$

where the partial derivative in the second equation implies that I^A is held fixed. These are known as the "osculating conditions". If we were to take the derivative of Eq. (3.2a) with respect to τ, one would obtain:

$$\frac{dx^\alpha}{d\tau} = \frac{dx^\alpha_G}{d\tau} = \frac{\partial x^\alpha_G}{\partial\tau} + \frac{\partial x^\alpha_G}{\partial I^A}\frac{dI^A}{d\tau} \tag{3.3}$$

Comparing this with Eq. (3.2b) gives us the first osculating geodesic equation:

$$\frac{\partial x^\alpha_G}{\partial I^A}\frac{dI^A}{d\tau} = 0. \tag{3.4}$$

Next we note that x^α_G must satisfy the geodesic equation:

$$\frac{d^2x^\alpha_G}{d\tau^2} + \Gamma^\alpha_{\beta\gamma}\frac{dx^\beta_G}{d\tau}\frac{dx^\gamma_G}{d\tau} = 0. \tag{3.5}$$

Thus if we subtract this from Eq. (3.1), we obtain

$$\frac{d^2x^\alpha}{d\tau^2} + \Gamma^\alpha_{\beta\gamma}\frac{dx^\beta}{d\tau}\frac{dx^\gamma}{d\tau} - \frac{d^2x^\alpha_G}{d\tau^2} - \Gamma^\alpha_{\beta\gamma}\frac{dx^\beta_G}{d\tau}\frac{dx^\gamma_G}{d\tau} = a^\alpha. \tag{3.6}$$

Using Eq. (3.2b), we can get rid of the Christoffel terms and the equation simplifies to

$$\frac{d^2x^\alpha}{d\tau^2} = \frac{\partial^2 x^\alpha_G}{\partial\tau^2} + a^\alpha. \tag{3.7}$$

If we now compare this to the τ derivative of Eq. (3.2b):

$$\frac{d^2x^\alpha}{d\tau^2} = \frac{d}{d\tau}\left(\frac{\partial x^\alpha_G}{\partial\tau}\right) = \frac{\partial^2 x^\alpha_G}{\partial\tau^2} + \frac{\partial}{\partial I^A}\left(\frac{\partial x^\alpha_G}{\partial\tau}\right)\frac{dI^A}{d\tau} \tag{3.8}$$

From this, we can infer the second osculating equation

$$\frac{\partial u^\alpha_G}{\partial I^A}\frac{dI^A}{d\tau} = a^\alpha \tag{3.9}$$

where the geodesic four-velocity is given by $\partial x^\alpha_G/\partial\tau$.

We prefer to work in terms of (Carter-)Mino time λ as opposed to propertime τ. Using the chain rule with Eq. (2.8) and using the Kerr metric to lower the indices on the second osculating equation (3.9), we can recast the osculating geodesic equations as

$$\frac{\partial x^\alpha_G}{\partial I^A}\frac{dI^A}{d\lambda} = 0 \quad \text{and} \quad \frac{\partial u^G_\alpha}{\partial I^A}\frac{dI^A}{d\lambda} = \Sigma a_\alpha. \tag{3.10a-b}$$

From these, evolution equations for each of the orbital elements can be calculated. This was first demonstrated in Ref. [136] for Schwarzschild inspirals, and has been utilized many times for evolving orbits in Schwarzschild [70, 145, 154, 156]. This was then extended to generic Kerr orbits in Ref. [137], which lays out four different formulations of the equations. Three of these formulations use $\vec{\mathcal{J}} = \{\mathcal{E}, \mathcal{L}, \mathcal{K}\}$ as the orbital elements and the quasi-Keplerian angles ψ and χ as the orbital phases, and two of these were numerically implemented and shown to agree with each other. The "null tetrad" formulation which we will briefly outline is the most numerically efficient of the different formulations and so is the primary method for evolving orbits in Kerr spacetime [68, 143, 160]. However, this formulation is not as suited to deriving averaged equations of motion due to the use of use of ψ and χ as orbital phases. We instead utilize a modified version of the action angle formulation for the first time where we use $\vec{P} = \{p, e, x\}$ as the orbital elements and the geodesic actions angels q_r and q_z as the orbital phases.

3.2 Evolution equations of the orbital elements

3.2.1 Generic case

We first find the evolution equations for the integrals of motion $\vec{P} = \{p, e, x\}$. In order to do so, we must consider how a different set of integrals of motion $\vec{\mathcal{J}} = \{\mathcal{E}, \mathcal{L}, \mathcal{K}\}$ evolve in terms of the covariant components of the particle's 4-acceleration $\{a_t, a_r, a_z, a_\phi\}$. Using the second osculating geodesic equation (3.10b) along with definitions of $\mathcal{E}$ and $\mathcal{L}$ in Eq. (2.5), we can obtain the evolution equations for $\mathcal{E}$ and $\mathcal{L}$:

$$\frac{d\mathcal{E}}{d\lambda} = -\Sigma a_t, \quad \text{and} \quad \frac{d\mathcal{L}}{d\lambda} = \Sigma a_\phi. \tag{3.11a-b}$$

To find the the evolution of $\mathcal{K}$, we note that the contravariant components of the Killing tensor can be written as [137]

$$\mathcal{K}^{\alpha\beta} = 2\Sigma l^{(\alpha}n^{\beta)} + r^2 g^{\alpha\beta}, \tag{3.12}$$

where l and n are null vectors with components

$$l = \frac{\varpi^2}{\Delta}\partial_t + \partial_r + \frac{a}{\Delta}\partial_\phi \quad \text{and} \quad n = \frac{\varpi^2}{2\Sigma}\partial_t - \frac{\Delta}{2\Sigma}\partial_r + \frac{a}{2\Sigma}\partial_\phi. \tag{3.13a-b}$$

Taking the derivative of $\mathcal{K}$ from Eq. (2.5c) with respect to proper time gives us

$$\frac{d\mathcal{K}}{d\tau} = \mathcal{K}^{\alpha\beta}u_\alpha a_\beta. \tag{3.14}$$

Expanding this out explicitly while making use of the orthogonality condition, $g^{\alpha\beta}u_\alpha a_\beta = 0$ and converting to Mino time gives us:

$$\frac{d\mathcal{K}}{d\lambda} = \frac{d\mathcal{E}}{d\lambda}\frac{2}{\Delta}\left(\varpi^4\mathcal{E} - a\varpi^2\mathcal{L}\right) + \frac{d\mathcal{L}}{d\lambda}\frac{2}{\Delta}\left(a^2\mathcal{L} - a\varpi^2\mathcal{E}\right) - 2\Sigma\Delta u_r a_r. \tag{3.15}$$

Using the above equations, we can now express how the roots $\{r_1, r_2, z_-\}$ evolve with Mino time by exploiting the same trick as in Appendix A.2 and A.3 of Ref. [137]. First, we note that, using the chain rule, we can express the rate of change of r_1 or r_2 as

$$\frac{dr_{1,2}}{d\lambda} = \frac{\partial r_{1,2}}{\partial \mathcal{J}_j}\frac{d\mathcal{J}_j}{d\lambda}. \tag{3.16}$$

We then find expressions for $\partial r_{1,2}/\partial\mathcal{J}_j$ be differentiating $V_r(r)$ with respect to $\mathcal{J}_j$.

$$\frac{\partial V_r}{\partial \mathcal{J}_j}\Big|_{r=r_1} = (1-\mathcal{E}^2)(r_1-r_2)(r_1-r_3)(r_1-r_4)\frac{\partial r_1}{\partial \mathcal{J}_j}, \tag{3.17a}$$

$$\frac{\partial V_r}{\partial \mathcal{J}_j}\Big|_{r=r_2} = -(1-\mathcal{E}^2)(r_1-r_2)(r_2-r_3)(r_2-r_4)\frac{\partial r_2}{\partial \mathcal{J}_j}. \tag{3.17b}$$

We then note that the coefficients proceeding $\partial r_{1,2}/\partial\mathcal{J}_j$ are also obtained by differentiating V_r with respect to r and then evaluating at $r_{1,2}$, i.e.,

$$\frac{\partial V_r}{\partial \mathcal{J}_j}\Big|_{r_{1,2}} = -\kappa(r_{1,2})\frac{\partial r_{1,2}}{\partial \mathcal{J}_j}, \tag{3.18}$$

where we have defined

$$\kappa(r) := \frac{dV_r}{dr} = 4\mathcal{E}F(r)r - 2r\Delta(r) - 2(r-M)(r^2+\mathcal{K}), \tag{3.19}$$

$$F(r) := \varpi(r)^2\mathcal{E} - a\mathcal{L}. \tag{3.20}$$

Combining equations (3.16) and (3.18) and using the appropriate definition of V_r from equation (2.7a) to calculate the partial derivatives gives us our evolution equations for r_1 and r_2:

$$\frac{dr_{1,2}}{d\lambda} = -\frac{2F(r_{1,2})}{\kappa(r_{1,2})}\left(\varpi(r_{1,2})^2\frac{d\mathcal{E}}{d\lambda} - a\frac{d\mathcal{L}}{d\lambda}\right) + \frac{\Delta(r_{1,2})}{\kappa(r_{1,2})}\frac{d\mathcal{K}}{d\lambda}. \tag{3.21}$$

We can use similar steps as above to find the evolution of z_-. Again, the chain rule tells us that the evolution of z_- follows

$$\frac{dz_-}{d\lambda} = \frac{\partial z_-}{\partial \mathcal{J}_j}\frac{d\mathcal{J}_j}{d\lambda}. \tag{3.22}$$

We then use $\frac{\partial V_z}{\partial \mathcal{J}_j}\Big|_{z_-}$ along with the second definition V_z in equation (2.7b) to find an expression for $\frac{\partial z_-}{\partial \mathcal{J}_j}$

$$\frac{\partial V_z}{\partial \mathcal{J}_j}\Big|_{z_-} = -2z_-(\beta z_-^2 - z_+^2)\frac{\partial z_-}{\partial \mathcal{J}_j}. \tag{3.23}$$

However, using the first definition of V_z in terms of $\{\mathcal{E}, \mathcal{L}, \mathcal{Q}\}$ gives us the following explicit expressions for

$$\frac{\partial V_z}{\partial \mathcal{E}}\Big|_{z_-} = 2a^2\mathcal{E}z_-^2(1-z_-^2), \quad \frac{\partial V_z}{\partial \mathcal{L}}\Big|_{z_-} = -2\mathcal{L}z_-^2, \quad \text{and} \quad \frac{\partial V_z}{\partial \mathcal{Q}}\Big|_{z_-} = 1 - z_-^2. \tag{3.24a-c}$$

Combining the results from equations (3.22), (3.23) and (3.24) gives us

$$\frac{2z_-(z_+ - \beta z_-)}{(1-z_-^2)}\frac{dz_-}{d\lambda} = \frac{d\mathcal{Q}}{d\lambda} - 2\mathcal{L}\left(\frac{z_-^2}{1-z_-^2}\right)\frac{d\mathcal{L}}{d\lambda} + 2a^2\mathcal{E}z_-^2\frac{d\mathcal{E}}{d\lambda}. \tag{3.25}$$

Since we have expressions for the evolution of $\{\mathcal{E}, \mathcal{L}, \mathcal{K}\}$, we can derive and expression for the evolution of $\mathcal{Q}$ by taking the derivative of equation (2.6) with respect to λ:

$$\frac{d\mathcal{Q}}{d\lambda} = \frac{d\mathcal{K}}{d\lambda} - 2(\mathcal{L} - a\mathcal{E})\left(\frac{d\mathcal{L}}{d\lambda} - a\frac{d\mathcal{E}}{d\lambda}\right). \tag{3.26}$$

Combing these two results and simplifying yields our final expression for the evolution of z_-

$$\frac{dz_-}{d\lambda} = \frac{1}{2z_-(z_+^2 - \beta z_-^2)}\left(x^2\frac{d\mathcal{K}}{d\lambda} - 2\left(\mathcal{L} - ax^2\mathcal{E}\right)\left(\frac{d\mathcal{L}}{d\lambda} - ax^2\frac{d\mathcal{E}}{d\lambda}\right)\right), \tag{3.27}$$

where we've used Eq. (2.12c) to tidy up the final expression.

Now that we know how $\{r_1, r_2, z_-\}$ evolve, determining the evolution of $\{p, e, x\}$ is straightforward since we can convert from one set to the other using the relations (2.9) and (2.10). We can then take the derivative of these equations with respect to λ and use the chain rule to obtain

$$\frac{dp}{d\lambda} = \frac{2}{M(r_1+r_2)^2}\left(r_2^2\frac{dr_1}{d\lambda} + r_1^2\frac{dr_2}{d\lambda}\right) := F_p, \tag{3.28a}$$

$$\frac{de}{d\lambda} = \frac{2}{(r_1+r_2)^2}\left(r_2\frac{dr_1}{d\lambda} - r_1\frac{dr_2}{d\lambda}\right) := F_e, \tag{3.28b}$$

$$\frac{dx}{d\lambda} = -\frac{z_-}{x}\frac{dz_-}{d\lambda} := F_x. \tag{3.28c}$$

3.2.2 Spherical case

The above expressions for equations for F_p and F_e become numerically singular in the spherical limit (i.e. when $e = 0$). These are not physical singularities, as one can derive alternative expressions which are regular in this limit by considering the evolution of the integrals of motion $\vec{\mathcal{J}}$ and then relating this to the evolution of p.

We begin by using the chain rule on the rate of change of r which gives us

$$\frac{dr}{d\lambda} = \frac{\partial r}{\partial \mathcal{J}_j}\frac{d\mathcal{J}_j}{d\lambda} \tag{3.29}$$

To find the partial derivatives $\partial r/\partial \mathcal{J}_j$, make use of the value of radial potential V_r defined as

$$\begin{aligned} V_r &= F^2 - \Delta\left(r^2 + \mathcal{K}\right) \\ &= -(1-\mathcal{E}^2)(r-r_1)(r-r_2)(r-r_3)(r-r_4) \\ &= -2(1-\mathcal{E}^2)(r-p)(r-r_3)(r-r_4). \end{aligned} \tag{3.30}$$

where r_1, r_2, r_3, r_4 are the roots of V_r and in the spherical case: $r_1 = r_2 = pM$. We then define the derivative of V_r with respect to r as

$$\begin{aligned} \kappa :=& \frac{dV_r}{dr} = 4\mathcal{E}Fr - 2\Delta r - 2(r-M)(r^2+K) \\ =& -2(1-\mathcal{E}^2)\Big[(r-r_3)(r-r_4) + (r-pM)(r-r_4) + (r-pM)(r-r_3)\Big]. \end{aligned} \tag{3.31}$$

If we take the derivative of κ with respect to r and evaluate at $r = pM$, we obtain

$$\frac{d\kappa}{dr}\bigg|_{r=pM} = -2(1-\mathcal{E}^2)(pM-r_3)(pM-r_4). \tag{3.32}$$

However, if we now take the partial derivatives of κ with respect to $\vec{\mathcal{J}}$ and evaluate at $r = pM$ we are left with

$$\frac{\partial\kappa}{\partial\mathcal{J}_j}\bigg|_{r=pM} = 2(1-\mathcal{E}^2)(pM-r_3)(pM-r_4)\frac{\partial r}{\partial\mathcal{J}_j} = \frac{d\kappa}{dr}\bigg|_{r=pM}\frac{\partial r}{\partial\mathcal{J}_j}. \tag{3.33}$$

Subbing the above result into Eq. (3.29), explicitly calculating $\partial\kappa/\partial\mathcal{J}_j$ using Eq. (3.31), evaluating at $r = pM$ and simplifying gives us our final result

$$\frac{dp}{d\lambda} = \frac{-4p}{d\kappa/dr}\left((F + \varpi^2\mathcal{E})\frac{dE}{d\lambda} - a\mathcal{E}\frac{d\mathcal{L}}{d\lambda} + \frac{M(1-p)}{2p}\frac{d\mathcal{K}}{d\lambda}\right) := F_p\Big|_{e=0} \tag{3.34}$$

By contrast, finding the evolution of the orbital eccentricity e in this limit is straightforward. It has been well established that under radiation reaction we expect quasi-circular inspirals with $e = 0$ to remain quasi-circular [180, 181], and so

$$\frac{de}{d\lambda}\bigg|_{e=0} = F_e\Big|_{e=0} = 0. \tag{3.35}$$

The evolution equation for the orbital inclination x is regular in the spherical limit. As such, we can continue to use the generic expression of F_x.

3.3 Evolution of the orbital phases

We will now discuss how to evolve the orbital phases. Since we make use of two different formulations of the osculating geodesic equations which differ by the choice of orbital phases, we will now outline the evolution equations for each.

3.3.1 Evolution of quasi-Keplerian angles

The most computational efficient formulations of the osculating geodesic equations for Kerr inspirals that have been implemented thus far employ the quasi-Keplerian angles ψ and χ. Evolution equations for these phases were derived in three different ways in Ref. [137]. By far the most convenient method is the null tetrad formulation where one first derives the equations of motion in terms of components of the force aligned with the Kinnersley null tetrad. One then relates the null tetrad components back to the Boyer-Lindquist components. The resulting evolution equations are finite at all points in the orbit, including the turning points. A full derivation of this formulation can be found in Ref. [137], but it is important to note that Ref. [137] makes use of different notation where $z = \cos^2\theta$ instead of $\cos\theta$, z_- is equivalent to z_-^2, and z_+ is equivalent to βz_+^2. We will only briefly summarize their findings here.

First define the following combinations of acceleration components [1]:

$$\mathcal{A}_I = a_\theta = \frac{-a_z}{\sqrt{1-z^2}} \tag{3.36a}$$

$$\mathcal{A}_{II} = -a\sin\theta a_t - \frac{a_\phi}{\sin\theta} \tag{3.36b}$$

$$\mathcal{A}_{III} = \frac{a\mathcal{H}}{\Sigma}a_t + \frac{u_z}{\Sigma}a_z + \frac{\mathcal{H}}{\Sigma\sin^2\theta}a_\phi, \tag{3.36c}$$

where $\mathcal{H} = \mathcal{L} - a\mathcal{E}\sin^2\theta$. We also find the component of the secondary's four-acceleration and four-velocity that is aligned with the null vector n can be written in terms of the Boyer-Lindquist components as

$$u_n = -\frac{F}{2\Sigma} - \frac{\Delta}{2\Sigma}u_r, \tag{3.37a}$$

$$a_n = \frac{\varpi^2}{2\Sigma}a_t - \frac{\Delta}{2\Sigma}a_r + \frac{a}{2\Sigma}. \tag{3.37b}$$

With these relations in hand, the evolution equation of the polar quasi-Keplerian angle χ is given by

$$\frac{d\chi}{d\lambda} = \hat{\Upsilon}_z^{(0)}\left(1 + \frac{(1-z_-^2)\Sigma\mathcal{A}_I\cos\chi}{z_-(z_+^2 - \beta z_-^2)\sin\theta)}\right) + \frac{\cos\chi\sin\chi\mathcal{H}a\Delta(\mathcal{A}_{III} - 2u_r a_n)}{2(z_+^2 - \beta z_-^2)u_n} + \frac{\cos\chi\sin\chi\mathcal{G}\mathcal{A}_{II}}{z_+^2 - \beta z_-^2}, \tag{3.38}$$

where $\mathcal{G} = \varpi^2\mathcal{L}\csc\theta - a^3x^2\mathcal{E}\sin\theta$. The evolution equation for the radial quasi-Keplerian angle ψ is given by

$$\begin{aligned}\frac{d\psi}{d\lambda} =& \hat{\Upsilon}_r^{(0)} + \frac{\mathcal{C}\mathcal{A}_{III}\sin\psi}{2(1+e\cos\psi)u_n} + \frac{\mathcal{D}\Sigma\mathcal{A}_{III}\hat{\Upsilon}_r^{(0)}}{2(1+e\cos\psi)^2u_n} - \frac{a\mathcal{B}\sin\theta\sin\psi\mathcal{A}_{II}}{1+e\cos\psi}\\ &+ \frac{\hat{\Upsilon}_r^{(0)}a_n}{u_n(1+e\cos\psi)^2}\left((1-e)^2(1-\cos\psi)\frac{\Sigma_1F_1}{\kappa_1} + (1+e)^2(1+\cos\psi)\frac{\Sigma_2F_2}{\kappa_2}\right),\end{aligned} \tag{3.39}$$

where

$$\mathcal{B} = \frac{F_1(1-e)(r+r_1)}{\kappa_1} - \frac{F_2(1+e)(r+r_2)}{\kappa_2}, \tag{3.40}$$

$$\mathcal{C} = \frac{\mathcal{Q}_1(1-e)}{\kappa_1} - \frac{\mathcal{Q}_2(1+e)}{\kappa_2}, \tag{3.41}$$

$$\mathcal{D} = (1-e)^2(1-\cos\psi)\frac{\Delta_1}{\kappa_1} + (1+e)^2(1+\cos\psi)\frac{\Delta_2}{\kappa_2}, \tag{3.42}$$

and

$$\begin{aligned}\mathcal{Q}_{1/2} =& -2a\mathcal{L}rr_{1/2} - a^4\mathcal{E}(r+r_{1/2}) + a^3\mathcal{L}(r+r_{1/2})\\ &- a^2\mathcal{E}(r^3 + r^2r_{1/2} + r_{1/2}^3 + rr_{1/2}(r_{1/2}-2))\\ &- \mathcal{E}rr_{1/2}(rr_{1/2}(r+r_{1/2}) - 2(r^2 + rr_{1/2} + r_{1/2}^2))\\ &- a^2(2a^2\mathcal{E} - 2\mathcal{E}rr_{1/2} + a\mathcal{L}(r+r_{1/2}-2))\cos^2\theta.\end{aligned} \tag{3.43}$$

---

[1] Note that these expressions will not match those found in Ref. [137] as we have fixed a typo in these expressions.

Deriving averaging transformations for these evolution equations for the quasi-Keplerian phases requires a zeroth order term in the phase transformations to account for the fact that the leading order rates of change depend on the phases i.e. $\hat{\Upsilon}_r(\psi)$ and $\hat{\Upsilon}_z(\chi)$ as demonstrated in (2.18). While this is not an insurmountable challenge as shown in Ref. [113], it is inconvenient to implement. Instead, we chose to employ an action angle prescription of the geodesic motion and use the action angles q_r and q_z as our radial and polar phases respectively. This allows us to avoid this zeroth order term in the averaging transformations and allows us take advantage of analytic solutions to the geodesic equations outlined in Chapter 2.

3.3.2 Evolution of action angles

We now look to derive evolution equations for the action angles q_r and q_z. The action angles will still evolve with their respective Mino-time frequencies, but now pick up a correction due to the evolution of the initial values. If we take the Mino time derivative of the solution for the orbital phases at geodesic order given by Eq.(2.23), one obtains:

$$\frac{dq_i}{d\lambda} = \Upsilon_i^{(0)} + \frac{dq_{i,0}}{d\lambda}. \tag{3.44}$$

To find the evolution equations for the initial values for the orbital phases, we can re-arrange the first osculating condition (3.10a) and exploit the fact that the evolution of r is independent of q_z, and the evolution z is independent of q_r, to get

$$\frac{dq_{i,0}}{d\lambda} = -\frac{1}{\partial x^i_G/\partial q_i}\left(\frac{\partial x^i_G}{\partial P_j}\frac{dP_j}{d\lambda}\right) := f_i, \tag{3.45}$$

where x^i_G is the geodesic expression for r or z given by Eqs. (2.24a) and (2.24b) respectively.

Unfortunately, this expression is difficult to evaluate numerically at the orbital turning points where $\partial x^i_G/\partial q_i$ goes to zero. This is not a physical singularity as we can derive an equivalent evolution equation that is finite at turning points be following the procedure outlined in Ref. [137]. We start by considering the definition of the geodesic potentials:

$$V_i(x^i, \vec{P}) = \left(\frac{dx^i}{d\lambda}\right)^2. \tag{3.46}$$

If we add together the derivative of both sides of this expression with respect to P_j and then multiply both sides by $\dot{P}_j$, one obtains:

$$\frac{\partial V_i}{\partial x^i}\left(\frac{\partial x^i}{\partial P_j}\dot{P}_j\right) + \frac{\partial V_i}{\partial P_j}\dot{P}_j = 2\dot{x}^i\left(\frac{\partial \dot{x}^i}{\partial P_j}\dot{P}_j\right). \tag{3.47}$$

Rearranging and plugging in Eq. (3.45) allows us to write

$$\frac{\partial V}{\partial x^i}\frac{\partial x^i}{\partial q_i}\dot{q}_{i0} = \frac{\partial V_i}{\partial P_j}\dot{P}_j - 2\dot{x}^i\left(\frac{\partial \dot{x}^i}{\partial P_j}\dot{P}_j\right). \tag{3.48}$$

We also note that taking the derivative of (3.46) with respect to Mino time λ and rearranging yields

$$\frac{\partial V_i}{\partial P_j}\dot{P}_j = 2\dot{x}^i\ddot{x}^i - \frac{\partial V_i}{\partial x^i}\dot{x}^i. \tag{3.49}$$

Rearranging this and subbing into equation (3.48) and simplifying gives us

$$\frac{\partial V}{\partial x_i}\frac{\partial x^i}{\partial q_i}\dot{q}_{i0} = 2\dot{x}^i\left(\left[\ddot{x}^i - \frac{1}{2}\frac{\partial V_i}{\partial x^i}\right] - \left(\frac{\partial \dot{x}^i}{\partial P_j}\dot{P}_j\right)\right). \tag{3.50}$$

We note that the square bracket term will vanish for geodesics. For perturbed orbits this means that this term will be proportional to the component of the four-acceleration a^i scaled by a factor of Σ^2 to compensate for taking derivatives with respect to λ instead of τ. When evaluating this expression, we make use of the osculating conditions (3.2) which leads to the simplification

$$\dot{x}^i = \dot{x}^i_G = \frac{\partial x^i_G}{\partial q_i}\frac{dq_i}{d\lambda} = \frac{\partial x^i_G}{\partial q_i}\Upsilon^{(0)}_i. \tag{3.51}$$

Combining these results with equation (3.50) gives us our final expression for the evolution of the initial phases:

$$\frac{dq_{i,0}}{d\lambda} = \frac{2\Upsilon^{(0)}_i}{\partial V_i/\partial x^i_G}\left(\Sigma^2 a^i - \left(\frac{\partial \dot{x}^i_G}{\partial P_j}\dot{P}_j\right)\right) := f_i. \tag{3.52}$$

This expression instead has a singularity whenever $\partial V_i/\partial x^i_G = 0$. Thus, for our numerical implementation, we use Eq. (3.45) for the majority of the orbital cycle and switch to Eq. (3.52) in the vicinity of turning points.

3.4 Equations of motion for self-forced generic Kerr inspirals

Finally, we also require evolution equations for "extrinsic quantities" that don't show up on the right hand side of the equations of motion, but are necessary to compute the trajectory and the waveform. These are the time and azimuthal coordinates of the secondary which, as a set, we denote by $\vec{S} = \{t, \phi\}$. Their evolution is given by the geodesic equations for t and ϕ, i.e., equations (2.7c) and (2.7d). Putting it all together, the equations of motion take the form:

$$\dot{P}_j = F_j(\vec{P}, \vec{q}), \tag{3.53a}$$

$$\dot{q}_i = \Upsilon^{(0)}_i(\vec{P}) + f_i(\vec{P}, \vec{q}), \tag{3.53b}$$

$$\dot{S}_k = f^{(0)}_k(\vec{P}, \vec{q}). \tag{3.53c}$$

These equations of motion are valid for generic inspirals under the influence of an unspecified perturbing force. However, we are interested in the case where the perturbing force is the GSF which scales with the small mass ratio $\epsilon = \mu/M$, meaning that the secondary's four-acceleration can be expressed as $a^\alpha = \epsilon a^\alpha_{(1)} + \epsilon^2 a^\alpha_{(2)} + \mathcal{O}(\epsilon^2)$. Factoring out this scaling, the equations of motion for generic Kerr inspirals become

$$\dot{P}_j = \epsilon F^{(1)}_j(\vec{P}, \vec{q}) + \epsilon^2 F^{(2)}_j(\vec{P}, \vec{q}), \tag{3.54a}$$

$$\dot{q}_i = \Upsilon^{(0)}_i(\vec{P}) + \epsilon f^{(1)}_i(\vec{P}, \vec{q}), \tag{3.54b}$$

$$\dot{S}_k = f^{(0)}_k(\vec{P}, \vec{q}). \tag{3.54c}$$

We find that the action angle implementation produces inspiral trajectories that are identical to inspirals calculated using the null tetrad formulation with quasi-Keplerian angles. We have implemented both the action angle and null tetrad osculating element equations into a Mathematica package that will be publicly available on the Black Hole Perturbation Toolkit [182]. We find numerically that the null tetrad formulation is more computationally efficient as it does not have any singular equations that necessitate switching between different evolution equations twice every orbital period. As such, for direct comparisons between OG and NIT inspirals we make use of the null tetrad formulation, but use the action angle formulation as the starting point for our averaging procedure.

Chapter 4

Near Identity Transformations

Near identity (averaging) transformations (NITs) are well known technique in applied mathematics and celestial mechanics [155]. This technique involves making small transformations to the equations of motion, such that the short timescale physics is averaged out, while retaining information about the long term evolution of a system.

In Ref. [154], these transformations were derived for a generic EMRI system which we review in Sec. 4.1. However, such transformations become ill-behaved in the presence of an orbital resonance. For this reason, we derive a version of the NIT that is valid in the vicinity of an orbital resonance in Sec. 4.2.

When we perform these transformations and solve the resulting averaged equitations of motion, the solutions are parametrized in terms of Mino time λ. It would be much more convenient for data analysis if the final solutions were instead parametrized by the Boyer-Lindquist coordinate time t as this can be easily to the retarded time at the detector. Doing so either requires numerically interpolating $\lambda(t)$ which could take minutes, or adding orbital timescale oscillations back into our NIT equations of motion. However, it was noted in Ref . [113] that these oscillations can be averaged out if one performs an additional averaging transformation. We outline this details of this procedure in Sec. 4.3.

Finally, in Sec. 4.4 we briefly describe a mathematically equivalent averaging technique that makes use of the natural separation of timescales in the EMRI problem, the two-timescale expansion.

4.1 A review of near-identity transformations for generic EMRI systems

4.1.1 Near Identity Transformation

The NIT variables, $\tilde{P}_j$, $\tilde{q}_i$ and $\tilde{S}_k$, are related to the OG variables P_j, q_i and S_k via

$$\tilde{P}_j = P_j + \epsilon Y_j^{(1)}(\vec{P},\vec{q}) + \epsilon^2 Y_j^{(2)}(\vec{P},\vec{q}) + \mathcal{O}(\epsilon^3), \tag{4.1a}$$

$$\tilde{q}_i = q_i + \epsilon X_i^{(1)}(\vec{P},\vec{q}) + \epsilon^2 X_i^{(2)}(\vec{P},\vec{q}) + \mathcal{O}(\epsilon^3), \tag{4.1b}$$

$$\tilde{S}_k = S_k + Z_k^{(0)}(\vec{P},\vec{q}) + \epsilon Z_k^{(1)}(\vec{P},\vec{q}) + \mathcal{O}(\epsilon^2). \tag{4.1c}$$

Here, the transformation functions $Y_j^{(n)}$, $X_i^{(n)}$, and $Z_k^{(n)}$ are required to be smooth, periodic functions of the orbital phases $\vec{q}$. At leading order, Eqs. (4.1) are identity transformations for P_k and q_i but not for S_k due to the presence of a zeroth order transformation term $Z_k^{(0)}$. The inverse transformations can be found for P_k and q_i by requiring that their composition with the transformations in Eqs.(4.1) must give the identity transformation. Expanding order by order in ϵ, this gives us

$$\begin{aligned} P_j = {} & \tilde{P}_j - \epsilon Y_j^{(1)}(\vec{\tilde{P}},\vec{\tilde{q}}) \\ & - \epsilon^2 \left(Y_j^{(2)}(\vec{\tilde{P}},\vec{\tilde{q}}) - \frac{\partial Y_j^{(1)}(\vec{\tilde{P}},\vec{\tilde{q}})}{\partial \tilde{P}_k} Y_k^{(1)}(\vec{\tilde{P}},\vec{\tilde{q}}) - \frac{\partial Y_j^{(1)}(\vec{\tilde{P}},\vec{\tilde{q}})}{\partial \tilde{q}_k} X_k^{(1)}(\vec{\tilde{P}},\vec{\tilde{q}}) \right) + \mathcal{O}(\epsilon^3), \end{aligned} \tag{4.2a}$$

$$\begin{aligned} q_i = {} & \tilde{q}_i - \epsilon X_i^{(1)}(\vec{\tilde{P}},\vec{\tilde{q}}) \\ & - \epsilon^2 \left(X_i^{(2)}(\vec{\tilde{P}},\vec{\tilde{q}}) - \frac{\partial X_i^{(1)}(\vec{\tilde{P}},\vec{\tilde{q}})}{\partial \tilde{P}_j} Y_j^{(1)}(\vec{\tilde{P}},\vec{\tilde{q}}) - \frac{\partial X_i^{(1)}(\vec{\tilde{P}},\vec{\tilde{q}})}{\partial \tilde{q}_k} X_k^{(1)}(\vec{\tilde{P}},\vec{\tilde{q}}) \right) + \mathcal{O}(\epsilon^3). \end{aligned} \tag{4.2b}$$

4.1.2 Transformed Equations of Motion

By taking the time derivative of the NIT (4.1), substituting the EMRI equations of motion (3.54) and inverse NIT (4.2), and expanding in powers of ϵ we obtain the NIT transformed equations of motions

$$\frac{d\tilde{P}_j}{d\lambda} = \epsilon \tilde{F}_j^{(1)}(\vec{\tilde{P}},\vec{\tilde{q}}) + \epsilon^2 \tilde{F}_j^{(2)}(\vec{\tilde{P}},\vec{\tilde{q}}) + \mathcal{O}(\epsilon^3), \tag{4.3a}$$

$$\frac{d\tilde{q}_i}{d\lambda} = \Upsilon_i^{(0)}(\vec{\tilde{P}}) + \epsilon \Upsilon_i^{(1)}(\vec{\tilde{P}},\vec{\tilde{q}}) + \mathcal{O}(\epsilon^2), \tag{4.3b}$$

where

$$\tilde{F}_j^{(1)} = F_j^{(1)} + \frac{\partial Y_j^{(1)}}{\partial \tilde{q}_i} \Upsilon_i^{(0)}, \qquad \Upsilon_i^{(1)} = f_i^{(1)} + \frac{\partial X_i^{(1)}}{\partial \tilde{q}_k} \Upsilon_k^{(0)} - \frac{\partial \Upsilon_i^{(0)}}{\partial \tilde{P}_j} Y_j^{(1)}, \tag{4.4a-b}$$

and

$$\tilde{F}_j^{(2)} = F_j^{(2)} + \frac{\partial Y_j^{(2)}}{\partial \tilde{q}_i} \Upsilon_i^{(0)} + \frac{\partial Y_j^{(1)}}{\partial \tilde{q}_i} f_i^{(1)} + \frac{\partial Y_j^{(1)}}{\partial \tilde{P}_k} F_k^{(1)} - \frac{\partial \tilde{F}_j^{(1)}}{\partial \tilde{P}_k} Y_k^{(1)} - \frac{\partial \tilde{F}_j^{(1)}}{\partial \tilde{q}_i} X_i^{(1)} \tag{4.5}$$

Note that all functions on the right hand side are evaluated at $\vec{\tilde{P}}$ and $\vec{\tilde{q}}$ and that we have adopted the convention that all repeated roman indices are summed over.

4.1.3 Cancellation of oscillating terms at $\mathcal{O}(\epsilon)$

Since each of the functions on the right hand side of our equations of motion are 2π periodic in the orbital phases, it will be useful to decompose these functions into Fourier series where we use the convention.

$$A(\vec{P},\vec{q}) = \sum_{\vec{\kappa} \in \mathbb{Z}^{i_{\mathsf{max}}}} A_{\vec{\kappa}}(\vec{P}) e^{i\vec{\kappa}\cdot\vec{q}}, \tag{4.6}$$

where i_{max} is the number of orbital phases. Based on this, we can split the function into an averaged piece $\langle A \rangle (\vec{P})$ given by

$$\langle A \rangle (\vec{P}) = A_{\vec{0}}(\vec{P}) = \frac{1}{(2\pi)^{i_{\mathsf{max}}}} \int \cdots \int_{\vec{q}} A(\vec{P},\vec{q}) dq_1 \ldots dq_{i_{\mathsf{max}}}, \tag{4.7}$$

and an oscillating piece given by

$$\breve{A}(\vec{P},\vec{q}) = A(\vec{P},\vec{q}) - \langle A \rangle (\vec{P}) = \sum_{\vec{\kappa} \neq \vec{0}} A_{\vec{\kappa}}(\vec{P}) e^{i\vec{\kappa}\cdot\vec{q}}. \tag{4.8}$$

Using this split, we can recast the expression for $\tilde{F}_j^{(1)}$ as

$$\begin{aligned}\tilde{F}_j^{(1)} &= F_j^{(1)} + \frac{\partial Y_j^{(1)}}{\partial \tilde{q}_i} = F_j^{(1)} + \frac{\partial \breve{Y}_j^{(1)}}{\partial \tilde{q}_i} \\ &= \left\langle F_j^{(1)} \right\rangle + \sum_{\vec{\kappa}\neq\vec{0}} \left(F_{j,\vec{\kappa}}^{(1)} + i(\vec{\kappa}\cdot\vec{\Upsilon}^{(0)})\breve{Y}_{j,\vec{\kappa}}^{(1)} \right) e^{i\vec{\kappa}\cdot\vec{q}}\end{aligned} \tag{4.9}$$

As such, we can cancel the oscillatory pieces of $\tilde{F}_j^{(1)}$ by choosing the oscillatory part of $Y_j^{(1)}$ to be

$$\breve{Y}_{j,\vec{\kappa}}^{(1)} := \frac{i}{\vec{\kappa}\cdot\vec{\Upsilon}^{(0)}} F_{j,\vec{\kappa}}^{(1)}(\vec{P}). \tag{4.10}$$

Clearly, one can only make this choice so long as there is no $\vec{\kappa}_{\mathsf{res}} = \{k_r, k_z\}$ where $k_r, k_z \in \mathbb{Z}$, such that $\vec{\kappa}_{\mathsf{res}} \cdot \vec{\Upsilon}^{(0)} = k_r \Upsilon_r^{(0)} + k_z \Upsilon_z^{(0)} = 0$. While this is usually the case, this breaks down in the presence of orbital resonances when the radial and polar frequencies become commensurate. We will deal with this case in Sec. 4.2, but for now we will assume that we are far from any orbital resonances.

Using the above choice for $\breve{Y}_j^{(1)}$, the equation for $\Upsilon_i^{(1)}$ becomes

$$\begin{aligned}\Upsilon_i^{(1)} &= f_i^{(1)} - \frac{\partial \Upsilon_i^{(0)}}{\partial \tilde{P}_j} Y_j^{(1)} + \frac{\partial X_i^{(1)}}{\partial \tilde{q}_k} \Upsilon_k^{(0)} \\ &= \left\langle f_i^{(1)} \right\rangle - \frac{\partial \Upsilon_i^{(0)}}{\partial \tilde{P}_j} \left\langle Y_j^{(1)} \right\rangle + \sum_{\vec{\kappa}\neq\vec{0}} \left(f_{i,\vec{\kappa}}^{(1)} - \frac{i}{\vec{\kappa}\cdot\vec{\Upsilon}^{(0)}} \frac{\partial \Upsilon_i^{(0)}}{\partial \tilde{P}_j} F_{j,\vec{\kappa}}^{(1)} + i(\vec{\kappa}\cdot\vec{\Upsilon}^{(0)}) \breve{X}_{i,\vec{\kappa}}^{(1)} \right) e^{i\vec{\kappa}\cdot\vec{q}}\end{aligned} \tag{4.11}$$

As a result, we can remove the oscillating pieces of $\Upsilon_i^{(1)}$ by choosing

$$\breve{X}_{i,\vec{\kappa}}^{(1)} := \frac{i}{\vec{\kappa}\cdot\vec{\Upsilon}^{(0)}} f_{i,\vec{\kappa}}^{(1)} + \frac{1}{(\vec{\kappa}\cdot\vec{\Upsilon}^{(0)})^2} \frac{\partial \Upsilon_i^{(0)}}{\partial P_j} F_{j,\vec{\kappa}}^{(1)}. \tag{4.12}$$

4.1.4 Cancellation of oscillating terms at $\mathcal{O}(\epsilon^2)$

Using the above choice for $\breve{Y}_j^{(1)}$, we can express the oscillatory part of the expression for $\tilde{F}_j^{(2)}$ as

$$\begin{aligned}\breve{\tilde{F}}_j^{(2)} =& \breve{F}_j^{(2)} + \frac{\partial \breve{Y}_j^{(2)}}{\partial \tilde{q}_i} \Upsilon_i^{(0)} + \left\{ \frac{\partial Y_j^{(1)}}{\partial \tilde{q}_i} f_i^{(1)} \right\} + \left\{ \frac{\partial Y_j^{(1)}}{\partial \tilde{P}_k} F_k^{(1)} \right\} - \frac{\partial \left\langle F_j^{(1)} \right\rangle}{\partial \tilde{P}_k} \breve{Y}_k^{(1)} \\ =& \sum_{\vec{\kappa}\neq\vec{0}} \Bigg(F_{j,\vec{\kappa}}^{(2)} + i(\vec{\kappa}\cdot\vec{\Upsilon}^{(0)}) \breve{Y}_{j,\vec{\kappa}}^{(2)} + \frac{\partial \left\langle Y_j^{(1)} \right\rangle}{\partial \tilde{P}_k} F_{k,\kappa}^{(1)} - i \frac{\partial \left\langle F_j^{(1)} \right\rangle}{\partial \tilde{P}_k} \frac{F_{k,\vec{\kappa}}}{\vec{\kappa}\cdot\vec{\Upsilon}^{(0)}} \\ &+ \sum_{\vec{\kappa}'\neq\vec{0}} \left(i \frac{F_{k,\vec{\kappa}-\vec{\kappa}'}^{(1)}}{\vec{\kappa}'\cdot\vec{\Upsilon}^{(0)}} \left(\frac{\partial F_{j,\vec{\kappa}'}^{(1)}}{\partial \tilde{P}_k} - \frac{F_{j,\vec{\kappa}'}^{(1)}}{\vec{\kappa}'\cdot\vec{\Upsilon}^{(0)}} \frac{\partial (\vec{\kappa}'\cdot\vec{\Upsilon}^{(0)})}{\partial \tilde{P}_k} \right) - \frac{\vec{\kappa}'\cdot\vec{f}_{\vec{\kappa}-\vec{\kappa}'}^{(1)}}{\vec{\kappa}'\cdot\vec{\Upsilon}^{(0)}} F_{j,\vec{\kappa}'}^{(1)} \right) \Bigg) e^{i\kappa\cdot\vec{q}},\end{aligned} \tag{4.13}$$

where we introduced the additional notation $\{\cdot\}$ to denote the oscillatory part of a product of functions. Thus we can remove the oscillatory part of $\tilde{F}_j^{(2)}$ by choosing

$$\begin{aligned}\breve{Y}_{j,\vec{\kappa}}^{(2)} =& \frac{i}{\vec{\kappa}\cdot\vec{\Upsilon}^{(0)}}\left(F_{j,\vec{\kappa}}^{(2)} + \frac{\partial\left\langle Y_j^{(1)}\right\rangle}{\partial\tilde{P}_k}F_{k,\vec{\kappa}}^{(1)} - i\frac{\partial\left\langle F_j^{(1)}\right\rangle}{\partial\tilde{P}_k}\frac{F_{k,\vec{\kappa}}^{(1)}}{\vec{\kappa}\cdot\vec{\Upsilon}^{(0)}}\right.\\ &\left. + \sum_{\vec{\kappa}'\neq\vec{0}}\left(i\frac{F_{k,\vec{\kappa}-\vec{\kappa}'}^{(1)}}{\vec{\kappa}'\cdot\vec{\Upsilon}^{(0)}}\left(\frac{\partial F_{j,\vec{\kappa}'}^{(1)}}{\partial\tilde{P}_k} - \frac{F_{j,\vec{\kappa}'}^{(1)}}{\vec{\kappa}'\cdot\vec{\Upsilon}^{(0)}}\frac{\partial(\vec{\kappa}'\cdot\vec{\Upsilon}^{(0)})}{\partial\tilde{P}_k}\right) - \frac{\vec{\kappa}'\cdot\vec{f}_{\vec{\kappa}-\vec{\kappa}'}^{(1)}}{\vec{\kappa}'\cdot\vec{\Upsilon}^{(0)}}F_{j,\vec{\kappa}'}^{(1)}\right)\right).\end{aligned} \tag{4.14}$$

4.1.5 Freedom in the averaged pieces

With the oscillatory pieces of the NIT equations of motion removed, terms in the equations of motion become

$$\tilde{F}_j^{(1)} = \left\langle F_j^{(1)}\right\rangle, \quad \Upsilon_i^{(1)} = \left\langle f_i^{(1)}\right\rangle - \frac{\partial\Upsilon_i^{(0)}}{\partial\tilde{P}_j}\left\langle Y_j^{(1)}\right\rangle, \tag{4.15a-b}$$

and

$$\tilde{F}_j^{(2)} = \left\langle F_j^{(2)}\right\rangle + \left\langle\frac{\partial\breve{Y}_j^{(1)}}{\partial\tilde{q}_i}\breve{f}_i^{(1)}\right\rangle + \left\langle\frac{\partial\breve{Y}_j^{(1)}}{\partial\tilde{P}_k}F_k^{(1)}\right\rangle + \frac{\partial\left\langle Y_j^{(1)}\right\rangle}{\partial\tilde{P}_k}\left\langle F_k^{(1)}\right\rangle - \frac{\partial\left\langle F_j^{(1)}\right\rangle}{\partial\tilde{P}_k}\left\langle Y_k^{(1)}\right\rangle \tag{4.16}$$

Note that we still have freedom to set the averaged pieces of the transformation functions $\left\langle Y_j^{(1)}\right\rangle$, $\left\langle Y_j^{(2)}\right\rangle$ and $\left\langle X_i^{(1)}\right\rangle$ to be anything we choose. There are many valid and interesting choices that one could make that are explored in Ref. [154]. For this work, we make use of the simplest choice: $\left\langle Y_j^{(1)}\right\rangle = \left\langle Y_j^{(2)}\right\rangle = \left\langle X_i^{(1)}\right\rangle = 0$, as this makes it easy to compare between OG and NIT inspirals. It also has the benefit of drastically simplifying equations of motion to

$$\tilde{F}_j^{(1)} = \left\langle F_j^{(1)}\right\rangle, \quad \Upsilon_i^{(1)} = \left\langle f_i^{(1)}\right\rangle, \tag{4.17a-b}$$

and

$$\tilde{F}_j^{(2)} = \left\langle F_j^{(2)}\right\rangle + \left\langle\frac{\partial\breve{Y}_j^{(1)}}{\partial\tilde{q}_i}\breve{f}_i^{(1)}\right\rangle + \left\langle\frac{\partial\breve{Y}_j^{(1)}}{\partial\tilde{P}_k}F_k^{(1)}\right\rangle \tag{4.18}$$

4.1.6 Evolution of extrinsic quantities

Now we look to remove the oscillatory pieces of the evolution equations for the extrinsic quantities $\vec{S}$:

$$\frac{dS}{d\lambda} = f_k^{(0)}(\vec{P},\vec{q}). \tag{4.19}$$

By substituting for $\vec{P}$ and $\vec{q}$ using the inverse transformation (4.2) and re-expanding in ϵ we can write this as an equation involving only the NIT variables $\vec{\tilde{P}}$ and $\vec{\tilde{q}}$,

$$\frac{dS}{d\lambda} = f_k^{(0)} - \epsilon\left(\frac{\partial f_k^{(0)}}{\partial\tilde{P}_j}Y_j^{(1)} + \frac{\partial f_k^{(0)}}{\partial\tilde{q}_i}X_i^{(1)}\right) + \mathcal{O}(\epsilon^2) \tag{4.20}$$

where all of the functions on the right hand side are now functions of $\vec{\tilde{P}}$ and $\vec{\tilde{q}}$. In order to remove the oscillatory pieces of the equations, we make use of a new set of extrinsic quantities $\vec{\tilde{S}}$ that are related to the original quantities by Eq. (4.1c). We note that since this transformation has a zeroth order in mass ratio term $Z_k^{(0)}$, it is not a near-identity transformation. Thus when we produce waveforms it will be necessary to be able to calculate $Z_k^{(0)}$ explicitly.

We then take the time derivative of Eq. (4.1c), substitute the equations of motion for $\vec{S}$ and expand order by order to obtain equations of motion for for $\vec{\tilde{S}}$:

$$\frac{d\tilde{S}_k}{d\lambda} = \Upsilon_k^{(0)} + \epsilon \Upsilon_k^{(1)} + \mathcal{O}(\epsilon^2), \tag{4.21}$$

where

$$\Upsilon_k^{(0)} = f_k^{(0)} + \frac{\partial Z_k^{(0)}}{\partial \tilde{q}_i} \Upsilon_i^{(0)} \tag{4.22a}$$

$$\Upsilon_k^{(1)} = \frac{\partial Z_k^{(0)}}{\partial \tilde{q}_i} \Upsilon_i^{(1)} + \frac{\partial Z_k^{(0)}}{\partial \tilde{P}_j} \tilde{F}_j^{(1)} + \frac{\partial Z_k^{(1)}}{\partial \tilde{q}_i} \Upsilon_i^{(0)} - \frac{\partial f_k^{(0)}}{\partial \tilde{P}_j} Y_j^{(1)} - \frac{\partial f_k^{(0)}}{\partial \tilde{q}_i} X_i^{(1)} \tag{4.22b}$$

We can now remove the oscillating pieces of the functions $\Upsilon_k^{(0)}$ by solving the equations

$$\breve{f}_k^{(0)} + \frac{\partial \breve{Z}_k^{(0)}}{\partial \tilde{q}_i} \Upsilon_i^{(0)} = 0 \tag{4.23a}$$

$$\frac{\partial \breve{Z}_k^{(0)}}{\partial \tilde{q}_i} \Upsilon_i^{(1)} + \frac{\partial \breve{Z}_k^{(0)}}{\partial \tilde{P}_j} \tilde{F}_j^{(1)} + \frac{\partial \breve{Z}_k^{(1)}}{\partial \tilde{q}_i} \Upsilon_i^{(0)} - \left\{ \frac{\partial f_k^{(0)}}{\partial \tilde{P}_j} Y_j^{(1)} \right\} - \left\{ \frac{\partial f_k^{(0)}}{\partial \tilde{q}_i} X_i^{(1)} \right\} = 0 \tag{4.23b}$$

for the oscillatory parts of the transformation $\breve{Z}_k^{(0)}$ and $\breve{Z}_k^{(1)}$. The first of these is satisfied by using the oscillating pieces for the analytic solutions for the geodesic motion of t and ϕ,

$$\breve{Z}_k^{(0)} = -\breve{S}_{k,r}(q_r) - \breve{S}_{k,z}(q_z). \tag{4.24}$$

It is unclear whether the equation for $Z_k^{(1)}$ would analytic solutions, but it can be solved numerically. Since we only need to know the extrinsic quantities to $\mathcal{O}(\epsilon)$ to generate waveforms, we do not need to be able to calculate this explicitly and it is sufficient to know that a solution exists.

Now the forcing functions only depend only on $\vec{\tilde{P}}$ and are given by

$$\Upsilon_k^{(0)} = \left\langle f_k^{(0)} \right\rangle \tag{4.25a}$$

$$\Upsilon_k^{(1)} = \frac{\partial \left\langle Z_k^{(0)} \right\rangle}{\partial \tilde{P}_j} \tilde{F}_j^{(1)} - \frac{\partial \left\langle f_k^{(0)} \right\rangle}{\partial \tilde{P}_j} \left\langle Y_j^{(1)} \right\rangle - \left\langle \frac{\partial \breve{f}_k^{(0)}}{\partial \tilde{P}_j} \breve{Y}_j^{(1)} \right\rangle - \left\langle \frac{\partial \breve{f}_k^{(0)}}{\partial \tilde{q}_i} \breve{X}_i^{(1)} \right\rangle \tag{4.25b}$$

Again, we have freedom in the average pieces of the transformation functions which we use to simplify this problem further. As before, we chose the simplest option and set $\left\langle Z_k^{(0)} \right\rangle = 0$ which along with our previous choices simplifies the expression for $\Upsilon_k^{(1)}$ to be

$$\Upsilon_k^{(1)} = -\left\langle \frac{\partial \breve{f}_k^{(0)}}{\partial \tilde{P}_j} \breve{Y}_j^{(1)} \right\rangle - \left\langle \frac{\partial \breve{f}_k^{(0)}}{\partial \tilde{q}_i} \breve{X}_i^{(1)} \right\rangle. \tag{4.26}$$

4.1.7 Summary of NIT Results

In summary, the equations of motion for the NIT variables $\tilde{P}_j, \tilde{q}_i$ and $\tilde{S}_k$ now take the form

$$\frac{d\tilde{P}_j}{d\lambda} = \epsilon \tilde{F}_j^{(1)}(\vec{\tilde{P}}) + \epsilon^2 \tilde{F}_j^{(2)}(\vec{\tilde{P}}) + \mathcal{O}(\epsilon^3), \tag{4.27a}$$

$$\frac{d\tilde{q}_i}{d\lambda} = \Upsilon_i^{(0)}(\vec{\tilde{P}}) + \epsilon \Upsilon_i^{(1)}(\vec{\tilde{P}}) + \mathcal{O}(\epsilon^2), \tag{4.27b}$$

$$\frac{d\tilde{S}_k}{d\lambda} = \Upsilon_k^{(0)}(\vec{\tilde{P}}) + \epsilon \Upsilon_k^{(1)}(\vec{\tilde{P}}) + \mathcal{O}(\epsilon^2). \tag{4.27c}$$

Crucially, these equations of motion are now independent of the orbital phases $\vec{q}$.

We chose the average pieces of the transformation terms to be $\left\langle Y_j^{(1)} \right\rangle = \left\langle Y_j^{(2)} \right\rangle = \left\langle X_i^{(1)} \right\rangle = \left\langle Z_k^{(0)} \right\rangle = \left\langle Z_k^{(1)} \right\rangle = 0$ and so the transformed forcing functions are related to the original functions by

$$\tilde{F}_j^{(1)} = \left\langle F_j^{(1)} \right\rangle, \quad \Upsilon_i^{(1)} = \left\langle f_i^{(1)} \right\rangle, \quad \Upsilon_k^{(0)} = \left\langle f_k^{(0)} \right\rangle, \tag{4.28a-c}$$

$$\tilde{F}_j^{(2)} = \left\langle F_j^{(2)} \right\rangle + \left\langle \frac{\partial \breve{Y}_j^{(1)}}{\partial \tilde{q}_i} \breve{f}_i^{(1)} \right\rangle + \left\langle \frac{\partial \breve{Y}_j^{(1)}}{\partial \tilde{P}_k} \breve{F}_k^{(1)} \right\rangle, \tag{4.28d}$$

$$\Upsilon_k^{(1)} = -\left\langle \frac{\partial \breve{f}_k^{(0)}}{\partial \tilde{P}_j} \breve{Y}_j^{(1)} \right\rangle - \left\langle \frac{\partial \breve{f}_k^{(0)}}{\partial \tilde{q}_i} \breve{X}_i^{(1)} \right\rangle. \tag{4.28e}$$

In deriving these equations of motion, we have constrained the oscillating pieces of the NIT transformation functions to be

$$\breve{Y}_j^{(1)} := \sum_{\vec{\kappa} \neq \vec{0}} \frac{i}{\vec{\kappa} \cdot \vec{\Upsilon}} F_{j,\vec{\kappa}}^{(1)} e^{i\vec{\kappa} \cdot \vec{q}}, \tag{4.29}$$

$$\breve{X}_i^{(1)} := \sum_{\vec{\kappa} \neq \vec{0}} \left(\frac{i}{\vec{\kappa} \cdot \vec{\Upsilon}} f_{i,\vec{\kappa}}^{(1)} + \frac{1}{(\vec{\kappa} \cdot \vec{\Upsilon})^2} \frac{\partial \Upsilon_i}{\partial P_j} F_{j,\vec{\kappa}}^{(1)} \right) e^{i\vec{\kappa} \cdot \vec{q}}. \tag{4.30}$$

After numerically solving the equations of motion, computing a waveform only requires knowledge of the transformations in Eq. (4.1) to zeroth order in the mass ratio so that the error is $\mathcal{O}(\epsilon)$ i.e.,

$$P_j = \tilde{P}_j + \mathcal{O}(\epsilon), \tag{4.31a}$$

$$q_i = \tilde{q}_i + \mathcal{O}(\epsilon), \tag{4.31b}$$

$$S_k = \tilde{S}_k - Z_k^{(0)}(\vec{\tilde{P}}, \vec{\tilde{q}}) + \mathcal{O}(\epsilon). \tag{4.31c}$$

where the zeroth order transformation term for the extrinsic quantities $\breve{Z}_k^{(0)}$ is known analytically as it related to the analytic solutions for the geodesic equations for t and ϕ by

$$\breve{Z}_k^{(0)} = -\breve{S}_{k,r}(q_r) - \breve{S}_{k,z}(q_z). \tag{4.32}$$

Furthermore, to be able to directly compare between OG and NIT inspirals, we will need to match their initial conditions to sufficient accuracy. To maintain an overall phase difference of $\mathcal{O}(\epsilon)$ in the course of an inspiral, the initial values of the phases and extrinsic quantities need only be known to zeroth order in ϵ. However, we need to know the initial values of the orbital elements $\vec{P}$ to linear order in ϵ and so we use

$$\tilde{P}_j(0) = P_j(0) + \epsilon Y_j^{(1)}(\vec{P}(0), \vec{q}(0)) + \mathcal{O}(\epsilon^2). \tag{4.33}$$

4.2 Near-resonant near identity transformations

As alluded to in the previous section, this technique will break down in the presence of orbital resonances where the radial and polar frequencies become commensurate i.e. $\kappa_r \Upsilon_r + \kappa_z \Upsilon_z = 0$ where $\kappa_r, \kappa_z \in \mathbb{Z}$. As such, we will introduce a new averaging procedure which averages almost all dependence on the orbital phases, except for the resonant phase $q_\perp := \kappa_r q_r + \kappa_z q_z$. This will mean that our equations will oscillate, and so will be slower to solve than the off-resonance NIT equations of motion. However, $q_\perp$ oscillates on a timescale between that of the slow evolution of the orbital elements and the rapidly oscillating orbital phases, and so can be thought of as a "semi-fast" variable.

4.2.1 Near Identity Transformation

We will first focus on the evolution of the orbital elements and orbital phases and so we once again introduce the transformation

$$\tilde{P}_j = P_j + \epsilon Y_j^{(1)}(\vec{P}, \vec{q}) + \epsilon^2 Y_j^{(2)}(\vec{P}, \vec{q}) + \mathcal{O}(\epsilon^3), \tag{4.34a}$$

$$\tilde{q}_i = q_i + \epsilon X_i^{(1)}(\vec{P}, \vec{q}) + \epsilon^2 X_i^{(2)}(\vec{P}, \vec{q}) + \mathcal{O}(\epsilon^3), \tag{4.34b}$$

$$\tilde{q}_\perp = q_\perp + \epsilon W^{(1)}(\vec{P}, \vec{q}) + \epsilon^2 W^{(2)}(\vec{P}, \vec{q}) + \mathcal{O}(\epsilon^3), \tag{4.34c}$$

where we have implicitly imposed that none of the functions of the right hand side depend on the resonant phase $q_\perp$. This transformation has an inverse that is given by

$$\begin{aligned} P_j = {} & \tilde{P}_j - \epsilon Y_j^{(1)}(\vec{\tilde{P}}, \vec{\tilde{q}}) \\ & - \epsilon^2 \left(Y_j^{(2)}(\vec{\tilde{P}}, \vec{\tilde{q}}) - \frac{\partial Y_j^{(1)}(\vec{\tilde{P}}, \vec{\tilde{q}})}{\partial \tilde{P}_k} Y_k^{(1)}(\vec{\tilde{P}}, \vec{\tilde{q}}) - \frac{\partial Y_j^{(1)}(\vec{\tilde{P}}, \vec{\tilde{q}})}{\partial \tilde{q}_k} X_k^{(1)}(\vec{\tilde{P}}) \right) + \mathcal{O}(\epsilon^3), \end{aligned} \tag{4.35a}$$

$$\begin{aligned} q_i = {} & \tilde{q}_i - \epsilon X_i^{(1)}(\vec{\tilde{P}}, \vec{\tilde{q}}) \\ & - \epsilon^2 \left(X_i^{(2)}(\vec{\tilde{P}}, \vec{\tilde{q}}) - \frac{\partial X_i^{(1)}(\vec{\tilde{P}}, \vec{\tilde{q}})}{\partial \tilde{P}_j} Y_j^{(1)}(\vec{\tilde{P}}, \vec{\tilde{q}}) - \frac{\partial X_i^{(1)}(\vec{\tilde{P}}, \vec{\tilde{q}})}{\partial \tilde{q}_k} X_k^{(1)}(\vec{\tilde{P}}, \vec{\tilde{q}}) \right) + \mathcal{O}(\epsilon^3). \end{aligned} \tag{4.35b}$$

$$\begin{aligned} q_\perp = {} & \tilde{q}_\perp - \epsilon W_i^{(1)}(\vec{\tilde{P}}, \vec{\tilde{q}}) \\ & - \epsilon^2 \left(W_i^{(2)}(\vec{\tilde{P}}, \vec{\tilde{q}}) - \frac{\partial W_i^{(1)}(\vec{\tilde{P}}, \vec{\tilde{q}})}{\partial \tilde{P}_j} Y_j^{(1)}(\vec{\tilde{P}}, \vec{\tilde{q}}) - \frac{\partial W_i^{(1)}(\vec{\tilde{P}}, \vec{\tilde{q}})}{\partial \tilde{q}_k} X_k^{(1)}(\vec{\tilde{P}}, \vec{\tilde{q}}) \right) + \mathcal{O}(\epsilon^3). \end{aligned} \tag{4.35c}$$

4.2.2 Transformed Equations of Motion

By taking the time derivative of the NIT (4.34), substituting the EMRI equations of motion (3.54) and inverse NIT (4.35), and expanding in powers of ϵ we obtain the NIT transformed equations of motions

$$\frac{d\tilde{P}_j}{d\lambda} = \epsilon \tilde{F}_j^{(1)}(\vec{\tilde{P}}, \vec{\tilde{q}}, \tilde{q}_\perp) + \epsilon^2 \tilde{F}_j^{(2)}(\vec{\tilde{P}}, \vec{\tilde{q}}, \tilde{q}_\perp) + \mathcal{O}(\epsilon^3), \tag{4.36a}$$

$$\frac{d\tilde{q}_i}{d\lambda} = \Upsilon_i^{(0)}(\vec{\tilde{P}}) + \epsilon \Upsilon_i^{(1)}(\vec{\tilde{P}}, \vec{\tilde{q}}, \tilde{q}_\perp) + \mathcal{O}(\epsilon^2), \tag{4.36b}$$

$$\frac{d\tilde{q}_\perp}{d\lambda} = \vec{\kappa}_{\text{res}} \cdot \vec{\Upsilon}^{(0)}(\vec{\tilde{P}}) + \epsilon \vec{\kappa}_{\text{res}} \cdot \vec{\Upsilon}^{(1)}(\vec{\tilde{P}}, \vec{\tilde{q}}, \tilde{q}_\perp) + \mathcal{O}(\epsilon^2), \tag{4.36c}$$

where

$$\tilde{F}_j^{(1)} = F_j^{(1)} + \frac{\partial Y_j^{(1)}}{\partial \tilde{q}_i}\Upsilon_i^{(0)}, \quad \Upsilon_i^{(1)} = f_i^{(1)} + \frac{\partial X_i^{(1)}}{\partial \tilde{q}_k}\Upsilon_k^{(0)} - \frac{\partial \Upsilon_i^{(0)}}{\partial \tilde{P}_j}Y_j^{(1)}, \tag{4.37a-b}$$

and

$$\tilde{F}_j^{(2)} = F_j^{(2)} + \frac{\partial Y_j^{(2)}}{\partial \tilde{q}_i}\Upsilon_i^{(0)} + \frac{\partial Y_j^{(1)}}{\partial \tilde{q}_i}f_i^{(1)} + \frac{\partial Y_j^{(1)}}{\partial \tilde{P}_k}F_k^{(1)} - \frac{\partial \tilde{F}_j^{(1)}}{\partial \tilde{P}_k}Y_k^{(1)} - \frac{\partial \tilde{F}_j^{(1)}}{\partial \tilde{q}_i}X_i^{(1)} - \frac{\partial \tilde{F}_j^{(1)}}{\partial \tilde{q}_\perp}W^{(1)} \tag{4.38}$$

Note that all functions on the right hand side are evaluated at $\tilde{\vec{P}}$, $\tilde{\vec{q}}$ and $\tilde{q}_\perp$.

4.2.3 Cancellation of oscillating terms at $\mathcal{O}(\epsilon)$

We note that we can decompose any 2π periodic function into its averaged, resonant oscillatory and non-resonant oscillatory pieces using a Fourier expansion:

$$A(\vec{P}, \vec{q}, q_\perp) = \langle A \rangle (\vec{P}) + \sum_{N\neq 0} A_{N\vec{\kappa}_{\text{res}}}(\vec{P})e^{iNq_\perp} + \sum_{\vec{\kappa}\in R} A_{\vec{\kappa}}(\vec{P})e^{i\vec{\kappa}\cdot\vec{q}}. \tag{4.39}$$

where R is the set $\{\vec{\kappa} \in \mathbb{Z}^2 | \vec{\kappa} \neq N\vec{\kappa}_{\text{res}}, \forall N \in \mathbb{Z}\}$ of all non-resonant 2-tuples and $\kappa_{\text{res}} = (\kappa_r, \kappa_z)$ is such that $\kappa_{\text{res}} \cdot \vec{\Upsilon}^{(0)} = 0$.

$$\begin{aligned}
\tilde{F}_j^{(1)} &= F_j^{(1)} + \frac{\partial Y_j^{(1)}}{\partial \tilde{q}_i}\Upsilon_i^{(0)} = F_j^{(1)} + \frac{\partial \breve{Y}_j^{(1)}}{\partial \tilde{q}_i}\Upsilon_i^{(0)} \\
&= \left\langle F_j^{(1)} \right\rangle + \sum_{N\neq 0} F_{j,N\vec{\kappa}_{\text{res}}}^{(1)} e^{iNq_\perp} + \sum_{\vec{\kappa}\in R} \left(F_{j,\vec{\kappa}}^{(1)} + i\left(\vec{\kappa}\cdot\vec{\Upsilon}^{(0)}\right)\breve{Y}_{j,\vec{\kappa}}^{(1)} \right) e^{i\vec{\kappa}\cdot\vec{q}}
\end{aligned} \tag{4.40}$$

As such, we can cancel the non-resonant oscillatory pieces of $\tilde{F}_j^{(1)}$ by choosing the oscillatory part of $Y_j^{(1)}$ to be

$$\breve{Y}_{j,\vec{\kappa}}^{(1)} := \frac{i}{\vec{\kappa}\cdot\vec{\Upsilon}^{(0)}} F_{j,\vec{\kappa}}^{(1)}(\vec{P}). \tag{4.41}$$

for $\vec{\kappa} \neq N\vec{\kappa}_{\text{res}}$ and 0 when $\vec{\kappa} = N\vec{\kappa}_{\text{res}}$. Using the above choice for $\breve{Y}_j^{(1)}$, the equation for $\Upsilon_i^{(1)}$ becomes

$$\begin{aligned}
\Upsilon_i^{(1)} =& f_i^{(1)} - \frac{\partial \Upsilon_i^{(0)}}{\partial \tilde{P}_j}Y_j^{(1)} + \frac{\partial X_i^{(1)}}{\partial \tilde{q}_k}\Upsilon_k^{(0)} \\
=& \left\langle f_i^{(1)} \right\rangle - \frac{\partial \Upsilon_i^{(0)}}{\partial \tilde{P}_j}\left\langle Y_j^{(1)} \right\rangle + \sum_{N\neq 0} f_{i,N\vec{\kappa}_{\text{res}}}^{(1)} e^{iN\tilde{q}_\perp} \\
&+ \sum_{\vec{\kappa}\in R} \left(f_{i,\vec{\kappa}}^{(1)} - \frac{i}{\vec{\kappa}\cdot\vec{\Upsilon}^{(0)}}\frac{\partial \Upsilon_i^{(0)}}{\partial \tilde{P}_j}F_{j,\vec{\kappa}} + i(\vec{\kappa}\cdot\vec{\Upsilon}^{(0)})X_{i,\vec{\kappa}}^{(1)} \right) e^{i\vec{\kappa}\cdot\vec{q}}
\end{aligned} \tag{4.42}$$

As a result, we can remove the oscillating pieces of $\Upsilon_i^{(1)}$ by choosing

$$\breve{X}_{i,\vec{\kappa}}^{(1)} := \frac{i}{\vec{\kappa}\cdot\vec{\Upsilon}^{(0)}} f_{i,\vec{\kappa}}^{(1)} + \frac{1}{(\vec{\kappa}\cdot\vec{\Upsilon}^{(0)})^2}\frac{\partial \Upsilon_i^{(0)}}{\partial P_j}F_{j,\vec{\kappa}}^{(1)}. \tag{4.43}$$

for $\vec{\kappa} \neq N\vec{\kappa}_{\text{res}}$. Moreover, we can determine the transformation term W by examining the equation the equation for $\vec{\kappa} \cdot \vec{\Upsilon}^{(1)}$:

$$\begin{aligned}
\vec{\kappa} \cdot \vec{\Upsilon}^{(1)} =& \vec{\kappa} \cdot \vec{f}^{(1)} - \frac{\partial \vec{\kappa} \cdot \vec{\Upsilon}^{(0)}}{\partial \tilde{P}_j} Y_j^{(1)} + \frac{\partial W^{(1)}}{\partial \tilde{q}_k} \Upsilon_k^{(0)} \\
=& \left\langle \vec{\kappa} \cdot \vec{f}^{(1)} \right\rangle - \frac{\partial (\vec{\kappa} \cdot \vec{\Upsilon}^{(0)})}{\partial \tilde{P}_j} \left\langle Y_j^{(1)} \right\rangle + \sum_{N \neq 0} (\vec{\kappa} \cdot \vec{f}^{(1)})_{N\vec{\kappa}_{\text{res}}} e^{iN\tilde{q}_\perp} \\
&+ \sum_{\vec{\kappa} \in R} \left((\vec{\kappa} \cdot \vec{f}^{(1)})_{\vec{\kappa}} - \frac{i}{\vec{\kappa} \cdot \vec{\Upsilon}^{(0)}} \frac{\partial (\vec{\kappa} \cdot \vec{\Upsilon}^{(0)})}{\partial \tilde{P}_j} F_{j,\vec{\kappa}} + i(\vec{\kappa} \cdot \vec{\Upsilon}^{(0)}) W_{\vec{\kappa}} \right) e^{i\vec{\kappa} \cdot \vec{q}}
\end{aligned} \tag{4.44}$$

As such, to remove the oscillating pieces of $\vec{\kappa} \cdot \vec{\Upsilon}^{(1)}$, the oscillatory piece of $W^{(1)}$ must take the form

$$\breve{W}_{\vec{\kappa}}^{(1)} := \frac{i}{\vec{\kappa} \cdot \vec{\Upsilon}^{(0)}} (\vec{\kappa} \cdot \vec{f}^{(1)})_{\vec{\kappa}} + \frac{1}{(\vec{\kappa} \cdot \vec{\Upsilon}^{(0)})^2} \frac{\partial (\vec{\kappa} \cdot \vec{\Upsilon}^{(0)})}{\partial P_j} F_{j,\vec{\kappa}}^{(1)} = \vec{\kappa} \cdot \vec{\breve{X}}^{(1)}. \tag{4.45}$$

for $\vec{\kappa} \neq N\vec{\kappa}_{\text{res}}$. Note that this derivation is consistent with the fact that since $\tilde{q}_\perp = \vec{\kappa} \cdot \vec{\tilde{q}}$ then by Eq. (4.34), $W_{\vec{\kappa}}^{(1)} = \vec{\kappa} \cdot \vec{X}^{(1)}$.

4.2.4 Cancellation of oscillating terms at $\mathcal{O}(\epsilon^2)$

Using the above choice for $\breve{Y}_j^{(1)}$, we can express the non-resonant oscillatory part of the expression for $\tilde{F}_j^{(2)}$ as

$$\begin{aligned}
\breve{\tilde{F}}_j^{(2)} =& \breve{F}_j^{(2)} + \frac{\partial \breve{Y}_j^{(2)}}{\partial \tilde{q}_i} \Upsilon_i^{(0)} + \left\{ \frac{\partial Y_j^{(1)}}{\partial \tilde{q}_i} f_i^{(1)} \right\} + \left\{ \frac{\partial Y_j^{(1)}}{\partial \tilde{P}_k} F_k^{(1)} \right\} - \frac{\partial \left\langle F_j^{(1)} \right\rangle}{\partial \tilde{P}_k} \breve{Y}_k^{(1)} \\
=& \sum_{\vec{\kappa} \in R} \left(F_{j,\vec{\kappa}}^{(2)} + i(\vec{\kappa} \cdot \vec{\Upsilon}^{(0)}) Y_{j,\vec{\kappa}}^{(2)} + \frac{\partial \left\langle Y_j^{(1)} \right\rangle}{\partial \tilde{P}_k} F_{k,\kappa}^{(1)} - i \frac{\partial \left\langle F_j^{(1)} \right\rangle}{\partial \tilde{P}_k} \frac{F_{k,\vec{\kappa}}^{(1)}}{\vec{\kappa} \cdot \vec{\Upsilon}^{(0)}} \right. \\
&+ \left. \sum_{\vec{\kappa}' \in R} \left(i \frac{F_{k,\vec{\kappa}-\vec{\kappa}'}^{(1)}}{\vec{\kappa}' \cdot \vec{\Upsilon}^{(0)}} \left(\frac{\partial F_{j,\vec{\kappa}'}^{(1)}}{\partial \tilde{P}_k} - \frac{F_{j,\vec{\kappa}'}^{(1)}}{\vec{\kappa}' \cdot \vec{\Upsilon}^{(0)}} \frac{\partial (\vec{\kappa}' \cdot \vec{\Upsilon}^{(0)})}{\partial \tilde{P}_k} \right) - \frac{\vec{\kappa}' \cdot \vec{f}_{\vec{\kappa}-\vec{\kappa}'}^{(1)}}{\vec{\kappa}' \cdot \vec{\Upsilon}^{(0)}} F_{j,\vec{\kappa}'}^{(1)} \right) \right) e^{i\kappa \cdot \vec{q}},
\end{aligned} \tag{4.46}$$

where $\{\cdot\}$ is used to denote the non-resonant oscillatory part of a product of functions. Thus we can remove the oscillatory part of $\tilde{F}_j^{(2)}$ by choosing

$$\begin{aligned}
\breve{Y}_{j,\vec{\kappa}}^{(2)} =& \frac{i}{\vec{\kappa} \cdot \vec{\Upsilon}^{(0)}} \left(F_{j,\vec{\kappa}}^{(2)} + \frac{\partial \left\langle Y_j^{(1)} \right\rangle}{\partial \tilde{P}_k} F_{k,\vec{\kappa}}^{(1)} - i \frac{\partial \left\langle F_j^{(1)} \right\rangle}{\partial \tilde{P}_k} \frac{F_{k,\vec{\kappa}}^{(1)}}{\vec{\kappa} \cdot \vec{\Upsilon}^{(0)}} \right. \\
&+ \left. \sum_{\vec{\kappa}' \in R} \left(i \frac{F_{k,\vec{\kappa}-\vec{\kappa}'}^{(1)}}{\vec{\kappa}' \cdot \vec{\Upsilon}^{(0)}} \left(\frac{\partial F_{j,\vec{\kappa}'}^{(1)}}{\partial \tilde{P}_k} - \frac{F_{j,\vec{\kappa}'}^{(1)}}{\vec{\kappa}' \cdot \vec{\Upsilon}^{(0)}} \frac{\partial (\vec{\kappa}' \cdot \vec{\Upsilon}^{(0)})}{\partial \tilde{P}_k} \right) - \frac{\vec{\kappa}' \cdot \vec{f}_{\vec{\kappa}-\vec{\kappa}'}^{(1)}}{\vec{\kappa}' \cdot \vec{\Upsilon}^{(0)}} F_{j,\vec{\kappa}'}^{(1)} \right) \right).
\end{aligned} \tag{4.47}$$

4.2.5 Freedom in the averaged pieces

With the non-resonant oscillatory pieces of the NIT equations of motion removed, terms in the equations of motion become

$$\tilde{F}_j^{(1)} = \left\langle F_j^{(1)} \right\rangle + \sum_{N \neq 0} F_{j,N\vec{\kappa}_{\text{res}}}^{(1)} e^{iNq_\perp}, \quad \Upsilon_i^{(1)} = \left\langle f_i^{(1)} \right\rangle + \sum_{N \neq 0} f_{i,N\vec{\kappa}_{\text{res}}}^{(1)} e^{iNq_\perp} - \frac{\partial \Upsilon_i^{(0)}}{\partial \tilde{P}_j} \left\langle Y_j^{(1)} \right\rangle, \tag{4.48a-b}$$

and

$$
\begin{aligned}
\tilde{F}_j^{(2)} = {}& \left\langle F_j^{(2)} \right\rangle + \sum_{N\neq 0} F^{(2)}_{j,N\vec{\kappa}_{\text{res}}} e^{iNq_\perp} + \left\langle \frac{\partial \breve{Y}_j^{(1)}}{\partial \tilde{q}_i} \breve{f}_i^{(1)} \right\rangle + \left\langle \frac{\partial \breve{Y}_j^{(1)}}{\partial \tilde{P}_k} F_k^{(1)} \right\rangle \\
&+ \frac{\partial \left\langle Y_j^{(1)} \right\rangle}{\partial \tilde{P}_k} \left\langle F_k^{(1)} \right\rangle - \frac{\partial \left\langle F_j^{(1)} \right\rangle}{\partial \tilde{P}_k} \left\langle Y_k^{(1)} \right\rangle
\end{aligned}
\tag{4.49}
$$

Note that we still have freedom to set the averaged pieces of the transformation functions $\left\langle Y_j^{(1)} \right\rangle$, $\left\langle Y_j^{(2)} \right\rangle$, $\left\langle X_i^{(1)} \right\rangle$, and $\left\langle W^{(1)} \right\rangle$ to be anything we choose. As before, we make the simplest choice: $\left\langle Y_j^{(1)} \right\rangle = \left\langle Y_j^{(2)} \right\rangle = \left\langle X_i^{(1)} \right\rangle = \left\langle W^{(1)} \right\rangle = 0$, as this makes it easy to compare between OG and NIT inspirals. It also has the benefit of drastically reducing the terms in our equations of motion to

$$
\tilde{F}_j^{(1)} = \left\langle F_j^{(1)} \right\rangle + \sum_{N\neq 0} F^{(1)}_{j,N\vec{\kappa}_{\text{res}}} e^{iNq_\perp}, \quad \Upsilon_i^{(1)} = \left\langle f_i^{(1)} \right\rangle + \sum_{N\neq 0} f^{(1)}_{i,N\vec{\kappa}_{\text{res}}} e^{iNq_\perp}, \tag{4.50a-b}
$$

and

$$
\tilde{F}_j^{(2)} = \left\langle F_j^{(2)} \right\rangle + \sum_{N\neq 0} F^{(2)}_{j,N\vec{\kappa}_{\text{res}}} e^{iNq_\perp} + \left\langle \frac{\partial \breve{Y}_j^{(1)}}{\partial \tilde{q}_i} \breve{f}_i^{(1)} \right\rangle + \left\langle \frac{\partial \breve{Y}_j^{(1)}}{\partial \tilde{P}_k} F_k^{(1)} \right\rangle \tag{4.51}
$$

4.2.6 Evolution of extrinsic quantities

The last thing to add this formulation is the evolution of the extrinsic quantities. Thankfully, both the t and ϕ geodesic equations are separable with respect to r and z and so

$$
f_k^{(0)} = \sum_{\vec{\kappa}\in\mathbb{Z}^N} f^{(0)}_{k,\vec{\kappa}} e^{i\vec{\kappa}\cdot\vec{q}} = \sum_N \left(f^{(0)}_{k,(N,0)} e^{iNqr} + f^{(0)}_{k,(0,N)} e^{iNqz} \right). \tag{4.52}
$$

This means that $f_k^{(0)}$ have no dependence on $q_\perp$ and so any term in the NIT transformations or equations of motion proportional to $1/\left(\kappa_{\text{res}} \cdot \Upsilon^{(0)}\right)$ will be multiplied by 0 and so all of our terms remain finite. As such, we can continue using the terms calculated in Sec. 4.1.6.

4.2.7 Summary of Near Resonant NIT

In summary, the equations of motion for the NIT variables $\tilde{P}_j, \tilde{q}_i$ and $\tilde{S}_k$ now take the form

$$
\frac{d\tilde{P}_j}{d\lambda} = \epsilon \tilde{F}_j^{(1)}(\vec{\tilde{P}}, \tilde{q}_\perp) + \epsilon^2 \tilde{F}_j^{(2)}(\vec{\tilde{P}}, \tilde{q}_\perp) + \mathcal{O}(\epsilon^3), \tag{4.53a}
$$

$$
\frac{d\tilde{q}_i}{d\lambda} = \Upsilon_i^{(0)}(\vec{\tilde{P}}) + \epsilon \Upsilon_i^{(1)}(\vec{\tilde{P}}, \tilde{q}_\perp) + \mathcal{O}(\epsilon^2), \tag{4.53b}
$$

$$
\frac{d\tilde{S}_k}{d\lambda} = \Upsilon_k^{(0)}(\vec{\tilde{P}}) + \epsilon \Upsilon_k^{(1)}(\vec{\tilde{P}}) + \mathcal{O}(\epsilon^2). \tag{4.53c}
$$

Crucially, these equations of motion only depend on the slowly evolving orbital elements $\vec{P}$ and the semi-fast resonant phase $q_\perp$ but not on any of the other rapidly oscillating orbital phases $\vec{q}$.

We chose the average pieces of the transformation terms to be $\left\langle Y_j^{(1)} \right\rangle = \left\langle Y_j^{(2)} \right\rangle = \left\langle X_i^{(1)} \right\rangle = \left\langle Z_k^{(0)} \right\rangle = \left\langle Z_k^{(1)} \right\rangle = \langle W^{(1)} \rangle = 0$ and so the transformed forcing functions are related to the original functions by

$$\tilde{F}_j^{(1)} = \sum_N F^{(1)}_{j,N\vec{\kappa}_{\text{res}}} e^{iNq_\perp}, \quad \Upsilon_i^{(1)} = \sum_N f^{(1)}_{i,N\vec{\kappa}_{\text{res}}} e^{iNq_\perp}, \quad \Upsilon_k^{(0)} = \left\langle f_k^{(0)} \right\rangle, \tag{4.54a-c}$$

$$\tilde{F}_j^{(2)} = \sum_N F^{(2)}_{j,N\vec{\kappa}_{\text{res}}} e^{iNq_\perp} + \left\langle \frac{\partial \breve{Y}_j^{(1)}}{\partial \tilde{q}_i} \breve{f}_i^{(1)} \right\rangle + \left\langle \frac{\partial \breve{Y}_j^{(1)}}{\partial \tilde{P}_k} \breve{F}_k^{(1)} \right\rangle, \tag{4.54d}$$

$$\Upsilon_k^{(1)} = -\left\langle \frac{\partial \breve{f}_k^{(0)}}{\partial \tilde{P}_j} \breve{Y}_j^{(1)} \right\rangle - \left\langle \frac{\partial \breve{f}_k^{(0)}}{\partial \tilde{q}_i} \breve{X}_i^{(1)} \right\rangle. \tag{4.54e}$$

In deriving these equations of motion, we have constrained the oscillating pieces of the NIT transformation functions to be

$$\breve{Y}_j^{(1)} := \sum_{\vec{\kappa} \in R} \frac{i}{\vec{\kappa} \cdot \vec{\Upsilon}} F^{(1)}_{j,\vec{\kappa}} e^{i\vec{\kappa}\cdot\vec{q}}, \tag{4.55}$$

$$\breve{X}_i^{(1)} := \sum_{\vec{\kappa} \in R} \left(\frac{i}{\vec{\kappa} \cdot \vec{\Upsilon}} f^{(1)}_{i,\vec{\kappa}} + \frac{1}{(\vec{\kappa} \cdot \vec{\Upsilon})^2} \frac{\partial \Upsilon_i}{\partial P_j} F^{(1)}_{j,\vec{\kappa}} \right) e^{i\vec{\kappa}\cdot\vec{q}}, \tag{4.56}$$

where R is the set $\{\vec{\kappa} \in \mathbb{Z}^2 | \vec{\kappa} \neq N\vec{\kappa}_{\text{res}}, \forall N \in \mathbb{Z}\}$ of all non-resonant 2-tuples and $\kappa_{\text{res}} = (\kappa_r, \kappa_z)$ is such that $\kappa_{\text{res}} \cdot \vec{\Upsilon}^{(0)} = 0$. After numerically solving the equations of motion, computing a waveform only requires knowledge of the transformations in Eq. (4.1) to zeroth order in the mass ratio, i.e.,

$$P_j = \tilde{P}_j + \mathcal{O}(\epsilon), \tag{4.57a}$$

$$q_i = \tilde{q}_i + \mathcal{O}(\epsilon), \tag{4.57b}$$

$$S_k = \tilde{S}_k - Z_k^{(0)}(\vec{\tilde{P}}, \vec{\tilde{q}}) + \mathcal{O}(\epsilon). \tag{4.57c}$$

where the zeroth order transformation term for the extrinsic quantities $\breve{Z}_k^{(0)}$ is unchanged in the presence of a resonance and is still given by

$$\breve{Z}_k^{(0)} = -\breve{S}_{k,r}(q_r) - \breve{S}_{k,z}(q_z). \tag{4.58}$$

Furthermore, to be able to directly compare between OG and NIT inspirals, we will need to match their initial conditions to sufficient accuracy. To maintain an overall phase difference of $\mathcal{O}(\epsilon)$ in the course of an inspiral, the initial values of the phases and extrinsic quantities need only be known to zeroth order in ϵ. However, we need to know the initial values of the orbital elements $\vec{P}$ to linear order in ϵ and so we use

$$\tilde{P}_j(0) = P_j(0) + \epsilon Y_j^{(1)}(\vec{P}(0), \vec{q}(0)) + \mathcal{O}(\epsilon^2). \tag{4.59}$$

4.3 Averaging transformations for motion parametrized by Boyer-Lindquist coordinate time

Solving the above equations will result in solutions for $\vec{\tilde{P}}$, $\vec{\tilde{q}}$ and $\vec{\tilde{S}}$ as functions of Mino-time λ. While this would include $t(\lambda)$, the transformation to $\lambda(t)$ is non-trivial, and in practice is done via interpolation which can be costly for long inspirals. It would be significantly more convenient for the solutions to be functions of t from the start so that one can produce time-domain waveforms for data analysis without this post-processing step. This can be accomplished for the OG equations by simply using the chain rule:

$$\frac{dP_j}{dt} = \frac{1}{f_t^{(0)}(\vec{P},\vec{q})}\left(\epsilon F_j^{(1)}(\vec{P})\right), \tag{4.60a}$$

$$\frac{dq_i}{dt} = \frac{1}{f_t^{(0)}(\vec{P},\vec{q})}\left(\Upsilon_i(\vec{P}) + \epsilon f_i^{(1)}(\vec{P},\vec{q})\right), \tag{4.60b}$$

$$\frac{d\phi}{dt} = \frac{1}{f_t^{(0)}(\vec{P},\vec{q})}\left(f_\phi^{(0)}(\vec{P},\vec{q})\right). \tag{4.60c}$$

Notice that we have one less equation of motion to solve. However, using the same approach to the NIT equations of motion results in

$$\frac{d\tilde{P}_j}{dt} = \frac{1}{f_t^{(0)}(\vec{P},\vec{q})}\left(\epsilon\tilde{F}_j^{(1)}(\vec{\tilde{P}}) + \epsilon^2\tilde{F}_j^{(2)}(\vec{\tilde{P}})\right), \tag{4.61a}$$

$$\frac{d\tilde{q}_i}{dt} = \frac{1}{f_t^{(0)}(\vec{P},\vec{q})}\left(\Upsilon_i(\vec{\tilde{P}}) + \epsilon\Upsilon_i^{(1)}(\vec{\tilde{P}})\right), \tag{4.61b}$$

$$\frac{d\tilde{\phi}}{dt} = \frac{1}{f_t^{(0)}(\vec{P},\vec{q})}\left(\Upsilon_\phi(\vec{\tilde{P}}) + \epsilon\Upsilon_\phi^{(1)}(\vec{\tilde{P}})\right). \tag{4.61c}$$

As we can see, we have now re-introduced a dependence on the orbital phases $\vec{q}$, defeating the purpose of our original NIT. Thankfully, as outlined in Ref. [113], these oscillations can also be averaged out by performing another transformation:

$$\mathcal{P}_j = \tilde{P}_j + \epsilon\Pi_j^{(1)}(\vec{\tilde{P}},\vec{\tilde{q}}) + \epsilon^2\Pi_j^{(2)}(\vec{\tilde{P}},\vec{\tilde{q}}) + \mathcal{O}(\epsilon^3), \tag{4.62a}$$

$$\varphi_\alpha = \tilde{Q}_\alpha + \Delta\varphi_\alpha + \epsilon\Phi_\alpha^{(1)}(\vec{\tilde{P}},\vec{\tilde{q}}) + \mathcal{O}(\epsilon^2), \tag{4.62b}$$

where $\vec{Q} = \{\vec{\tilde{q}}, \tilde{\phi}\}$, $\Delta\varphi_\alpha = \Omega_\alpha^{(0)}(\vec{\tilde{P}})\Delta t^{(0)}$ and $\Omega_\alpha^{(0)}$ is the Boyer-Lindquist fundamental frequency of the tangent geodesic.

To obtain the equations of motion for $\mathcal{P}_j$ and φ_α, we take the time derivative of Eq. (4.62), substitute the expression for the NIT equations of motion, and then use the inverse transformation of Eq. (4.62) to ensure that all functions are expressed in terms of $\vec{\mathcal{P}}$ and $\vec{q}$ and the expand order by order in ϵ. We then chose the oscillatory functions Δt, $\Phi_i^{(1)}$, $\Pi_j^{(1)}$ and $\Pi_j^{(2)}$ such that they cancel out any oscillatory terms that appear at each order ϵ. This results in averaged equations of motion that take the following form:

$$\frac{d\mathcal{P}_j}{dt} = \epsilon\Gamma_j^{(1)}(\vec{\mathcal{P}}) + \epsilon^2\Gamma_j^{(2)}(\vec{\mathcal{P}}) + \mathcal{O}(\epsilon^3), \tag{4.63a}$$

$$\frac{d\varphi_\alpha}{dt} = \Omega_\alpha^{(0)}(\vec{\mathcal{P}}) + \epsilon\Omega_\alpha^{(1)}(\vec{\mathcal{P}}) + \mathcal{O}(\epsilon^2). \tag{4.63b}$$

These equations of motion are related to the Mino time averaged equations of motion (4.27) with the adiabatic terms given by

$$\Gamma_j^{(1)} = \frac{\tilde{F}_j^{(0)}}{\Upsilon_t^{(0)}}, \quad \Omega_\alpha^{(0)} = \frac{\Upsilon_\alpha^{(0)}}{\Upsilon_t^{(0)}}, \tag{4.64a-b}$$

and the post-adiabatic terms given by

$$\Gamma_j^{(2)} = \frac{1}{\Upsilon_t}\Big(\tilde{F}_j^{(2)} + \tilde{F}_j^{(1)}\frac{\partial}{\partial \mathcal{P}_j}\left\langle \Pi_j^{(1)} \right\rangle - \left\langle f_t^{(0)}\Pi_k^{(1)} \right\rangle \frac{\partial \Gamma_j^{(1)}}{\partial \mathcal{P}_k} - \Upsilon_t^{(1)}\Gamma_j^{(1)}\Big), \tag{4.65a}$$

$$\Omega_\alpha^{(1)} = \frac{1}{\Upsilon_t^{(0)}}\Big(\Upsilon_\alpha^{(1)} + \tilde{F}_j^{(1)}\left\langle \frac{\partial \Delta\varphi_\alpha}{\partial \mathcal{P}_j} \right\rangle - \left\langle f_t^{(0)}\Pi_k^{(1)} \right\rangle \frac{\partial \Omega_\alpha^{(0)}}{\partial \mathcal{P}_k} - \Upsilon_t^{(1)}\Omega_i^{(1)}\Big). \tag{4.65b}$$

This constrains the oscillating pieces of our transformation to be

$$\Delta t = \sum_{\kappa\neq 0} \frac{f_{t,\vec{\kappa}}^{(0)}}{-i\vec{\kappa}\cdot\vec{\Upsilon}^{(0)}} = -\breve{Z}_t, \tag{4.66a}$$

$$\breve{\Pi}_j^{(1)} = \sum_{\kappa\neq 0} \frac{f_{t,\vec{\kappa}}^{(0)}}{-i\vec{\kappa}\cdot\vec{\Upsilon}^{(0)}}\Gamma_j^{(1)} = -\breve{Z}_t\Gamma_j^{(1)}, \quad \text{and} \tag{4.66b}$$

$$\begin{aligned}\Phi_{\alpha,\vec{\kappa}}^{(1)} =& \frac{i}{\vec{\kappa}\cdot\vec{\Upsilon}^{(0)}}\left(\frac{\partial \Delta\varphi_{\alpha,\vec{\kappa}}}{\partial \mathcal{P}_j}\tilde{F}_j^{(1)} - \frac{f_{t,\vec{\kappa}}^{(0)}}{\Upsilon_t^{(0)}}\Upsilon_t^{(1)}\Omega_t^{(0)}\right.\\ &\left. + \sum_{\vec{\kappa}'\neq\vec{0}}\left[\left(i\vec{\kappa}'\cdot\vec{X}_{\vec{\kappa}-\vec{\kappa}'}^{(1)}f_{t,\vec{\kappa}'}^{(0)} + Y_{j,\vec{\kappa}-\vec{\kappa}'}^{(1)}\frac{\partial f_{t,\vec{\kappa}'}^{(0)}}{\partial \mathcal{P}_j}\right)\Omega_\alpha^{(0)} - \Pi_{j,\vec{\kappa}-\vec{\kappa}'}^{(1)}f_{t,\vec{\kappa}'}^{(0)}\frac{\partial \Omega_\alpha^{(0)}}{\partial \mathcal{P}_j}\right]\right).\end{aligned} \tag{4.66c}$$

We are free to chose the averaged pieces of $\Pi_j^{(1)}$ and we make the simplification that $\left\langle \Pi_j^{(1)} \right\rangle = 0$. With this and the identity $\left\langle f_t^{(0)}(\int f_t^{(0)} d\vec{q}) \right\rangle = 0$, we get the simplification $\left\langle f_t^{(0)}\Pi_j^{(1)} \right\rangle = 0$. Thus the expressions for $\Gamma_j^{(2)}$ and $\Omega_\alpha^{(1)}$ simplify to

$$\Gamma_j^{(2)} = \frac{1}{\Upsilon_t^{(0)}}\Big(\tilde{F}_j^{(2)} - \Upsilon_t^{(1)}\Gamma_j^{(1)}\Big), \tag{4.67a}$$

$$\Omega_\alpha^{(1)} = \frac{1}{\Upsilon_t^{(0)}}\Big(\Upsilon_\alpha^{(1)} - \Upsilon_t^{(1)}\Omega_\alpha^{(0)}\Big). \tag{4.67b}$$

Alternatively, one could derive the the above equations of motion in a more straightforward manner by using the chain rule and expanding in powers of ϵ:

$$\begin{aligned}\frac{d\mathcal{P}_j}{dt} &= \frac{d\tilde{P}_j}{d\lambda}\left(\frac{d\tilde{t}}{d\lambda}\right)^{-1} + \mathcal{O}(\epsilon^3)\\ &= \frac{\epsilon\tilde{F}_j^{(1)} + \epsilon^2\tilde{F}_j^{(2)}}{\Upsilon_t^{(0)} + \epsilon\Upsilon_t^{(1)}} + \mathcal{O}(\epsilon^3)\\ &= \frac{\epsilon\tilde{F}_j^{(1)}}{\Upsilon_t^{(0)}} + \frac{\epsilon^2}{\Upsilon_t^{(0)}}\left(\tilde{F}_j^{(2)} - \frac{\Upsilon_t^{(1)}\tilde{F}_j^{(1)}}{\Upsilon_t^{(0)}}\right) + \mathcal{O}(\epsilon^3)\\ &= \epsilon\Gamma_j^{(1)} + \epsilon^2\Gamma_j^{(2)} + \mathcal{O}(\epsilon^3)\end{aligned} \tag{4.68a}$$

$$\begin{aligned}\frac{d\varphi_\alpha}{dt} &= \frac{d\tilde{q}_i}{d\lambda}\left(\frac{d\tilde{t}}{d\lambda}\right)^{-1} + \mathcal{O}(\epsilon^2)\\ &= \frac{\Upsilon_\alpha^{(0)} + \epsilon\Upsilon_\alpha^{(1)}}{\Upsilon_t^{(0)} + \epsilon\Upsilon_t^{(1)}} + \mathcal{O}(\epsilon^2)\\ &= \frac{\Upsilon_\alpha^{(0)}}{\Upsilon_t^{(0)}} + \frac{\epsilon}{\Upsilon_t^{(0)}}\left(\Upsilon_\alpha^{(1)} - \frac{\Upsilon_t^{(1)}\Upsilon_\alpha^{(0)}}{\Upsilon_t^{(0)}}\right) + \mathcal{O}(\epsilon^2)\\ &= \Omega_\alpha^{(0)} + \epsilon\Omega_\alpha^{(1)} + \mathcal{O}(\epsilon^2)\end{aligned} \tag{4.68b}$$

What is most useful about these equations of motion is that their solutions $\vec{\mathcal{P}}(t)$ and $\vec{\varphi}(t)$ is exactly what is required to feed into waveform generating schemes. We show the equivalence between $\vec{\varphi}(t)$ and the relationship derived in Ref. [156] between the solutions to the original NIT EOM and the waveform "voices" Φ_{mn} in Chapter. 5.

We note that in principle, one could also perform this transformation on the near-resonant NIT. However, the right hand side of equations of motion will still depend on the resonant phase $q_\perp = \kappa_r\tilde{q}_r + \kappa_z\tilde{q}_z$. While we have an analytic solution for $\varphi_\alpha(\vec{\tilde{q}})$ to leading order, one must numerically find the relationship $\tilde{q}_i(\vec{\varphi})$, making this approach impractical for inspiral calculations. This may be remedied in future if an analytic solution for $\tilde{q}_i(\vec{\varphi})$ is found.

In fact, this additional averaging step could be avoided entirely if we could describe the geodesic system directly in terms of action angles associated with Boyer-Lindquist frequencies, i.e. φ_α. This has been done in the Schwarzschild case as an analytic expansion in eccentricity [183]. However, an analytic expression for generic orbits in Kerr spacetime in terms of these has phases has yet to be found.

4.4 Two-timescale expansion

We note that there is an equivalent way of obtaining the above solutions via the two-timescale expansion (TTE). We exploit the difference between the timescales of the system by defining $\mathcal{T} := \epsilon t$ as the slow time which governs the long term, secular behaviour of the system and defining t as the fast time of the system that governs the short term, orbital dynamics and treating these two times independent variables. As such, we expand the transformed variables as

$$\mathcal{P}_j(\mathcal{T},\epsilon) = \mathcal{P}_j^{(0)}(\mathcal{T}) + \epsilon\mathcal{P}_j^{(1)}(\mathcal{T}) + \mathcal{O}(\epsilon^2), \tag{4.69a}$$

$$\varphi_\alpha(\mathcal{T},\epsilon) = \frac{1}{\epsilon}\left[\varphi_\alpha^{(0)}(\mathcal{T}) + \epsilon\varphi_\alpha^{(1)}(\mathcal{T})\right] + \mathcal{O}(\epsilon^1). \tag{4.69b}$$

Applying this expansion to the t parametrized NIT equations of motion, one finds that the equations of motion for the two-timescale expanded variables takes the form:

$$\frac{d\mathcal{P}_j^{(0)}}{d\mathcal{T}} = \Gamma_j^{(0)}(\vec{\mathcal{P}}^{(0)}), \tag{4.70a}$$

$$\frac{d\varphi_\alpha^{(0)}}{d\mathcal{T}} = \Omega_\alpha^{(0)}(\vec{\mathcal{P}}^{(0)}), \tag{4.70b}$$

$$\frac{d\mathcal{P}_j^{(1)}}{d\mathcal{T}} = \Gamma_j^{(2)}(\vec{\mathcal{P}}^{(0)}) + \mathcal{P}_k^{(1)}\left(\frac{\partial \Gamma_j^{(1)}}{\partial \mathcal{P}_k}(\vec{\mathcal{P}}^{(0)})\right), \tag{4.70c}$$

$$\frac{d\varphi_\alpha^{(1)}}{d\mathcal{T}} = \Omega_\alpha^{(1)}(\vec{\mathcal{P}}^{(0)}) + \mathcal{P}_k^{(1)}\left(\frac{\partial \Omega_\alpha^{(0)}}{\partial \mathcal{P}_k}(\vec{\mathcal{P}}^{(0)})\right). \tag{4.70d}$$

There is a trade-off for solving these equations of motion. We now have to solve a system of coupled differential equations that is twice the size and thus is more expensive to solve numerically, but the solutions are independent of ϵ and one can construct a solution that is valid for any value of ϵ using equations 4.69. Thus if one wants to compute multiple inspirals with varying mass ratios, the TTE can be more efficient overall. However, there is also an issue where the inspiral will stop prematurely, as the variables $\mathcal{P}_j^{(0)}$ typically reach values which correspond to the ISSO before $\mathcal{P}_j$ do. In this regime, one should instead employ a transition to plunge as the this is where the adiabaticity assumptions of the OG equations and two-timescale expansions is expected to break down. We will use both the NIT and the TTE of motion to produce waveforms and compare them to waveforms generated using the OG equations in Chapter 7 to assess which is the more practical framework for producing post-adiabatic EMRI waveforms.

Chapter 5

Waveforms

We wish to examine the effect of these averaging procedures on the final gravitational waveform. In order to do this, we first must establish a relationship between the solutions for the inspiral quantities and the waveform. We first examine how to do this with fully relativistic waveforms generated from solutions to the Teukolsky equation, as implemented in `FastEMRIWaveform` package [120, 135]. Unfortunately, this implementation is not currently available for Kerr inspirals, and so we use "semi-relativistic" quadrupole formula used by the numerical kludge models, which we discuss in depth. While such an approximation may not be accurate enough for LISA data science, such waveforms still fair surprisingly well compared to Teukolsky snapshot waveforms, even in the strong field regime [116]. What is most important for this work, where we are testing different calculations of inspiral trajectories, is that all waveforms are calculated using the same formula so that any difference in the waveforms is due to differences in the inspiral trajectories. We then end this chapter with a brief discussion of the waveform analysis employed in this work.

5.1 Waveform Generation

5.1.1 Teukolsky Based Waveforms

We start by assigning the origin of our coordinate scheme to be the position of the primary MBH. We then set the Cartesian coordinates of our observer to be $x = (t_{\text{ret}}, \vec{x})$. It is useful to express the observer's coordinates in terms of spherical coordinates (R, θ, Φ), where R is the distance from the observer to origin given by $R^2 = \vec{x} \cdot \vec{x}$, Θ is the observer's latitude and Φ is the observer's azimuth. We assign the Cartesian coordinates of the secondary, which we model as a point particle, to be $x_p = (t_p, \vec{x}_p)$. As such, we can express the retarded time as measured by the observer in terms of Boyer-Lindquist coordinate time t via $t_{\text{ret}} = t - |\vec{x} - \vec{x}_p| \approx t - R$ in the limit where R is large. The complex waveform strain can be decomposed onto a basis of spin-weighted spherical harmonics ${}_{-2}Y_{lm}$ given by:

$$h(t_{\text{ret}}) = h_+ - ih_\times = \frac{1}{R} \sum_{lm} H_{lm}(t_{\text{ret}})_{-2}Y_{lm}(\Theta, \Phi). \tag{5.1}$$

If we assume the secondary is moving on a geodesic, the waveform modes H_{lm} exhibit a discrete frequency spectrum and so can be Fourier decomposed into:

$$H_{lm}(t_{\text{ret}}) = \sum_{\vec{\kappa}} A_{lm\vec{\kappa}}(\vec{P}) e^{-i(\vec{\kappa} \cdot \vec{\Omega}^{(0)}) t_{\text{ret}}} \tag{5.2}$$

where the complex amplitudes $A_{lm\vec{\kappa}}$ can be related to non-homogeneous Teukolsky amplitudes, which can be precomputed for a given set of orbital elements $\vec{P}$ [134]. Such waveforms are known as "snapshot" waveforms, as they only capture a small section of the total waveform [17].

For a full EMRI waveform, one needs to account for the fact that the frequencies and the orbital elements will slowly evolve with time, resulting in a continuous frequency spectrum. As such, this Fourier mode decomposition becomes a "mulit-voice" decomposition [134]:

$$H_{lm}(t_{\text{ret}}) = \sum_{\vec{\kappa}} A_{lm\vec{\kappa}}(\vec{P}(t_{\text{ret}}))e^{-i\varPhi_{m\vec{\kappa}}(t_{\text{ret}})}, \tag{5.3}$$

where the waveform "voices" $\varPhi_{m\vec{\kappa}}$ [1] are given by

$$\varPhi_{m\vec{\kappa}} = m\int_0^t \Omega_\phi(t)dt + \kappa_r\int_0^t \Omega_r(t)dt + \kappa_z\int_0^t \Omega_z(t)dt. \tag{5.4}$$

Let us recall the equations of motion for the orbital phases obtains after performing the NIT and the additional transformation such that our solutions are in terms of Boyer-Linquist time (4.63). When we express this in integral form, one obtains

$$\varphi_\alpha = \int_0^t \left(\Omega_\alpha^{(0)}(t) + \epsilon\Omega_\alpha^{(1)}(t) + \mathcal{O}(\epsilon^2)\right)dt = \int_0^t \left(\Omega_\alpha(t) + \mathcal{O}(\epsilon^2)\right)dt \tag{5.5}$$

As such, we can finally assign a physical meaning to these phases, as at leading order in ϵ, they are directly related to the waveform voices:

$$\varPhi_{m,\vec{\kappa}} = m\varphi_\phi + \kappa_r\varphi_r + \kappa_z\varphi_z + \mathcal{O}(\epsilon) \tag{5.6}$$

This is further supported by Ref. [156], where a relationship between the waveform phases and the NIT phases $\vec{\tilde{q}}$ was independently derived. When expressed in our notation with polar motion included, this relationship is given by:

$$\varPhi_{m\vec{\kappa}} = m\tilde{\phi} + \kappa_r\tilde{q}_r + \kappa_z\tilde{q}_z + (m\Omega_\phi^{(0)} + \kappa_r\Omega_r^{(0)} + \kappa_z\Omega_z^{(0)})(t-\tilde{t}) + \mathcal{O}(\epsilon) \tag{5.7}$$

Using the relationship between t and $\tilde{t}$ given by Eq. (4.1c), the result for Δt given by Eq. (4.66a), and the relationship between the NIT action angles given by Eq. (4.62), one can obtain:

$$\begin{aligned}
\varPhi_{m\vec{\kappa}} &= m(\tilde{\phi} - \Omega_\phi(t-\tilde{t})) + \kappa_r(\tilde{q}_r - \Omega_r(t-\tilde{t})) + \kappa_z(\tilde{q}_z - \Omega_z(t-\tilde{t})) + \mathcal{O}(\epsilon)\\
&= m(\tilde{\phi} - \Omega_\phi Z_t^{(0)}) + \kappa_r(\tilde{q}_r - \Omega_r Z_t^{(0)}) + \kappa_z(\tilde{q}_z - \Omega_z Z_t^{(0)}) + \mathcal{O}(\epsilon)\\
&= m(\tilde{\phi} + \Omega_\phi\Delta t) + \kappa_r(\tilde{q}_r + \Omega_r\Delta t) + \kappa_z(\tilde{q}_z + \Omega_z\Delta t) + \mathcal{O}(\epsilon)\\
&= m(\tilde{\phi} + \Delta\varphi_\phi) + \kappa_r(\tilde{q}_r + \Delta\varphi_r) + \kappa_z(\tilde{q}_z + \Delta\varphi_z) + \mathcal{O}(\epsilon)\\
&= m\varphi_\phi + \kappa_r\varphi_r + \kappa_z\varphi_z + \mathcal{O}(\epsilon).
\end{aligned} \tag{5.8}$$

One can also freely replace the dependence of the evolving orbital elements $\vec{P}(t)$ with either $\vec{\tilde{P}}(t)$ or $\vec{\mathcal{P}}(t)$, as at leading order in ϵ they are identical. As a result, we can now relate a solution to the EMRI's inspiral trajectory to its associated Teukolsky based waveform.

The biggest obstacle to overcome when producing these waveforms is pre-computing the amplitudes $A_{lm\vec{\kappa}}$ across the four dimensional generic Kerr parameter space. Moreover, once those are known, each time one wishes to sample the waveform, one has to sum over a large number of both Fourier and spherical harmonic modes. This can get very expensive when one has to sample a multi-year long waveform every few seconds i.e. $\mathcal{O}(10^6)$ times. The `FastEMRIWaveforms` package overcomes these challenges by using a neural net to interpolate the amplitudes and GPUs to speed up the mode summation process [120]. However, thus far this package can only compute fully relativistic waveforms for eccentric Schwarzschild EMRIs. As such, we make use of this package when we dealing with the Schwarzschild case and make use of a semi-relativistic waveform scheme when we require waveforms in Kerr spacetime.

---

[1] These are not to be confused with the waveform's phase as the coefficients $A_{lm\vec{\kappa}}$ are complex and thus make up part of the overall waveform phase.

5.1.2 Semi-Relativistic Quadrupole Waveforms

Since we make heavy use of the flat space-quadrupole waveform generation scheme, we will discuss it in depth and follow the derivation given in Ref. [116]. We start by making the approximation that we are in a weak field situation and so we can describe the spacetime as $g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}$ where $\eta_{\mu\nu}$ is the flat space (Minkowski) metric and $h_{\mu\nu}$ is a perturbation. The trace reversed metric perturbation is defined as $\bar{h}^{\mu\nu} = h^{\mu\nu} - \frac{1}{2}\eta^{\mu\nu}h$ where h is the scalar one gets by contracting the perturbation with the flat space metric $h = \eta^{\mu\nu}h_{\mu\nu}$. We then impose the Lorenz gauge condition on the metric perturbation i.e. $\bar{h}^{\mu\alpha}_{;\alpha} = 0$ where semi-colon subscript denotes covariant derivative i.e. $f_{;\alpha} = \nabla_\alpha f$. This allows us to write the the linearized Einstein field equations as

$$\Box \bar{h}^{\mu\nu} = -16\pi \mathcal{T}^{\mu\nu} \tag{5.9}$$

where $\Box$ is the usual flat space wave operator and the effective energy momentum tensor $\mathcal{T}^{\mu\nu}$ satisfies:

$$\mathcal{T}^{\mu\nu}{}_{,\nu} = 0, \tag{5.10}$$

where a comma subscript denotes a partial derivative. The wave-equation (5.9) has the retarded time solution:

$$\bar{h}^{ij}(t_{\text{ret}}, \vec{x}) = 4\int \frac{\mathcal{T}^{ij}(t - |\vec{x} - \vec{x}'|, \vec{x}')}{|\vec{x} - \vec{x}'|} d^3x', \tag{5.11}$$

where the coordinate $\vec{x}'$ is the integration variable which goes over all space locations where the effective stress energy tensor is non-zero. By repeatedly substituting (5.10) into (5.11), dropping terms that are sub-leading in $1/R$ and assuming that the stress energy tensor either has compact support or falls off as $R \to \infty$ fast enough to neglect the boundary terms, one can derive the Press formula which is valid for extended and fast moving sources:

$$\bar{h}^{ij}(t_{\text{ret}}, \vec{x}) = \frac{2}{R}\frac{d^2}{dt^2}\int \left[\left(\mathcal{T}^{00} - 2\mathcal{T}^{0l}n_l + \mathcal{T}^{lm}n_l n_m\right) x'^i x'^j\right]_{t'=t_{\text{ret}}} d^3x', \tag{5.12}$$

where $\vec{n}$ is the unit vector given by $\vec{n} = \vec{x}/R$ [184]. If the source motion is only negligibly influenced by gravity, we can approximate the effective stress energy with the energy-momentum tensor of the matter source $T_{\mu\nu}$. Combining this with the slow motion limit of the Press formula recovers the quadrupole formula:

$$\bar{h}^{ij}(t, \vec{x}) = \frac{2}{R}\frac{d^2 I^{ij}}{dt^2}, \tag{5.13}$$

where I^{ij} is the source's mass quadrupole moment which is given by

$$I^{ij}(t') = \int T^{00} x'^i x'^j d^3x'. \tag{5.14}$$

We now specialize to the case of a point-particle of mass μ moving along a worldline $x_p(\tau)$, which means that the energy momentum tensor is given by:

$$\begin{aligned} T^{\mu\nu}(t', \vec{x}') &= \mu \int_{-\infty}^{\infty} \frac{dx_p'^{\mu}}{d\tau}\frac{dx_p'^{\nu}}{d\tau}\delta^4(x' - x_p'(\tau))d\tau, \\ &= \mu \frac{d\tau}{dt_p'}\frac{dx_p'^{\mu}}{d\tau}\frac{dx_p'^{\nu}}{d\tau}\delta^3(\vec{x}' - \vec{x}_p'(t')) \end{aligned} \tag{5.15}$$

where τ is the particle's proper-time and is related to the coordinate time of the particle by $dt_p/d\tau = 1 + \mathcal{O}(v^2/c^2)$, where the speed v is given by $v^2 = |d\vec{x}'_p/dt'_p|^2$. In the slow motion approximation, we assume that $v^2 << c^2$ and so we can express the mass quadrupole moment as

$$\begin{aligned} I^{ij}(t') = \int T^{00} x'^i x'^j d^3x' &= \int \mu \frac{d\tau}{dt'_p}\frac{dt'_p}{d\tau}\frac{dt'_p}{d\tau} x'^i x'^j \delta^3(\vec{x}' - \vec{x}'_p(t')) d^3x' \\ &\approx \mu \int x'^i x'^j \delta^3(\vec{x}' - \vec{x}'_p(t')) d^3x' = \mu x'^i x'^j \end{aligned} \tag{5.16}$$

We then make the semi-relativistic approximation that the flat-space Cartesian coordinates of the particle $\vec{x}_p$ can be mapped onto the Boyer-Lindquist coordinates of Kerr spacetime via:

$$x_p^1 := r \sin\theta \cos\phi, \quad x_p^2 := r \sin\theta \sin\phi, \quad x_p^3 := r \cos\theta \tag{5.17}$$

where r is given by the radial geodesic solution (2.24a), θ is given by $\theta = \arccos(z)$ where z is polar geodesic solution (2.24b) and ϕ is given by the solution for azimuthal coordinate that one obtains for the inspiral trajectory.

Moreover, in equation (5.13), we use the chain rule to replace the derivatives with respect to coordinate time t with derivatives with respect to Mino-time λ as we have analytic expressions for the Mino-time derivatives of the coordinates (assuming that the trajecotry at a given instant is on a tangent geodesic). We also neglect the rates of change of the orbital elements as these are of order $\mathcal{O}(\epsilon)$. As such, we can express the Mino-time derivatives of the radial coordinate as

$$\frac{dr}{d\lambda} = \frac{\partial r}{\partial q_r}\Upsilon_r^{(0)} + \mathcal{O}(\epsilon), \quad \frac{d^2 r}{d\lambda^2} = \frac{\partial^2 r}{\partial q_r^2}\left(\Upsilon_r^{(0)}\right)^2 + \mathcal{O}(\epsilon). \tag{5.18}$$

Similarly, the Mino-time derivatives of the polar coordinates can be expressed as

$$\frac{d\theta}{d\lambda} = \frac{-1}{\sqrt{1-z^2}}\frac{\partial z}{\partial q_z}\Upsilon_z^{(0)} + \mathcal{O}(\epsilon), \quad \frac{d^2\theta}{d\lambda^2} = \frac{-1}{\sqrt{1-z^2}}\left(\frac{\partial^2 z}{\partial q_z^2} + \frac{\partial z}{\partial q_z}\frac{z}{1-z^2}\right)\left(\Upsilon_z^{(0)}\right)^2 + \mathcal{O}(\epsilon). \tag{5.19}$$

We also need to know the Mino-time derivatives of the rates of change the extrinsic quantities t and ϕ. The first derivative of these with respect to Mino time are given by the geodesic expressions for $f_t^{(0)}$ in equation (2.7c) and $f_\phi^{(0)}$ in equation (2.7d) respectively. To leading order in ϵ, their second derivatives can be expressed using the chain rule as:

$$\frac{d^2 t}{d\lambda^2} = \frac{\partial f_t^{(0)}}{\partial r}\frac{dr}{d\lambda} + \frac{\partial f_t^{(0)}}{\partial z}\frac{dz}{d\lambda} + \mathcal{O}(\epsilon), \quad \frac{d^2\phi}{d\lambda^2} = \frac{\partial f_\phi^{(0)}}{\partial r}\frac{dr}{d\lambda} + \frac{\partial f_\phi^{(0)}}{\partial z}\frac{dz}{d\lambda} + \mathcal{O}(\epsilon), \tag{5.20}$$

From this, we express the first derivative of the particle's flat space Cartesian coordinates with respect to Mino time as:

$$\dot{x}_p^1 = \dot{r}\cos\phi\sin\theta + r\left(\dot{\theta}\cos\theta\cos\phi - \dot{\phi}\sin\theta\sin\phi\right), \tag{5.21a}$$

$$\dot{x}_p^2 = \dot{r}\sin\phi\sin\theta + r\left(\dot{\theta}\cos\theta\sin\phi - \dot{\phi}\sin\theta\cos\phi\right), \tag{5.21b}$$

$$\dot{x}_p^3 = \dot{r}\cos\theta - r\dot{\theta}\sin\theta, \tag{5.21c}$$

where we have used overdots to denote derivatives with respect to λ. Their second derivatives are expressed as:

$$\begin{aligned}\ddot{x}_p^1 =& \ddot{r}\sin\theta\cos\phi + 2\dot{r}\left(\dot{\theta}\cos\theta\cos\phi - \dot{\phi}\sin\theta\sin\phi\right), \\ & - r\left(\left(\ddot{\theta}^2 + \dot{\phi}^2\right)\sin\theta\cos\phi + 2\dot{\theta}\dot{\phi}\cos\theta\sin\phi - \ddot{\theta}\cos\theta\cos\phi + \ddot{\phi}\sin\theta\sin\phi\right),\end{aligned} \tag{5.22a}$$

$$\begin{aligned}\ddot{x}_p^2 =& \ddot{r}\sin\theta\sin\phi + 2\dot{r}\left(\dot{\theta}\cos\theta\sin\phi + \dot{\phi}\sin\theta\cos\phi\right) \\ & + r\left(-\left(\dot{\theta}^2 + \dot{\phi}^2\right)\sin\theta\sin\phi + 2\dot{\theta}\dot{\phi}\cos\theta\cos\phi - \ddot{\theta}\cos\theta\sin\phi + \ddot{\phi}\sin\theta\cos\phi\right),\end{aligned} \tag{5.22b}$$

$$\ddot{x}_p^3 = \ddot{r}\cos\theta - 2\dot{r}\sin\theta - r\left(\dot{\theta}^2\cos\theta + \ddot{\theta}\sin\theta\right). \tag{5.22c}$$

The components of the metric perturbation $\bar{h}^{ij}$ can then be expressed in terms of these derivatives via:

$$\bar{h}^{ij} = \frac{2}{\dot{t}^2}\left(\frac{\ddot{t}}{\dot{t}}\left(\dot{x}_p^i x_p^j + x_p^i \dot{x}_p^j\right) + \ddot{x}_p^i x_p^j + x_p^i \ddot{x}_p^j + 2\dot{x}_p^i \dot{x}_p^j\right) \tag{5.23}$$

As such, we end up with a final expression for $\bar{h}^{ij}$ at a given instant t which can be expressed analytically so long as one supplies the solutions for $(p(t), e(t), x(t), q_r(t), q_z(t), \phi(t))$.

Since we assume that the observer is a large distance from the source, we are only concerned with the transverse and traceless (TT) parts of $\bar{h}^{ij}$. As such we perform a TT projection of the quadrupole formula onto an orthonormal spherical coordinate system given by:

$$\hat{e}_R = \frac{\partial}{\partial R}, \quad \hat{e}_\Theta = \frac{1}{R}\frac{\partial}{\partial\Theta}, \quad \hat{e}_\Phi = \frac{1}{R\sin\Theta}\frac{\partial}{\partial\Phi} \tag{5.24}$$

Thus the waveform in transverse traceless gauge is given by

$$\bar{h}_{TT}^{i,j} = \frac{1}{2}\begin{pmatrix} 0 & 0 & 0 \\ 0 & h^{\Theta\Theta} - h^{\Phi\Phi} & h^{\Phi\Theta} \\ 0 & h^{\Phi\Theta} & h^{\Theta\Theta} - h^{\Phi\Phi} \end{pmatrix}, \tag{5.25}$$

where

$$h^{\Theta\Theta} = \cos^2\Theta\left(\bar{h}^{11}\cos^2\Phi + \bar{h}^{12}\sin 2\Phi + \bar{h}^{22}\sin^2\Phi\right) + \bar{h}^{33}\sin^2\Theta - \sin 2\Theta\left(\bar{h}^{13}\cos\Phi + \bar{h}^{23}\sin\Phi\right), \tag{5.26a}$$

$$h^{\Phi\Theta} = \frac{1}{2}\cos\Theta\left(-\bar{h}^{11}\sin 2\Phi + 2\bar{h}^{12}\cos 2\Phi + \bar{h}^{22}\sin 2\Phi\right) + \sin 2\Theta\left(\bar{h}^{13}\sin\Phi - \bar{h}^{23}\cos\Phi\right), \tag{5.26b}$$

$$h^{\Phi\Phi} = \bar{h}^{11}\sin 2\Phi - \bar{h}^{12}\sin 2\Phi + \bar{h}^{22}\cos 2\Phi. \tag{5.26c}$$

We can then recover the plus and cross polarizations by noting that $h_+ = h^{\Theta\Theta} - h^{\Phi\Phi}$ and $h_\times = h^{\Phi\Theta}$, and which allows us to express the complex waveform strain as

$$h = h_+ - ih_\times. \tag{5.27}$$

5.2 Waveform Analysis

Now that we can calculate waveforms from our inspiral trajectories, we need to be able to quantitatively compare them. Note that for this subsection, we make use of SI units as opposed to geometrized units.

We imagine that we wish to compare two time-domain waveforms h_1 and h_2 which differ by an amount $\delta h = h_2 - h_1$. The noise weighted inner product of these two waveforms is given by

$$\langle h_1 | h_2 \rangle = 2 \int_0^\infty \frac{\tilde{h}_1^*(f)\tilde{h}_2(f) + \tilde{h}_1(f)\tilde{h}_2^*(f)}{S_n(f)} df, \tag{5.28}$$

where $\tilde{h}(f)$ denotes the Fourier transform of the time domain waveform $h(t)$ and h^* is the complex conjugate of h. For our analysis is Chapter 6, we take the power spectral density (PSD) of the detector noise $S_n(f)$ to be a flat noise curve. In subsequent chapters, we instead use an approximate LISA sensitivity curve derived in Ref. [185] which is given by:

$$S_n(f) = \frac{10}{3L^2}\left(P_{\mathsf{OMS}}(f) + \frac{4P_{\mathsf{acc}}(f)}{(2\pi f)^4}\right)\left(1 + \frac{6}{10}\left(\frac{f}{f_*}\right)^2\right), \tag{5.29}$$

where $L = 2.5 \times 10^6$m, $f_* = c/(2\pi L) = 19.09$ mHz. The single-link optical metrology noise P_{OMS} is taken to be $P_{\mathsf{OMS}}(f) = (1.5 \times 10^{-11}\text{m})^2\text{Hz}^{-1}$ and the single test mass acceleration noise P_{acc} is taken to be $P_{\mathsf{OMS}}(f) = (3 \times 10^{-15}\text{m s}^{-2})^2 \left(1 + \left(\frac{0.4\text{mHz}}{f}\right)^2\right)\text{Hz}^{-1}$. On top of instrument noise, unresolved galactic binaries will also be a source of effective noise. As such, we take the galactic confusion noise S_c to be

$$S_c(f) = A f^{-7/3} e^{-f^{C_1} + C_2 \sin(C_3 f)} \left(1 + \tanh(C_4(f_k - f))\right) \text{Hz}^{-1} \tag{5.30}$$

with the parameters $A = 9 \times 10^{-45}$, $C_1 = 1.38$, $C_2 = -221$ $C_3 = 521$, $C_4 = 1680$ and $f_k = 1.13 \times 10^{-3}$ such that the confusion noise is consistent with a 4 year observation time. From this, one can define the singal to noise ratio (SNR) ρ to be:

$$\rho^2 = \langle h | h \rangle \tag{5.31}$$

Similarly, the fractional waveform overlap $\mathcal{O}$ can be calculated as:

$$\mathcal{O} = \frac{\langle h_1 | h_2 \rangle}{\sqrt{\langle h_1 | h_1 \rangle \langle h_2 | h_2 \rangle}} = \frac{\langle h_1 | h_2 \rangle}{\rho_1 \rho_2}. \tag{5.32}$$

This overlap ranges from 1, when the two waveforms are identical, to 0, when the waveforms are perfectly orthogonal. When dealing with waveform overlaps close to 1, it is often useful to talk in terms of the fractional waveform mismatch $\mathcal{M}$, which is simply $\mathcal{M} = 1 - \mathcal{O}$. In practice, we make use of the `WaveformMatch` function from the SimulationTools package to calculate waveform overlaps [186].

Now that we can compare waveforms, it is worth asking how high the fractional overlap should be for LISA data science. One benchmark that is often quoted is a fractional overlap of $\mathcal{O} = 0.97$. This was first derived in Ref. [187] and corresponds to the minimum overlap needed to construct a template bank that corresponds to a 90% -ideal observed event rate.

While we do reference this benchmark on occasion in this work, we prefer to make sense of the mismatch in a different way. As outlined in Ref. [188], two waveforms are said to be "indistinguishable" if the inner product of the waveform difference satisfies:

$$\langle \delta h | \delta h \rangle < 1. \tag{5.33}$$

Following the analysis in Ref. [189], we relate this condition to the waveform mismatch:

$$\begin{aligned}
\langle \delta h | \delta h \rangle &= \langle h_1 - h_2 | h_1 - h_2 \rangle \\
&= \langle h_1 | h_1 \rangle + \langle h_2 | h_2 \rangle - 2\langle h_1 | h_2 \rangle \\
&= \rho_1^2 + \rho_2^2 - 2\rho_1\rho_2 + 2\rho_1\rho_2 - 2\langle h_1 | h_2 \rangle \\
&= (\rho_1 - \rho_2)^2 + 2\rho_1\rho_2 \left(1 - \frac{\langle h_1 | h_2 \rangle}{\rho_1\rho_2}\right) \\
&= (\rho_1 - \rho_2)^2 - 2\rho_1\rho_2\mathcal{M}
\end{aligned} \tag{5.34}$$

In the limit where $\rho_1 \approx \rho_2$, this simplifies further to

$$\langle \delta h | \delta h \rangle \approx 2\rho_1\rho_2\mathcal{M}. \tag{5.35}$$

As such, for a given mismatch $\mathcal{M}$, the two waveforms are said to be indistinguishable for SNRs

$$\rho \lessapprox \frac{1}{\sqrt{2\mathcal{M}}}. \tag{5.36}$$

Since the accumulated SNR for EMRIs is only expected to be as large as $\rho \sim 50$ [28], waveforms with $\mathcal{M} \lessapprox 2 \times 10^{-4}$ will be indistinguishable for most EMRI signals.

Chapter 6

Eccentric self-forced inspirals into a rotating black hole

In this chapter, we specialize our formulation to eccentric orbits that lie in the equatorial plane of the primary. We show how this affects the OG equations of motion in Sec. 6.1. We then introduce our interpolated model for the gravitational self force in Sec. 6.2 , which combined with the OG equations of motion allows us calculate the very first self-forced inspirals in Kerr spacetime. To speed up these inspiral calculations, we explicitly derive the averaged equations of motion for the case of eccentric Kerr inspirals in Sec. 6.3. We describe our numerical implementation for calculating the various terms in the NIT equations of motion in Sec. 6.4 and present the results of this implementation in Sec. 6.5.

6.1 Osculating Geodesics

We now look to specialize the osculating geodesic equations of motion to the special case of eccentric inspirals in the equatorial plane under the influence of the first-order ratio gravitational self force (GSF). This corresponds to setting $x = \pm 1$ for prograde and retrograde orbits, respectively. Due to symmetry, motion in the equatorial plane will stay in the equatorial plane, and thus $\dot{x} = 0$. As such, we only need to track the evolution of $\vec{P} = \{p, e\}$. Similarly, the equations of motion no longer depend on the polar phase q_z, and so we only need to evolve the radial phase $\vec{q} = \{q_r\}$. As such, the system of equations simplifies to

$$\frac{dp}{d\lambda} = \epsilon F_p^{(1)}(a, p, e, q_r) + \epsilon^2 F_p^{(2)}(a, p, e, q_r), \tag{6.1a}$$

$$\frac{de}{d\lambda} = \epsilon F_e^{(1)}(a, p, e, q_r) + \epsilon^2 F_e^{(2)}(a, p, e, q_r), \tag{6.1b}$$

$$\frac{dq_r}{d\lambda} = \Upsilon_r^{(0)}(a, p, e) + \epsilon f_r^{(1)}(a, p, e, q_r), \tag{6.1c}$$

$$\frac{dt}{d\lambda} = f_t^{(0)}(a, p, e, q_r), \tag{6.1d}$$

$$\frac{d\phi}{d\lambda} = f_\phi^{(0)}(a, p, e, q_r). \tag{6.1e}$$

Moreover, since we only have access to first order in the mass ratio GSF results, we set $F_p^{(2)} = F_e^{(2)} = 0$.

6.2 Interpolated GSF model

Driving inspirals requires a model for the self-force that can be rapidly evaluated at each instant during the inspiral. To achieve this we tile the parameter space with GSF data which we can then interpolate. This has been done in two different ways for eccentric Schwarzschild inspirals [70, 145]. However, both of these works required computing the self-force at tens of thousands of points in the parameter space. While this might not pose much of a problem for the 2D parameter space of eccentric, Schwarzschild inspirals, these approaches would not scale well to the 4D parameter space for generic Kerr inspirals. Motivated by this, as well as the computational expense of the eccentric Kerr self-force code, we build an interpolation model based on Chebyshev polynomials that is accurate to percent level across a 2D slice of the EMRI parameter space using only a few hundred points.

We start by fixing the value of the spin parameter of the primary, which we choose to be $a = 0.9M$ for Kerr inspirals or $a = 0$ for Schwarzschild inspirals and set the inclination x to be either 1 or -1 for prograde orbits or retrograde orbits respectively. This reduces our parameter space to two parameters; the semilatus rectum p and the eccentricity e. We then define a parameter y_{eq} using the p and the position of the last stable equatorial orbit p_{LSEO}. For Kerr orbits, we chose y to be

$$y_{\mathsf{Kerr}} = \sqrt{\frac{p_{\mathsf{LSEO}}(a, e, x)}{p}}. \tag{6.2}$$

With this parametrization we found that the accuracy of the Chebyshev interpolation is limited by the appearance of cusps at the LSEO in the data. To ameliorate their impact we instead used a parameter y_{eq} given by

$$y_{\mathsf{Schwarz}} = 1 - \left(1 - \frac{p_{\mathsf{LSEO}}(0, e, x)}{p}\right)^{1/3} \tag{6.3}$$

for later runs in Schwarzschild spacetime. In either case, tiling the parameter space in y instead of p will concentrate more points near the separatrix where the self force varies the most.

We let y range from $y_{\mathsf{min}} = 0$ $(0.01$ for Schwarzschild$)$ to $y_{\mathsf{max}} = 1$ and e range from $e_{\mathsf{min}} = 0$ to $e_{\mathsf{max}} = 0.5$ for Kerr and $e_{\mathsf{min}} = 0$ to $e_{\mathsf{max}} = 0.3$ for Schwarzschild. We define parameters u and v which cover this parameter space as they range from $(-1, 1)$

$$u := \frac{y_{eq} - (y_{\mathsf{min}} + y_{\mathsf{max}})/2}{(y_{\mathsf{min}} - y_{\mathsf{max}})/2} \quad \text{and} \quad v := \frac{e - (e_{\mathsf{min}} + e_{\mathsf{max}})/2}{(e_{\mathsf{min}} - e_{\mathsf{max}})/2}, \tag{6.4a-b}$$

This parametrization is convenient when using Chebyshev polynomials of the first kind, where the order n polynomial is defined by $T_n(\cos\varphi) := \cos(n\varphi)$. The Chebyshev nodes are the roots these polynomials, and the location of the kth root of nth polynomial is given by

$$N_k = \cos\left(\frac{2k-1}{2n}\pi\right) \tag{6.5}$$

We then calculate the GSF on a 15×7 grid of Chebyshev nodes, with the u values given by the roots of the 15th order polynomial and the v values given by the roots of the 7th order polynomial. At each point on our grid, we Fourier decompose each component of the force with respect to the radial action angle q_r. We then multiply the data for each Fourier coefficient by a factor of $(1 - y_{eq})/(1 - e^2)$, as we find empirically that this smooths the behaviour of the force near the separatrix and improves the accuracy of our interpolation. Next, we use Chebyshev polynomials to interpolate each Fourier coefficient across the (u, v) grid. We then resum the modes to reconstruct our interpolated gravitational self force model:

$$a_\alpha = \frac{1-e^2}{1-y_{eq}} \sum_{n=0}^{15} A^n_\alpha(y,e)\cos(nq_r) + B^\kappa_\alpha(y,e)\sin(nq_r), \tag{6.6}$$

where

$$A^n_\alpha(y,e) = \sum_{i=0}^{14}\sum_{j=0}^{6} A^{nij}_\alpha T_i\,(u)\,T_j\,(v) \quad \text{and} \quad B^n_\alpha(y,e) = \sum_{i=0}^{14}\sum_{j=0}^{6} B^{nij}_\alpha T_i\,(u)\,T_j\,(v) \tag{6.7}$$

Using this procedure forces each component to become singular at the LSEO. While the GSF changes rapidly as one approaches the last stable orbit, we do not expect the components of the self force to diverge at the LSEO. Understanding the analytic structure of the self-force in this region would likely improve future interpolation models.

We note that the GSF should satisfy the orthogonality condition with the geodesic four-velocity, i.e., $a_\alpha u^\alpha = 0$. Interpolation will bring with it a certain amount of error which can cause this condition to be violated. We find empirically that we can reduce this interpolation error by projecting the force so that this condition is always satisfied, i.e.,

$$a^\perp_\alpha = a_\alpha + a_\beta u^\beta u_\alpha. \tag{6.8}$$

This procedure allows us to create a smooth, continuous model for the gravitational self force with relative errors less than 5×10^{-3} in the strong field – see Fig. 6.1. The variation in the accuracy of the model is primarily a by-product of how close a given test point (green cross) is to the data points (white dots) used to create the model. We note that this level of precision would not be sufficient for production grade waveforms for LISA, as we would need the relative error of the orbit averaged dissipative self-force to be less than $\sim \epsilon$, whereas the oscillatory pieces of the self-force only need to be interpolated to an accuracy of a fraction of a percent [70]. These are somewhat crude estimates for the accuracy requirements as the effect on parameter estimation has yet to be quantified. Our present interpolation model already likely reaches the latter criteria and a future hybrid method that combines flux and self-force data, similar to the one constructed in Ref. [70], can likely reach the overall accuracy goal. Nonetheless, our present model is more than sufficient to test our averaging procedure and to explore the effects of the GSF for eccentric Kerr inspirals. This will now be treated as the underlying forcing model for both the OG and NIT inspirals.

6.3 Near Identity Transformations

We now apply the near identity averaging transformation procedure to the equations of motion for equatorial Kerr inspirals to obtain:

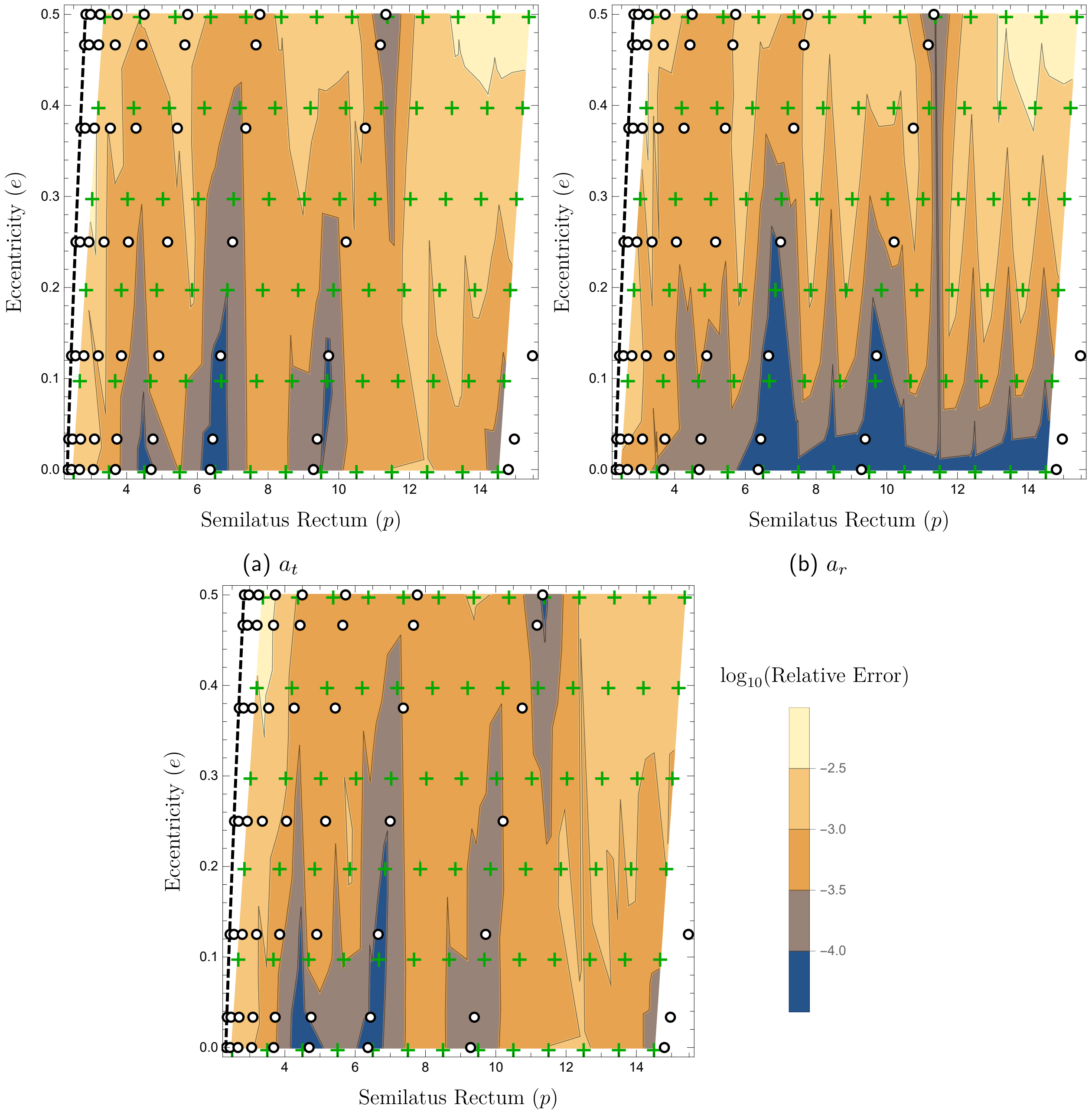
$$\frac{d\tilde{p}}{d\lambda} = \epsilon \tilde{F}^{(1)}_p(a,\tilde{p},\tilde{e}) + \epsilon^2 \tilde{F}^{(2)}_p(a,\tilde{p},\tilde{e}), \tag{6.9a}$$

$$\frac{d\tilde{e}}{d\lambda} = \epsilon \tilde{F}^{(1)}_e(a,\tilde{p},\tilde{e}) + \epsilon^2 \tilde{F}^{(2)}_e(a,\tilde{p},\tilde{e}), \tag{6.9b}$$

$$\frac{d\tilde{q_r}}{d\lambda} = \Upsilon^{(0)}_r(a,\tilde{p},\tilde{e}) + \epsilon \Upsilon^{(1)}_r(a,\tilde{p},\tilde{e}), \tag{6.9c}$$

$$\frac{d\tilde{t}}{d\lambda} = \Upsilon^{(0)}_t(a,\tilde{p},\tilde{e}) + \epsilon \Upsilon^{(1)}_t(a,\tilde{p},\tilde{e}), \tag{6.9d}$$

$$\frac{d\tilde{\phi}}{d\lambda} = \Upsilon^{(0)}_\phi(a,\tilde{p},\tilde{e}) + \epsilon \Upsilon^{(1)}_\phi(a,\tilde{p},\tilde{e}). \tag{6.9e}$$



(a) a_t

(b) a_r

(c) a_ϕ

Figure 6.1: The relative error of the components of the interpolated gravitational self force model for prograde equatorial orbits with $a = 0.9M$. The white dots represent the data points that were interpolated. The green crosses represent the data set that the model was tested against. The black dashed line represents the location of the last stable equatorial orbit. The relative error was calculated using the normalised L^2 error over a single orbital cycle.

The leading order terms in each equation of motion are simply the original function averaged over a single geodesic orbit, i.e.,

$$\tilde{F}_p^{(1)} = \left\langle F_p^{(1)} \right\rangle, \quad \tilde{F}_e^{(1)} = \left\langle F_e^{(1)} \right\rangle, \quad \Upsilon_r^{(1)} = \left\langle f_r^{(1)} \right\rangle, \tag{6.10}$$

$$\Upsilon_t^{(0)} = \left\langle f_t^{(0)} \right\rangle, \quad \Upsilon_\phi^{(0)} = \left\langle f_\phi^{(0)} \right\rangle, \tag{6.11}$$

where $\Upsilon_t^{(0)}$ and $\Upsilon_\phi^{(0)}$ are the Mino-time t and ϕ fundamental frequencies. The remaining terms are more complicated and require Fourier decomposing the original functions and their derivatives with respect to to the orbital elements (p, e). To express the result, we define the operator

$$\mathcal{N}(A) = \sum_{n \neq 0} \frac{-1}{\Upsilon_r^{(0)}} \left[A_n f_{r,-n}^{(1)} - \frac{i}{n} \left(\frac{\partial A_n}{\partial \tilde{p}} F_{p,-n}^{(1)} + \frac{\partial A_n}{\partial \tilde{e}} F_{e,-n}^{(1)} - \frac{A_n}{\Upsilon_r^{(0)}} \left(\frac{\partial \Upsilon_r^{(0)}}{\partial \tilde{p}} F_{p,-n}^{(1)} + \frac{\partial \Upsilon_r^{(0)}}{\partial \tilde{e}} F_{e,-n}^{(1)} \right) \right) \right] \tag{6.12}$$

With this in hand, the remaining terms in the equations of motion are found to be

$$\tilde{F}_p^{(2)} = \left\langle F_p^{(2)} \right\rangle + \mathcal{N}(F_p^{(1)}), \quad \tilde{F}_e^{(2)} = \left\langle F_p^{(2)} \right\rangle + \mathcal{N}\left(F_e^{(1)} \right), \quad \Upsilon_t^{(1)} = \mathcal{N}\left(f_t^{(0)} \right), \quad \Upsilon_\phi^{(1)} = \mathcal{N}\left(f_\phi^{(1)} \right). \tag{6.13}$$

Note that since we currently only have access to first order GSF effects, we set $\left\langle F_p^{(2)} \right\rangle = \left\langle F_e^{(2)} \right\rangle = 0$.

6.4 Implementation

Combining the above model with our action angle formulation of the osculating geodesic equations provides us with everything required to calculate the NIT equations of motion. We first evaluate and interpolate the various terms in the NIT equations of motion across the parameter space. This offline process is costly but it only needs needs to be completed once. By contrast, the online steps are computationally cheap, which allows us to rapidly compute eccentric self-forced inspirals into a Kerr black hole.

6.4.1 Offline Steps

To make the offline calculation we complete the following steps.

1. We start by selecting a grid to evaluate the NIT functions upon. We chose y values between 0.2 and 0.998 in 320 equally spaced steps and e values from 0.001 to 0.5 in 500 equally spaced steps (160,000 points) in the case of Kerr, or use the same spacing in y_{eq} but only grid in e from 0.001 to 0.3 in 300 equally spaced steps (96,000 points) in Schwarzschild.[1]
2. For each point in the parameter space (a, y_{eq}, e) we evaluate the functions $F^{(1)}_{p\backslash e}, f^{(1)}_r$ and $f^{(0)}_{t\backslash\phi}$ along with their derivatives with respect to p and e for 30 equally spaced values of q_r from 0 to 2π.
3. We then perform a fast Fourier transform on the output data to obtain the Fourier coefficients of the forcing functions and their derivatives.

---

[1] Evaluating the NIT functions is computationally cheap so using a dense grid does not significantly increase the computational burden. Using a dense grid also allows us to use Mathematica's default Hermite polynomial interpolation method for convenience of implementation. The grid spacing is chosen to be sufficiently dense that interpolation error is a negligible source of error for our comparisons between the OG and NIT inspirals, though a less dense grid may also achieve this.

4. With these, we then use Eqs. (6.10), (6.11), (6.12) and (6.13) to construct $\tilde{F}^{(1\backslash 2)}_{p\backslash e}$ and $\Upsilon^{(1)}_{r\backslash t\backslash \phi}$ for that point in the parameter space.

5. We also use Eqs. (4.55) and (4.56) to construct the Fourier coefficients of the first-order transformation functions $Y^{(1)}_{p\backslash e}$ and $X^{(1)}_{r}$.

6. We then repeat this procedure across the parameter space for each point in our grid.

7. Finally we interpolate the values for $\tilde{F}^{(1\backslash 2)}_{p\backslash e}$ and $\Upsilon^{(1)}_{r\backslash t\backslash \phi}$ along with the coefficients of $\breve{Y}^{(1)}_{p\backslash e}$ and $\breve{X}^{(1)}_{r}$ across this grid using Hermite interpolation and store the interpolants for future use.

We implemented the above algorithm in Mathematica 12.2 and find, parallelized across 20 CPU cores takes, the calculation takes about one day to complete. This is a small price to pay, since these offline steps need only be completed once.

6.4.2 Online Steps

The online steps are required for every inspiral calculation, but are comparatively inexpensive. The online steps for computing an NIT inspiral are as follows.

1. We load in the interpolants for $\tilde{F}^{(1\backslash 2)}_{p\backslash e}$ and $\Upsilon^{(1)}_{r\backslash t\backslash \phi}$, define the NIT equations of motion.

2. In order to make comparisons between NIT and OG inspirals, we also load interpolants of the Fourier coefficients of $\breve{Y}^{(1)}_{p/e}$ and $\breve{X}^{(1)}_{r}$ and Eq. (4.1) to construct first order near-identity transformations.[2]

3. We state the initial conditions of the inspiral $(p_0, e_0, q_{r,0})$ and use the NIT to leading order in the mass ratio to transform these into initial conditions for the NIT equations of motion, i.e., $(\tilde{p}_0, \tilde{e}_0, \tilde{q}_{r0})$.

4. We then evolve the NIT equations of motion using an ODE solver (in this case Mathematica's `NDSolve`).

As with the offline steps we implement the online steps in Mathematica. Note that steps (ii) and (iii) are only necessary because we want to make direct comparisons between NIT and OG inspirals with the same initial conditions. In general, the difference between the NIT and OG variables will always be $\mathcal{O}(\epsilon)$, and so performing the NIT transformation or inverse transformation to greater than zeroth order in mass ratio will not be necessary when producing waveforms to post adiabatic order, i.e. with phases accurate to $\mathcal{O}(\epsilon)$.

6.5 Results

In this section we present the results from the NIT equations of motion. We first perform some consistency checks in Sec. 6.5.1. We then show that our NIT and OG inspirals agree to the relevant order in the mass ratio in Sec. 6.5.2. Here we also compute, for the first time, self-forced inspirals in Kerr spacetime. With our fast NIT model we then explore the impact of the conservative effects of the first-order GSF as calculated in radiation gauge for Kerr inspirals in Sec. 6.5.3. Finally, in Sec. 6.5.4, we compare Schwarzschild inspirals calculated using a radiation gauge GSF model and a Lorenz gauge GSF model.

---

[2] Note that while including $\breve{X}^{(1)}_{r}$ in the transformation is not strictly necessary, we do so anyway to further reduce the initial difference between the two inspirals.

6.5.1 Consistency checks

Before computing inspirals, we perform a series of consistency checks on the NIT equations of motion. A useful feature of the NIT is how it separates adiabatic and post-adiabatic effects of the gravitational self-force. At first order in the mass ratio, this corresponds to the dissipative and conservative pieces respectively. We note that when we substitute $a^\alpha \to a^\alpha_{\text{diss}}$, we find that $\tilde{F}^{(2)}_{p\backslash e}$ and $\Upsilon^{(1)}_{r\backslash t\backslash \phi}$ are numerically consistent with zero, while $\tilde{F}^{(1)}_{p\backslash e}$ remains unchanged. Similarly, when we substitute $a^\alpha \to a^\alpha_{\text{cons}}$, $\tilde{F}^{(1)}_{p\backslash e}$ and $\tilde{F}^{(2)}_{p\backslash e}$ become consistent with zero, while $\Upsilon^{(1)}_{r\backslash t\backslash \phi}$ remain the same as before. The functions $\tilde{F}^{(2)}_{p\backslash e}$ only becomes non-zero when both dissipative and conservative effects of the first order self-force are present.

From $\tilde{F}^{(1)}_{p\backslash e}$, one can calculate the average rate of change of energy and angular momentum via the following relation:

$$\left\langle \frac{d\mathcal{E}}{dt} \right\rangle = \frac{\epsilon}{\Upsilon^{(0)}_t} \left(\frac{\partial \mathcal{E}}{\partial p} \tilde{F}^{(1)}_p + \frac{\partial \mathcal{E}}{\partial e} \tilde{F}^{(1)}_e \right) \tag{6.14a}$$

$$\left\langle \frac{d\mathcal{L}}{dt} \right\rangle = \frac{\epsilon}{\Upsilon^{(0)}_t} \left(\frac{\partial \mathcal{L}}{\partial p} \tilde{F}^{(1)}_p + \frac{\partial \mathcal{L}}{\partial e} \tilde{F}^{(1)}_e \right). \tag{6.14b}$$

We compared these to the energy and angular momentum fluxes at infinity tabulated in the Black Hole Perturbation Toolkit [182] and generated with a variant of the `Gremlin` code [79, 80] and found that the balance laws were upheld up to relative errors $< 10^{-3}$ throughout the parameter space which is consistent with the interpolation error of our self-force model.

From all of this, we can infer the significance of each of the terms in Eq. (7.12): $\Upsilon^{(0)}_r$, $\Upsilon^{(0)}_t$ and $\Upsilon^{(0)}_\phi$ capture the background geodesic motion, $\tilde{F}^{(1)}_p$ and $\tilde{F}^{(1)}_e$ capture the adiabatic effects due to the first order dissipative self-force, $\Upsilon^{(1)}_r$, $\Upsilon^{(1)}_t$, and $\Upsilon^{(1)}_\phi$ capture the post-adiabatic effects due to the first order conservative self-force, and $\tilde{F}^{(2)}_p$, $\tilde{F}^{(2)}_e$ capture the interplay between the first order dissipative and conservative self-force, as well as the effect of the orbit averaged contribution from the second order self-force.

6.5.2 Comparison between OG and NIT inspirals

In order to test the accuracy of our implementation, we compare inspirals calculated using the OG equations of motion found in Ref. [137] to those calculated using the near-identity transformed equations of motion. To demonstrate these results, we choose a binary with a primary of mass $M = 10^6 M_\odot$ and a secondary of mass $\mu = 10 M_\odot$ for a typical EMRI mass ratio of $\epsilon = 10^{-5}$. To push our procedure to the limit, we chose the initial conditions of our prograde inspiral to be deep in the strong field and highly eccentric with $p_0 = 7.1$ and $e_0 = 0.48$ such that the resulting inspiral would take approximately 1 year to plunge. We also set $q_{r,0} = t_0 = \phi_0 = 0$ for simplicity.

Figure 6.2 shows the evolution of p and e over time. The trajectories calculated with the OG equations of motion have order ϵ oscillations on the orbital timescale which requires the numerical integrator to take small time steps to accurately resolve. The NIT trajectory does not have these oscillations so the numerical integrator can take much larger steps and still faithfully track the averaged trajectory throughout the entire inspiral. The inverse NIT given in Eq. (4.2) through $\mathcal{O}(\epsilon)$ can be used to add the oscillations back on to the NIT trajectory. We find that while this is unnecessary for computing accurate waveforms, it demonstrates that the NIT trajectory remains in phase with the OG trajectory – see the insets of Fig. 6.2.

The accuracy of our NIT model is further demonstrated by Fig. 6.3 which shows the absolute difference in the orbital phase q_r and the extrinsic quantities t and ϕ between the NIT and OG evolutions. Over the course of the year long inspiral, $|t - (\tilde{t} - Z^{(0)}_t)| \leq 5 \times 10^{-3}$,$|\phi - (\tilde{\phi} - Z^{(0)}_\phi)| \leq 10^{-5}$ and

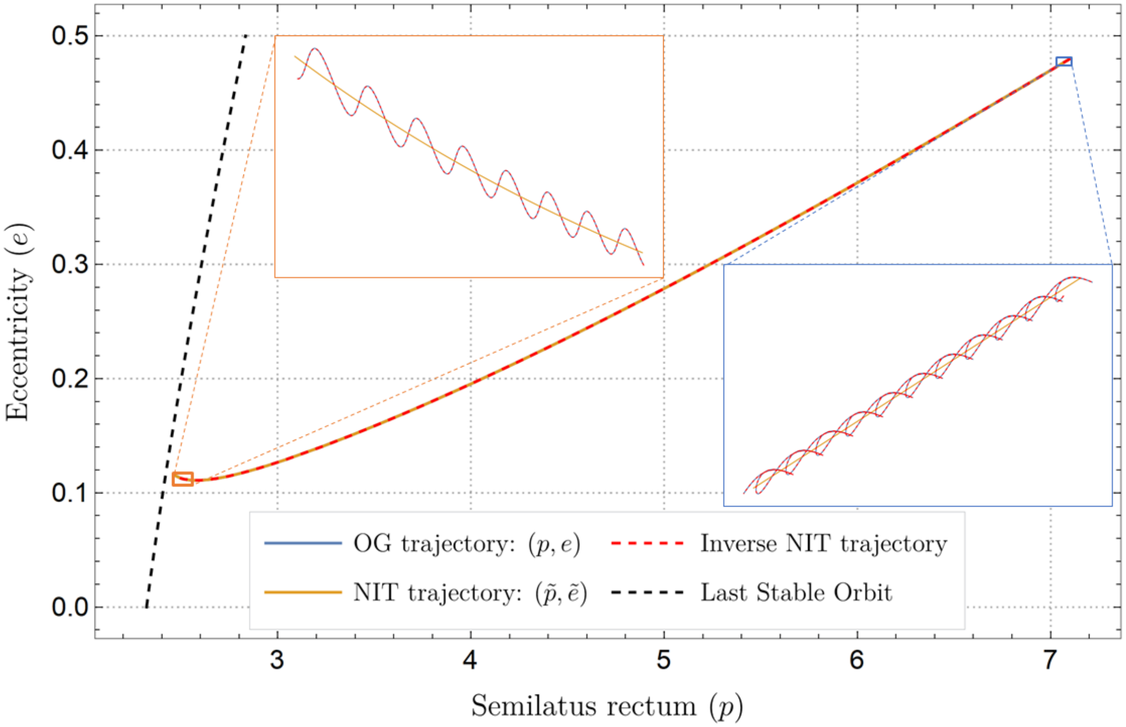


Figure 6.2: The trajectory through (p, e) space for an inspiral with $\epsilon = 10^{-5}$, $a = 0.9M$, and initial conditions $(p_0 = 7.1, e_0 = 0.48)$. We show the inspiral computed using the osculating geodesic equations, the NIT equations of motion and the inverse NIT to first order in ϵ. The insets zoom into the start and end of the inspiral to reveal the small orbital timescale oscillations. The NIT averages through these oscillations, and when using the inverse NIT to add the oscillations back on, we see that the NIT trajectory remains almost perfectly in phase with the OG trajectory throughout the inspiral.

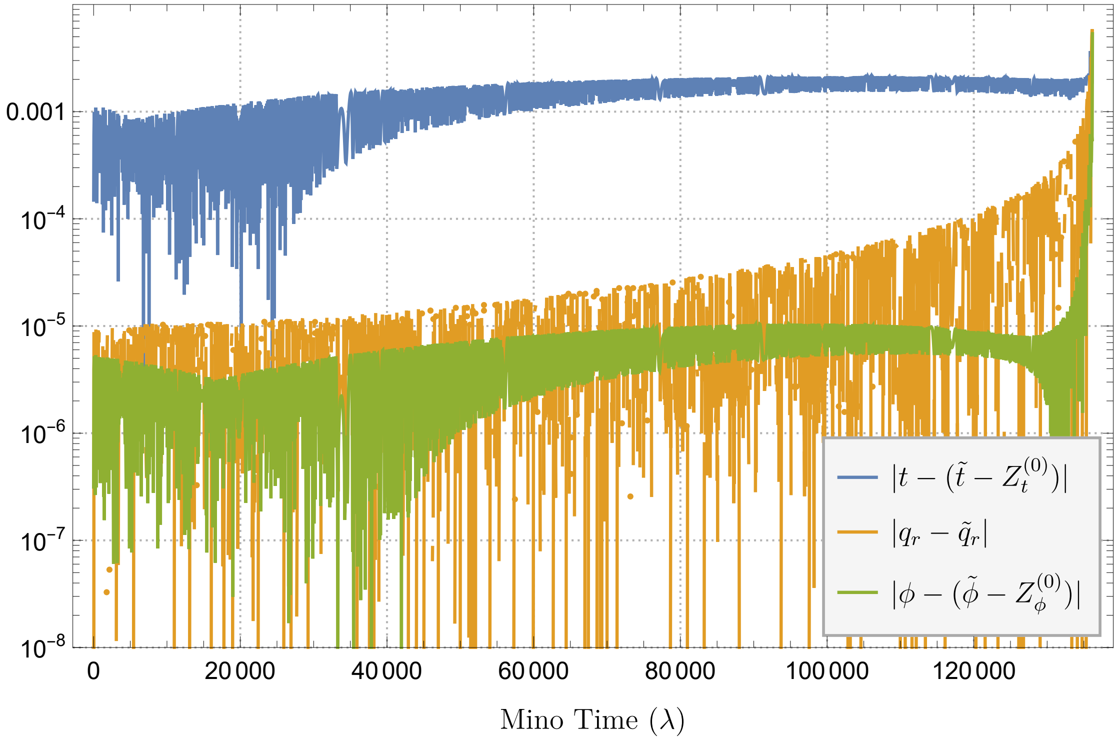


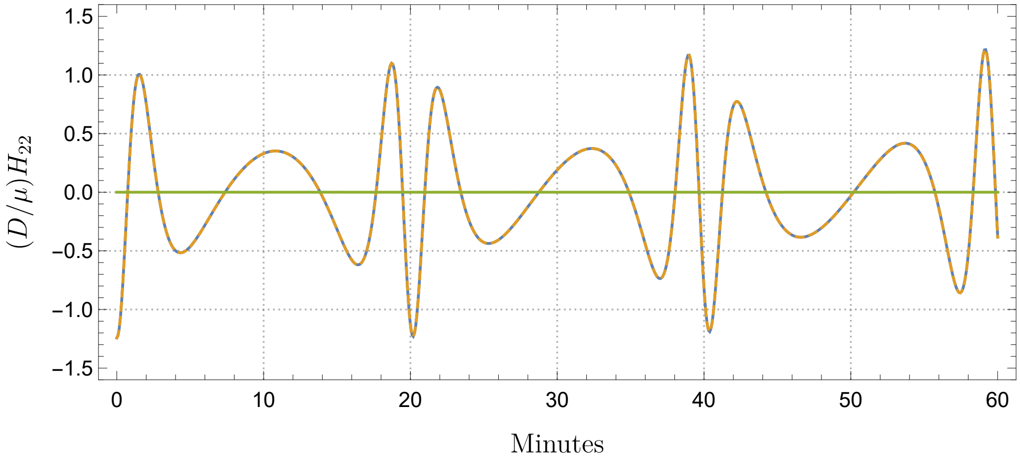
Figure 6.3: The difference in the orbital phase and extrinsic quantities for a equatorial Kerr inspiral with $\epsilon = 10^{-5}$ and $a = 0.9M$ calculated using the OG and NIT equations of motion with initial conditions $p_0 = 7.1, e_0 = 0.48$. We find that the differences remain small throughout the inspiral, only becoming large as the secondary approaches the last stable orbit where the adiabatic approximation breaks down.

$|q_r - \tilde{q}_r| \leq 10^{-3}$ with the differences only spiking to $\leq 10^{-2}$ just as the trajectories reach the separatrix where the adiabatic approximation breaks down.

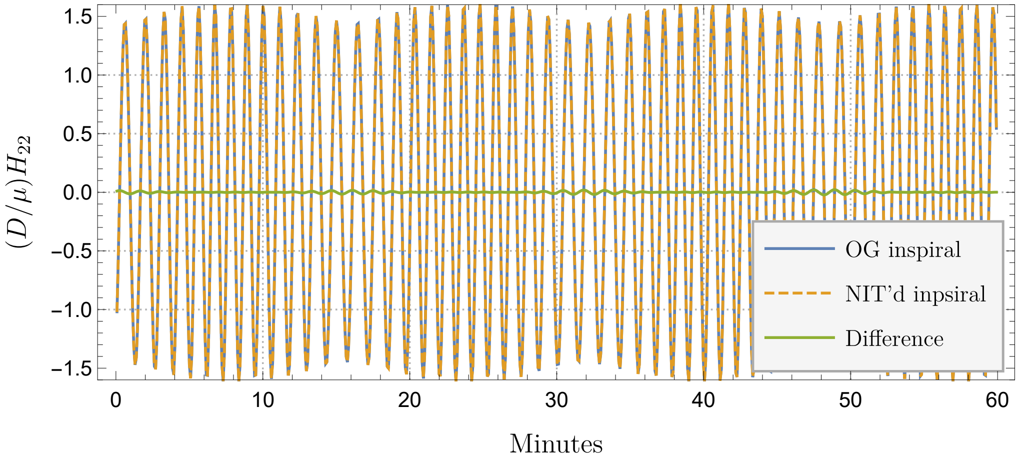
Finally, we test the effect the NIT procedure has on the waveform. In principle, we could use our averaged equations of motion in conjunction with the `FastEMRIWaveforms` (FEW) framework to rapidly compute waveforms with relativistic amplitudes. However, currently, the FEW framework only has amplitude data for Schwarzschild inspirals. As such, we make use of the same procedure as the Numerical Kludge [116] by mapping the Boyer-Lindquist coordinates $\{t, r, \theta, \phi\}$ to flat space coordinates and using the quadrupole formula to generate the waveform. The resulting waveforms are only an approximation to the true waveforms, but since both inspiral trajectories are being fed through the same waveform generation scheme this should not bias the results when finding the difference in the waveform as a result of using the NIT trajectory instead of the OG trajectory.

From Fig. 6.4, we can see that the waveforms generated by each evolution scheme,sampled every $t = 1M \approx 5s$, are almost identical by eye. We can further quantify this by calculating the waveform mismatch using the `WaveformMatch` function from the `SimulationTools` [186] Mathematica package and assuming a flat noise curve. From Fig. 6.5, we see that the mismatch remains below 5×10^{-8} throughout the inspiral. Using Eq.(5.36), these two waveforms would be completely indistinguishable for EMRIs with signal-to-noise ratio (SNR) of upto $1000\sqrt{10} \approx 3162$.

Next, the difference between the OG and NIT quantities should scale linearly with the mass ratio. This is illustrated in Fig. 6.6, where starting with initial conditions $p_0 = 4$ and $e_0 = 0.2$ we evolved the inspiral until it reached $p = 3$ for mass ratios ranging from 10^{-1} to 10^{-5}. While working with only machine precision arithmetic we found that for smaller mass ratios the numerical error of the solver of



(a) First hour of the waveform.



(b) Last hour of the waveform.

Figure 6.4: Two snapshots of the dominant $(l, m) = (2, 2)$ mode of the quadrupole waveform for our prograde, equatorial Kerr inspiral with $(a, \epsilon, p_0, e_0) = (0.9M, 10^{-5}, 7.1, 0.48)$. These snapshots correspond to the first and last hours of the inspiral. This shows that the waveform generated using the NIT trajectory almost perfectly overlaps with the waveform generated using the OG trajectory. It also demonstrates how dramatically an EMRI waveform evolves throughout the inspiral.

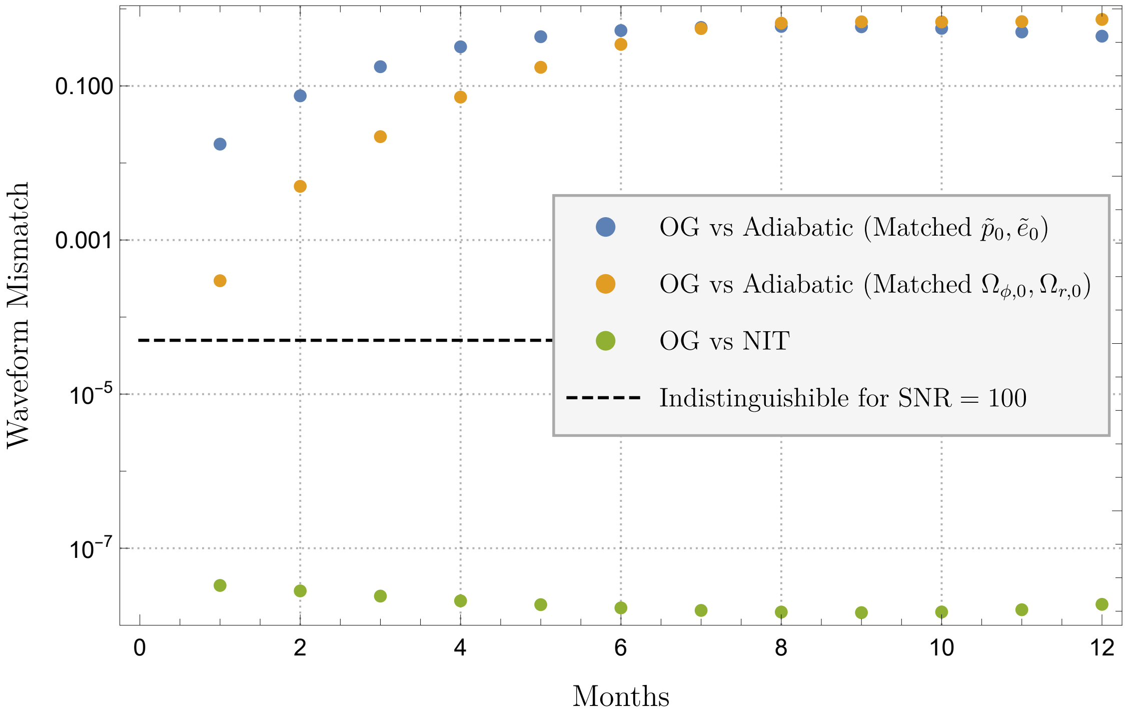


Figure 6.5: The mismatch between the semi-relativistic quadrupole waveforms between inspirals calculated using the OG equations with $(a, \epsilon, p_0, e_0) = (0.9M, 10^{-5}, 7.1, 0.48)$ and the adiabatic EOM matched initial conditions, the adiabatic EOM calculated with matched initial frequencies, and the near-identity transformed EOM. We also mark the mismatch that would be indistinguishable for signals with SNR $= 100$.

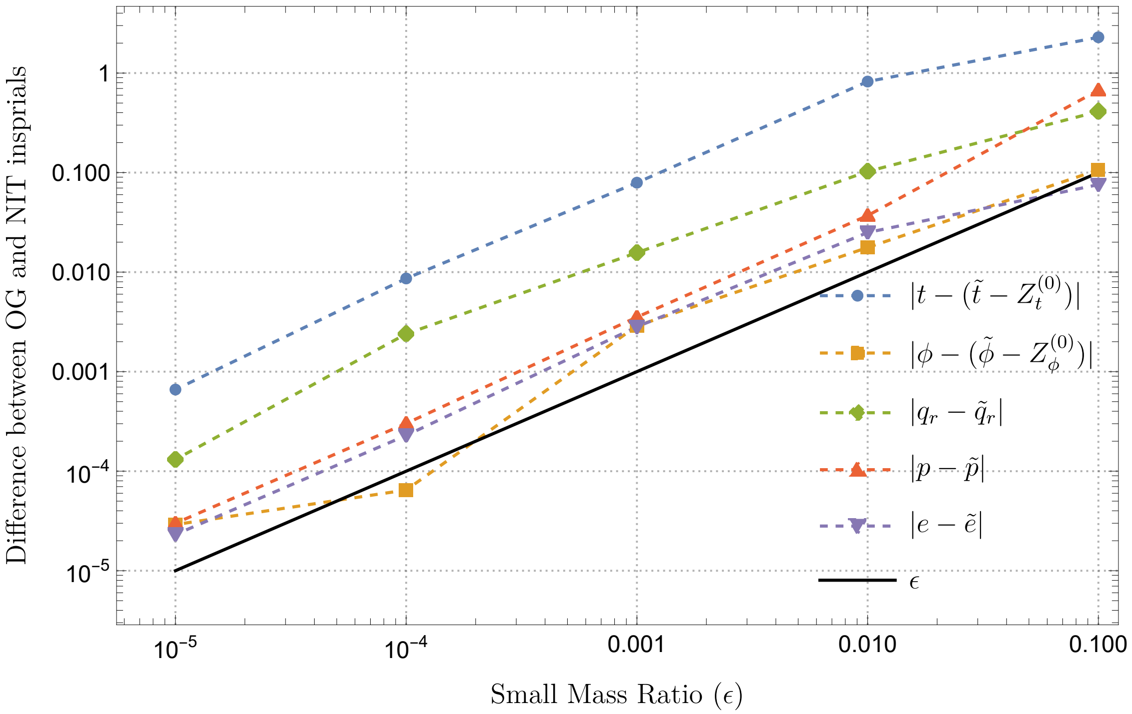


Figure 6.6: The absolute difference in the quantities of a prograde inspiral with $a = 0.9M$ and $e_0 = 0.2$ after evolving from $p = 4$ to $p = 3$ using either the OG or NIT equations of motion. We observe that all the differences follow the solid, black ϵ reference curve, as expected.

the OG inspiral became dominant over the difference with the NIT. To rectify this, we increased the working precision of our solver to 30 significant digits and found that the difference does, in fact, scale linearly with the mass ratio. This requirement for higher precision only affected the OG solver, the NIT equations of motion can be solved with machine precision arithmetic without introducing any significant error.

Since the difference between OG and NIT quantities scales with the mass ratio, it is natural to ask how large can the mass ratio be before the NIT and OG waveforms differ enough to affect data analysis. Following the procedure outlined in Ref. [156], we used our fast NIT inspiral code along with a root-finding algorithm to find the initial value of p that corresponds to a year long inspiral for a given value of the mass ratio and initial eccentricity, and assuming a primary mass of $10^6 M_\odot$. We use these initial conditions to calculate the overlap between year-long NIT and OG waveforms. This calculation is repeated with mass ratios $\epsilon = \{1, 3, 5, 7, 9\} \times 10^{-3}$ and initial eccentricities e_0 ranging from 0.05 to 0.45 in equally spaced steps of 0.05. The result of this analysis can be seen in Fig. 6.7. This demonstrates that NIT and OG waveforms have overlaps larger than the benchmark of 0.97 [187] for mass ratios less than $\approx 3 \times 10^{-3}$, but these overlaps decrease substantially for mass ratios larger than this. We also see that the overlap generally decreases as the initial eccentricity increases, though this effect is not as strong as the effect demonstrated by a similar analysis in Ref. [156] for NITs applied to highly eccentric inspirals in Schwarzschild. They also found that the mismatch between NIT and OG waveforms became substantial for mass ratios larger than 2×10^{-4}. These differences between the two analyses are most likely the result of our inspirals being deeper in the strong field and driven by a self-force computed in a different gauge (Ref. [156] uses the Lorenz gauge self-force). Such mismatches should not be an issue for EMRI data analysis as EMRIs have mass ratios that range from 10^{-7} to 10^{-4}. However, these mismatches become significant for intermediate mass ratio inspirals, with mass ratios between 10^{-4} to

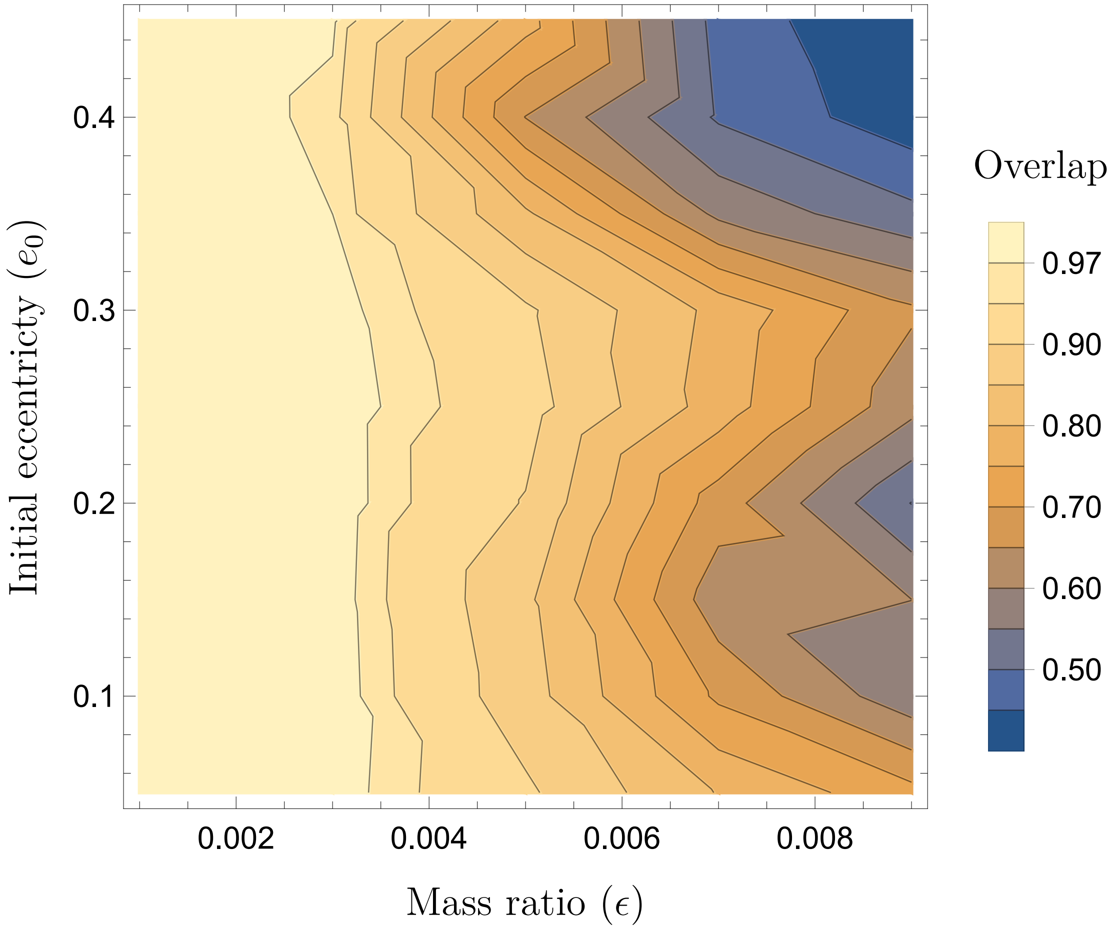


Figure 6.7: The overlap between OG and NIT waveforms for year-long, prograde, $a = 0.9M$, equatorial Kerr inspirals as a function of the mass ratio and initial eccentricity. The difference between the two waveforms is less than the accuracy benchmark of 0.97 for mass ratios $\leq 3 \times 10^{-3}$, but not for mass ratios larger than this. While increasing eccentricity does have an effect on the overlap, this effect is not as strong as the effect observed in Fig. 9 of Ref. [156].

ϵ	OG Inspiral	NIT Inspiral	Speed-up
10^{-2}	44s	0.85s	~ 37
10^{-3}	6m 48s	0.78s	~ 491
10^{-4}	54m 12s	0.81s	~ 3782
10^{-5}	6hrs 16m	0.76s	~ 29655

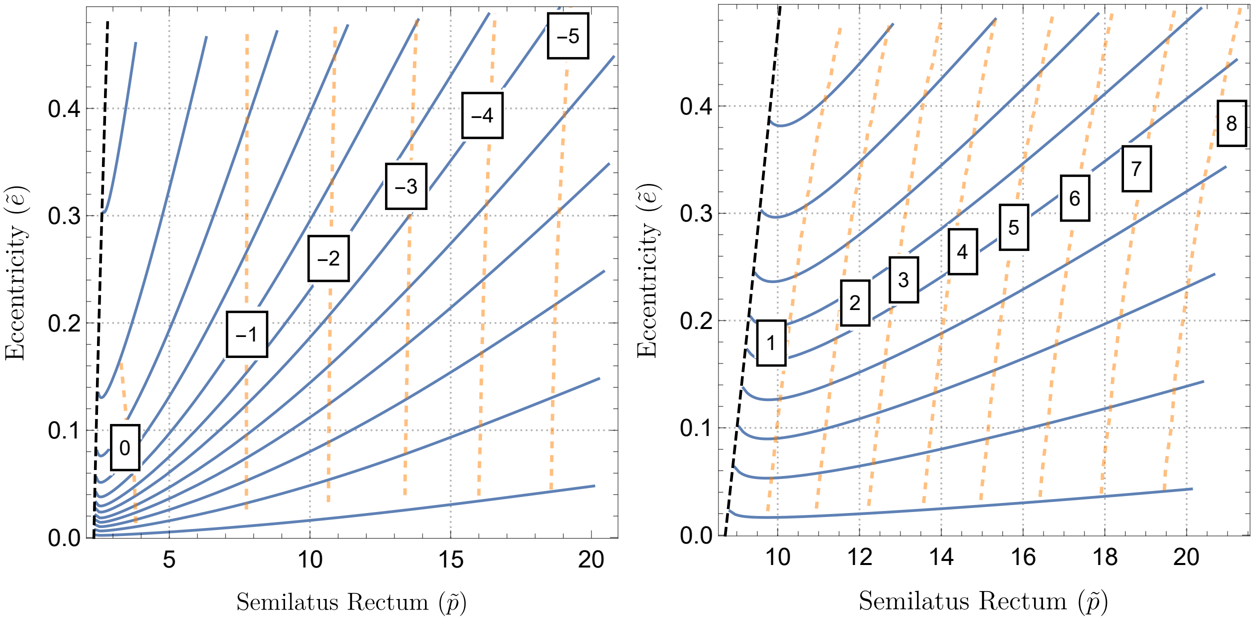
Table 6.1: Computational time required to evolve an inspiral from its initial conditions of $p_0 = 7.1$ and $e_0 = 0.48$ to the last stable orbit for different values of the mass ratio, as calculated in Mathematica 12.2 on an Intel Core i7 @ 2.2GHz. The computational time for the OG inspiral scales inversely with the mass ratio, whereas the computational time for NIT inspirals is independent of the mass ratio. This demonstrates how the smaller the mass ratio of the inspiral, the greater speed-up one obtains from using the NIT equations of motion.

10^{-1}. Since both the OG and NIT equations of motion are formally valid to the same order in the mass ratio, it is not clear a priori which of the two would be closer to the true inspiral. When completed at 1-post-adiabtatic (1PA) order the two sets of equations represent different resummations of the 1PA equations of motion, differing only in their higher order (2+) PA terms. The fact that we are seeing a significant difference between these two resummations for intermediate mass ratios suggests that such higher order PA terms might become relevant. However, in this case it might just be signalling the importance of the missing orbit-averaged dissipative self-force term at 1PA order.

Finally, we note that using the NIT equations of motion produces a substantial speed-up over using the OG equations. From Table 6.1, we see the typical computation time for an inspiral starting at $p_0 = 7.1$ and $e_0 = 0.48$ and evolved until the inspiral reaches the last stable orbit for different values of the mass ratio. We see that as we decrease the mass ratio by an order of magnitude, the OG inspiral takes roughly an order of magnitude longer to compute, as it would have to resolve an order of magnitude more orbital cycles before reaching last stable orbit. The NIT inspirals all take roughly the same amount of time to evolve to the last stable orbit, regardless of the mass ratio. Using our current Mathematica implementation, the NIT inspirals can be computed in less than a second. This time could be further reduced tens of milliseconds if one uses a compiled language such as C/C++, as was done in Paper I [154]. We see that using the NIT equations of motion is most advantageous for long inspirals with small mass ratios. Another benefit of using the NIT is that the inspiral requires taking fewer time steps, which results in less numerical error, making it easier achieve a given target accuracy.

The only disadvantage of our formulation is that our final trajectory is parametrized in terms Mino time λ, whereas LISA data analysis applications will need waveforms parametrized Boyer-Lindquist retarded time t. Since our formulation also outputs $t(\lambda)$, we can numerically invert this to get $\lambda(t)$ which allows us to resolve this issue at the cost of additional computation time. This was also a problem with the NIT formulation in Schwarzschild where the final trajectory is outputted as a function of the quasi-Keplerian angle χ [154, 156]. We have found that this problem can be circumvented entirely by performing an additional transformation to our NIT equations of motion which would produce averaged equations of motion parametrized by t as outlined in Sec. 4.3 which we implement for spherical inspirals in chapter 7.

Since we are now satisfied that our formulation can produce fast and accurate self-force driven trajectories, we can now use this procedure to explore the phenomenology of eccentric, equatorial Kerr inspirals.



(a) Prograde inspirals

(b) Retrograde inspirals

Figure 6.8: Sample trajectories through $(\tilde{p}, \tilde{e})$ space for prograde and retrograde equatorial Kerr inspirals with $\epsilon = 10^{-5}$ and $a = 0.9M$. From these plots, we see the familiar behaviour of EMRIs losing eccentricity as the compact object approaches the primary and then gaining eccentricity just before crossing the last stable equatorial orbit (dashed black line). The dashed orange curves are contours that mark the number of radians $\tilde{q}_{r,0}$ will evolve from a given point until plunge. The conservative self-force for retrograde orbits has a similar effect to the non-spinning case as it causes $\tilde{q}_{r,0}$ to increase throughout the inspiral. In the prograde case, $\tilde{q}_{r,0}$ decreases for most of the inspiral and then slightly increases shortly before plunge.

6.5.3 Impact of adiabatic and post-adiabatic effects

With the ability to generate fast and accurate inspirals, we can survey the physics of equatorial Kerr inspirals and examine how this differs from the Schwarzschild case. From Fig. 6.8a, we see the familiar effect of gravitational radiation reaction on the semilatus rectum, $\tilde{p}$, and eccentricity, $\tilde{e}$, whereby $\tilde{p}$ and $\tilde{e}$ both decrease over the inspiral with $\tilde{e}$ growing a little as the last stable orbit is approached [75–77]. As the inspiral approaches the last stable equatorial orbit adiabaticity breaks down and the inspiral undergoes a transition to plunge [138–140, 142]. As such, we stop our inspirals just before the LSEO. Our results are the first inspirals to include conservative self-force corrections to the equations of motion in Kerr spacetime. The initial phase $q_{r,0}$ only evolves secularly when conservative self-force corrections are present and so we use this as a measure of the influence of these corrections [145]. This is illustrated by the dashed orange curves in Fig. 6.8a, which mark the number of radians $\tilde{q}_{r,0}$ will evolve from a given pair of initial conditions $(\tilde{p}_0, \tilde{e}_0)$ until the last stable orbit. For retrograde Kerr (and Schwarzschild orbits in Fig. 6.10), we find that $\tilde{q}_{r,0}$ increases throughout the inspiral, whereas for prograde Kerr $\tilde{q}_{r,0}$ decreases during the inspiral before increasing slightly just before plunge. This is consistent with the change of sign in the correction to the rate of periapsis advance induced by the conservative self force as a function of spin in the circular orbit limit [85] – see Appendix A for further details.

As discussed in Sec. 6.5.1, one can readily calculate adiabatic inspirals using the NIT equations of motion by simply neglecting the post-adiabatic terms. However, when trying to determine how post-

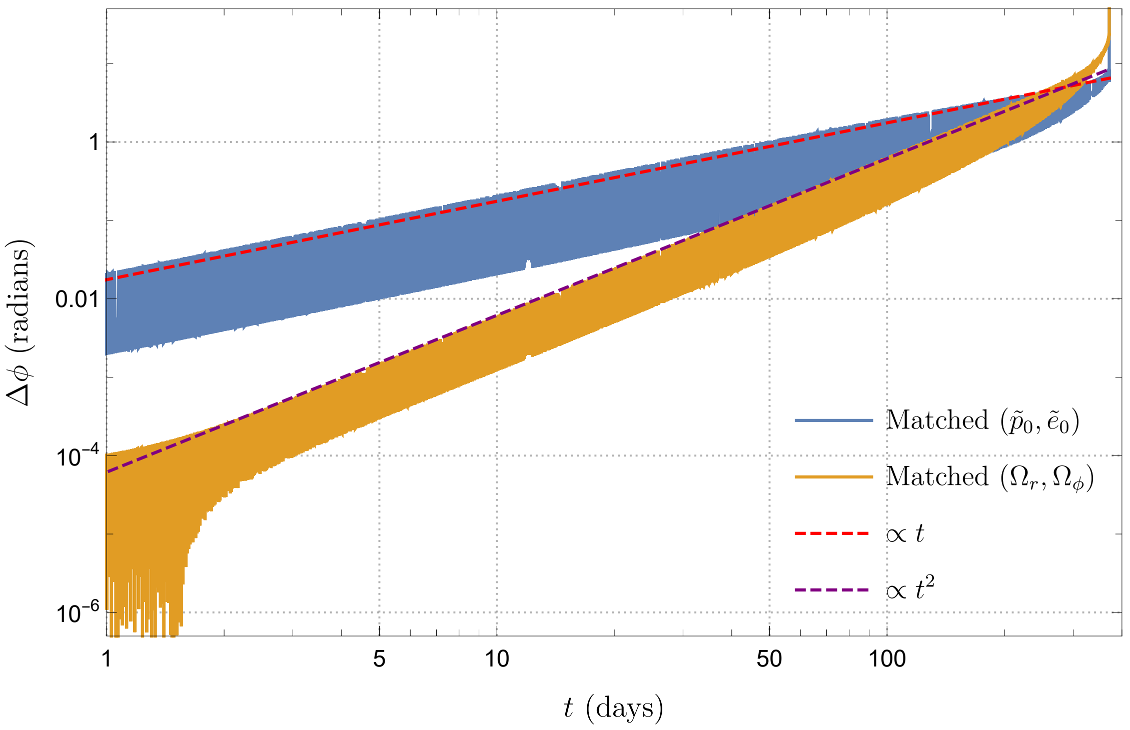


Figure 6.9: Difference in ϕ as a function of t between an adiabatic and a first order self-forced inspiral when either matching initial conditions or matching the initial Boyer-Lindquist frequencies. The self-forced inspiral has initial conditions $(\tilde{p}_0, \tilde{e}_0) = (7.1, 0.48)$ with mass ratio $\epsilon = 10^{-5}$. Matching initial conditions results in an error that grows linearly with t, while matching frequencies produces an error that is initially constant and then grows quadratically with t.

adiabatic corrections effect the inspiral, one must be mindful of how one matches up an adiabatic inspiral with its post-adiabatic counterpart. Following the argument found in Refs. [145] and [70], matching the initial conditions $(\tilde{p}_0, \tilde{e}_0)$ results in an error in the orbital phases that grows linearly in t as the conservative self-force changes the orbital frequencies [83]. Instead, one should instead match the Boyer-Lindquist time fundamental frequencies Ω_r and Ω_ϕ. For an adiabatic inspiral, these are directly related to the Mino-time fundamental frequencies via $\Omega^{(0)}_{r\backslash\phi} = \Upsilon^{(0)}_{r\backslash\phi}/\Upsilon^{(0)}_t$ [174]. To calculate these frequencies as perturbed by the conservative self-force, one can either follow the method outlined in Ref. [70], or one can calculate them directly from the NIT equations of motion:

$$\Omega^{\mathrm{SF}}_r = \frac{\Upsilon^{(0)}_r + \epsilon\Upsilon^{(1)}_r}{\Upsilon^{(0)}_t + \epsilon\Upsilon^{(1)}_t} + \mathcal{O}(\epsilon^2) \quad \text{and} \quad \Omega^{\mathrm{SF}}_\phi = \frac{\Upsilon^{(0)}_\phi + \epsilon\Upsilon^{(1)}_\phi}{\Upsilon^{(0)}_t + \epsilon\Upsilon^{(1)}_t} + \mathcal{O}(\epsilon^2). \tag{6.15}$$

We find that both approaches give the same result up to an error that scales as ϵ^2. With this in hand, we can now choose a value for our initial conditions $(\tilde{p}^{\mathrm{SF}}_0, \tilde{e}^{\mathrm{SF}}_0)$ for our self-forced inspiral, and then root find for initial conditions $(\tilde{p}^{\mathrm{Ad}}_0, \tilde{e}^{\mathrm{Ad}}_0)$ that satisfy the simultaneous equations

$$\Omega^{\mathrm{SF}}_r(\tilde{p}^{\mathrm{SF}}_0, \tilde{e}^{\mathrm{SF}}_0) - \Omega^{(0)}_r(\tilde{p}^{\mathrm{Ad}}_0, \tilde{e}^{\mathrm{Ad}}_0) = 0, \tag{6.16a}$$

$$\Omega^{\mathrm{SF}}_\phi(\tilde{p}^{\mathrm{SF}}_0, \tilde{e}^{\mathrm{SF}}_0) - \Omega^{(0)}_\phi(\tilde{p}^{\mathrm{Ad}}_0, \tilde{e}^{\mathrm{Ad}}_0) = 0. \tag{6.16b}$$

Using this procedure to match the initial frequencies we find that the linear-in-t growth of the difference in the orbital phases is removed and the phase difference grows quadratically in t as expected – see Fig. 6.9.

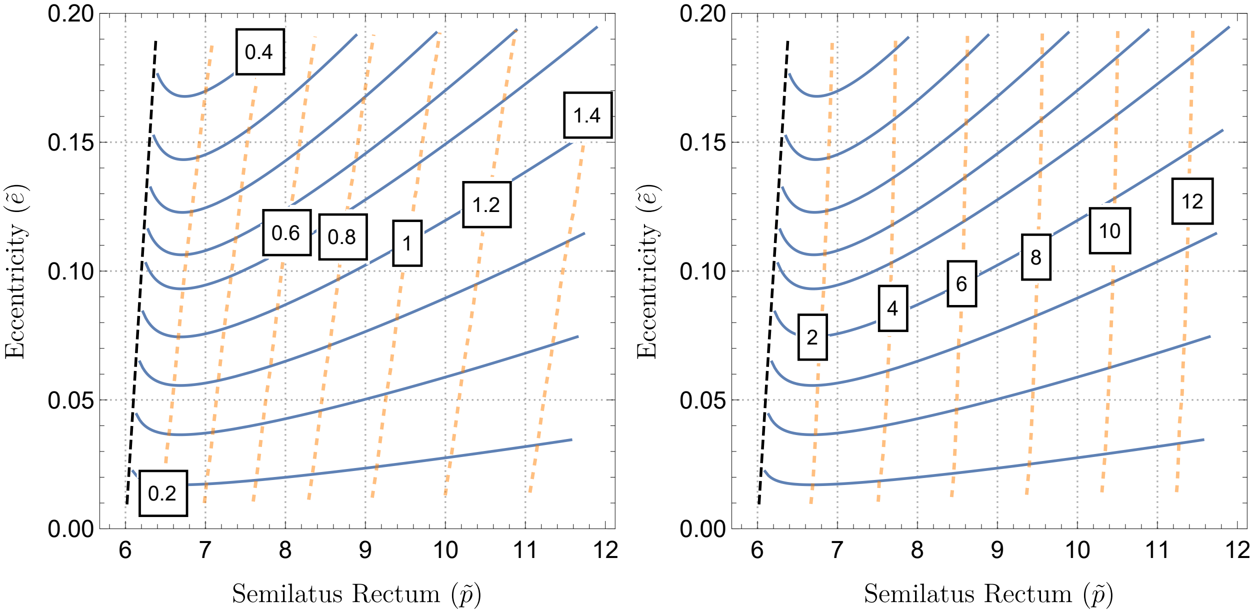
6.5.4 Comparing inspirals driven using radiation gauge and Lorenz gauge self-force in Schwarzschild spacetime

We now turn our attention to the special case of Schwarzschild $(a = 0)$, where we now have interpolated GSF models calculated in two different gauges. In addition to our outgoing radiation gauge self-force model, we make use of an interpolated Lorenz gauge self-force from Ref. [145], which is valid in the domain $6 \leq p \leq 12$ and $0 \leq e \leq 0.2$. We apply the same NIT procedure to inspirals driven by this force model, and find agreement with inspirals calculated in Paper I, up to the precision of the numerical solver.

To assess the accuracy of the dissipative self-force, we calculate the orbit averaged energy and angular momentum fluxes, and find that they agree with values from the literature with a relative error less than 10^{-3} for both models across the parameter space. To assess the accuracy of the conservative self-force, we calculate the periapsis advance in the circular orbit limit as outlined in [82] using the formula found in [85]. We find that both models show good agreement with the literature across the Lorenz gauge model's domain of validity, with the Lorenz gauge model producing errors less than 10^{-3} and the radiation gauge model producing relative errors less than 10^{-2}.

While we find good agreement between the two results for gauge invariant quantities, we see from Fig. 6.10 that the inspirals experience dramatically different conservative effects, depending on the gauge used. While in both cases, the conservative self-force acts against geodesic periapsis advance, we see that the evolution of $q_{r,0}$ depends heavily on the gauge involved, while the trajectories through p and e space are less affected. This is to be expected as the leading order averaged rates of change of p and e are related to the gauge invariant asymptotic fluxes, while the change in $q_{r,0}$ is induced entirely by the (gauge dependent) conservative self-force [99].

Just as when comparing adiabatic and self-forced inspirals it is important to match the initial frequencies (rather than the initial (p, e) values). We note that for the Lorenz gauge model, we must



(a) Outgoing radiation gauge inspirals

(b) Lorenz gauge inspirals

Figure 6.10: Sample Schwarzschild trajectories through $(\tilde{p}, \tilde{e})$ space using either a radiation gauge or a Lorenz gauge model, accompanied by contours denoting the change in $\tilde{q}_{r,0}$ (in radians) by the end of the inspiral if the inspiral had started in that point of the parameter space. While there are only slight differences in the $(\tilde{p}, \tilde{e})$ trajectories, there is a stark difference in the evolution of $\tilde{q}_{r,0}$ induced by each model.

account for the fact that the perturbed time coordinate, $\hat{t}$, is not asymptotically flat [190]. We can define an asymptotically flat time coordinate for Lorenz gauge inspirals via the following rescaling

$$t = (1 + \epsilon\alpha)\hat{t}. \tag{6.17}$$

where α is given by

$$\alpha(p, e) = -\frac{1}{2} h^{(1)}_{tt}(r \to \infty). \tag{6.18}$$

We make use of a code provided to us by S. Akcay to numerically calculate this quantity for Lorenz gauge values of p and e [97, 191]. Equation (6.17) means the perturbed Boyer-Lindquist frequencies must also be rescaled by:

$$\Omega^{(LG)}_r = (1 - \epsilon\alpha)\frac{\Upsilon^{(0)}_r + \epsilon\Upsilon^{(1)}_{r(LG)}}{\Upsilon^{(0)}_t + \epsilon\Upsilon^{(1)}_{t(LG)}} \quad \text{and} \quad \Omega^{(LG)}_\phi = (1 - \epsilon\alpha)\frac{\Upsilon^{(0)}_\phi + \epsilon\Upsilon^{(1)}_{\phi(LG)}}{\Upsilon^{(0)}_t + \epsilon\Upsilon^{(1)}_{t(LG)}}. \tag{6.19}$$

In the radiation gauge model, the corresponding subtleties have been dealt with by including the gauge completion corrections, so the frequencies can be calculated using Eq. (6.15) as before. Thus, we can choose a value for $\tilde{p}^{(LG)}_0$ and $\tilde{e}^{(LG)}_0$ in Lorenz gauge and root find for values of $\tilde{p}^{(RG)}_0$ and $\tilde{e}^{(RG)}_0$ in radiation gauge that satisfy:

$$\Omega^{(RG)}_r(\tilde{p}^{(RG)}_0, \tilde{e}^{(RG)}_0) - \Omega^{(LG)}_r(\tilde{p}^{(LG)}_0, \tilde{e}^{(LG)}_0) = 0, \tag{6.20a}$$

$$\Omega^{(RG)}_\phi(\tilde{p}^{(RG)}_0, \tilde{e}^{(RG)}_0) - \Omega^{(LG)}_\phi(\tilde{p}^{(LG)}_0, \tilde{e}^{(LG)}_0) = 0. \tag{6.20b}$$

This allows us to make comparisons between inspirals driven by self-force models calculated in different gauges. We use an inspiral driven by the Lorenz-gauge force model with initial conditions $(p_0, e_0) = (11, 0.18)$, mass ratio $\epsilon = 10^{-5}$ as our reference inspiral which should last just over two and a half years for a $10^6 M_\odot$ primary.

In Fig. 6.11, we see the difference in the phase of the waveform Φ as a function of time between the Lorenz gauge NIT inspiral, and a number of reference models. We make use of the relations between the NIT quantities and the waveform voices derived in Ref. [156] to find

$$\mathit{\Phi}_r = \tilde{q}_r - \Omega^{(0)}_r Z^{(0)}_t + \mathcal{O}(\epsilon) \quad \text{and} \quad \mathit{\Phi}_\phi = \tilde{\phi} - \Omega^{(0)}_\phi Z^{(0)}_t + \mathcal{O}(\epsilon). \tag{6.21}$$

We then feed the solutions for $\{\tilde{p}(t), \tilde{e}(t), \Phi_r(t), \Phi_\phi(t)\}$ into the FastEMRIWaveforms package to generate these eccentric Schwarzschild waveforms [135]. Finally, we make use of the `SimulationTools` Mathematica package to calculate the mismatches and decompose the waveforms into a single evolving amplitude $A(t)$ and phase $\Phi(t)$. This allows us to find the difference in the waveform phase $\Delta\Phi(t)$ between the Lorenz gauge inspiral and the other inspiral calculations. We use this as our point of comparison as the waveform phase is an observable and thus a gauge invariant quantity.

We note that in each case, we see constant error which gives way to quadratic growth with t just as in Fig. 6.9. As we discussed in Sec. 6.5.3, this shows that the initial frequencies were correctly matched. From the blue curve, we see that the NIT radiation gauge inspiral quickly goes out of phase with the Lorenz gauge NIT inspiral, resulting in a very large mismatch of 0.93. We found that the largest source of error here is due to interpolation error for in the adiabatic pieces of the NIT. Since these are related to the gauge invariant fluxes, these pieces should be identical in both models. As such, we can rectify this error by using the Lorenz gauge functions for the adiabatic pieces and continue to use the radiation gauge functions for the conservative pieces of the NIT equations of motion. The improvement is evident in the green curve, which shows much better agreement with the Lorenz gauge NIT inspiral, with the

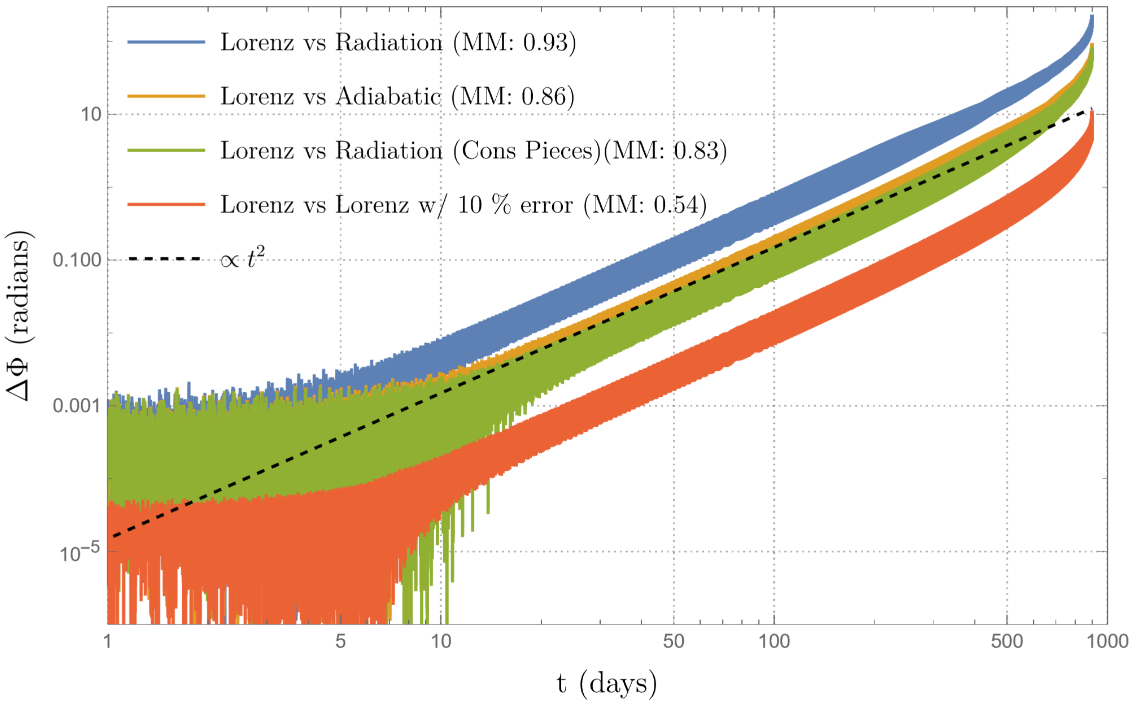


Figure 6.11: The difference in the waveform phase Φ for various inspirals as a function of t when compared to NIT inspiral driven by a Lorenz gauge self-force model, with initial conditions $(a, p_0, e_0) = (0, 11, 0.18)$, mass ratio $\epsilon = 10^{-5}$, viewing angles $\Theta = \pi/4$ and $\varPhi = 0$, and sampled every $\Delta t = 1M \sim 5s$. We also show the mismatch (MM) between the waveforms in each case. By matching the initial frequencies, we compare an inspiral calculated using a radiation gauge self-force model, an adiabatic inspiral, an inspiral with the adiabatic pieces of the Lorenz gauge model and conservative pieces from the radiation gauge model, and a Lorenz gauge model with a 10 % relative error added to each conservative piece. In all cases the difference grows quadratically in time. This plot suggests that post-adiabatic waveforms calculated using only the first-order self-force differ significantly depending on the gauge used.

mismatch falling to 0.83. However, it is only slightly better than matching an adiabatic inspiral (orange curve) using Eq. (6.16) resulting in a mismatch of 0.86. Both radiation gauge and adiabatic inspirals go out of phase by almost 100 radians by the time they reach the last stable orbit.

In order to rule out the possibility of interpolation error of the conservative effects being the primary cause of this difference, we repeat the Lorenz gauge inspiral, but this time we manually add a relative error of $\delta = 0.1$ to all of the conservative pieces of both the NIT equation of motion and our matching procedure for the initial conditions, e.g., $\dot{\tilde{q}}_r = \Upsilon_r^{(0)} + \epsilon \Upsilon_r^{(1)} \to \Upsilon_r^{(0)} + \epsilon \left(\Upsilon_r^{(1)} + \delta \Upsilon_r^{(1)} \right)$ etc. We note that this is an order of magnitude larger than the 10^{-2} error produced by the radiation gauge model when calculating the gauge invariant quasi-circular periapsis advance. From the red curve, we see that manually adding a constant 10% relative error results in phase difference and a mismatch (0.54) which is significantly smaller than what we observe between the two self-forced inspirals. This gives us confidence that this difference is not dominated by numerical error.

From these investigations, we infer that the trajectories driven using only the first order self-force are gauge dependent, and thus, so too are their waveforms. Since post-adiabatic waveforms are an observable quantity, this leads us to conclude that incorporating the orbit-averaged dissipative second-order self-force will be necessary to obtain gauge invariant, post-adiabatic waveforms. Moreover, since the difference between the Radiation and Lorenz gauge self-forced inspirals is of the same magnitude as the difference with the adiabatic inspiral, we further conclude that the impact of the orbit-averaged dissipative second-order self-force must be of a similar magnitude in at least one of the two gauges.

Chapter 7

Spherical self-forced inspirals into a rotating black hole

For this chapter, we look exclusively at the special case of spherical (circular and inclined) inspirals under the influence of the first-order gravitational self-force. We show how this affects the OG equations of motion in Sec. 7.1. We then introduce our interpolated model for the energy and angular momentum fluxes and gravitational self force in Sec. 7.2 , which combined with the OG equations of motion allows us calculate the very first spherical self-forced inspirals. To speed up these inspiral calculations, we explicitly derive the averaged equations of motion for the case of spherical Kerr inspirals by employing near-identity averaging transformations, an additional averaging transformation to parametrize our inspiral solutions in terms of Boyer-Lindquist coordinate time, and the two timescale expansion in Sec. 7.3. We describe our numerical implementation for calculating the various terms in the averaged equations of motion in Sec. 7.4. We present the results of this implementation in Sec. 7.5 and test the accuracy and speed-up, while also investigating the effect of using higher accuracy flux models and the impact of the first order self-force on the resulting inspiral trajectories.

7.1 Spherical osculating geodesics

We now restrict ourselves to the special case of spherical inspirals which means setting $e = 0$. As stated in Sec. 3.2.2, when $e = 0$ the rate of change $\dot{e} = 0$, and so we no longer have any dependence on the eccentricity. Thus we only need to track the evolution of $\vec{P} = \{p, x\}$. Similarly, the equations of motion no longer depend on the polar phase q_r, and so we only need to evolve the radial phase $\vec{q} = \{q_z\}$. As such, the equations of motion simplify to

$$\frac{dp}{d\lambda} = \epsilon F_p^{(1)}(a,p,x,q_z) + \epsilon^2 F_p^{(2)}(a,p,x,q_z), \tag{7.1a}$$

$$\frac{dx}{d\lambda} = \epsilon F_x^{(1)}(a,p,x,q_z) + \epsilon^2 F_x^{(2)}(a,p,x,q_z), \tag{7.1b}$$

$$\frac{dq_z}{d\lambda} = \Upsilon_z^{(0)}(a,p,x) + \epsilon f_z^{(1)}(a,p,x,q_z), \tag{7.1c}$$

$$\frac{dt}{d\lambda} = f_t^{(0)}(a,p,x,q_z). \tag{7.1d}$$

$$\frac{d\phi}{d\lambda} = f_\phi^{(0)}(a,p,x,q_z). \tag{7.1e}$$

Again, since we only have access to first order in the mass ratio GSF results, we set $F_p^{(2)} = F_x^{(2)} = 0$.

7.2 Interpolated models

7.2.1 Interpolated gravitational wave fluxes for spherical Kerr inspirals

Using a similar interpolation procedure to that outlined in Sec. 6.2, we will first interpolate the energy and angular momentum fluxes for spherical orbits. Since flux calculations are significantly cheaper than calculating the GSF, it is much more feasible to densely tile a large section of the parameter space with flux data than GSF data, resulting in better interpolation of the leading-order adiabatic effects. It will also allow us to carry out consistency checks on our GSF model.

We start by fixing the value of the spin parameter of the primary, which we choose to be $a = 0.9M$ and setting the eccentricity to zero. This reduces our parameter space to two parameters; the semilatus rectum p and the inclination x. We then define a parameter y_{sph} using the p and the position of the innermost stable spherical orbit r_{ISSO}. We chose y_{sph} to be

$$y_{sph} = \sqrt{\frac{r_{\text{ISSO}}(a,x)}{p}}. \tag{7.2}$$

Tiling the parameter space with y_{sph} instead of p will concentrate more points near the ISSO where the fluxes and the GSF experience the most variation. We let y_{sph} range from 0 to 1 and x range from -1 to 1 thus covering all inclinations for both prograde and retrograde orbits. We then calculate the GSF on a 18×19 grid of Chebyshev nodes, with the y_{sph} values given by the roots of the 18th order polynomial and the x values given by the roots of the 19th order polynomial.

At each point on this grid, the energy and angular momentum flux both at infinity and at the event horizon was calculated using a modified version of the code from Ref. [100] for spherical orbits. From these fluxes, one can infer the leading order orbit averaged rate of change of the energy and angular momentum of the secondary via the following balance laws:

$$\left\langle \frac{d\mathcal{E}}{dt} \right\rangle = -\epsilon(\mathcal{F}^{\infty}_{\mathcal{E}} + \mathcal{F}^{\mathcal{H}}_{\mathcal{E}}) + \mathcal{O}(\epsilon^2) \tag{7.3a}$$

$$\left\langle \frac{d\mathcal{L}}{dt} \right\rangle = -\epsilon(\mathcal{F}^{\infty}_{\mathcal{L}} + \mathcal{F}^{\mathcal{H}}_{\mathcal{L}}) + \mathcal{O}(\epsilon^2) \tag{7.3b}$$

We could interpolate the rates of change of energy and angular momentum, but we find it more convenient to work with the orbital elements p and x. As such, we find their rates of change via the chain rule:

$$\left\langle \frac{dp}{dt} \right\rangle = \frac{\partial p}{\partial \mathcal{E}} \left\langle \frac{d\mathcal{E}}{dt} \right\rangle + \frac{\partial p}{\partial \mathcal{L}} \left\langle \frac{d\mathcal{L}}{dt} \right\rangle = \epsilon \Gamma^{(1)}_{p} \tag{7.4a}$$

$$\left\langle \frac{dx}{dt} \right\rangle = \frac{\partial x}{\partial \mathcal{E}} \left\langle \frac{d\mathcal{E}}{dt} \right\rangle + \frac{\partial x}{\partial \mathcal{L}} \left\langle \frac{d\mathcal{L}}{dt} \right\rangle = \epsilon \Gamma^{(1)}_{x}. \tag{7.4b}$$

The partial derivatives can be found using the analytic expressions for $\mathcal{E}(p,x)$ and $\mathcal{L}(p,x)$ to construct the Jacobian

$$J = \begin{bmatrix} \frac{d\mathcal{E}}{dp} & \frac{d\mathcal{E}}{dx} \\ \frac{d\mathcal{L}}{dp} & \frac{d\mathcal{L}}{dx} \end{bmatrix}. \tag{7.5}$$

This can then be inverted to give

$$J^{-1} = \begin{bmatrix} \frac{dp}{d\mathcal{E}} & \frac{dp}{d\mathcal{L}} \\ \frac{dx}{d\mathcal{E}} & \frac{dx}{d\mathcal{L}} \end{bmatrix} = \frac{1}{\det J} \begin{bmatrix} \frac{d\mathcal{L}}{dx} & -\frac{d\mathcal{E}}{dx} \\ -\frac{d\mathcal{L}}{dp} & \frac{d\mathcal{E}}{dp} \end{bmatrix}. \tag{7.6}$$

To improve the accuracy of our interpolation, we rescale the data for $\Gamma_p^{(1)}$ and $\Gamma_x^{(1)}$ by a factor of $p^3(1-y_{sph})$ and $p^{11/2}(1-x^2)$ respectively. This scaling comes from the leading order PN term for $\Gamma_p^{(1)}$ and $\Gamma_x^{(1)}$, times a term that is zero for the limiting cases of either the separatrix or the equatorial plane respectively. Finally, we take a discrete cosine transformation of the data on our Chebyshev grid to obtain the Chebyshev polynomial coefficients $C_{p/x}^{ij}$. Summing these coefficients together with Chebyshev polynomials gives us the following approximate expressions for $\Gamma_p^{(1)}$ and $\Gamma_x^{(1)}$:

$$\Gamma_p^{(1)} = \frac{1}{p^3(1-y_{sph})}\sum_{i=0}^{17}\sum_{j=0}^{18} C_p^{ij} T_i\left(2y-1\right) T_j\left(x\right), \tag{7.7a}$$

$$\Gamma_x^{(1)} = \frac{1}{p^{11/2}(1-x^2)}\sum_{i=0}^{17}\sum_{j=0}^{18} C_x^{ij} T_i\left(2y-1\right) T_j\left(x\right). \tag{7.7b}$$

Using the largest coefficient for $i = 17$ and $j = 18$ to estimate the relative error, we infer that these interpolants should have a relative error of $\sim 10^{-6}$. This should also be verified against flux data calculated from an independent code to more accurately determine the relative accuracy of the interpolant.

7.2.2 Interpolated gravitational self-force model for spherical Kerr inspirals

We now use this same interpolation scheme to create a model for the gravitational self-force that is continuous throughout the parameter space and fast to evaluate which can be used with the OG equations to calculate spherical Kerr inspirals.

However, given the cost of our GSF code, we restrict ourselves to a 2D slice of the EMRI parameter space using only a few hundred points. Once again, we restrict $a = 0.9$ and let y_{sph} range from $y_{\mathsf{min}} = 0.1$ to $y_{\mathsf{max}} = 1$, but instead of x, we opt to tile in $z_-^2 = 1 - x^2$ and let $z^2_{-,\mathsf{min}} = 0$ to $z^2_{-,\mathsf{max}} = 0.5$. This allows us to cover moderate inclination angles $\theta_{\mathsf{inc}} \leq 45^\circ$ while excluding high inclinations, where the mode sum regularization for the GSF becomes prohibitively expensive, in a manner that is convenient for Chebyshev interpolation.

We define parameters u and v which cover this parameter space as they range from $(-1, 1)$

$$u := \frac{y_{sph} - (y_{\mathsf{min}} + y_{\mathsf{max}})/2}{(y_{\mathsf{min}} - y_{\mathsf{max}})/2} \text{ and,} \tag{7.8a}$$

$$v := \frac{z_-^2 - (z_{-,\mathsf{min}} + z^2_{-,\mathsf{max}})/2}{(z^2_{-,\mathsf{min}} - z^2_{-,\mathsf{max}})/2}. \tag{7.8b}$$

We then calculate the GSF on a 18×9 grid of Chebyshev nodes, with the u values given by the roots of the 18th order polynomial and the v values given by the roots of the 9th order polynomial. At each point on our grid, we Fourier decompose each component of the force with respect to the polar action angle q_z. We then multiply the data for each Fourier coefficient by a factor of $(1-y_{sph})^2$, as we find that this smooths the behaviour of the force near the separatrix and improves the accuracy of our interpolation. Next, we use Chebyshev polynomials to interpolate each Fourier coefficient across the (u, v) grid. We then resum the modes to reconstruct our interpolated gravitational self force model:

$$a_\mu = \sum_{m=0}^{24} \frac{C_\mu^m(y_{sph}, z_-^2)\cos(mq_z) + D_\mu^m(y_{sph}, z_-^2)\sin(mq_z)}{(1-y_{sph})^2}, \tag{7.9a}$$

$$C_{\mu}^{m}(y,e) = \sum_{i=0}^{17}\sum_{j=0}^{8} C_{\mu}^{mij} T_i\left(u\right) T_j\left(v\right), \tag{7.9b}$$

$$D_{\mu}^{m}(y,e) = \sum_{i=0}^{17}\sum_{j=0}^{8} D_{\mu}^{mij} T_i\left(u\right) T_j\left(v\right). \tag{7.9c}$$

We note that this choice of rescaling causes each component to become singular at the ISSO, and while the components of the GSF change rapidly as one approaches the ISSO, we still expect them to be finite at the ISSO. A greater understanding of the analytic structure of the GSF in this region would greatly improve this and any future interpolated GSF models.

We note that the GSF should satisfy the orthogonality condition with the geodesic four-velocity, i.e., $a_{\mu}u^{\mu} = 0$. Interpolation will bring with it a certain amount of error which can cause this condition to be violated. Since the OG equations are derived assuming this condition to be true, we project the force so that this condition is always satisfied, i.e.,

$$a_{\mu}^{\perp} = a_{\mu} + a_{\nu}u^{\nu}u_{\nu}. \tag{7.10}$$

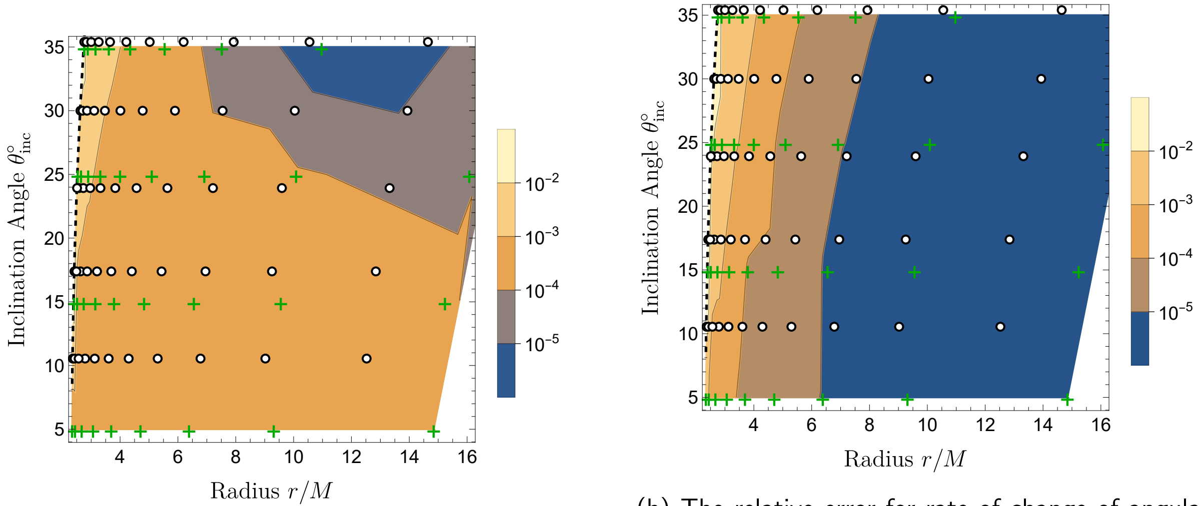
To verify the accuracy of our interpolated model, we employ the flux balance laws to compare the local energy and angular momentum lost by the secondary against the energy and angular momentum fluxes radiated at infinity and down the horizon:

$$\frac{\Upsilon_{\tau}^{(0)}}{\Upsilon_{t}^{(0)}}\left\langle a_{t}^{\perp}\right\rangle = \left(\mathcal{F}_{\mathcal{E},\infty} + \mathcal{F}_{\mathcal{E},H}\right), \tag{7.11a}$$

$$\frac{\Upsilon_{\tau}^{(0)}}{\Upsilon_{t}^{(0)}}\left\langle a_{\phi}^{\perp}\right\rangle = -\left(\mathcal{F}_{\mathcal{L},\infty} + \mathcal{F}_{\mathcal{L},H}\right). \tag{7.11b}$$

The relative error in the fluxes across the parameter space is shown in Fig. 7.1. The white circles indicate the points used to generate the interpolated GSF model, whereas the green crosses indicate the points in the parameter space where the fluxes were calculated and thus where the comparisons were made. We find that the relative error in the fluxes is less than 10^{-2} across the parameter space and is typically of the order $10^{-3} - 10^{-4}$ in most of the parameter space, with the ecxception being in a region very close to the ISSO. This is comparable to previous methods which required tens of thousands of points to achieve the same level of accuracy [70, 145]. While this would not be sufficiently accurate for LISA data analysis which would require fluxes accurate to $\sim 1/\epsilon$, such an interpolating error may be permissible for the post-adiabatic contributions of the first order GSF [70].

We can then use this model in conjunction with the OG equations of motion (7.1) to calculate inspiral trajectories. However, we find these trajectories take minutes to hours to compute due to the need to resolve hundreds of thousands of orbital cycles. We will now look to leverage averaging transformations which can remove the dependence of the orbital phases from the equations of motion while retaining the accuracy required to produce post-adiabatic EMRI waveforms.



(a) The relative error for rate of change of energy.

(b) The relative error for rate of change of angular momentum.

Figure 7.1: The relative error in the rate of energy and angular momentum loss between the asymptotic fluxes and the interpolated GSF model. The white dots indicate the locations of the data points the model is interpolating. The green crosses indicate the locations of the flux calculations and thus where the comparisons are made. The relative error is always $< 10^{-2}$ and is typically of the order $\sim 10^{-3}$–10^{-4}.

7.3 Averaged equations of motion for spherical Kerr inspirals

When specialized to spherical Kerr inspirals, the NIT equations of motion, as parametrized by Mino time λ, take the form:

$$\frac{d\tilde{p}}{d\lambda} = \epsilon \tilde{F}^{(1)}_p(a,\tilde{p},\tilde{x}) + \epsilon^2 \tilde{F}^{(2)}_p(a,\tilde{p},\tilde{x}), \tag{7.12a}$$

$$\frac{d\tilde{x}}{d\lambda} = \epsilon \tilde{F}^{(1)}_x(a,\tilde{p},\tilde{x}) + \epsilon^2 \tilde{F}^{(2)}_x(a,\tilde{p},\tilde{x}), \tag{7.12b}$$

$$\frac{d\tilde{q}_z}{d\lambda} = \Upsilon^{(0)}_z(a,\tilde{p},\tilde{x}) + \epsilon \Upsilon^{(1)}_z(a,\tilde{p},\tilde{x}), \tag{7.12c}$$

$$\frac{d\tilde{t}}{d\lambda} = \Upsilon^{(0)}_t(a,\tilde{p},\tilde{x}) + \epsilon \Upsilon^{(1)}_t(a,\tilde{p},\tilde{x}), \tag{7.12d}$$

$$\frac{d\tilde{\phi}}{d\lambda} = \Upsilon^{(0)}_\phi(a,\tilde{p},\tilde{x}) + \epsilon \Upsilon^{(1)}_\phi(a,\tilde{p},\tilde{x}). \tag{7.12e}$$

The leading order terms in each equation of motion are simply the original function averaged over a single geodesic orbit, i.e.,

$$\tilde{F}^{(1)}_p = \left\langle F^{(1)}_p \right\rangle, \quad \tilde{F}^{(1)}_x = \left\langle F^{(1)}_x \right\rangle, \quad \Upsilon^{(1)}_z = \left\langle f^{(1)}_z \right\rangle, \tag{7.13a-c}$$

$$\Upsilon^{(0)}_t = \left\langle f^{(0)}_t \right\rangle \quad \Upsilon^{(0)}_\phi = \left\langle f^{(0)}_\phi \right\rangle, \tag{7.13d-e}$$

where $\Upsilon^{(0)}_t$ and $\Upsilon^{(0)}_\phi$ are the Mino-time t and ϕ fundamental frequencies which are known analytically [174]. The remaining terms are more complicated and require Fourier decomposing the original functions and their derivatives with respect to to the orbital elements (p, x). To express the result, we define the operator

$$\mathcal{N}(A) = \sum_{m\neq 0} \frac{-1}{\Upsilon_z^{(0)}} \left[A_m f^{(1)}_{z,-m} - \frac{i}{m} \left(\frac{\partial A_m}{\partial \tilde{p}} F^{(1)}_{p,-m} + \frac{\partial A_m}{\partial \tilde{x}} F^{(1)}_{x,-m} - \frac{A_m}{\Upsilon_z^{(0)}} \left(\frac{\partial \Upsilon_z^{(0)}}{\partial \tilde{p}} F^{(1)}_{p,-m} + \frac{\partial \Upsilon_z^{(0)}}{\partial \tilde{x}} F^{(1)}_{x,-m} \right) \right) \right] \tag{7.14}$$

With this in hand, the remaining terms in the equations of motion are found to be

$$\begin{aligned} \tilde{F}_p^{(2)} &= \left\langle F_p^{(2)} \right\rangle + \mathcal{N}(F_p^{(1)}), & \tilde{F}_x^{(2)} &= \left\langle F_x^{(2)} \right\rangle + \mathcal{N}(F_x^{(1)}), \\ \Upsilon_t^{(1)} &= \mathcal{N}(f_t^{(0)}), & \Upsilon_\phi^{(1)} &= \mathcal{N}(f_\phi^{(0)}). \end{aligned} \tag{7.14a-d}$$

Note that since we currently only have access to the first order GSF, we set $\left\langle F_p^{(2)} \right\rangle = \left\langle F_x^{(2)} \right\rangle = 0$. Combining these results with Eqs. (4.63) one can find the NIT equations of motion parametrized by Boyer-Lindquist time t for the phases $\vec{\varphi} = \{\varphi_z, \varphi_\phi\}$ as defined in Eq. (4.62b) and orbital elements $\vec{\mathcal{P}} = \{p_\varphi, x_\varphi\}$ as defined in Eq. (4.62a)

$$\frac{dp_\varphi}{dt} = \epsilon \Gamma_p^{(1)}(a, p_\varphi, x_\varphi) + \epsilon^2 \Gamma_p^{(2)}(a, p_\varphi, x_\varphi), \tag{7.15a}$$

$$\frac{dx_\varphi}{dt} = \epsilon \Gamma_x^{(1)}(a, p_\varphi, x_\varphi) + \epsilon^2 \Gamma_x^{(2)}(a, p_\varphi, x_\varphi), \tag{7.15b}$$

$$\frac{d\varphi_z}{dt} = \Omega_z^{(0)}(a, p_\varphi, x_\varphi) + \epsilon \Omega_z^{(1)}(a, p_\varphi, x_\varphi), \tag{7.15c}$$

$$\frac{d\varphi_\phi}{dt} = \Omega_\phi^{(0)}(a, p_\varphi, x_\varphi) + \epsilon \Omega_\phi^{(1)}(a, p_\varphi, x_\varphi). \tag{7.15d}$$

The leading order terms in these equations are given by

$$\Gamma_p^{(1)} = \tilde{F}_p^{(1)} / \Upsilon_t^{(0)}, \quad \Gamma_x^{(1)} = \tilde{F}_x^{(1)} / \Upsilon_t^{(0)}, \tag{7.16a-b}$$

$$\tilde{\Omega}_z^{(0)} = \Upsilon_z^{(0)} / \Upsilon_t^{(0)}, \quad \tilde{\Omega}_\phi^{(0)} = \Upsilon_\phi^{(0)} / \Upsilon_t^{(0)}. \tag{7.16c-d}$$

The sub-leading terms are given by

$$\Gamma_p^{(2)} = \frac{1}{\Upsilon_t^{(0)}} \left(\tilde{F}_p^{(2)} - \Upsilon_t^{(1)} \Gamma_p^{(1)} \right), \tag{7.17a}$$

$$\Gamma_x^{(2)} = \frac{1}{\Upsilon_t^{(0)}} \left(\tilde{F}_x^{(2)} - \Upsilon_t^{(1)} \Gamma_x^{(1)} \right), \tag{7.17b}$$

$$\Omega_z^{(1)} = \frac{1}{\Upsilon_t^{(0)}} \left(\Upsilon_z^{(1)} - \Upsilon_t^{(1)} \Omega_z^{(0)} \right), \tag{7.17c}$$

$$\Omega_\phi^{(1)} = \frac{1}{\Upsilon_t^{(0)}} \left(\Upsilon_\phi^{(1)} - \Upsilon_t^{(1)} \Omega_\phi^{(0)} \right). \tag{7.17d}$$

Finally, using the two timescale expansion equations of motion, as defined in Eq. 4.70, the adiabatic equations of motion are given by

$$\frac{dp_\varphi^{(0)}}{d\mathcal{T}} = \Gamma_p^{(1)}, \quad \frac{dx_\varphi^{(0)}}{d\mathcal{T}} = \Gamma_x^{(1)}, \tag{7.18a-b}$$

$$\frac{d\varphi_z^{(0)}}{d\mathcal{T}} = \Omega_z^{(0)}, \quad \frac{d\varphi_\phi^{(0)}}{d\mathcal{T}} = \Omega_\phi^{(0)}. \tag{7.18c-d}$$

The post-adiabatic contributions to the equations of motion are given by

$$\frac{dp_\varphi^{(1)}}{d\mathcal{T}} = \Gamma_p^{(2)} + p_\varphi^{(1)}\frac{\partial\Gamma_p^{(1)}}{\partial p_\varphi^{(0)}} + x_\varphi^{(1)}\frac{\partial\Gamma_p^{(1)}}{\partial x_\varphi^{(0)}}, \tag{7.19a}$$

$$\frac{dx_\varphi^{(1)}}{d\mathcal{T}} = \Gamma_x^{(2)} + p_\varphi^{(1)}\frac{\partial\Gamma_x^{(1)}}{\partial p_\varphi^{(0)}} + x_\varphi^{(1)}\frac{\partial\Gamma_x^{(1)}}{\partial x_\varphi^{(0)}}, \tag{7.19b}$$

$$\frac{d\varphi_z^{(1)}}{d\mathcal{T}} = \Omega_p^{(1)} + p_\varphi^{(1)}\frac{\partial\Omega_z^{(0)}}{\partial p_\varphi^{(0)}} + x_\varphi^{(1)}\frac{\partial\Omega_z^{(0)}}{\partial x_\varphi^{(0)}}, \tag{7.19c}$$

$$\frac{d\varphi_\phi^{(1)}}{d\mathcal{T}} = \Omega_\phi^{(1)} + p_\varphi^{(1)}\frac{\partial\Omega_\phi^{(0)}}{\partial p_\varphi^{(0)}} + x_\varphi^{(1)}\frac{\partial\Omega_\phi^{(0)}}{\partial x_\varphi^{(0)}}, \tag{7.19d}$$

Solutions for post-adiabatic inspirals can be obtained by solving Eqs. (7.18) and (7.19) simultaneously and using Eqs. (4.69) along with a value for the mass ratio ϵ to recover $p_\varphi(t), x_\varphi(t), \varphi_z(t)$ and $\varphi_z(t)$. With the averaged equations for spherical Kerr inspirals in hand, we will now outline our numerical implementation for rapidly computing spherical self-forced inspirals.

7.4 Implementation

Combining the interpolated GSF model along with our action angle formulation of the OG equations provides us with all the information required to calculate the NIT and TTE equations of motion. We first evaluate and interpolate the various terms in the NIT and TTE equations of motion across the parameter space. While this offline process can be expensive, it only need to be completed once. Once completed, the online process of calculating self-forced inspirals can be completed in less than a second.

7.4.1 Offline Steps

The offline calculation consists of the following steps.

1. We start by selecting a grid which covers the parameter space. We chose y values between 0.099 and 0.999 in 451 equally spaced steps and z_-^2 values from 0.002 to 0.5 in 250 equally spaced steps (112,750 points).[1]

2. For each point in the parameter space (a, y_{sph}, z_-^2) we evaluate the functions $F^{(1)}_{p\backslash x}$, $f^{(1)}_z$ and$f^{(0)}_{t\backslash\phi}$ along with their derivatives with respect to p and x for 49 equally spaced values of q_z from 0 to 2π.

3. We then perform a fast Fourier transform on the output data to obtain the Fourier coefficients of the forcing functions and their derivatives.

---

[1] Evaluating the NIT functions is computationally cheap so using a dense grid does not significantly increase the computational burden. Using an equally spaced grid also allows us to use Mathematica's default Hermite polynomial interpolation method for convenience of implementation. The grid spacing is chosen to be sufficiently dense that interpolation error is a negligible source of error for our comparisons between the OG, NIT and TTE inspirals, though a less dense grid may also achieve this.

4. With these, we then use Eqs. (7.13a-c), (7.13d-e), (7.14) and (7.14a-d) to construct $\tilde{F}^{(1\backslash 2)}_{p\backslash x}$, $\partial\tilde{F}^{(1)}_{p\backslash x}/\partial p$, $\partial\tilde{F}^{(1)}_{p\backslash x}/\partial x$, and $\Upsilon^{(1)}_{t\backslash z\backslash\phi}$ at a given point in the parameter space. Note that all other terms needed for the NIT and TTE can be derived from these terms or are already known analytically.
5. We also use Eqs. (4.55) and (4.56) to construct the Fourier coefficients of the first order transformation functions $Y^{(1)}_{p\backslash x}$ and $X^{(1)}_{z}$.
6. We then repeat this procedure across the parameter space for each point in our grid.
7. Finally we interpolate the values for $\tilde{F}^{(1\backslash 2)}_{p\backslash x}$, $\partial\tilde{F}^{(1)}_{p\backslash x}/\partial p$, $\partial\tilde{F}^{(1)}_{p\backslash x}/\partial p$, and $\Upsilon^{(1)}_{t\backslash z\backslash\phi}$ along with the coefficients of $Y^{(1)}_{p\backslash x}$ across this grid using Hermite interpolation and store the interpolants for future use.

We implemented the above algorithm in Mathematica 12.2 and find, parallelized across 20 CPU cores, the calculation takes about 5 hours to complete. Since these offline steps need only be completed once, this is a comparatively small price to pay.

7.4.2 Online Steps

The online steps are required for every inspiral calculation are, by contrast, computationally inexpensive and are as follows:

1. We load in the interpolants for $\tilde{F}^{(1\backslash 2)}_{p\backslash x}$, and $\Upsilon^{(1)}_{t\backslash z\backslash\phi}$, and define the NIT equations of motion. For TTE inspirals, one also needs to load in the interpolated derivatives $\partial\tilde{F}^{(1)}_{p\backslash x}/\partial p$ and $\partial\tilde{F}^{(1)}_{p\backslash x}/\partial x$ then define the TTE equations of motion.
2. In order to make comparisons between OG, NIT, and TTE inspirals, we also load interpolants of the Fourier coefficients of $\breve{Y}^{(1)}_{p/x}$ and Eqs. (4.1), (4.62) and (4.66) to construct first order near-identity transformations.
3. We state the initial conditions of the inspiral $(p(0), x(0), q_z(0))$ and use the NIT to leading order in the mass ratio to transform these into initial conditions for the NIT/TTE equations of motion, i.e., $(p_\varphi(0), x_\varphi(0), \varphi_z(0))$.
4. We then evolve the NIT or TTE equations of motion using an ODE solver (in our case Mathematica's `NDSolve`).

As with the offline steps we implement the online steps in Mathematica. Note that steps (ii) and (iii) are only necessary because we want to make direct comparisons between NIT and OG inspirals with the same initial conditions. In general, the difference between the NIT and OG variables will always be $\mathcal{O}(\epsilon)$, and so performing the NIT transformation or inverse transformation to greater than zeroth order in mass ratio will not be necessary when producing waveforms to post adiabatic order, i.e. with phases accurate to $\mathcal{O}(\epsilon)$.

Using this procedure, one can compute year-long inspiral trajectories in less than a second. By comparison, using the quasi-Keplerian OG equations with sufficiently high relative and absolute accuracy requirements takes on the order of hours to compute the same trajectory. Since we do not yet have interpolated Teukolsky amplitudes for Kerr inspirals, we use the semi-relativistic quadrupole formula as outlined in 5.1.2 to generate the waveform strain for the NIT, TTE and OG inspiral trajectories, sampled at 10s intervals. In order to quantify the effect that using each of these methods has on the resulting waveform, we calculate calculate the fractional overlap $\mathcal{O}(h_1, h_2)$ between two waveform strains h_1 and h_2 we using `WaveformMatch` function from the `SimulationTools` package [186] and use the approximate LISA sensitivity curve as our noise curve from Eq. (5.29) [185].

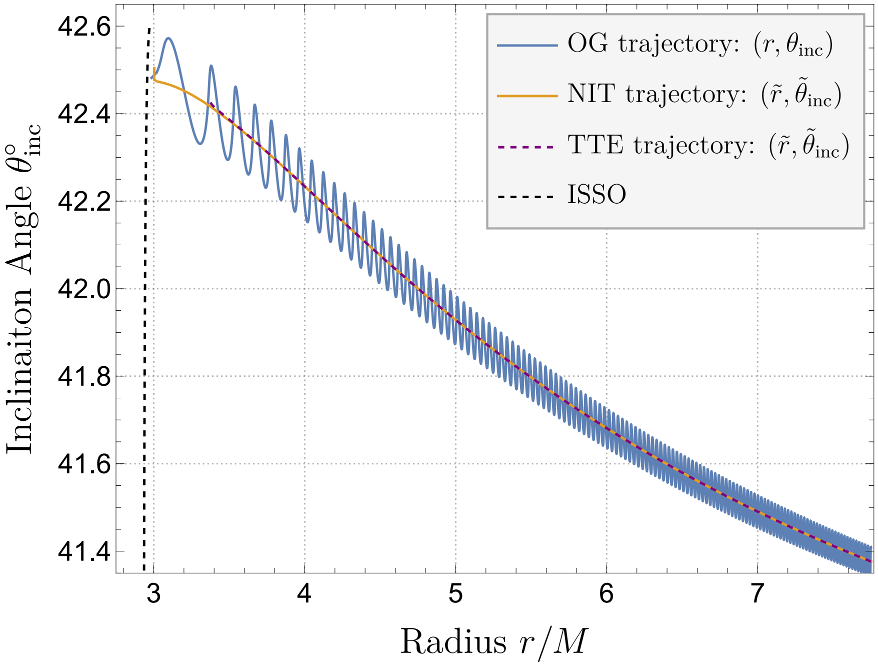


Figure 7.2: The trajectory through (r, θ_{inc}) space for an inspiral with $\epsilon = 10^{-2}$, $a = 0.9$, and initial conditions $(p_0 = 7.75, x_0 = 0.75)$. We use such a large mass ratio to show off the orbital timescale oscillations one encounters when using the OG equations. Using the NIT or TTE equations of motion averages out these oscillations and results in almost identical inspirals. We also see that the TTE equations of motion break down further from the ISSO than the OG or NIT equations of motion.

7.5 Results

7.5.1 OG vs NIT and TTE inspirals

To test the accuracy of our NIT and TTE implementations, we compare inspirals calculated using the OG equations of motion against inspirals calculated using the NIT or TTE equations of motion. We chose a case study of a typical system EMRI with a $10^6 M_\odot$ primary and initial conditions for inclination of $x = 0.75$ and semilatus rectum of $p = 7.75$. This highly inclined, strong field inspiral should provide a strenuous test of our numerical implementations. We also set the initial phases $q_z(0) = \phi(0) = 0$ for simplicity.

Fig. 7.2 shows the trajectories produced by each method. We have purposely chosen an unusually large mass ratio for an EMRI of $\epsilon = 10^{-2}$ so that the orbital timescale oscillations are clearly visible on the plot. We see that for each trajectory the orbital separation decreases substantially, while the orbit becomes slightly more inclined. However, the OG trajectory oscillates on the orbital timescale, while the NIT and TTE trajectories faithfully capture the average evolution of the trajectory. While the NIT and TTE trajectories may appear at first glance to be identical, it is important to note that the TTE equations of motion break down sooner than the NIT equations of motion. This is due to the evolution of $p_\varphi^{(0)}$ and $x_\varphi^{(0)}$ reaching the location of the ISSO sooner than p_φ and x_φ. This is not as important of an issue as one might expect, as this close to the ISSO one should swap over to a "transition to plunge" scheme [138–141].

Fig. 7.3 shows the difference in the phases between the OG trajectory and the NIT and TTE trajectories for a year long inspiral with a mass ratio of $\epsilon = 10^{-5}$. In both cases, we find that the difference in the phases stays below 10^{-3} throughout the inspiral, spiking only when the inspiral approaches the ISSO

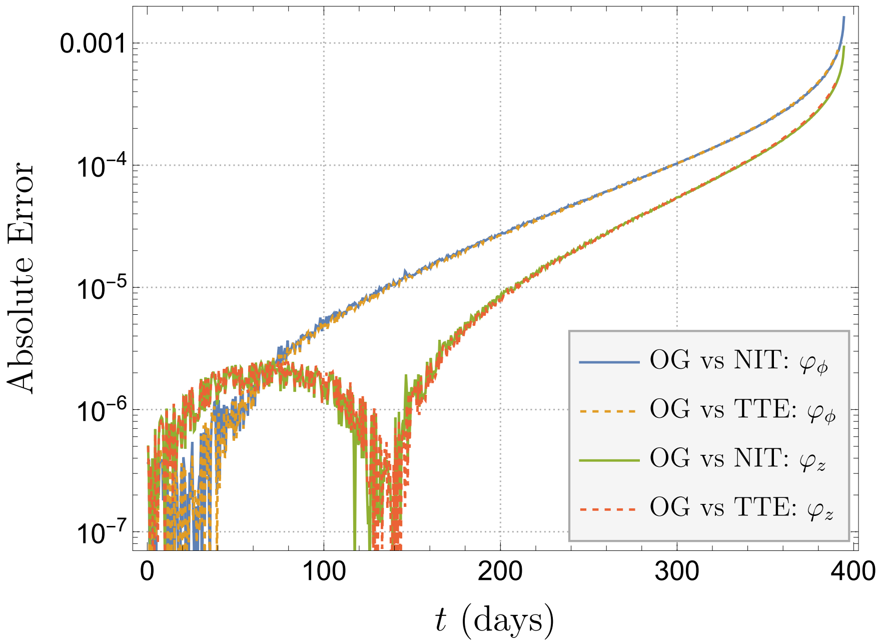


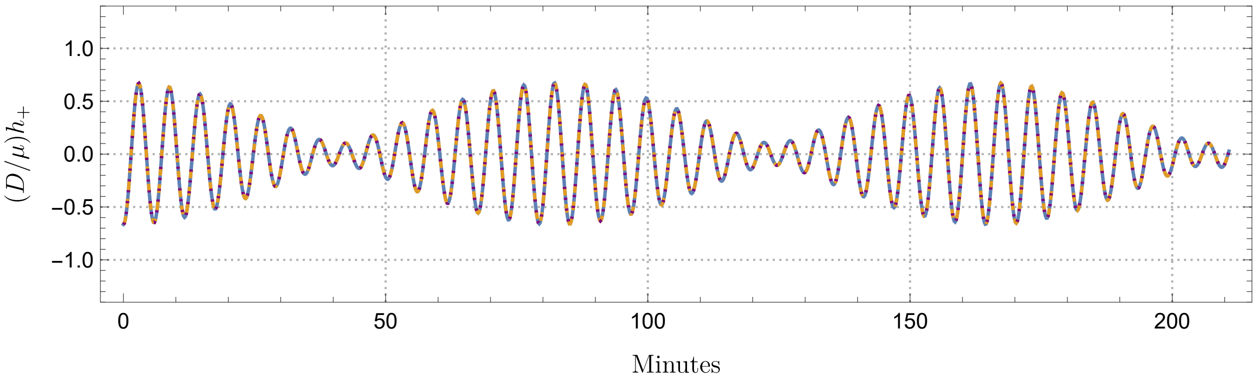
Figure 7.3: The difference in the orbital phases for a spherical Kerr inspiral with $\epsilon = 10^{-5}$ and $a = 0.9M$ when using the OG equations verse the NIT or TTE equations of motion. In both cases, the difference stays small throughout the inspiral, only becoming large as the secondary approaches the ISSO when the adiabatic assumption breaks down.

where we expect the adiabatic assumption implicit in the OG, NIT, and TTE equations of motion to start to break down. Even then, the difference in the phases is substantially lower than subradian accuracy requirement needed for LISA data analysis. We have also found that the growth in the error over time is most closely correlated with the interpolation error for the terms in the NIT and TTE equations of motion, and so interpolating on a denser grid should reduce this error even more.

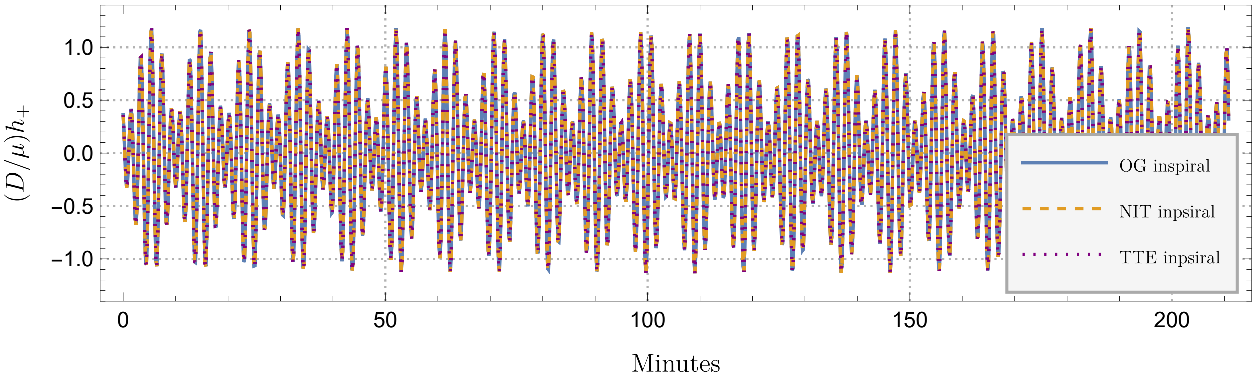
We then use the flat semi-relativistic approximation to produce quadrupole waveforms which can be seen in Figs. 7.4a and 7.4b. These figures show the first and last 3 and a half hours of the waveforms produced by the OG and NIT trajectories respectively. From the plots it is clear that the two waveforms overlap heavily, with the differences between them being miniscule, as demonstrated by the green curve. The waveform mismatch between the OG waveforms and both the NIT and TTE waveforms only $\sim 2.5 \times 10^{-8}$. This means that they meet the indistinguishability criteria from the OG waveforms for signal-to-noise ratios (SNRs) of up to ≈ 4500 according to (5.36).

We note that formally both the NIT and the TTE should induce an error in both the phases and the orbital elements that scales linearly with the mass ratio. To ensure that our implementation is converging correctly, we evolve inspirals from initial conditions $p(0) = 5$ and $x(0) = 0.75$ until they reach $p = 4$ with different values of the mass ratios and find the largest error between the OG and NIT/TTE phases and orbital elements found during the last 3 orbital cycles of the inspirals. The results of this test can be seen in Fig. 7.5. We see that for larger mass ratios the errors converge linearly as expected. However, for mass ratios smaller than 10^{-3}, the formal error in the NIT and TTE is no longer the dominant source of error for the evolution of the phases. Instead the error is dominated by either interpolation error or the error in the ODE solver. While both of these sources can be further suppressed at the cost of more expensive offline and online steps respectively, we find this to be unnecessary for producing trajectories accurate enough for LISA data science.

Intuitively, one would expect the error in our averaging procedure to scale with the size of the



(a) The first 3.5 hours of the waveform.



(b) The last 3.5 hours of the waveform.

Figure 7.4: The "semi-relativistic" quadrupole waveform strains for an ERMI with $\epsilon = 10^{-5}$, $p(0) = 7.75$ and $x(0) = 0.75$ as viewed from edge on sampled once every 10s. The blue curve is the waveform generated from the OG inspiral, the dashed yellow curve is the waveform generated from the NIT inspiral, and the purple dotted curve is the waveform generated from the TTE inspiral.

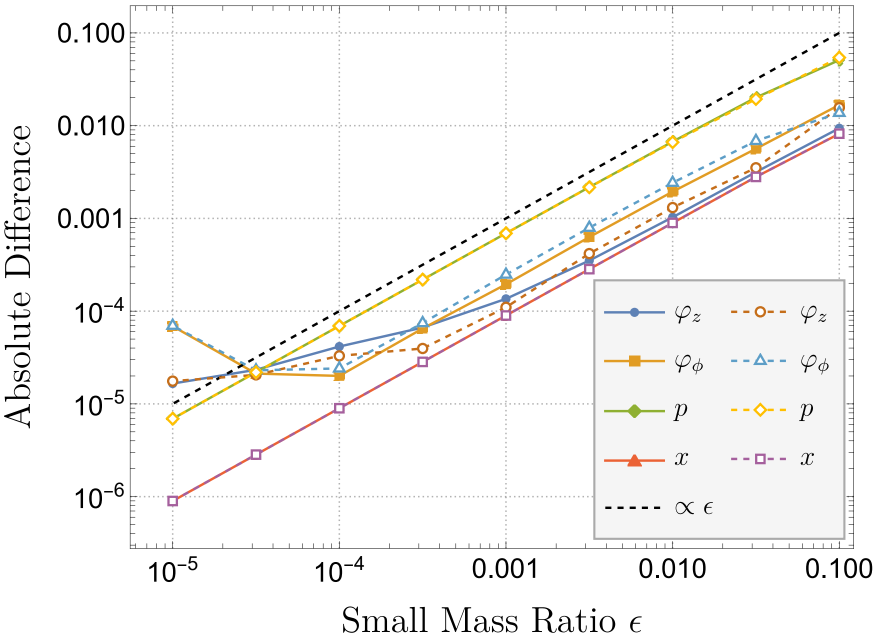
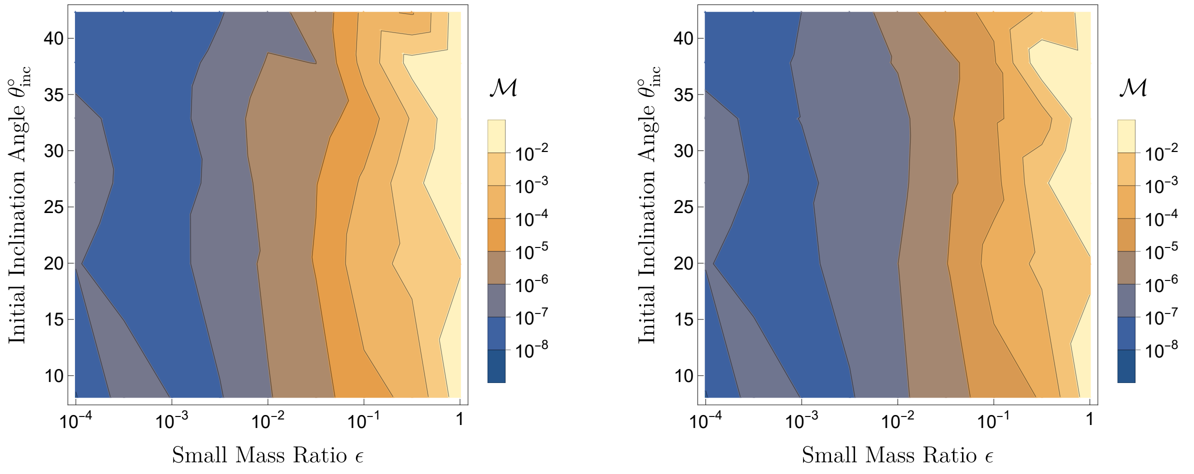


Figure 7.5: The absolute difference in the quantities of a prograde inspiral with $a = 0.9M$ and initial inclination of $x_0 = 0.75$ after evolving from $p = 5$ to $p = 4$ when using different evolution equations. The solid lines are the difference between OG and NIT equations of motion and the dashed lines are the difference between using the OG and TTE equations of motion. We see that the differences generally follow the black ϵ curve and so converge linearly with the mass ratio until we reach mass ratios $\leq 10^{-3}$ where the error in the phases becomes dominated by interpolation and numerical error.



(a) Mismatch between OG and NIT waveforms.

(b) Mismatch between OG and TTE waveforms.

Figure 7.6: The mismatch between year long OG and NIT/TTE waveforms as a function of orbital inclination and mass ratio. We do not see a substantial difference in accuracy between using either NIT or TTE waveforms. We also see that the mismatch remains smaller than 0.03 for mass ratios as large $\epsilon \approx 0.5$.

ϵ	OG	NIT	Speedup	TTE	Speedup
10^{-2}	24.7s	0.48s	~ 51	22ms	$\sim 1.1 \times 10^6$
10^{-3}	4m 20s	0.17s	~ 1530	39ms	$\sim 6.7 \times 10^6$
10^{-4}	43m 19s	0.43s	~ 6044	27ms	$\sim 9.6 \times 10^7$
10^{-5}	7hrs 42m	0.49s	~ 54380	22ms	$\sim 1.2 \times 10^9$

Table 7.1: Computational time required to evolve an inspiral from its initial conditions of $p_0 = 7.75$ and $x_0 = 0.75$ to the last stable orbit for different values of the mass ratio, as calculated in Mathematica 13 on an Intel Core i7 @ 2.2GHz. The computational time for the OG inspiral scales inversely with the mass ratio, whereas the computational time for NIT inspirals is independent of the mass ratio. The computation for the TTE inspiral was $0.53s$, which is slightly longer than any of the NIT inspirals. However, if we consider the TTE calculation as an offline step, one can then immediately recover the solution for any value of ϵ in a matter of milliseconds.

orbital timescale oscillations which themselves scale with the mass ratio and the orbital inclination. As such, we wish to determine the section of the parameter space where the difference in the waveforms between the OG and the NIT/TTE inspirals would be small enough for LISA data science. To do this, we first fix the mass of the primary to be $10^6 M_\odot$ and created a function which uses adiabatic inspirals and root-finding to numerically compute the initial orbital separation for an inspiral which will take one year to reach the ISSO for a given inclination and mass ratio. We then create a grid of mass ratios $\epsilon = \{1, 10^{-1/2}, 10^{-1}, 10^{-3/2}, 10^{-2}, 10^{-5/2}, 10^{-3}, 10^{-7/2}, 10^{-4}\}$ and inclinations $x = \{0.74, 0.79, 0.84, 0.89, 0.94, 0.99\}$. For each point on this grid, we calculate a year long inspiral using the OG equations of motion, the NIT equations of motion, and the TTE equations of motion and then generate a waveform from each.

The waveform mismatch between the OG and NIT or TTE waveforms is displayed in Fig. 7.6. From these plots we see that there does not seem to be a substantial difference in terms of accuracy from using either the NIT or TTE equations of motion. While there may be some correlation between mismatch and inclination, this does not appear to be a strong effect at least for $\theta_{\text{inc}} \leq 45°$. The strongest effect on the waveform mismatch comes from the mass ratio. It is worth noting that 3×10^{-2} is a commonly chosen maximum mismatch for a waveform template bank that corresponds to a 90 %-ideal observed event rate [187]. Our results suggests that one could in principle produce such a template bank of spherical inspirals using NIT or TTE waveforms even for mass ratios as large as $\epsilon \approx 0.5$. In practice however, we are still missing important post-adiabatic contributions such as second-order effects and contributions from the spin of the secondary, and so such a template bank would have substantial systematic biases.

Now that we have established the accuracy of our averaging procedure, we now demonstrate the speed up that one enjoys from using either the NIT or TTE equations of motion instead of using the OG equations of motion in Table. 7.1. For each of these calculations, the initial conditions are set to $p(0) = 7.75$ and $x(0) = 0.75$ the inspirals are evolved to the ISSO. All calculations were done using machine precision numbers and an accuracy goal of 7 for Mathematica's NDSolve function. We find that using the OG inspiral calculation takes longer as the mass ratio gets smaller since the solver will have to resolve more orbital cycles before the inspiral reaches the ISSO. However, we find that the NIT inspiral takes roughly the same amount of time regardless of the mass ratio. The TTE inspiral takes longer than any NIT inspiral, which is to be expected as one has to simultaneously solve for twice as many equations. However, one only has to solve this system once and then one can reconstruct the solution for any mass ratio thereafter. As a result, if one needs to compute multiple inspirals with varying mass ratios, the TTE equations of motion are significantly more computationally efficient.

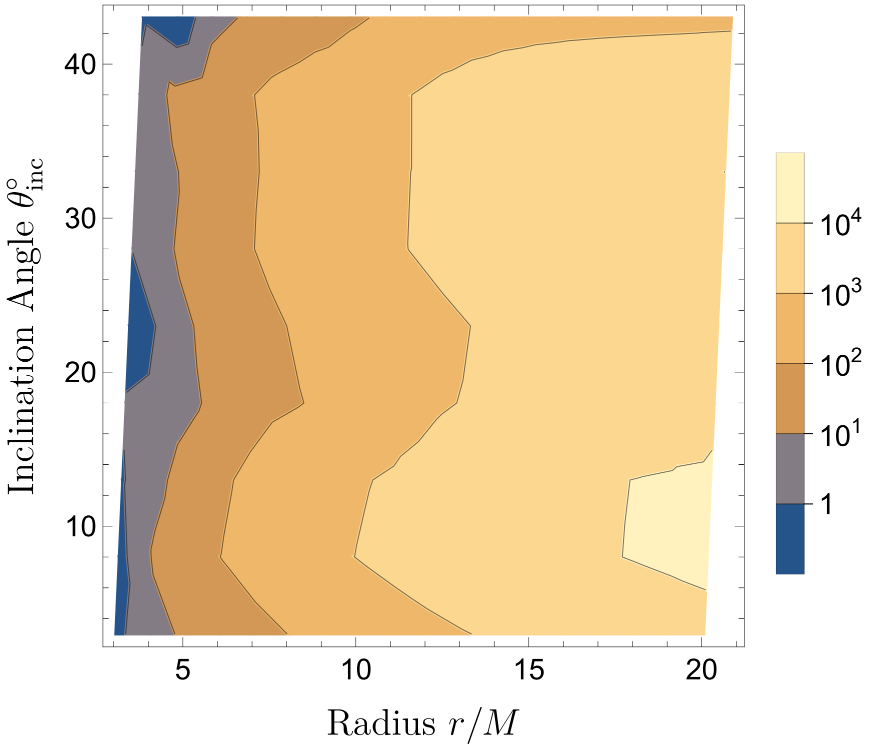


Figure 7.7: The absolute difference in φ_ϕ for a NIT inspiral evolved from a point in the parameter space to the ISSO with a typical EMRI mass ratio of $\epsilon = 10^{-5}$, depending on whether one incorporates high precision fluxes or not. Using the high precision fluxes results in a more faithful inspiral, with phase differences ranging from $\sim 10 - 10^4$ radians.

7.5.2 Using Higher Precision Fluxes

Now that we can compute fast and accurate inspirals, it is worth recalling that the relative accuracy of our interpolated GSF model is currently too low for production level waveforms. This is primarily due to the stringent relative accuracy requirements for the adiabatic pieces of $\lesssim \epsilon \times 10^{-2}$ for subradian accuracy in the phases. This can be improved by incorporating information from the asymptotic fluxes, which can be interpolated to a much higher accuracy across the parameter space due to the much cheaper cost of flux calculations compared to GSF calculations. This is why our interpolated flux model is accurate to $\sim 10^{-6}$ while our interpolated GSF model is only accurate to $\sim 10^{-3}$. Such information can be incorporated into the GSF model itself to improve the accuracy of the OG inspirals as well as the resulting NIT and TTE inspirals [70]. But since the NIT and TTE equations of motion are naturally split into adiabatic and post-adiabatic pieces, one can calculate $\Gamma^{(1)}_j$ directly from the fluxes via equations (7.3) and (7.4), interpolate them to higher precision than the GSF model, and then substitute these improved adiabatic terms into the averaged equations of motion.

To test the difference this would make to the overall accuracy of our post-adiabatic inspirals, we looked at the final error in the φ_ϕ phase when evolved from one point in the parameter space to the ISSO when using the flux+GSF model verses using only the GSF model for inspirals with $\epsilon = 10^{-5}$. From Fig. 7.7, we see that improving the adiabatic pieces results in a phase difference that can range for from tens to tens of thousands of radians for multi-year long inspirals and this difference gets larger as one moves away from the ISSO. As such, using high precision fluxes for the adiabatic contributions to the equations of motion is vital for obtaining accurate adiabatic and post-adiabatic inspirals.

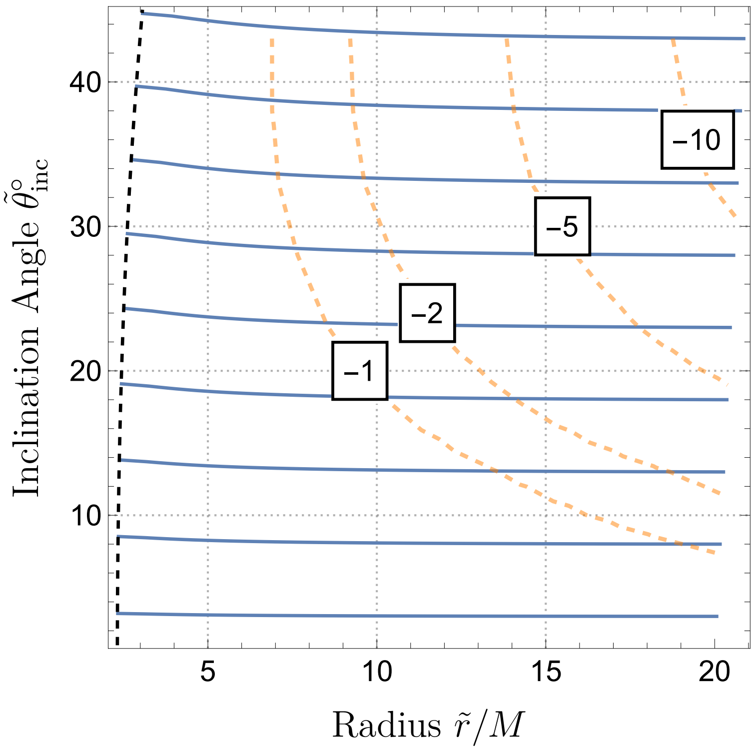


Figure 7.8: The blue curves show the adiabatic trajectories through $\{r, \theta_{\text{inc}}\}$ space, while the orange contours denote the final value of $\varphi^{(1)}_{\phi}$ when evolved from that point in the parameter space to the ISSO.

7.5.3 Impact of the self-force on spherical inspirals

Since this is the first time the first-order GSF has been computed for spherical Kerr inspirals, we examine the impact that the adiabatic and (a subset of) the post-adiabatic contributions have on the inspiral in Fig. 7.8. It is worth recalling that without second-order contributions to the 1PA inspiral, this is not a gauge invariant measurement, as we demonstrated in Chapter 6, and so these results are only representative of the outgoing radiation gauge. The blue curves show typical adiabatic trajectories through $\{r, \theta_{\text{inc}}\}$ space. From this we see that the self-force causes the orbital radius to decrease over time, but also cause the inclination angle to increase slightly throughout the inspiral. This is consistent with previous work on adiabatic spherical inspirals [78, 80]. The post-adiabatic contributions to solutions for r and θ_{inc} do not change this trend in any significant way since these contributions to the orbital elements is $\mathcal{O}(\epsilon)$. However, the post-adiabatic contribution to the orbital phases is $\mathcal{O}(1)$ and is thus very significant. This is demonstrated by the dashed contours on Fig. 7.8, which indicate the final value of the post-adiabatic piece of the azimuthal phase $\varphi^{(1)}_{\phi}$, when evolved from a given point in the parameter space to the ISSO. Since the azimuthal phase most strongly impacts the final waveform phase, this gives a good estimate of how many radians one can expect the waveform to be out of phase in the absence of post-adiabatic contributions without having to specify the mass ratio of the binary. It is worth restating that this does not include the second-order GSF contributions to $\varphi^{(1)}_{\phi}$ as these are not currently known.

Chapter 8

Fast generic inspirals into a rotating black hole

We now look to apply near-identity averaging transformations to the case of generic Kerr inspirals, while incorporating the phenomenon of orbital resonances. We start by restating the form of the osculating geodesics equations for generic Kerr inspirals in Sec. 8.1. We then outline the construction of a toy GSF model using information from our equatorial and spherical interpolated GSF models in Sec. 8.2.

In Sec. 8.3, we apply near-identity averaging transformations to generic Kerr inspirals in the absence of any low order orbital resonances and test how the accuracy and speed up converge with the mass ratio. We perform a similar analysis in Sec. 8.4, but this time for the near-resonant NIT in the vicinity of the $|\kappa_z|/|\kappa_r| = 2/3$ resonance.

Having confirmed that these two averaging procedures are performing as expected, we then prescribe how to transition from the non-resonant NIT to the near-resonant NIT and back again in Sec. 8.5. We test this transition procedure on two year long EMRIs, one which evolves through a single order resonance in Sec. 8.5.1 and one which evolves through two low order resonances in Sec. 8.5.2. While this implementation must be tested and optimised further before it could be recommenced for data analysis applications, these tests confirm that the overall strategy of combing non-resonant and near-resonant NIT equations of motion can accurately capture resonance crossings while drastically speeding up EMRI trajectory calculations.

8.1 Generic Kerr Osculating Geodesics

For generic Kerr orbits with both eccentricity and inclination, the osculating geodesic equations from Eq. (3.53) take the form

$$\frac{dp}{d\lambda} = \epsilon F_p^{(1)}(a,p,e,x,q_r,q_z) + \epsilon^2 F_p^{(2)}(a,p,e,x,q_r,q_z), \tag{8.1a}$$

$$\frac{de}{d\lambda} = \epsilon F_e^{(1)}(a,p,e,x,q_r,q_z) + \epsilon^2 F_e^{(2)}(a,p,e,x,q_r,q_z), \tag{8.1b}$$

$$\frac{dx}{d\lambda} = \epsilon F_x^{(1)}(a,p,e,x,q_r,q_z) + \epsilon^2 F_x^{(2)}(a,p,e,x,q_r,q_z), \tag{8.1c}$$

$$\frac{dq_r}{d\lambda} = \Upsilon_r^{(0)}(a,p,e,x) + \epsilon f_r^{(1)}(a,p,e,x,q_r,q_z), \tag{8.1d}$$

$$\frac{dq_z}{d\lambda} = \Upsilon_z^{(0)}(a,p,e,x) + \epsilon f_z^{(1)}(a,p,e,x,q_r,q_z), \tag{8.1e}$$

$$\frac{dt}{d\lambda} = f_t^{(0)}(a,p,e,x,q_r,q_z), \tag{8.1f}$$

$$\frac{d\phi}{d\lambda} = f_\phi^{(0)}(a,p,e,x,q_r,q_z). \tag{8.1g}$$

These equations of motion are valid everywhere, including through orbital resonances, and accurately capture the effects of transient resonances on the inspiral trajectory. The same is true of the quasi-Keplerian version of these equations, which we use in practice to calculate the OG inspirals as it is more numerically efficient than the action angle version. Even still, these equations are slow to evaluate due to the need to resolve orbital timescale oscillations. As such, we treat inspirals using the OG equations as if they are the "true" inspiral, and use them as the point of comparison as we implement more numerically efficient methods of evolving generic Kerr inspirals.

8.2 Generic Kerr GSF toy model

In order to start calculating inspirals, we first need a model for the secondary's four-acceleration. Unfortunately, creating an interpolated GSF model for generic Kerr inspirals is computationally unfeasible at this time, due to the cost of computing the generic Kerr self-force for a single point in the parameter space and the need to tile in three dimensions instead of two dimensions required for the equatorial and spherical cases. Instead we construct a self-force inspired toy model for generic orbits by combining our interpolated self-force models for eccentric and spherical orbits in such a way that we have radial and polar cross terms in the Fourier expansion of the force components which will give rise to resonant effects.

We first recall that our first order eccentric orbit self-force model takes the form:

$$a_\mu^{(1)} = A_\mu^0 + \sum_{n=1}^{15} A_\mu^n \cos(nq_r) + B_\mu^n \sin(nq_r), \tag{8.2}$$

where we have absorbed the rescaling factor of $(1-e^2)/(1-y_{eq})$ into the coefficients A_μ^n and B_μ^n. Note that we truncate the series at $n = 15$ as this provides sufficient accuracy. Furthermore, our first order spherical orbit self-force model takes the form:

$$a_\mu^{(1)} = C_\mu^0 + \sum_{m=1}^{24} C_\mu^m \cos(mq_z) + D_\mu^m \sin(mq_z) \tag{8.3}$$

where again we have absorbed the rescaling factor of $1/(1-y_{eq})^2$ into the coefficients C_μ^m and D_μ^m.

While it would be easier to combine these terms together to get cross terms if the Fourier series was expressed as a complex exponential series instead of a $\sin$ and $\cos$ series, it is possible to derive a Fourier series for a real valued 2D function my making use of trig identities and simplifying:

$$\begin{aligned} a_\mu^{(1)}(q_r,q_z) &= \sum_{n=0}^{\infty}\sum_{m=0}^{\infty} \mathcal{A}_\mu^{nm}\cos(nq_r)\cos(mq_z) \\ &+ \sum_{n=0}^{\infty}\sum_{m=0}^{\infty} \mathcal{B}_\mu^{nm}\cos(nq_r)\sin(mq_z) \\ &+ \sum_{n=0}^{\infty}\sum_{m=0}^{\infty} \mathcal{C}_\mu^{nm}\sin(nq_r)\cos(mq_z) \\ &+ \sum_{n=0}^{\infty}\sum_{m=0}^{\infty} \mathcal{D}_\mu^{nm}\sin(nq_r)\sin(mq_z) \end{aligned}$$

In our toy model, we calculate the cross terms using an outer product, meaning that the coefficients $\mathcal{A}_\mu^{nm}$, $\mathcal{B}_\mu^{nm}$, $\mathcal{C}_\mu^{nm}$, and $\mathcal{D}_\mu^{nm}$ are given by:

$$\mathcal{A}_\mu^{nm} = A_\mu^n C_\mu^m \tag{8.5a}$$

$$\mathcal{B}_\mu^{nm} = A_\mu^n D_\mu^m \tag{8.5b}$$

$$\mathcal{C}_\mu^{nm} = B_\mu^n C_\mu^m \tag{8.5c}$$

$$\mathcal{D}_\mu^{nm} = B_\mu^n D_\mu^m \tag{8.5d}$$

from our equatorial and spherical GSF models. Using what we know from the $n=0$ and $m=0$ cases, we express our generic orbit force components as:

$$\begin{aligned} a_\mu^{(1)}(q_r,q_z) = {} & A_\mu^0 + C_\mu^0 \\ & + \sum_{n=1}^{15} A_\mu^n\cos(nq_r) + B_\mu^n\sin(nq_r) + \sum_{m=1}^{24} C_\mu^m\cos(mq_z) + D_\mu^m\sin(mq_z) \\ & + \sum_{n=1}^{15}\sum_{m=1}^{24} A_\mu^n C_\mu^m\cos(nq_r)\cos(mq_z) + \sum_{n=1}^{15}\sum_{m=1}^{24} A_\mu^n D_\mu^m\cos(nq_r)\sin(mq_z) \\ & + \sum_{n=1}^{15}\sum_{m=1}^{24} B_\mu^n C_\mu^m\sin(nq_r)\cos(mq_z) + \sum_{n=1}^{15}\sum_{m=1}^{24} B_\mu^n D_\mu^m\sin(nq_r)\sin(mq_z) \end{aligned}$$

Note that all of the inclined orbit terms will vanish in the equatorial limit except for C_μ^0 and all the eccentric orbit terms will vanish in the circular limit except for A_μ^0. We wish to weight these orbit averaged terms so that we can recover the two limit cases accurately.

To do this, we note that our equatorial model covers eccentricities ranging from $0 < e < 0.5$. Our spherical model is tiled in terms of a parameter v where $v = z_-^2 = 1 - x^2$ such that our model covers $0 < v < 0.5$. This similarity will come in handy. We want the weighting factors to be smooth, recover the two limit factors, and we make the choice that when $e = v$ that the weighting factors are both 0.5

so that the orbit averaged piece will be the mean of the equatorial and spherical contributions. As such, we chose following weighting functions:

$$\alpha(e,v) = \begin{cases} \frac{e}{e+v} & e \leq v \\ 1 - \frac{v}{e+v} & e > v \end{cases}, \quad \beta(e,v) = \begin{cases} \frac{v}{e+v} & e \leq v \\ 1 - \frac{e}{e+v} & e > v \end{cases}. \tag{8.7}$$

Despite their piecewise definition, these functions are smooth and continuous everywhere except for the point $(e, v) = (0, 0)$. We can now write our generic force components as

$$\begin{aligned} a^{(1)}_\mu(q_r, q_z) =& \alpha A^0_\mu + \beta C^0_\mu \\ &+ \sum_{n=1}^{15} A^n_\mu \cos(nq_r) + B^n_\mu \sin(nq_r) + \sum_{m=1}^{24} C^m_\mu \cos(mq_z) + D^m_\mu \sin(mq_z) \\ &+ \sum_{n=1}^{15}\sum_{m=1}^{24} A^n_\mu C^m_\mu \cos(nq_r)\cos(mq_z) + \sum_{n=1}^{15}\sum_{m=1}^{24} A^n_\mu D^m_\mu \cos(nq_r)\sin(mq_z) \\ &+ \sum_{n=1}^{15}\sum_{m=1}^{24} B^n_\mu C^m_\mu \sin(nq_r)\cos(mq_z) + \sum_{n=1}^{15}\sum_{m=1}^{24} B^n_\mu D^m_\mu \sin(nq_r)\sin(mq_z). \end{aligned}$$

Before now, we only ever needed the orbit averaged contribution from the second order self-force, which in the absence of any results for Kerr inspirals, we simply set to zero. However, when evolving near an orbital resonance, Eqs. (4.53) show that the oscillatory part of the second order self-force also contributes at post adiabatic order. Since we are already using a toy model for the first order self-force, we also choose to create a toy model for the second order self-force $a^{(2)}_\mu$. However, we do not have any data for generic Kerr inspirals that can help inform such a model. As such, we have opted to use this first order toy model to inform the second order toy model, which we take to be:

$$a^{(2)}_\mu = \frac{a^{(1)}_\mu}{r^2\sqrt{1-z^2}}, \tag{8.9}$$

where the factor of $1/r^2$ is used to ensure that the second order self-force corresponds to the correct post-Newtonian order and the factor $1/\sqrt{1-z^2}$ prescribes the effect of the inclination, and implicitly spin of the primary, on the second order self-force. With these both of these terms, we express the self-force as $a_\mu = a^{(1)}_\mu + \epsilon a^{(2)}_\mu$.

Finally, in order for this toy model to work with the method of osculating geodesics, we require that it must satisfy the orthogonality condition with the geodesic four velocity, i.e., $a_\mu u^\mu = 0$. To enforce this relationship we project off any parts of the force that violate this condition using the following relationship:

$$a^\perp_\mu = a_\mu + a_\nu u^\nu u_\mu. \tag{8.10}$$

Using the projected force components with our osculating geodesic equations of motion we find inspirals that have qualitatively correct p, e, and x evolution for an EMRI under the effect of the gravitational self-force as well as strong resonant effects.

8.3 Non-Resonant NIT

8.3.1 Non-resonant NIT equations of motion

In the absence of orbital resonances, one can use near identity averaging transformations to derive equations of motion which have no dependence on the orbital phases:

$$\frac{d\tilde{p}}{d\lambda} = \epsilon \tilde{F}_p^{(1)}(a,\tilde{p},\tilde{e},\tilde{x}) + \epsilon^2 \tilde{F}_p^{(2)}(a,\tilde{p},\tilde{e},\tilde{x}), \tag{8.11a}$$

$$\frac{d\tilde{e}}{d\lambda} = \epsilon \tilde{F}_e^{(1)}(a,\tilde{p},\tilde{e},\tilde{x}) + \epsilon^2 \tilde{F}_e^{(2)}(a,\tilde{p},\tilde{e},\tilde{x}), \tag{8.11b}$$

$$\frac{d\tilde{x}}{d\lambda} = \epsilon \tilde{F}_x^{(1)}(a,\tilde{p},\tilde{e},\tilde{x}) + \epsilon^2 \tilde{F}_x^{(2)}(a,\tilde{p},\tilde{e},\tilde{x}), \tag{8.11c}$$

$$\frac{d\tilde{q}_r}{d\lambda} = \Upsilon_r^{(0)}(a,\tilde{p},\tilde{e},\tilde{x}) + \epsilon \Upsilon_r^{(1)}(a,\tilde{p},\tilde{e},\tilde{x}), \tag{8.11d}$$

$$\frac{d\tilde{q}_z}{d\lambda} = \Upsilon_z^{(0)}(a,\tilde{p},\tilde{e},\tilde{x}) + \epsilon \Upsilon_z^{(1)}(a,\tilde{p},\tilde{e},\tilde{x}), \tag{8.11e}$$

$$\frac{d\tilde{t}}{d\lambda} = \Upsilon_t^{(0)}(a,\tilde{p},\tilde{e},\tilde{x}) + \epsilon \Upsilon_t^{(1)}(a,\tilde{p},\tilde{e},\tilde{x}), \tag{8.11f}$$

$$\frac{d\tilde{\phi}}{d\lambda} = \Upsilon_\phi^{(0)}(a,\tilde{p},\tilde{e},\tilde{x}) + \epsilon \Upsilon_\phi^{(1)}(a,\tilde{p},\tilde{e},\tilde{x}). \tag{8.11g}$$

The leading order terms in each equation of motion are simply the original function averaged over a single geodesic orbit, i.e.,

$$\tilde{F}_p^{(1)} = \left\langle F_p^{(1)} \right\rangle, \quad \tilde{F}_e^{(1)} = \left\langle F_e^{(1)} \right\rangle, \quad \tilde{F}_x^{(1)} = \left\langle F_x^{(1)} \right\rangle, \tag{8.12}$$

$$\Upsilon_r^{(1)} = \left\langle f_r^{(1)} \right\rangle, \quad \Upsilon_z^{(1)} = \left\langle f_z^{(1)} \right\rangle, \tag{8.13}$$

$$\Upsilon_t^{(0)} = \left\langle f_t^{(0)} \right\rangle, \quad \Upsilon_\phi^{(0)} = \left\langle f_\phi^{(0)} \right\rangle. \tag{8.14}$$

The remaining terms are more complicated and require Fourier decomposing the original functions and their derivatives with respect to to the orbital elements (p,e,x). To express the result, we define the operator

$$\begin{aligned}
\mathcal{N}(A) = \sum_{(n,m)\neq(0,0)} \frac{i}{n\Upsilon_r^{(0)} + m\Upsilon_z^{(0)}} \Bigg[& i A_{n,m}\left(n f^{(1)}_{r,-n,-m} + m f^{(1)}_{z,-n,-m} \right) + \\
& \frac{\partial A_{n,m}}{\partial \tilde{p}} F^{(1)}_{p,-n,-m} + \frac{\partial A_{n,m}}{\partial \tilde{e}} F^{(1)}_{e,-n,-m} + \frac{\partial A_{n,m}}{\partial \tilde{x}} F^{(1)}_{x,-n,-m} - \frac{A_{n,m}}{n\Upsilon_r^{(0)} + m\Upsilon_z^{(0)}} \times \\
& \left(\left(n\frac{\partial \Upsilon_r^{(0)}}{\partial \tilde{p}} + m\frac{\partial \Upsilon_z^{(0)}}{\partial \tilde{p}} \right) F^{(1)}_{p,-n,-m} + \left(n\frac{\partial \Upsilon_r^{(0)}}{\partial \tilde{e}} + m\frac{\partial \Upsilon_z^{(0)}}{\partial \tilde{e}} \right) F^{(1)}_{e,-n,-m} + \left(n\frac{\partial \Upsilon_r^{(0)}}{\partial \tilde{x}} + m\frac{\partial \Upsilon_z^{(0)}}{\partial \tilde{x}} \right) F^{(1)}_{x,-n,-m} \right) \Bigg]
\end{aligned} \tag{8.15}$$

With this in hand, the remaining terms in the equations of motion are found to be

$$\tilde{F}_p^{(2)} = \left\langle F_p^{(2)} \right\rangle + \mathcal{N}\left(F_p^{(1)} \right), \quad \tilde{F}_e^{(2)} = \left\langle F_e^{(2)} \right\rangle + \mathcal{N}\left(F_e^{(1)} \right), \quad \tilde{F}_x^{(2)} = \left\langle F_x^{(2)} \right\rangle + \mathcal{N}\left(F_x^{(1)} \right) \tag{8.16a}$$

$$\Upsilon_t^{(1)} = \mathcal{N}\left(f_t^{(0)} \right), \quad \Upsilon_\phi^{(1)} = \mathcal{N}\left(f_\phi^{(1)} \right). \tag{8.16b}$$

8.3.2 Implementation

Combining the toy generic GSF model along with our action angle formulation of the OG equations gives us all the information required to calculate the NIT equations of motion. We first evaluate and interpolate the various terms in the NIT equations of motion across the parameter space. While this offline process can be expensive, it only needs to be completed once. By contrast, the online process of calculating self-forced inspirals can be run in less than a second.

8.3.3 Offline Steps

The offline calculation consists of the following steps.

1. We start by selecting a grid which covers the parameter space. We fix the spin of the primary to be $a = 0.9M$ chose an equally spaced gird of $\vec{P} = (p, e, x)$ values on which to evaluate the terms in the NIT equations of motion. To ensure that we are not in the vicinity of a low order resonance, we only tile a small patch of the parameter space where no such resonances are present. As such, we allow p to range from 9 to 10.1 in steps of 0.1, e to range from 0.25 to 0.41 in steps of 0.01, and x range from 0.79 to 0.81 in steps of 0.005 for a total of 1020 grid points.

2. For each of these grid points we evaluate the functions $F^{(1)}_{p\backslash x\backslash e}$,$F^{(2)}_{p\backslash x\backslash e}$, $f^{(1)}_{r\backslash z}$, and $f^{(0)}_{t\backslash\phi}$ along with their derivatives with respect to p,e, and x for 21 equally spaced values of both q_r and q_z ranging from 0 to 2π each for a total of 441 evaluations for each function.

3. We then perform a fast Fourier transform on the output data to obtain the Fourier coefficients of the forcing functions and their derivatives.

4. With these, we then use Eqs. (8.12), (8.13), (8.15) and (8.14) to construct $\tilde{F}^{(1\backslash 2)}_{p\backslash e\backslash x}$, and $\Upsilon^{(1)}_{t\backslash z\backslash\phi}$.

5. We also use Eqs. (4.55) and (4.56) to construct the Fourier coefficients of the first order transformation functions $Y^{(1)}_{p\backslash e\backslash x}$ and $X^{(1)}_{r\backslash z}$.

6. We then repeat this procedure across the parameter space for each point in our grid.

7. Finally we interpolate the values for $\tilde{F}^{(1\backslash 2)}_{p\backslash e\backslash x}$ and $\Upsilon^{(1)}_{t\backslash z\backslash\phi}$ along with the coefficients of $Y^{(1)}_{p\backslash e\backslash x}$ and $X^{(1)}_{r\backslash z}$ across this grid using Hermite interpolation and store the interpolants for future use.

We implemented the above algorithm in Mathematica 13 and find, parallelized across 40 CPU cores takes, the calculation takes about 12 hours to complete.

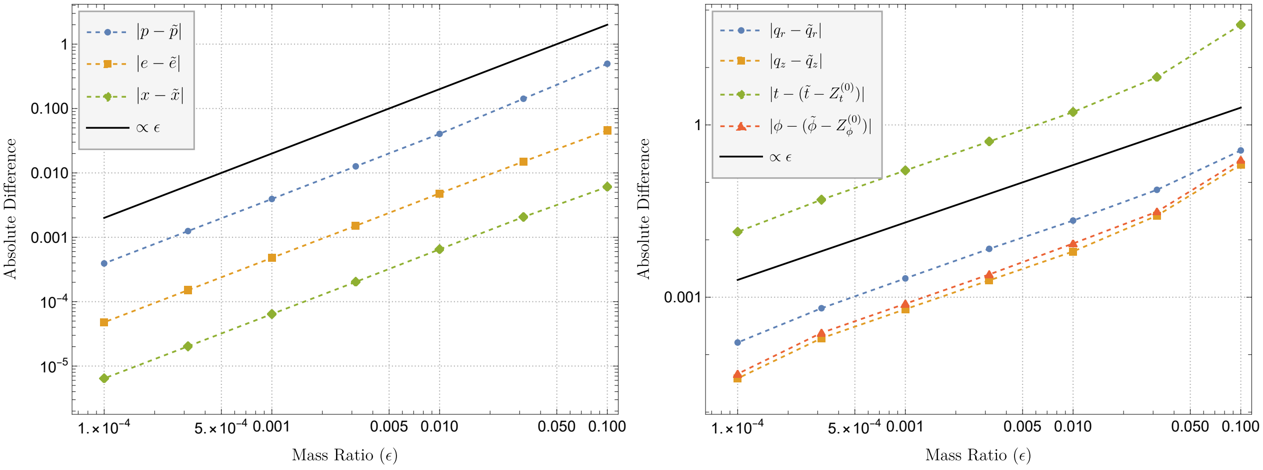
8.3.4 Online Steps

The computationally inexpensive online steps needed for every inspiral calculation are as follows

1. We load in the interpolants for $\tilde{F}^{(1\backslash 2)}_{p\backslash e\backslash x}$, and $\Upsilon^{(1)}_{t\backslash r\backslash z\backslash\phi}$, and define the NIT equations of motion.

2. In order to make comparisons between OG and NIT inspirals we also load interpolants of the Fourier coefficients of $Y^{(1)}_{p\backslash e\backslash x}$.

3. We then state the initial conditions of the inspiral $(p(0), e(0), x(0), q_r(0), q_z(0))$ and use Eq. (4.33) to ensure that the initial conditions are matched to linear order in the mass ratio.



(a) Orbital elements.

(b) Orbital phases and extrinsic quantities.

Figure 8.1: The absolute difference in the quantities of an inspiral with $a = 0.9M$ and initial conditions $(e_0, x_0) = (0.38, 0.8)$ and evolved from $p = 9.5$ to $p = 9$ for difference values of the mass ratio when calculated using either the OG or NIT equations of motion. As expected, the differences scale linearly with the mass ratio.

4. We then evolve the NIT equations of motion using an ODE solver (in our case Mathematica's `NDSolve`).

As with the offline steps we implement the online steps in Mathematica. Note that steps (2) and (3) are only necessary because we want to make direct comparisons between NIT and OG inspirals with the same initial conditions. In general, the difference between the NIT and OG variables will always be $\mathcal{O}(\epsilon)$, and so performing the NIT transformation or inverse transformation to greater than zeroth order in mass ratio will not be necessary when producing waveforms to post adiabatic order, i.e. with phases accurate to $\mathcal{O}(\epsilon)$.

8.3.5 Results

In order to test that we have implemented the NIT equations of motion correctly, we wish to examine how the differences between the OG and NIT inspiral quantities vary with the mass ratio. If implemented correctly, the difference between the two should scale linearly with the mass ratio, and so any deviation from this would indicate either a bug in our code, a large interpolation error for the terms in our NIT equations of motion, or error in the numerical solver. As such, we start an inspiral at $(p(0), e(0), x(0)) = (9.5, 0.4, 0.8)$ and evolve the inspiral until $p = 9$ using both the OG and NIT equations of motion and varying the mass ratio from 10^{-1} to 10^{-4}. These initial conditions were specifically chosen to avoid encountering any low order resonances during the inspiral.

The results of this test are shown in Fig. 8.1. The figure demonstrates that the differences between the OG and NIT orbital elements, the phases, and the extrinsic values all scale linearly with the mass ratio. This indicates that in the absence of low order orbital resonances, the standard NIT formulation is valid for generic Kerr inspirals.

Table. 8.1 displays the speed up one obtains from using the NIT equations of motion instead of the OG equations. We see that as the mass ratio decreases, the longer it takes for the OG inspirals to compute the solver now has to resolve more orbital timescale oscillations. By contrast, the runtime of

ϵ	OG Inspiral	NIT Inspiral	Speed-up
10^{-1}	9.33s	0.127s	~ 73.5
$10^{-1.5}$	13.2s	0.099s	~ 133
10^{-2}	29.7s	0.088s	~ 338
$10^{-2.5}$	88.3s	0.103s	~ 857
10^{-3}	281s	0.117s	~ 2402
$10^{-3.5}$	860s	0.14s	~ 6143
10^{-4}	2697s	0.254s	$\sim 10,618$

Table 8.1: Computational time required to evolve an inspiral from its initial conditions of $(p_0, e_0, x_0) = (9.5, 0.38, 0.8)$ to $p = 9$ for different values of the mass ratio. These inspirals were calculated numerically using NDSolve with an accuracy goal of 8 as implemented in Mathematica 13 on an Intel Core i7 @ 2.2GHz. As before, the computational time for the OG inspiral scales inversely with the mass ratio, whereas the computational time for NIT inspirals is independent of the mass ratio. This demonstrates how the smaller the mass ratio of the inspiral, the greater speed-up one obtains from using the NIT equations of motion.

the NIT inspirals is independent of the mass ratio as there are no orbital oscillations to resolve. As a result, the NIT works best, in terms of both accuracy and relative speed-up, for inspirals with very small mass ratios.

8.4 Near-resonant NIT

While we have shown that the NIT can be applied for generic orbits away from resonances, we expect most EMRIs to evolve through at least one low order resonance while in the LISA band [160]. In the presence of one of these low order resonances, $\tilde{F}^{(2)}_{p\backslash e\backslash x}$, $\Upsilon^{(1)}_{r\backslash z}$, $Y^{(1)}_{p\backslash e\backslash x}$ and $X^{(1)}_{r\backslash z}$ all exhibit singular behaviour. As such, we adopt a near-resonant NIT formulation when in the vicinity of an orbital resonance. While the resulting inspirals won't be as quick to compute, the resulting inspiral quantities should still be accurate to the OG inspiral to linear order in mass ratio.

8.4.1 Near -resonant NIT equations of motion

Assuming that we are in the vicinity of an orbital resonance where $\vec{\kappa}_{\text{res}} \cdot \vec{\Upsilon}^{(0)} = \kappa_r \Upsilon_r^{(0)} + \kappa_z \Upsilon_z^{(0)} = 0$, we make use of the following partially averaged equations of motion

$$\frac{d\tilde{p}}{d\lambda} = \epsilon \tilde{F}_p^{(1)}(a, \tilde{p}, \tilde{e}, \tilde{x}, \tilde{q}_\perp) + \epsilon^2 \tilde{F}_p^{(2)}(a, \tilde{p}, \tilde{e}, \tilde{x}, \tilde{q}_\perp), \tag{8.17a}$$

$$\frac{d\tilde{e}}{d\lambda} = \epsilon \tilde{F}_e^{(1)}(a, \tilde{p}, \tilde{e}, \tilde{x}, \tilde{q}_\perp) + \epsilon^2 \tilde{F}_e^{(2)}(a, \tilde{p}, \tilde{e}, \tilde{x}, \tilde{q}_\perp), \tag{8.17b}$$

$$\frac{d\tilde{x}}{d\lambda} = \epsilon \tilde{F}_x^{(1)}(a, \tilde{p}, \tilde{e}, \tilde{x}, \tilde{q}_\perp) + \epsilon^2 \tilde{F}_x^{(2)}(a, \tilde{p}, \tilde{e}, \tilde{x}, \tilde{q}_\perp), \tag{8.17c}$$

$$\frac{d\tilde{q}_r}{d\lambda} = \Upsilon_r^{(0)}(a, \tilde{p}, \tilde{e}, \tilde{x}) + \epsilon \Upsilon_r^{(1)}(a, \tilde{p}, \tilde{e}, \tilde{x}, \tilde{q}_\perp), \tag{8.17d}$$

$$\frac{d\tilde{q}_z}{d\lambda} = \Upsilon_z^{(0)}(a, \tilde{p}, \tilde{e}, \tilde{x}) + \epsilon \Upsilon_z^{(1)}(a, \tilde{p}, \tilde{e}, \tilde{x}, \tilde{q}_\perp), \tag{8.17e}$$

$$\frac{d\tilde{t}}{d\lambda} = \Upsilon_t^{(0)}(a, \tilde{p}, \tilde{e}, \tilde{x}) + \epsilon \Upsilon_t^{(1)}(a, \tilde{p}, \tilde{e}, \tilde{x}), \tag{8.17f}$$

$$\frac{d\tilde{\phi}}{d\lambda} = \Upsilon_\phi^{(0)}(a, \tilde{p}, \tilde{e}, \tilde{x}) + \epsilon \Upsilon_\phi^{(1)}(a, \tilde{p}, \tilde{e}, \tilde{x}), \tag{8.17g}$$

where now the equations of motion depend only on the slowly evolving orbital elements $(\tilde{p}, \tilde{e}, \tilde{x})$ and the resonant phase $q_\perp = \kappa_r q_r + \kappa_z q_z$. The leading order terms in each equation of motion are given by only the resonant modes of the Fourier sum of the corresponding OG term i.e.,

$$\tilde{F}_p^{(1)} = \sum_N F_{p,Nk_r,Nk_z}^{(1)} e^{iNq_\perp}, \quad \tilde{F}_e^{(1)} = \sum_N F_{e,N\kappa_r,N\kappa_z}^{(1)} e^{iNq_\perp}, \quad \tilde{F}_x^{(1)} = \sum_N F_{x,N\kappa_r,N\kappa_z}^{(1)} e^{iNq_\perp}, \tag{8.18}$$

$$\Upsilon_r^{(1)} = \sum_N f_{r,N\kappa_r,N\kappa_z}^{(1)} e^{iNq_\perp}, \quad \Upsilon_z^{(1)} = \sum_N f_{z,N\kappa_r,N\kappa_z}^{(1)} e^{iNq_\perp}. \tag{8.19}$$

The remaining terms are more complicated and require Fourier decomposing the original functions and their derivatives with respect to to the orbital elements (p, e, x). To express the result, we define the operator

$$\begin{aligned}
\mathcal{N}_{\text{res}}(A) = \sum_{(n,m)\in R} \frac{i}{n\Upsilon_r^{(0)} + m\Upsilon_z^{(0)}} \Bigg[& iA_{n,m}\left(nf_{r,-n,-m}^{(1)} + mf_{z,-n,-m}^{(1)}\right) + \\
& \frac{\partial A_{n,m}}{\partial \tilde{p}} F_{p,-n,-m}^{(1)} + \frac{\partial A_{n,m}}{\partial \tilde{e}} F_{e,-n,-m}^{(1)} + \frac{\partial A_{n,m}}{\partial \tilde{x}} F_{x,-n,-m}^{(1)} - \frac{A_{n,m}}{n\Upsilon_r^{(0)} + m\Upsilon_z^{(0)}} \times \\
& \left(\left(n\frac{\partial \Upsilon_r^{(0)}}{\partial \tilde{p}} + m\frac{\partial \Upsilon_z^{(0)}}{\partial \tilde{p}}\right) F_{p,-n,-m}^{(1)} + \left(n\frac{\partial \Upsilon_r^{(0)}}{\partial \tilde{e}} + m\frac{\partial \Upsilon_z^{(0)}}{\partial \tilde{e}}\right) F_{e,-n,-m}^{(1)} + \left(n\frac{\partial \Upsilon_r^{(0)}}{\partial \tilde{x}} + m\frac{\partial \Upsilon_z^{(0)}}{\partial \tilde{x}}\right) F_{x,-n,-m}^{(1)}\right)\Bigg],
\end{aligned} \tag{8.20}$$

where R is the set $\{(n,m) \in \mathbb{Z}^2 | (n,m) \neq N(\kappa_r, \kappa_z), \forall N \in \mathbb{Z}\}$. Armed with this operator, we can express the remaining terms in the equations of motion as

$$\tilde{F}_p^{(2)} = \sum_N F_{p,Nk_r,Nk_z}^{(2)} e^{iNq_\perp} + \mathcal{N}_{\text{res}}\left(F_p^{(1)}\right) \tag{8.21a}$$

$$\tilde{F}_e^{(2)} = \sum_N F_{e,Nk_r,Nk_z}^{(2)} e^{iNq_\perp} + \mathcal{N}_{\text{res}}\left(F_e^{(1)}\right) \tag{8.21b}$$

$$\tilde{F}_x^{(2)} = \sum_N F_{x,Nk_r,Nk_z}^{(2)} e^{iNq_\perp} + \mathcal{N}_{\text{res}}\left(F_x^{(1)}\right) \tag{8.21c}$$

Due to the separability of the rates of change of the extrinsic quantities $f_t^{(1)}$ and $f_\phi^{(1)}$, we can continue to use the same equations averaged equations of motion as in the non-resonant case:

$$\Upsilon_t^{(0)} = \left\langle f_t^{(0)} \right\rangle, \quad \Upsilon_\phi^{(0)} = \left\langle f_\phi^{(0)} \right\rangle. \tag{8.22a}$$

$$\Upsilon_t^{(1)} = \mathcal{N}\left(f_t^{(0)} \right), \quad \Upsilon_\phi^{(1)} = \mathcal{N}\left(f_\phi^{(0)} \right). \tag{8.22b}$$

8.4.2 Implementation

The offline calculation consists of the following steps.

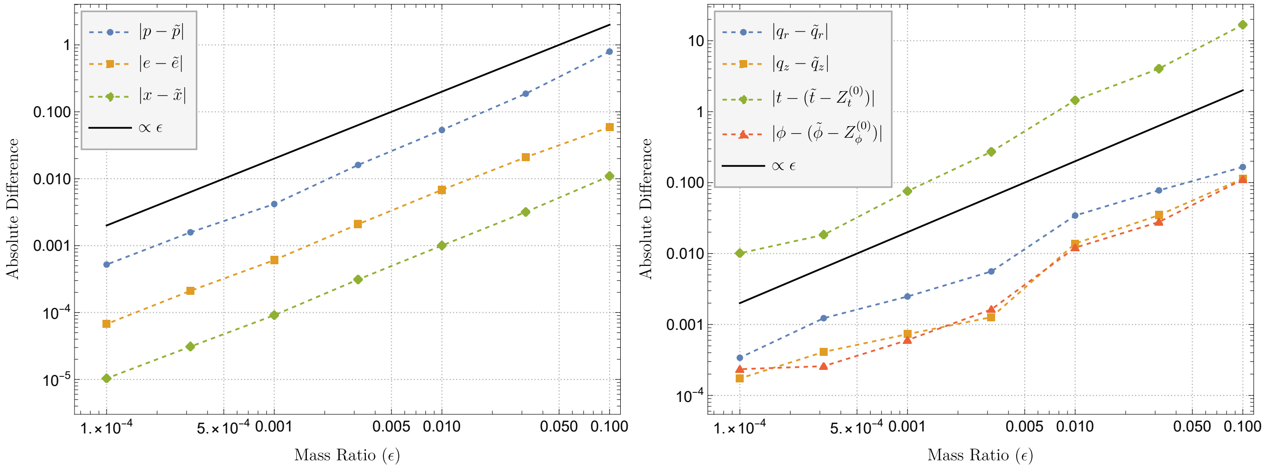
1. We start by selecting a grid which covers the parameter space. We fix the spin of the primary to be $a = 0.9M$ chose an equally spaced gird of $\vec{P} = (p, e, x)$ values on which to evaluate the terms in the NIT equations of motion. We now wish to cover much more of the parameter space and so we allow p to range from 3.75 to 7.25 in steps of 0.05, e to range from 0.25 to 0.41 in steps of 0.01, and x range from 0.79 to 0.81 in steps of 0.005 for a total of 1020 grid points.
2. For each of these grid points we evaluate the functions $F^{(1)}_{p\backslash e\backslash x}$,$F^{(2)}_{p\backslash e\backslash x}$, $f^{(1)}_{r\backslash z}$, and $f^{(0)}_{t\backslash\phi}$ along with their derivatives with respect to p,e,and x for 21 equally spaced values of both q_r and q_z ranging from 0 to 2π each for a total of 441 evaluations for each function.
3. We then perform a fast Fourier transform on the output data to obtain the Fourier coefficients of the forcing functions and their derivatives.
4. We repeat this across the parameter space and save the values of each of these Fourier coefficients.
5. We then interpolate the coefficients of $F^{(1\backslash 2)}_{p\backslash e\backslash x}$ and $f^{(1)}_{r\backslash z}$ using Hermite polynomials. This allows us to construct not only the transformation terms $Y^{(1)}_{p\backslash e\backslash x}$ and $X^{(1)}_{r\backslash z}$, but will also allow us to quickly evaluate the semi-oscillating terms in the equations of motion $\sum_N F^{(1\backslash 2)}_{p\backslash e\backslash x, Nk_r, N\kappa_z} e^{iNq_\perp}$ and $\sum_N f^{(1)}_{r\backslash z, N\kappa_r, Nk_z} e^{iNq_\perp}$.
6. Using the rest of the stored Fourier coefficients, one can then use the definition near-resonance $\mathcal{N}_{\text{res}}$ operator given by Eq. (8.20) to construct the remaining parts of $\tilde{F}^{(2)}_{p\backslash e\backslash x}$ at each grid point, which are then interpolated using Hermite polynomials and stored for future use.
7. The above step can be repeated if we wish to interpolate these functions for a different orbital resonance in the parameter space, or use the non-resonant NIT expressions in the absence of any orbital resonances.

We implemented the above algorithm in Mathematica 13.0 and find that when parallelized across 40 CPU cores, the calculation takes about 7 hours to calculate the Fourier coefficients, about an 30 minutes to interpolate Fourier coefficients of $F^{(1\backslash 2)}_{p\backslash e\backslash x}$ and $f^{(1)}_{r\backslash z}$ and then about an hour to interpolate the the near-resonant NIT terms for a single resonance.

8.4.3 Online Steps

The computationally inexpensive online steps needed for every inspiral calculation are as follows

1. We load in the interpolants for the near-resonant $\tilde{F}^{(1\backslash 2)}_{p\backslash e\backslash x}$, and $\Upsilon^{(1)}_{t\backslash r\backslash z\backslash\phi}$, and define the near-resonant NIT equations of motion.



(a) Orbital elements.

(b) Orbital phases and extrinsic quantities.

Figure 8.2: The absolute difference in the quantities of an inspiral with $a = 0.9M$ and initial conditions $(e_0, x_0) = (0.36, 0.8)$ and evolved from $p = 6.5$ to $p = 5.8$ for difference values of the mass ratio when calculated using either the OG or near-resonant NIT equations of motion. As expected, the differences scale linearly with the mass ratio, though these differences in the phases and extrinsic quantities oscillate due to the dependence of the near-resonance NIT equations of motion on the resonant phase $q_\perp$.

2. In order to make comparisons between OG and NIT inspirals we also load interpolants of the Fourier coefficients of $F^{(1)}_{p\backslash e\backslash x}$ so that can construct the term $Y^{(1)}_{p\backslash e\backslash x}$.

3. We then state the initial conditions of the inspiral $(p(0), e(0), x(0), q_r(0), q_z(0))$ and use Eq. (4.33) to ensure that the initial conditions are matched to linear order in the mass ratio.

4. We then evolve the NIT equations of motion using an ODE solver (in our case Mathematica's `NDSolve`).

8.4.4 Results

As before, we now test the implementation of the near-resonance NIT by investigating how the differences between the OG and NIT inspiral quantities vary with the mass ratio. Again, if implemented correctly, we expect the differences to scale linearly with the mass ratio. We start each inspiral at $(p(0), e(0), x(0)) = (6.5, 0.36, 0.8)$ and evolve the system until $p = 5.8$ while varying the mass ratio from 10^{-1} to 10^{-4}, which ensures that the inspirals always pass through the $|\kappa_z|/|\kappa_r| = 2/3$ orbital resonance.

The differences between using the OG and near-resonant NIT equations of motion are displayed in Fig. 8.2. We see that the difference in the orbital elements scales linearly with the mass ratio, as shown in Fig. 8.2a. The differences in the orbital phases and extrinsic quantities still scale linearly with the mass ratio but also tend to oscillate more than in the non-resonant case. This is due to these quantities being more sensitive than the orbital elements to the dependence on the near-resonant NIT equations of motion on the resonant phase $q_\perp$. Overall, this test demonstrates that the differences scale as expected and assures us that we have implemented the near-resonant NIT correctly.

In Table. 8.2 shows the runtime for each of these inspirals. As before, we see that the runtime of the OG inspirals scales inversely with the mass ratio. However, we also see that the runtime of the near-resonant NIT inspirals depends on the mass ratio, since the near-resonant NIT depends on

ϵ	OG Inspiral	NIT Inspiral	Speed-up
10^{-1}	3.44s	0.724s	~ 4.75
$10^{-1.5}$	7.21s	0.72s	~ 10
10^{-2}	18.7s	0.817s	~ 22.9
$10^{-2.5}$	48.7s	1.36s	~ 35.8
10^{-3}	160s	2.41s	~ 66.4
$10^{-3.5}$	516s	5.08s	~ 102
10^{-4}	1611s	14.35s	~ 112

Table 8.2: Computational time required to evolve an inspiral from its initial conditions of $(p_0, e_0, x_0) = (6.5, 0.38, 0.8)$ to $p = 5.8$ for different values of the mass ratio. These inspirals were calculated numerically using NDSolve with accuracy and precision goals both set to 13.5 as implemented in Mathematica 13 on an Intel Core i7 @ 2.2GHz. As before, the computational time for the OG inspiral scales inversely with the mass ratio, but now the computational time for near-resonant NIT inspirals also scales with the mass ratio due to the dependence on $q_\perp$. The end result that one does not obtain the same level of speed-up as in the non-resonant case, but using the near-resonant NIT still provides substantial speed-up over using the OG equations.

the orbital phase and thus has to resolve more orbital cycles, just like the OG inspirals. However, the near-resonant NIT inspirals do not have to resolve as many oscillations as the OG inspirals, resulting in considerable speed-up. While this is helpful, near-resonant NIT inspirals require seconds to minutes to evolve in Mathematica which is substantially slower than the sub-second calculation times achieved in the non-resonant case.

8.5 Resonance Transition

While we can now accurately evolve inspirals in the vicinity of an orbital resonance in a way that is an order of magnitude faster than using the OG equations, this is still not fast enough for LISA data analysis. One way this can be sped-up further is by using the non-resonant NIT equations when far away from a resonance and then switching to the near-resonant NIT equations of motion when in the vicinity of the resonance.

Where this transition should take place to minimise the time using the near-resonant NIT while retaining subradian phase accuracy was investigated in Ref. [192] using an error budget analysis. It found that assuming one could evaluate the second order self-force at the point of resonance, but not its derivatives with respect to the orbital elements, and assuming an error in the phase accuracy no greater than $\mathcal{O}(\epsilon^{1/2})$, the transition should take place when $\kappa_r \Upsilon_r^{(0)} + \kappa_z \Upsilon_z^{(0)} \propto \epsilon^{1/4}$. We leave verifying this choice of condition and how the phase error scales with the mass ratio as future work and assume that this analysis is correct. Furthermore, the analysis only gives the scaling with the mass ratio and no bounds for the leading coefficient. Empirically, we take this coefficient to be a constant of 2 for all resonances, and so we assume the near-resonance condition to be

$$\kappa_r \Upsilon_r + \kappa_z \Upsilon_z = 2\epsilon^{1/4}. \tag{8.23}$$

Investigating this factor and optimising the near-resonance condition further is also left as future work.

One must note that when transitioning from the non-resonant NIT to the near-resonant NIT, we must apply a small jump to the orbital elements and orbital phases in order to capture the change in the near-identity transformations. While normally we do need the explicit expressions of the transformation terms $Y_j^{(1)}$ and $X_i^{(1)}$ when evolving inspirals, unless we are comparing inspirals calculated using the OG equations, we find that including jumps in terms of these transformation has a noticeable impact on the accuracy of our inspiral calculations that require a resonance transition. As such, when entering the near-resonance region:

$$\tilde{P}_j \to \tilde{P}_j - \epsilon \sum_N Y^{(1)}_{j,N\kappa_r,N\kappa_z} e^{iq_\perp} + \mathcal{O}(\epsilon^2) \tag{8.24a}$$

$$\tilde{q}_i \to \tilde{q}_i - \epsilon \sum_N X^{(1)}_{i,N\kappa_r,N\kappa_z} e^{iq_\perp} + \mathcal{O}(\epsilon^2). \tag{8.24b}$$

Similarly, when exiting the near-resonance region, we also apply the jump condition:

$$\tilde{P}_j \to \tilde{P}_j + \epsilon \sum_N Y^{(1)}_{j,N\kappa_r,N\kappa_z} e^{iq_\perp} + \mathcal{O}(\epsilon^2) \tag{8.25a}$$

$$\tilde{q}_i \to \tilde{q}_i + \epsilon \sum_N X^{(1)}_{i,N\kappa_r,N\kappa_z} e^{iq_\perp} + \mathcal{O}(\epsilon^2). \tag{8.25b}$$

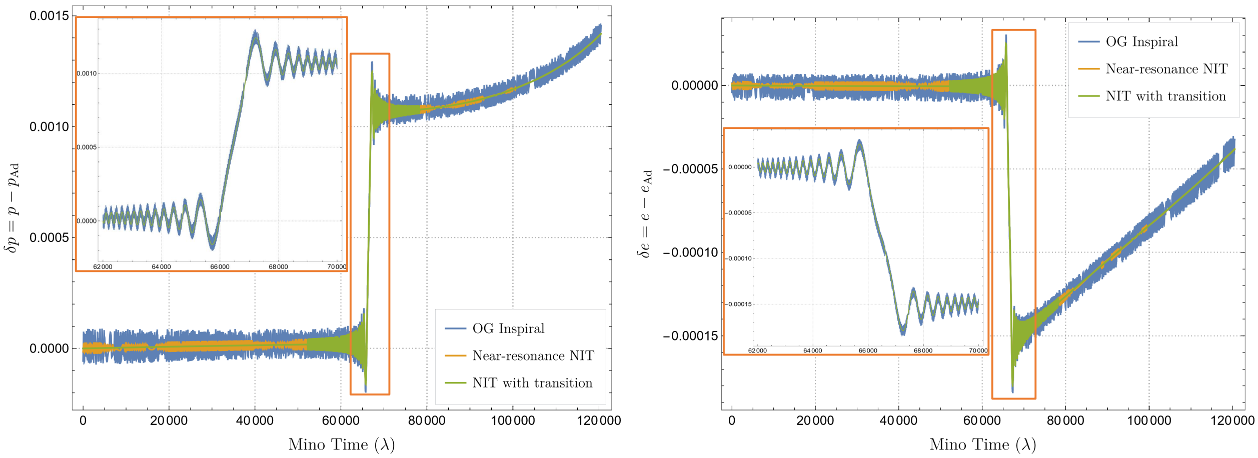
Using this prescription, we wish to demonstrate that transitioning from the non-resonant NIT to the near-resonant NIT equations of motion and back again can speed up the runtime for a semi-realistic EMRI system while introducing no significant error, and that we can use these transitions to account for multiple resonance crossings. To this end, we examine two year long ERMIs; one which passes through a single low order resonance and one which passes through two.

8.5.1 Transitioning through a single resonance

We now examine the case of a canonical EMRI consisting of a $10^6 M_\odot$ primary and a $10 M_\odot$ secondary for a mass ratio of $\epsilon = 10^{-5}$. The inspiral has initial conditions $(p_0, e_0, x_0) = (7.045, 0.45, 0.8)$ and evolved until the semilatus rectum reaches the value $p = 4.74$. These values were chosen so that the inspiral would last a little over one year, and so that the inspiral passes through the low-order $|\kappa_z|/|\kappa_r = 2/3$ resonance. It is worth noting that the inspiral crosses through many other resonances which, in order of appearance, are the $7/10$, $2/3$, $5/8$, $3/5$, and $4/7$ resonances. Technically, there are infinitely many resonance crossings since natural numbers are dense in the reals. However, we only list resonances who's highest value is less than 10 since we truncate our Fourier expansions after the 10th coefficient. Thus, in our numerical implementation, the terms in the NIT and averaged equations of motion are not directly effected by any resonance with a value larger than 10. Moreover, the lowest order resonance is the $2/3$ resonance and so we expect that to have the largest effect on the inspiral. As such, we only use the near-resonance NIT to account for the $2/3$ resonance and neglect all others in order to understand the effect this will have on the accuracy of our inspiral calculations.

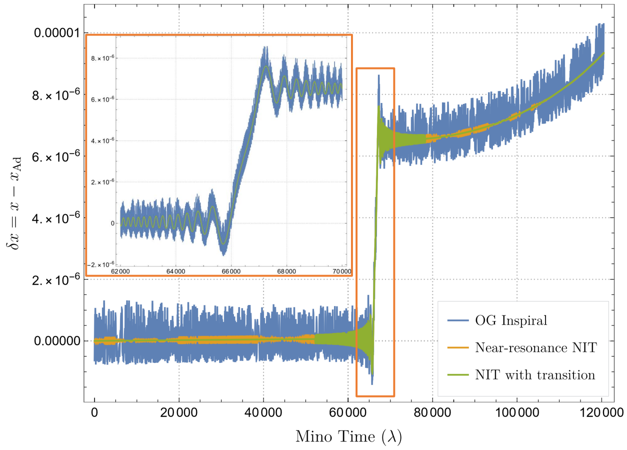
We first compute the year long inspiral using the OG equations using NDSolve with accuracy and precision goals set to 13.5 which took just over two days to compute on a single core of an Intel Xeon E5-2698V4 @ 2.20GHz. Using this as our point of comparison, we evolved inspirals with equivalent initial conditions and accuracy and precision goals utilizing the near-resonance equations of motion and using Eq.(8.23) to transition from the non-resonant NIT to the near-resonant NIT and back again.

We also evolved an adiabatic inspiral in order to subtract this contribution away from the post-adiabatic inspirals to show the effect of the resonance crossing on p,e and x which can be seen in Fig. 8.3. While the inspiral also crosses through other resonances, Fig. 8.3 makes it clear that the $2/3$ resonance has by far the largest effect on the orbital elements, with the effects of the other resonance crossings being far too small to resolve. Furthermore, this figure demonstrates how the near-resonance NIT and the NIT with a resonance transition both capture the "resonance jump" experienced by the OG inspiral while including far fewer orbital oscillations.



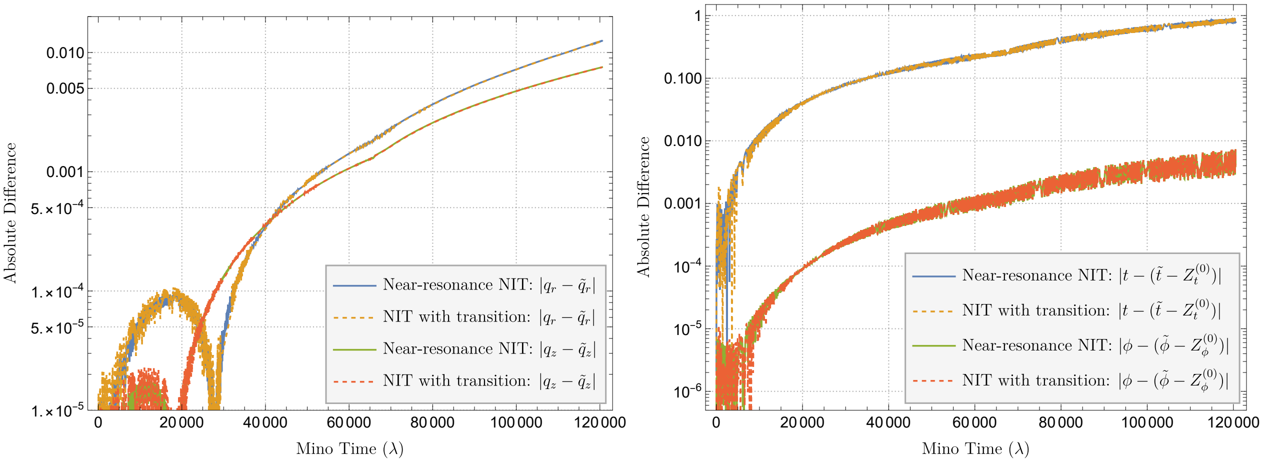
(a) The difference in p.

(b) The difference in e.



(c) The difference in x.

Figure 8.3: The difference between the evolution of orbital elements $\delta\vec{P} = \vec{P} - \vec{P}_{\textsf{Ad}}$, where $\vec{P} = (p, e, x)$, as a function of Mino time λ for a year long inspiral with $a = 0.9M$, $\epsilon = 10^{-5}$ and initial conditions $(p_0, e_0, x_0) = (7.045, 0.45, 0.8)$. One can see that all three methods for evolving the inspiral accurately capture the effect of the 2/3 orbital resonance, but the near-resonance NIT does so without resolving as many oscillations.



(a) Orbital phases.

(b) Extrinsic quantities.

Figure 8.4: The absolute difference in the phases and extrinsic quantities between the OG and NIT equations of motion for a year long inspiral with $a = 0.9M$, $\epsilon = 10^{-5}$ and initial conditions $(p_0, e_0, x_0) = (7.045, 0.45, 0.8)$. We observe that there is no difference in accuracy between using the near-resonance NIT for the entire inspiral and using the NIT away from resonances and transitioning to the resonant NIT in the vicinity of the resonance. $q_\perp$.

We also examine the effect that including a resonance transition has on the accuracy of the orbital phases and extrinsic quantities which is displayed in Fig. 8.4. We see a natural growth in the phase error over time, which may be due to accumulating numerical error due to the numerical solver, but may also be due to neglecting the effects of the other resonance crossings besides the $2/3$ resonance. Importantly, we see no difference in accuracy when using either the near-resonant NIT or using the NIT with a transition. Moreover, the end of a year long inspiral the difference in the phases is $< 2 \times 10^{-2}$ and the difference in t/M is less than 1, which should be accurate enough produce post-adiabatic waveforms fit for LISA data science.

This is confirmed by Table. 8.3. We see the runtime required to calculate inspirals using either the non-resonant NIT, the near-resonant NIT or the NIT with resonance transition. The table also lists the mismatch between the semi-relativistic quadrupole waveforms generated from these inspirals when compared to that produced by the OG inspiral. We find the non-resonant NIT to be the most computationally efficient but the resulting inspirals are far too inaccurate to be useful. However, the near-resonant NIT produces inspirals and waveforms that are accurate enough for LISA data science even when neglecting all other resonance crossings. Unfortunately, with a single inspiral taking 2097s or 35 minutes to compute, the near-resonance NIT is substantially faster than the OG inspiral but still much too slow for practical waveform generation for data analysis. Finally, we note that transitioning from the non-resonant NIT to the near-resonant NIT and back again provides the best of both approaches, producing inspirals and waveforms just as accurately as the near-resonant NIT while only taking $58.7s$ to compute an inspiral. This is still slower than the sub-second computation time that one would need for data analysis, but this can be reduced further by optimising the resonance condition and using more efficient numerical methods.

In conclusion, this test case has demonstrated two important insights. First, one does not need to account for every orbital resonance to produce waveforms that are sufficiently accurate for LISA science, only the lowest order ones. Second, using a combination of non-resonant and near-resonant

Inspiral	Runtime	Mismatch
Non-resonant NIT	4.02s	0.376
Near-resonant NIT	2097s	3.86×10^{-4}
NIT w/ transition	58.7s	3.84×10^{-4}

Table 8.3: A table of the time taken to compute a year long inspiral with mass ratio 10^{-5} and initial conditions $(p_0, e_0, x_0) = (7.045, 0.45, 0.8)$ using different equations of motion as calculated numerically using NDSolve with accuracy and precision goals both set to 13.5 as implemented in Mathematica 13 on an Intel Core i7 @ 2.2GHz. For reference, using the OG equations to compute the same inspiral took over two days to compute on a single core of an Intel Xeon E5-2698V4 @ 2.20GHz. The mismatch between the semi-relativistic quadrupole waveforms generated from these inspirals and the OG inspiral is also listed. We see that using the non-resonant NIT produces inspirals that are fast, but the resulting waveforms are completely inadequate. Taking into account the $2/3$ resonance using the near-resonance NIT produces significantly more accurate waveforms, but the inspiral calculation is much too slow for data analysis. Finally, using a transition between non-resonant and near-resonant NIT equations of motion retains all of the accuracy of just using the near-resonant NIT and is significantly faster to compute.

NIT equations of motion is the best strategy so far for accurately capturing resonant behaviour while reducing the computation time for calculating post-adiabatic inspiral trajectories.

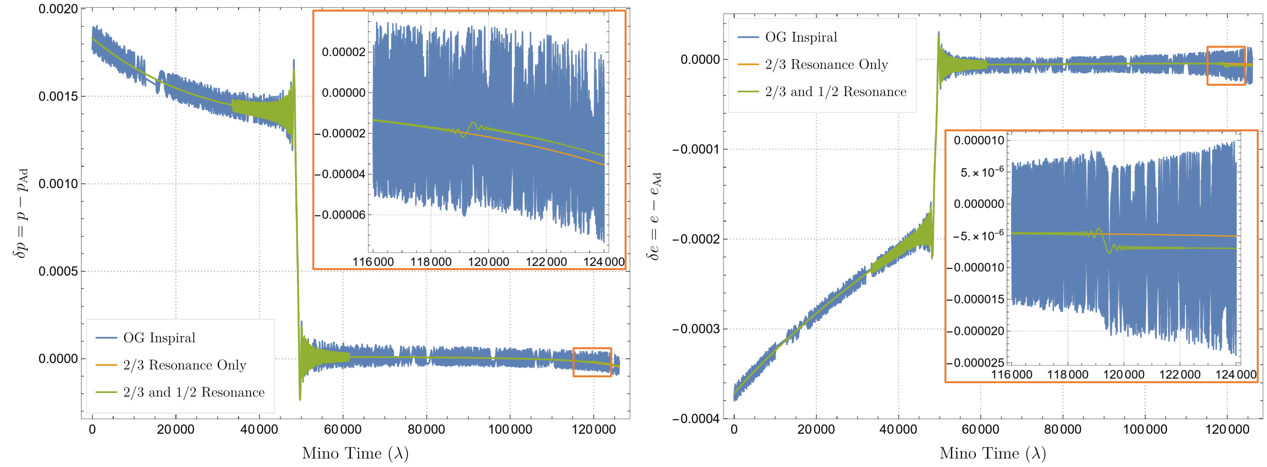
8.5.2 Transitioning through multiple resonances

Armed with these two insights, we now look to a case where there is more than one low-order resonance crossing. We wish to see if one can produce sufficiently accurate waveforms if one only employs a resonance transition for the $2/3$ resonance or if one needs to account for both resonances. Again, we pick a canonical EMRI mass ratio of $\epsilon = 10^{-5}$, and chose initial conditions $(p_0, e_0, x_0) = (6.8, 0.45, 0.8)$ and evolve until $p = 3.75$ such that the resulting inspiral lasts for just over 1 year. This time, however, the inspiral passes through the following of orbital resonances: $2/3, 5/8, 3/5, 4/7, 5/9, 1/2,$ and $4/9$. With the results of our last test in mind, we neglect all of the resonance crossings bar the $2/3$ and $1/2$ resonances. Using this inspiral, we wish to investigate if we can accurately transition through more than one resonance, and how much accuracy is lost if one accounts for the $2/3$ resonance but neglects the $1/2$ resonance.

We first compute the year long inspiral using the quasi-Keplerian OG equations using NDSolve with accuracy and precision goals set to 13.5 which took just over three days to compute on a single core of an Intel Xeon E5-2698V4 @ 2.20GHz. Using this as our point of comparison, we evolved inspirals with equivalent initial conditions and accuracy and precision goals utilizing the non-resonant NIT equations of motion with either a single near-resonance transition for the $2/3$ resonance, or with two near-resonance transitions for the $2/3$ and $1/2$ resonances.

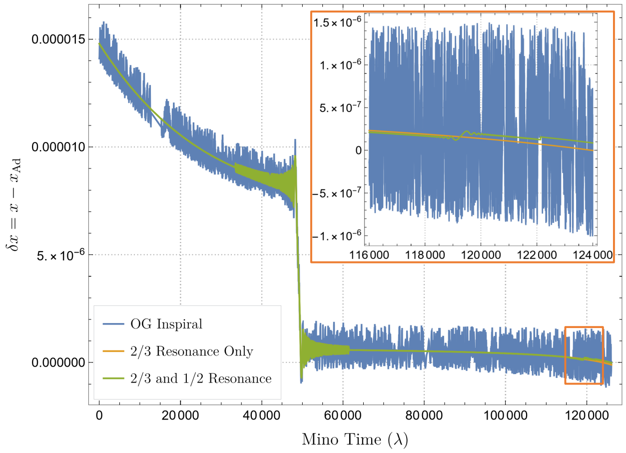
To demonstrate the effect of the two resonances on the evolution of the semilatus rectum p, we match an adiabatic inspiral to the OG inspiral after the $2/3$ resonance but before the $1/2$ resonance and evolve it both forward and backwards in time. We then subtract the adiabatic solution for the semilatus rectum p_{Ad} from the post-adiabatic inspirals solutions to illustrate the effects of both the $2/3$ and $1/2$ resonances as seen in Fig. 8.5. The figure shows how the $2/3$ has a significantly larger effect on the evolution of p than the $1/2$ resonance, and that while there is an error induced by neglecting the $1/2$ resonance, it is comparably small.

This is further supported when we look at the differences in the orbital phases and extrinsic quantities in Fig. 8.6. We see that neglecting the $1/2$ resonance induces a small but noticeable error in the orbital



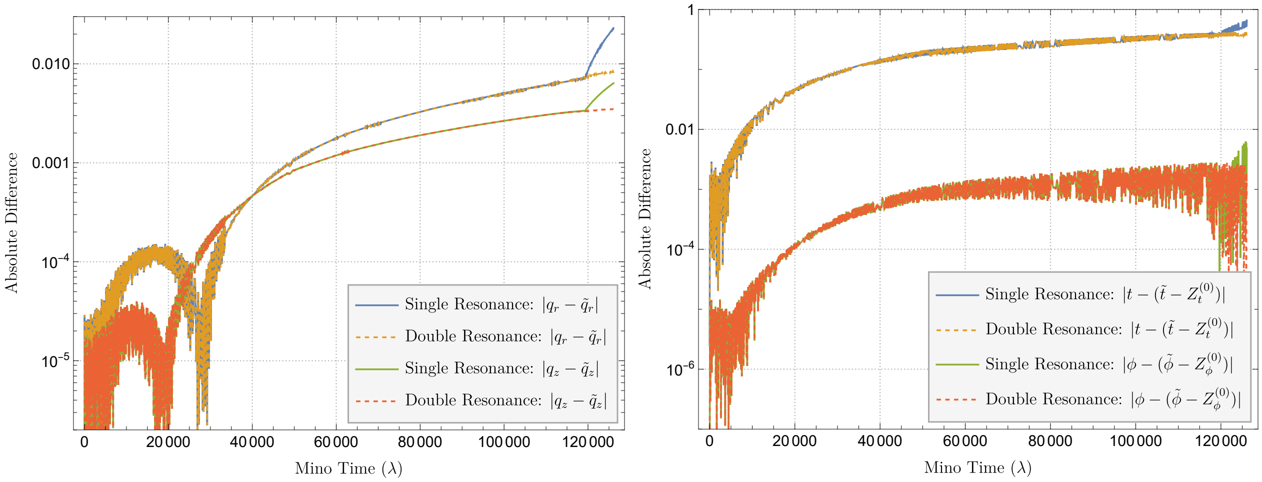
(a) The difference in p.

(b) The difference in e.



(c) The difference in x.

Figure 8.5: The difference between the evolution of orbital elements $\delta\vec{P} = \vec{P} - \vec{P}_{\text{Ad}}$ as a function of Mino time λ for a year long inspiral with $a = 0.9M$, $\epsilon = 10^{-5}$ and initial conditions $(p_0, e_0, x_0) = (6.8, 0.45, 0.8)$ using the OG equations of motion, the NIT equations of motion with a transition through the $2/3$ resonance and the NIT equations of motion with transitions for both the $2/3$ and $1/2$ resonances. This demonstrates importance of capturing the $2/3$ resonance and that the effect of neglecting the $1/2$ is minor by comparison.



(a) Orbital phases.

(b) Extrinsic quantities.

Figure 8.6: The absolute difference in the phases and extrinsic quantities between the OG and NIT equations of motion for a year long inspiral with $a = 0.9M$, $\epsilon = 10^{-5}$ and initial conditions $(p_0, e_0, x_0) = (6.8, 0.45, 0.8)$. One can clearly see the effect of neglecting the $1/2$ resonance.

phases and extrinsic quantities towards the end of the inspiral. Since the inspiral terminates shortly after the $1/2$ resonance crossing, this error remains small. However, if the inspiral were to be evolved for longer, this error will accumulate and may become substantial. This suggests that while incorporating the $1/2$ resonance may not provide a significant increase in accuracy for this particular inspiral, in general one may still need to account for it.

This is somewhat complicated by Table. 8.4, which shows the runtime and the waveform mismatch of each inspiral as compared with the waveform generated by the OG inspiral. We see that neglecting the $2/3$ resonance produces a waveform which agrees very poorly with the OG waveform. Including the transition through the $2/3$ resonance produces a waveform with a mismatch of only 2.67×10^{-4} which is significantly smaller than the 3×10^{-3} requirement to produce a waveform bank that can capture 90% of signals. Furthermore, including the transition through the $1/2$ resonance reduces this mismatch even further to 2.58×10^{-4}, but this minor improvement may not be worth more than doubling the runtime from $66s$ to $163s$.

In conclusion, both this and the previous test indicate that of the resonance crossings that we have examined, the most important to account for is the $2/3$ resonance. One may also need to account for other low order resonances such at the $1/2$ and $1/3$ resonances, but our preliminary results suggest that the $2/3$ resonance might be the only resonance that one must include to produce post-adiabatic waveforms accurate enough for LISA science. However, a more robust study of inspirals throughout more of the parameter space is needed before such a strong conclusion can be drawn.

Inspiral	Runtime	Mismatch
No resonance transition	6.31s	0.648
Single resonance transition	66.2s	2.67×10^{-4}
Double resonance transition	163.3s	2.58×10^{-4}

Table 8.4: A table of the time taken to compute a year long inspiral with mass ratio 10^{-5} and initial conditions $(p_0, e_0, x_0) = (6.8, 0.45, 0.8)$ using NIT equations of motion with different numbers of resonant transitions as calculated using NDSolve with accuracy and precision goals both set to 13.5 and implemented in Mathematica 13 on an Intel Core i7 @ 2.2GHz. For reference, using the OG equations to compute the same inspiral took just over three days to compute on a single core of an Intel Xeon E5-2698V4 @ 2.20GHz. The mismatch between the semi-relativistic quadrupole waveforms generated from these inspirals and the OG inspiral is also listed. Each resonance transition adds more computation time to the overall calculation, but it also increases the accuracy. The biggest increase in accuracy comes from doing a single resonance transition for the $2/3$ resonance. While accounting for the $1/2$ resonance does decrease the waveform mismatch, it only does so by a miniscule amount at the cost of more than doubling the computation time.

Chapter 9

Conclusion

The goal of this work was to develop a procedure that allowed for the rapid calculation of post-adiabatic EMRI trajectories with a spinning primary in a way that was both accurate and fast enough for LISA data science.

To this end, we first reviewed geodesic motion in Kerr spacetime. Using the analytic solutions to the geodesic motion in terms of action angles, we were able to implement an action angle formulation of the osculating geodesic equations for forced motion around a Kerr black hole. Inspirals evolved using these equations of motion were identical to those evolved using the null tetrad formulation derived and implemented in Ref. [137]. We plan to release a Mathematica package containing both versions of the osculating geodesics equations for both of these formulations which we be made available as part of the Black Hole Perturbation Toolkit [182].

Using the action angle osculating geodesics as a foundation, we reviewed the technique of near-identity averaging transformations for calculating self-forced inspirals. These transformations remove the dependence on the rapidly oscillating orbital phases, resulting in equations of motion that can be solved in less than a second. However, this technique had two weaknesses.

Firstly, the transformation ill-behaved in the presence of orbital resonances. We ameliorate this by deriving a near-resonance NIT which removes all oscillations apart from the resonant phase. While this means the equations of motion are slower to solve than using the original NIT, it still provides a substantial speed-up over using the OG equations while accurately capturing the resonant behaviour.

Secondly, the resulting solutions were not parametrized in terms of Boyer-Lindquist coordinate time, which is inconvenient for waveform production and data analysis. This was resolved by following the procedure outlined in Ref. [113] which involves applying an additional averaging transformation. Finally, while connections between NIT and the two-timescale expansion TTE were already well known [113,155], we explicitly show how one can recover the TTE equations of motion from the NIT equations of motion.

We now wish to investigate the effect of using these averaged equations of motion over the OG equations on the final waveform. and so we review various methods for calculating waveforms from an inspiral trajectory. In particular, we make heavy use of the semi-relativistic quadrupole formula as other methods have yet to be extended to Kerr inspirals. We also review various methods for comparing two different waveforms and benchmarks for waveform accuracy.

Up until this point, NITs had only ever been applied to self-forced inspirals in Schwarzschild spacetime [154, 156]. This work marks the first application of NITs to Kerr self-forced inspirals and investigated three important regions of the parameter space; eccentric and equatorial, circular and inclined (spherical), and generic with orbital resonances. We now present the conclusions from each of these investigations.

9.1 Eccentric Self-forced inspirals into a rotating black hole

In Chapter 6, we present the first self-forced inspirals in Kerr spacetime. We computed the self-force in the radiation gauge using the code of Ref. [99] and interpolated it over a region of the parameter space of eccentric, equatorial orbits using Chebyshev interpolation. Our model achieves sub-percent accuracy for the self-force across the two dimensional parameter space using only 105 points. This is a substantial improvement previous methods require $\mathcal{O}(10^3)$ points to achieve a comparable level of accuracy.

So far we have applied our method to strong-field regions of the parameter space for three values of the primary's spin ($a = 0, \pm 0.9M$). It remains as future work to interpolate over the spin of the primary, however, the Chebyshev interpolation method appears to be a promising approach to tiling data from expensive gravitational self-force codes across the 4-dimensional generic Kerr parameter space. This method could be further improved with the aid of a detailed of the study of the analytic structure of the GSF near the last stable orbit.

Combining this model with the OG equations allows us to produce the very first self-forced Kerr inspirals, but numerically calculating even a single inspiral can take minutes to hours. To overcome this, we combined the action angle OG formulation with the NIT and this method to the case of eccentric, equatorial inspirals, allowing us to calculate inspirals in less than a second.

We showed that our NIT quantities remain close to the original evolution variables throughout the inspiral at the expected order in the mass-ratio. When the mass ratio is less than $1:300$, we find the difference between year-long NIT and OG inspirals becomes significant for data analysis, reinforcing the findings of Ref. [156]. Note, however, that a priori it is not known which (the NIT or OG inspiral) is closer to the true inspiral, since both are accurate to the same order in the mass-ratio.

With our efficient NIT model of eccentric, equatorial inspirals we explored the effects of the gravitational self force. We find that prograde inspirals around a rapidly rotating black hole generally experience an additional periastron advance on top of the periastron advance induced by geodesic motion. This is in contrast to the "periastron retreat" experienced by retrograde inspirals and inspirals around non-rotating black holes [193].

The NIT equations of motion make it convenient to compare inspirals both with and without post-adiabatic effects included and we confirmed that without post-adiabatic effects, the orbital phases of a typical EMRI will incur an error of order $\mathcal{O}(\epsilon^0)$.

Moreover, by comparing inspirals under the influence of self-force models calculated in different gauges, we find that the resulting trajectories are gauge dependent. This difference due to gauge causes a de-phasing that is comparable in magnitude to not including any post-adiabatic effects. This suggests that in order to obtain gauge invariant post-adiabatic waveforms, one must also include second order self-force results.

9.2 Spherical self-forced inspirals into a rotating black hole

In Chapter 7, we present the very first calculations of the first order gravitational self-force in the radiation gauge for spherical orbits in Kerr spacetime by utilizing a modified version of the code found in refs [99, 100].

We then interpolated this data using Fourier decomposition and Chebyshev interpolation to produce a continuous model that for inclinations upto $0^\circ \leq \theta_{\text{inc}} \leq 45^\circ$ with sub-percent accuracy using only 162 points. This same method also allowed us to interpolate orbit averaged rate of change of energy and angular momentum from the asymptotic fluxes for all inclinations (both prograde and retrograde) with an accuracy of $\sim 10^{-6}$ using only 342 points. This is a significant improvement over other interpolation methods found in the literature which require at least an order of magnitude more points to achieve a comparable level of accuracy [70, 145].

This could be further improved with a better choice when rescaling the data before interpolation, ideally choosing a function informed by the leading order PN contributions in the weak fields and/or the analytic structure of the GSF near the ISSO. So far this model is only valid for orbits where the primary's spin is $a = 0.9M$ and so interpolating over the other values of spin is left for future work. However the work in this chapter, along with Chapter. 6, show that the Chebyshev interpolation methods are a promising approach for interpolating information from expensive GSF and flux codes across the vast 4-dimensional parameter space of generic Kerr inspirals.

Using our interpolated GSF model along with the osculating geodesics (OG) formulation specialised to the spherical case, as derived in Sec. 3.2.2, allows us to calculate the first ever spherical self-forced inspirals around a Kerr black hole. However numerically evolving these inspirals can take minutes to hours due to the need to resolve the $\sim 1/\epsilon$ orbital oscillations.

Once again, we overcome this by employing the technique of near-identity averaging transformations (NITs), which produce equations of motion that capture the correct long-term secular evolution of the binary but can also be rapidly numerically solved. Following the methodology of Ref. [113], we improve upon this formulation by employing a second averaging transformation such that the solutions to our equations of motion are parametrized by Boyer-linquist coordinate time instead of Mino time. Since Boyer-linquist time can be related to the time at the detector, this improved NIT procedure is much more convenient for generating waveforms and for data analysis. We also employ a two-timescale expansion (TTE) of the NIT equations of motion which factors out the dependence of the mass ratio, at the cost of doubling the number of equations to be solved.

We found that quantities calculated from either the NIT or TTE equations of motion remained close to the original evolution variables throughout the inspiral to the expected order in the mass-ratio. This technique works particularly well for spherical inspirals as we find the mismatch between waveforms calculated from NIT or TTE inspirals and waveforms calculated from OG inspirals is $\leq 10^{-3}$ even binaries with $\epsilon \sim 10^{-1}$.

Using our efficient inspiral trajectories, we investigate the effect of first order self-force on inspirals across the spherical Kerr parameter space. We find that the orbital separation decreases with time while the orbit becomes more inclined over time, which is consistent with the findings of adiabatic evolutions of spherical Kerr inspirals. We also find that neglecting the conservative effects can result in an $\mathcal{O}(1)$ radian de-phasing in the azimuthal phase, with this effect becoming larger with the inclination angle of the orbit.

Finally, we note that both NIT and TTE equations of motion allow us to further improve the accuracy of our waveforms by replacing the adiabatic terms with higher accuracy interpolants calculated from the asymptotic fluxes. Making this improvement can result in phase differences ranging from tens to tens of thousands of radians for multi-year long EMRIs. This highlights the necessity of efficient gravitational wave flux codes for both adiabatic and post-adiabatic EMRI waveforms [134].

9.3 Fast generic inspirals into a rotating black hole

In Chapter 8, we present the first application of near-identity averaging transformations to generic Kerr inspirals both far from and in the vicinity of low order orbital resonances. Generic Kerr GSF codes are too computational expensive to feasibly create an interpolated self-force model as was done in the previous two chapters. However, by combining the interpolated models for the eccentric equatorial GSF and the spherical GSF, we created a toy model for the generic GSF which is quick to evaluate and qualitatively has the behaviours one would expect of the generic GSF. It is this toy model, along with the OG equations, that we use to drive our generic Kerr inspirals.

We then use near-identity averaging transformations to speed up these calculations. We investigate how the accuracy and speed-up scales with the mass ratio and confirm that this technique works as expected, so long as one is not in the presence of a low order orbital resonance. Since the NIT becomes singular in the event of an orbital resonance, we implement a near-resonance NIT for the first time and test the scaling of the accuracy and speed-up with the mass ratio in the presence of the $(3,-2)$ resonance. From this, we find that while the difference between the near-resonance NIT and OG inspirals scale linearly with the mass ratio as we would expect, we find that the near-resonance NIT to be about an order of magnitude faster than the OG inspiral, but could still take seconds to minutes to compute a single inspiral.

We note that this could be reduced further by utilizing the non-resonant NIT when far from a resonance and then transitioning to the near-resonance NIT when in the vicinity of a resonance. We test this procedure using two different year long EMRI trajectories; one which only evolves through the $2/3$ resonance, and one which also evolves through the $1/2$ resonance. From these tests, we confirm that one can use the transition procedure with no apparent loss of accuracy compared to only using the near-resonance NIT. Moreover, one can safely neglect higher order resonances without any significant loss of accuracy. Our results even suggest that one might be able to ignore all resonances bar the $2/3$ resonance since it has by far the largest impact on the inspiral, though further investigation is needed to ensure that this is really the case across the parameter space.

However, even with the speed-up from utilising both the non-resonant and near-resonant NITs, our current implementation still takes 1-2 minutes to compute a year long EMRI evolving through a single resonance, which, while a drastic improvement over the days required by the OG equations, is still not fast enough for LISA data analysis. This can be reduced further by investigating and optimising the condition for what is considered "near-resonance", or by implementing the above procedure in a compiled language such as C/C++, though this still may not be fast enough. It is possible that a NIT model with an empirically fitted resonant "jump" derived from the near-resonant NIT (similar to that used for tidal resonances [163]) could account for resonances while minimising runtime, but this will be left as future work.

9.4 Future Work

Currently our first order self-force models are only valid in a small slice of the EMRI parameter space. In particular, it will be important to interpolate over different values of the primary spin and investigate how this effects the GSF. We are currently limited by expensive codes and so future flux and GSF codes will need to become more efficient for the practical production of fast and extensive EMRI waveforms.

We also note that our work thus far only incorporates first order GSF results. Post-adiabatic waveforms will require second order results to not only reach $\mathcal{O}(1)$ accuracy in the phases [67], but also to produce gauge invariant waveforms, as was noted in Chapter 6. Thus, the inspirals and waveforms produced here are only representative of the gauge in which the self-force was calculated, which in our case is the out-going radiation gauge. While second order results are not currently available for Kerr orbits, our framework can accommodate them when they become available. However, since second order results are calculated in the two-timescale framework, some translation may be needed.

We are also missing any post-adiabatic contributions due to the spin of the secondary [194–199]. Incorporating the linear in spin contribution to the energy and angular momentum fluxes into our averaged equations should be straightforward once they become available for generic orbits. Similarly, now that spin perturbed frequencies due to conservative effects from the Mathisson-Papapetrou-Dixon equations are available for generic Kerr orbits [198, 199], it should be possible to fold these effects into our framework, which will be the subject of a future project.

Currently, we terminate our inspiral just before the last stable orbit before plunge, as after this the adiabaticity assumption of the method of osculating geodesics breaks down. While the final section of EMRI would make up a small fraction of the total SNR, for completeness one could also attach a "transition to plunge" procedure [138–141], especially if one wishes to generalise this approach to IMRIs or comparable mass binaries, though this would make the overall model more expensive.

Moreover, we are assuming that the EMRI is not subject to any environmental effects which may have an effect on the inspiral, such as the presence of matter [200] or an accretion disk [201]. If such environmental effects can be expressed in terms of the secondary's four-acceleration, then it should be possible to incorporate effects into our efficient inspiral calculations and understand their impact on EMRI trajectories and waveforms. This sort of investigation has already been done in the case of an external third body perturbing an EMRI system [162], and it was found to only have a noticeable effect at "tidal resonances" when all three orbital frequencies are related by a small number ratio [143]. These tidal resonance effects have now been modelled [163] and can in principal be incorporated into our framework.

We note we have used the semi-relativistic quadrupole formula to generate the waveforms from the OG, NIT and TTE inspirals. This is sufficient for this work as we only wish to compare the difference in the waveforms caused by different inspiral calculations. However LISA data analysis will require fully-relativistic waveform amplitudes such as those currently in the FastEMRIWaveforms (FEW) package for Schwarzschild inspirals [135]. Currently FEW only uses adiabatic inspirals, but this can be improved by employing either our NIT or TTE equations of motion. Once the waveform amplitudes have been interpolated for Kerr inspirals, they can be combined immediately with the implementation presented in this work.

Interfacing with FEW will also necessitate reimplementing the procedures outlined in this work into a compiled language such as C/C++ with a Python wrapper, as opposed to our current implementation which is in Wolfram Mathematica. Not only would this make efficient EMRI inspiral calculations more accessible to the data analysis community, but it will also come with an order of magnitude worth of speed-up. This should make our inspiral calculations efficient enough for Bayesian parameter estimation to be feasible, which will allow us to more rigorously investigate the systematic biases introduced by adding or neglecting certain effects on the inspiral [120, 201, 202].

While there is still more work that needs to be done, the work presented in this thesis stands as an important step towards bridging the gap between the modelling of extreme mass ratio inspirals and LISA data science.

Appendix A

Self-force corrections to the periapsis advance around a spinning black hole

The periapsis advance is a observable quantity that has been used to compare models of compact binary dynamics [193]. The effect of the gravitational self-force on this observable for quasi-circular EMRIs around a rotating primary was explored in Ref. [85]. One important insight, that is present in the supplemental material of that work, is the effect that the spin of the primary and orbital radius has on the self-force correction to the rate of periapsis advance. For completeness, we highlight this result in this appendix.

For quasi-circular inspirals the relation between the dimensionless quantity $W = \Omega_r^2/\Omega_\phi^2$ and $\Omega_\phi^{(0)}$ is an important benchmark for comparing between different calculational approaches to the two-body problem. [85, 193, 203]. The linear in mass ratio correction to the quantity is defined via

$$W(\epsilon; a, \Omega_\phi^{(0)}) = W(0; a, \Omega_\phi^{(0)}) + \epsilon\rho(a, \Omega_\phi^{(0)}) + \mathcal{O}(\epsilon^2), \tag{A.1}$$

where $W(0; a, \Omega_\phi^{(0)})$ is the background value for the periapsis advance, and $\rho(a, \Omega_\phi^{(0)})$ is the correction induced by the first-order gravitational self-force.

Fig. A.1 demonstrates how ρ varies as a function of orbital radius r and the spin of the primary a. We plot the ratio r_{ISCO}/r, where r_{ISCO} is the radius of the innermost stable circular orbit (ISCO). This ratio is convenient for plotting the results as goes from 1 at the ISCO for all spin values and asymptotically approaches zero as r grows large. As one would expect, the plot demonstrates that this correction grows larger as the radius of the inspiral approaches the ISCO. This correction is positive for all retrograde orbits and in the strong field for prograde orbits. This means that self-force typically acts against the periapsis advance caused by the background geodesic motion, resulting in a reduction of the observed periapsis advance of the binary. However, for positive spins and at large radii, there is a region of the parameter space (in blue) where this correction is negative, meaning that the self-force increases the observed rate of periapsis advance compared to the background geodesic motion. The larger the spin, the smaller the radii at which this effect occurs. As such, this effect is most prominent for prograde orbits around rapidly rotating black holes.

We find that the effect of the conservative self-force on the orbital phase for eccentric inspirals is consistent with the sign of the self-force induced rate of periastron advance, ρ, in the quasi-circular limit – see Sec. 6.5.3.

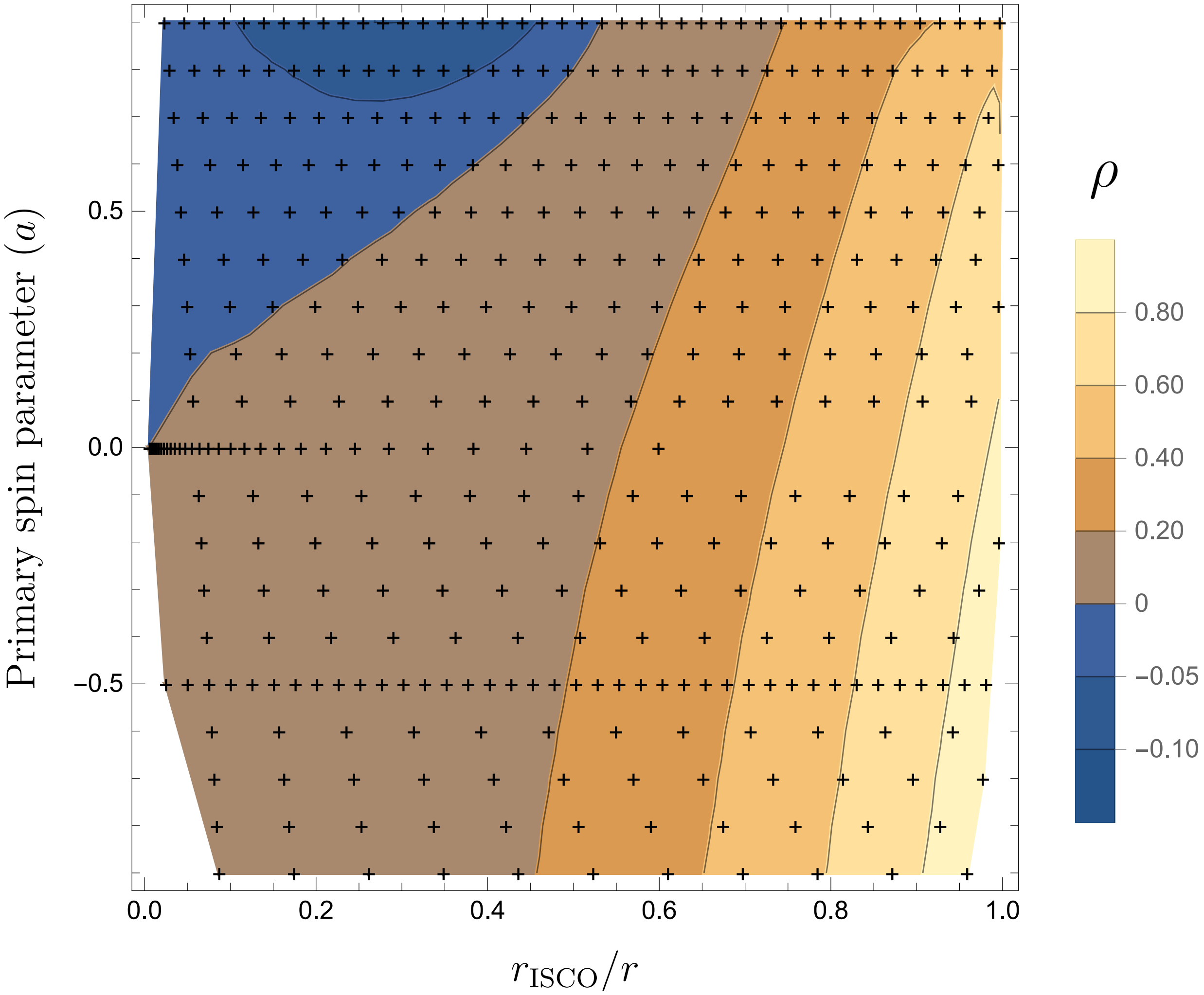


Figure A.1: The linear in mass ratio correction, ρ, to the periapsis advance, W, as a function of distance from the innermost stable circular orbit (ISCO) r_{ISCO}/r and the spin of the primary, a. The contours show that ρ grows larger as the radius of the inspiral approaches the ISCO. ρ is positive for most of the parameter space, including for all retrograde orbits, implying that in these regions the self-force acts against the geodesic periapsis advance. However, for all prograde orbits, there is a region (in blue) where this correction is negative meaning the self-force instead increases the rate of periapsis advance. The region where this occurs grows larger as a increases The black crosses mark the location of the underlying data from Ref. [85] used to calculate the contour plot.

Bibliography

[1] P. Lynch, M. van de Meent, and N. Warburton, "Eccentric self-forced inspirals into a rotating black hole," *Class. Quant. Grav.* **39** no. 14, (2022) 145004, `arXiv:2112.05651 [gr-qc]`.

[2] **LIGO Scientific, Virgo** Collaboration, B. P. Abbott *et al.*, "Observation of Gravitational Waves from a Binary Black Hole Merger," *Phys. Rev. Lett.* **116** no. 6, (2016) 061102, `arXiv:1602.03837 [gr-qc]`.

[3] C. W. Misner, K. S. Thorne, and J. A. Wheeler, *Gravitation*. W. H. Freeman, San Francisco, 1973.

[4] **LIGO Scientific, Virgo** Collaboration, B. P. Abbott *et al.*, "GWTC-1: A Gravitational-Wave Transient Catalog of Compact Binary Mergers Observed by LIGO and Virgo during the First and Second Observing Runs," *Phys. Rev. X* **9** no. 3, (2019) 031040, `arXiv:1811.12907 [astro-ph.HE]`.

[5] **LIGO Scientific, Virgo** Collaboration, R. Abbott *et al.*, "GWTC-2: Compact Binary Coalescences Observed by LIGO and Virgo During the First Half of the Third Observing Run," *Phys. Rev. X* **11** (2021) 021053, `arXiv:2010.14527 [gr-qc]`.

[6] **LIGO Scientific, VIRGO, KAGRA** Collaboration, R. Abbott *et al.*, "GWTC-3: Compact Binary Coalescences Observed by LIGO and Virgo During the Second Part of the Third Observing Run," `arXiv:2111.03606 [gr-qc]`.

[7] **LIGO Scientific, VIRGO, KAGRA** Collaboration, R. Abbott *et al.*, "The population of merging compact binaries inferred using gravitational waves through GWTC-3," `arXiv:2111.03634 [astro-ph.HE]`.

[8] **LIGO Scientific, VIRGO, KAGRA** Collaboration, R. Abbott *et al.*, "Tests of General Relativity with GWTC-3," `arXiv:2112.06861 [gr-qc]`.

[9] M. Maggiore *et al.*, "Science Case for the Einstein Telescope," *JCAP* **03** (2020) 050, `arXiv:1912.02622 [astro-ph.CO]`.

[10] E. D. Hall *et al.*, "Gravitational-wave physics with Cosmic Explorer: Limits to low-frequency sensitivity," *Phys. Rev. D* **103** no. 12, (2021) 122004, `arXiv:2012.03608 [gr-qc]`.

[11] S. Kawamura *et al.*, "Current status of space gravitational wave antenna DECIGO and B-DECIGO," *PTEP* **2021** no. 5, (2021) 05A105, `arXiv:2006.13545 [gr-qc]`.

[12] J. Mei *et al.*, "The TianQin project: Current progress on science and technology," *Progress of Theoretical and Experimental Physics* **2021** no. 5, (08, 2020) , `https://academic.oup.com/ptep/article-pdf/2021/5/05A107/37953035/ptaa114.pdf`. `https://doi.org/10.1093/ptep/ptaa114`. 05A107.

[13] Z. Luo, Z. Guo, G. Jin, Y. Wu, and W. Hu, "A brief analysis to taiji: Science and technology," *Results in Physics* **16** (2020) 102918. `https://www.sciencedirect.com/science/article/pii/S2211379719321928`.

[14] J. Baker *et al.*, "The Laser Interferometer Space Antenna: Unveiling the Millihertz Gravitational Wave Sky," `arXiv:1907.06482 [astro-ph.IM]`.

[15] P. Amaro-Seoane *et al.*, "Low-frequency gravitational-wave science with eLISA/NGO," *Class. Quant. Grav.* **29** (2012) 124016, `arXiv:1202.0839 [gr-qc]`.

[16] S. Babak, J. Gair, A. Sesana, E. Barausse, C. F. Sopuerta, C. P. Berry, E. Berti, P. Amaro-Seoane, A. Petiteau, and A. Klein, "Science with the space-based interferometer LISA. V. Extreme mass-ratio inspirals," *Phys. Rev. D* **95** no. 10, (Mar, 2017) , `arXiv:1703.09722`. `http://arxiv.org/abs/1703.09722http://dx.doi.org/10.1103/PhysRevD.95.103012`.

[17] S. Drasco and S. A. Hughes, "Gravitational wave snapshots of generic extreme mass ratio inspirals," *Phys. Rev. D* **73** no. 2, (2006) 024027, `arXiv:gr-qc/0509101`. [Erratum: Phys.Rev.D 88, 109905 (2013), Erratum: Phys.Rev.D 90, 109905 (2014)].

[18] J. Kormendy and D. Richstone, "Inward bound: The Search for supermassive black holes in galactic nuclei," *Ann. Rev. Astron. Astrophys.* **33** (1995) 581.

[19] J. Kormendy and L. C. Ho, "Coevolution (Or Not) of Supermassive Black Holes and Host Galaxies," *Ann. Rev. Astron. Astrophys.* **51** (2013) 511–653, `arXiv:1304.7762 [astro-ph.CO]`.

[20] A. M. Ghez, S. Salim, S. D. Hornstein, A. Tanner, M. Morris, E. E. Becklin, and G. Duchene, "Stellar orbits around the galactic center black hole," *Astrophys. J.* **620** (2005) 744–757, `arXiv:astro-ph/0306130`.

[21] R. Genzel, F. Eisenhauer, and S. Gillessen, "The Galactic Center massive black hole and nuclear star cluster," *Reviews of Modern Physics* **82** no. 4, (Oct., 2010) 3121–3195, `arXiv:1006.0064 [astro-ph.GA]`.

[22] **Event Horizon Telescope** Collaboration, K. Akiyama *et al.*, "First M87 Event Horizon Telescope Results. VI. The Shadow and Mass of the Central Black Hole," *Astrophys. J. Lett.* **875** no. 1, (2019) L6, `arXiv:1906.11243 [astro-ph.GA]`.

[23] **Event Horizon Telescope** Collaboration, K. Akiyama *et al.*, "First Sagittarius A* Event Horizon Telescope Results. VI. Testing the Black Hole Metric," *Astrophys. J. Lett.* **930** no. 2, (2022) L17.

[24] S. A. Hughes, "Trust but verify: The Case for astrophysical black holes," *eConf* **C0507252** (2005) L006, `arXiv:hep-ph/0511217`.

[25] M. J. Rees, "Tidal disruption of stars by black holes of 10 to the 6th-10 to the 8th solar masses in nearby galaxies," *Nature* **333** (1988) 523–528.

[26] M. Camenzind, *Compact objects in astrophysics : white dwarfs, neutron stars, and black holes*. 2007.

[27] W. R. Brown, "Hypervelocity stars," *Annual Review of Astronomy and Astrophysics* **53** no. 1, (2015) 15–49, `https://doi.org/10.1146/annurev-astro-082214-122230`. `https://doi.org/10.1146/annurev-astro-082214-122230`.

[28] S. Babak, J. R. Gair, and R. H. Cole, "Extreme mass ratio inspirals: perspectives for their detection," *Fund. Theor. Phys.* **179** (2015) 783–812, `arXiv:1411.5253 [gr-qc]`.

[29] P. Amaro-Seoane, "The gravitational capture of compact objects by massive black holes," `arXiv:2011.03059 [gr-qc]`.

[30] J. R. Gair, L. Barack, T. Creighton, C. Cutler, S. L. Larson, E. S. Phinney, and M. Vallisneri, "Event rate estimates for LISA extreme mass ratio capture sources," *Class. Quant. Grav.* **21** (2004) S1595–S1606, `arXiv:gr-qc/0405137`.

[31] J. R. Gair, M. Vallisneri, S. L. Larson, and J. G. Baker, "Testing General Relativity with Low-Frequency, Space-Based Gravitational-Wave Detectors," *Living Rev. Rel.* **16** (2013) 7, `arXiv:1212.5575 [gr-qc]`.

[32] E. Barausse *et al.*, "Prospects for Fundamental Physics with LISA," *Gen. Rel. Grav.* **52** no. 8, (2020) 81, `arXiv:2001.09793 [gr-qc]`.

[33] C. P. L. Berry, S. A. Hughes, C. F. Sopuerta, A. J. K. Chua, A. Heffernan, K. Holley-Bockelmann, D. P. Mihaylov, M. C. Miller, and A. Sesana, "The unique potential of extreme mass-ratio inspirals for gravitational-wave astronomy," `arXiv:1903.03686`. `http://arxiv.org/abs/1903.03686`.

[34] D. Laghi, N. Tamanini, W. Del Pozzo, A. Sesana, J. Gair, S. Babak, and D. Izquierdo-Villalba, "Gravitational-wave cosmology with extreme mass-ratio inspirals," *Mon. Not. Roy. Astron. Soc.* **508** no. 3, (2021) 4512–4531, `arXiv:2102.01708 [astro-ph.CO]`.

[35] A. Einstein, "Die grundlage der allgemeinen relativitätstheorie," *Annalen der Physik* **354** no. 7, (1916) 769–822, `https://onlinelibrary.wiley.com/doi/pdf/10.1002/andp.19163540702`. `https://onlinelibrary.wiley.com/doi/abs/10.1002/andp.19163540702`.

[36] E. Poisson and C. M. Will, *Gravity: Newtonian, Post-Newtonian, Relativistic*. Cambridge University Press, 2014.

[37] D. R. Lorimer, "Binary and Millisecond Pulsars," *Living Rev. Rel.* **11** (2008) 8, `arXiv:0811.0762 [astro-ph]`.

[38] S. Isoyama, R. Sturani, and H. Nakano, "Post-Newtonian templates for gravitational waves from compact binary inspirals," `arXiv:2012.01350 [gr-qc]`.

[39] J. Blümlein, A. Maier, P. Marquard, and G. Schäfer, "Fourth post-Newtonian Hamiltonian dynamics of two-body systems from an effective field theory approach," *Nucl. Phys. B* **955** (2020) 115041, `arXiv:2003.01692 [gr-qc]`.

[40] L. Blanchet, "Gravitational Radiation from Post-Newtonian Sources and Inspiralling Compact Binaries," *Living Rev. Rel.* **17** (2014) 2, `arXiv:1310.1528 [gr-qc]`.

[41] T. Damour, "Gravitational scattering, post-Minkowskian approximation and Effective One-Body theory," *Phys. Rev. D* **94** no. 10, (2016) 104015, `arXiv:1609.00354 [gr-qc]`.

[42] J. Vines, "Scattering of two spinning black holes in post-Minkowskian gravity, to all orders in spin, and effective-one-body mappings," *Class. Quant. Grav.* **35** no. 8, (2018) 084002, `arXiv:1709.06016 [gr-qc]`.

[43] A. Guevara, A. Ochirov, and J. Vines, "Scattering of Spinning Black Holes from Exponentiated Soft Factors," *JHEP* **09** (2019) 056, `arXiv:1812.06895 [hep-th]`.

[44] B. Maybee, D. O'Connell, and J. Vines, "Observables and amplitudes for spinning particles and black holes," *JHEP* **12** (2019) 156, `arXiv:1906.09260 [hep-th]`.

[45] Z. Bern, C. Cheung, R. Roiban, C.-H. Shen, M. P. Solon, and M. Zeng, "Black Hole Binary Dynamics from the Double Copy and Effective Theory," *JHEP* **10** (2019) 206, `arXiv:1908.01493 [hep-th]`.

[46] L. Smarr, A. Čadež, B. DeWitt, and K. Eppley, "Collision of two black holes: Theoretical framework," *Phys. Rev. D* **14** (Nov, 1976) 2443–2452. `https://link.aps.org/doi/10.1103/PhysRevD.14.2443`.

[47] R. A. Matzner, H. E. Seidel, S. L. Shapiro, L. Smarr, W.-M. Suen, S. A. Teukolsky, and J. Winicour, "Geometry of a black hole collision," *Science* **270** no. 5238, (1995) 941–947, `https://www.science.org/doi/pdf/10.1126/science.270.5238.941`. `https://www.science.org/doi/abs/10.1126/science.270.5238.941`.

[48] F. Pretorius, "Evolution of binary black hole spacetimes," *Phys. Rev. Lett.* **95** (2005) 121101, `arXiv:gr-qc/0507014`.

[49] M. Campanelli, C. O. Lousto, P. Marronetti, and Y. Zlochower, "Accurate evolutions of orbiting black-hole binaries without excision," *Phys. Rev. Lett.* **96** (2006) 111101, `arXiv:gr-qc/0511048`.

[50] J. G. Baker, J. Centrella, D.-I. Choi, M. Koppitz, and J. van Meter, "Gravitational wave extraction from an inspiraling configuration of merging black holes," *Phys. Rev. Lett.* **96** (2006) 111102, `arXiv:gr-qc/0511103`.

[51] M. Boyle *et al.*, "The SXS Collaboration catalog of binary black hole simulations," *Class. Quant. Grav.* **36** no. 19, (2019) 195006, `arXiv:1904.04831 [gr-qc]`.

[52] H. P. Pfeiffer, "Numerical simulations of compact object binaries," *Class. Quant. Grav.* **29** (2012) 124004, `arXiv:1203.5166 [gr-qc]`.

[53] V. Varma, S. E. Field, M. A. Scheel, J. Blackman, L. E. Kidder, and H. P. Pfeiffer, "Surrogate model of hybridized numerical relativity binary black hole waveforms," *Phys. Rev. D* **99** no. 6, (2019) 064045, `arXiv:1812.07865 [gr-qc]`.

[54] V. Varma, S. E. Field, M. A. Scheel, J. Blackman, D. Gerosa, L. C. Stein, L. E. Kidder, and H. P. Pfeiffer, "Surrogate models for precessing binary black hole simulations with unequal masses," *Phys. Rev. Research.* **1** (2019) 033015, `arXiv:1905.09300 [gr-qc]`.

[55] M. Dhesi, H. R. Rüter, A. Pound, L. Barack, and H. P. Pfeiffer, "Worldtube excision method for intermediate-mass-ratio inspirals: Scalar-field toy model," *Phys. Rev. D* **104** no. 12, (2021) 124002, `arXiv:2109.03531 [gr-qc]`.

[56] J. Yoo, V. Varma, M. Giesler, M. A. Scheel, C.-J. Haster, H. P. Pfeiffer, L. E. Kidder, and M. Boyle, "Targeted large mass ratio numerical relativity surrogate waveform model for GW190814," *Phys. Rev. D* **106** no. 4, (2022) 044001, `arXiv:2203.10109 [gr-qc]`.

[57] C. O. Lousto and J. Healy, "Study of the Intermediate Mass Ratio Black Hole Binary Merger up to 1000:1 with Numerical Relativity," `arXiv:2203.08831 [gr-qc]`.

[58] A. Buonanno and T. Damour, "Effective one-body approach to general relativistic two-body dynamics," *Phys. Rev. D* **59** (1999) 084006, `arXiv:gr-qc/9811091`.

[59] T. Damour and A. Nagar, "The Effective One Body description of the Two-Body problem," *Fundam. Theor. Phys.* **162** (2011) 211–252, `arXiv:0906.1769 [gr-qc]`.

[60] A. Buonanno, G. B. Cook, and F. Pretorius, "Inspiral, merger and ring-down of equal-mass black-hole binaries," *Phys. Rev. D* **75** (2007) 124018, `arXiv:gr-qc/0610122`.

[61] **LIGO Scientific, Virgo** Collaboration, B. P. Abbott *et al.*, "A guide to LIGO–Virgo detector noise and extraction of transient gravitational-wave signals," *Class. Quant. Grav.* **37** no. 5, (2020) 055002, `arXiv:1908.11170 [gr-qc]`.

[62] A. Albertini, A. Nagar, A. Pound, N. Warburton, B. Wardell, L. Durkan, and J. Miller, "Comparing second-order gravitational self-force and effective one body waveforms from inspiralling, quasi-circular and nonspinning black hole binaries II: the large-mass-ratio case," `arXiv:2208.02055 [gr-qc]`.

[63] A. Nagar and S. Albanesi, "Towards a gravitational self force-informed effective-one-body waveform model for nonprecessing, eccentric, large-mass-ratio inspirals," `arXiv:2207.14002 [gr-qc]`.

[64] E. Poisson, A. Pound, and I. Vega, "The Motion of point particles in curved spacetime," *Living Rev. Rel.* **14** (2011) 7, `arXiv:1102.0529 [gr-qc]`.

[65] M. Van De Meent and H. P. Pfeiffer, "Intermediate Mass-Ratio Black Hole Binaries: Applicability of Small Mass-Ratio Perturbation Theory," *Phys. Rev. Lett.* **125** no. 18, (Jun, 2020) , `arXiv:2006.12036`. `https://arxiv.org/abs/2006.12036v2`.

[66] B. Wardell, A. Pound, N. Warburton, J. Miller, L. Durkan, and A. Le Tiec, "Gravitational waveforms for compact binaries from second-order self-force theory," `arXiv:2112.12265 [gr-qc]`.

[67] T. Hinderer and É. É. Flanagan, "Two-timescale analysis of extreme mass ratio inspirals in Kerr spacetime: Orbital motion," *Phys. Rev. D - Part. Fields, Gravit. Cosmol.* **78** no. 6, (Sep, 2008) , `arXiv:0805.3337`.

[68] E. E. Flanagan and T. Hinderer, "Transient resonances in the inspirals of point particles into black holes," *Phys. Rev. Lett.* **109** (2012) 071102, `arXiv:1009.4923 [gr-qc]`.

[69] J. Thornburg, "The Capra Research Program for Modelling Extreme Mass Ratio Inspirals," *GW Notes* **5** (2011) 3–53, `arXiv:1102.2857 [gr-qc]`.

[70] T. Osburn, N. Warburton, and C. R. Evans, "Highly eccentric inspirals into a black hole," *Phys. Rev. D* **93** no. 6, (Mar, 2016) , `arXiv:1511.01498`.

[71] L. Barack, "Gravitational self force in extreme mass-ratio inspirals," *Class. Quant. Grav.* **26** (2009) 213001, `arXiv:0908.1664 [gr-qc]`.

[72] A. Pound, "Conservative effect of the second-order gravitational self-force on quasicircular orbits in Schwarzschild spacetime," *Phys. Rev. D* **90** no. 8, (2014) 084039, `arXiv:1404.1543 [gr-qc]`.

[73] D. Bini and T. Damour, "Conservative second-order gravitational self-force on circular orbits and the effective one-body formalism," *Phys. Rev. D* **93** no. 10, (2016) 104040, `arXiv:1603.09175 [gr-qc]`.

[74] C. Cutler, D. Kennefick, and E. Poisson, "Gravitational radiation reaction for bound motion around a Schwarzschild black hole," *Phys. Rev. D* **50** (1994) 3816–3835.

[75] K. Glampedakis and D. Kennefick, "Zoom and whirl: Eccentric equatorial orbits around spinning black holes and their evolution under gravitational radiation reaction," *Phys. Rev. D* **66** no. 4, (2002) , `arXiv:0203086 [gr-qc]`.

[76] R. Fujita and M. Shibata, "Extreme mass ratio inspirals on the equatorial plane in the adiabatic order," *Phys. Rev. D* **102** no. 6, (2020) 064005, `arXiv:2008.13554 [gr-qc]`.

[77] S. Isoyama, R. Fujita, A. J. K. Chua, H. Nakano, A. Pound, and N. Sago, "Adiabatic waveforms from extreme-mass-ratio inspirals: an analytical approach," `arXiv:2111.05288 [gr-qc]`.

[78] F. D. Ryan, "Effect of gravitational radiation reaction on nonequatorial orbits around a Kerr black hole," *Phys. Rev. D* **53** (1996) 3064–3069, `arXiv:gr-qc/9511062`.

[79] S. A. Hughes, "Evolution of circular, nonequatorial orbits of Kerr black holes due to gravitational-wave emission," *Phys. Rev. D - Part. Fields, Gravit. Cosmol.* **61** no. 8, (Mar, 2000) 084004, `arXiv:9910091 [gr-qc]`. `https://link.aps.org/doi/10.1103/PhysRevD.61.084004`.

[80] S. A. Hughes, "Evolution of circular, nonequatorial orbits of Kerr black holes due to gravitational-wave emission. II. Inspiral trajectories and gravitational waveforms," *Phys. Rev. D* **64** no. 6, (Apr, 2001) 15, `arXiv:0104041 [gr-qc]`. `http://arxiv.org/abs/gr-qc/0104041http://dx.doi.org/10.1103/PhysRevD.64.064004`.

[81] L. Barack and N. Sago, "Gravitational self-force correction to the innermost stable circular orbit of a schwarzschild black hole," *Phys. Rev. Lett.* **102** no. 19, (Feb, 2009) , `arXiv:0902.0573`. `https://arxiv.org/abs/0902.0573v2`.

[82] L. Barack, T. Damour, and N. Sago, "Precession effect of the gravitational self-force in a Schwarzschild spacetime and the effective one-body formalism," *Phys. Rev. D - Part. Fields, Gravit. Cosmol.* **82** no. 8, (Aug, 2010) , `arXiv:1008.0935`. `http://arxiv.org/abs/1008.0935http://dx.doi.org/10.1103/PhysRevD.82.084036`.

[83] L. Barack and N. Sago, "Beyond the geodesic approximation: Conservative effects of the gravitational self-force in eccentric orbits around a Schwarzschild black hole," *Phys. Rev. D - Part. Fields, Gravit. Cosmol.* **83** no. 8, (Jan, 2011) , `arXiv:1101.3331`. `http://arxiv.org/abs/1101.3331http://dx.doi.org/10.1103/PhysRevD.83.084023`.

[84] N. Warburton and L. Barack, "Self-force on a scalar charge in Kerr spacetime: Eccentric equatorial orbits," *Phys. Rev. D - Part. Fields, Gravit. Cosmol.* **83** no. 12, (Mar, 2011) , `arXiv:1103.0287`. `https://arxiv.org/abs/1103.0287v3`.

[85] M. Van De Meent, "Self-Force Corrections to the Periapsis Advance around a Spinning Black Hole," *Phys. Rev. Lett.* **118** no. 1, (Oct, 2017) , `arXiv:1610.03497`. `http://arxiv.org/abs/1610.03497http://dx.doi.org/10.1103/PhysRevLett.118.011101`.

[86] J. Vines and E. E. Flanagan, "Is motion under the conservative self-force in black hole spacetimes an integrable Hamiltonian system?," *Phys. Rev. D* **92** (2015) 064039, `arXiv:1503.04727 [gr-qc]`.

[87] R. Fujita, S. Isoyama, A. Le Tiec, H. Nakano, N. Sago, and T. Tanaka, "Hamiltonian Formulation of the Conservative Self-Force Dynamics in the Kerr Geometry," *Class. Quant. Grav.* **34** no. 13, (2017) 134001, `arXiv:1612.02504 [gr-qc]`.

[88] A. Pound, B. Wardell, N. Warburton, and J. Miller, "Second-Order Self-Force Calculation of Gravitational Binding Energy in Compact Binaries," *Phys. Rev. Lett.* **124** no. 2, (Aug, 2020) , `arXiv:1908.07419`. `http://arxiv.org/abs/1908.07419`.

[89] N. Warburton, A. Pound, B. Wardell, J. Miller, and L. Durkan, "Gravitational-wave energy flux for compact binaries through second order in the mass ratio," `arXiv:2107.01298`. `http://arxiv.org/abs/2107.01298`.

[90] L. Durkan and N. Warburton, "Slow evolution of the metric perturbation due to a quasicircular inspiral into a Schwarzschild black hole," `arXiv:2206.08179 [gr-qc]`.

[91] V. Toomani, P. Zimmerman, A. Spiers, S. Hollands, A. Pound, and S. R. Green, "New metric reconstruction scheme for gravitational self-force calculations," *Class. Quant. Grav.* **39** no. 1, (2022) 015019, `arXiv:2108.04273 [gr-qc]`.

[92] S. D. Upton and A. Pound, "Second-order gravitational self-force in a highly regular gauge," *Phys. Rev. D* **103** no. 12, (Jan, 2021) , `arXiv:2101.11409`. `https://arxiv.org/abs/2101.11409v2`.

[93] P. Diener, I. Vega, B. Wardell, and S. Detweiler, "Self-consistent orbital evolution of a particle around a schwarzschild black hole," *Phys. Rev. Lett.* **108** no. 19, (Dec, 2012) , `arXiv:1112.4821`. `https://arxiv.org/abs/1112.4821v3`.

[94] S. R. Dolan and L. Barack, "Self-force via m-mode regularization and 2+1D evolution. III. Gravitational field on Schwarzschild spacetime," *Phys. Rev. D - Part. Fields, Gravit. Cosmol.* **87** no. 8, (Nov, 2013) , `arXiv:1211.4586`. `https://arxiv.org/abs/1211.4586v1`.

[95] L. Barack and N. Sago, "Gravitational self-force on a particle in circular orbit around a Schwarzschild black hole," *Phys. Rev. D - Part. Fields, Gravit. Cosmol.* **75** no. 6, (Jan, 2007) , `arXiv:0701069 [gr-qc]`. `http://arxiv.org/abs/gr-qc/0701069http://dx.doi.org/10.1103/PhysRevD.75.064021`.

[96] L. Barack and N. Sago, "Gravitational self-force on a particle in eccentric orbit around a Schwarzschild black hole," *Phys. Rev. D - Part. Fields, Gravit. Cosmol.* **81** no. 8, (Feb, 2010) , `arXiv:1002.2386`. `https://arxiv.org/abs/1002.2386v3`.

[97] S. Akcay, N. Warburton, and L. Barack, "Frequency-domain algorithm for the Lorenz-gauge gravitational self-force," *Phys. Rev. D - Part. Fields, Gravit. Cosmol.* **88** no. 10, (Aug, 2013) , `arXiv:1308.5223`. `https://arxiv.org/abs/1308.5223v2`.

[98] T. Osburn, E. Forseth, C. R. Evans, and S. Hopper, "Lorenz gauge gravitational self-force calculations of eccentric binaries using a frequency domain procedure," *Phys. Rev. D - Part. Fields, Gravit. Cosmol.* **90** no. 10, (Sep, 2014) , `arXiv:1409.4419`. `http://arxiv.org/abs/1409.4419http://dx.doi.org/10.1103/PhysRevD.90.104031`.

[99] M. Van De Meent, "Gravitational self-force on eccentric equatorial orbits around a Kerr black hole," *Phys. Rev. D* **94** no. 4, (Jun, 2016) , `arXiv:1606.06297`. `http://arxiv.org/abs/1606.06297http://dx.doi.org/10.1103/PhysRevD.94.044034`.

[100] M. Van De Meent, "Gravitational self-force on generic bound geodesics in Kerr spacetime," *Phys. Rev. D* **97** no. 10, (May, 2018) , `arXiv:1711.09607`.

[101] L. Barack and A. Ori, "Gravitational selfforce and gauge transformations," *Phys. Rev. D* **64** (2001) 124003, `arXiv:gr-qc/0107056`.

[102] S. Akcay, "A Fast Frequency-Domain Algorithm for Gravitational Self-Force: I. Circular Orbits in Schwarzschild Spacetime," *Phys. Rev. D* **83** (2011) 124026, `arXiv:1012.5860 [gr-qc]`.

[103] S. Isoyama, L. Barack, S. R. Dolan, A. Le Tiec, H. Nakano, A. G. Shah, T. Tanaka, and N. Warburton, "Gravitational Self-Force Correction to the Innermost Stable Circular Equatorial Orbit of a Kerr Black Hole," *Phys. Rev. Lett.* **113** no. 16, (2014) 161101, `arXiv:1404.6133 [gr-qc]`.

[104] S. A. Teukolsky, "Perturbations of a rotating black hole. 1. Fundamental equations for gravitational electromagnetic and neutrino field perturbations," *Astrophys. J.* **185** (1973) 635–647.

[105] P. L. Chrzanowski, "Vector Potential and Metric Perturbations of a Rotating Black Hole," *Phys. Rev. D* **11** (1975) 2042–2062.

[106] L. S. Kegeles and J. M. Cohen, "CONSTRUCTIVE PROCEDURE FOR PERTURBATIONS OF SPACE-TIMES," *Phys. Rev. D* **19** (1979) 1641–1664.

[107] R. M. Wald, "Construction of Solutions of Gravitational, Electromagnetic, Or Other Perturbation Equations from Solutions of Decoupled Equations," *Phys. Rev. Lett.* **41** (1978) 203–206.

[108] S. R. Dolan, C. Kavanagh, and B. Wardell, "Gravitational perturbations of rotating black holes in Lorenz gauge," `arXiv:2108.06344 [gr-qc]`.

[109] A. Pound, C. Merlin, and L. Barack, "Gravitational self-force from radiation-gauge metric perturbations," *Phys. Rev. D* **89** no. 2, (2014) 024009, `arXiv:1310.1513 [gr-qc]`.

[110] C. Merlin and A. G. Shah, "Self-force from reconstructed metric perturbations: numerical implementation in Schwarzschild spacetime," *Phys. Rev. D* **91** no. 2, (2015) 024005, `arXiv:1410.2998 [gr-qc]`.

[111] M. van de Meent and A. G. Shah, "Metric perturbations produced by eccentric equatorial orbits around a Kerr black hole," *Phys. Rev. D* **92** no. 6, (2015) 064025, `arXiv:1506.04755 [gr-qc]`.

[112] J. Miller and A. Pound, "Two-timescale evolution of extreme-mass-ratio inspirals: Waveform generation scheme for quasicircular orbits in Schwarzschild spacetime," *Phys. Rev. D* **103** no. 6, (Jun, 2021) , `arXiv:2006.11263`. `http://arxiv.org/abs/2006.11263`.

[113] A. Pound and B. Wardell, "Black hole perturbation theory and gravitational self-force," `arXiv:2101.04592`. `http://arxiv.org/abs/2101.04592`.

[114] L. Barack and C. Cutler, "LISA capture sources: Approximate waveforms, signal-to-noise ratios, and parameter estimation accuracy," *Phys. Rev. D - Part. Fields, Gravit. Cosmol.* **69** no. 8, (Oct, 2004) 24, `arXiv:0310125 [gr-qc]`. `https://arxiv.org/abs/gr-qc/0310125v3`.

[115] D. Kennefick, "Approximating the inspiral of test bodies into Kerr black holes," *Phys. Rev. D - Part. Fields, Gravit. Cosmol.* **66** no. 6, (May, 2002) .

[116] S. Babak, H. Fang, J. R. Gair, K. Glampedakis, and S. A. Hughes, ""kludge" gravitational waveforms for a test-body orbiting a Kerr black hole," *Phys. Rev. D - Part. Fields, Gravit. Cosmol.* **75** no. 2, (2007) , `arXiv:0607007 [gr-qc]`.

[117] C. F. Sopuerta and N. Yunes, "New Kludge scheme for the construction of approximate waveforms for extreme-mass-ratio inspirals," *Phys. Rev. D - Part. Fields, Gravit. Cosmol.* **84** no. 12, (Sep, 2011) , `arXiv:1109.0572`. `https://arxiv.org/abs/1109.0572v2`.

[118] A. J. Chua and J. R. Gair, "Improved analytic extreme-mass-ratio inspiral model for scoping out eLISA data analysis," *Class. Quantum Gravity* **32** no. 23, (Oct, 2015) , `arXiv:1510.06245`. `https://arxiv.org/abs/1510.06245v1`.

[119] P. C. Peters and J. Mathews, "Gravitational radiation from point masses in a Keplerian orbit," *Phys. Rev.* **131** (1963) 435–439.

[120] M. L. Katz, A. J. K. Chua, L. Speri, N. Warburton, and S. A. Hughes, "FastEMRIWaveforms: New tools for millihertz gravitational-wave data analysis," `arXiv:2104.04582`. `http://arxiv.org/abs/2104.04582`.

[121] A. J. Chua, C. J. Moore, and J. R. Gair, "Augmented kludge waveforms for detecting extreme-mass-ratio inspirals," *Phys. Rev. D* **96** no. 4, (Aug, 2017) .

[122] K. A. Arnaud, S. Babak, J. G. Baker, M. J. Benacquista, N. J. Cornish, C. Cutler, S. L. Larson, B. S. Sathyaprakash, M. Vallisneri, A. Vecchio, and J. Y. Vinet, "An overview of the mock LISA data challenges," in *AIP Conf. Proc.*, vol. 873, pp. 619–624. Sep, 2006. `arXiv:0609105 [gr-qc]`. `https://arxiv.org/abs/gr-qc/0609105v1`.

[123] S. Babak, J. G. Baker, M. J. Benacquista, N. J. Cornish, S. L. Larson, I. Mandel, S. T. McWilliams, A. Petiteau, E. K. Porter, E. L. Robinson, M. Vallisneri, A. Vecchio, M. Adams, K. A. Arnaud, A. Błaut, M. Bridges, M. Cohen, C. Cutler, F. Feroz, J. R. Gair, P. Graff, M. Hobson, J. S. Key, A. Królak, A. Lasenby, R. Prix, Y. Shang, M. Trias, J. Veitch, and J. T. Whelan, "The Mock LISA Data Challenges: From challenge 3 to challenge 4," *Class. Quantum Gravity* **27** no. 8, (Dec, 2010) , `arXiv:0912.0548`. `https://arxiv.org/abs/0912.0548v2`.

[124] A. J. K. Chua and C. J. Cutler, "Non-local parameter degeneracy in the intrinsic space of gravitational-wave signals from extreme-mass-ratio inspirals," `arXiv:2109.14254 [gr-qc]`.

[125] S. Drasco, E. E. Flanagan, and S. A. Hughes, "Computing inspirals in Kerr in the adiabatic regime. I. The Scalar case," *Class. Quant. Grav.* **22** (2005) S801–846, `arXiv:gr-qc/0505075`.

[126] S. A. Hughes, S. Drasco, E. E. Flanagan, and J. Franklin, "Gravitational radiation reaction and inspiral waveforms in the adiabatic limit," *Phys. Rev. Lett.* **94** no. 22, (2005) .

[127] N. Sago, T. Tanaka, W. Hikida, K. Ganz, and H. Nakano, "The Adiabatic evolution of orbital parameters in the Kerr spacetime," *Prog. Theor. Phys.* **115** (2006) 873–907, `arXiv:gr-qc/0511151`.

[128] Y. Mino, "Perturbative approach to an orbital evolution around a supermassive black hole," *Phys. Rev. D - Part. Fields, Gravit. Cosmol.* **67** no. 8, (Feb, 2003) , `arXiv:0302075v1 [gr-qc]`. `http://arxiv.org/abs/gr-qc/0302075http://dx.doi.org/10.1103/PhysRevD.67.084027`.

[129] S. Isoyama, R. Fujita, H. Nakano, N. Sago, and T. Tanaka, ""Flux-balance formulae" for extreme mass-ratio inspirals," *PTEP* **2019** no. 1, (2019) 013E01, `arXiv:1809.11118 [gr-qc]`.

[130] T. Regge and J. A. Wheeler, "Stability of a Schwarzschild singularity," *Phys. Rev.* **108** (1957) 1063–1069.

[131] E. Poisson, "Gravitational radiation from a particle in circular orbit around a black hole. 1: Analytical results for the nonrotating case," *Phys. Rev. D* **47** (1993) 1497–1510.

[132] S. L. Detweiler, "Black holes and gravitational waves. I. Circular orbits about a rotating hole.," *Astrophysical Journal* **225** (Nov., 1978) 687–693.

[133] L. S. Finn and K. S. Thorne, "Gravitational waves from a compact star in a circular, inspiral orbit, in the equatorial plane of a massive, spinning black hole, as observed by LISA," *Phys. Rev. D* **62** (2000) 124021, `arXiv:gr-qc/0007074`.

[134] S. A. Hughes, N. Warburton, G. Khanna, A. J. Chua, and M. L. Katz, "Adiabatic waveforms for extreme mass-ratio inspirals via multivoice decomposition in time and frequency," *Phys. Rev. D* **103** no. 10, (Feb, 2021) , `arXiv:2102.02713`. `https://arxiv.org/abs/2102.02713v3`.

[135] A. J. Chua, M. L. Katz, N. Warburton, and S. A. Hughes, "Rapid Generation of Fully Relativistic Extreme-Mass-Ratio-Inspiral Waveform Templates for LISA Data Analysis," *Phys. Rev. Lett.* **126** no. 5, (Aug, 2021) , `arXiv:2008.06071`. `https://arxiv.org/abs/2008.06071v1`.

[136] A. Pound and E. Poisson, "Osculating orbits in Schwarzschild spacetime, with an application to extreme mass-ratio inspirals," *Phys. Rev. D - Part. Fields, Gravit. Cosmol.* **77** no. 4, (Feb, 2008) , `arXiv:0708.3033`.

[137] J. R. Gair, E. E. Flanagan, S. Drasco, T. Hinderer, and S. Babak, "Forced motion near black holes," *Phys. Rev. D - Part. Fields, Gravit. Cosmol.* **83** no. 4, (Feb, 2011) , `arXiv:1012.5111`.

[138] K. S. Thorne, "Transition from inspiral to plunge for a compact body in a circular equatorial orbit around a massive, spinning black hole," *Phys. Rev. D - Part. Fields, Gravit. Cosmol.* **62** no. 12, (Mar, 2000) 8, `arXiv:0003032 [gr-qc]`. `https://arxiv.org/abs/gr-qc/0003032v1`.

[139] O. Burke, J. Gair, and J. Simón, "Transition from inspiral to plunge: A complete near-extremal trajectory and associated waveform," *Phys. Rev. D* **101** no. 6, (Sep, 2020) , `arXiv:1909.12846`. `http://arxiv.org/abs/1909.12846http://dx.doi.org/10.1103/PhysRevD.101.064026`.

[140] G. Compère and L. Küchler, "Self-Consistent Adiabatic Inspiral and Transition Motion," *Phys. Rev. Lett.* **126** no. 24, (Feb, 2021) , `arXiv:2102.12747`. `http://arxiv.org/abs/2102.12747`.

[141] G. Compère and L. Küchler, "Asymptotically matched quasi-circular inspiral and transition-to-plunge in the small mass ratio expansion," `arXiv:2112.02114 [gr-qc]`.

[142] P. A. Sundararajan, "Transition from adiabatic inspiral to geodesic plunge for a compact object around a massive Kerr black hole: Generic orbits," *Phys. Rev. D - Part. Fields, Gravit. Cosmol.* **77** no. 12, (Mar, 2008) , `arXiv:0803.4482v3`. `http://arxiv.org/abs/0803.4482http://dx.doi.org/10.1103/PhysRevD.77.124050`.

[143] P. Gupta, B. Bonga, A. J. K. Chua, and T. Tanaka, "Importance of tidal resonances in extreme-mass-ratio inspirals," *Phys. Rev. D* **104** no. 4, (2021) 044056, `arXiv:2104.03422 [gr-qc]`.

[144] N. Warburton, T. Osburn, and C. R. Evans, "Evolution of small-mass-ratio binaries with a spinning secondary," *Phys. Rev. D* **96** no. 8, (Aug, 2017) , `arXiv:1708.03720`. `http://arxiv.org/abs/1708.03720http://dx.doi.org/10.1103/PhysRevD.96.084057`.

[145] N. Warburton, S. Akcay, L. Barack, J. R. Gair, and N. Sago, "Evolution of inspiral orbits around a Schwarzschild black hole," *Phys. Rev. D - Part. Fields, Gravit. Cosmol.* **85** no. 6, (Mar, 2012) , `arXiv:1111.6908`.

[146] S. Mano, H. Suzuki, and E. Takasugi, "Analytic solutions of the Teukolsky equation and their low frequency expansions," *Prog. Theor. Phys.* **95** (1996) 1079–1096, `arXiv:gr-qc/9603020`.

[147] S. Mano and E. Takasugi, "Analytic solutions of the Teukolsky equation and their properties," *Prog. Theor. Phys.* **97** (1997) 213–232, `arXiv:gr-qc/9611014`.

[148] R. Fujita and H. Tagoshi, "New numerical methods to evaluate homogeneous solutions of the Teukolsky equation," *Prog. Theor. Phys.* **112** (2004) 415–450, `arXiv:gr-qc/0410018`.

[149] R. Fujita, W. Hikida, and H. Tagoshi, "An Efficient Numerical Method for Computing Gravitational Waves Induced by a Particle Moving on Eccentric Inclined Orbits around a Kerr Black Hole," *Prog. Theor. Phys.* **121** (2009) 843–874, `arXiv:0904.3810 [gr-qc]`.

[150] C. Merlin, A. Ori, L. Barack, A. Pound, and M. van de Meent, "Completion of metric reconstruction for a particle orbiting a Kerr black hole," *Phys. Rev. D* **94** no. 10, (2016) 104066, `arXiv:1609.01227 [gr-qc]`.

[151] M. van De Meent, "The mass and angular momentum of reconstructed metric perturbations," *Class. Quant. Grav.* **34** no. 12, (2017) 124003, `arXiv:1702.00969 [gr-qc]`.

[152] M. van de Meent and A. Pound, "Gauge completion in radiation gauge metric reconstruction." (unpublished).

[153] L. Barack and A. Ori, "Mode sum regularization approach for the selfforce in black hole space-time," *Phys. Rev. D* **61** (2000) 061502, `arXiv:gr-qc/9912010`.

[154] M. Van De Meent and N. Warburton, "Fast self-forced inspirals," *Class. Quantum Gravity* **35** no. 14, (Feb, 2018) , `arXiv:1802.05281`. `http://arxiv.org/abs/1802.05281http://dx.doi.org/10.1088/1361-6382/aac8ce`.

[155] J. Kevorkian, "Perturbation Techniques for Oscillatory Systems With Slowly Varying Coefficients.," *SIAM Rev.* **29** no. 3, (Mar, 1987) 391–461.

[156] J. McCart, T. Osburn, and J. Y. J. Burton, "Highly eccentric EMRI waveforms via fast self-forced inspirals," `arXiv:2109.00056`. `http://arxiv.org/abs/2109.00056`.

[157] M. Van De Meent, "Conditions for sustained orbital resonances in extreme mass ratio inspirals," *Phys. Rev. D - Part. Fields, Gravit. Cosmol.* **89** no. 8, (Nov, 2014) , `arXiv:1311.4457`. `http://arxiv.org/abs/1311.4457http://dx.doi.org/10.1103/PhysRevD.89.084033`.

[158] U. Ruangsri and S. A. Hughes, "Census of transient orbital resonances encountered during binary inspiral," *Phys. Rev. D* **89** no. 8, (2014) 084036, `arXiv:1307.6483 [gr-qc]`.

[159] E. E. Flanagan, S. A. Hughes, and U. Ruangsri, "Resonantly enhanced and diminished strong-field gravitational-wave fluxes," *Phys. Rev. D* **89** no. 8, (2014) 084028, `arXiv:1208.3906 [gr-qc]`.

[160] C. P. Berry, R. H. Cole, P. Cañizares, and J. R. Gair, "Importance of transient resonances in extreme-mass-ratio inspirals," *Phys. Rev. D* **94** no. 12, (Dec, 2016) .

[161] L. Speri and J. R. Gair, "Assessing the impact of transient orbital resonances," *Phys. Rev. D* **103** no. 12, (Mar, 2021) , `arXiv:2103.06306`. `https://arxiv.org/abs/2103.06306v2`.

[162] B. Bonga, H. Yang, and S. A. Hughes, "Tidal resonance in extreme mass-ratio inspirals," *Phys. Rev. Lett.* **123** no. 10, (2019) 101103, `arXiv:1905.00030 [gr-qc]`.

[163] P. Gupta, L. Speri, B. Bonga, A. J. K. Chua, and T. Tanaka, "Modeling transient resonances in extreme-mass-ratio inspirals," `arXiv:2205.04808 [gr-qc]`.

[164] B. Carter, "Global structure of the kerr family of gravitational fields," *Phys. Rev.* **174** (Oct, 1968) 1559–1571. `https://link.aps.org/doi/10.1103/PhysRev.174.1559`.

[165] S. Drasco and S. A. Hughes, "Rotating black hole orbit functionals in the frequency domain," *Phys. Rev. D - Part. Fields, Gravit. Cosmol.* **69** no. 4, (Feb, 2004) , `arXiv:0308479 [astro-ph]`.

[166] F. D. Ryan, "Effect of gravitational radiation reaction on circular orbits around a spinning black hole," *Phys. Rev. D* **52** (1995) R3159–R3162, `arXiv:gr-qc/9506023`.

[167] S. Drasco and S. A. Hughes, "Erratum: Gravitational wave snapshots of generic extreme mass ratio inspirals (Physical Review D - Particles, Fields, Gravitation and Cosmology (2006) 73 (024027))," Nov, 2014.

[168] W. Schmidt, "Celestial mechanics in Kerr spacetime," Tech. Rep. 10, 2002. `arXiv:0202090 [gr-qc]`.

[169] L. C. Stein and N. Warburton, "Location of the last stable orbit in Kerr spacetime," *Phys. Rev. D* **101** no. 6, (Dec, 2020) , `arXiv:1912.07609`. `http://arxiv.org/abs/1912.07609`.

[170] O. Long and L. Barack, "Time-domain metric reconstruction for hyperbolic scattering," *Phys. Rev. D* **104** no. 2, (2021) 024014, `arXiv:2105.05630 [gr-qc]`.

[171] L. Barack and O. Long, "Self-force correction to the deflection angle in black-hole scattering: a scalar charge toy model," `arXiv:2209.03740 [gr-qc]`.

[172] S. Drasco and S. A. Hughes, "Rotating black hole orbit functionals in the frequency domain," *Phys. Rev. D* **69** (2004) 044015, `arXiv:astro-ph/0308479`.

[173] Z. Nasipak, T. Osburn, and C. R. Evans, "Repeated faint quasinormal bursts in extreme-mass-ratio inspiral waveforms: Evidence from frequency-domain scalar self-force calculations on generic Kerr orbits," *Phys. Rev. D* **100** no. 6, (2019) 064008, `arXiv:1905.13237 [gr-qc]`.

[174] R. Fujita and W. Hikida, "Analytical solutions of bound timelike geodesic orbits in Kerr spacetime," *Class. Quantum Gravity* **26** no. 13, (2009) , `arXiv:0906.1420`.

[175] M. Van De Meent, "Analytic solutions for parallel transport along generic bound geodesics in Kerr spacetime," *Class. Quantum Gravity* **37** no. 14, (Jul, 2020) , `arXiv:1906.05090`.

[176] J. M. Bardeen, W. H. Press, and S. A. Teukolsky, "Rotating black holes: Locally nonrotating frames, energy extraction, and scalar synchrotron radiation." *Astrophys. J. 178: No. 2, 347-69(1 Dec 1972).* . `https://www.osti.gov/biblio/4585183`.

[177] J. Levin and G. Perez-Giz, "Homoclinic Orbits around Spinning Black Holes. I. Exact Solution for the Kerr Separatrix," *Phys. Rev. D* **79** (2009) 124013, `arXiv:0811.3814 [gr-qc]`.

[178] J. Levin and G. Perez-Giz, "A Periodic Table for Black Hole Orbits," *Phys. Rev. D* **77** (2008) 103005, `arXiv:0802.0459 [gr-qc]`.

[179] J. Brink, M. Geyer, and T. Hinderer, "Astrophysics of resonant orbits in the Kerr metric," *Phys. Rev. D* **91** no. 8, (2015) 083001, `arXiv:1501.07728 [gr-qc]`.

[180] D. Kennefick and A. Ori, "Radiation reaction induced evolution of circular orbits of particles around Kerr black holes," *Phys. Rev. D* **53** (1996) 4319–4326, `arXiv:gr-qc/9512018`.

[181] C. M. Will, "Compact binary inspiral: Nature is perfectly happy with a circle," *Class. Quant. Grav.* **36** no. 19, (2019) 195013, `arXiv:1906.08064 [gr-qc]`.

[182] "Black Hole Perturbation Toolkit." (bhptoolkit.org).

[183] V. Witzany, "Action-angle coordinates for black-hole geodesics I: Spherically symmetric and Schwarzschild," `arXiv:2203.11952 [gr-qc]`.

[184] W. H. Press, "Gravitational Radiation from Sources Which Extend Into their Own Wave Zone," *Phys. Rev. D* **15** (1977) 965–968.

[185] T. Robson, N. J. Cornish, and C. Liu, "The construction and use of LISA sensitivity curves," *Class. Quant. Grav.* **36** no. 10, (2019) 105011, `arXiv:1803.01944 [astro-ph.HE]`.

[186] "SimulationTools." (simulationtools.org).

[187] B. J. Owen, "Search templates for gravitational waves from inspiraling binaries: Choice of template spacing," *Phys. Rev. D - Part. Fields, Gravit. Cosmol.* **53** no. 12, (Nov, 1996) 6749–6761, `arXiv:9511032 [gr-qc]`. `https://arxiv.org/abs/gr-qc/9511032v1`.

[188] L. Lindblom, B. J. Owen, and D. A. Brown, "Model waveform accuracy standards for gravitational wave data analysis," *Phys. Rev. D - Part. Fields, Gravit. Cosmol.* **78** no. 12, (Sep, 2008) , `arXiv:0809.3844`. `http://arxiv.org/abs/0809.3844http://dx.doi.org/10.1103/PhysRevD.78.124020`.

[189] S. T. McWilliams, B. J. Kelly, and J. G. Baker, "Observing mergers of nonspinning black-hole binaries," *Phys. Rev. D - Part. Fields, Gravit. Cosmol.* **82** no. 2, (Apr, 2010) , `arXiv:1004.0961`. `https://arxiv.org/abs/1004.0961v1`.

[190] N. Sago, L. Barack, and S. L. Detweiler, "Two approaches for the gravitational self force in black hole spacetime: Comparison of numerical results," *Phys. Rev. D* **78** (2008) 124024, `arXiv:0810.2530 [gr-qc]`.

[191] S. Akcay, A. Le Tiec, L. Barack, N. Sago, and N. Warburton, "Comparison between self-force and post-Newtonian dynamics: Beyond circular orbits," *Phys. Rev. D - Part. Fields, Gravit. Cosmol.* **91** no. 12, (Mar, 2015) , `arXiv:1503.01374v2`. `http://arxiv.org/abs/1503.01374http://dx.doi.org/10.1103/PhysRevD.91.124014`.

[192] G. Lukes-Gerakopoulos and V. Witzany, "Nonlinear Effects in EMRI Dynamics and Their Imprints on Gravitational Waves," *Handb. Gravitational Wave Astron.* (Mar, 2021) 1–44, `arXiv:2103.06724`. `http://arxiv.org/abs/2103.06724`.

[193] A. Le Tiec, A. H. Mroue, L. Barack, A. Buonanno, H. P. Pfeiffer, N. Sago, and A. Taracchini, "Periastron Advance in Black Hole Binaries," *Phys. Rev. Lett.* **107** (2011) 141101, `arXiv:1106.3278 [gr-qc]`.

[194] V. Skoupý and G. Lukes-Gerakopoulos, "Spinning test body orbiting around a Kerr black hole: Eccentric equatorial orbits and their asymptotic gravitational-wave fluxes," *Phys. Rev. D* **103** no. 10, (2021) 104045, `arXiv:2102.04819 [gr-qc]`.

[195] G. A. Piovano, A. Maselli, and P. Pani, "Extreme mass ratio inspirals with spinning secondary: a detailed study of equatorial circular motion," *Phys. Rev. D* **102** no. 2, (2020) 024041, `arXiv:2004.02654 [gr-qc]`.

[196] G. A. Piovano, R. Brito, A. Maselli, and P. Pani, "Assessing the detectability of the secondary spin in extreme mass-ratio inspirals with fully relativistic numerical waveforms," *Phys. Rev. D* **104** no. 12, (2021) 124019, `arXiv:2105.07083 [gr-qc]`.

[197] J. Mathews, A. Pound, and B. Wardell, "Self-force calculations with a spinning secondary," *Phys. Rev. D* **105** no. 8, (2022) 084031, `arXiv:2112.13069 [gr-qc]`.

[198] L. V. Drummond and S. A. Hughes, "Precisely computing bound orbits of spinning bodies around black holes. I. General framework and results for nearly equatorial orbits," *Phys. Rev. D* **105** no. 12, (2022) 124040, `arXiv:2201.13334 [gr-qc]`.

[199] L. V. Drummond and S. A. Hughes, "Precisely computing bound orbits of spinning bodies around black holes. II. Generic orbits," *Phys. Rev. D* **105** no. 12, (2022) 124041, `arXiv:2201.13335 [gr-qc]`.

[200] L. Polcar, G. Lukes-Gerakopoulos, and V. Witzany, "Extreme mass ratio inspirals into black holes surrounded by matter," `arXiv:2205.08516 [gr-qc]`.

[201] L. Speri, A. Antonelli, L. Sberna, S. Babak, E. Barausse, J. R. Gair, and M. L. Katz, "Measuring accretion-disk effects with gravitational waves from extreme mass ratio inspirals," `arXiv:2207.10086 [gr-qc]`.

[202] M. L. Katz, S. Marsat, A. J. K. Chua, S. Babak, and S. L. Larson, "GPU-accelerated massive black hole binary parameter estimation with LISA," *Phys. Rev. D* **102** no. 2, (2020) 023033, `arXiv:2005.01827 [gr-qc]`.

[203] A. Le Tiec, A. Buonanno, A. H. Mroué, H. P. Pfeiffer, D. A. Hemberger, G. Lovelace, L. E. Kidder, M. A. Scheel, B. Szilágyi, N. W. Taylor, and S. A. Teukolsky, "Periastron Advance in Spinning Black Hole Binaries: Gravitational Self-Force from Numerical Relativity," **88** no. 12, (2013) 124027, `arXiv:1309.0541 [gr-qc]`.