Using perturbation theory to understand the two body problem in general relativity

by

Pranesh Adhyam Sundararajan

M. Sc. (Hons) Physics, B. E. (Hons) Electrical & Electronics
Engineering
Birla Institute of Technology & Science, Pilani, India (2005)

Submitted to the Department of Physics
in partial fulfillment of the requirements for the degree of

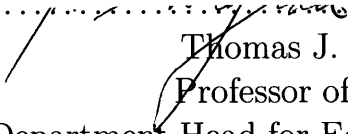Doctor of Philosophy in Physics

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

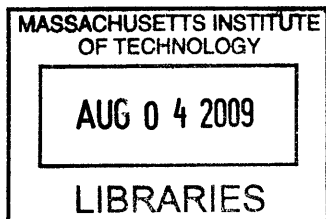September 2009

Author . . . .
Department of Physics
June 30, 2009

Certified by. . . . .
Scott A. Hughes
Associate Professor of Physics
Thesis Supervisor

Accepted by . . . .
Thomas J. Greytak
Professor of Physics
Associate Department Head for Education

Using perturbation theory to understand the two body problem in general relativity

by

Pranesh Adhyam Sundararajan

Abstract

Binary systems composed of compact objects (neutron stars and black holes) radiate gravitational waves (GWs). The prospect of detecting these GWs using ground and space based experiments has made it imperative to understand the dynamics of such compact binaries. This work describes several advances in our ability to model compact binaries and extract the rich science they encode.

A major part of this dissertation focuses on the subset of binaries composed of a massive, central black hole ($10^5 - 10^8 M_\odot$) and a much smaller compact object ($1 - 100 M_\odot$). The emission of gravitational energy from such extreme mass ratio inspirals (EMRIs) forces the separation between the two components to shrink, leading to their merger. We treat the smaller object as a point-like particle on the stationary space-time of the larger black hole. The EMRI problem can be broken down into two related parts: *(i)* A determination of the *inspiral trajectory* followed by the smaller object, and *(ii)* A characterization of the *gravitational waveforms* that result from such an inspiral.

The initial part of this work discusses the development of a numerical algorithm that solves for the GWs that result from the perturbations generated by the smaller object. It accepts *any* reasonable inspiral trajectory as an input and produces the resulting waveforms with an accuracy greater than 99%.

Next, we present a technique to model the part of the inspiral trajectory that immediately precedes the final plunge of smaller object into the massive black hole. Along with earlier research, this enables us to compute the smaller object's complete inspiral trajectory. We now have a versatile toolkit that can model GWs from EMRIs.

Finally, we present another application of this work. GWs carry linear momentum away from a binary. Integrating the lost momentum leaves an asymmetric binary with a non-zero recoil velocity after merger. We compute the recoils from EMRIs and extrapolate them to comparable mass binaries. We find that extrapolating perturbation theory gives results that agree well with those from numerical relativity, but require far less computation time.

Thesis Supervisor: Scott A. Hughes
Title: Associate Professor of Physics

Acknowledgments

Contents

9

List of Figures

15

List of Tables

18

Chapter 1

Introduction and overview

This chapter provides motivation for the work in subsequent chapters by placing it in the context of current research in the field.

The dynamics of two astrophysical objects under their mutual gravitational attractive force (the "two-body problem") has been a subject of research for several centuries. Pillars of modern science like Newton, Kepler, Galileo, and Einstein have devoted significant portions of their work in trying to understand gravitational two body interactions. A detailed study of these interactions has improved our understanding of phenomena ranging from the Earth-Moon system to binaries composed of supermassive black holes (BHs). The richness of these two-body interactions make them interesting to scientists even today.

1.1 Gravitational waves, the two-body problem, and general relativity

The two body problem in Newtonian gravity has a particularly clean analytic solution [49]. A wide variety of astrophysical binaries can be described by Newtonian gravity. However, general relativistic effects begin to enter when the objects are large, dense and separated by short distances. Close binaries composed of compact objects (neutron stars or black holes) exhibit relativistic effects, making them ideal candidates to

explore deviations from Newtonian gravity. Relativistic effects significantly alter the dynamics of compact binaries. The situation becomes even more complicated when one (or both) of the members of the binary is a black hole. We need to account for a host of additional complications such as spacetimes, horizons, coordinate systems, and observers. The equations that describe the two body problem become rich and complex, and necessitate a combination of numerical and analytical techniques.

Contrary to Newtonian binaries, relativistic binaries radiate energy and angular momentum in the form of gravitational waves (GWs) [26]. GWs travel at the speed of light in vacuum and transport energy from the source. Loss of energy and angular momentum through GWs forces the orbit to shrink and the binary to ultimately coalesce. The Hulse-Taylor pulsar, PSR 1913+16 [55] is the most famous example of such a relativistic binary. Over 30 years of observational data show that the evolution of the binary's orbit is in excellent accord with general relativistic predictions [105].

The significance of gravitational waves to general relativity is vital. In most cases, relativistic effects are corrections to the dominant Newtonian effect. For example, the relativistic precession of the perihelion of Mercury is only a few arc seconds every century — a tiny correction to its Keplerian motion. On the other hand GWs are a testable prediction of general relativity that do not appear in Newtonian gravity. Moreover, GWs encode information about the spacetime of the source. A detailed map of the spacetime in the vicinity of a massive and dense astrophysical system may be the most stringent test of general relativity ever.

Formally, we need a description of the curved spacetime that embeds the binary. One needs to solve the non-linear Einstein field equations for the metric. For black hole binaries, each black hole's horizon presents an additional complication. In addition to orbital motion, members of the binary may spin around their own axes. The equations that describe these complex effects are highly coupled and non-linear. Unlike the Newtonian case, a generic solution to the two body problem that includes all effects does not exist. However, there are several successful approaches that are tailored to address the problem in certain interesting limits:

- Post Newtonian (PN) theory [17]: The PN expansion is obtained by iterating

Einstein's field equations from the Newtonian limit to higher order in GM/rc^2. Ref. [17] is an excellent review of PN methods applicable to the binary problem. The symbol M represents the characteristic mass of the system and r represents the characteristic separation between its members. It is designed to describe the binary's dynamics and GWs that it emits. The advantage of this approach is that it provides closed-form analytical solutions that are applicable to a wide variety of binaries. The disadvantage is that they are an accurate description only when the separation is greater than a few Schwarzchild radii, $2GM/c^2$. Nevertheless, PN theory has greatly improved our understanding of several potential sources of GWs. It is widely used in data-analysis routines in gravitational wave detectors.

- Numerical relativity [85]: It has now become possible to apply numerical methods to solve Einstein's equations on large computer clusters. Great progress in numerical relativity has been achieved over the last few years. The non-linear field equations can now be cast into a form amenable to numerical methods. The main advantage here is that no approximation scheme is used; only errors due to the numerical scheme arise. Numerical relativity has furnished us with templates for GWs from the final stages of black hole mergers. It works very well when the masses of the binary's constituents are comparable. The major disadvantage here is that it is very computationally intensive. This makes long evolutions (for large initial separations) cumbersome. Moreover, a different simulation needs to be performed for each set of parameters (eg. mass, spin, initial separation) that characterizes the binary.

- Black hole perturbation theory [98, 53, 36, 97]: BH perturbation theory has been successfully used to describe binaries involving a massive central black hole (of mass M) and a much less massive companion (of mass μ). We treat the smaller companion as a perturbation on the stationary spacetime of the larger BH. The advantages here are: (a) the equations from perturbation theory are much simpler than the full Einstein field equations, and (b) It is valid even

when the separation between the binary's constituents is small. The major disadvantage is that it does not include terms of $\mathcal{O}(\mu/M)^2$ and higher. Black hole perturbation theory is the subject of most of this dissertation. A more detailed description is given in Sec. 1.5.

It is important to point out that we get consistent results from all approaches to general relativity. Results from numerical relativity in weak fields match those from post-Newtonian theory. Perturbation theory at high mass ratios is in concordance with numerical relativity.

1.2 Gravitational waves, the two-body problem, and Astronomy

Observations of electromagnetic radiation from astrophysical sources have increased our understanding of the universe immensely. For example, studying the sun at various wavelengths allows us to infer the processes that take place within it. Observations of highly regular pulses from a point source in the sky led to the discovery of pulsars. Today, we observe the cosmic microwave background (CMB), which is relic radiation from an epoch approximately 400 million years after the big bang. These are just a few examples from an endless list.

Gravitational waves equip us with a different, complementary tool kit to observe the universe with. Just as electromagnetic waves, GWs carry information pertaining to the properties of their respective sources. In particular, GWs contain direct information about bulk properties (such as mass and spin) of relativistic systems that are very difficult to obtain otherwise. The weakness of the interaction between GWs and intervening matter (eg. dust) means that it does not get attenuated akin to electromagnetic waves. For the same reason, detecting gravitational waves requires extreme sensitivity — detectors are made of the same matter that GWs barely interact with.

There are several postulated sources of GWs [30]. We list a few of the most interesting and potentially observable ones:

- Stellar mass compact binaries: Stellar mass black holes and neutron stars ($1 - 100M_\odot$) are dense objects formed as a result of the death of massive main sequence stars. The orbits of binaries composed of black holes and neutron stars shrink as a result of GW emission. The members of the binary may even emit electromagnetic radiation (eg. they could be pulsars, X-ray binaries). During the last stages, we expect the frequency of these GWs to be $\sim 10 - 100$ Hz. Such systems are potential sources of GWs for ground-based detectors like LIGO, GEO and VIRGO [66, 104, 46].

- Short hard gamma ray bursts: A number of studies (eg. Ref. [45]) conclude that neutron star - neutron star/black hole mergers may be responsible for the emission of short (≤ 2 sec) bursts of hard gamma rays. A key feature of this mechanism is the tidal disruption of the neutron star. A coincident observation of the electromagnetic signal and GWs from such an event will allow us to probe the internal structure of the neutron star during its disruption.

- Core collapse supernovae: The death of massive stars ($M > 25M_\odot$) is accompanied by an explosion that expunges its outer envelopes of gas. The Crab nebula [106] is an example of the remnant of one such supernova. A small fraction of the energy released during the explosion is emitted in GWs. Ground-based detectors may see GWs during these events.

- Stochastic GW background: Much like the CMB, we expect a stochastic distribution of background GWs. Matter dynamics prior to galaxy formation is believed to produce a large portion of this GW background. Ground based detectors have already placed impressive upper limits on the energy density contained in the GW background [1]. The power spectrum of primordial GWs can help us explain events in the very early universe that shaped cosmological evolution.

- Extreme mass ratio inspirals (EMRIs): The centers of most galaxies are believed to host supermassive black holes ($M \sim 10^6 - 10^8 M_\odot$). The black hole associated

with Sgr A*, the bright radio source at the center of the Milky Way has a mass of about $4 \times 10^6 M_\odot$. The inspiral of stellar mass compact objects into these supermassive black holes lead to EMRIs. GWs from the last year of an EMRI should be detectable by the proposed space based detector, LISA [67] up to redshifts of $z \sim 2$. EMRIs will be discussed in more detail in chapters 2, 3, and 4.

- Supermassive black hole binaries: Galaxy mergers are thought to be common at high redshifts ($z \geq 3$). These mergers contribute to hierarchical structure growth and are responsible for the formation of supermassive black hole binaries. These binaries radiate GWs in LISA's sensitive bandwidth. The galaxy merger may also ignite quasar activity. This is a particularly interesting scenario because it allows for independent measurements of redshift and distance. The redshift can be inferred from the electromagnetic signature whereas the GWs furnish the distance to the source [65]. Such measurements can effectively constrain cosmological models.

- White dwarf binaries: Low mass main sequence stars ($M < 5 - 10 M_\odot$) eventually evolve to white dwarfs (WDs). GWs from galactic WD binaries should be visible to LISA. In fact, it is likely that there are so many of them that it creates a low frequency "confusion" noise.

Our ability to theoretically understand GW signatures from these sources is vital in order to fully extract the potential science from GWs. This has motivated many theorists over the last few decades to model GW sources from first principles. Of the several sources mentioned above, systems composed of compact binaries have been identified as the cleanest and simplest to model. This is largely because gravitational interaction between the two bodies is the only overwhelmingly dominant force. (Core collapse supernovae are examples of sources where internal processes like chemical reactions, convection, shocks etc. pose serious complications to theorists trying to model them.) We now have a vibrant community of researchers working on various approaches to the two-body problem mentioned in the previous section.

1.3 The simplest relativistic binary

This section reviews the lowest order relativistic effects that a simple binary composed of two point masses exhibits. We also sketch some important results pertaining to gravitational waves; detailed derivations are available in any general relativity textbook (eg. Ref. [26]). We set $G = c = 1$ everywhere.

The metric of a spacetime that is slightly perturbed from its flat Minkowski background is given by,

$$g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu} , \tag{1.1}$$

where $\eta_{\mu\nu} = (-1, 1, 1, 1)$ and $h \ll \eta$. Gauge freedom in general relativity loosely means that we can choose a coordinate system wisely. In this case, we choose the "Lorenz" gauge where,

$$\partial^\mu \bar{h}_{\mu\nu} = 0 , \tag{1.2}$$

$$\bar{h}_{\mu\nu} = h_{\mu\nu} - \frac{1}{2}\eta_{\mu\nu}h , \tag{1.3}$$

$h \equiv \eta^{\mu\nu}h_{\mu\nu}$ is the trace of $h_{\mu\nu}$. With this clever choice of coordinates, we compute the curvature and Ricci tensors upto $\mathcal{O}(h)$. Substituting this into Einstein's field equations give,

$$\Box \bar{h}_{\mu\nu}(\mathbf{x}, t) = -16\pi T_{\mu\nu}(\mathbf{x}, t) \tag{1.4}$$

where $\Box = \partial_i \partial^i$ and $T_{\mu\nu}$ is the stress energy tensor. For example, $T_{\mu\nu} = 0$ for vacuum; $T_{\mu\nu} = (\rho + p)u_\mu u_\nu + pg_{\mu\nu}$ for a perfect fluid, u^μ is the 4-velocity of a fluid element, ρ and p are the rest frame energy density and isotropic pressure respectively. Eq. (1.4) shows that the metric perturbations satisfy the *wave equation* in vacuum and hence are *gravitational waves*. This wave equation can be solved using a radiative Green's

function (eg. Ref. [57]) to yield,

$$\bar{h}_{\mu\nu}(\mathbf{x}, t) \quad = \quad 4 \int d^3 x' \frac{T_{\mu\nu}(\mathbf{x}', \mathbf{t} - |\mathbf{x} - \mathbf{x}'|)}{|\mathbf{x} - \mathbf{x}'|} \ . \tag{1.5}$$

Here, $\mathbf{x}'$ is the position of the source (regions where $T_{\mu\nu} \neq 0$) and $\mathbf{x}$ is the location of the field.

Eq. (1.5) suggests that every component of $h_{\mu\nu}$ is radiative. However, this is only an artifact of our gauge choice. It turns out that only the spatial, transverse, and traceless parts of $h_{\mu\nu}$ are gauge invariant and radiative degrees of freedom [40]. To obtain them, we "project out" the transverse and traceless part ($h_{\mu\nu}^{\mathrm{TT}}$) using the projection tensor P_{ij},

$$h_{ij}^{\mathrm{TT}} \quad = \quad h_{kl} \left(P_{ki} P_{lj} - \frac{1}{2} P_{kl} P_{ij} \right) \quad \text{where,} \tag{1.6}$$

$$P_{ij} \quad = \quad \delta_{ij} - n_i n_j \quad \text{and} \tag{1.7}$$

n_i is the unit normal along the wave's propagation. Note that $\delta_{ij} h_{ij}^{\mathrm{TT}} = 0$ (traceless), $n_i h_{ij}^{\mathrm{TT}} = 0$ (transverse) and $\bar{h}_{ij}^{\mathrm{TT}} = h_{ij}^{\mathrm{TT}}$. If we choose Cartesian coordinates (x, y, z) and the direction of propagation along $\hat{z}$, h_{ij}^{TT} reduces to,

$$h_{ij}^{\mathrm{TT}} = \begin{pmatrix} h_+ & h_\times & 0 \\ h_\times & -h_+ & 0 \\ 0 & 0 & 0 \end{pmatrix} \ . \tag{1.8}$$

In essence, there are only two independent degrees of freedom (which are identified as GW polarizations), h_+ and $h_\times$.

When $|\mathbf{x}' - \mathbf{x}| \simeq R$ is large enough, we can massage Eq. (1.5) to obtain the

quadrupole formula for gravitational radiation,

$$h_{ij}^{TT}(\mathbf{x}, t) \;=\; \frac{2}{R} \frac{d^2 \mathcal{I}_{kl}(t-R)}{dt^2} \left[P_{ki} P_{lj} - \frac{1}{2} P_{kl} P_{ij} \right] \;,\; \text{where} \tag{1.9}$$

$$\mathcal{I}_{ij} \;=\; I_{ij} - \frac{1}{3} \delta_{ij} I_{kk} \;\; \text{and} \tag{1.10}$$

$$I_{ij}(t) \;=\; \int d^3 x' \, x'^{i} x'^{j} T_{tt}(\mathbf{x}', t) \tag{1.11}$$

is the mass quadrupole moment of the source.

We can ascribe a stress-energy tensor to describe the energy and momentum contained within $h_{\mu\nu}$. It can be shown that [56],

$$T_{\mu\nu}^{\text{GW}} \;=\; \frac{1}{32\pi} \left\langle \nabla_\mu h_{\alpha\beta} \, \nabla_\nu \, h^{\alpha\beta} \right\rangle \;. \tag{1.12}$$

Here, ∇ denotes covariant differentiation. Note that this expression is valid even when the background is not flat, although $h_{\mu\nu}$ is still defined in the transverse-traceless gauge. The angular brackets denote spatial and temporal averaging over an appropriate scale. Specializing this expression for energy flux (GW luminosity) in the transverse, traceless gauge gives,

$$\frac{dE^{\text{GW}}}{dt} \;=\; \int dA \, T_{tk} \, n^k \;, \tag{1.13}$$

$$\;=\; \int dA \, \frac{1}{32\pi} \left\langle \partial_t h_{ij}^{\text{TT}} \partial_k h_{ij}^{\text{TT}} \right\rangle n^k \;, \tag{1.14}$$

$$\;=\; \frac{1}{5} \left\langle \frac{d^3 \mathcal{I}_{ij}}{dt^3} \frac{d^3 \mathcal{I}_{ij}}{dt^3} \right\rangle \;. \tag{1.15}$$

As the notation suggests, dA is an area element and the integral is performed over a sphere bounding the source.

As an illustration, we apply the quadrupole formula, Eq. (1.9) and Eq. (1.15) to a binary composed of two equal masses M, separated by a distance $2a$, and in a circular orbit about their common center of mass [79]. Let the orbit of the binary lie in the

$x - y$ plane. The Keplerian energy of the orbit is,

$$E^{\text{orb}} = -\frac{\mu M}{4a} \tag{1.16}$$

and the orbital frequency is,

$$\Omega = \sqrt{\frac{M}{8a^3}} . \tag{1.17}$$

If the binary is at a distance r, we find that,

$$h_{ij}^{\text{TT}}(t, \mathbf{x}) = \frac{8M}{r}\Omega^2 a^2 \begin{pmatrix} -\cos 2\Omega t_r & -\sin 2\Omega t_r & 0 \\ -\sin 2\Omega t_r & \cos 2\Omega t_r & 0 \\ 0 & 0 & 0 \end{pmatrix} \text{ and} \tag{1.18}$$

$$\frac{dE^{\text{GW}}}{dt} = -\frac{32}{5}M^2(2a)^4\Omega^6 , \tag{1.19}$$

where $t_r = t - |\mathbf{x}' - \mathbf{x}| \simeq t - r$. The energy carried away by GWs reduces the binary's orbital energy. Imposing conservation of energy,

$$E_{\text{tot}} = E^{\text{orb}} + E^{\text{GW}} = \text{const} , \tag{1.20}$$

$$\Rightarrow \frac{dE^{\text{orb}}}{dt} = -\frac{dE^{\text{GW}}}{dt} , \tag{1.21}$$

$$\Rightarrow \frac{dE^{\text{orb}}}{da}\frac{da}{dt} = -\frac{32}{5}M^2(2a)^4\Omega^6 , \tag{1.22}$$

$$\Rightarrow \frac{da}{dt} = -\frac{128M^3}{5a^3} . \tag{1.23}$$

This simple calculation contains several aspects of what is to follow in chapters 2, 3, 4, and 5. Indeed, any solution to the two-body problem in general relativity should reduce to these results in the appropriate limit.

1.4 Gravitational wave detection

The direct detection of GWs is at the forefront of modern research in Physics. The term "direct" here refers to the interaction between GWs and multiple test masses.

Although we do not have a direct detection yet, we expect ground-based instruments like LIGO, VIRGO, GEO and TAMA to make the breakthrough within the next decade or so. The LIGO Scientific Collaboration has already completed five science runs. These ground-based detectors are most sensitive at ~ 100 Hz. The proposed space based GW observatory, LISA is expected to measure GW signals at frequencies centered around ~ 0.01 Hz.

GW detection is not the subject of this dissertation, and thus this section is necessarily brief. However, we illustrate the phenomenon that experimentalists strive to measure. In essence, GWs are metric perturbations which stretch and squeeze the background spacetime. As an example, consider GWs propagating along the $\hat{z}$ axis. According to Eq. (1.1) and Eq. (1.8),

$$g_{\mu\nu} \ = \ \eta_{\mu\nu} + h_{\mu\nu} \ , \tag{1.24}$$

$$\Rightarrow g_{ij} \ = \ \begin{pmatrix} 1 + h_+(t-r) & h_\times(t-r) & 0 \\ h_\times(t-r) & 1 - h_+(t-r) & 0 \\ 0 & 0 & 1 \end{pmatrix} \ . \tag{1.25}$$

Now consider two point masses along the x-axis, separated by a distance L. At any given instant in time, the proper distance between them is,

$$s \ = \ \int \sqrt{g_{ij} dx^i dx^j} \ , \tag{1.26}$$

$$= \ \int_0^L dx \sqrt{(1 + h_+) + 2h_\times \frac{dy}{dx} + (1 - h_+)\left(\frac{dy}{dx}\right)^2 + \left(\frac{dz}{dx}\right)^2} \ , \tag{1.27}$$

$$\simeq \ L + \frac{Lh_+}{2} \tag{1.28}$$

Thus, the effect is to stretch and squeeze the proper distance between the two masses by a fractional amount $\Delta L/L = (s - L)/L \simeq h_+/2$. Experiments are designed to measure precisely this.

Let us now estimate the strain $\Delta L/L$ for radiation from the binary considered in

31

Sec. 1.3. From Eq. (1.18),

$$\Delta L/L \simeq \frac{R_s^2}{ar} \; , \qquad\qquad (1.29)$$

where $R_s = 2M$ is the Schwarzchild radius of each mass. For a $10 M_\odot$ black hole binary in the Coma cluster ($r \sim 100 \text{Mpc}$) separated by $a \sim 10 R_s$, the strain is only $\Delta L/L = 10^{-21}$. To put this in the context of atomic physics, two test masses have to be separated by a distance of $\sim 10^6$ m to experience a change in separation of ~ 1 Fermi (typical size of a nucleus).

Several sources of noise contaminate a GW detector's output. It has been shown that the noise can be subtracted most efficiently using a technique called *matched filtering*. This involves matching a detector's output to theoretical templates of anticipated signals in order to detect the GWs and measure the parameters that characterize the source [81, 5]. This has been the major motivation for several theoretical GW research programs. Indeed, one of the goals of this dissertation is to construct templates for GW signals from extreme mass ratio inspirals.

1.5 Black hole perturbation theory

An important subset of binaries are those in which a stellar mass compact object ($\mu \simeq 1 - 100 M_\odot$) radiates gravitational energy and inspirals into its massive black hole ($M \simeq 10^6 - 10^8 M_\odot$) companion. Such extreme mass ratio inspirals (EMRIs) can be ideally fit within the framework of black hole perturbation theory. We assume that the background spacetime is that of a static massive black hole and treat the smaller object as a spacetime perturbation of $\mathcal{O}(\mu/M)$. (Note that the background in Sec. 1.3 was flat.)

The most "natural" way to solve for these perturbations is to expand the metric about the stationary BH background ($g_{\mu\nu}^{BH}$) as,

$$g_{\mu\nu}(x^\alpha) = g_{\mu\nu}^{BH}(x^\alpha) + h_{\mu\nu}(x^\alpha) \; . \qquad\qquad (1.30)$$

Next, this expansion can be substituted into Einstein's equations and linearized.

When the background spacetime is spherically symmetric (Shwarzschild), $h_{\mu\nu}$ can be expanded in a basis of spherical harmonics. This simplifies the problem dramatically, leading to tractable differential equations for $h_{\mu\nu}$.

Unfortunately, astronomical black holes are not spherically symmetric. They rotate, making them cylindrically symmetric about their axis of spin. Such a spacetime is described by the Kerr metric ,

$$(ds^2)^{\text{Kerr}} = -\left(1 - \frac{2Mr}{\rho^2}\right)dt^2 - \frac{4aMr\sin^2\theta}{\rho^2}dtd\phi + \frac{\rho^2}{\Delta}dr^2$$
$$+\rho^2 d\theta^2 + \left(r^2 + a^2 + \frac{2Mra^2\sin^2\theta}{\rho^2}\right)d\phi^2 \ . \tag{1.31}$$

Here, (r, θ, ϕ, t) are the BL coordinates, which reduce to spherical coordinates in flat space. The symbol M is the black hole's mass, a parametrizes its spin ($-M < a < M$), $\Delta = r^2 - 2Mr + a^2$ and $\rho^2 = r^2 + a^2\cos^2\theta$. The absence of spherical symmetry leads to extremely complicated equations for the metric $h_{\mu\nu}$. Teukolsky [99] resolved this by considering *curvature* perturbations of the form,

$$R_{\alpha\beta\gamma\delta} = R_{\alpha\beta\gamma\delta}^{\text{BH}} + \delta R_{\alpha\beta\gamma\delta} \ , \tag{1.32}$$

instead. A brief outline of his approach follows: Rearranging the Bianchi identity yields,

$$\nabla_\gamma R_{\alpha\beta\mu\nu} = - \nabla_\beta R_{\gamma\alpha\mu\nu} - \nabla_\alpha R_{\beta\gamma\mu\nu} \ . \tag{1.33}$$

Taking another covariant derivative gives,

$$\Box R_{\alpha\beta\mu\nu} = - \nabla^\gamma \nabla_\beta R_{\gamma\alpha\mu\nu} - \nabla^\gamma \nabla_\alpha R_{\beta\gamma\mu\nu} \ . \tag{1.34}$$

Here, $\Box = g_{\mu\nu}^{\text{BH}} \nabla^\mu \nabla^\nu = \nabla_\gamma \nabla^\gamma$ is the covariant wave operator. Substituting Eq. (1.32) in Eq. (1.34) results in a wave equation for $\delta R_{\alpha\beta\mu\nu}$. To extract the radiative degrees of freedom for metric perturbations far away from the source, we project out

the Weyl curvature scalar ψ_4,

$$\psi_4 = -C_{\alpha\beta\gamma\delta} n^\alpha \bar{m}^\beta n^\gamma \bar{m}^\delta \qquad (1.35)$$

where $C_{\alpha\beta\gamma\delta}$ is the Weyl tensor, n^α and $\bar{m}^\beta$ are legs of the Newman-Penrose null tetrad. The Newman-Penrose formalism defines a complete basis at each point in spacetime. Within this formalism, the metric becomes,

$$g^{\mu\nu} = l^\mu n^\nu + n^\nu l^\mu - m^\mu \bar{m}^\nu - m^\nu \bar{m}^\mu , \qquad (1.36)$$

where

$$l^\mu = \left(\frac{r^2 + a^2}{\Delta}, 1, 0, \frac{a}{\Delta} \right) , \qquad (1.37)$$

$$n^\mu = \left(r^2 + a^2, -\Delta, 0, a \right) / (2\rho^2) , \text{ and} \qquad (1.38)$$

$$m^\mu = \left(ia\sin\theta, 0, 1, \frac{i}{\sin\theta} \right) / \left(\sqrt{2} \left(r + ia\cos\theta \right) \right) . \qquad (1.39)$$

The quantities l^μ and n^μ lie along the ingoing and outgoing null directions; m^μ and $\bar{m}^\mu$ span the angular directions.

Remember, the Weyl tensor is the traceless component of the Riemann curvature tensor. Loosely,

$$\text{"Riemann} = \text{Weyl} + \text{Ricci"} . \qquad (1.40)$$

Far away from an EMRI, Ricci= 0 and the background is static,

$$\Rightarrow \psi_4 = -\delta R_{\alpha\beta\gamma\delta} n^\alpha \bar{m}^\beta n^\gamma \bar{m}^\delta . \qquad (1.41)$$

The upshot of this exercise is the Teukolsky equation for the Weyl scalars, Eq. (2.5) of chapter 2. Almost magically, the Teukolsky equation is variable separable. However, we will see in chapters 2 and 3 that separating the variables is not the most optimum method to solve the Teukolsky equation for all applications. Far from the

source, the metric perturbations are very simply related to ψ_4,

$$\psi_4 \;=\; \frac{1}{2}\left(\frac{\partial^2 h_+}{\partial t^2} - i\frac{\partial^2 h_\times}{\partial t^2}\right) . \tag{1.42}$$

Thus, the metric perturbations are the final outcome of this procedure, rather than something with which we begin our analysis.

1.6 This dissertation

As mentioned earlier, the aim of this dissertation is to exploit BH perturbation theory to understand certain crucial aspects of the two body problem. The treatment of EMRIs using BH perturbation theory has been the subject of much recent research (eg: Ref. [98, 53, 36, 97] and references therein). To a large extent, the need to model EMRIs accurately has been motivated by the prospect of LISA, the proposed space based GW detector. Signals from a number of GW sources are likely to be present in LISA's science data. Researchers are currently working on developing the tools and investigating the capability of LISA data analysis routines to extract the rich science encoded in these signals. Accurate theoretical models of GWs are a critical input to the development of a robust mock data bank, and to the development of accurate LISA measurement techniques. This is the major motivation for most of this dissertation (chapters 2, 3, and 4). In the process however, we have developed a versatile semi-analytical/semi-numerical toolkit that can be used for a wide variety of applications involving perturbation theory, not just LISA science. For example, this toolkit could be used in a comparison of results between perturbation theory and full numerical relativity. Chapter 5 explores this to some extent.

The EMRI problem can be broken down into two related parts: (*i*) A determination of the *inspiral trajectory* followed by the smaller object, and (*ii*) A characterization of the *gravitational waveforms* that result from such an inspiral. Chapters 2 and 3 develop techniques to generate theoretical inspiral waveforms. Chapter 4 generalizes an earlier approach to model a crucial part of the inspiral trajectory. Chapter 5

uses the tools developed in chapters 2, 3, and 4 along with earlier work to estimate the residual linear momentum after black hole mergers by extrapolating results from perturbation theory to higher mass ratios. The rest of this section highlights the goals and important results from each chapter.

1.6.1 Building gravitational waveforms

The Teukolsky perturbation equation, Eq. 2.5 describes the gravitational waveforms, $h_+(t)$ and $h_\times(t)$ that result from a given perturbation. It is a second order, linear and inhomogeneous partial differential equation that describes perturbations due to scalar, vector and tensor fields in the vicinity of Kerr black holes. The goal of chapters 2 and 3 is to build a versatile numerical code to study EMRIs by solving the Teukolsky equation. This code accepts *any* reasonable world-line of the smaller object as an input and computes the resulting GWs.

The source term of the Teukolsky equation contains Dirac delta functions (and its derivatives) due to the point-like nature of the perturbing mass. Chapter 2 develops an inherently discrete model of the singular source term and uses it to build a (2+1)D finite-difference numerical code to solve the Teukolsky equation. [The azimuthal dependence is factored out, resulting in an equation in two space and one time dimensions, (r, θ, t).] Chapter 2 shows that the code can calculate gravitational waveforms when the smaller object is confined to a simple bound, circular orbit on the equatorial plane of the central black hole. Tables 2.1 – 2.6 show that the agreement in the fluxes of the radiated energy with previous calculations is greater than 99%. Additionally, this improved model of the source term leads to an order of magnitude gain in accuracy and performance (computational speed) over similar efforts in the past. Further, Fig. 2-1 shows that the code reproduces the well known quasi-normal ringing of the black hole in response to a perturbation in the *absence* of the smaller object. In summary, the code developed in chapter 2 greatly improves the speed and accuracy with which we can solve the Teukolsky equation, despite the singular source term.

The goal of chapter 3 is to extend the code to accommodate more physically

interesting world lines of the smaller object. Models of the source term used in chapter 2 lead to spurious high-frequency numerical instabilities if the world line of the smaller object is non-equatorial and eccentric. Chapter 3 generalizes the formalism in chapter 2 such that these instabilities can be suppressed. One of the key steps in this extension is to incorporate a Gaussian filter within the source term. When the smaller object is confined to a bound geodesic, Refs. [53, 36] have developed a code to solve the Teukolsky equation using a formalism based on the decomposition of the Teukolsky equation into Fourier frequency modes. In cases where it was possible, chapter 3 validates waveforms with extensive comparisons to this frequency-domain based formalism. Tables 3.1 – 3.6 show that the correlation between waveforms from these two completely different procedures is greater than 99% for a large fraction of parameter space. Fig. 3-2 shows a typical result of such a comparison. The code is now in a position to calculate the waveforms corresponding to any reasonable world line of the smaller object. Fig. 3-4 shows GWs generated during the early inspiral portion of a typical EMRI. The evolution of the amplitude and frequency of the radiation during the inspiral is clearly visible.

1.6.2   Generating inspiral trajectories

The trajectory of the smaller object in an EMRI can be broken down into three regimes: (a) An early adiabatic inspiral phase, where the inspiral time-scale is much larger than the orbital period. In this stage, the trajectory can be approximated as a sequence of bound geodesics; (b) A late-time radial infall, which can be approximated as a single unstable, plunging geodesic; and (c) A regime where the body transitions from inspiral to plunge [20, 74]. In [74], Ori and Thorne introduced a method to model the trajectory during the transition when the smaller object starts from an approximately circular, equatorial trajectory. Chapter 4 generalizes this procedure to include non-equatorial and eccentric trajectories. I describe the relevance and some details of this generalization here.

Up to initial conditions, a set of three constants, the energy (E), the component of angular momentum along the spin axis (L_z), and the Carter constant (Q) define a

geodesic around a rotating black hole. The procedure in chapter 4 to predict the trajectory of the smaller object during the transition consists of expanding the geodesic equations about the last stable (geodesic) orbit and subjecting them to evolving "constants", $[E(t), L_z(t), Q(t)]$. Tables 4.1 and 4.2 show that the procedure in chapter 4 reproduces the results in [74] for the special case of an approximately circular, equatorial trajectory. Further, Figs. 4-1 and 4-3 show that that the trajectory exhibits the generally expected qualitative behavior arising from effects of both radiation reaction and unstable geodesics.

This prescription is meant to serve as a stopgap for many other open and important problems. Although it neglects the conservative part of the self-force, it can be combined with earlier work to make a reasonable prediction of the entire world line of the compact object from inspiral to plunge. This world line can be fed to the code developed in chapters 2 and 3 to produce the resulting gravitational waveforms from an EMRI. A number of researchers are working towards solving the radiation reaction problem exactly. Although approximate, the results in [74] may serve as an independent check for these solutions. This work will also be a good point of comparison for inspirals predicted by full (3+1)D numerical relativity at small mass ratios.

1.6.3 Recoil velocities from black hole mergers

Comparable mass black hole binaries radiate gravitational energy as they spiral into each other and merge. An integration of the momentum carried away by gravitational waves from asymmetric binaries results in a non-zero recoil velocity of the merged object. Earlier calculations have shown that perturbation theory extrapolated to mass ratios (μ/M) of $\mathcal{O}(0.1)$ yields reliable estimates for the recoil velocity. In chapter 5 we use the numerical toolkit developed in chapters 2, 3, and 4 to improve earlier estimates for the recoil velocities from black hole mergers. We compare these results with (3+1) dimensional numerical relativity. Such a comparison is both a calibration of the reliability of our code and a consistency check for numerical relativity.

The most important results in this chapter are the recoil velocities shown in tables

5.1-5.5 and Fig. 5-5. The tables show recoil velocities for binaries over a range of mass ratios and black hole spins. These estimates are in excellent agreement with earlier calculations from numerical relativity, perturbation theory and PN theory. This is significant because: (a) Exploration of parameter space is computationally easier in BH perturbation theory than in numerical relativity, and (b) BH perturbation theory is more accurate for small mass ratios and in strong fields than PN theory.

Chapter 2

Towards adiabatic waveforms for inspiral into Kerr black holes: I. A new model of the source for the time domain perturbation equation

*This chapter is based on Physical Review D **71**, 084008 (2007), which was written in collaboration with Gaurav Khanna and Scott A. Hughes.*

2.1   Introduction

2.1.1   Background

The extreme mass ratio limit of binary systems — binaries with one mass far smaller than the other — has been a special focus of research in gravitation in recent years. This is in part because this problem is, at least formally, particularly clean and beautiful: the mass ratio allows us to treat the binary as an exact black hole solution plus a perturbation due to the secondary mass. Perturbative techniques can be used to analyze the system, making it (in principle at least) much more tractable than the general two-body problem in general relativity.

This limit is also of great astrophysical interest, as it perfectly describes *capture binaries*: binary systems created by the capture of stellar mass compact objects onto relativistic orbits of massive black holes in galaxy cores. Post formation, the evolution of such binaries is driven by gravitational-wave (GW) emission — the GW backreaction circularizes and shrinks the binaries, eventually driving the smaller body to plunge and merge with its larger companion. Such events are now believed to be relatively abundant (see Ref. [52] for up-to-date discussion and review of the relevant literature). Since the last year or so of the inspiral is likely to generate GWs that lie in the low-frequency band of space-based GW antennae such as *LISA* [67], *extreme mass ratio inspirals* (or EMRIs) are key targets for future GW observations.

This chapter builds on earlier work in perturbation theory to lay the foundation for the development of *adiabatic EMRI waveforms*. "Adiabatic" refers to the fact that they are computed using an approximation to the true equations of motion that takes advantage of the nearly periodic nature of the smaller body's motion on "short" timescales. This approximation fails to capture certain important aspects of the binary's evolution. In particular, adiabatic waveforms only incorporate *dissipative* effects of the small body's perturbation — effects which cause radiation of energy and angular momentum to distant observers and down the hole, driving the orbit to decay. *Conservative* effects — effects which conserve energy and angular momentum, but push the orbit away from the geodesic trajectory of the background spacetime — are missed in this approach. It has been convincingly demonstrated [83] that conservative effects change orbital phasing in a way that could be observationally significant. The dissipative-only adiabatic approach to EMRI waveform generation is thus, by construction, somewhat deficient.

In our view, this deficiency is outweighed by the fact that it will produce waveforms that capture the spectral features of true waveforms — a complicated shape "colored" by the three fundamental orbital frequencies and their harmonics. Also, the adiabatic approach is likely to produce these waveforms on a relatively short timescale. Though not perfectly accurate, adiabatic waveforms will be an invaluable tool in the short term for workers developing a data analysis architecture for measur-

ing EMRI events. In the long term, these waveforms may even be accurate enough to serve as "detection templates" for EMRI events. *Measuring* the characteristics of EMRI sources will require matching data with as accurate a model as can be made, and over as long a timespan as possible — perhaps a year or more. By contrast, *detecting* EMRI events does not require matching a signal with a template for such a long time [42]. For the short integration times needed for detection, work in progress indicates that conservative effects do not shift the phase so badly that the signal fails to match a template. What shift does accumulate due to conservative effects can be accommodated by systematic errors in source parameters, allowing detection to occur. (This is discussed in Appendix A of Ref. [34].)

2.1.2 Our approach to adiabatic inspiral

The approach which we advocate for building adiabatic waveforms uses a hybrid of frequency-domain and time-domain perturbation theory techniques. These two techniques have complementary strengths and weaknesses; by combining the best features of both toolsets, we hope to make waveforms that are as accurate as possible. Though a diversion from the main topic of this paper, this approach is a key motivation for our work. We thus ask the reader to indulge us as we briefly describe our rationale.

In the adiabatic limit and neglecting conservative effects, the separation of timescales means that orbits are, to high accuracy, simply geodesic trajectories of the spacetime on short timescales. The orbital decay that is driven by backreaction amounts to the system evolving from one geodesic orbit to another. Computing the effect of radiation reaction thus amounts to computing the sequence of orbits through which the system passes en route to the final plunge of the smaller body into the large black hole [54].

A geodesic orbit is characterized (up to initial conditions) by three constants: energy E; axial angular momentum L_z; and "Carter constant" Q (see, e.g., [72], Chap. 33). The Carter constant can be approximately interpreted as the square of the component of angular momentum perpendicular to the spin axis. It is a consequence of the presence of a Killing tensor in Kerr geometry. Computing this sequence of orbits is equivalent to computing the rate at which these constants change due to

radiative backreaction. In this picture, it is useful to regard each orbit (E, L_z, Q) as a point in an orbital phase space, and to regard the rates at which they evolve, $(\dot{E}, \dot{L}_z, \dot{Q})$, as defining a tangent vector to the trajectory an evolving system traces through this phase space. Adiabatic radiation reaction thus amounts to calculating this tangent at all orbits.

In the extreme mass ratio limit, the smaller body moves very slowly through orbit space — it spends many orbits in the vicinity of each (E, L_z, Q). This slow evolution means that the tangent vector is most accurately represented by the *average* rate at which these constants evolve: $(\langle \dot{E} \rangle, \langle \dot{L}_z \rangle, \langle \dot{Q} \rangle)$, where the angle brackets denote an appropriate averaging with respect to the orbits. Such an averaging is defined in Ref. [34].

Once adiabatic radiation reaction data has been found for all orbits, it is straightforward to choose initial conditions and compute the worldline $\mathbf{z}(t)$ which an inspiralling body follows. In this framework, it is just a geodesic worldline with the constants slowly evolving:

$$\mathbf{z}(t) = \mathbf{z}_{\text{geod}}[E(t), L_z(t), Q(t)] \ . \tag{2.1}$$

This worldline can then be used to build the source term for the wave equation, allowing us to compute the gravitational waves generated as the small body spirals in. We note here that this approach is conceptually identical to the "kludge" presented in Ref. [8]. Indeed, the almost unreasonable success of kludge waveforms served as an inspiration for this formulation of inspiral[1].

In the hybrid approach, a frequency-domain code would be used for the adiabatic radiation reaction, and a time-domain code used to generate the waves from a small body following the worldline that radiation reaction defines. Since any function built from bound Kerr black hole orbits has a spectrum that is fully described by three easily computed frequencies and their harmonics [90, 35], the averaging needed in

---

[1]The major difference between the hybrid inspiral described here and the kludge is that the hybrid inspiral aims to correctly solve a wave equation at all points along the orbit. The kludge instead uses a physically motivated approximate wave formula based on variation of the source's multipole moments, defined in a particular coordinate system.

this prescription is extremely fast and easy to compute in the frequency domain. Many harmonics may be needed, but each harmonic is independent of all others. Frequency-domain codes are thus easily parallelized and the calculation can be done very rapidly. In the time domain, averaging is much more cumbersome — a geodesic orbit and the radiation it generates must be followed over many orbits to insure that all beatings between different harmonics have been sampled. Convergence to the true average for quantities like $\langle \dot{E} \rangle$ will be slow for generic (inclined and eccentric) Kerr black hole orbits (the most interesting case, astrophysically).

By contrast, building the associated gravitational waveform with a frequency-domain code is rather cumbersome. One must build the Fourier expansion of the waves from many coefficients, and accurately sum them to produce the wave at any moment of time. The benefit of each harmonic being independent of all others is lost. In the time domain, building the waveform is automatic — modulo two time derivatives, the waveform *is* the observable that the code produces. Given a worldline, it is straightforward to build a source for the time-domain wave equation; one then cannot help but compute the waveform that source generates.

We are thus confident that by using both frequency and time-domain perturbation techniques, we can get the best of both worlds — letting each technique's complementary strengths shine to build EMRI waveforms that are as accurate as possible, in the context of the adiabatic approximation.

2.1.3 This chapter

Key to the success of the hybrid approach is the development of fast, accurate codes for both frequency and time-domain approaches to black hole perturbation theory. First results from a frequency-domain code which can handle generic orbits have recently been presented [36], and the last major formal step (understanding the adiabatic evolution of Carter's constant Q due to GW emission) is essentially in hand [34, 89, 88]. The frequency-domain side of this program is thus in a good state. Our goal now is to develop time-domain tools sufficiently robust and generic to handle the case of interest.

The major difficulty in building a time-domain perturbation code is the source term, representing the smaller member of the binary which perturbs the large black hole's spacetime. In the frequency domain, the small body is usually approximated as having zero spatial extent, and can be represented using delta functions (and their derivatives). One then constructs a Green's function from solutions of the source-free perturbation equation and integrates over the source. Thanks to the delta nature of the source in this representation, this integral can be done analytically. This trick cannot be done in the time domain — one must choose a functional form of the source which can be represented on a finite difference grid. The challenge is to pick a representation that accurately captures the very narrow spatial extent of the source, but is sufficiently smooth that the source does not seed excessive amounts of numerical error. This is particularly difficult for sources representing highly dynamic, generic Kerr black hole orbits in which the source rapidly moves across the grid.

Much recent success in this approach has come from representing the source as a truncated, narrow Gaussian [68]. Khanna [59] and Burko and Khanna [23] have so far examined some orbit classes (equatorial orbits, both circular and eccentric) and found that they can quickly and robustly generate waveforms from orbits around Kerr black holes. As a diagnostic of this technique, they compute the flux of energy carried by this radiation and find agreement with pre-existing frequency-domain calculations at the few percent level.

An interesting recent development is the use of finite element techniques to represent time-domain sources. Such methods are tailor made for resolving problems with multiple lengthscales, and as such may be ideal for the EMRI problem. Sopuerta and Laguna [94] have found that a finite element code makes it possible to represent the source with amazing accuracy — agreement with frequency-domain calculations at the few hundredths of a percent level seems common. To date, they have only examined binaries in which the larger black hole is non-rotating, but they argue convincingly [93] that the difficulties required to model Kerr perturbations should not be terribly difficult to surmount. These techniques are an extremely promising direction that is sure to develop extensively in the next several years.

Our goal in this paper is to develop another representation of the source term that is simpler (and concomitantly less accurate) than finite element methods, but that is developed somewhat more systematically than the truncated Gaussian. The key ingredient of this approach is an extension of a finite impulse representation of the Dirac delta function [37, 102]. In essence, one writes the discrete delta as a series of spikes on the finite difference grid, with the largest spike centered at the argument of the delta, and with the spikes rapidly falling off away from this center. One chooses the magnitude of the spikes such that the delta function's integral properties are preserved, particularly the rule that

$$\int dx\, f(x)\delta(x - x_0) = f(x_0) \,.$$ (2.2)

The discrete delta described in Ref. [37] allows one to make a tradeoff between localization and smoothness — one can smear the delta over k points, choosing k to be small if source sharpness is the key property needed, or allowing k to expand if too much sharpness causes numerical problems. This representation introduces a kind of optimization parameter which one can engineer as needed to find the best compromise between smoothness and localization.

We extend the finite impulse representation of the delta described in [37, 102] in two important ways. First, the source term of the Teukolsky equation requires not just the delta, but also the delta's first and second derivatives. We therefore generalize this procedure to develop discrete delta derivatives. If the delta is represented by k points, then both derivatives will require $k + 2$ points. The guiding principle of this extension is again the notion that the integral properties of these functions must be preserved:

$$\int dx\, f(x)\delta'(x - x_0) = -f'(x_0) \,,$$
$$\int dx\, f(x)\delta''(x - x_0) = f''(x_0) \,.$$ (2.3)

(Here, prime means d/dx.)

If the discrete delta function does not lie precisely on a grid point, then one must use interpolation to appropriately weight impulse functions from the neighboring grid points. Our second extension of Ref. [37] is to introduce higher order interpolation (cubic) which offers another way to trade smoothness for localization. This is particularly valuable when (as in our application) the source is coupled to a wave equation.

We test this representation by developing a new time-domain Teukolsky equation solver which uses this form of the delta for its source (the "δ-code") and comparing to a well-established code (see, e.g., [68, 59, 23]) which uses a truncated Gaussian (the "G-code"). The G-code has been described in detail in a previous publication [68]; for the purpose of this paper, the most salient feature of this code is how it represents the source term. The G-code begins with the following approximation to the Dirac delta function:

$$\delta[x - x(t)] \simeq \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{[x - x(t)]^2}{2\sigma^2}\right) . \tag{2.4}$$

[Cf. Ref. [68], Eq. (19).] The width σ is chosen to be small enough that this delta only spreads across a few grid zones. The Teukolsky equation source is then built from this Gaussian representation and its derivatives.

The δ-code by contrast uses the representation described in detail in the following sections of this paper — a representation that is discrete by design, rather than a discretization of a continuous delta approximation. The principle advantage of this form seems to be that it makes it possible to rigorously enforce integral identities involving the delta *plus* its derivatives.

There are a few other minor differences between these two codes, which are artifacts of the codes' independent developments. Chief among these differences are the use of slightly different axial coordinates (the G-code uses the usual Boyer-Lindquist coordinate ϕ; the δ-code follows Ref. [64] and uses a coordinate $\tilde{\phi}$ defined in Sec. 2.2.1), and the use of slightly different fundamental "fields" (i.e., slightly different representations of the Weyl curvature scalar ψ_4 which the Teukolsky equation governs). There are also some differences in the way the two codes implement boundary conditions. We present a detailed comparison of the results from the two codes in Sec.

2.4. It's worth pointing out that that we also have taken the G-code and replaced its source term with that used by the δ-code. This exercise confirmed all of the results we obtained with the δ-code, demonstrating that these minor differences had no impact on our results.

As a proof-of-principle check of this idea's validity, we restrict our present analysis to circular, equatorial orbits. The results from both codes are then compared against frequency-domain results. Flux of energy carried by gravitational waves is a very useful benchmark with which to diagnose a perturbation theory code's accuracy (especially for very simple orbits when averaging is easy both in time and frequency domains). In all cases, we find (after some experimentation to optimize our discrete delta) that this new source form is more accurate (typically by factors of $2-5$) and faster (often by factors of about 10) than the truncated Gaussian. For our purpose, it appears that this form of source function will be very well-suited to serve as the core of the time-domain portion of our hybrid approach to EMRI waveforms.

Chapters 3 and 5 will then apply this technique to flesh out the hybrid approach. Chapter. 3 will examine how well this source works for highly dynamical trajectories — generic (inclined and eccentric) geodesic orbits and non-geodesic trajectories that evolve due to radiation reaction. Chapter. 5 will use this toolkit to compute recoil velocities from black hole mergers. We find that the discrete delta source term handles such orbits very robustly, validating earlier results for generic orbits. We are now in a position to develop hybrid EMRI waveforms in earnest, using frequency-domain tools to compute the effects of radiation reaction, building an inspiral worldline from those effects, and finally computing the waveform with our time-domain code.

2.1.4   Organization of this chapter

The remainder of this chapter is organized as follows: Section 2.2 reviews how one solves the Teukolsky equation in the time domain, introducing the equation itself, specializing to the form that we use for our calculations, and showing how to extract waveforms and fluxes from its solutions. We first review in Sec. 2.2.1 how one solves for the homogeneous (source-free) form of the Teukolsky equation, an important first

step to developing a robust solver for the sourced case. We follow very closely the procedure laid out in Ref. [64]; this section is thus largely a review and summary of that paper (with a few minor corrections noted). Section 2.2.2 then describes in detail the form of the source term that applies when perturbations arise from an orbiting body.

The need to model this source using a delta function motivates Sec. 2.3, our model for a discrete delta and its derivatives. This section presents the key new idea of this paper. After describing the basic idea behind our discrete delta, we first present in some detail (Sec. 2.3.1) an extremely simple two-point discrete delta function. This illustrates the concepts and principles of this approach. We then generalize this idea to a multiple point delta in Sec. 2.3.2, and then show how to smooth things with higher order interpolation in Sec. 2.3.3. Some preliminary issues related to the convergence of quantities computed using the discrete delta are introduced in Sec. 2.3.4.

We test this delta representation in Sec. 2.4, examining how the various methods we develop work at describing the Teukolsky source function. Section 2.4.1 first compares the different discrete delta functions with each other, demonstrating how the different approaches change the quality and accuracy of our results. Based on this analysis, we choose to use the high order (cubic) delta described in Sec. 2.3.3 in the remainder of our work. We then examine the convergence of our code, demonstrating second-order convergence in Sec. 2.4.2. Finally, in Sec. 2.4.3 we compare the discrete delta with the Gaussian source function, demonstrating explicitly how this new representation improves both the code's speed and accuracy.

Our benchmark for evaluating our results is to compare the energy flux carried by the system's emitted gravitational waves to results obtained using a frequency-domain code [53, 36]. This operation requires us to extract these waves at a particular finite radius. Section 2.5 examines the dependence of these fluxes as a function of extraction radius, and finds that they are very well fit by a simple power law. Using this law, we can easily extrapolate our results to very large radius; doing so greatly improves agreement with frequency-domain results, typically indicating that our errors are significantly smaller than 1% for a large fraction of parameter space.

Concluding discussion in given in Sec. 2.6. Besides summarizing the major findings of this analysis, we discuss areas where this new computational technology can be applied.

2.2 Numerical implementation of the Teukolsky equation in the time domain

Here we describe the evolution algorithm used in the δ-Code, built using a two step Lax-Wendroff algorithm. Our notation and approach closely follow that used in [64]; some of this section therefore can be considered a summary of that paper. All details related to the G-Code were described in [68].

Teukolsky derived a master equation that describes perturbations due to scalar, vector and tensor fields in the vicinity of Kerr black holes in [100, 99]. In Boyer-Lindquist coordinates, this equation reads

$$
-\left[\frac{(r^2 + a^2)^2}{\Delta} - a^2 \sin^2 \theta \right] \partial_{tt} \Psi - \frac{4Mar}{\Delta} \partial_{t\phi} \Psi
$$
$$
-2s \left[r - \frac{M(r^2 - a^2)}{\Delta} + ia \cos \theta \right] \partial_t \Psi
$$
$$
+ \Delta^{-s} \partial_r \left(\Delta^{s+1} \partial_r \Psi \right) + \frac{1}{\sin \theta} \partial_\theta \left(\sin \theta \partial_\theta \Psi \right) +
$$
$$
\left[\frac{1}{\sin^2 \theta} - \frac{a^2}{\Delta} \right] \partial_{\phi\phi} \Psi
$$
$$
+ 2s \left[\frac{a(r - M)}{\Delta} + \frac{i \cos \theta}{\sin^2 \theta} \right] \partial_\phi \Psi
$$
$$
- \left(s^2 \cot^2 \theta - s \right) \Psi = -4\pi \left(r^2 + a^2 \cos^2 \theta \right) T, \tag{2.5}
$$

where M is the mass of the black hole, a its angular momentum per unit mass, $\Delta = r^2 - 2Mr + a^2 = (r - r_+)(r - r_-)$, $r_\pm = M \pm \sqrt{M^2 - a^2}$ and s is the "spin weight" of the field. The $s = \pm 2$ versions of these equations describe perturbations to the Weyl curvature tensor, in particular the radiative degrees of freedom ψ_0 and ψ_4. That is, $\Psi = \psi_0$ for $s = +2$, and $\Psi = \rho^{-4} \psi_4$ for $s = -2$, with $\rho = -1/(r - ia \cos \theta)$. The T in the RHS of this equation depends on the details of the perturbing source.

It is here that the Dirac delta function and its derivatives enter. A discussion of T is postponed to the latter half of this section, after we discuss the numerical evolution of the homogeneous Teukolsky equation.

Gravitational waves, h_+ and $h_\times$ as well as the energy flux dE/dt [60, 24], can be obtained far away from the system by using $s = -2$ in Eq. (2.5) and then identifying

$$\psi_4 = \frac{1}{2}\left(\frac{\partial^2 h_+}{\partial t^2} - i\frac{\partial^2 h_\times}{\partial t^2}\right) , \tag{2.6}$$

$$\frac{dE}{dt} = \lim_{r\to\infty}\left[\frac{1}{4\pi r^6}\int_\theta\int_\phi \sin\theta\, d\theta\, d\phi\left|\int_{-\infty}^t d\tilde{t}\,\Psi(\tilde{t},r,\theta,\phi)\right|^2\right]$$

$$= \lim_{r\to\infty}\left[\frac{1}{2}\int_\theta \sin\theta\, d\theta\left|\int_{-\infty}^t d\tilde{t}\,\Phi(\tilde{t},r,\theta)\right|^2\right] . \tag{2.7}$$

The θ and ϕ directions are taken with respect to the black hole's spin axis; the function $\Phi(t,r,\theta)$ is a reweighting of the field Ψ which we define precisely in Eq. (2.12) below.

2.2.1 Homogeneous Teukolsky equation

Reference [64] demonstrated stable numerical evolution of (2.5) for $s = -2$. The δ-code has been built using the algorithm presented in [64], after accounting for some typographical errors, which are also discussed in [77]. The contents of this section are largely review of the results presented in [64]; as such, our discussion is particularly brief here.

Our code uses the tortoise coordinate r^* in the radial direction, and azimuthal coordinate $\tilde{\phi}$; these coordinates are related to the usual Boyer Lindquist quantities by

$$dr^* = \frac{r^2 + a^2}{\Delta}dr \tag{2.8}$$

$$\Rightarrow r^* = r + \frac{2Mr_+}{r_+ - r_-}\ln\frac{r - r_+}{2M}$$

$$- \frac{2Mr_-}{r_+ - r_-}\ln\frac{r - r_-}{2M} , \tag{2.9}$$

52

and

$$d\tilde{\phi} = d\phi + \frac{a}{\Delta}dr \qquad (2.10)$$

$$\Rightarrow \tilde{\phi} = \phi + \frac{a}{r_+ - r_-}\ln\frac{r - r_+}{r - r_-}. \qquad (2.11)$$

Following [64], we factor out the azimuthal dependence and use the ansatz,

$$\Psi(t, r^*, \theta, \tilde{\phi}) = e^{im\tilde{\phi}}r^3\Phi(t, r^*, \theta). \qquad (2.12)$$

Defining

$$\Pi \equiv \partial_t\Phi + b\,\partial_{r^*}\Phi, \qquad (2.13)$$

$$b \equiv \frac{r^2 + a^2}{\Sigma}, \qquad (2.14)$$

and

$$\Sigma^2 \equiv (r^2 + a^2)^2 - a^2\,\Delta\,\sin^2\theta \qquad (2.15)$$

allows the Teukolsky equation to be rewritten as

$$\partial_t\boldsymbol{u} + \boldsymbol{M}\partial_{r_*}\boldsymbol{u} + \boldsymbol{L}\boldsymbol{u} + \boldsymbol{A}\boldsymbol{u} = \boldsymbol{T}, \qquad (2.16)$$

where

$$\boldsymbol{u} \equiv \{\Phi_R, \Phi_I, \Pi_R, \Pi_I\} \qquad (2.17)$$

is the solution vector. The subscripts R and I refer to the real and imaginary parts respectively. (Note that the Teukolsky function Ψ is a complex quantity.) The matrices $\boldsymbol{M}$, $\boldsymbol{A}$ and $\boldsymbol{L}$ are

$$\boldsymbol{M} \equiv \begin{pmatrix} b & 0 & 0 & 0 \\ 0 & b & 0 & 0 \\ m_{31} & m_{32} & -b & 0 \\ -m_{32} & m_{31} & 0 & -b \end{pmatrix}, \qquad (2.18)$$

53

$$
\boldsymbol{A} \equiv \begin{pmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \\ a_{31} & a_{32} & a_{33} & a_{34} \\ -a_{32} & a_{31} & -a_{34} & a_{33} \end{pmatrix} , \tag{2.19}
$$

and

$$
\boldsymbol{L} \equiv \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ l_{31} & 0 & 0 & 0 \\ 0 & l_{31} & 0 & 0 \end{pmatrix} , \tag{2.20}
$$

where

$$
m_{31} = -bc_1 + b\partial_{r^*}b + c_2 , \tag{2.21}
$$

$$
m_{32} = bc_3 + 2am(r^2 + a^2)/\Sigma^2 , \tag{2.22}
$$

$$
\begin{aligned}
a_{31} = {} & \Delta \frac{m^2 + 2\cos\theta s m + \cos^2\theta s^2 - \sin^2\theta s}{\Sigma^2 \sin^2\theta} \\
& -6\Delta \frac{a^2 + r\,(r\,(s+2) - M\,(s+3))}{r^2\Sigma^2} ,
\end{aligned} \tag{2.23}
$$

$$
a_{32} = \frac{4M(r-1)smaM + (6am\Delta)/r}{\Sigma^2} , \tag{2.24}
$$

$$
a_{33} = c_1 , \tag{2.25}
$$

$$
a_{34} = -c_3 , \tag{2.26}
$$

$$
l_{31} = -\frac{\Delta}{\Sigma^2}\frac{\partial^2}{\partial\theta^2} - \cot\theta\frac{\Delta}{\Sigma^2}\frac{\partial}{\partial\theta} , \tag{2.27}
$$

$$
c_1 = 2s(-3Mr^2 + Ma^2 + r^3 + ra^2)/\Sigma^2 , \tag{2.28}
$$

$$
c_2 = -2\frac{r\Delta\,(1+s) - (a^2 - r^2)Ms}{\Sigma^2} - \frac{6\Delta b}{r\Sigma} , \tag{2.29}
$$

$$
c_3 = 2a(2rMm + \Delta s\cos\theta)/\Sigma^2 . \tag{2.30}
$$

The equations above have been written such that the typographical errors in [64]'s a_{31}, a_{32}, a_{34} and c_2 are obvious. It turns out that the coefficients listed in [64] are correct when the ansatz $\Psi(t, r^*, \theta, \tilde{\phi}) = e^{im\tilde{\phi}}\Phi(t, r^*, \theta)$ is used. $\boldsymbol{T}$ is a quantity constructed from the source term and is discussed in the latter half of this section.

54

Rewriting Eq. (2.16) as

$$\partial_t \boldsymbol{u} + \boldsymbol{D} \partial_{r^*} \boldsymbol{u} = \boldsymbol{S} \; , \tag{2.31}$$

where

$$\boldsymbol{D} \equiv \begin{pmatrix} b & 0 & 0 & 0 \\ 0 & b & 0 & 0 \\ 0 & 0 & -b & 0 \\ 0 & 0 & 0 & -b \end{pmatrix} , \tag{2.32}$$

$$\boldsymbol{S} = \boldsymbol{T} - (\boldsymbol{M} - \boldsymbol{D}) \partial_{r^*} \boldsymbol{u} - \boldsymbol{L}\boldsymbol{u} - \boldsymbol{A}\boldsymbol{u}, \tag{2.33}$$

and subjecting it to Lax-Wendroff iterations produces stable time-evolutions. Each Lax-Wendroff iteration consists of two steps. In the first step, the solution vector between grid points is obtained from

$$\boldsymbol{u}_{i+1/2}^{n+1/2} = \frac{1}{2} \left(\boldsymbol{u}_{i+1}^n + \boldsymbol{u}_i^n \right) - \tag{2.34}$$
$$\frac{\delta t}{2} \left[\frac{1}{\delta r^*} \boldsymbol{D}_{i+1/2}^n \left(\boldsymbol{u}_{i+1}^n - \boldsymbol{u}_i^n \right) - \boldsymbol{S}_{i+1/2}^n \right] \; .$$

This is used to compute the solution vector at the next time step,

$$\boldsymbol{u}_i^{n+1} = \boldsymbol{u}_i^n - \delta t \left[\frac{1}{\delta r^*} \boldsymbol{D}_i^{n+1/2} \left(\boldsymbol{u}_{i+1/2}^{n+1/2} - \boldsymbol{u}_{i-1/2}^{n+1/2} \right) - \boldsymbol{S}_i^{n+1/2} \right] \; . \tag{2.35}$$

The angular subscripts are dropped here for clarity. All angular derivatives were computed using second order centered finite difference expressions. Notice that the matrices $\boldsymbol{D}$, $\boldsymbol{A}$ and $\boldsymbol{M}$ are time independent. In addition, the time stepping must satisfy the Courant-Friedrichs-Lewy condition [64], $\delta t \leq \max\{\, \delta r^*, 5M\, \delta\theta \,\}$, where δt is the time step[2].

Following [64], we set Φ and Π to zero on the inner and outer radial boundaries.

---

[2]Conducting a von Neumann local stability analysis on all the points of our numerical grid yields that this condition is sufficient for stable evolutions. See reference [64] for more details.

55

While the asymptotic behavior

$$\lim_{r^* \to -\infty} |\Psi| \propto \Delta^{-s} \tag{2.36}$$

makes this condition reasonably accurate at the inner boundary, it is clearly unphysical at the outer boundary. By placing our outer boundary sufficiently far, error due to our outer boundary condition can be made unimportant; reflections from the outer boundary have no important impact on our results. Symmetry of the spheroidal harmonics is used to determine the angular boundary conditions. For even $|m|$ modes, we have $\partial_\theta \Phi = 0$ at $\theta = 0, \pi$. On the other hand, $\Phi = 0$ at $\theta = 0, \pi$ for modes of odd $|m|$.

As a test of our evolution equation, we have examined source-free field evolution (setting $\boldsymbol{T} = 0$) for a variety of initial data, in particular comparing extensively with the results of [64]. As an example of one of our tests, we take initial data corresponding to an ingoing, narrow Gaussian pulse. This data perturbs the black hole, causing it to ring down according to its characteristic quasi-normal frequencies. We find extremely good agreement in mode amplitude and evolution (typically $\sim 1\%$ error) with results from [64]. Figure 2-1 shows the result of such a test, illustrating the quasi-normal ringing and power law tail for the $l = 2$, $m = 0$ mode of a black hole with spin parameter $a = 0.9$.

2.2.2 The source term

We now consider the source term, T, of Eq. (2.1). It is given by

$$T = 2\rho^{-4}T_4 , \tag{2.37}$$

$$\begin{aligned}
T_4 = {} & (\tilde{\Delta} + 3\gamma - \bar{\gamma} + 4\mu + \bar{\mu})(\tilde{\Delta} + 2\gamma - 2\bar{\gamma} + \bar{\mu})T_{\bar{m}\bar{m}} \\
& -(\tilde{\Delta} + 3\gamma - \bar{\gamma} + 4\mu + \bar{\mu})(\bar{\delta} - 2\bar{\tau} + 2\alpha)T_{n\bar{m}} \\
& +(\bar{\delta} - \bar{\tau} + \bar{\beta} + 3\alpha + 4\pi)(\bar{\delta} - \bar{\tau} + 2\bar{\beta} + 2\alpha)T_{nn} \\
& -(\bar{\delta} - \bar{\tau} + \bar{\beta} + 4\pi)(\tilde{\Delta} + 2\gamma + 2\bar{\mu})T_{n\bar{m}} .
\end{aligned} \tag{2.38}$$

56
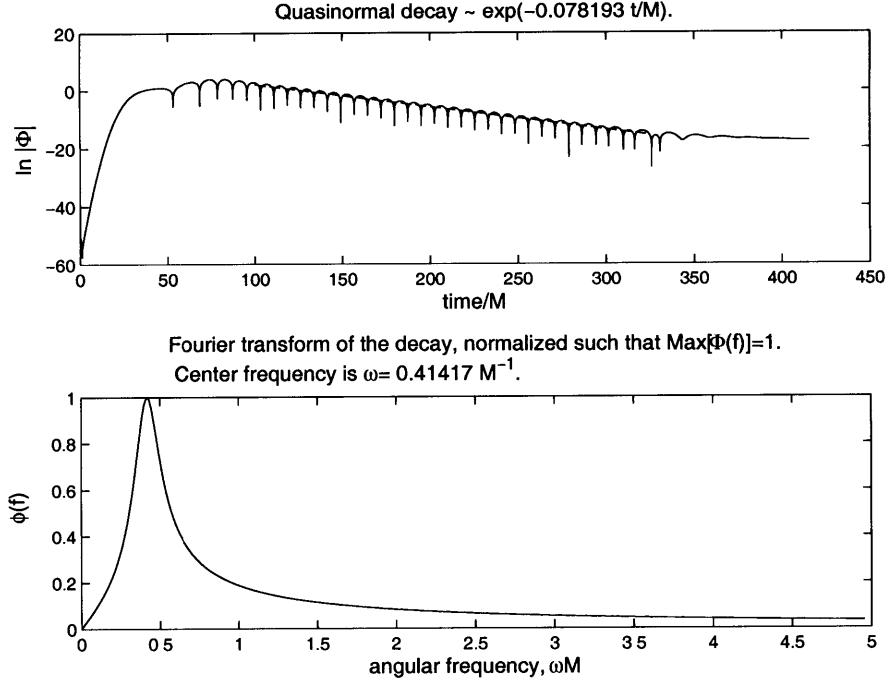
Figure 2-1: Illustrations of quasi-normal ringing for a black hole with $a/M = 0.9$; the $l = 2$, $m = 0$ mode is shown here. Top panel: Evolution of the magnitude of the Teukolsky function, extracted at $r = 20M$, $\theta = \pi/2$. We plot the time evolution of $\ln|\Phi|$ at this position. Overplotted on this curve (dashed line) is a function $\propto \exp(-0.078193t/M)$, demonstrating that we recover the expected decay law with a damping time $\tau = 12.789M$. Bottom panel: Magnitude of the Fourier transform of $\Phi(t)$. Notice that it peaks at $\omega = 0.41417/M$. These results for ω and τ are in excellent agreement with the expected values of $(\omega M, M/\tau) = (0.41, -0.078)$ from Ref. [61] for quasi-normal ringing of the $l = 2$, $m = 0$ mode for $a = 0.9$.

Reference [100] provides definitions for the various quantities which appear in Eqs. (2.38) and (2.38). Of particular importance are the quantities T_{nn}, $T_{n\bar{m}}$, and $T_{\bar{m}\bar{m}}$, given by contracting the stress energy tensor for the orbiting body with the Newman-Penrose null-tetrad legs n^μ and $\bar{m}^\mu$:

$$n^\mu \doteq \left(\frac{\rho^2 (r^2 + a^2)}{2}, -\frac{\Delta \rho^2}{2}, 0, \frac{a\rho^2}{2} \right) , \qquad (2.39)$$

$$\bar{m}^\mu \doteq \frac{1}{\sqrt{2}(r - ia\cos\theta)} \left(-ia\sin\theta, 0, 1, -i\csc\theta \right) , \qquad (2.40)$$

where $\Delta = r^2 - 2Mr + a^2$ and $\rho^{-2} = r^2 + a^2 \cos^2 \theta$.

Also of great importance here are the Newman-Penrose operators $\tilde{\Delta}$ and $\bar{\delta}$:

$$
\begin{aligned}
\tilde{\Delta} &= n^\mu \frac{d}{dx^\mu} \\
&= \frac{\rho^2 \left(r^2 + a^2\right)}{2} \frac{d}{dt} - \frac{\rho^2 \Delta}{2} \frac{d}{dr} + \frac{aim\rho^2}{2} \\
&\equiv \Delta_t + \Delta_r + \Delta_\phi \; ; \\
\bar{\delta} &= \bar{m}^\mu \frac{d}{dx^\mu} \\
&= -\frac{ia \sin \theta \rho^2 (r + ia \cos \theta)}{\sqrt{2}} \frac{d}{dt} \\
&\quad + \frac{(r + ia \cos \theta)\rho^2}{\sqrt{2}} \frac{d}{d\theta} + \frac{m\rho^2 (r + ia \cos \theta)}{\sqrt{2} \sin \theta} \\
&\equiv \bar{\delta}_t + \bar{\delta}_\theta + \bar{\delta}_\phi \; .
\end{aligned}
$$

(2.41)

(2.42)

(The operator $\tilde{\Delta}$ is normally written without the tilde; we have added it here to avoid confusion with $\Delta = r^2 - 2Mr + a^2$.)

To proceed, we next must analyze the stress energy tensor describing the small body. A point body of mass μ disturbing the Kerr spacetime is given by

$$
T_{\mu\nu} = \frac{\mu u_\mu u_\nu \rho^2}{u^t \sin \theta} \delta[r - R(t)] \delta[\theta - \Theta(t)] \delta[\phi - \Phi(t)] \; ,
$$

(2.43)

where $u^\mu = dx^\mu/d\tau$, and where $[R(t), \Theta(t), \Phi(t)]$ describe the Boyer-Lindquist coordinate worldline of the small body. Due to axial symmetry of the Kerr spacetime, the only ϕ dependence in the stress-energy tensor comes from $\delta[\phi - \Phi(t)]$. Writing

$$
T_{\mu\nu} = \sum_m T_{\mu\nu}^m e^{im\phi}
$$

(2.44)

and using the fact that

$$
\delta[\phi - \Phi(t)] = \frac{1}{2\pi} \sum_m e^{im[\phi - \Phi(t)]} \; ,
$$

(2.45)

58

we find

$$T_{\mu\nu}^m = \frac{\mu u_\mu u_\nu \rho^2}{2\pi u^t \sin\theta} \delta[r - R(t)]\delta[\theta - \Theta(t)]e^{-im\Phi(t)} . \tag{2.46}$$

Using this expansion for $T_{\mu\nu}$, it is a simple matter to construct $T_{nn} = n^\mu n^\nu T_{\mu\nu}$, $T_{n\bar{m}} = n^\mu \bar{m}^\nu T_{\mu\nu}$, and $T_{\bar{m}\bar{m}} = \bar{m}^\mu \bar{m}^\nu T_{\mu\nu}$. We then insert these terms into Eqs. (2.38). Using the chain rule repeatedly leaves us with a (rather complicated) expression involving radial and θ derivatives of Dirac delta functions. We thus face the task of representing the delta function and its derivatives accurately on a numerical grid. This is the major innovation of this paper, and is discussed in detail in the following section.

It should be noted at this point that, since the Teukolsky equation is most naturally written in terms of the tortoise coordinate r^*, we must describe the radial behavior of the source term in r^* as well. To this end, we replace the radial delta function and all radial derivatives as follows:

$$\delta[r - R(t)] = \frac{\delta[r^* - R^*(t)]}{|dr/dr^*|} , \tag{2.47}$$

$$\frac{d}{dr} = \frac{r^2 + a^2}{\Delta} \frac{d}{dr^*} . \tag{2.48}$$

Finally, in our numerical implementation, we define the vector $\boldsymbol{T}$ appearing in Eq. (2.7) as

$$\boldsymbol{T} \doteq \left[0, 0, \mathrm{Re}(\hat{T}), \mathrm{Im}(\hat{T})\right] , \tag{2.49}$$

where

$$\hat{T} = \frac{4\pi\Delta\left(r^2 + a^2\cos^2\theta\right)}{r^3\left[(r^2 + a^2)^2 - a^2\Delta\sin^2\theta\right]} \times \tag{2.50}$$

$$\exp\left(-im\frac{a}{r_+ - r_-}\ln\left[\frac{r - r_+}{r - r_-}\right]\right)\left[2\rho^{-4}T_4\right] .$$

The exponential factor in this expression corrects for the fact that the evolution code uses the azimuthal variable $\tilde{\phi}$, but the source term is expanded in ϕ.

59

2.3 The discrete delta function and its derivatives

As pointed out in the previous section, the Dirac delta function enters the Teukolsky equation because we approximate the perturbing mass by a point particle. By its definition as an integrable singularity, the delta function is very difficult to represent on a finite difference grid. The best we can hope to do is to develop a model function that captures its most important features, particularly localization to a very small spatial region, as well as integrability and derivative properties. The following three subsections describe the model for the delta function we have developed. We first describe a very basic model that demonstrates how to satisfy our criteria in Sec. 2.3.1. In Sec. 2.3.2 and 2.3.3, we then refine this basic model. These refinements have been found to improve the overall accuracy of the code.

The discrete delta function approach we use is inspired by the work presented in Ref. [37]. Our technique can be considered an extension of that used in [37]; in particular, they do not develop delta function derivatives, nor do they implement all the refinements discussed in 2.3.2 and 2.3.3. Nonetheless, Ref. [37] played an extremely important role in developing the foundations of our work.

Before turning to a detailed discussion of our techniques for modeling the delta function on a numerical grid, we first mention some general considerations pertaining to delta functions on a finite difference grid. For concreteness, consider a function and delta combination, $f(x)\delta(x - \alpha)$. The function $f(x)$ is taken to be known, and can be calculated for any x. For the sake of argument, let the delta be modeled by two coefficients δ_k and δ_{k+1} on grid points x_k and x_{k+1} respectively; the delta is taken to be zero everywhere else. (This in fact pertains to the form of the delta discussed in Sec. 2.3.1.)

Now imagine integrating $f(x)\delta(x - \alpha)$ over all x. Analytically, we know that this should give us $f(\alpha)$. Numerically integrating this on our grid gives us

$$h \sum_i f(x_i)\delta(x_i - \alpha) = h[f(x_k)\delta_k$$
$$+ f(x_{k+1})\delta_{k+1}]. \tag{2.51}$$

This equation suggests that the numerical integral approximates $f(\alpha)$ by interpolating between grid points x_k and x_{k+1}. If $f(x)$ is rapidly varying, this interpolation may not be accurate enough; this is sure to be a source of error as we integrate our PDE forward in time.

Great improvement can be achieved by enforcing the well-known identity

$$f(x)\delta(x - \alpha) = f(\alpha)\delta(x - \alpha). \tag{2.52}$$

The numerical integral now becomes

$$h \sum_i f(\alpha)\,\delta(x_i - \alpha) = h\left[f(\alpha)\delta_k + f(\alpha)\delta_{k+1}\right]$$
$$= f(\alpha)h\left(\delta_k + \delta_{k+1}\right)$$
$$= f(\alpha)\,, \tag{2.53}$$

and the identity is preserved exactly. In the last step, we use the discrete analog of the property

$$\int dx\,\delta(x - \alpha) = 1\,. \tag{2.54}$$

Similar identities can be used on the delta function derivatives:

$$f(x)\delta'(x - \alpha) = f(\alpha)\delta'(x - \alpha)$$
$$- f'(\alpha)\delta(x - \alpha)\,, \tag{2.55}$$
$$f(x)\delta''(x - \alpha) = f(\alpha)\delta''(x - \alpha) - 2f'(\alpha)\delta'(x - \alpha)$$
$$+ f''(\alpha)\delta(x - \alpha)\,. \tag{2.56}$$

We recommend using these identities as much as possible when numerically implementing the algorithms sketched in the following three subsections.

2.3.1   A simple numerical delta function

Consider the function $\delta(x - \alpha)$, where $x_k \leq \alpha \leq x_{k+1}$; i.e, α lies between two discrete grid points. Let $h = x_{k+1} - x_k = x_k - x_{k-1}$ be the grid resolution. We use the following integral to define the delta function:

$$\int_{\alpha-\epsilon}^{\alpha+\epsilon} dx\ f(x)\ \delta(x - \alpha) \;=\; f(\alpha)\,, \tag{2.57}$$

where $\epsilon > 0$ and $f(x)$ is any well behaved function. This means that $\delta(x - \alpha)$ is zero everywhere, except at $x = \alpha$, where it is singular. Translating this integral to a summation, we have:

$$\int_{\alpha-\epsilon}^{\alpha+\epsilon} dx\ f(x)\ \delta(x - \alpha) \;\simeq\; h \sum_{i} f(x_i)\,\delta_i \tag{2.58}$$

$$\Rightarrow f(\alpha) \;\simeq\; h \sum_{i} f(x_i)\,\delta_i\,, \tag{2.59}$$

where δ_i is the discrete delta defined on the grid. Since α does not necessarily lie on a gridpoint, we can linearly interpolate to find:

$$f(\alpha) \;=\; \frac{f(x_{k+1}) - f(x_k)}{h}(\alpha - x_k) + f(x_k)$$
$$+ O(h^2)\,. \tag{2.60}$$

Substituting this back into our earlier expression and comparing coefficients, we have

$$
\begin{aligned}
\delta_i &= \frac{\alpha - x_k}{h^2} \text{ for } i = k+1 \\
&= \frac{x_{k+1} - \alpha}{h^2} \text{ for } i = k \\
&= 0 \text{ everywhere else .}
\end{aligned}
\tag{2.61}
$$

Notice that if $\alpha = x_k$, then $\delta_i = 1/h$ for $i = k$, but is zero everywhere else; a similar result holds if $\alpha = x_{k+1}$. This reproduces our intuitive notion that the delta function is zero everywhere except at a single point, and that it integrates to unity. We take the viewpoint that the integrability of the delta function is its key defining property, using this rule to derive the results presented below. This is the approach that was used in Ref. [37].

Another approach to defining a numerical delta function, suggested in [59], is to first define a step function on the grid, and then use finite differencing to obtain the delta and its derivatives. The approach described above matches this proposal when α lies exactly on a grid point.

We can proceed in a similar fashion to find formulae for the derivatives. Let us define

$$
\gamma = \frac{x_{k+1} - \alpha}{h} = 1 - \frac{\alpha - x_k}{h} .
\tag{2.62}
$$

Again, we start from the defining integrals,

$$
\int dx \, f(x) \, \delta'(x - \alpha) = -f'(\alpha) ,
\tag{2.63}
$$

$$
\int dx \, f(x) \, \delta''(x - \alpha) = f''(\alpha) .
\tag{2.64}
$$

Note that a prime denotes d/dx. Our goal is to derive a form which enforces these

integrals in summation form:

$$h \sum_i f(x_i)\, \delta_i' \simeq -f'(\alpha)$$

$$= -h \sum_i f'(x_i)\delta_i$$

$$= -\gamma f'(x_k)$$

$$-(1-\gamma)f'(x_{k+1}) + O(h^2); \qquad (2.65)$$

$$h \sum_i f(x_i)\, \delta_i'' \simeq f''(\alpha)$$

$$= h \sum_i f''(x_i)\delta_i$$

$$= \gamma f''(x_k) +$$

$$(1-\gamma)f''(x_{k+1}) + O\left(h^2\right). \qquad (2.66)$$

The derivatives of $f(x_k)$ are given by the finite difference formulae,

$$f'(x_k) = \frac{f(x_{k+1}) - f(x_{k-1})}{2h} + O(h^2), \qquad (2.67)$$

$$f''(x_k) = \frac{f(x_{k+1}) - 2f(x_k) + f(x_{k-1})}{h^2}$$

$$+O(h^2). \qquad (2.68)$$

Substitution of these approximations in (2.65) and (2.66) and a comparison of coefficients yields for the derivative:

$$\delta_i' = \frac{\gamma}{2h^2} \text{ for } i = k - 1,$$

$$= \frac{1-\gamma}{2h^2} \text{ for } i = k,$$

$$= -\frac{\gamma}{2h^2} \text{ for } i = k+1,$$

$$= \frac{\gamma-1}{2h^2} \text{ for } i = k+2,$$

$$= 0 \text{ everywhere else }. \qquad (2.69)$$

64

For the second derivative:

$$\begin{aligned}
\delta_i'' &= \frac{\gamma}{h^3} \text{ for } i = k - 1 \ , \\
&= \frac{1 - 3\gamma}{h^3} \text{ for } i = k \ , \\
&= \frac{3\gamma - 2}{h^3} \text{ for } i = k + 1 \ , \\
&= \frac{1 - \gamma}{h^3} \text{ for } i = k + 2 \ , \\
&= 0 \text{ everywhere else } .
\end{aligned} \tag{2.70}$$

Notice that we need four points to represent the derivatives of the delta function in this scheme.

2.3.2   A multiple point delta function

The procedure described in Sec. 2.3.1 can be extended to represent the delta function over a larger number of points. On the one hand, this spreads out the delta, moving us away from our ideal of a function that is non-zero in as small a region as possible; on the other hand, it allows us to represent it more smoothly on our grid. The number of points $(2n + 2)$ that we use can thus be considered an optimization parameter, allowing us to trade localization for smoothness. As we shall see, there is typically a value of n that represents a very good compromise.

We start off with the 'linear hat' delta function defined in [37] and [102]

$$\delta_i = \begin{cases} \gamma_i/h & \text{for } |x_i - \alpha| \leq \epsilon = nh \\ 0 & \text{otherwise} \end{cases} \ , \tag{2.71}$$

where

$$\gamma_i = \frac{1}{n}\left(1 - \frac{|x_i - \alpha|}{nh}\right) \ , \tag{2.72}$$

and n is an integer. Note that when $n = 1$, γ_i reduces to the γ that was defined in Sec. 2.3.1. Note also that γ_i is non-zero only for $i \in [k, ..., k + 2n - 1]$(so that there

65

are a total of $2n$ points), and that α lies between the grid points x_{k+n-1} and x_{k+n}. In this labeling scheme, x_k is the smallest gridpoint where δ_i is nonzero.

Substituting this form of the delta function into our defining integral relation,

$$h \sum_i f(x_i) \delta_i \simeq f(\alpha) \,, \tag{2.73}$$

we find

$$\sum_i f(x_i) \gamma_i \simeq f(\alpha) \,, \tag{2.74}$$

The quantity γ_i is thus a weighting factor whose weight depends on the distance of x_i from α. Setting $f(x) = 1$, we find the property

$$\sum_i \gamma_i = 1 \,. \tag{2.75}$$

Now consider the derivative of the delta function. Our goal is again to enforce the rule

$$h \sum_i \delta_i' f(x_i) \simeq -f'(\alpha) \simeq -h \sum_i \delta_i f'(x_i) \,. \tag{2.76}$$

Inserting the finite difference formulae for the derivatives of f, Eqs. (2.67) and (2.68), into this relation, we find

$$
\begin{aligned}
f'(\alpha) \;\simeq\; & -\sum_{i=k}^{k+2n-1} \gamma_i \left[\frac{f(x_{i+1}) - f(x_{i-1})}{2h} \right] \\
= \; & -\frac{1}{2h} \left[\left(\sum_{i=k+1}^{k+2n-2} \gamma_{i-1} f(x_i) \right) \right] \\
& + \frac{1}{2h} \left[\gamma_{k+2n-2} f(x_{k+2n-1}) \right] \\
& + \frac{1}{2h} \left[\gamma_{k+2n-1} f(x_{k+2n}) \right] \\
& - \frac{1}{2h} \left[\left(\sum_{i=k+1}^{k+2n-2} \gamma_{i+1} f(x_i) \right) \right] \\
& + \frac{1}{2h} \left[-\gamma_k f(x_{k-1}) - \gamma_{k+1} f(x_k) \right] \,.
\end{aligned}
\tag{2.77}
$$

66

Comparing coefficients, we read off

$$\delta'_{k-1} = \frac{\gamma_k}{2h^2},$$ (2.78)

$$\delta'_k = \frac{\gamma_{k+1}}{2h^2},$$ (2.79)

$$\delta'_{k+j} = -\frac{\gamma_{k+j-1} - \gamma_{k+j+1}}{2h^2}$$

$$\text{for } j \in [1, 2n-2],$$ (2.80)

$$\delta'_{k+2n-1} = -\frac{\gamma_{k+2n-2}}{2h^2},$$ (2.81)

$$\delta'_{k+2n} = -\frac{\gamma_{k+2n-1}}{2h^2}.$$ (2.82)

The formulas for the delta derivative coefficients can be understood intuitively. The n-point generalization approximates the delta function as an isosceles triangle centered at α and sampled at $2n$ points. The derivative is simply the slope of this isosceles triangle at all points, except at the center and the edges, where the derivative is discontinuous. The discontinuity is replaced by coefficients that ensure the integral properties of the derivative. The delta derivative takes a particularly simple formula in the "bulk":

$$\begin{aligned}
\delta'_{k+j} &= -\frac{\gamma_{k+j-1} - \gamma_{k+j+1}}{2h^2} \text{ for } j \in [1, 2n-2] \\
&= \frac{1}{2nh^2} \left[\frac{|x_{k+j-1} - \alpha|}{nh} - \frac{|x_{k+j+1} - \alpha|}{nh} \right] \\
&= \begin{cases} \frac{1}{2n^2h^2} & \text{for } x_{k+j+1} < \alpha - h \\ -\frac{1}{2n^2h^2} & \text{for } x_{k+j+1} > \alpha + h \end{cases}
\end{aligned}$$ (2.83)

Notice that the delta derivative coefficients are non-zero for $i \in [k-1, k+2n]$ — one point wider in each direction than the span of the delta on the grid.

67

A similar analysis can be done for the second derivatives. We start off with

$$f''(\alpha) \simeq \sum_{i=k}^{k+2n-1} \gamma_i f''(x) , \tag{2.84}$$

$$h \sum_i \delta_i'' f(x_i) = \sum_{i=k}^{k+2n-1} \gamma_i \left[\frac{f(x_{i+1}) - 2f(x_i) + f(x_{i-1})}{h^2} \right] . \tag{2.85}$$

Reading off the coefficients leaves us with

$$\delta_{k-1}'' = \frac{\gamma_k}{h^3} , \tag{2.86}$$

$$\delta_k'' = \frac{\gamma_{k+1} - 2\gamma_k}{h^3} , \tag{2.87}$$

$$\delta_{k+j}'' = \frac{\gamma_{k+j+1} - 2\gamma_{k+j} + \gamma_{k+j-1}}{h^3} = 0$$

$$\text{for } j \in [1, 2n-2] , \tag{2.88}$$

$$\delta_{k+2n-1}'' = \frac{\gamma_{k+2n-2} - 2\gamma_{k+2n-1}}{h^3} , \tag{2.89}$$

$$\delta_{k+2n}'' = \frac{\gamma_{k+2n-1}}{h^3} . \tag{2.90}$$

Notice that the second derivative is zero in the "bulk" — it corresponds to the second derivative of a line, with constant slope. Like the first derivative, these coefficients are non-zero for $i \in [k-1, k+2n]$ — two points broader than the delta itself.

We have found that a very sharp delta function, like the two-point model described in the previous section, leads to instabilities for orbits with varying r or θ (e.g., for eccentric orbits). Using a smoother n-point representation suppresses these instabilities; this will be discussed in greater detail in chapter 3. Since r and θ are constant for circular, equatorial orbits, these instabilities do not arise for the cases examined in detail here. Thus for the case at hand, our numerical errors originating from the finite representation of the delta are smallest when $n = 1$. This is demonstrated in detail in Sec. 2.4.

2.3.3 Higher order interpolation for smoothness

Finally, we present a representation of the delta which uses a higher order interpolation scheme. This again spreads the "stencil" of the delta function over a wider patch of the grid, but improves our ability to reproduce the integral formulation of the delta identities.

Using cubic interpolation (which requires a total of four points), we find the rule

$$h \sum_i f(x_i)\delta_i \;=\; f(\alpha) \tag{2.91}$$

$$\Rightarrow h \sum_i f(x_i)\delta_i \;=\; -\frac{(\alpha - x_{k+1})(\alpha - x_{k+2})(\alpha - x_{k+3})}{6h^3} f(x_k)$$
$$+\frac{(\alpha - x_k)(\alpha - x_{k+2})(\alpha - x_{k+3})}{2h^3} f(x_{k+1})$$
$$-\frac{(\alpha - x_k)(\alpha - x_{k+1})(\alpha - x_{k+3})}{2h^3} f(x_{k+2})$$
$$+\frac{(\alpha - x_k)(\alpha - x_{k+1})(\alpha - x_{k+2})}{6h^3} f(x_{k+3}) \,. \tag{2.92}$$

The location α lies between grid points x_{k+1} and x_{k+2}. From this expression, we read off the coefficients

$$\delta_i \;=\; -\frac{(\alpha - x_{k+1})(\alpha - x_{k+2})(\alpha - x_{k+3})}{6h^4} \text{ at } x_k \,, \tag{2.93}$$

$$=\; +\frac{(\alpha - x_k)(\alpha - x_{k+2})(\alpha - x_{k+3})}{2h^4} \text{ at } x_{k+1} \,, \tag{2.94}$$

$$=\; -\frac{(\alpha - x_k)(\alpha - x_{k+1})(\alpha - x_{k+3})}{2h^4} \text{ at } x_{k+2} \,, \tag{2.95}$$

$$=\; +\frac{(\alpha - x_k)(\alpha - x_{k+1})(\alpha - x_{k+2})}{6h^4} \text{ at } x_{k+3} \,. \tag{2.96}$$

69

A similar analysis for the first derivatives yields

$$\delta_i' = \frac{(x_{k+1} - \alpha)(\alpha - x_{k+2})(\alpha - x_{k+3})}{12h^5} \text{ at } x_{k-1} , \qquad (2.97)$$

$$= \frac{(\alpha - x_k)(\alpha - x_{k+2})(\alpha - x_{k+3})}{4h^5} \text{ at } x_k , \qquad (2.98)$$

$$= -\frac{(\alpha - x_{k+1})(2\alpha - 3x_k + x_{k+2})(\alpha - x_{k+3})}{12h^5} \text{ at } x_{k+1} , \qquad (2.99)$$

$$= -\frac{(\alpha - x_k)(\alpha - x_{k+2})(2\alpha + x_{k+1} - 3x_{k+3})}{12h^5} \text{ at } x_{k+2} , \qquad (2.100)$$

$$= \frac{(\alpha - x_k)(\alpha - x_{k+1})(\alpha - x_{k+3})}{4h^5} \text{ at } x_{k+3} , \qquad (2.101)$$

$$= -\frac{(\alpha - x_k)(\alpha - x_{k+1})(\alpha - x_{k+2})}{12h^5} \text{ at } x_{k+4} . \qquad (2.102)$$

Note, we have used the second order finite difference formula, Eq. (2.67), to derive this result. In principle, higher order formulas for the derivative could have been used. We kept to the second order formula in order to keep the derivative stencil narrow, and also for consistency with our time-stepping algorithm.

Finally, for the second derivatives we find

$$\delta_i'' = \frac{(x_{k+1} - \alpha)(\alpha - x_{k+2})(\alpha - x_{k+3})}{6h^6}$$

$$\text{at } x_{k-1} , \qquad (2.103)$$

$$= \frac{(5\alpha - 3x_k - 2x_{k+1})(\alpha - x_{k+2})(\alpha - x_{k+3})}{6h^6}$$

$$\text{at } x_k , \qquad (2.104)$$

$$= -\frac{\left(10\alpha^2 - (9x_k + 4x_{k+1} + 7x_{k+2})\alpha + x_{k+1}x_{k+2} + 3x_k(x_{k+1} + 2x_{k+2})\right)(\alpha - x_{k+3})}{6h^6}$$

$$\text{at } x_{k+1} , \qquad (2.105)$$

$$= \frac{(\alpha - x_k)\left(10\alpha^2 - (7x_{k+1} + 4x_{k+2} + 9x_{k+3})\alpha + 3x_{k+2}x_{k+3} + x_{k+1}(x_{k+2} + 6x_{k+3})\right)}{6h^6}$$

$$\text{at } x_{k+2} , \qquad (2.106)$$

$$= -\frac{(\alpha - x_k)(\alpha - x_{k+1})(5\alpha - 2x_{k+2} - 3x_{k+3})}{6h^6}$$

$$\text{at } x_{k+3} , \qquad (2.107)$$

$$= \frac{(\alpha - x_k)(\alpha - x_{k+1})(\alpha - x_{k+2})}{6h^6}$$

$$\text{at } x_{k+4} . \qquad (2.108)$$

As discussed in more detail in the following section, our analysis suggests that this cubic interpolation method works best.

We emphasize at this point that, although we are motivated by Teukolsky equation applications, our discussion here was not specialized to the Teukolsky equation in any way. The delta models sketched here can be used in any finite-difference numerical algorithm. We also note that one does not need to stop at cubic-order interpolation; the basic idea of that scheme could easily be extended to higher order if the application warranted it. As the order is increased, the "stencil" of the delta is likewise increased, pushing us away from the intuitive notion of a structureless impulse. This leads us to believe that there may be a certain interpolation order beyond which the model ceases to work well.

2.3.4 Convergence with the discrete delta function

The non-smooth nature of the discrete delta function makes understanding the convergence properties of a code based on this function somewhat subtle. Here we briefly summarize some key issues related to convergence with the discrete delta. This summary is based on detailed discussion of discretization errors given in Ref. [102]. The punchline of this discussion is that the discrete delta function is typically at least second-order convergent, and thus we expect our code to likewise be second-order convergent.

Let δ_i be the discretized version of $\delta(x - \alpha)$ defined on a discrete grid x_i, let $h = x_{i+1} - x_i$ be the grid spacing, and let δ_i be non-zero at $x_k, x_{k+1}, \ldots, x_{k+2n-1}$. The continuous variable α lies between x_{k+n-1} and x_{k+n}. This allows us to define a parameter η such that

$$\alpha = x_{k+n-1} + \eta h. \tag{2.109}$$

This quantity is a measure of how close α is to a grid point; clearly, $0 \leq \eta \leq 1$.

We now define the moments of the discrete delta by

$$M_r(\delta, \alpha, h) = h \sum_{i=k}^{k+2n-1} \delta_i (x_i - \alpha)^r , \tag{2.110}$$

71

where r is an integer. In the continuum limit, this definition becomes

$$
\begin{aligned}
M_r &\to \int \delta(x - \alpha)\,(x - \alpha)^r\;dx\;, \\
&= 1 \quad r = 0 \\
&= 0 \quad r > 0\;.
\end{aligned}
\tag{2.111}
$$

A discrete representation will clearly have the correct zeroth moment; however, it will only have $M_{r>0} = 0$ up to some maximum r. Reference [102] proves that, if q is the lowest integer such that $M_q \neq 0$, then

$$
\left| f(\alpha) - h \sum \delta_i f(x_i) \right| \le C h^q \;,
\tag{2.112}
$$

where C is approximately[3] a constant. This delta representation is then qth-order convergent.

For the multiple point discrete delta discussed in Sec. 2.3.2, we find $M_0 = 1$, $M_1 = 0$, $M_2 \neq 0$. When we use this discrete delta, we therefore expect our code to be second-order convergent. We demonstrate this behavior in Sec. 2.4.2. For the cubic delta function, we find $M_0 = 1$, $M_{1,2,3} = 0$, $M_4 \neq 0$. In this case, errors due to the delta representation are expected to be fourth order. However, since our stepping algorithm is itself second-order, we expect second-order convergence overall.

2.4 Numerical Implementation and Evaluation of the Discrete Delta Function

We now implement the Teukolsky equation's source term using the techniques discussed in Sec. 2.3 for the simple case of a point particle in a circular, equatorial orbit around a massive black hole. Our goal is to compare the different forms of the discrete delta discussed in the previous section and to evaluate which is likely to work best for practical modeling of radiation from astrophysical systems. We also compare our

---

[3]In our application, C varies slightly depending on how close the delta peak is to a grid point.

discrete delta model to the Gaussian approximation that has been used in previous work, illustrating the power of this new model.

We obviously require some "standard" to compare our results against. Frequency-domain codes provide extremely accurate results for circular, equatorial orbits, largely since their emitted radiation is concentrated in a small number of multipoles; as such, they make an excellent standard against which to compare our results. We use the code described in [53] as our standard.

Tables 2.1 and 2.2 shows energy fluxes obtained from our code for the most dominant azimuthal modes, $|m| = 2$ and 3 respectively. We compare these figures with those obtained from the code used in [53][4].

There are two major reasons that the fluxes we compute depart from those computed by frequency-domain codes. First, the time-domain code must extract fluxes at some finite radius. The FD approach produces, by construction, the waveforms and fluxes that would be measured infinitely far from the generating binary; this simply cannot be done on a finite radial grid. A detailed discussion of the impact of finite extraction radius is given in Sec. 2.5. In brief, we find by varying the extraction radius that fluxes can be fit to a very simple power law. This power law then allows us to infer the flux that would be measured by distant observers. The second source of error is simply numerical — finite difference errors plus the approximate nature of our discrete delta. Roughly speaking, accounting for finite extraction radius reduces our errors by about a factor of 2 − 5; the residual error is thus most likely simply numerical error. This is described in much greater detail in Sec. 2.5.

Figures 2-2 and 2-3 illustrate a typical example of the structure for the Teukolsky function Ψ that we find. We show the $m = 2$ mode of an orbit with radius $r_0 = 7.9456M$ around a Schwarzschild black hole; this orbit was selected in order to compare with results published in [94]. Note that the orbital period at this radius is $T = 2\pi\sqrt{r_0^3/M} \simeq 140M$. The data is read out at radius $R \equiv r_{\text{extract}} = 250M$, $\theta = \pi/2$; our numerical grid runs from $-100M \leq r^* \leq 500M$, with a resolution

---

[4]Symmetry in the azimuthal direction results in equal fluxes for $+m$ and $−m$ modes. Thus, $|m|$ refers to the sum of the fluxes from the $+m$ and $−m$ modes which is equal to twice the flux from either the $+m$ or the $−m$ mode.
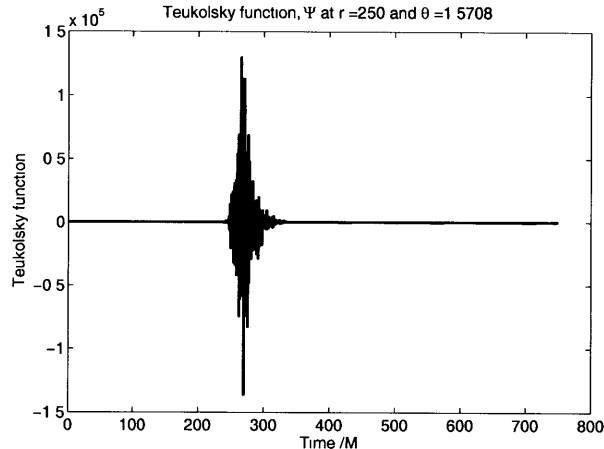
Figure 2-2: The real part of the $m = 2$ mode of the Teukolsky function Ψ as a function of time for a point particle of mass $\mu/M = 0.01$ in a circular orbit of radius $r_0/M = 7.9456$. These data were extracted in the equatorial plane $(\theta = \pi/2)$ at radius $R = 250M$. At this location the Teukolsky function is zero by construction until $t \simeq 250M$, at which point a spurious burst reaches the extraction radius. This burst is due to our unphysical initial conditions; it quickly propagates off the grid, leaving a reasonable physical solution for all time afterwords.

$\delta r^* = 0.0625M$, and from $0 \le \theta \le \pi$ with $\delta\theta = \pi/40$.

Figure 2-2 shows the real part of Ψ over a broad span of time, from the beginning of our simulation to $t \simeq 800M$. At $t \sim 250M$, a very high amplitude, unphysical burst of radiation reaches the extraction radius. This spurious burst is due to our initial conditions: We initially set $\Psi = 0$ and $\partial_t \Psi = 0$, which is not consistent with our source function. The time at which this burst reaches the extraction radius is perfectly consistent with radiation propagating at the speed of light ($c = 1$ in our units) across our numerical grid. The burst quickly propagates off the grid, and the solution for Ψ settles down to a simple oscillation. This is shown in Fig. 2-3, which zooms in on the behavior of Ψ for $t \ge 350M$. Notice that Ψ_R has an oscillation period of about $70M$, precisely what we expect for the $m = 2$ mode of a source whose orbital frequency is $140M$. The energy flux we find from this mode is $\dot{E}/\mu^2 = 1.708 \times 10^{-4}$, in excellent agreement with results published in Ref. [94] (compare Table II of [94], noting that our results require summing over all l for fixed m).

74

Teukolsky function, Ψ at r =250 and θ =1 5708.
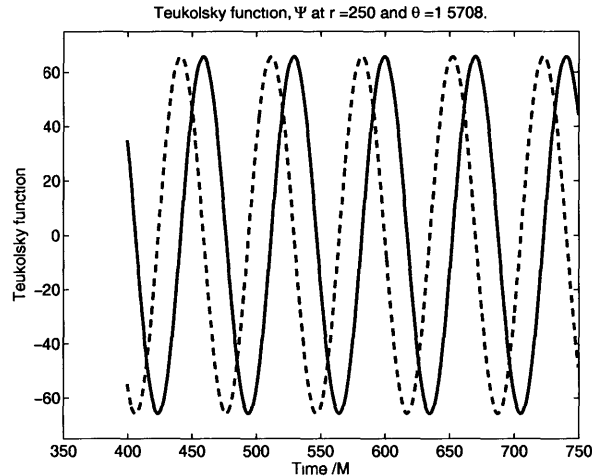
Figure 2-3: The same as Fig. 2-2, but zooming in on the data for $t \geq 350M$. Solid and dashed lines are the real and imaginary parts of Ψ respectively. The Teukolsky function oscillates with a period of about $70M$; since the source has a period of about $140M$, this is exactly what we expect for the $m = 2$ mode. We measure the total flux of energy carried by this mode to be $\dot{E}/\mu^2 = 1.708 \times 10^{-4}$, in agreement with previous results (see, e.g., Ref. [94]).

Figure 2-4 illustrates the spatial behavior of the real part of Ψ at a particular moment in time $(t = 312M)$. This plot illustrates the behavior of Re Ψ as a function r^* and θ over a wide span of our grid. Clearly visible are the $m = 2$ mode of the radiation propagating to large radius as well as the nearly singular delta function source itself.

2.4.1 Comparison of different discrete delta functions

In this section, we compare results from the various models for the delta function presented in Sec. 2.3. The variable point approximation for the linear delta function, presented in Sec. 2.3 provides us with a nice handle to study the convergence of our results. As we increase n, the half-width of the delta function, the singularity spreads out and its sharpness decreases. Notice that the physical spread of the source term is $2(n+1)\delta r^*$(due to the spread of the delta derivatives). Thus, decreasing the resolution has the same effect on the physical width as decreasing n.
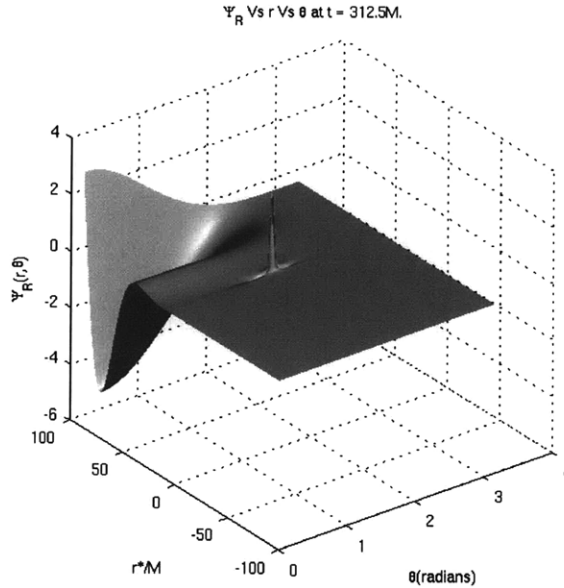
Figure 2-4: The same as Fig. 2-2, but now showing the data for a given moment in time ($t = 312.5M$) for a wide range of r^* and θ. Along with the outward propagating radiation packet visible at large radius, the nearly singular delta function source is clearly visible at the particle's position.

In Tables 2.3 and 2.4, we present results describing two equatorial circular orbits, one in the extremely strong field ($r_0/M = 2.32$, $a/M = 0.9$), and another at weaker field ($r_0/M = 12$, $a/M = 0$). We show the variation in flux with n, the half-width of the radial delta function. The angular delta function is represented using two points (i.e., $n_{\mathrm{ang}} = 1$), the minimum number of non-trivial points. The resolutions ($\delta r^*, \delta\theta$) are held fixed at $(0.0625M, \pi/40)$

The third and fourth columns of Tables 2.3 and 2.4 compare the flux in energy carried by GWs as computed using the δ-code to flux computed using our frequency-domain standard. The third column gives a "raw" comparison — we extract the time-domain fluxes at radius $R = 250M$ and compare to the frequency-domain result. In the fourth column, we extrapolate the time-domain data, $R \to \infty$, using the algorithm described in Sec. 2.5. The fourth column thus contains the most relevant data for assessing which delta representation is "best". We include the third column

76

to demonstrate that the difference before extrapolating is not terribly large, but that it is large enough that the gain due to this extrapolation is significant. It also illustrates that not performing the extrapolation can mislead regarding which form of the delta is most accurate.

Data from Table 2.4 indicate that, among the n-point representations, $n = 1$ gives the best results for circular, equatorial orbits. The cubic representation, however, is even better — the smoothness of this technique apparently reduces error even more. We choose the cubic delta for the remainder of our analysis because it is both smooth and accurate.

By contrast, the most accurate flux in Table 2.3 occurs when $n = 7$ (total of 8 points to represent the source), rather than $n = 1$. The reason is due to a competition between errors from finite differencing and errors from our delta representation. In particular, we have noticed experimentally that finite difference errors tend, on average, to spuriously decrease our measured flux; errors from spreading the delta over the grid tend to augment the flux. (We emphasize that this is merely a rule-of-thumb tendency we have noted; we also emphasize that we do not as of yet have a good explanation for these effects.)

As we approach the horizon, finite difference errors tend to become more important. This can be compensated by increasing the width of our delta representations. At $n = 7$, the spread of our source is just enough to accurately compensate for finite difference errors. At larger radius (e.g., the $r_0 = 12M$ orbit shown in Table 2.4), finite difference errors are so small that we do best using the minimum number of points possible in our model.

2.4.2   Convergence of our code

As discussed in Sec. 2.3.4, we generally expect a code built using the discrete delta on grid with spacing h to exhibit $O(h^2)$ convergence. In particular, we expect the Weyl scalar ψ_4 to show second-order convergence. We check this expectation by examining the flux of energy carried by gravitational waves. Since we expect $\psi_4 = \psi_4^{\text{true}} + O(h^2)$,

we likewise expect $\dot{E}$ to exhibit second-order convergence:

$$\dot{E} \sim |\psi_4|^2 \sim |\psi_4^{\text{true}}|^2 + O(h^2) \ . \tag{2.113}$$

To demonstrate this convergence, we show the energy flux measured at $R = 250M$ for two different strong-field[5] orbits: $r_0 = 5M$, $a = 0.8M$ and $r_0 = 4.64M$, $a = 0.9M$. The radial and angular grids are set to

$$\delta r^* = 0.0625 \times 2^{-b/4} \ , \tag{2.114}$$

$$\delta\theta = \pi/30 \times 2^{-b/4} \ . \tag{2.115}$$

Actually, $\delta\theta$ was modified slightly from this to insure that $\theta = \pi/2$ lies exactly on a grid point. This reduced variations about the main h^2 trend owing to the slight dependence of the proportionality "constant" on the delta's peak (cf. discussion in Sec. 2.3.4). Figure 2-5 shows the results our runs for $b \in [-1, \dots, 4]$. Convergence is shown by examining the fractional error with respect to our densest grid,

$$\text{error} \equiv \frac{|\dot{E}_b - \dot{E}_4|}{\dot{E}_4} \ , \tag{2.116}$$

where $\dot{E}_b$ is the flux inferred at grid parameter b. We normalize to $b = 4$ since it is the densest grid available to us. Modulo some slight oscillations, the overall trend of our data is in very good agreement with second-order convergence.

2.4.3 Comparison of discrete and Gaussian approximations for the numerical delta

The work in [68, 59, 23] approximates the delta-function by a narrow Gaussian such that it integrates to unity over the numerical grid. The Gaussian smears out the

---

[5]It's worth noting that we found it to be rather difficult to demonstrate convergence using weak-field orbits. For such orbits, the differences in our computed fluxes were quite small as we varied our grid density. We need strong-field orbits in order for the errors to be large enough that the convergence trend is apparent.
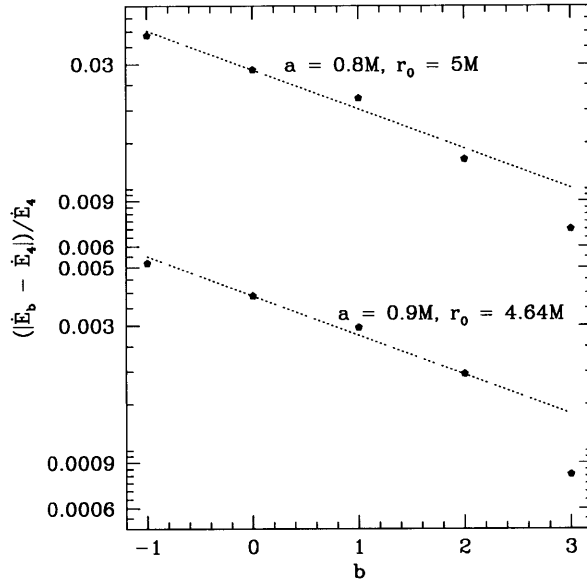
Figure 2-5: An illustration of our code's convergence behavior. We show the fractional deviation in energy flux in the $|m| = 2$ mode, measured at $R = 250M$ as a function of grid spacing. The grid is controlled by the integer b using $\delta r^* = 0.0625 \times 2^{-b/4}$, $\delta\theta = \pi/30 \times 2^{-b/4}$, with $b \in [-1, \ldots, 4]$. The upper data set is for fluxes measured from an orbit with $a = 0.8M$, $r_0 = 5M$; the lower set is for $a = 0.9M$, $r_0 = 4.64M$. For each data set, the dotted line represents what we would expect for perfect second-order convergence (fit arbitrarily to the data for $b = 0$); the large dots represent our actual convergence data.

singularity and thus the source term is non zero at a large number of points on the numerical grid. In contrast, the models presented in Sec. 2.3 use only a few points to depict the delta-function and thus do not share this disadvantage. A comparison on the same hardware and software platform showed that the techniques used here (the δ-code) are about *twelve* times faster than the ones that use a smeared Gaussian (the G-code). The last two columns in Tables 2.1 and 2.2 show the fluxes from the Gaussian-approximated delta function. Note that the errors in these fluxes are about $2 - 3\%$, quite a bit higher than errors from the δ-code. Both codes were run with identical parameters and grid resolutions. The accuracy of both codes improves with higher resolution, and improvement in both is seen by moving the extraction radius farther out. However, when these parameters are fixed, we find that the δ-code is

faster and demonstrates higher accuracy.

2.5 Accounting for finite extraction radius

When one discusses the gravitational-wave fluxes which a system generates, one is normally interested in their asymptotic value infinitely far away. It is of course not possible for a finite coordinate grid to reach all the way into this distant zone, so it is of great importance to understand how our fluxes vary with respect to our finite extraction radius R.

In flat spacetime, the extraction radius is not very important; it just needs to be sufficiently far away that the field it measures is purely radiative (i.e, not contaminated by near-field effects). Conservation laws guarantee that fluxes follow a $1/r^2$ law in this region, and so the integrated flux is independent of extraction radius.

Things are not so simple in a curved spacetime — radiation is effectively scattered off of spacetime curvature, modifying its propagation characteristics compared to flat spacetime intuition. This is responsible for the late time "tails" that are seen when a radiation packet propagates away from a black hole (cf. the late time behavior seen in Fig. 2-1). These tails can be regarded, heuristically, as radiation whose propagation to large radius was delayed by scattering off the spacetime. It also causes the integrated flux to depend on and vary with the radius at which it is measured.

We now examine how our fluxes vary with respect to extraction radius. Tables 2.5 and 2.6 present the fluxes measured for four representative strong-field orbits. In each case, we measure $\dot{E}$ for the $|m| = 2$ and $|m| = 3$ modes at extraction radii $R/M = 100, 200, 300, 400, 500$, and 600. These data are then fit to the ansatz

$$\dot{E} = \dot{E}_\infty \left[1 - q \left(m \Omega_{\text{orb}} R \right)^{-p} \right] . \tag{2.117}$$

The parameters q, p, and $\dot{E}_\infty$ are determined by the fit. Notice that $\dot{E}_\infty$ represents the flux that (according to this ansatz) would be measured infinitely far away. Note that this form was suggested to us by L. M. Burko [21], and replaces a previous

version which used $(r_{\mathrm{orb}}/R)^p$ rather than $(m\Omega_{\mathrm{orb}}R)^{-p}$. The two forms can be easily related to one another; however, the form involving $m\Omega_{\mathrm{orb}}$ emphasizes that it is the asymptotic behavior of the mode, rather than a property of the orbit, that sets $\dot{E}$. This form should also be more readily extendable to non-circular orbits.

Figure 5 shows an example of how well this fit works for one of the cases given in Table 2.5 ($r_0/M = 10$, $a/M = 0.99$, $m = 2$). Pragmatically, this ansatz appears to fit the data quite well; the quality shown in Fig. 2-6 is typical for the data that we examined. Interestingly, we find in all cases that the exponent $p \simeq 2$, independent of black hole spin, orbit radius, or mode number. Detailed calculations which we will present elsewhere [22] shows that this corresponds to the *dominant* correction to the asymptotic behavior of ψ_4. Such behavior was also demonstrated by Newman and Unti [73].
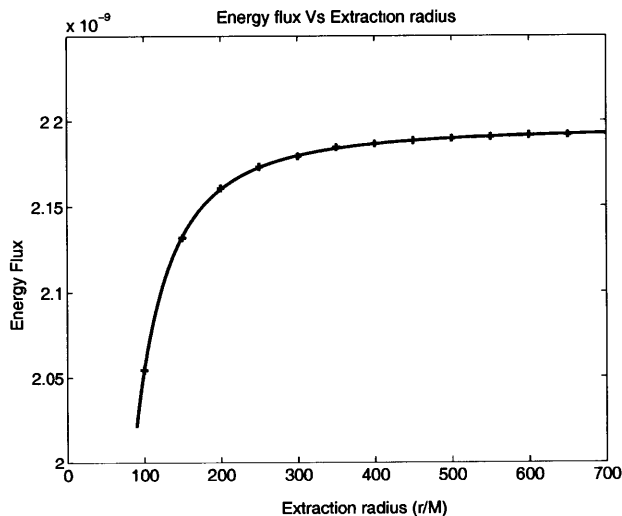
Figure 2-6: A power law fit to our numerically extracted energy fluxes for the case $r_0/M = 10$, $a/M = 0.99$, $m = +2$. Numerical data is indicated by the dots; the curve is the best fit we obtain for the ansatz given by Eq. (2.117). For this case, the best fit parameters are $q = 7.45$, $p = 2.06$, $\dot{E}_\infty = 2.197 \times 10^{-5}$.

We now use the fit (2.117) to compare the extrapolated measured flux $\dot{E}_\infty$ to frequency-domain results $\dot{E}_{\mathrm{FD}}$. This is shown in the last column of Table 2.6. In all cases, the error we find is less than 1%, sometimes substantially less. A similar fit can

be performed on the angular momentum flux, with similar results. We conclude that the fit (2.117) accounts for finite extraction radius, providing an accurate estimate for the fluxes that a particle radiates to infinity. Residual errors are thus much more likely to be true measures of numerical error in our calculation, and not an artifact of the extraction.

2.6 Summary and Future work

We have presented a simple, new technique for modeling the Dirac delta function and its derivatives on a finite difference grid. This technique requires that the source be modeled only on a handful of points on the grid. Our particular goal in this analysis is to model a pointlike source function for the time-domain Teukolsky equation, appropriate to describe the smaller member of an extreme mass ratio binary. We emphasize that our models for the discrete delta and its derivatives are more broadly applicable than just the Teukolsky equation — these techniques can be used in any context that requires modeling a sharp, delta-like function on a finite difference grid.

We test this approach by solving the Teukolsky equation for a test body in a circular, equatorial orbit of a Kerr black hole. Comparing with a well tested time-domain code that treats the orbiting body using a truncated Gaussian, we find that this new approach is extremely fast (often by a factor of ~ 10) and accurate. Using a frequency-domain code as a benchmark to compare the flux of energy carried by gravitational waves, we find that the code which uses the discrete delta function is typically a factor of $2-5$ more accurate than the Gaussian treatment most commonly used previously. This accuracy can be improved still further (at least for fluxes) by using a simple fit that accounts for the variation of the flux with the extraction radius. Combining our new source function with this fitting law, we find that our code agrees with the frequency-domain benchmark with errors smaller than 1% for a large fraction of parameter space, sometimes significantly smaller.

Since the goal of this analysis is to contribute to the modeling of EMRI gravitational-wave sources, the restriction to circular and equatorial orbits, though a useful, illustra-

tive test, is not astrophysically realistic. Since such binaries form through scattering processes, they are expected to have substantial eccentricity [42], and the secondary's orbit should have no special alignment with the spin axis of the large black hole. Chapter 3 will study how well this new technique handles such orbits. This realistic case is substantially more difficult to treat than circular, equatorial orbits, since the orbiting body (and our discrete delta model) very rapidly crosses back and forth over grid points in both the radial and latitudinal directions. The tables and figures in chapter 3 show that we get very good results even when we move our discrete delta model rapidly in a dynamical orbit. Chapters 3 and 5 also examine wave emission from inspiral (non-geodesic) sequences, including the merger and ringdown phases.

The final goal of this work will be to compute adiabatic inspiral waveforms using a hybrid of frequency-domain and time-domain, as described in the introduction. With a robust time-domain code for computing waves from nearly arbitrary physical worldlines and with a robust frequency-domain code capable of "mass producing" radiation reaction data for generic Kerr black hole orbits this problem should boil down to simple a matter of available CPU resources. Once we are in this state, we hope to produce waveforms efficiently enough that they can be used by workers looking at problems in *LISA* data analysis and waveform measurement (e.g., the "Mock *LISA* Data Challenge" [4, 3, 29, 2]). These waveforms are likely to be useful for other astrophysical problems, such as computing radiation recoil from both the slow inspiral and the dynamic plunge. Computing this effect in the extreme mass ratio limit may serve as a precision check on recent work looking at this problem in full numerical relativity [10, 50, 51, 63, 25].

2.7 Acknowledgments

Table 2.1: Energy flux extracted at $R \equiv r_{\text{extract}} = 250M$ for circular, equatorial orbits for the $m = |2|$ mode of a particle with mass $\mu/M = 1$. a/M is the BH spin, r_0/M is the orbital radius. The labels "δ" and "G" refer to the results from δ-code and G-code respectively. Values listed under "FD" are the corresponding fluxes from the frequency-domain code used in [53].

a/M	r_0/M	$\dot{E}_\delta$	$\dot{E}_{FD}$	$\frac{\dot{E}_\delta - \dot{E}_{FD}}{\dot{E}_{FD}}$	$\dot{E}_G$	$\frac{\dot{E}_G - \dot{E}_{FD}}{\dot{E}_{FD}}$
0	6	7.385×10^{-4}	7.368×10^{-4}	0.0023	7.246×10^{-4}	-0.017
0	8	1.650×10^{-4}	1.651×10^{-4}	-0.0055	1.623×10^{-4}	-0.016
0	10	5.344×10^{-5}	5.374×10^{-5}	-0.0004	5.290×10^{-5}	-0.016
0.5	6	5.551×10^{-4}	5.539×10^{-4}	0.0022	5.437×10^{-4}	-0.018
0.5	8	1.399×10^{-4}	1.401×10^{-4}	-0.0015	1.375×10^{-4}	-0.019
0.5	10	4.781×10^{-5}	4.812×10^{-5}	-0.0065	4.691×10^{-5}	-0.025
0.9	4	2.654×10^{-3}	2.662×10^{-3}	-0.0030	2.611×10^{-3}	-0.019
0.9	6	4.614×10^{-4}	4.621×10^{-4}	-0.0016	4.531×10^{-4}	-0.020
0.9	8	1.249×10^{-4}	1.254×10^{-4}	-0.0039	1.230×10^{-4}	-0.019
0.9	10	4.419×10^{-5}	4.456×10^{-5}	-0.0084	4.339×10^{-5}	-0.026
0.99	4	2.469×10^{-3}	2.484×10^{-3}	-0.0059	2.434×10^{-3}	-0.020
0.99	6	4.450×10^{-4}	4.461×10^{-4}	-0.0024	4.372×10^{-4}	-0.020
0.99	8	1.221×10^{-4}	1.226×10^{-4}	-0.0041	1.201×10^{-4}	-0.020
0.99	10	4.346×10^{-5}	4.386×10^{-5}	-0.0090	4.270×10^{-5}	-0.026

Table 2.2: Energy flux extracted at $250M$ for circular, equatorial orbits for the $m = |3|$ mode of a particle with mass $\mu/M = 1$. All symbols and notation are as in Table 2.1.

a/M	r_0/M	$\dot{E}_\delta$	$\dot{E}_{FD}$	$\frac{\dot{E}_\delta - \dot{E}_{FD}}{\dot{E}_{FD}}$	$\dot{E}_G$	$\frac{\dot{E}_G - \dot{E}_{FD}}{\dot{E}_{FD}}$
0	6	1.465×10^{-4}	1.460×10^{-4}	0.0035	1.431×10^{-4}	-0.020
0	8	2.445×10^{-5}	2.449×10^{-5}	-0.0017	2.399×10^{-5}	-0.020
0	10	6.383×10^{-6}	6.435×10^{-6}	-0.0080	6.291×10^{-6}	-0.022
0.5	6	1.015×10^{-4}	1.014×10^{-4}	0.0011	0.992×10^{-4}	-0.021
0.5	8	1.993×10^{-5}	1.980×10^{-5}	0.0066	1.935×10^{-5}	-0.023
0.5	10	5.521×10^{-6}	5.572×10^{-6}	-0.0090	5.410×10^{-6}	-0.029
0.9	4	6.485×10^{-4}	6.467×10^{-4}	0.0028	6.336×10^{-4}	-0.020
0.9	6	8.031×10^{-5}	8.043×10^{-5}	-0.0015	7.865×10^{-5}	-0.022
0.9	8	1.710×10^{-5}	1.717×10^{-5}	-0.0043	1.677×10^{-5}	-0.023
0.9	10	4.992×10^{-6}	5.044×10^{-6}	-0.0103	4.893×10^{-6}	-0.030
0.99	4	5.932×10^{-4}	5.924×10^{-4}	0.0014	5.805×10^{-4}	-0.021
0.99	6	7.688×10^{-5}	7.688×10^{-5}	-4.8×10^{-5}	7.528×10^{-5}	-0.022
0.99	8	1.6542×10^{-5}	1.669×10^{-5}	-0.0086	1.628×10^{-5}	-0.025
0.99	10	4.879×10^{-6}	4.942×10^{-6}	-0.0127	4.792×10^{-6}	-0.030

Table 2.3: Comparison of several implementations of the discrete delta function. We show results for the linear hat delta described in Sec. 2.3.2, as well as the cubic delta function described in Sec. 2.3.3. All fluxes are measured at $R = 250M$ for the $|m| = 2$ mode. For the linear hat delta, the total number of points in the function is $2(n+1)$. The cubic delta uses 6 points in all. These results are for orbits of radius $r_0 = 2.32M$ about a black hole with $a = 0.9M$. The total flux in $|m| = 2$ modes according to our frequency-domain standard is $\dot{E}_{\rm FD}/\mu^2 = 2.061 \times 10^{-2}$.

Total points, $2(n+1)$	$\dot{E}_{250}$	$(\dot{E}_{250} - \dot{E}_{FD})/\dot{E}_{FD}$	$\dot{E}_\infty$	$(\dot{E}_\infty - \dot{E}_{FD})/\dot{E}_{FD}$
64	2.889×10^{-2}	4.0×10^{-1}	2.890×10^{-2}	4.0×10^{-1}
32	2.194×10^{-2}	6.5×10^{-2}	2.195×10^{-2}	6.5×10^{-2}
16	2.055×10^{-2}	-2.8×10^{-3}	2.056×10^{-2}	-2.4×10^{-3}
8	2.027×10^{-2}	-1.6×10^{-2}	2.028×10^{-2}	-1.6×10^{-2}
4	2.023×10^{-2}	-1.9×10^{-2}	2.024×10^{-2}	-1.8×10^{-2}
cubic	2.024×10^{-2}	-1.8×10^{-2}	2.024×10^{-2}	-1.8×10^{-2}

Table 2.4: Same as Table 2.3, but now for an orbit with $r_0 = 12M$ about a black hole with $a = 0$. The frequency-domain flux for $|m| = 2$ modes in this case is $\dot{E}_{\rm FD}/\mu^2 = 2.172 \times 10^{-5}$.

Total points, $2(n+1)$	$\dot{E}_{250}$	$(\dot{E}_{250} - \dot{E}_{FD})/\dot{E}_{FD}$	$\dot{E}_\infty$	$(\dot{E}_\infty - \dot{E}_{FD})/\dot{E}_{FD}$
64	2.342×10^{-5}	7.8×10^{-2}	2.376×10^{-5}	9.4×10^{-1}
32	2.191×10^{-5}	8.7×10^{-3}	2.224×10^{-5}	2.4×10^{-1}
16	2.156×10^{-5}	-7.5×10^{-3}	2.187×10^{-5}	7.1×10^{-3}
8	2.148×10^{-5}	-1.1×10^{-2}	2.179×10^{-5}	3.3×10^{-3}
4	2.146×10^{-5}	-1.2×10^{-2}	2.177×10^{-5}	2.5×10^{-3}
cubic	2.145×10^{-5}	-1.2×10^{-2}	2.177×10^{-5}	2.3×10^{-3}

Table 2.5: Fluxes extracted at a sequence of radii on the numerical grid. a/M is the BH spin, r_0/M is the orbital radius and $|m|$ is the azimuthal mode. $\dot{E}_R$ is the flux measured at radius RM.

$|m|$	a/M	r_0/M	$\dot{E}_{100}$	$\dot{E}_{200}$	$\dot{E}_{300}$	$\dot{E}_{400}$	$\dot{E}_{500}$
2	0.99	4	2.4567×10^{-3}	2.4681×10^{-3}	2.4702×10^{-3}	2.4709×10^{-3}	2.4712×10^{-3}
2	0.99	10	4.1032×10^{-5}	4.3209×10^{-5}	4.3598×10^{-5}	4.3767×10^{-5}	4.3828×10^{-5}
2	0.90	10	4.1729×10^{-5}	4.3930×10^{-5}	4.4322×10^{-5}	4.4456×10^{-5}	4.4517×10^{-5}
2	0.00	12	1.9584×10^{-5}	2.1256×10^{-5}	2.1554×10^{-5}	2.1654×10^{-5}	2.1699×10^{-5}
3	0.99	4	5.8962×10^{-4}	5.9278×10^{-4}	5.9334×10^{-4}	5.9353×10^{-4}	5.9361×10^{-4}
3	0.99	10	4.5778×10^{-6}	4.8558×10^{-6}	4.9051×10^{-6}	4.9220×10^{-6}	4.9297×10^{-6}
3	0.90	10	4.6791×10^{-6}	4.9588×10^{-6}	5.0085×10^{-6}	5.0255×10^{-6}	5.0333×10^{-6}
3	0.00	12	1.9528×10^{-6}	2.1326×10^{-6}	2.1650×10^{-6}	2.1761×10^{-6}	2.1812×10^{-6}

Table 2.6: Best fit parameters, $\dot{E}_\infty$, p, q [appearing in Eq. (2.117)] for data presented in Table 2.5.

$|m|$	a/M	r_0/M	$\dot{E}_{600}$	$\dot{E}_\infty$	p	q	$\dot{E}_{FD}$	$\frac{\dot{E}_\infty - \dot{E}_{FD}}{\dot{E}_{FD}}$
2	0.99	4	2.4714×10^{-3}	2.4718×10^{-3}	2.04	3.40	2.4836×10^{-3}	-0.0048
2	0.99	10	4.3861×10^{-5}	4.3953×10^{-5}	1.96	2.31	4.3948×10^{-5}	0.0001
2	0.90	10	4.4550×10^{-5}	4.462×10^{-5}	2.05	2.70	4.4560×10^{-5}	0.0014
2	0.00	12	2.1723×10^{-5}	2.1779×10^{-5}	2.07	2.59	2.1722×10^{-5}	0.0026
3	0.99	4	5.9364×10^{-4}	5.9375×10^{-4}	2.09	10.74	5.9239×10^{-4}	0.0023
3	0.99	10	4.9339×10^{-6}	4.9428×10^{-6}	2.06	7.27	4.9417×10^{-6}	0.0002
3	0.90	10	5.0375×10^{-6}	5.0466×10^{-6}	2.06	7.16	5.0440×10^{-6}	0.0005
3	0.00	12	2.1839×10^{-6}	2.1900×10^{-6}	2.05	6.20	2.1890×10^{-6}	0.0004

Chapter 3

Towards adiabatic waveforms for inspiral into Kerr black holes: II. Dynamical sources and generic orbits

*This chapter is based on Physical Review D **78**, 024022 (2008), which was written in collaboration with Gaurav Khanna, Scott A. Hughes and Steve Drasco.*

3.1 Introduction

3.1.1 Background

The extreme mass ratio limit of general relativity's two-body problem has been a major focus of work in recent years. This limit corresponds to a stellar mass compact object that orbits and perturbs a massive black hole. The system generates gravitational waves (GWs) which drive the small body to inspiral into the large black hole. Measuring such "extreme mass ratio inspiral," or *EMRI*, events is a major goal for space-based GW antennae, particularly the LISA mission[1]. EMRIs should be mea-

---

[1] http://lisa.nasa.gov, http://sci.esa.int/lisa

surable to a redshift $z \sim 0.5 - 1$. The event rate at this range is estimated to be high enough that a multiyear LISA mission should measure dozens to hundreds of EMRI events [42]. Because the smaller body only slightly perturbs the larger black hole's spacetime, EMRI GWs are expected to provide an exceptionally clean probe of black hole properties. We expect to use EMRIs to measure black hole masses and spins with extremely good accuracy [11], and even to test how well the spacetime meets the rather stringent constraints that the "no-hair" theorems of general relativity impose on black holes [28, 48, 12, 44].

Understanding EMRI sources will require us to compare measured waves with theoretical models that are as accurate as possible. This goal motivates much recent EMRI work. The waves are sufficiently complicated that simply *detecting* them in LISA's datastream will be a challenge. Techniques for finding these events are currently being developed and tested through the "Mock LISA Data Challenges", or MLDCs (see Refs. [6, 7] for overviews of recent MLDCs). An important input to these challenges (and to the development of EMRI measurement techniques more generally) are waveform models that capture the true complexity of EMRI events (see [7] for discussion of recent work to include EMRI waves in the MLDCs).

This chapter presents a further step in our program to construct accurate EMRI wave models. As discussed in the introduction to chapter 2, our goal is to make "adiabatic" waveforms — waveforms built by separately treating the long-time dissipative evolution and the short-time conservative motion. In our present analysis, we take the short-time motion to be a geodesic orbit of the background spacetime; our approach thus amounts to approximating the inspiral trajectory as a sequence of geodesic orbits. As discussed by Pound and Poisson [82], this limit is more properly a "radiative" or "dissipative" approximation, since we do not include conservative self-interactions. It may be possible to augment this analysis with at least some conservative effects [41], so we believe the program we are developing is capable of building truly adiabatic inspiral waveforms as described in [82]. We will describe our goal as "adiabatic" waveforms, but the reader should bear in mind that the approximation we are currently developing is more restricted than this.

Geodesic orbits are described (up to initial conditions) by three conserved constants: energy E, axial angular momentum L_z, and "Carter constant" Q. Using black hole perturbation theory, we compute the rate at which these three constants evolve; fast and accurate frequency-domain codes make it possible to compute these rates of change fairly easily [36, 88, 34]. We then build the parameter-space trajectory $[E(t), L_z(t), Q(t)]$ followed by the small body; choosing initial conditions, it is simple to build the coordinate-space worldline $[r(t), \theta(t), \phi(t)]$ of a particular inspiral. From this worldline, we build the source to a time-domain code. The output of this code is, at last, our model EMRI wave.

3.1.2 Time-domain black hole perturbation theory

Since the frequency-domain portion of this program is already well in hand, our current focus is on the time-domain code. In essence, our goal is to build a code which takes as input any physically reasonable worldline, and provides as output the waveform produced by a small body on this worldline. In chapter 2 , we demonstrated an accurate (2+1)D numerical code to solve, in the time domain, the wave equation for curvature perturbations to a black hole — the Teukolsky equation [100]. Our code evolves the Weyl curvature scalar Ψ_4, constructed by projecting the vacuum curvature onto appropriate components of a null tetrad; see chapter 2 for details. The azimuthal dependence of Ψ_4 is separated out (due to the ϕ symmetry of black holes); the dependence on the Boyer-Lindquist coordinates r, θ, and t is found by evolving Ψ_4 on a numerical r-θ grid.

As is common in black hole perturbation theory, we treat the smaller body as a Dirac-delta point particle, leading to a singular source for the Teukolsky equation. In the frequency domain, the delta can be dealt with analytically, and presents no great challenge. By contrast, accurately computing the effect of a sharp source on the time-domain code's numerical grid can be extremely challenging. In chapter 2 , we presented a new technique for treating the singular source term. Our innovation was to model the delta as a series of finite impulses, with the largest impulse located close to the delta's argument, falling off rapidly as we move away from this "central"

spike. Importantly, this approach allows us to accurately model the *derivatives* of the delta function. Since the Teukolsky equation source depends on first and second derivatives of the delta (as well as the delta itself), this appears to give us an accuracy boost relative to other finite-difference delta representations (such as a truncated Gaussian), which may accurately capture the delta's behavior, but not do so well with the derivatives.

3.1.3   This chapter

Chapter 2 focused on the properties of this new source representation. To clarify this focus, we studied very simple orbits: We only considered the (astrophysically unlikely) case of circular, equatorial black hole orbits. We now extend this to include inclined, eccentric and generic orbits, as well as non-geodesic inspiral sequences.

A particle in a circular, equatorial orbit has constant radial and angular coordinate, confining it to a fixed location on the r-θ grid. Eccentricity means that the orbit oscillates radially, crossing radial grid zones. Similarly, orbital inclination results in angular grid crossing. We quickly discovered that these new motions introduce high frequency numerical noise. This noise can be controlled by combining a low pass filter with a higher order discretization of the delta function; details are given in Sec. 3.2. Aside from this mild extension of the basic formalism presented in chapter 2 , it was not terribly difficult to use our new source term to handle a broad class of astrophysically interesting orbits. To validate our results, we present in Sec. 3.3 extensive comparisons with waveform snapshots computed in the frequency domain [36], demonstrating graphically and quantitatively (with appropriate overlap integrals) excellent agreement between the two techniques.

As extensively discussed in the introduction to chapter 2 and here, our goal is to compute the waves from inspiral of a small body through a sequence of orbits. As a proof-of-concept demonstration of the feasibility of this idea, we present a simple example of inspiral in Sec. 3.4. In this example, we evolve through our geodesic sequence using a "kludged" approximation to the rates of change of orbital constants, using the code described in Refs. [43, 8]. These waveforms are not reliable EMRI

models, but they illustrate the ease with which we can handle the effect of radiation emission on the motion of the source. Computing waves from an inspiral is no more of a computational challenge than computing waves from a bound geodesic.

The next step will be to combine accurate radiative backreaction with our time-domain solver to compute "adiabatic" EMRI waveforms (albeit ones that still neglect conservative self interactions). Plans for this next step are described in our final summary, Sec. 3.5.

3.2 Dynamically varying discrete delta functions

In chapter 2 , we presented a method for representing a Dirac delta function and its derivatives on a discrete numerical grid. In that chapter, we only considered a delta with fixed radial and angular position. Naive application of the discrete delta models presented in chapter 2 leads to instabilities when the particle moves in the numerical grid. The following argument outlines the root cause of these instabilities. Consider the function $\delta[x-\alpha(t)]$, where $x_k \le \alpha(t) \le x_{k+1}$; i.e., the delta's peak varies with time and lies between two discrete grid points. Let x_i represent any discrete point on our grid, and let $h = x_{k+1} - x_k = x_k - x_{k-1}$ be the grid resolution. Naive application of the results from chapter 2 might lead us to model the delta function with the impulse weights

$$\delta_i(t_n) = \frac{\alpha(t_n) - x_k}{h^2} \text{ for } i = k + 1 \tag{3.1}$$

$$= \frac{x_{k+1} - \alpha(t_n)}{h^2} \text{ for } i = k \tag{3.2}$$

$$= 0 \text{ everywhere else } . \tag{3.3}$$

(This "two impulse" delta is in fact just the simplest representation we developed in chapter 2 , but is useful for the following discussion.) Each t_n defines a time slice of our r-θ grid. As α varies from one time slice to another, so do the coefficients at x_k and x_{k+1}. The frequency spectrum of $\delta_k(t_n)$ and $\delta_{k+1}(t_n)$ will reflect the amount of variation in α. A large variation in α will produce a high frequency component in the

Fourier transform of the time series of each weight. These variations couple to the time derivatives in the homogeneous part of the Teukolsky equation. Consequently, the solution contains spurious high frequency features of numerical origin.

Consider the extreme limit of this effect: α changes so rapidly that the delta's peak moves across a grid zone in a single time slice:

$$\alpha(t_1) = \alpha(t_0) - h \ , \qquad (3.4)$$

so that

$$x_k \leq \alpha(t_0) \leq x_{k+1} \qquad (3.5)$$

but

$$x_{k-1} \leq \alpha(t_1) \leq x_k \ . \qquad (3.6)$$

The weight of the delta function very suddenly becomes zero at x_{k+1} as we step from $t = t_0$ to $t = t_1$; likewise, the weight at x_{k-1} very suddenly jumps from non-zero to zero in this step. The coupling of this sudden change to numerical time derivatives drives instabilities in our code, in a manner reminiscent of the initial burst of radiation that occurs due to the sudden appearance of the particle at the start of our evolution; see Fig. 2 of chapter 2 .

This problem is substantially mitigated by using a delta representation with a wider stencil; examples of this are described in chapter 2 . Wide stencils reduce the amount by which each weight changes from step to step, thereby reducing numerical noise. Another useful tool is to increase the order of the delta representation, thereby increasing the smoothness of the delta and its derivatives. This is particularly important since the Teukolsky equation is a second-order differential equation; some smoothness in the derivatives is necessary to prevent the differential operator from seeding excessive noise. Finally, residual high frequency noise can be removed by convolving the source with a low pass filter[2]. These three techniques are each

---

[2]An obvious brute force workaround left off this list is to simply make the grid extremely fine and use tiny time steps. This does not address the root cause of instabilities seeded by particle motion, though it is certainly something used in practice (to the extent that computational limits allow).

described in the following subsections.

Each of these techniques smear out the delta function, pushing us away from the idealization of a zero width singularity. Choosing between stability (which tends to push us to a wider delta) and faithful representation of the singularity (which pushes us to a narrow delta) leads us to an optimization problem; we tune our delta representation in a way that (hopefully) minimizes numerical noise and maximizes accuracy. Note also that, in addition to high-frequency noise generated by abrupt movement of the delta across the grid, spurious excitations of the quasinormal modes of the black hole also appear due this motion. This source of "noise" appears to be controlled by grid resolution — wider grids lead to less pointlike deltas, which spuriously excite these modes. This spurious contribution to the EMRI waves can be mitigated with a form of Richardson extrapolation [84]. We discuss this further in Sec. 3.3 and the Appendix.

3.2.1 Higher order delta functions

Discrete delta representations based on linear and cubic interpolation were derived in chapter 2 . We now extend this process to arbitrary polynomial order, equipping us with an entire family of discrete delta functions.

As in chapter 2 , we start from the defining integral,

$$\int_{\alpha(t)-\epsilon}^{\alpha(t)+\epsilon} dx \ f(x) \ \delta[x - \alpha(t)] \ = \ f[\alpha(t)] \ . \tag{3.7}$$

Let $x_{k+n-1} \leq \alpha \leq x_{k+n}$; the reason for our somewhat idiosyncratic choice of subscripts will become clear as we proceed. For clarity, we will not explicitly write out the time dependence of α; the reader should bear in mind that $\alpha = \alpha(t)$ in all that follows. Rewriting Eq. (3.7) as a sum over a finite step size, we have

$$\int_{\alpha-\epsilon}^{\alpha+\epsilon} dx \ f(x) \ \delta(x - \alpha) \ \simeq \ h \sum_i f(x_i) \, \delta_i$$
$$\Rightarrow f(\alpha) \ \simeq \ h \sum_i f(x_i) \, \delta_i \ . \tag{3.8}$$

The function $f(\alpha)$ can be approximated by the Lagrange interpolating polynomial,

$$f(\alpha) = \sum_{i=k}^{k+2n-1} \frac{\Pi(\alpha)}{(\alpha - x_i)\Pi'(x_i)} f(x_i) , \qquad (3.9)$$

where $2n$ is the order of interpolation and

$$\Pi(\alpha) = \prod_{i=k}^{k+2n-1} (\alpha - x_i) = \prod_{i=0}^{2n-1} (\alpha - x_{k+i}) \qquad (3.10)$$

$$\Pi'(x_j) = \left[\frac{d\Pi}{d\alpha}\right]_{\alpha=x_j} = \prod_{i=k,i\neq j}^{k+2n-1} (x_j - x_k) . \qquad (3.11)$$

Inserting this in Eq. (3.8) leaves us with

$$\sum_{i=k}^{k+2n-1} \frac{\Pi(\alpha)}{(\alpha - x_i)\Pi'(x_i)} f(x_i) = h \sum_i f(x_i) \delta_i ; \qquad (3.12)$$

comparing coefficients of $f(x_i)$ allows us to read off δ_i,

$$\delta_i = \frac{\Pi(\alpha)}{h(\alpha - x_i)\Pi'(x_i)} . \qquad (3.13)$$

We thus see that δ_i is non-zero for $i \in [k, k + 2n - 1]$.

The weights for derivatives of the delta function can be obtained similarly. Writing the identities

$$\int dx\, f(x)\, \delta'(x - \alpha) = -f'(\alpha) \qquad (3.14)$$

$$\int dx\, f(x)\, \delta''(x - \alpha) = f''(\alpha) \qquad (3.15)$$

96

as sums gives us

$$h \sum_i f(x_i)\,\delta_i' \simeq -f'(\alpha)$$

$$= -h \sum_i f'(x_i)\delta_i$$

$$\Rightarrow \sum_i f(x_i)\,\delta_i' = -\sum_{i=k}^{k+2n-1} \frac{\Pi(\alpha)f'(x_i)}{h(\alpha - x_i)\Pi'(x_i)} , \qquad (3.16)$$

$$h \sum_i f(x_i)\,\delta_i'' \simeq f''(\alpha)$$

$$= h \sum_i f''(x_i)\delta_i$$

$$\Rightarrow \sum_i f(x_i)\,\delta_i'' = \sum_{i=k}^{k+2n-1} \frac{\Pi(\alpha)f''(x_i)}{h(\alpha - x_i)\Pi'(x_i)} . \qquad (3.17)$$

We now insert centered finite difference formulae for the derivatives of $f(x_i)$ to obtain

$$\sum_i f(x_i)\,\delta_i' = -\sum_{i=k}^{k+2n-1} \frac{\Pi(\alpha)}{h(\alpha - x_i)\Pi'(x_i)} \times$$

$$\left[\frac{f(x_{i+1}) - f(x_{i-1})}{2h} \right] , \qquad (3.18)$$

$$\sum_i f(x_i)\,\delta_i'' = \sum_{i=k}^{k+2n-1} \frac{\Pi(\alpha)}{h(\alpha - x_i)\Pi'(x_i)} \times$$

$$\left[\frac{f(x_{i+1}) - 2f(x_i) + f(x_{i-1})}{h^2} \right]. \qquad (3.19)$$

Expressions (3.18) and (3.19) are in a form that makes it simple to read off δ_i' and δ_i''. For example, δ_j' can be calculated by setting $f(x_j) = 1$ and $f(x_l) = 0$, $l \neq j$. It is straightforward to verify that setting $n = 1$ and $n = 2$ reproduces the weights given by the two-point linear hat and the cubic formulae (described in chapter 2) respectively. We also note that the delta derivative coefficients are non-zero for $i \in [k-1, k+2n]$.

3.2.2 Wider stencils at a given interpolation order

In chapter 2 , we generalized the two-point linear hat delta function such that it can be represented over a larger number of points. Similarly, we develop a procedure to widen the stencil of the generalized model obtained from Eqs. (3.13), (3.18), and (3.19).

Consider a model for δ_i obtained from Eq. (3.13) for some $n = m$. Then, $\delta_i \neq 0$ for $i \in [k, \ldots, k+2m-1]$. Our goal is to widen this representation by some integer factor w such that the coefficients are non-zero for a wider range of grid points. Let us label the weights of this wider representation by δ_i^w, with $\delta_i^w \neq 0$ for $i \in [k, \ldots, k+2wm-1]$. It should be emphasized that this is different from simply using Eq. (3.13) with $n = wm$; we have not changed the polynomial order, it remains fixed at $2m$.

For concreteness, let us choose $w = 2$, doubling the number of points in the delta representation. We infer the coefficients δ_i^2 at gridpoints $i = k$, $k + 2$, $k + 4$, ..., $k + 4m - 2$, by widening the grid by a factor of two: We evaluate δ_i with $h \rightarrow 2h$, $x_{k+j} \rightarrow x_{k+2j}$ to get

$$
\begin{aligned}
\delta_{k+2j}^2 &= \delta_{k+j} \rfloor_{h \rightarrow 2h, x_{k+j} \rightarrow x_{k+2j}} \\
&= \frac{\Pi(\alpha)}{2h(\alpha - x_{k+2j})\Pi'(x_{k+2j})} ,
\end{aligned} \tag{3.20}
$$

where

$$
\Pi(\alpha) = \prod_{i=0}^{2m-1} (\alpha - x_{k+2i}) . \tag{3.21}
$$

Finally, we need δ_i^2 at the intermediate points $i = k+1$, $k+3$, ..., $k+4m-1$. We do this by exploiting the translational symmetry of the problem. Momentarily reinsert the time dependence of the δ's and α. Now consider the hypothetical situation where

$$
\alpha(t_0) = \alpha_0 ,
$$

$$
\alpha(t_1) = \alpha_0 - h ; \tag{3.22}
$$

i.e, $\alpha(t)$ changes by a grid spacing from t_0 to t_1. We must have

$$
\begin{aligned}
\delta^2_{k+2j}(t_1) &= \delta^2_{k+2j+1}(t_0) \\
\Rightarrow \delta^2_{k+2j}(t_0)\rfloor_{\alpha(t)\to\alpha_0-h} &= \delta^2_{k+2j+1}(t_0)\rfloor_{\alpha(t)\to\alpha_0} .
\end{aligned}
\tag{3.23}
$$

We can turn this equation the other way around to read off the coefficient δ^2_{k+2j+1} at t_0: Simply replace $\alpha(t)$ with $\alpha(t) - h$ in the formula for $\delta^2_{k+2j}(t_0)$ to obtain $\delta^2_{k+2j+1}(t_0)$. Since there was nothing special about our time slice, t_0, we find

$$
\delta^2_{k+2j+1}(t_n) = \delta^2_{k+2j}(t_n)\rfloor_{\alpha(t)\to\alpha(t)-h}
\tag{3.24}
$$

for *any* moment t_n.

Though we chose $w = 2$ for concreteness, the above argument can be generalized to any integer w. Since our result holds for all time slices, we again suppress the time dependence to obtain expressions for any integer w:

$$
\Pi(\alpha) = \prod_{i=0}^{2m-1} (\alpha - x_{k+wi}) ,
\tag{3.25}
$$

$$
\delta^w_{k+wj} = \delta_{k+j}\rfloor_{h\to wh, x_{k+j}\to x_{k+wj}}
\tag{3.26}
$$

$$
= \frac{\Pi(\alpha)}{wh(\alpha - x_{k+wj})\Pi'(x_{k+wj})} ,
\tag{3.27}
$$

$$
\delta^w_{k+wj+l} = \delta^2_{k+wj}\rfloor_{\alpha(t)\to\alpha(t)-lh}
$$

$$
\text{for } l \in [1, 2, \ldots, w - 1] .
\tag{3.28}
$$

These techniques carry over to the derivatives as well:

$$
\delta'^w_{k+wj} = \delta'_{k+j}\rfloor_{h\to wh, x_{k+j}\to x_{k+wj}}
\tag{3.29}
$$

$$
\delta'^w_{k+wj+l} = \delta'^w_{k+wj}\rfloor_{\alpha(t)\to\alpha(t)-lh}
$$

$$
\text{for } l \in [1, 2, \ldots, w - 1] ;
\tag{3.30}
$$

99

and

$$\delta_{k+wj}''^{w} = \delta_{k+j}''|_{h\to wh, x_{k+j}\to x_{k+wj}} \tag{3.31}$$

$$\delta_{k+wj+l}''^{w} = \delta_{k+wj}''^{w}|_{\alpha(t)\to\alpha(t)-lh}$$

$$\text{for } l \in [1, 2, \ldots, w-1] . \tag{3.32}$$

These should be used with Eqs. (3.13), (3.18) and (3.19) to widen the Teukolsky source term by any factor w.

3.2.3 Smoothing the source with a Gaussian filter

Further control of numerical noise can be achieved by filtering high frequency components in the source term. This requires a convolution of the source with a discrete low pass filter. We use a Gaussian filter because it maximizes the uncertainty principle — it can be localized in both position and frequency with greatest efficiency.

Consider a source of the form

$$s(x) = f_1(x)\delta(x - \alpha) + f_2(x)\delta'(x - \alpha) + f_3(x)\delta''(x - \alpha) . \tag{3.33}$$

Delta function identities allow us to rewrite this as

$$s(x) = g_1(\alpha)\delta(x - \alpha) + g_2(\alpha)\delta'(x - \alpha) + g_3(\alpha)\delta''(x - \alpha) , \tag{3.34}$$

where

$$g_1(\alpha) = f_1(\alpha) - f_2'(\alpha) + f_3''(\alpha) ,$$
$$g_2(\alpha) = f_2(\alpha) - 2f_3'(\alpha) ,$$
$$g_3(\alpha) = f_3''(\alpha) . \tag{3.35}$$

On a discrete grid, this becomes

$$s(x_i) = s_i = g_1(\alpha)\delta_i + g_2(\alpha)\delta_i' + g_3(\alpha)\delta_i'' . \tag{3.36}$$

100

If the delta function and its derivatives span $2n + 2$ grid points, with $x_{k+n-1} \leq \alpha \leq x_{k+n}$, then $s_i \neq 0$ for $i \in [k-1, \dots, k+2n]$. The source s_i is zero everywhere else on the grid.

The Gaussian filter is given by

$$c_k = \frac{\exp[-(kh/b)^2/2]}{\sum_{i=-p}^{p} \exp[-(ih/b)^2/2]} \,, \tag{3.37}$$

where $k \in [-p, -p+1, \dots, p]$ and b is the width of the filter. The quantities p and b are adjustable parameters. Typically, we use $p = 30$ and $b = 1.5h$. Notice that

$$\sum_{i=-p}^{p} c_i = 1 \,; \tag{3.38}$$

this normalization guarantees that the integrated value of any function convolved with the filter is unchanged.

We now convolve the source with the filter to obtain

$$sg_k = \sum_{i=-p}^{p} c_i s_{k+i} \,, \tag{3.39}$$

where sg_k is the smoothed source term. This indicates that $sg_k \neq 0$ for $k \in [k - p, \dots, k + 2n + p - 1]$.

A wide filter spreads the source over a large domain on the numerical grid and thus increases errors, although it eliminates spurious harmonics. We have found that using a wide stencil followed by a narrow Gaussian smoother works very well to reduce numerical noise and minimize errors from an insufficiently pointlike source.

3.2.4 Order of convergence of the filtered delta

Chapter 2 discussed in detail the convergence of a code that uses a discrete delta. Crucial background is given by Ref. [102] and summarized in chapter 2 . The key

point is that the moment

$$M_r = h \sum_{i=k}^{k+2n-1} \delta_i (x_i - \alpha)^r \tag{3.40}$$

controls the delta's convergence properties. Clearly, $M_0 = 1$ (otherwise the delta is not properly normalized); in the continuum limit, $M_r = 0$ for $r > 0$. For the discrete delta, the smallest non-zero value of r for which $M_r \neq 0$ sets the order of convergence. In particular, if $M_r \neq 0$, then a code which uses this delta will be no higher than rth-order convergent.

We now show that, if a delta representation is second-order convergent before smoothing with the Gaussian filter ($M_0 = 1$, $M_1 = 0$, $M_2 \neq 0$), it will remain second-order convergent after smoothing. Upon convolving the discrete delta with the Gaussian smoother, we find

$$\delta g_i = \sum_{j=-p}^{p} c_j \delta_{i+j} \, . \tag{3.41}$$

Let us denote the moments of the smoothed delta by M_r^g. As discussed in Sec. 3.2.3, the convolution does not change the delta's normalization as long as the Gaussian filter is itself properly normalized; thus

$$M_0^g \equiv h \sum_{i=k-p}^{k+2n+p-1} \delta g_i = 1 \, . \tag{3.42}$$

We now examine the next higher moment of the smoothed delta:

102

$$M_1^g \equiv h \sum_{i=k-p}^{k+2n+p-1} \delta g_i(x_i - \alpha) = h \sum_{j=-p}^{p} \sum_{i=k-p}^{k+2n+p-1} c_j \delta_{i+j}(x_i - \alpha) ,$$

$$= h \sum_{j=-p}^{p} c_j \sum_{i=k-p}^{k+2n+p-1} \delta_{i+j}(x_i - \alpha) ,$$

$$= h \sum_{j=-p}^{p} c_j \sum_{i=k-p}^{k+2n+p-1} \delta_{i+j}(x_{i+j} - \alpha - jh) ,$$

$$= h \sum_{j=-p}^{p} c_j \sum_{i=k-p}^{k+2n+p-1} \delta_{i+j}(x_{i+j} - \alpha)$$

$$- h \sum_{j=-p}^{p} hjc_j \sum_{i=k-p}^{k+2n+p-1} \delta_{i+j} . \tag{3.43}$$

The first term on the final line of (3.43) gives zero: Since $\sum \delta_l x_l = \alpha$,

$$h \sum_{j=-p}^{p} c_j \sum_{i=k-p}^{k+2n+p-1} \delta_{i+j}(x_{i+j} - \alpha) = h \sum_{j=-p}^{p} c_j \sum_{l=k-p+j}^{k+2n+p+j-1} \delta_l(x_l - \alpha)$$

$$= 0 . \tag{3.44}$$

The second line follows because $|j| <= p$, $\delta_i = 0$ if i lies outside $[k, k + 2n - 1]$ and $M_1 = 0$.

The second term on the final line of (3.43) also yields zero:

$$h \sum_{j=-p}^{p} hjc_j \sum_{i=k-p}^{k+2n+p-1} \delta_{i+j} = h^2 \sum_{j=-p}^{p} jc_j \sum_{l=k-p+j}^{k+2n+p+j-1} \delta_l ,$$

$$= h^2 \sum_{j=-p}^{p} jc_j$$

$$= 0 . \tag{3.45}$$

The Gaussian filter's symmetry property $c_j = c_{-j}$ has been applied in the last step. Hence, we find $M_1^g = M_1 = 0$.

Evaluating the second moment proceeds similarly, but we find in the end terms involving $\sum_{j=-p}^{p} j^2 c_j$ which do not vanish. Thus, M_2^g is the first non-vanishing moment

of the discrete delta, demonstrating that the Gaussian-filtered discrete delta function exhibits second-order convergence. The argument can be extended to the delta derivatives as well. The smoothed Teukolsky source term will thus be second-order convergent.

3.3 Waveforms and comparisons for generic geodesic Kerr orbits

We now present the waveforms generated by a point particle in a geodesic orbit around a Kerr black hole. The code used to generate these waves is discussed in detail in chapter 2; the only important change to that discussion is that the source term uses the techniques presented in Sec. 3.2 above. We begin by reviewing Kerr black hole geodesics, sketching the numerical scheme used to solve the equations of motion. We then examine different classes of eccentric and inclined orbits and compare the waveforms against those obtained from a frequency-domain code whose details are given in Ref. [36]. We compute the correlation between the two waveforms in order to measure our level of agreement with frequency-domain waveforms.

Our numerical grid is laid out in Boyer-Lindquist coordinates and uses $(\delta r, \delta \theta, \delta t) = (0.04M, \pi/60, 0.02M)$ for the radial, angular and temporal resolutions. The source term is constructed using Eqs. (3.13), (3.18) and (3.19) with n_θ in the range 3–9 (depending on the orbit) for the angular delta-function and $n_r = 2$ for the radial delta. We use a Gaussian filter of width $b = 1.5\delta\theta$ to smooth higher harmonic noise.

3.3.1 Geodesics in Kerr spacetime

The source term for the time-domain code takes as input the worldline of the perturbation's source. Here, we neglect radiation reaction and assume that the point particle follows a bound geodesic trajectory around the central massive black hole. This bound trajectory can be computed by numerically integrating the geodesic equations. We now briefly review how we massage the geodesic equations to put them

into a form that makes for accurate numerical calculation; this material is presented in greater depth in Sec. IIC of Ref. [8].

The normal "textbook" presentation of the equations governing Kerr black hole geodesics is

$$\Sigma^2 \left(\frac{dr}{d\tau}\right)^2 = \left[E(r^2 + a^2) - aL_z\right]^2$$
$$-\Delta \left[r^2 + (L_z - aE)^2 + Q\right]$$
$$\equiv R(r) \tag{3.46}$$

$$\Sigma^2 \left(\frac{d\theta}{d\tau}\right)^2 = Q - \cos^2\theta \left[a^2(1 - E^2) + L_z^2/\sin^2\theta\right] \tag{3.47}$$

$$\Sigma\frac{d\phi}{d\tau} = \frac{L_z}{\sin^2\theta} - aE + \frac{a}{\Delta}\left[E(r^2 + a^2) - aL_z\right] \tag{3.48}$$

$$\Sigma\frac{dt}{d\tau} = a(L_z - aE\sin^2\theta)$$
$$+\frac{r^2 + a^2}{\Delta}\left[E(r^2 + a^2) - aL_z\right] . \tag{3.49}$$

[See, e.g., Ref. [72], Eqs. (33.32a–d).] Here, $\Sigma = r^2 + a^2\cos^2\theta$, $\Delta = r^2 - 2Mr + a^2$ (where $a = |\vec{S}|/M$ is the black hole's spin angular momentum per unit mass). The constants of motion are orbital energy E, axial angular momentum L_z, and Carter constant Q.

This form of the equations of motion is not well suited to numerical studies; in particular, $dr/d\tau$ and $d\theta/d\tau$ pass through zero and change sign when the orbiting body goes through turning points associated with those motions. A handy way to eliminate these problems is to eliminate the turning points by remapping the coordinates r and θ to parameters which accumulate secularly. The following parametrization, inspired by the Newtonian limit, has been found to work extremely well even deep in the strong field of rapidly rotating black holes:

$$r = \frac{p}{1 + e\cos\psi} , \tag{3.50}$$
$$\cos\theta = \cos\theta_{\text{min}}\cos\chi . \tag{3.51}$$

105

In the Newtonian limit, p is the orbit's semi-latus rectum, and e is its eccentricity; $\theta_{\min}$ is the minimum value of θ reached by the orbiting body, and is used to define the orbit's inclination θ_{inc}

$$\theta_{\mathrm{inc}} = \frac{\pi}{2} - \mathrm{sgn}(L_z)\theta_{\min} \ . \tag{3.52}$$

Once E, L_z, and Q are specified, p, e, and θ_{inc} are fully determined. It is then a straightforward matter to turn Eqs. (3.46) and (3.47) into expressions for $d\psi/d\tau$ and $d\chi/d\tau$; see Ref. [8] for details. The resulting expressions behave extremely well for all bound orbits outside the black hole's event horizon. A numerical integrator for these variables allows us to compute the dynamics of our orbiting body's Teukolsky equation source term.

Before moving on, we note that, within the context of the dissipative-only or radiative approximation to inspiral, it is simple to modify these equations to build the worldline of an inspiralling body: We simply allow the orbital "constants" (E, L_z, and Q; or, p, e, and θ_{inc}) to evolve according to the inspiral law. Reference [8] uses approximate radiation reaction, based on fits to strong-field radiation reaction calculations in regimes where it is well understood, to compute the inspiral worldlines which underlie the "kludge" waveforms. We use this prescription for evolving the constants in Sec. 3.4 to demonstrate this code's ability to compute inspiral waves.

3.3.2 Comparison with frequency-domain waveforms

To validate our waveforms, we compare with the "snapshots" generated using the frequency-domain code described in Ref. [36]. This code uses the fact that bound Kerr geodesics are fully described by three frequencies (radial Ω_r, latitudinal Ω_θ, and axial Ω_ϕ) to build the waveform from a geodesic orbit as a sum over harmonics of these frequencies [35]. Since both the time-domain and frequency-domain codes solve the same master equation, they should produce identical waveforms for identical orbits, so long as each code is sufficiently accurate.

To quantify the accuracy with which a time-domain waveform $X = (x_1, x_2, \ldots, x_n)$ agrees with a frequency-domain waveform $Y = (y_1, y_2, \ldots, y_n)$, we use the following

106

correlation measure:

$$r_{XY} \equiv \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2}\sqrt{\sum (y_i - \bar{y})^2}} \, . \tag{3.53}$$

(The sums in all cases are from $i = 1$ to $i = n$.) This coefficient is identical to the *match* between two waveforms defined by Owen [76] in the white noise limit [noise spectral density $S_h(f) = $ constant]. One might expect the waveforms' mean values $\bar{x}$ and $\bar{y}$ to equal zero. However, finite duration effects can make these quantities slightly non-zero, so it is useful to explicitly do this subtraction.

A useful reformulation of Eq. (3.53) is

$$r_{XY} \equiv \frac{n \sum x_i y_i - \sum x_i \sum y_i}{\sqrt{n \sum x_i^2 - \left(\sum x_i\right)^2}\sqrt{n \sum y_i^2 - \left(\sum y_i\right)^2}} \, . \tag{3.54}$$

Note that r_{XY} is always between -1 and 1; a value close to 1 indicates that the two waveforms are well correlated. Note also that the correlation depends on how many points n are used in comparing the two waveforms (or equivalently, the span of time over which we compare the waves). We have found that as long as $n \geq$ several hundred, we get consistent results: Changing n for a given comparison only causes small variations in the fourth significant digit of r_{XY}.

It is of course possible to concoct other measures of how well two waveforms agree. Ideally, disagreements between waveforms should be quantified in terms of their observational significance. For example, Cutler and Vallisneri have demonstrated that it is not unusual for waveforms with a match of 0.9999 to differ significantly in their estimates of the parameters which describe the source [31]. For our present purpose, r_{XY} is sufficient to demonstrate that our time-domain code produces high quality waveforms; whether they are sufficiently high quality to be used for GW measurement purposes will need to be re-examined at a later time.

An important step in producing accurate waveforms is to perform runs at multiple resolutions, then estimate (and eliminate) the waveform error using a form of Richardson extrapolation [84]. This plays a crucial role in reducing "noise" from

spurious excitation of the large black hole's quasinormal modes. The details of this extrapolation technique are described in Appendix 3.6.
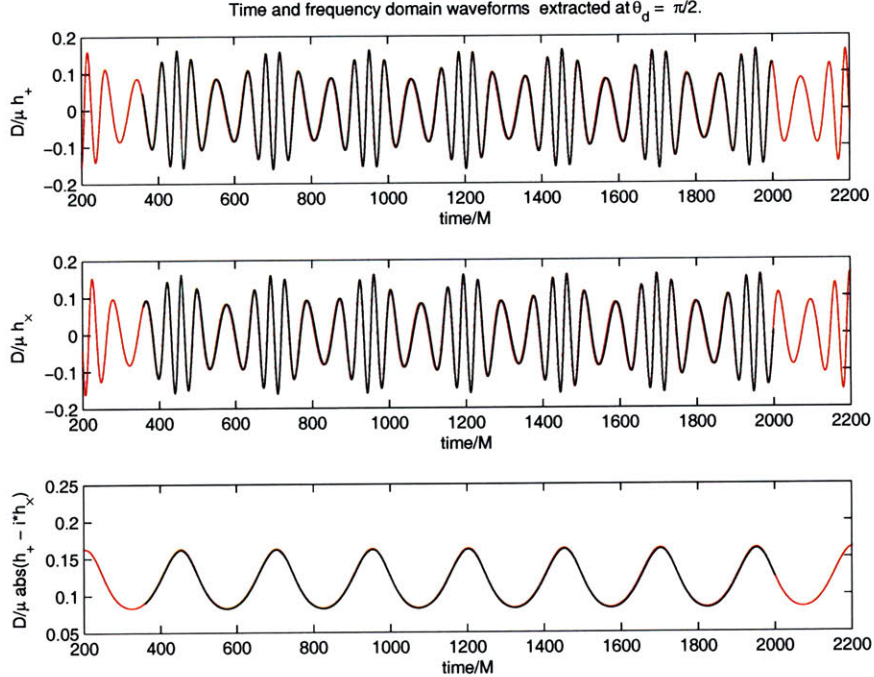
Figure 3-1: Comparison of time- and frequency-domain waveforms. We show waves for the $m = 2$ mode from a point particle with orbital parameters $p = 6.472M$, $e = 0.3$ and $\theta_{\mathrm{inc}} = 0$ orbiting a black hole with spin $a/M = 0.3$. The angle between the spin axis of the black hole and the line of sight is $\theta_d = \pi/2$. Time-domain results are in black, frequency-domain results in red. Top panel: "plus" polarizations in dimensionless units. Middle: "cross" polarizations. Bottom: Comparison of $|h_+ - ih_\times|$. This last quantity gives a good visual measure of the level of agreement between the two waveforms. The correlations between the two waveforms are 0.9974 (plus) and 0.9975 (cross).

Tables 3.1, 3.2, 3.3, 3.4, 3.5 and 3.6 list the correlation coefficients for the $m = 2$ and $m = 3$ azimuthal modes of different classes of orbits. The coefficient is greater than 0.99 for a large fraction of parameter space. Time domain runs corresponding to each column required about 125 CPU hours on an Apple MacPro processor. That code was compiled using the Intel C++ compiler. The frequency domain code's cost is about 3-4 CPU hours per waveform when attempting to get both asymptotic energy fluxes to accuracies of about 0.1% to 1% on a machine using a 3.2 GHz Pentium
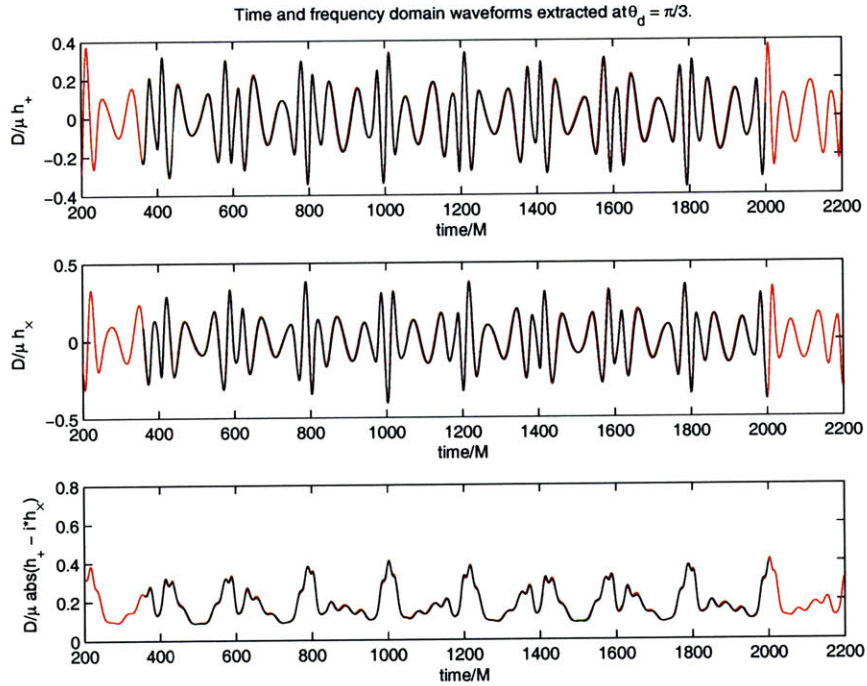
Figure 3-2: Comparison of time- and frequency-domain waveforms. Here, we show waves for the $m = 2$ mode for a geodesic with $p = 6M$, $e = 0.3$ and $\theta_{\mathrm{inc}} = \pi/3$ about a black hole with spin $a/M = 0.9$; black is time-domain results, red is frequency domain. The correlations in this case are 0.9961 (plus) and 0.9962 (cross).

4 Xeon processor. We also show (Figs. 3-1, 3-2, and 3-3) examples of the waves, computed with both time- and frequency-domain codes, to give the reader a visual sense of the overlap.

3.4 Inspiral waveforms

Having demonstrated that the finite-impulse source works well for astrophysically relevant generic black hole orbits, we now examine how well we do evolving through a sequence of such orbits. Since each orbit in the sequence is no different than the orbits that we validated against in Sec. 3.3.2, we anticipate no great difficulty here. Indeed, the biggest challenge is choosing a method to evolve through our sequence. Our goal is to do this with a frequency-domain code to build the orbital-constant
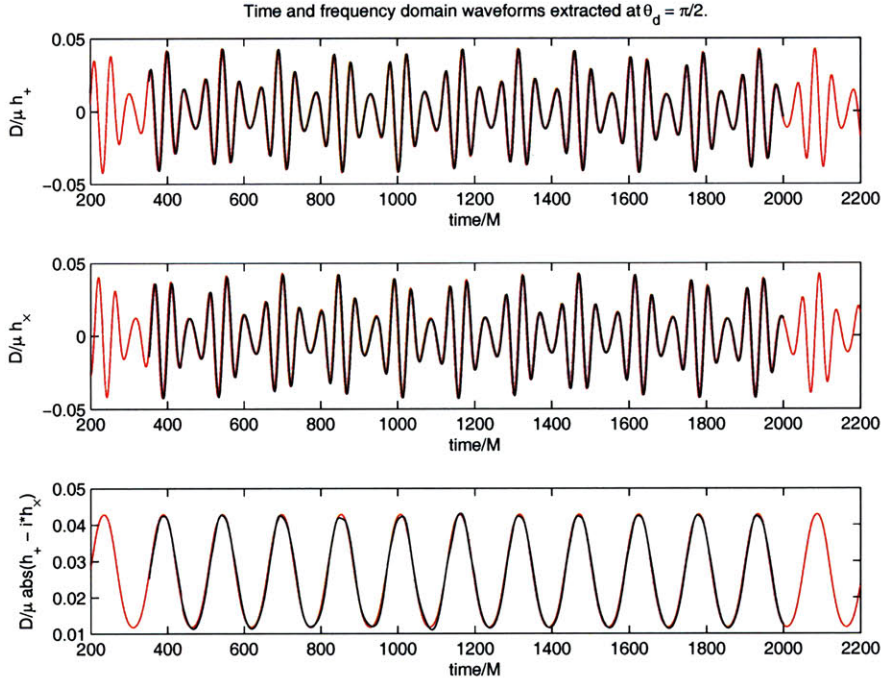
Figure 3-3: Comparison of time- and frequency-domain waveforms. These waves are for the $m = 3$ mode from a circular geodesic with orbital parameters $p = 6M$, and $\theta_{\mathrm{inc}} = \pi/4$ around a hole with spin $a/M = 0.9$. All symbols have the same meaning as in Fig. 3-1. The correlations are 0.9769 (plus) and 0.9770 (cross).

trajectory $[E(t), L_z(t), Q(t)]$. To quickly produce results that are qualitatively correct, we presently make this trajectory using the "kludge" inspiral treatment described in Ref. [43], and used to make model waveforms in Ref. [8]. The "kludge" uses a somewhat idiosyncratic mix of post-Newtonian backreaction formulae combined with numerical results from frequency-domain backreaction in the circular, inclined ($e = 0$, $\theta_{\mathrm{inc}} \neq 0$) and eccentric, equatorial ($e \neq 0$, $\theta_{\mathrm{inc}} = 0$) limits to estimate the properties of EMRI waves. By construction, the results agree very well with Teukolsky-based inspirals in those limits; for the generic case, they produce plausible inspirals.

Figure 3-4 shows our waveform for a "kludge" inspiral. We took the large black hole to have spin $a = 0.5M$, and set the mass ratio to $\mu/M = 0.016$. The orbit was initially chosen to have semi-latus rectum $p = 10M$, eccentricity $e = 0.5$, and inclination $\theta_{\mathrm{inc}} = 0.5$ radians. This figure shows features reminiscent of the geodesic

Table 3.1: Correlation between time- and frequency-domain waveforms for the $m = 2$ mode for a range of equatorial, eccentric orbits. The parameters p, e and $\theta_{\rm inc}$ are semi-latus rectum, eccentricity, and inclination of the geodesic orbit, a/M is the black hole spin and θ_d is the angle between the spin axis and the line of sight to the observer. The last two columns show correlations for the plus and cross polarizations.

p/M	e	$\theta_{\rm inc}$ (deg)	a/M	θ_d (deg)	h_+ corr.	$h_\times$ corr.
6.472	0.3	0	0.3	30	0.9961	0.9962
6.472	0.3	0	0.3	60	0.9969	0.9969
6.472	0.3	0	0.3	90	0.9974	0.9975
5.768	0.3	0	0.7	30	0.9971	0.9971
5.768	0.3	0	0.7	60	0.9977	0.9978
5.768	0.3	0	0.7	90	0.9983	0.9983
6.472	0.7	0	0.3	30	0.9915	0.9911
6.472	0.7	0	0.3	60	0.9911	0.9908
6.472	0.7	0	0.3	90	0.9900	0.9901
5.768	0.7	0	0.7	30	0.9625	0.9607
5.768	0.7	0	0.7	60	0.9621	0.9601
5.768	0.7	0	0.7	90	0.9596	0.9578

snapshots shown in Figs. 3-1, 3-2, and 3-3; in addition, one can clearly see evolution of the wave's properties. The increase in the wave's frequency, largely due to the decay of the orbit's semi-latus rectum, is quite clear. Perhaps less obvious is a signature of the eccentricity's decay. This is illustrated most clearly by comparing the lower left and lower right panels of Fig. 3-4, which zoom onto early and late portions of the inspiral. Early on, the waveform is dominated by a series of high-frequency bursts; these occur when the small body passes through periapsis and "whirls" most rapidly about the massive black hole. There is then a relatively quiet section while the body "zooms" out to apoapsis, and then comes in to "whirl" at periapsis again. As eccentricity shrinks, the difference between periapsis and apoapsis becomes smaller. The high-frequency bursts crowd closer and closer together, approaching a continuum sinusoid as the eccentricity approaches zero.

Although this inspiral model is somewhat unphysical, we expect that it shares many properties with true adiabatic inspiral waveforms. In particular, the spectral evolution of a wave like that in Fig. 3-4 should be quite similar to the evolution of real EMRI waveforms. It should be emphasized that computing the waveform shown

Table 3.2: Correlation between time- and frequency-domain waveforms for the $m = 2$ mode for a range of inclined nearly circular orbits. All symbols have the same meaning as in Table 3.1.

p/M	e	θ_{inc} (deg)	a/M	θ_d (deg)	h_+ corr.	$h_\times$ corr.
6	10^{-4}	45	0.5	60	0.9968	0.9967
6	10^{-4}	45	0.5	90	0.9961	0.9960
8	10^{-4}	45	0.5	60	0.9923	0.9919
8	10^{-4}	45	0.5	90	0.9908	0.9903
6	10^{-4}	45	0.9	60	0.9967	0.9967
6	10^{-4}	45	0.9	90	0.9961	0.9961
8	10^{-4}	45	0.9	60	0.9920	0.9919
8	10^{-4}	45	0.9	90	0.9905	0.9907
6	10^{-4}	60	0.5	60	0.9964	0.9965
6	10^{-4}	60	0.5	90	0.9952	0.9952
8	10^{-4}	60	0.5	60	0.9917	0.9910
8	10^{-4}	60	0.5	90	0.9888	0.9882
6	10^{-4}	60	0.9	60	0.9986	0.9986
6	10^{-4}	60	0.9	90	0.9981	0.9982
8	10^{-4}	60	0.9	60	0.9917	0.9915
8	10^{-4}	60	0.9	90	0.9891	0.9890

in Fig. 3-4 required about as much computational effort as computing the geodesic snapshot waves, Figs. 3-1, 3-2, and 3-3 (modulo a factor $\sim$ 4–5 since the waveform in Fig. 3-4 lasts $\sim$ 4–5 times longer than the others). Given a robust code to generate the inspiral worldline of EMRI systems, the waveforms that our code produces should be a useful tool for examining issues in LISA measurement and data analysis.

3.5 Summary and future work

We have now shown that the finite impulse delta representation of the time-domain Teukolsky equation's source works very well for complicated and astrophysically relevant orbits. In our previous analysis ([97], chapter 2), we confined ourselves to the simplest circular, equatorial black hole orbits. The basic ideas from chapter 2 work well even when the source arises from highly inclined and highly eccentric orbits, and when the source evolves through a sequence of those orbits. It is now a relatively straightforward matter to compute the waves arising from a body following

Table 3.3: Correlation between time- and frequency-domain waveforms for the $m = 2$ mode for a range of generic orbits. All symbols have the same meaning as in Table 3.1.

p/M	e	θ_{inc} (deg)	a/M	θ_d (deg)	h_+ corr.	$h_\times$ corr.
6	0.3	40	0.9	60	0.9978	0.9978
6	0.3	40	0.9	90	0.9976	0.9976
8	0.3	40	0.5	60	0.9898	0.9897
8	0.3	40	0.5	90	0.9910	0.9910
6	0.7	40	0.9	60	0.9898	0.9906
6	0.7	40	0.9	90	0.9889	0.9891
6	0.7	60	0.9	60	0.9905	0.9868
6	0.7	60	0.9	90	0.9895	0.9866
6	0.3	60	0.9	60	0.9961	0.9962
6	0.3	60	0.9	90	0.9950	0.9954
8	0.3	60	0.5	60	0.9906	0.9890
8	0.3	60	0.5	90	0.9884	0.9866

any reasonably behaved worldline in the spacetime of a black hole.

The primary complication arising from these more generic orbit classes is that the orbiting body will cross zones within our numerical grid. The source thus becomes dynamical; the finite-impulse delta must likewise be dynamical to represent it. The evolution of the impulses that we use to represent the delta can seed numerical noise, reducing the calculation's accuracy. We have found that minor extensions of chapter 2 's basic techniques greatly mitigate the impact of this source of numerical noise. In particular, by using a higher-order representation (Sec. 3.2.1), the delta is smoothed enough that the coupling to the Teukolsky equation's second-order differential operators does not seed much error. Widening the delta's stencil (Sec. 3.2.2) also helps, since the fractional change in a given impulse will be less if the delta is represented by more impulses. Finally, residual high frequency noise not removed by these techniques can be taken out by convolving the Teukolsky source term with a low-pass (Gaussian) filter (Sec. 3.2.3). It's worth emphasizing that we smooth the entire source term, not just the delta function (which would arguably make our delta rather similar to the truncated Gaussian [59, 23] which this technique was designed to improve upon).

Comparison with results from the frequency-domain [36] demonstrates that the

Table 3.4: Correlation between time- and frequency-domain waveforms for the $m = 3$ mode for a range of equatorial eccentric orbits. All symbols are as in Table 3.1.

p/M	e	θ_{inc} (deg)	a/M	θ_d (deg)	h_+ corr.	$h_\times$ corr.
6.472	0.3	0	0.3	30	0.9908	0.9909
6.472	0.3	0	0.3	60	0.9922	0.9922
6.472	0.3	0	0.3	90	0.9930	0.9931
5.768	0.3	0	0.7	30	0.9934	0.9935
5.768	0.3	0	0.7	60	0.9943	0.9944
5.768	0.3	0	0.7	90	0.9948	0.9948
6.472	0.7	0	0.3	30	0.9931	0.9931
6.472	0.7	0	0.3	60	0.9905	0.9906
6.472	0.7	0	0.3	90	0.9923	0.9923
5.768	0.7	0	0.7	30	0.9928	0.9929
5.768	0.7	0	0.7	60	0.9932	0.9930
5.768	0.7	0	0.7	90	0.9920	0.9921

waveforms generated with this source term are of very high quality (Sec. 3.3). Visually, the waveforms lie on top of one another in every case that we have examined; a quantitative overlap integral demonstrates that waveforms from the two calculations are often more than 99% correlated. A key step in achieving such high quality results is to estimate the largest errors in our time-domain calculations, and then subtract that estimate from our result. We do this by performing these calculations at two different grid resolutions; under the assumption that our dominant error is quadratic in grid spacing, we then estimate the magnitude of our error (Appendix 3.6). The excellent agreement we achieve with frequency-domain results validates this approach, at least for all the cases we have considered.

So far, our main physics accomplishment is excellent agreement between time- and frequency-domain approaches to waveform calculation. It should be emphasized, however, that for waveform calculations, there will be a large set of circumstances in which time-domain codes are more efficient. For generic orbits, a frequency-domain code may require the calculation and summation of many thousand multipoles and Fourier modes. A time-domain code "automatically" sums over all modes (except the m index), so that (in principle) it is no more difficult to compute the waves from a highly inclined, highly eccentric black hole orbit than from an orbit with modest

114

Table 3.5: Correlation between time- and frequency-domain waveforms for the $m = 3$ mode for a range of inclined nearly circular orbits. All symbols are as in Table 3.1.

p/M	e	θ_{inc} (deg)	a/M	θ_d (deg)	h_+ corr.	$h_\times$ corr.
6	10^{-4}	45	0.5	60	0.9918	0.9918
6	10^{-4}	45	0.5	90	0.9907	0.9907
8	10^{-4}	45	0.5	60	0.9798	0.9798
8	10^{-4}	45	0.5	90	0.9773	0.9772
6	10^{-4}	45	0.9	60	0.9912	0.9913
6	10^{-4}	45	0.9	90	0.9905	0.9906
8	10^{-4}	45	0.9	60	0.9787	0.9790
8	10^{-4}	45	0.9	90	0.9769	0.9770
6	10^{-4}	60	0.5	60	0.9884	0.9884
6	10^{-4}	60	0.5	90	0.9876	0.9876
8	10^{-4}	60	0.5	60	0.9636	0.9640
8	10^{-4}	60	0.5	90	0.9674	0.9675
6	10^{-4}	60	0.9	60	0.9665	0.9661
6	10^{-4}	60	0.9	90	0.9680	0.9678
8	10^{-4}	60	0.9	60	0.9463	0.9473
8	10^{-4}	60	0.9	90	0.9608	0.9641

inclination and eccentricity.

The real payoff of this tool will come when we allow the source to radiatively decay, evolving through a sequence of orbits. As a demonstration that this can be done, we use a "kludged" inspiral to compute a body's inspiral, and then use that inspiral as the source for our time-domain solver in Sec. 3.4. Though not a physically accurate inspiral, this scenario shares many properties with the actual adiabatic inspiral. In particular, it demonstrates the computational advantage of a robust time-domain code for computing inspiral waveforms, given the worldline the inspiraling body follows.

Future work will address our goal of complete waveforms for the EMRI problem, in the context of the dissipation-only approximation to EMRI dynamics. We have recently extended our frequency-domain code to include the evolution of Carter's constant in the radiative backreaction limit [88], and will use this code to produce the radiation reaction data describing an inspiraling body. With this step in hand, no issue of principle stands in the way of coupling the time- and frequency-domain approaches to make usefully accurate EMRI waveforms.

Table 3.6: Correlation between time- and frequency-domain waveforms for the $m = 3$ mode for a range of generic orbits. All symbols are as in Table 3.1.

p/M	e	θ_{inc} (deg)	a/M	θ_d (deg)	h_+ corr.	$h_\times$ corr.
6	0.3	40	0.9	60	0.9917	0.9916
6	0.3	40	0.9	90	0.9915	0.9914
8	0.3	40	0.5	60	0.9801	0.9803
8	0.3	40	0.5	90	0.9785	0.9785
6	0.7	40	0.9	60	0.9906	0.9981
6	0.7	40	0.9	90	0.9899	0.9895
6	0.7	60	0.9	60	0.9862	0.9862
6	0.7	60	0.9	90	0.9819	0.9821
6	0.3	60	0.9	60	0.9790	0.9788
6	0.3	60	0.9	90	0.9840	0.9839
8	0.3	60	0.5	60	0.9788	0.9791
8	0.3	60	0.5	90	0.9747	0.9744

3.6 Appendix: Waveform extrapolation

Here we describe the variation of Richardson extrapolation which we use to estimate and eliminate the largest errors arising from our finite difference scheme. In Ref. [97], we showed that our algorithm is second order convergent. This means that we can write the solution at any given resolution as

$$\Psi_c = \Psi_t + a_1 \delta r^2 + a_2 \delta \theta^2 + a_3 \delta r \delta \theta + \mathcal{O}(\delta^3) \,, \tag{3.55}$$

where Ψ_c is the computed solution and Ψ_t is the "true" solution. The final term $\mathcal{O}(\delta^3)$ indicates that additional error terms will be third order in the grid spacing (and higher). The spatial and temporal dependencies of Ψ_c and Ψ_t have been suppressed. We now perform runs at two different resolutions, $(\delta r_1, \delta \theta_1)$ and $(\delta r_2, \delta \theta_2)$, with all other parameters fixed. The resolutions are chosen such that

$$\frac{\delta r_1}{\delta r_2} = \frac{\delta \theta_1}{\delta \theta_2} = n \,. \tag{3.56}$$

116

Neglecting higher order terms, the two results can be written

$$\Psi_{c1} \simeq \Psi_t + a_1 \delta r_1^2 + a_2 \delta \theta_1^2 + a_3 \delta r_1 \delta \theta_1 \,, \qquad (3.57)$$

$$\Psi_{c2} \simeq \Psi_t + a_1 \delta r_2^2 + a_2 \delta \theta_2^2 + a_3 \delta r_2 \delta \theta_2 \,. \qquad (3.58)$$

The relation between the two resolutions, Eq. (3.56), allows us to write

$$\Psi_{c2} = \Psi_t + 1/n^2 (a_1 \delta r_1^2 + a_2 \delta \theta_1^2 + a_3 \delta r_1 \delta \theta_1) \,. \qquad (3.59)$$

Subtracting Eq. (3.59) from Eq. (3.57) leaves us with

$$\Psi_{c1} - \Psi_{c2} = (1 - 1/n^2)(a_1 \delta r_1^2 + a_2 \delta \theta_1^2 + a_3 \delta r_1 \delta \theta_1) \,;$$

$$(3.60)$$

rearranging, we find

$$(a_1 \delta r_1^2 + a_2 \delta \theta_1^2 + a_3 \delta r_1 \delta \theta_1) = \frac{\Psi_{c1} - \Psi_{c2}}{1 - 1/n^2} \,. \qquad (3.61)$$

To the extent that neglect of higher-order errors is warranted, this estimates the largest source of error. Using Eq. (3.57) we can now estimate the "true" value:

$$\Psi_t \simeq \Psi_{c1} - (a_1 \delta r_1^2 + a_2 \delta \theta_1^2 + a_3 \delta r_1 \delta \theta_1)$$

$$= \Psi_{c1} - \frac{\Psi_{c1} - \Psi_{c2}}{1 - 1/n^2} \,. \qquad (3.62)$$

Figure 3-5 illustrates the improvement that this variant of Richardson extrapolation can yield. We plot h_+ at two different resolutions: $(\delta r_1, \delta \theta_1) = (0.04, \pi/60)$ and $(\delta r_2, \delta \theta_2) = (0.026667, \pi/90)$. We also show the extrapolated waveform, and the frequency-domain prediction. The particle is in a geodesic orbit with parameters $p = 6M$, $\theta_{\text{inc}} = 45°$, $e = 10^{-4}$ and the black hole has a spin of $a = 0.9M$. The two time-domain calculations each differ noticeably from the frequency-domain result; the extrapolated waveform by contrast agrees very well. This excellent agreement can be

117

regarded as a modified three-level convergence test, whose first two levels are the time domain waveforms and third level is the frequency domain waveform. If the code were not second order convergent, our assumption for the functional form of the errors in Eq. (3.61) would be erroneous. This would lead to a substantial disagreement between the extrapolated and frequency domain waveforms.

3.7   Acknowledgments

Figure 3-4: Waveform ($m = 2$ mode) of a small body spiraling into a massive black hole. We use "kludge" backreaction to evolve through a sequence of orbits, but compute the waves with our time-domain solver. The large black hole has spin $a = 0.5M$; the small body's orbit initially has parameters $p = 10M$, $e = 0.5$, and $\theta_{\text{inc}} = 0.5$ radians. The mass ratio of the system is $\mu/M = 0.016$. The top panel shows the full span that we simulated; the bottom two panels are zooms on early (bottom left) and late (bottom right) segments. Note the clear evolution of the wave's frequency as the orbit's mean radius shrinks.

119
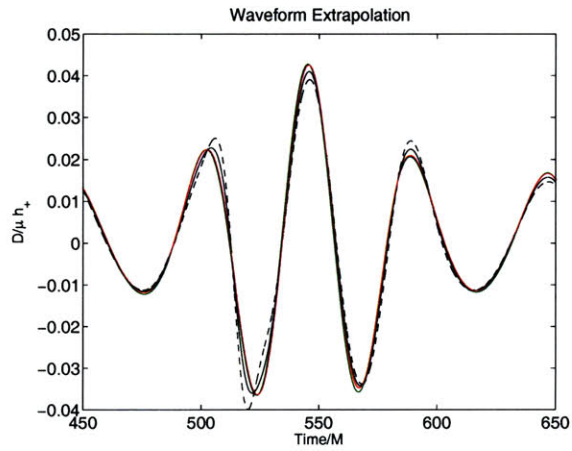
Figure 3-5: Extrapolation applied to h_+ for the $m = 3$ mode from a point particle in a nearly circular geodesic with orbital parameters $e = 10^{-4}$, $p = 6M$, and $\theta_{\mathrm{inc}} = \pi/4$ around a rotating black hole with spin $a/M = 0.9$. The dashed and solid black lines denote h_+ obtained with resolutions $(\delta r, \delta \theta) = (0.04, \pi/60)$ and $(0.026667, \pi/90)$ respectively. The solid red line is the extrapolated waveform; the solid green line is the equivalent frequency-domain waveform. Notice how well the extrapolated time-domain wave agrees with the frequency-domain result (which is nearly hidden by the red curve).

Chapter 4

The transition from adiabatic inspiral to geodesic plunge for a compact object around a massive Kerr black hole: Generic orbits

*This chapter is based on Physical Review D **77**, 124050 (2008).*

4.1 Introduction and motivation

Extreme mass ratio inspirals (EMRIs), in which stellar mass compact objects radiate gravitational energy and fall into their massive black hole companions, are promising sources of gravitational waves. LISA [67], the proposed space based gravitational wave detector should detect waves from the last stages of such inspirals. A clear theoretical understanding of the dynamics of EMRIs is vital to the detection of these gravitational waves.

The small mass ratios, which typically lie in the range $\mu/M = 10^{-5} - 10^{-8}$, allow EMRIs to be treated within the framework of perturbation theory. The trajectory of the compact object can be roughly broken into three regimes: (a) An adiabatic inspiral phase, during which the dominant inspiral mechanism arises from the radi-

ation reaction force on the smaller object. In this stage, the time scale over which the characteristic radial separation (between the compact object and its black hole companion) changes is large compared to the orbital period. This allows us to approximate the trajectory as a sequence of bound geodesics. (b) A plunge phase, during which stable geodesics do not exist. It has been shown [74] that the effect of radiation reaction is negligible during the plunge and that this phase can be modeled as a geodesic infall. (c) A regime where the spiraling compact object transitions from adiabatic inspiral to geodesic plunge. The course of motion at this juncture shows aspects of both, the self-force from radiation reaction and the effects of unstable geodesics.

In [74], Ori and Thorne introduce a method to predict the motion when the object is constrained to an approximately circular, equatorial orbit. We generalize this procedure to include inclined and eccentric trajectories. A few modifications to the prescription in [74] are introduced to handle such generic orbits. The results from our generalized prescription are in excellent agreement with [74].

The simple calculation described in this chapter is meant to serve as a stopgap for many other open and important problems. Chapters 2 and 3 described ([97, 98] and references therein) the development of a code to solve the Teukolsky equation in the time-domain. The world line of the compact object serves as an input to this code. While the world line in the adiabatic phase can be calculated from a frequency-domain based Teukolsky equation solver [53, 36], the trajectory in the transition regime for completely generic obits remains unknown. This calculation will provide the missing link needed to generate a complete inspiral trajectory.

A number of researchers are working towards solving the self-force problem exactly [80, 13]. Such an exact solution can be separated (at least qualitatively) into time-reversal symmetric and asymmetric components. The symmetric component (the "conservative self-force") conserves the integrals of motion. On the other hand, the asymmetric component (the "dissipative self-force") leads to non-zero time derivatives of the integrals of motion. Recent advances demonstrate that we are making steady progress on this problem. For example, the self force is now essentially understood

for circular orbits around Schwarzschild black holes [13]. Although approximate, the results in this chapter may serve as an independent check for these solutions. It is worth noting that if it becomes possible to include the conservative force in a simple way, we should be able to build its impact into the formalism developed here. This work may also be of interest for numerical relativity — a perturbative inspiral constructed by the techniques discussed here may be an accurate point of comparison for full numerical inspirals for small ratios (and may even be useful, if not so accurate, for mass ratios that are not strictly perturbative).

Ref. [75] discusses the transition when the compact object is in an eccentric, equatorial orbit. However, the focus of that paper is to calculate the transit time and estimate the probability for LISA to observe such a transition. Our intent is to generate the world line during the transition. We also choose our initial conditions differently than they are chosen in Ref. [75]; we discuss these differences in more detail in Sec. 4.3.

The rest of the chapter is organized as follows: Section 4.2 discusses circular orbits with arbitrary inclination. Sec. 4.3 generalizes the formalism developed in Sec. 4.2 to include eccentricity. Finally, we summarize our results in Sec. 4.4.

4.2   The transition trajectory for circular orbits

Up to initial conditions, a set of three constants, the energy, E, the component of the angular momentum along the spin axis, L_z, and the Carter constant, Q define a geodesic. The Carter constant has an approximate interpretation of being the square of the component of angular momentum perpendicular to the spin axis. As the compact object radiates, the "constants" that define its geodesic will gradually evolve. (We will refer to $[E(t), L_z(t), Q(t)]$ as the "constants", although they are slowly evolving.) A common approach to model the adiabatic regime consists of treating the motion as the sequence of geodesics [53, 36] defined by these evolving constants. As pointed out in [82], this limit amounts to a "radiative" or "dissipative" approximation. A true adiabatic approximation would be a sequence of orbits in

which each orbit included conservative self corrections. Since we currently use purely geodesic orbits as our background motion (in lieu of a self-force enhanced description), we will refer to a sequence of geodesics as an "adiabatic inspiral" throughout this chapter. Thus, within the adiabatic approximation, the world line of a particle is computed by mapping $[E(t), L_z(t), Q(t)]$ to $[r(t), \theta(t), \phi(t)]$. The symbols r, θ and ϕ are the usual Boyer-Lindquist coordinates.

In contrast, the plunge can be treated as a single unstable geodesic with almost constant E, L_z and Q. Thus, the passage from adiabatic inspiral to geodesic plunge must contain both these features — slowly evolving "constants" and marginal stability.

4.2.1 Kerr Geodesics

The following system of first order equations describes geodesics in a Kerr [14, 27] geometry:

$$\Sigma \frac{dr}{d\tau} = \pm\sqrt{R} \,, \tag{4.1}$$

$$\Sigma \frac{d\theta}{d\tau} = \pm\sqrt{V_\theta} \,, \tag{4.2}$$

$$\Sigma \frac{d\phi}{d\tau} = V_\phi \,, \tag{4.3}$$

$$\Sigma \frac{dt}{d\tau} = V_t \,. \tag{4.4}$$

The potentials can be expressed as:

$$R = \frac{1}{\mu^2} \left[E\left(a^2 + r^2\right) - aL_z \right]^2 -$$

$$+ \frac{\Delta}{\mu^2} \left[(L_z - aE)^2 + r^2\mu^2 + Q \right] , \tag{4.5}$$

$$V_\theta = \frac{1}{\mu^2} [Q -$$

$$\cos^2\theta \left(a^2 \left(\mu^2 - E^2 \right) + L_z^2 / \sin^2\theta \right)] , \tag{4.6}$$

$$V_\phi = \frac{1}{\mu} \left[L_z / \sin^2\theta - aE \right]$$

$$+ \frac{a}{\mu\Delta} \left[E\left(r^2 + a^2\right) - L_z a \right] , \tag{4.7}$$

$$V_t = \frac{1}{\mu} \left[a \left(L_z - aE\sin^2\theta \right) \right] +$$

$$\frac{r^2 + a^2}{\mu\Delta} \left[E\left(r^2 + a^2\right) - L_z a \right] . \tag{4.8}$$

The parameters (r, θ, ϕ, t) are the Boyer-Lindquist coordinates, M is the black hole mass, μ is the perturbing mass, $\Sigma = r^2 + a^2 \cos^2\theta$, $\Delta = r^2 - 2Mr + a^2$ and a is the spin parameter of the black hole. The constants (E, L_z, Q) represent the *actual* energy, momentum and Carter constant (in units of M, M^2 and M^4 respectively), not the dimensionless versions of them. By introducing the perturbing mass explicitly, our notation deviates from previous literature. We do this in order to show the dependence of the transition phase on the mass of the perturbing object. We also set $G = c = 1$ everywhere.

4.2.2 The last stable orbit

A standard but not unique definition of the "inclination" of a Kerr geodesic is given by

$$\cos\iota = \frac{L_z}{\sqrt{L_z^2 + Q}} , \tag{4.9}$$

$$\Rightarrow Q = \frac{L_z^2}{\cos^2\iota} - L_z^2 . \tag{4.10}$$

125

It is possible to use ι to eliminate the Carter constant. Thus, any circular orbit can be parametrized by its radius (r) and inclination (ι).

The last stable orbit (LSO) serves as an important reference point — the inspiral is adiabatic well before the compact object crosses the LSO and is approximately a plunge well after the crossing. Since the transition occurs in the vicinity of the LSO, a preliminary step in our computation is to determine r and (E, L_z, Q) at the LSO for a given inclination at the LSO, ι_{LSO}. Note that ι changes with time because it is a function of $[E(t), L_z(t), Q(t)]$.

Circular orbits satisfy

$$R = 0 \text{ and} \tag{4.11}$$

$$R' = \frac{dR}{dr} = 0 \ . \tag{4.12}$$

We must have $R' = 0$ because the LSO lies at an extremum of R. We also require that

$$R'' = \frac{d^2 R}{dr^2} > 0 \ , \tag{4.13}$$

for the extremum to be stable. This implies that the orbit will be marginally stable if $R'' = 0$. Thus, the three equations $R = R' = R'' = 0$ can be numerically solved for a given ι_{LSO} to yield r, E, L_z and Q at the LSO.

4.2.3 The constants in the transition regime

We need a model of the phase space trajectory, $[E(t), L_z(t), Q(t)]$ near the LSO in order to compute the world line of the compact object as it transitions from inspiral

to plunge. To this end, we Taylor expand about the LSO to obtain

$$E(t) \simeq E_{\mathrm{LSO}} + (t - t_{\mathrm{LSO}})\dot{E}_{\mathrm{LSO}} ,\tag{4.14}$$

$$L_z(t) \simeq L_{z,\mathrm{LSO}} + (t - t_{\mathrm{LSO}})\dot{L}_{z,\mathrm{LSO}} ,\tag{4.15}$$

$$Q(t) \simeq Q_{\mathrm{LSO}} + (t - t_{\mathrm{LSO}})(\dot{Q}_{\mathrm{LSO}} + \delta\dot{Q})$$

$$+\delta Q ,\tag{4.16}$$

which are natural generalizations of equations (3.4) and (3.5) of Ref. [74]. The overdot denotes differentiation with respect to t. We will later see that our initial condition for t amounts to choosing t_{LSO}, the instant at which the compact object crosses the LSO. This choice is consistent with the procedure in Ref. [74] — Eq. (3.14) of Ref. [74] implies a choice of $t_{\mathrm{LSO}} = 0$.

The constant terms in Eq. (4.16), δQ and $\delta\dot{Q}$, are needed to guarantee that the trajectory remains circular as we enter the transition. As the notation suggests, these constants are small compared to Q_{LSO} and $\dot{Q}_{\mathrm{LSO}}$. They are discussed in more detail when we discuss initial conditions for the transition in Sec. 4.2.6.

The expressions (4.14), (4.15) and (4.16) do not include conservative effects of the self force. Pound and Poisson [82] have demonstrated that this omission will lead to observationally significant changes. Inclusion of these effects would effectively alter the potentials, Eq. (4.5) - Eq. (4.8) leading to slight deviations of $(E, L_z, Q)_{\mathrm{LSO}}$ and r_{LSO} (for a given ι_{LSO}) from their geodesic values. The exact impact of these effects will not be known until we know what the corrections are. We will later see that our results possess all the expected qualitative features despite this handicap. Moreover, the prescription in [74] and its generalization presented here can easily incorporate these effects once they are known.

The fluxes at the LSO remain a parameter in our code. We use the code developed in [53] to provide us the dimensionless fluxes, $(M/\mu)^2\dot{E}$, $(M/\mu^2)\dot{L}_z$ and $(1/\mu^3)\dot{Q}$ at the LSO. Equivalently, we can use the expressions in [43] (with zero eccentricity) for the dimensionless fluxes.

127

4.2.4 Reparametrization of the θ-equation

Numerical integration of the θ-equation warrants some care. The issue arises because $d\theta/dt$ vanishes at the turning points, $\theta_{\max}$ and $\theta_{\min}$, where

$$0 \leq \theta_{\min} \leq \theta_{\max} \leq \pi \ . \tag{4.17}$$

The potential problems posed by the turning points can be eliminated by reparametrizing θ. Following Ref. [53], we use

$$z = \cos^2\theta \ = \ z_- \cos^2\chi \ , \tag{4.18}$$

where

$$\beta(z - z_+)(z - z_-) \ = \ \beta z^2 - z\frac{Q + L_z^2 + a^2(\mu^2 - E^2)}{\mu^2}$$
$$+ \frac{Q}{\mu^2} \ , \tag{4.19}$$

and $\beta = a^2(\mu^2 - E^2)/\mu^2$. The θ-equation of motion now becomes

$$\frac{d\chi}{dt} = \frac{\sqrt{\beta(z_+ - z)}}{\gamma + a^2 E z(\chi)/\mu} \ , \tag{4.20}$$

where

$$\gamma = \frac{E}{\mu}\left[\frac{(r^2 + a^2)^2}{\Delta} - a^2\right] - \frac{2MraL_z}{\Delta\mu} \ . \tag{4.21}$$

Equation (4.20) can now be integrated without turning points because χ varies from 0 to π to 2π as θ varies from $\theta_{\min}$ to $\theta_{\max}$ and back to $\theta_{\min}$.

4.2.5 The prescription

In keeping with our main objective of obtaining the world line $[r(t), \theta(t), \phi(t)]$ through the transition regime, we eliminate τ by dividing equation (4.1) by (4.4) and squaring

128

the result to obtain

$$\left(\frac{dr}{dt}\right)^2 = \frac{R(r,\chi)}{V_t(r,\chi)^2} \equiv F \ .$$ (4.22)

One more time derivative gives the acceleration:

$$\frac{d^2r}{dt^2} = \frac{1}{2}\left[\frac{\partial}{\partial r}\left(\frac{R}{V_t^2}\right) + \frac{\partial}{\partial \chi}\left(\frac{R}{V_t^2}\right)\frac{d\chi/dt}{dr/dt}\right] \ .$$ (4.23)

Ideally, Eq. (4.23) must have other additive terms proportional to non-zero powers of μ. This is analogous to Eq. (3.10) of [74]. Excluding this term amounts to ignoring the conservative self force.

Since the transition phase is in the proximity of the LSO, we can Taylor expand F about r_{LSO}, E_{LSO}, $L_{z,\text{LSO}}$ and Q_{LSO} to obtain

$$\begin{aligned}
F(r, L_z, E, \chi, \iota) \simeq{} & \frac{1}{6}\frac{\partial^3 F}{\partial r^3}\bigg|_{\text{LSO}}(r - r_{\text{LSO}})^3 + \frac{\partial^2 F}{\partial r \partial L_z}\bigg|_{\text{LSO}}(L_z - L_{z,\text{LSO}})(r - r_{\text{LSO}}) \\
& + \frac{\partial^2 F}{\partial r \partial E}\bigg|_{\text{LSO}}(E - E_{\text{LSO}})(r - r_{\text{LSO}}) \\
& + \frac{\partial^2 F}{\partial r \partial Q}\bigg|_{\text{LSO}}(Q - Q_{\text{LSO}})(r - r_{\text{LSO}}) \ .
\end{aligned}$$ (4.24)

Thus, the acceleration now becomes [1]:

$$\begin{aligned}
\frac{d^2r}{dt^2} ={} & \frac{1}{2}\left[\frac{1}{2}\frac{\partial^3 F}{\partial r^3}\bigg|_{\text{LSO}}(r - r_{\text{LSO}})^2 + \frac{\partial^2 F}{\partial r \partial L_z}\bigg|_{\text{LSO}}(L_z - L_{z,\text{LSO}}) + \frac{\partial^2 F}{\partial r \partial E}\bigg|_{\text{LSO}}(E - E_{\text{LSO}})\right. \\
& \left. + \frac{\partial^2 F}{\partial r \partial Q}\bigg|_{\text{LSO}}(Q - Q_{\text{LSO}}) + \frac{\partial F}{\partial \chi}\frac{d\chi/dt}{dr/dt}\right] \ .
\end{aligned}$$ (4.25)

We have not expanded the second term in Eq. (4.23) because we do not know the value of χ at $r = r_{\text{LSO}}$ a priori. Similarly, the ϕ-equation takes the form

$$\frac{d\phi}{dt} = \frac{V_\phi(r,\chi)}{V_t(r,\chi)} \ .$$ (4.26)

---

[1]Note that Eq. (4.24) ignores terms of order $(\mu/M)^2$ and higher.

129

The trajectory in the transition phase can now be computed by integrating equations (4.25), (4.26) and (4.20) from some starting point outside the LSO to some ending point inside the LSO, for a given ι_{LSO}, with time varying E, L_z and Q.

4.2.6 Initial conditions

The angles, ϕ and χ can be set to zero without loss of generality. Setting $\chi = 0$ corresponds to starting the inspiral at $\theta = \theta_{\min}$.

The choice of initial radius depends explicitly on μ. In Ref. [74], the authors define parameters, α, β, κ, τ_0 and R_0. These are used to scale out the perturbing mass from the equation of motion and initial conditions. Although we prefer to retain dimensions in the equations of motion, we specify initial conditions in a dimensionless form, independent of μ. This will be useful be in interpreting our results and making comparisons with Ref. [74]. Following Ref. [74], we define

$$X = \left(\frac{\mu}{M}\right)^{2/5} \frac{r - r_{\mathrm{LSO}}}{R_0} \,, \tag{4.27}$$

$$R_0 = (\beta\kappa_0)^{2/5}\alpha^{-3/5} \,, \tag{4.28}$$

$$T = \left(\frac{\mu}{M}\right)^{1/5} \frac{\tilde{t} - \tilde{t}_{\mathrm{LSO}}}{\tau_0} \left.\frac{d\tau}{dt}\right|_{\mathrm{LSO}} \,, \tag{4.29}$$

where

$$\alpha = -\frac{1}{4} \frac{\partial^3}{\partial\tilde{r}^3}\left[\frac{R}{\Sigma^2}\right]_{\mathrm{LSO}} \,, \tag{4.30}$$

$$\beta = \frac{1}{2}\left[\frac{\partial^2}{\partial\tilde{L}_z\partial\tilde{r}}\left(\frac{R}{\Sigma^2}\right) + \frac{\dot{\tilde{E}}}{\dot{\tilde{L}}_z}\frac{\partial^2}{\partial\tilde{E}\partial\tilde{r}}\left(\frac{R}{\Sigma^2}\right)\right.$$
$$\left.+\frac{\dot{\tilde{Q}}}{\dot{\tilde{L}}_z}\frac{\partial^2}{\partial\tilde{Q}\partial\tilde{r}}\left(\frac{R}{\Sigma^2}\right)\right]_{\mathrm{LSO}} \tag{4.31}$$

$$\kappa(t) = -\frac{1}{\mu/M}\frac{d\tilde{L}_z}{d\tilde{\tau}} = -\frac{d\tilde{L}_z/d\tilde{t}}{(\mu/M)(d\tau/dt)} \,, \tag{4.32}$$

$$\kappa_0 = \kappa|_{\mathrm{LSO}} \,, \tag{4.33}$$

$$\tau_0 = (\alpha\beta\kappa_0)^{-1/5} \,, \tag{4.34}$$

130

with $\tilde{r} = r/M$, $\tilde{t} = t/M$, $\tilde{E} = E/\mu$, $\tilde{L}_z = L_z/(\mu M)$ and $\tilde{Q} = Q/(\mu M)^2$.

These definitions reduce to those presented in Ref. [74] when $\iota = 0$. It is useful to observe that κ does not scale with μ. We evaluate $d\tau/dt$, α, β and κ_0 at $\theta = \pi/2 - \iota_{\text{LSO}}$ because we do not know θ_{LSO} a priori. Notice that X and T are dimensionless.

The smoothness of the transition implies that there is no fixed instant at which the transition starts or ends. Motivated by the choices in Ref. [74], we set $T \simeq -1$ at $t = 0$ and stop the numerical integrator when $X \leq X_e = -5$.

In summary, our initial conditions are $T = -1$, $\phi = 0$ and $\chi = 0$ at $t = 0$ [2]. Setting $T = -1$ at $t = 0$ allows us to calculate t_{LSO} and hence $E(0)$ and $L_z(0)$ from equations (4.14) and (4.15). We then solve $R(E, L_z, Q, r) = 0$ and $dR/dr = 0$ to obtain $r(0)$ and $Q(0)$. This is analogous to Sec. IIIC of Ref. [74] where they enforce $X = \sqrt{-T}$ to determine X at $t = 0$.

The trajectory is adiabatic before the start of the transition. At $t = 0$, we must impose the condition [58, 87, 71] that circular orbits remain circular even under adiabatic radiation reaction. Thus, requiring that $\dot{R} = dR/dt = 0$ and $\dot{R}' = d^2R/drdt = 0$ leads to expressions (3.5) and (3.6) of [53] for $\dot{r}(0)$ and $\dot{Q}(0)$ respectively.

We can now substitute $Q(0)$ and $\dot{Q}(0)$ in Eq. (4.16) to obtain two independent equations,

$$Q(0) = Q_{\text{LSO}} - t_{\text{LSO}}(\dot{Q}_{\text{LSO}} + \delta\dot{Q}) + \delta Q \text{ and} \qquad (4.35)$$

$$\dot{Q}(0) = \dot{Q}_{\text{LSO}} + \delta\dot{Q}, \qquad (4.36)$$

which can be used to evaluate δQ and $\delta\dot{Q}$.

4.2.7 Code algorithm and numerical results

The previous sections developed the steps required to calculate the compact body's trajectory as it transitions from inspiral to plunge. We now summarize the algorithm that was actually used to implement this prescription:

---

[2]It is important to keep $|T|$ small enough that our Taylor expansion about the LSO remains a valid approximation.

(1) Take ι_{LSO} as input.

(2) Compute E and L_z at the LSO.

(3) Obtain $\dot{E}$ and $\dot{L}_z$ at the LSO from the code developed in [53]. We may also use the expressions in [43] (which reduce to the results in [53] for circular orbits), which will be particularly useful when we generalize to eccentric orbits.

(4) Choose initial conditions $T \simeq -1$, $\phi = 0$ and $\chi = 0$ at $t = 0$.

(5) Calculate $E(0)$ and $L_z(0)$ from equations (4.14) and (4.15).

(6) Solve for $r(0)$ and $Q(0)$ by imposing $R = 0$ and $dR/dr = 0$ at $t = 0$.

(7) Compute $\dot{r}(0)$ and $\dot{Q}(0)$ from equations (3.5) and (3.6) of [53].

(8) Substitute $Q(0)$ and $\dot{Q}(0)$ in Eq. (4.16) to evaluate δQ and $\delta \dot{Q}$.

(9) Use a Runge-Kutta integrator on (4.25), (4.26) and (4.20) to compute the coordinates at the next step. A time step of $\delta t \simeq 0.05 M$ works well.

(10) Update the "constants", $E_{i+1} = E_i + \dot{E}\delta t$, $L_{z,i+1} = L_{z,i} + \dot{L}_z \delta t$ and $Q_{i+1} = Q_i + (\dot{Q} + \delta \dot{Q})\delta t$. The subscript i denotes a discrete time instant.

(11) Repeat steps (9)-(11) until $X(t) \simeq -5$.

The primary objective of this calculation is to compute the world line of the compact object during the transition. Figures 4-1 and 4-2 illustrate r, θ and ϕ motions of the compact object for a typical set of parameters. We also show a plunging geodesic matched to the end of the transition.

Table 4.1 shows the parameters and transit times for a range of inclination angles. In general, we find that the transit time increases with inclination. However, the dimensionless transit time ΔT remains approximately constant,

$$\Delta T \simeq 3.3 - 3.4 \,, \tag{4.37}$$

when $X_e = -5$ for all values of a and ι. Again, this is a consistent generalization of the result in Ref. [74] where they find $\Delta T \simeq 3.3$ for all circular, equatorial orbits.

132

Table 4.1: Fluxes and transit times for different inclinations. We set $a = 0.5M$, $\mu = 10^{-6}M$, $M = 1$, $T(0) = -1$ and $X_e = -5$.

$\iota^{\circ}_{\mathrm{LSO}}$	$\frac{r_{\mathrm{LSO}}}{M}$	$\frac{M^2}{\mu^2}\dot{E}_{\mathrm{LSO}}$	$\frac{M}{\mu^2}\dot{L}_{z,\mathrm{LSO}}$	$\frac{1}{\mu^3}\dot{Q}_{z,\mathrm{LSO}}$	α	β	R_0	κ_0	τ_0	t/M	ΔT
10^{-3}	4.23	−0.00457	−0.0422	−0.000572	0.00311	0.0327	0.0699	0.603	2.80	944.9	3.36
10	4.26	−0.00446	−0.0409	−0.00684	0.00304	0.0327	0.0677	0.604	2.81	952.4	3.36
20	4.32	−0.00415	−0.0375	−0.0241	0.00284	0.0329	0.0615	0.610	2.82	974.9	3.36
30	4.43	−0.00368	−0.0323	−0.0481	0.00254	0.0333	0.0523	0.618	2.84	1012.6	3.36
40	4.59	−0.00314	−0.0262	−0.0733	0.00219	0.0342	0.0416	0.630	2.86	1065.9	3.35
50	4.78	−0.002594	−0.0198	−0.0946	0.00184	0.0363	0.0309	0.643	2.88	1134.4	3.35
60	5.01	−0.00208	−0.0139	−0.108	0.00152	0.0403	0.0211	0.657	2.90	1217.9	3.35

Table 4.2: Variation of transit time with perturbing mass, μ/M. We set $a = 0.9M$, $\iota_{\mathrm{LSO}} = 0.001^{\circ}$, $M = 1$, $T_s = -1$ and $X_e = -5$. Note that $r_{\mathrm{LSO}} = 2.32M$.

μ/M	t/M	ΔT
10^{-3}	118.9	3.449
10^{-4}	185.6	3.397
10^{-5}	292.2	3.375
10^{-6}	461.9	3.367
10^{-7}	731.3	3.363
10^{-8}	1158.6	3.362

4.2.8 Comparison with Ref. [74]

The results in Ref. [74] provide an important sanity check for the case of circular, equatorial orbits. However, we have to account for the minor differences between the two approaches. Ref. [74] makes the approximations

$$\frac{d\phi}{dt} \simeq \left.\frac{d\phi}{dt}\right|_{\mathrm{ISCO}} \quad \text{and} \tag{4.38}$$

$$\frac{d\tau}{dt} \simeq \left.\frac{d\tau}{dt}\right|_{\mathrm{ISCO}}, \tag{4.39}$$

which lead to

$$\kappa = -\frac{d\tilde{L}_z/d\tilde{t}}{(\mu/M)(d\tau/dt)},$$

$$\simeq -\frac{d\tilde{L}_z/d\tilde{t}|_{\mathrm{ISCO}}}{(\mu/M)(d\tau/dt)_{\mathrm{ISCO}}}, \tag{4.40}$$

133
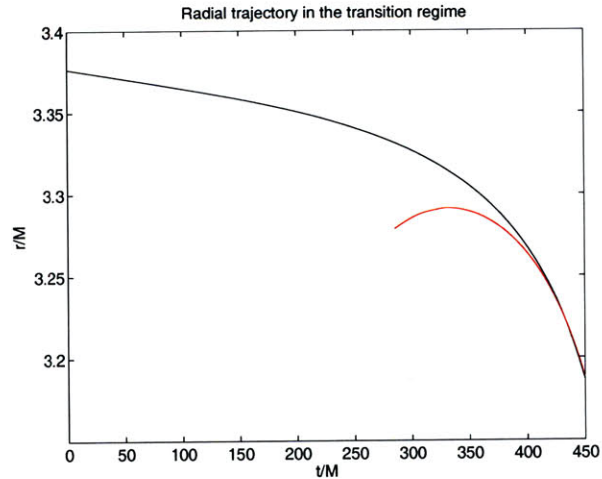
Figure 4-1: Radial trajectory during the transition (black line) from inspiral to plunge for a compact object of mass $\mu = 10^{-5}M$ in a nearly circular orbit around a black hole with spin $a = 0.8M$. The compact object crosses the LSO at time $t_{\mathrm{LSO}} = 137.5M$. The inclination of the orbit at t_{LSO} is $\iota_{\mathrm{LSO}} = 37°$. The red line is a plunging geodesic matched to the end of the transition.

which is a dimensionless constant. In our prescription, $d\tau/dt$ varies with time. This time dependence has to be enforced because $d\tau/dt$ is a function of θ, whose value at the LSO is not known a priori. The circular, equatorial case in Ref. [74] does not suffer from this pathology because $\theta = \pi/2$ at all times. Thus, we treat κ as a slowly varying function of time. Table 4.2 shows the transit times for a nearly equatorial orbit ($\iota_{\mathrm{LSO}} = 0.001$) and a range of mass ratios. As the mass ratio becomes smaller, the variation in κ decreases, and the dimensionless transit time converges to the limit where κ is constant.

Our initial conditions differ slightly from those used in Ref. [74]. Effectively, they use the Taylor expansion of $R(r)$ to solve $dR/dr = 0$ and $d^2R/(dr\,dt) = 0$ for $r(0)$ and $\dot{r}(0)$ respectively. In contrast, we solve the equations exactly. This leads to differences of less than 1%.

134
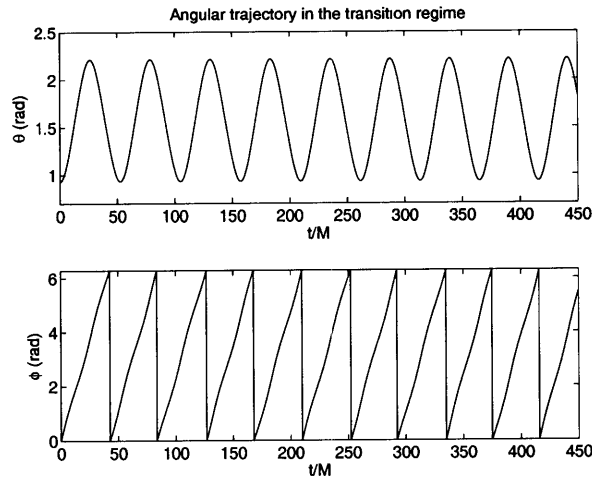
Figure 4-2: Angular motion during the transition for a compact object around a spinning black hole with identical parameters as in Fig. 4-1.

4.3 Eccentric orbits

The methods developed thus far only discussed circular orbits. We now extend this technique to include non-zero eccentricity. In the absence of radiation reaction, the geodesic equations admit bound eccentric orbits. These orbits are conventionally parametrized by the semi-latus rectum, p, and the eccentricity, e. The radial coordinate can now be expressed as

$$r(t) = \frac{p}{1 + e \cos \psi(t)} \, . \tag{4.41}$$

The angle $\psi(t)$ is analogous to the eccentric anomaly and can be solved for numerically. The geodesic has turning points at $\psi = 0, \pi$. Deep in the adiabatic inspiral, the compact object's trajectory is well approximated by a sequence of orbits with slowly varying $p(t)$ and $e(t)$.

Geodesics beyond the LSO do not have turning points (where $dr/dt = 0$). This changes the situation considerably because the parameters, p and e are not well-defined anymore. Thus, the trajectory ceases to have turning points somewhere during the transition from inspiral to plunge. We will later show that this feature is

135

naturally buried in our model of the transition.

4.3.1 The last stable orbit

As with circular orbits, the last stable bound geodesic is an important reference in our procedure. The inner and outer turning points (r_{min} and r_{max}) of the LSO are related to e_{LSO} and p_{LSO} through

$$r_{\text{min}} = \frac{p_{\text{LSO}}}{1 + e_{\text{LSO}}} \text{ and} \qquad (4.42)$$

$$r_{\text{max}} = \frac{p_{\text{LSO}}}{1 - e_{\text{LSO}}} . \qquad (4.43)$$

Our goal is to determine p_{LSO} and the constants (E, L_z, Q) at the LSO for a given ι_{LSO} and e_{LSO}. This can be achieved by requiring that

$$\frac{dR}{dr} = 0 \text{ at } r = r_{\text{min}} , \qquad (4.44)$$

$$R = 0 \text{ at } r = r_{\text{min}} \text{ and } r = r_{\text{max}} . \qquad (4.45)$$

Recall that the function R is given by Eq. (4.5) and ι is defined by Eq. (4.9). We require Eq. (4.44) to be satisfied because the inner most turning point corresponds to a local maximum of $(-R)$. Equation (4.45) enforces the compact object's velocity to vanish at the turning points. Equations (4.44), (4.45) and (4.9) can be solved numerically for p and (E, L_z, Q) at the LSO. Sec. 4.5 describes the details of this numerical procedure.

4.3.2 The constants during the transition

As with the circular case, our initial conditions are such that we effectively *choose* the LSO crossing to occur at $t = t_{\text{LSO}}$. This allows us to expand the constants about

136

the LSO to obtain

$$E(t) \simeq E_{\text{LSO}} + (t - t_{\text{LSO}})\dot{E}_{\text{LSO}} ,$$ (4.46)

$$L_z(t) \simeq L_{z,\text{LSO}} + (t - t_{\text{LSO}})\dot{L}_{z,\text{LSO}} ,$$ (4.47)

$$Q(t) \simeq Q_{\text{LSO}} + (t - t_{\text{LSO}})\dot{Q}_{\text{LSO}}.$$ (4.48)

Notice that we no longer need the corrections, δQ and $\dot{\delta Q}$ because there are no additional symmetries to constrain $Q(0)$ and $\dot{Q}(0)$; $E(t)$, $L_z(t)$ and $Q(t)$ are independent.

As discussed in Sec. 4.2.3, equations (4.46), (4.47) and (4.48) do not include conservative effects of the self force. Just as the circular case, this will lead to a slight shift of $(E, L_z, Q)_{\text{LSO}}$ and p_{LSO} (for a given e_{LSO} and ι_{LSO}) with respect to their geodesic values. Again, our motivation to stick with this approximation stems from the facts that: (a) These effects can be incorporated into our prescription once they are known, and (b) Our results show the generally expected behavior, at least qualitatively.

Numerical methods to calculate the change in the Carter constant due to gravitational-wave backreaction have recently become available [89, 88]. Work is in progress implementing that result in the code we use to compute the rate of change of orbital constants [33]. For now, we use the approximate expressions for $\dot{Q}$ described in [43]; it will be a simple matter to update our code when more accurate $\dot{Q}$ results are available.

4.3.3   The prescription for eccentric orbits

Our next task is to derive equations of motion to map the phase space trajectory to an actual world line. The angular equations, Eq. (4.20) and Eq. (4.26), remain unaffected. Our strategy for the radial equation is to expand the geodesic equation about $(E_{\text{LSO}}, L_{z,\text{LSO}}, Q_{\text{LSO}})$. This leaves us with

$$\frac{d^2r}{dt^2} = \frac{1}{2}\left[\frac{\partial F}{\partial r} + \frac{\partial F}{\partial \chi}\frac{d\chi/dt}{dr/dt}\right], \tag{4.49}$$

$$\frac{\partial F}{\partial r} \simeq \left[\frac{\partial^2 F}{\partial r \partial E}\bigg|_{\text{LSO}}(E - E_{\text{LSO}}) + \frac{\partial^2 F}{\partial r \partial L_z}\bigg|_{\text{LSO}}(L_z - L_{z,\text{LSO}}) + \frac{\partial^2 F}{\partial r \partial Q}\bigg|_{\text{LSO}}(Q - Q_{\text{LSO}})\right.$$
$$\left. + \frac{\partial F}{\partial r}(r, \chi; E_{\text{LSO}}, L_{z,\text{LSO}}, Q_{\text{LSO}})\right]. \tag{4.50}$$

Note that we only expand about the constants, not the r-coordinate, because there is no unique r at the LSO. In the absence of the first three terms in Eq. (4.50), the equation of motion is simply a geodesic at the LSO. This is consistent with our intuitive notion of "expanding about the LSO". The existence of turning points presents a complication while integrating Eq. (4.50) numerically. We present a method to tackle this in Sec. 4.6.

4.3.4 Initial conditions

We need initial conditions for r and dr/dt before we start the numerical integrator. Motivated by the initial conditions for circular orbits, we set $T \simeq -1$ at $t = 0$. This amounts to choosing t_{LSO}. We can now determine $[E(0), L_z(0), Q(0)]$, which can be mapped to (p, e, ι) at $t = 0$. This mapping is allowed because the trajectory is adiabatic before $t = 0$. The coordinates at any point on the geodesic defined by $[E(0), L_z(0), Q(0)]$ can serve as our initial conditions. For simplicity, we choose

$$r(0) = \frac{p}{1 + e}, \tag{4.51}$$

$$\frac{dr}{dt}(0) = 0, \tag{4.52}$$

$$\phi(0) = 0, \tag{4.53}$$

$$\chi(0) = 0. \tag{4.54}$$

The equations of motion can now be easily integrated across the LSO.

4.3.5 Code implementation and numerical results

Taking eccentricity into account changes our algorithm slightly. We summarize the code's algorithm as follows:

(1) Take ι_{LSO} and e_{LSO} as input.

(2) Compute E, L_z and Q at the LSO.

(3) Obtain $\dot{E}$, $\dot{L}_z$ and $\dot{Q}$ at the LSO from the expressions in Ref. [43].

(4) Choose initial conditions $T \simeq -1$, $\phi = 0$ and $\chi = 0$ at $t = 0$.

(5) Calculate $E(0)$, $L_z(0)$ and $Q(0)$ from equations (4.46), (4.47) and (4.48).

(6) Map $[E(0), L_z(0), Q(0)]$ to (p, e, ι).

(7) Set $r = p/(1 + e)$ and $dr/dt = 0$ at $t = 0$.

(8) Use a Runge-Kutta integrator on (4.49), (4.26) and (4.20) to compute the coordinates at the next step. A time step of $\delta t \simeq 0.05M$ works well.

(10) Update the "constants", $E_{i+1} = E_i + \dot{E}_{LSO}\delta t$, $L_{z,i+1} = L_{z,i} + \dot{L}_{z,LSO}\delta t$ and $Q_{i+1} = Q_i + \dot{Q}_{LSO}\delta t$. The subscript i refers to a discrete time instant.

(11) Repeat steps (9)-(11) until $X \simeq -5$.

Recall that the local minimum of the potential R is less than zero for bound orbits and is greater than zero for a plunging geodesic. The minimum is exactly zero at the LSO. These conditions can be used as sanity checks while performing the numerical integration.

Table 4.3: Fluxes and transit times for different eccentricities. We set $a = 0.8M$, $\mu = 10^{-5}$, $\iota_{LSO} = 45°$, $M = 1$, $T_s = -1$ and $X_e = -5$.

ι_{LSO}°	$\frac{r_{LSO}}{M}$	$\dot{E}_{LSO}$ (M^2/μ^2)	$\dot{L}_{z,LSO}$ (M/μ^2)	$\dot{Q}_{z,LSO}$ $(1/\mu^3)$	α	β	R_0	κ_0	τ_0	t/M	ΔT
10^{-4}	3.58	-0.00974	-0.0619	-0.153	0.00517	0.0530	3.04	0.113	7.98	486.3	3.34
0.1	3.70	-0.00857	-0.0545	-0.136	0.00351	0.0506	3.54	0.0969	8.97	448.2	2.81
0.2	3.84	-0.00795	-0.0479	-0.120	0.00220	0.0484	4.33	0.0832	10.2	373.5	2.10
0.3	3.96	-0.00751	-0.0419	-0.105	0.00117	0.0463	5.83	0.0714	12.1	341.1	1.66
0.4	4.09	-0.00693	-0.0361	-0.0900	0.000365	0.0442	10.8	0.0604	15.9	332.7	1.25
0.5	4.22	-0.00607	-0.0300	-0.0745	-0.000280	0.0420	11.5	0.0496	17.7	331.6	1.14
0.6	4.35	-0.00450	-0.0236	-0.0582	-0.000801	0.0401	5.41	0.0385	15.2	338.8	1.37
0.7	4.49	-0.00351	-0.0168	-0.0413	-0.00123	0.0381	3.57	0.0272	15.1	381.9	1.56
0.8	4.62	-0.00206	-0.0100	-0.0245	-0.00159	0.0362	2.44	0.0162	16.1	507.2	1.95

Figures 4-3, 4-4 and 4-5 show a typical trajectory during the transition from inspiral to plunge. The compact object starts at the minimum of the last bound geodesic

before the plunge. The radial coordinate increases until it reaches a maximum where $R = dr/dt = 0$. Subsequently, it turns around and heads toward the minimum. After executing a number of "whirls" near the minimum, the trajectory becomes unstable, and thus plunges into the central black hole. The whirls are evident from the angular trajectory plotted in Fig. 4-5. We also show a plunging geodesic matched to the end of the transition. Notice that the plunge spends quite a bit of time at $r \sim 2.8M$ — much more time than the transition trajectory. This is because the radiation emission built into the transition trajectory's construction pushes it off this marginally stable orbit rather quickly.

Table 4.3 shows the various parameters and transit times for a range of eccentricities. Note that the parameters α, β, R_0, κ_0 and τ_0 (which are defined in Sec. 4.2.6) are evaluated at p_{LSO}. In general, we find that the transit time is proportional to α. This is not surprising because α is the first term in the Taylor expansion of the potential, R. We also observe some degree of correlation between the transit time and τ_0, the parameter used to define the dimensionless time.
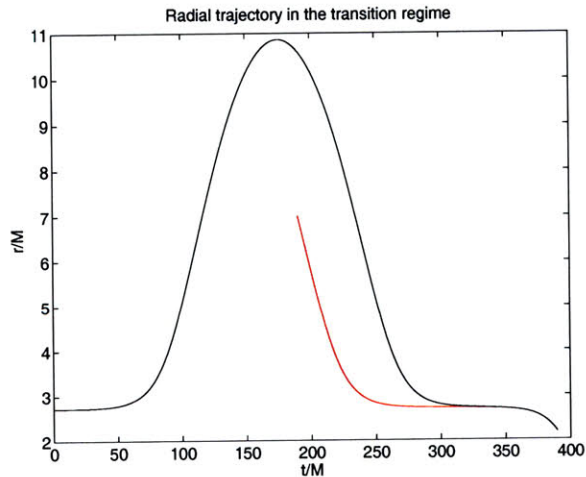
Figure 4-3: Radial trajectory during the transition (black line) from inspiral to plunge for a compact object of mass $\mu = 10^{-6}M$ in an eccentric orbit around a black hole with spin $a = 0.8M$. The compact object crosses the LSO at time $t_{\text{LSO}} = 196.7M$. The inclination and eccentricity of the orbit at t_{LSO} are $\iota_{\text{LSO}} = 45°$ and $e_{\text{LSO}} = 0.6$ respectively. The red line is an unstable geodesic matched to the end of the transition.
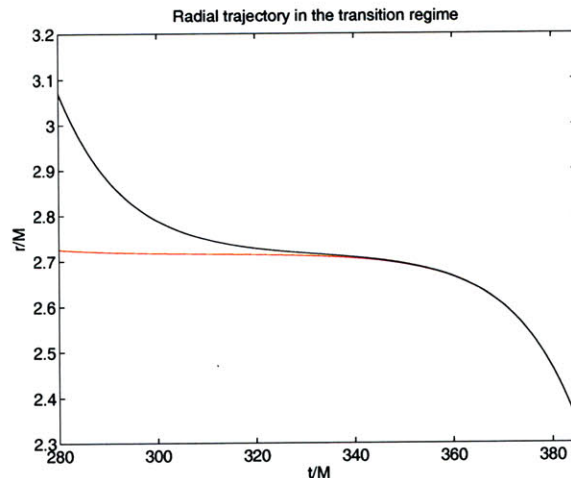
Figure 4-4: Same as Fig. 4-3, but zooming in on the final "whirls".

4.3.6 Comparison with Ref. [75]

As mentioned in the introduction, there are differences between our generalized prescription and the method developed in Ref. [75], which only models the transition when the compact object is in an eccentric, equatorial orbit. First, we set our initial conditions at the *start* of the LSO, whereas Ref. [75] sets the initial conditions at the *end* of the LSO. This educated choice allows Ref. [75] to derive an analytic form for the trajectory. Second, we differ in the choice of final conditions. [3]

In attempting to make comparisons with Ref. [75], we found a number of typographical errors. Thus, we extract the essence of the calculation in Ref. [75] and present it in a form that (hopefully) makes the errors obvious. We start by expressing the radial geodesic equation as

$$\left(\frac{dr}{d\tau}\right)^2 + V(r) = 0 , \tag{4.55}$$

where

$$V(r) = -\frac{R}{\Sigma^2} . \tag{4.56}$$

---

[3]See Sec. IID*3* and Ref. [20] of Ref. [75] for a description their choice of final conditions.
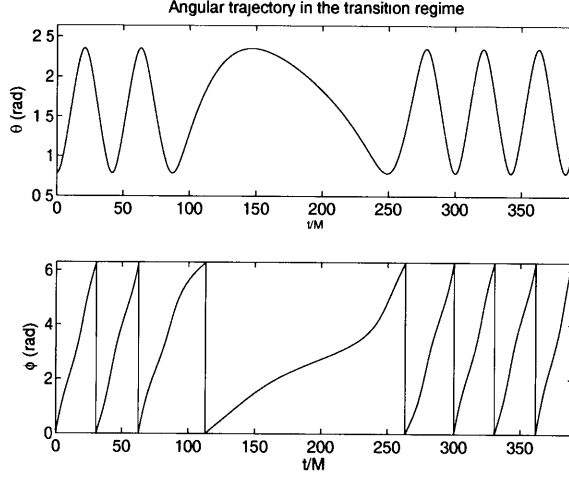
Figure 4-5: Angular trajectory during the transition for the same set of parameters as in Fig. 4-3.

The orbit is unstable if the local maximum of $V(r)$ is negative. Define

$$I = -\text{Max}\{V(r)\} = -V(r_{\text{max}}) \ . \tag{4.57}$$

Note that this implies $V'(r_{\text{max}}) = 0$ and $V''(r_{\text{max}}) < 0$. We Taylor expand Eq. (4.55) about the maximum of $V(r)$ corresponding to some (E, L_z, Q) just beyond the LSO to get

$$\left(\frac{d(\delta r)}{d\tau}\right)^2 + V(r_{\text{max}}) + \delta r V'(r_{\text{max}})$$

$$+\frac{1}{2}\delta r^2 V''(r_{\text{max}}) = 0 \ , \tag{4.58}$$

$$\Rightarrow \gamma^2 \left(\frac{d(\delta r)}{dt}\right)^2 + \frac{1}{2}\delta r^2 V''(r_{\text{max}}) = I$$

$$\Rightarrow \gamma^2 \left(\frac{d(\delta r)}{dt}\right)^2 - \frac{\delta r^2}{\tau_s^2} = I \tag{4.59}$$

142
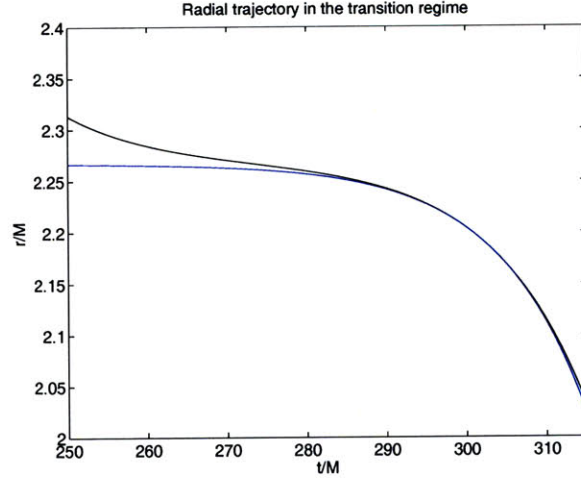
Figure 4-6: Comparison of our trajectory with approximate analytic results from Ref. [75]. The compact object is in an eccentric, equatorial trajectory with parameters $e_{\text{LSO}} = 0.6$ and $\mu = 10^{-6}M$. Its mass is $\mu = 10^{-6}M$ and is around a black hole with spin $a = 0.8M$. The black line shows our trajectory; the blue line is obtained from Ref. [75]. The observed deviation is because the approximation in Ref. [75] is somewhat more restrictive than ours.

where

$$\delta r = r(t) - r_{\text{max}}, \tag{4.60}$$

$$\gamma = \left.\frac{dt}{d\tau}\right|_{r_{\text{max}}} = \left.\frac{V_t}{\Sigma}\right|_{r_{\text{max}}}, \tag{4.61}$$

$$\tau_s^2 = 2/|V''(r_{\text{max}})|. \tag{4.62}$$

The solution of Eq. (4.59) in the regime of interest is

$$r(t) = r_{\text{max}} - \sqrt{I}\tau_s \sinh\left(\frac{t - t_c}{\gamma\tau_s}\right), \tag{4.63}$$

where t_c is an integration constant. We can compare our numerical solution with Eq. (4.63) by letting the two trajectories intersect at some arbitrary instant. This freedom is equivalent to choosing initial conditions. For example, Fig. 4-6 shows the two trajectories near r_{max} for which t_c is chosen such that they intersect at $t = 300M$. The compact object has mass $\mu = 10^{-6}M$ and is in an eccentric orbit with

143

$e_{LSO} = 0.6$ around a black hole with spin $a = 0.8M$. Notice that Eq. (4.63) is valid only in the *immediate* vicinity of r_{max} because $(dt/d\tau)$ and (E, L_z, Q) are assumed constant. Inclusion of the time-dependence of $(dt/d\tau)$ is crucial because it leads to time varying γ, which alters the natural timescale in Eq. (4.63). This explains the observed deviation at large values of $|\delta r|$.

4.4   Summary and Future work

The primary focus of this chapter is to provide an approximate model for the trajectory of a compact object as it transitions from an adiabatic inspiral to a geodesic plunge. We have presented a generalization of the procedure in Ref. [74], where circular, equatorial orbits are treated. We derive approximate equations of motion [Eq. (4.25) and Eq. (4.50)] by Taylor expanding the geodesic equations about the LSO and subjecting them to evolving E, L_z and Q. We can now readily integrate these equations numerically. Figures 4-1 and 4-2 show the radial and angular trajectories for a typical inclined, circular orbit. We also plot the plunging geodesic that it transitions to. Figures 4-3 and 4-5 are analogous plots for an eccentric orbit. Our numerical experiments suggest that the transit time is correlated with α, the coefficient of the first term in the Taylor expansion of the radial potential.

The code developed in chapters 2 and 3 solves the Teukolsky equation in the time-domain and thus computes gravitational waveforms for almost any given trajectory of the compact object. We intend to generate waveforms by feeding the world lines calculated using this prescription to the time-domain Teukolsky equation-solver. The resulting waveforms will be useful for LISA data analysis routines. An example of such a waveform from the last stages of an EMRI is shown in Fig. 5-1. Chapter 5 explores the possibility of using these waveforms to estimate recoil velocities from mergers of compact objects with black holes.

4.5 Appendix A: The LSO for eccentric orbits

The following set of equations need to be solved in order to compute p and (E, L_z, Q) at the LSO for a given inclination (ι) and eccentricity, (e):

$$R(r, E, L_z) = 0 ,\tag{4.64}$$

$$R\left(r\frac{1+e}{1-e}, E, L_z\right) = 0 \text{ and}\tag{4.65}$$

$$\frac{dR}{dr}(r, E, L_z) = 0 .\tag{4.66}$$

Recall that R is given by Eq. (4.5). The carter constant, Q can be eliminated using Eq. (4.9). Applying an iterative technique to solve the above equations directly can lead to problems because the terms that do not contain r are identical in equations (4.64) and (4.65). We can skirt around this problem by solving the equivalent set of equations,

$$R_1(r, E, L_z) = R(r, E, L_z) = 0 ,\tag{4.67}$$

$$R_2(r, E, L_z) = R\left(r\frac{1+e}{1-e}, E, L_z\right) - R_1(r, E, L_z)$$

$$= 0 \text{ and}\tag{4.68}$$

$$R_3(r, E, L_z) = \frac{dR}{dr}(r, E, L_z) = 0 ,\tag{4.69}$$

using the standard Newton-Raphson method described in [84]. This iterative procedure takes an initial guess for the solution as input. We use

$$r_0 = 6.1(1 - a/2) ,\tag{4.70}$$

$$L_{z,0} = r_0 v \cos \iota \frac{1 - 2qv^3 + q^2v^4}{\sqrt{1 - 3v^2 + 2qv^3}} \text{ and}\tag{4.71}$$

$$E_0 = \frac{1 - 2v^2 + qv^3}{\sqrt{1 - 3v^2 + 2qv^3}} .\tag{4.72}$$

where $q = a/M$, $r = \sqrt{M/r}$ and $S_0 = (r_0, L_{z,0}, E_0)^T$ is our initial guess for $S = (r, L_z, E)^T$. Let S_i denote the solution at any given iteration. The algorithm consists

145

of incrementing S_i as follows:

$$S_{i+1} = S_i + \lambda \times \delta S_i \; , \tag{4.73}$$

$$\tag{4.74}$$

where

$$\delta S_i \;\; = \;\; J_i^{-1} B_i \; , \tag{4.75}$$

$$J_i \;\; = \;\; \begin{pmatrix} \partial R_1/\partial r & \partial R_1/\partial E & \partial R_1/\partial L_z \\ \partial R_2/\partial r & \partial R_2/\partial E & \partial R_2/\partial L_z \\ \partial R_3/\partial r & \partial R_3/\partial E & \partial R_3/\partial L_z \end{pmatrix}_i \; , \tag{4.76}$$

$$B_i \;\; = \;\; (-R_{1,i}, -R_{2,i}, -R_{3,i})^T \; , \tag{4.77}$$

and $\lambda \simeq 0.1$. The subscript "i" denotes that the expressions are evaluated at $(r_i, L_{z,i}, E_i)$. We stop iterating when $|B_i| < x$, where $x \simeq 10^{-7}$. The method outlined here works well for a large fraction of parameter space.

4.6 Appendix B: Numerical integration across turning points

As mentioned in Sec. 4.3.3, Eq. (4.50) passes through turning points. The numerical integrator can accumulate error when $dr/dt \to 0$. This section describes our algorithm to resolve the issue.

Let t_p denote the instant at which $dr/dt = 0$. The radial motion is highly symmetric about the turning point. Thus, we must have,

$$\left. \frac{dr}{dt} \right|_{t_p+\epsilon} = \left. -\frac{dr}{dt} \right|_{t_p-\epsilon} \; , \tag{4.78}$$

where ϵ is an infinitesimal duration of time. When the radial velocity becomes very

small, we exploit this symmetry and set

$$\left.\frac{dr}{dt}\right|_{t_p+\delta t} = -\left.\frac{dr}{dt}\right|_{t_p-\delta t} \,,\qquad (4.79)$$

which is the discretized version of Eq. (4.78).

4.7 Acknowledgments

Chapter 5

Recoil velocities from black hole mergers

This chapter is based on a paper in preparation, written in collaboration with Gaurav Khanna and Scott A. Hughes. It will be submitted for publication in 2009.

5.1 Introduction and background

There is very strong evidence for the existence of supermassive black holes (SMBHs) at the centers of most massive galaxies. Galaxy mergers lead to the formation of SMBH binaries which ultimately coalesce. Astronomers believe that such mergers play an important part in large-scale structure formation, especially at high redshifts. There is already a growing catalog of candidates for such SMBH binaries. Examples include active galaxies with double cores [62, 70, 86], systems with doubly-peaked emission lines [47], and systems that appear to be periodic or semi-periodic, such as the blazar OJ287 [103]. The initial separation between the components of a SMBH binary formed in this manner is expected to be large enough that GW emission is weak. However, the binary loses energy through interaction with field stars, ejecting them away from the center. Ultimately, the binary ejects all the stars in its vicinity, driving the SMBHs close enough to each other that GW emission becomes the dominant radiation mechanism [15]. Thus, the last stages of the merger are primarily driven by

gravitational wave (GW) emission from the compact binary. GWs radiated during this process carry away energy and momentum (linear and angular) from the binary black hole system. If the binary is asymmetric, the net effect of the momentum loss is to impart a recoil velocity or "kick" to the merged object.

5.1.1 Background

Refs. [19, 78, 16] showed that GWs carry significant linear momentum away from non-spherical radiating systems. The theoretical estimation of recoil velocities from SMBH mergers has been the subject of much recent research. The first such computation was done by Fitchett [39]. He treated the gravitational interaction as Newtonian and included the lowest order mass and current multipoles needed for GW emission to compute the recoil velocity. This early calculation predicted that recoil velocities could approach 1000s of km/s, which is greater than the escape velocity for many galaxies.

There have been several successful efforts to improve this calculation by including additional general relativistic effects. The various approaches can be divided into the following categories:

- Estimates from black hole perturbation theory: BH perturbation theory is most widely used to describe binaries involving a massive central black hole (of mass M) and a much less massive companion (of mass μ). However, Ref. [38] has shown that recoil estimates from perturbation theory with $\mu/M \sim \mathcal{O}(0.1)$ fall within the right ballpark. In addition, Ref. [95] applies the close-limit approximation to comparable, but unequal mass binaries and obtains results that compare very well those from full numerical relativity simulations.

- Analytic estimates from post Newtonian (PN) theory: PN theory is a perturbative expansion about GM/rc^2 of the equations that govern the dynamics of the binary. The symbol M represents the characteristic mass of the system and r represents the characteristic separation between the BHs. In Refs. [18], the authors estimate the recoil velocities using PN theory.

- Estimates from numerical solutions of Einstein's equations: It has recently become possible to model the final stages of binary black hole mergers by solving Einstein's equations numerically. The authors in Refs. [9, 25] present recoil velocity estimates from these simulations.

- The effective one-body (EOB) approach: This approximation technique has been successfully applied to regimes spanning from the late-inspiral to the final plunge of a binary black hole system. The central concept here is to treat the motion of the two bodies (about one another) as the motion of a single body in an effective space-time. The effective space-time turns out to be that of a "deformed" black hole and thus, the approach requires an understanding of the motion of a test body in such a geometry. This EOB approach has recently had great success in being able to generate waveforms and recoil velocities that match well with those from full numerical relativity [32, 91].

Irrespective of the approach followed, all calculations predict recoil velocities in the range of a few hundred km/s for most SMBH binaries, which are much lower than Fitchett's original calculation.

The computed recoil velocities also serve another important purpose. They can be viewed as a common point of comparison for the four approaches to strong field gravity viz., post Newtonian theory, numerical relativity, perturbation theory, and the EOB approach. Computationally, the recoil velocity from a merging binary is calculated by integrating the emitted radiation over several orbits. Any significant systematic error in the approach (to strong field gravity) used will tend to magnify the error in the estimated recoil velocity. Thus, the evaluated recoils for a range of BH spins and mass ratios form a single set of numbers which serves as a good platform for comparing various approaches to strong field gravity.

5.1.2 This chapter

Ref. [38] predicts upper and lower bounds for recoil velocities from BH perturbation theory. In this chapter, we make use of the tools developed in earlier chapters to

build on their work and make precise predictions. This section is brief overview of the developments in BH perturbation theory leading to this work.

There are two components in exploiting perturbation theory to describe realistic inspirals. First, we need a method to compute the world line followed by the smaller object. Next, we need to compute the associated GWs from the inspiral. The Teukolsky equation describes radiation that arises due to scalar, vector and tensor perturbations. It is a second order, linear, and inhomogeneous partial differential equation for curvature perturbations around spinning BHs. We use the Teukolsky equation to describe BH binaries by treating the smaller companion as a perturbation to the central BH's spacetime.

The Teukolsky equation can be solved by two different techniques. The equation happens to be variable separable, leading to a decomposition of the solution into Fourier frequency modes. This frequency-domain based decomposition works especially well when the smaller object is in a bound geodesic around the central BH [53, 36]. (This is because GWs from bound geodesics show discrete spectra.) One of the primary strengths of frequency domain formalism is that it can be harnessed to produce the inspiral world-line from a large initial separation down to the last stable geodesic orbit (LSO). This is possible because the radiated fluxes in energy and angular momentum corresponding to geodesics can be computed to very high accuracy. However, the ability of the formalism to generate waveforms from realistic inspirals becomes inefficient because the spectrum becomes continuous and a large number of Fourier modes need to be computed.

An alternate approach to solve the Teukolsky equation is by treating it as a (2+1) dimensional PDE and numerically evolving the curvature perturbations with time on a spatial grid [97, 98]. The strength of this approach lies in its ability to produce GW solutions corresponding to inspiral world lines (which have continuous spectra). The disadvantage is that it is computationally more expensive than the FD formalism to calculate the radiated fluxes from most bound geodesics.

Ref. [74] and its generalization, Ref. [96] present a technique to compute the trajectory of the smaller object as it transitions from the LSO and plunges into the horizon

of the central BH. Combining this with the FD formalism, we can now compute the entire inspiral trajectory of the smaller object (within certain approximations).

In summary, we can now compute the entire inspiral world line of the smaller object and its associated GWs. This represents the culmination of several years of work in BH perturbation theory by many researchers. With this in hand, we now calculate the recoil velocity that results from the merger.

Our primary results, the recoil velocities shown in tables 5.1-5.5 and Fig. 5-5 are in excellent agreement with earlier calculations from numerical relativity, perturbation theory and PN theory. This is significant because: (a) Exploration of parameter space is computationally easier in BH perturbation theory than in numerical relativity, and (b) BH perturbation theory is more accurate for small mass ratios and in strong fields than PN theory.

On a slight (but important) detour, we mention that this work, along with the research leading to it presents the first simulations of GWs from the final stages of an EMRI in a Kerr spacetime. Waveforms like those shown in Figs. 5-1 and 5-2 will be useful in LISA data analysis routines.

The rest of this chapter is organized as follows: In Sec. 5.2, we summarize our approach to inspirals, highlighting the approximations made. In Sec. 5.3, we present the recoil velocities computed from our approach. We also compare these predictions with earlier results. In Sec. 5.4 we discuss the convergence of our results with respect to the number of azimuthal modes included in the waveforms. We use units in which $G = c = 1$ throughout this chapter.

5.2   Summary of our approach to inspirals

There are two steps in the computation of the recoil velocity. First, we compute the inspiral trajectory followed by the smaller object. Next, we utilize this trajectory to solve the Teukolsky equation and generate its associated GWs and recoil velocities. Our approach is strictly valid only in the limit of extremely small mass ratios.

5.2.1   Geodesics in a Kerr spacetime

Geodesics in a Kerr spacetime [14, 27] satisfy the following equations:

$$\Sigma \frac{dr}{d\tau} = \pm\sqrt{R} , \tag{5.1}$$

$$\Sigma \frac{d\theta}{d\tau} = \pm\sqrt{V_\theta} , \tag{5.2}$$

$$\Sigma \frac{d\phi}{d\tau} = V_\phi , \tag{5.3}$$

$$\Sigma \frac{dt}{d\tau} = V_t . \tag{5.4}$$

The potentials can be expressed as:

$$R = \left[E\left(a^2 + r^2\right) - aL_z \right]^2 -$$
$$+ \left[(L_z - aE)^2 + r^2\mu^2 + Q \right] , \tag{5.5}$$

$$V_\theta = [Q-$$
$$\cos^2\theta \left(a^2\left(1 - E^2\right) + L_z^2/\sin^2\theta \right)] , \tag{5.6}$$

$$V_\phi = \left[L_z/\sin^2\theta - aE \right]$$
$$+ \frac{a}{\Delta} \left[E\left(r^2 + a^2\right) - L_z a \right] , \tag{5.7}$$

$$V_t = \left[a\left(L_z - aE\sin^2\theta\right) \right] +$$
$$\frac{r^2 + a^2}{\Delta} \left[E\left(r^2 + a^2\right) - L_z a \right] . \tag{5.8}$$

The parameters (r, θ, ϕ, t) are the Boyer-Lindquist coordinates, M is the black hole mass, μ is the perturbing mass, $\Sigma = r^2 + a^2\cos^2\theta$, $\Delta = r^2 - 2Mr + a^2$ and a is the spin parameter of the black hole. Along with initial conditions, the dimensionless energy, momentum and Carter constant, (E, L_z, Q) define a geodesic.

The geodesic equations admit two types of solutions: (i) Bound orbits, where the radial and angular motions are periodic, and (ii) Unstable orbits, where the particle ultimately plunges into the central BH. The constants (E, L_z, Q) determine the stability of an orbit.

154

5.2.2 The Teukolsky equation

The Teukolsky equation describes perturbations due to scalar, vector and tensor fields in the vicinity of Kerr black holes [100, 99]. In Boyer-Lindquist coordinates, this equation is

$$
-\left[\frac{(r^2+a^2)^2}{\Delta} - a^2\sin^2\theta\right]\partial_{tt}\Psi - \frac{4Mar}{\Delta}\partial_{t\phi}\Psi
$$
$$
-2s\left[r - \frac{M(r^2-a^2)}{\Delta} + ia\cos\theta\right]\partial_t\Psi
$$
$$
+\Delta^{-s}\partial_r\left(\Delta^{s+1}\partial_r\Psi\right) + \frac{1}{\sin\theta}\partial_\theta\left(\sin\theta\partial_\theta\Psi\right) +
$$
$$
\left[\frac{1}{\sin^2\theta} - \frac{a^2}{\Delta}\right]\partial_{\phi\phi}\Psi
$$
$$
+2s\left[\frac{a(r-M)}{\Delta} + \frac{i\cos\theta}{\sin^2\theta}\right]\partial_\phi\Psi
$$
$$
-\left(s^2\cot^2\theta - s\right)\Psi = -4\pi\left(r^2+a^2\cos^2\theta\right)T, \tag{5.9}
$$

where M is the mass of the black hole, a its angular momentum per unit mass, $\Delta = r^2 - 2Mr + a^2 = (r-r_+)(r-r_-)$, $r_\pm = M \pm \sqrt{M^2-a^2}$ and s is the "spin weight" of the field. The $s = \pm 2$ versions of these equations describe perturbations to the Weyl curvature tensor, in particular the radiative degrees of freedom ψ_0 and ψ_4. That is, $\Psi = \psi_0$ for $s = +2$, and $\Psi = \rho^{-4}\psi_4$ for $s = -2$, with $\rho = -1/(r - ia\cos\theta)$. The T in the right hand side of the equation is a source term, constructed from the trajectory of the perturbing mass. The details of the perturbing object enter here.

The equations

$$
\psi_4 = \frac{1}{2}\left(\frac{\partial^2 h_+}{\partial t^2} - i\frac{\partial^2 h_\times}{\partial t^2}\right), \tag{5.10}
$$

$$
\frac{dP_i}{dt}(t) = \lim_{r\to\infty}\left[\frac{r^2}{4\pi}\int d\Omega\frac{x_i}{r}\left|\int_{-\infty}^t\psi_4\right|^2\right], \tag{5.11}
$$

relate ψ_4 to the "plus"(h_+)/"cross"$(h_\times)$ polarizations of the GWs, and the linear momentum flux dP_i/dt. The subscript i denotes a Cartesian coordinate (x, y or z).

The axial symmetry of the problem makes the azimuthal dependence simple. The

155

code developed in Ref. [97, 98] factors out this dependence and uses the ansatz,

$$\Psi(t,r,\theta,\phi) \;=\; \sum_m e^{im\phi} \exp\left(im\left[\frac{a}{r_+ - r_-} \ln\frac{r - r_+}{r - r_-} \right] \right) r^3 \Phi_m(t,r,\theta), \quad (5.12)$$

to solve the Teukolksy equation numerically as a (2+1) dimensional partial differential equation. The key result of Ref. [97, 98] is that we can generate GWs from generic non-geodesic orbits that are both inclined and eccentric. The error was shown to be less than 1% for a large fraction of parameter space.

5.2.3 The trajectory

The trajectory of the smaller object can be broken down into three regimes: (a) An adiabatic inspiral phase, where the inspiral time scale is much larger than the orbital period. In this phase, we approximate the trajectory as a sequence of bound geodesics; (b) A late-time radial infall, which can be approximated as a single unstable, plunging geodesic; and (c) A short regime where the body transitions from inspiral to plunge [74, 96].

Ref. [53, 36] solves the Teukolsky equation by expanding ψ_4 as a sum of Fourier frequency modes in r, θ, ϕ, and t. This frequency-domain based algorithm is very effective at computing waveforms and radiated fluxes $(\dot{E}, \dot{L}_z, \dot{Q})$ corresponding to bound geodesics. We use the frequency-domain code to construct the smaller object's "phase space" trajectory $[E(t), L_z(t), Q(t)]$ [53, 36] during the early adiabatic inspiral that defines the sequence of geodesics it evolves through. The geodesic equations map $[E(t), L_z(t), Q(t)]$ to an inspiral world line $[r(t), \theta(t), \phi(t)]$ for a given set of initial conditions.

Ref. [74] and its generalization, Ref. [96] present a prescription to model the trajectory during the transition from early inspiral to the final plunge. It also provides initial conditions for the plunging geodesic that it transitions to. In summary, we can combine the results of Ref. [53, 36, 74, 96] to construct the entire inspiral trajectory. We reiterate that the "piecewise" trajectory constructed in this manner is only approximate in that it neglects the conservative part of the gravitational self-force on

156

the smaller object.

5.3   Recoil velocities from black hole mergers

This section presents the numerically computed recoil velocities from our approach. Tables 5.1-5.5 show the recoil velocities for a number of approximately circular inspirals in the equatorial place. Each table shows data corresponding to a particular BH spin for a range of mass ratios. Ideally, we would like the initial separation to be ∞. Due to computational restrictions, we set the initial separation such that the effect of finite initial separation is minimal. In most cases, the major contribution to the recoil velocity is from the final stages of the merger. Hence, the error from starting out at a large but not infinite separation is small. However, this error increases with black hole spin. We discuss this in detail in Sec. 5.3.2. Since the inspiral is slower for smaller μ/M, our initial separation decreases with mass ratio.

We use Eq. (5.11) to compute the linear momentum flux in the x, y and z directions from ψ_4. We obtain the recoil velocity by integrating dP/dt,

$$v_{\text{rec,i}} \;=\; c \int_{t_0}^{t_e} dt \frac{dP_i}{dt} \;,$$ (5.13)

$$|v_{\text{rec}}| \;=\; \sqrt{v_x^2 + v_y^2 + v_z^2} \;.$$ (5.14)

The upper limit t_e is chosen after the merger, when the amplitude of the ringdown becomes substantially small. We choose the lower limit t_0 such that the effect of finite initial separation is minimal. In the absence of radiation reaction, $\dot{P}_i$ is purely sinusoidal [101]. This implies $P_i = 0$ when $\dot{P}_i$ is at a local extremum. At large enough separation, the effect of radiation reaction is small and we expect P_i and $\dot{P}_i$ to be approximately sinusoidal. This is seen in Fig. 5-3. Thus, we choose t_0 to be the first local extremum in our computed $\dot{P}_i(t)$. This assures that,

$$P_i(t_0) \;\ll\; \int_{t_0}^{t_e} dt \frac{dP_i}{dt} \;.$$ (5.15)

Fig. 5-1 plots $\psi_4(r = 200M, \theta = \pi/2, \phi = 0, t)$ just before and after merger. The sudden jump in amplitude during the final stages of merger is clearly visible. The exponential "ringdown" to the BH's quiescent state is also visible.
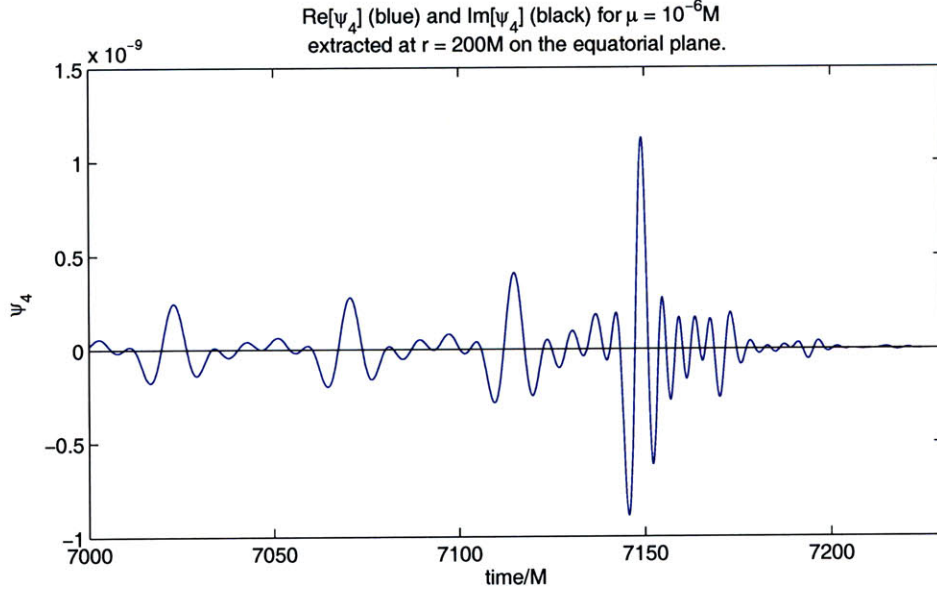
Figure 5-1: Waveforms from merger and ringdown of a binary consisting of a massive black hole of mass M and a much smaller companion of mass $\mu = 10^{-6}M$. The central BH has spin parameter $a/M = 0.6$. The complex field ψ_4 is a radiative component of the Weyl curvature tensor and is related to the metric perturbations via Eq. (5.10). The waveform is extracted at a radius of $200M$ on the equatorial plane of the central BH. Azimuthal modes with $|m| < 5$ are included in the waveform.

5.3.1 Variation with mass and spin

The data in tables 5.1-5.5 indicate that the recoil velocity scales approximately as $(\mu/M)^2$. This is easy to understand from the viewpoint of perturbation theory. The major contribution to the recoil comes from the final plunge, which is approximately a geodesic. Thus, the plunging trajectory remains approximately the same for all mass ratios. However, ψ_4 scales as μ/M which leads to $dP/dt \propto (\mu/M)^2$ from Eq. (5.11). For higher mass ratios, the contribution to the recoil from the inspiral and transition become more prominent. This leads to a deviation from the $(\mu/M)^2$ scaling.

158
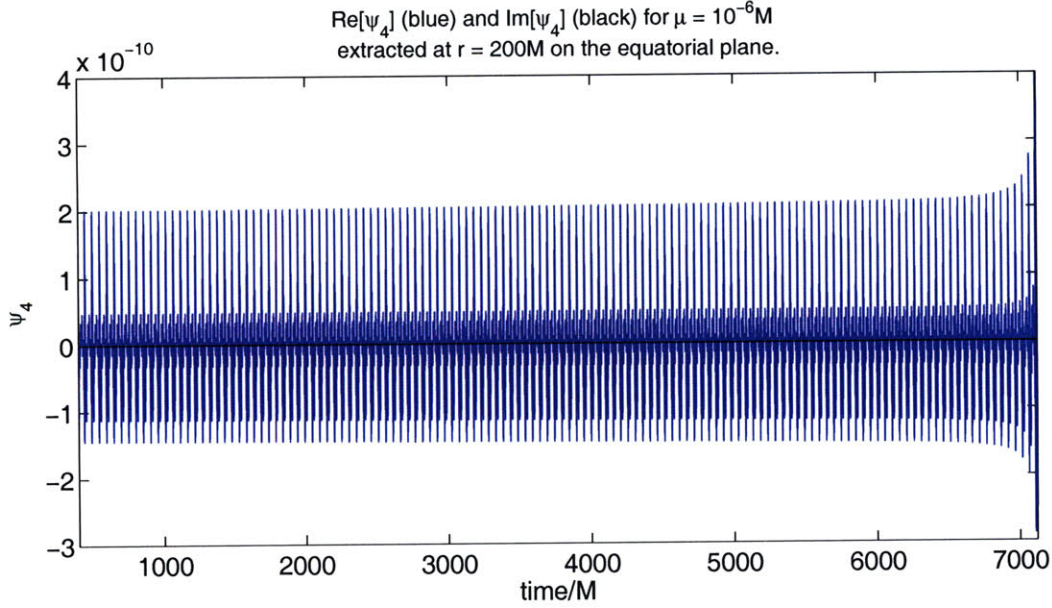
Figure 5-2: Same as Fig. 5-1, but for the time interval prior to the merger.

Tables 5.1-5.5 show that the recoil velocity decreases with increasing BH spin. This trend is consistent with observations from full numerical relativity simulations in Ref. [9]. Moreover, tables 5.1-5.5 suggest that recoils from retrograde motion $(a/M < 0)$ are significantly higher than recoils from prograde motion $(a/M > 0)$. This is qualitatively simple to explain. Recoils are the result of radial, non-circular motion. [1] In the presence of radiation reaction, the smaller object's velocity is approximately tangential, superposed with a small radial component. The largest amount of radial, non-circular motion happens during the final plunge phase. Thus, a longer plunge phase leads to a greater recoil. Now, the duration of the plunge phase depends on the radius of the innermost stable circular orbit (ISCO), which in turn depends on the spin of the central BH. It turns out that the radius of the ISCO decreases with a/M, which implies a longer plunge phase for retrograde motion. This leads to the observed decrease in recoils with increasing a/M. Thus, the net effect of alignment of the central BH's spin with the orbital angular momentum on frame

---

[1]The average $\dot{P}_i(t)$ from purely circular motion is exactly zero.

159
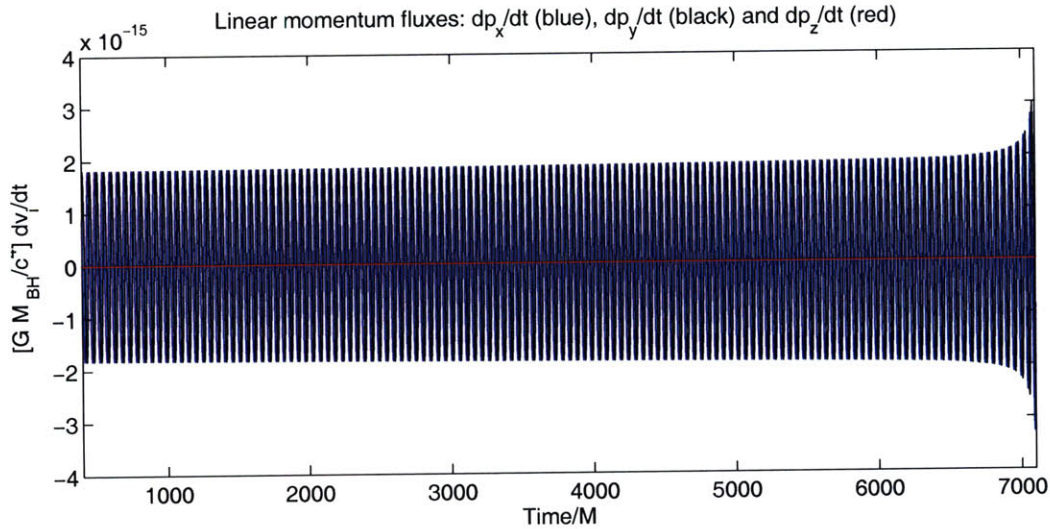
Figure 5-3: Momentum flux for a binary inspiral with parameters as in Fig. 5-1. The integrated recoil velocity is 1.0×10^{-6} km/sec.

dragging tends to decrease the recoil.

Table 5.1: Recoil velocities for a range of nearly circular inspirals in the equatorial plane. The central BH has spin $a/M = 0.0$. The quantity μ/M denotes the mass ratio, p is the initial separation, and $v_{\rm rec}$ is our numerically computed recoil velocity. These results use the azimuthal modes $|m| = 0, 1, 2, 3, 4, 5$. The column $v_{\rm nr}$ shows recoils obtained from Eq. (5.16) with best fit parameters quoted in Ref. [9]; $v_{\rm fit}$ is obtained from our best fit parameters.

μ/M	p/M	a/M	$v_{\rm rec}$ (km/sec)	$v_{\rm nr}$(km/sec)	$v_{\rm fit}$(km/sec)	$\frac{v_{\rm rec}-v_{\rm nr}}{v_{\rm rec}}$	$(v_{\rm rec}-v_{\rm fit})/v_{\rm rec}$
10^{-4}	6.4	0.0	1.18×10^{-4}	8.83×10^{-5}	1.14×10^{-4}	0.25	3.6×10^{-2}
10^{-5}	6.14	0.0	1.18×10^{-6}	8.83×10^{-7}	1.14×10^{-6}	0.25	3.6×10^{-2}
10^{-6}	6.04	0.0	1.18×10^{-8}	8.83×10^{-9}	1.14×10^{-8}	0.25	3.6×10^{-2}

5.3.2   Comparison with numerical relativity

It has now become possible to successfully simulate mergers (eg. Refs. [9, 25]) between comparable mass black holes by solving the Einstein field equations numerically. These simulations of the final stages of merger give excellent estimates of the recoil velocities. Several groups have proposed phenomenological formulae for the recoil

160

Table 5.2: Same as Table 5.1, except that the central BH has spin parameter $a/M = 0.3$.

μ/M	p/M	a/M	$v_{\rm rec}$ (km/sec)	$v_{\rm nr}$(km/sec)	$v_{\rm fit}$(km/sec)	$\frac{v_{\rm rec}-v_{\rm nr}}{v_{\rm rec}}$	$(v_{\rm rec} - v_{\rm fit})/v_{\rm rec}$
10^{-4}	5.45	0.3	9.64×10^{-5}	7.04×10^{-5}	9.55×10^{-5}	0.27	9.3×10^{-3}
10^{-5}	5.13	0.3	9.64×10^{-7}	7.05×10^{-7}	9.55×10^{-7}	0.27	9.1×10^{-3}
10^{-6}	5.025	0.3	1.05×10^{-8}	7.05×10^{-9}	9.55×10^{-9}	0.33	9.0×10^{-2}

Table 5.3: Same as Table 5.1, except that the central BH has spin parameter $a/M = 0.6$.

μ/M	p/M	a/M	$v_{\rm rec}$ (km/sec)	$v_{\rm nr}$(km/sec)	$v_{\rm fit}$(km/sec)	$\frac{v_{\rm rec}-v_{\rm nr}}{v_{\rm rec}}$	$(v_{\rm rec} - v_{\rm fit})/v_{\rm rec}$
10^{-4}	4.35	0.6	6.94×10^{-5}	5.61×10^{-5}	7.29×10^{-5}	0.19	-5.1×10^{-2}
10^{-5}	3.99	0.6	6.96×10^{-7}	5.61×10^{-7}	7.30×10^{-7}	0.19	-4.8×10^{-2}
10^{-6}	3.881	0.6	6.98×10^{-9}	5.61×10^{-9}	7.30×10^{-9}	0.20	-4.5×10^{-2}

velocity by fitting the available data from simulations to analytical expressions. We test the validity of these expressions in the extreme mass ratio limit.

For example, Ref. [9] proposes that,

$$v_{\rm rec} = f(q, a_1, a_2; V_0, k, \beta) , \tag{5.16}$$

$$= \frac{32 V_0 q^2}{(1+q)^5} \sqrt{(1-q)^2 + 2(1-q)\beta K + K^2}, \tag{5.17}$$

where $q = m_1/m_2$, m_1 and m_2 are the masses of the two BHs $(m_1 < m_2)$; $K = k(q\hat{a}_1 - \hat{a}_2)$; $\hat{a}_1 = a_1/m_1$, $\hat{a}_2 = a_2/m_2$; a_1 and a_2 are the spin parameters of the BHs with masses m_1 and m_2 respectively. The authors obtain the free parameters V_0, K and β by fitting their data to Eq. (5.16). Their best fit parameters are $V_0 = 276$ km/sec, $\beta = 0.84$ and $k = 0.85$.

We fit the data in tables 5.1-5.5 to Eq. (5.16) in order to provide estimates for V_0, K and β from BH perturbation theory. We make the association,

$$M = m_1 + m_2, \tag{5.18}$$

$$\mu = \frac{m_1 m_2}{M}, \tag{5.19}$$

because BH perturbation theory is formulated in the center of momentum frame whereas the numerical relativity simulations are performed in a "lab" frame.

Table 5.4: Same as Table 5.1, except that the central BH has spin parameter $a/M = -0.3$.

μ/M	p/M	a/M	$v_{\rm rec}$ (km/sec)	$v_{\rm nr}$(km/sec)	$v_{\rm fit}$(km/sec)	$\frac{v_{\rm rec}-v_{\rm nr}}{v_{\rm rec}}$	$(v_{\rm rec} - v_{\rm fit})/v_{\rm rec}$
10^{-4}	7.29	-0.3	1.29×10^{-4}	1.08×10^{-4}	1.30×10^{-4}	0.16	-4.3×10^{-3}
10^{-5}	7.06	-0.3	1.28×10^{-6}	1.08×10^{-6}	1.30×10^{-6}	0.16	-1.2×10^{-2}
10^{-6}	6.98	-0.3	1.28×10^{-8}	1.08×10^{-8}	1.30×10^{-8}	0.16	-1.2×10^{-2}

Table 5.5: Same as Table 5.1, except that the central BH has spin parameter $a/M = -0.6$.

μ/M	p/M	a/M	$v_{\rm rec}$ (km/sec)	$v_{\rm nr}$(km/sec)	$v_{\rm fit}$(km/sec)	$\frac{v_{\rm rec}-v_{\rm nr}}{v_{\rm rec}}$	$(v_{\rm rec} - v_{\rm fit})/v_{\rm rec}$
10^{-4}	7.89	-0.6	1.40×10^{-4}	1.28×10^{-4}	1.44×10^{-4}	8.2×10^{-2}	-2.6×10^{-2}
10^{-5}	7.87	-0.6	1.42×10^{-6}	1.29×10^{-6}	1.44×10^{-6}	9.5×10^{-2}	-1.2×10^{-2}
10^{-6}	7.86	-0.6	1.40×10^{-8}	1.29×10^{-8}	1.44×10^{-8}	8.2×10^{-2}	-2.6×10^{-2}

Sec. 5.6 outlines the chi-squared procedure adapted to fit our data to Eq. (5.16). Our best fit parameters are $V_0 = 355.59 \pm 5.1$ km/sec, $k = 0.09 \pm 0.67$ and $\beta = 5.48 \pm 40$. Fig. 5-4 shows our data along with the line of best fit. The parameter with dimensions of velocity, V_0 is in fairly good agreement with Ref. [9]. On the other hand, the dimensionless parameters k and β are poorly constrained. This is because recoil velocities are less sensitive to substantial changes in k and β at small mass ratios. However, the error bars on k and β are large enough to be consistent with Ref. [9]. Interestingly, the authors in Ref. [91] also find that the effective-one-body approach is in excellent agreement with Eq. (5.16).

Ref. [25] and Ref. [69] use another variation of Eq. (5.16),

$$\vec{v}_{\rm rec} = v_m \hat{e}_1 + v_\perp(\cos\xi\hat{e}_1 + \sin\xi\hat{e}_2), \tag{5.20}$$

$$v_m = A\frac{q^2(1-q)}{(1+q)^5}\left(1 + B\frac{q}{(1+q)^2}\right), \tag{5.21}$$

$$v_\perp = H\frac{q^2}{(1+q)^5}(\hat{a}_2 - q\hat{a}_1), \tag{5.22}$$

for equatorial orbits. The orthogonal unit vectors $\hat{e}_1$ and $\hat{e}_2$ lie in the orbital plane. Since BH perturbation theory is accurate only up to leading order in q, we ignore B
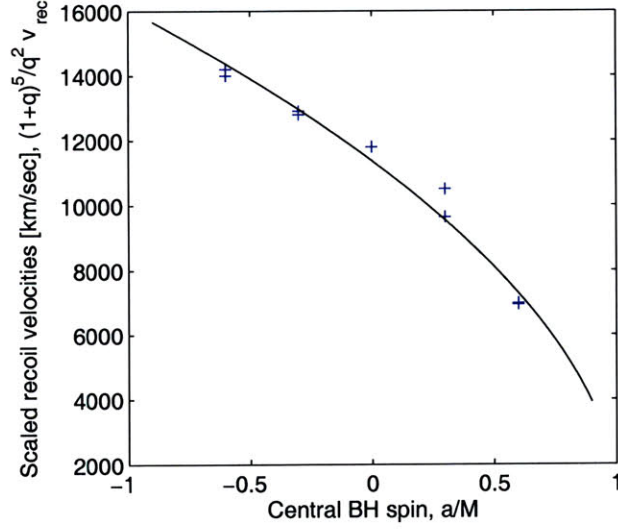
Figure 5-4: Line of best fit (solid black) for the data in tables 5.1-5.5. Our data ("+") are also shown.

and fit our data to,

$$v_{\text{rec}} = \sqrt{v_m^2 + 2v_m v_\perp \cos \xi + v_\perp^2} , \tag{5.23}$$

$$\simeq A \frac{q^2}{(1+q)^5} \Big[(1-q)^2 - $$

$$2(1-q)\frac{H}{A}(q\hat{a}_1 - \hat{a}_2)\cos \xi + \frac{H^2}{A^2}(q\hat{a}_1 - \hat{a}_2)^2 \Big]^{1/2} . \tag{5.24}$$

Comparing Eq. (5.23) to Eq. (5.16) gives best fit parameters $A = 32V_0 = 1.14 \times 10^4$ km/sec, $H = kA = 1024$ km/sec and $\cos \xi = -\beta = -5.48$. Ref. [9] obtains $A = 1.35 \times 10^4$ km/sec, $B = -1.48$, $H = 7540$ km/sec and $\cos \xi = 0.82$. On the other hand, Ref. [69] obtains $A = 1.2 \times 10^4$ km/sec, $B = -0.93$, $H = 6900$ km/sec and $\cos \xi = 0.82$. Our estimates for A, the parameter that is constrained well are in good agreement.

Fig. 5-5 is an alternate way to compare recoil velocities obtained from perturbation theory with those obtained from numerical relativity [9, 25, 69]. The solid black line is Eq. (5.16) for $\mu/M = 0.1$ and $-0.7 < a/M < 0.7$ with best fit parameters

163
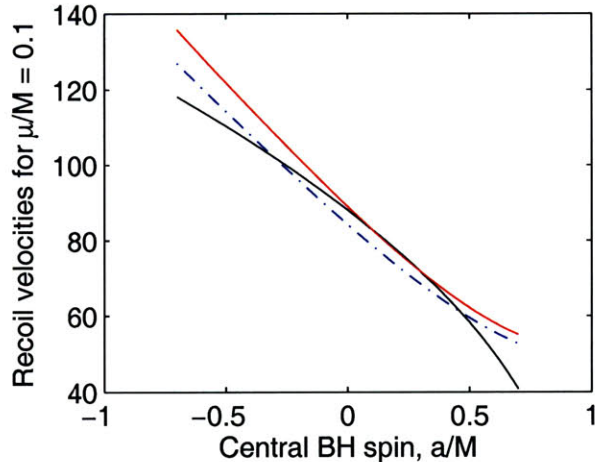
Figure 5-5: Comparison of recoil velocity estimates (km/sec) from various approaches for a binary with $\mu/M = 0.1$ ($q = 0.127$). The solid black line is our estimate. The blue (dot-dashed) line uses best fit parameters from Ref. [69]; the red (grey) line uses best fit parameters from Ref. [9].

obtained here. The red (grey) and blue (dot-dashed) lines are Eq. (5.23) with best fit parameters obtained from Refs. [9] and [69]. Interestingly, our estimates fall *between* the numerical relativity estimates for a large fraction of parameter space. Furthermore, our estimates are consistent with Fig. 2 of Ref. [38]. We notice that our disagreements are largest for high spins. We discuss this below:

For a given mass ratio, the duration of the adiabatic inspiral increases with BH spin a/M. This is because the radius of the innermost stable circular orbit decreases with a/M, which in turn leads to a longer adiabatic phase (slow) and a shorter plunge phase (rapid). The numerical integration time increases with the duration of the trajectory. This makes generation of recoil data from the time-domain Teukolsky solver (the code used here) for large spins ($a/M > 0.6$) computationally cumbersome. We are investigating the possibility of using the code developed in Ref. [38] to calculate the recoil from the early inspiral and integrating it with the code used here.

The decrease in recoils at large spins is associated with an "anti-kick" during the merger and ringdown phases. This effect arises as a result of the loss of linear momentum in a direction such that it cancels out most of the recoil accumulated over

164

several cycles. Ref. [92] examines the anti-kick for a range of binaries. The published version of this work will include a detailed exploration of this effect.

We incur a computational cost of approximately 1800 CPU hours on a X86_64 processor for the recoil velocity from each set of parameters. The duration of each simulation is about $7500M$.

5.4   Convergence with azimuthal mode

As mentioned in Sec. 5.2, we expand the solution of the Teukolsky equation as a sum of azimuthal modes,

$$\Psi(t, r, \theta, \phi) = \sum_m e^{im\phi} \exp\left(im\left[\frac{a}{r_+ - r_-}\ln\frac{r - r_+}{r - r_-}\right]\right) r^3 \Phi_m(t, r, \theta) . \tag{5.25}$$

If we include only one m-mode in the summation, the momentum flux from Eq. (5.11) becomes,

$$\dot{P}_x \propto \int_0^{2\pi} d\phi \cos\phi |\Phi_m(t, r, \theta)|^2 , \tag{5.26}$$

$$= 0 \text{ and} \tag{5.27}$$

$$\dot{P}_y \propto \int_0^{2\pi} d\phi \sin\phi |\Phi_m(t, r, \theta)|^2 , \tag{5.28}$$

$$= 0 . \tag{5.29}$$

In general, the integral over ϕ in Eq. (5.11) is proportional to,

$$\dot{P}_x \propto \sum_{m,m'} \int_0^{2\pi} d\phi \cos\phi \, e^{i(m-m')\phi}(\Phi_m \Phi_{m'}^*) , \tag{5.30}$$

$$\propto \sum_{m,m'} [\delta_{m+1}\delta_{m'} + \delta_m \delta_{m'+1}] (\Phi_m \Phi_{m'}^*) , \tag{5.31}$$

$$\dot{P}_x \propto \sum_{m,m'} \int_0^{2\pi} d\phi \sin\phi \, e^{i(m-m')\phi}(\Phi_m \Phi_{m'}^*) . \tag{5.32}$$

$$\propto \sum_{m,m'} [\delta_{m+1}\delta_{m'} - \delta_m \delta_{m'+1}] (\Phi_m \Phi_{m'}^*) , \tag{5.33}$$

Thus, contributions from terms with $m = m'$ vanish. This shows that the recoil velocity arises due to beating between adjacent $m-$modes.

Ideally, Eq. (5.25) should include an infinite number of $m-$modes. Due to computational constraints, we only include the modes $|m| \leq 5$. This is justified because we expect contributions to decrease exponentially for larger $m-$modes. In this section, we examine the convergence of our recoil velocities with m.

Table 5.6: Convergence with azimuthal mode. We show recoil velocities, v_{rec} for a binary with $\mu/M = 10 \times 10^{-5}$, $a/M = 0.3$ when a restricted number of azimuthal (m) modes are included in the waveform. Inclusion of modes beyond $|m| = 5$ changes v_{rec} only marginally.

$|m|-$modes included	v_{rec} (km/sec)
0,1,2	7.57×10^{-7}
0,1,2,3	8.72×10^{-7}
0,1,2,3,4	9.63×10^{-7}
0,1,2,3,4,5	9.64×10^{-7}

Table 5.6 shows the effect of including additional modes on the recoil velocity for a typical BH binary. As expected, we observe steady convergence in the recoils. The inclusion of the $m = +5, -5$ modes changes the recoil only by ~ 0.1 percent.

5.5 Conclusions and future work

This chapter is the culmination of approximately a decade of work in BH perturbation theory. We now have a versatile toolkit to fully exploit BH perturbation theory to generate gravitational waveforms, fluxes, and recoil velocities from BH binaries. Figs. 5-1 and 5-2 show GWs from a typical binary. Recoil velocities from binaries with a range of parameters are shown in tables 5.1-5.5 and Fig. 5-5. Most importantly, the results in this chapter are in concordance with other approaches to relativistic binaries.

A natural extension of this work will be to include completely generic inspirals, where the orbit of the smaller object is inclined with respect to the central BH's equator. In theory, we have all the tools required to compute recoils from such inspirals. However, world lines from these generic inspirals are more detailed and

require extensive computations. We are currently working on making the process more efficient and integrating it with the existing code base.

The results here are motivation to add additional features to perturbation theory. For example, the spin of the perturbing object is completely ignored. It will be worthwhile to include the effects of the spin angular momentum of the perturbing object to its dynamics and the associated GWs. The dynamics will require solving the Papapetreau equations of motion instead of the geodesic equations done here. The stress-energy tensor of the perturbing object (that appears in the Teukolsky equation) will also have to be suitably generalized to account for spin.

We emphasize that the trajectory used in computing the inspirals here do not account for the conservative part of the gravitational self force. When this becomes available, it should be straightforward to incorporate it into our model for the trajectory to yield improved waveforms and recoil velocity estimates.

5.6 Appendix: Best fit parameters

We outline our procedure to estimate the parameters V_0, k and β by fitting our data to the expression Eq. (5.16). For the most part, we use the standard chi-squared analysis [84]. We highlight some minor differences here.

For a given mass ratio (q_i) and spin (a_i), define

$$y_i = \frac{(1+q_i)^5}{q_i^2} v_{\text{rec},i} \ , \tag{5.34}$$

$$\mathbf{b} = (V_0, k, \beta) \tag{5.35}$$

$$F_i = \frac{(1+q_i)^5}{q_i^2} f(q_i, 0, a_i; \mathbf{b})$$

$$= F(q_i, a_i; \mathbf{b}) \ , \text{ and} \tag{5.36}$$

$$\chi^2(\mathbf{b}) = \frac{1}{\sigma^2} \sum_i [y_i - F(q_i, a_i; \mathbf{b})]^2 \ . \tag{5.37}$$

The subscript i denotes a data point. For example, the second row of table 5.2 gives $v_{\text{rec},i} = 9.64 \times 10^{-7}$ km/sec, $\mu_i = 10^{-5} M$, $a_{1,i} = 0$ and $a_{2,i} = 0.3$. Note that the spin of the massive black hole is $a_{2,i} = a_i$ and that of the perturbing mass is

$a_{1,i} = 0$ always. The parameter σ^2 is the variance of y_i *if* it were a Gaussian random variable. Our rescaling of the numerical recoil velocities [by introducing the factor $\propto (1 + q_i)^5/q_i^2 \sim 1/\mu^2$ in Eq. (5.37)] eliminates the effect of the scaling of $v_{\text{rec},i}$ with μ. Else, contributions from higher mass ratios will dominate χ^2.

Let $\hat{\mathbf{b}}$ be the (unknown) actual value of $\mathbf{b}$. Our best fit parameters, $\mathbf{b}^* = (V_0^*, k^*, \beta^*)$ are those that minimize χ^2. If y_i were a Gaussian random variable, the best fit parameters are also random and can shown to be maximum likelihood estimators for $\hat{\mathbf{b}}$.

We now need a measure of the error in our estimate $\mathbf{b}^*$. If some $\mathbf{b}^* + \delta\mathbf{b}^*$ were obtained for a set of data $\{y_i + \delta y_i\}$ with the same actual parameters $\hat{\mathbf{b}}$, then the random variable

$$\Delta\chi^2 \;\;=\;\; \chi^2(\mathbf{b}^* + \delta\mathbf{b}^*) - \chi^2(\hat{\mathbf{b}}) \tag{5.38}$$

follows a chi-squared probability distribution. Note that we can estimate σ^2 in Eq. (5.37) by computing

$$\sigma^2 = \sum_{i=0}^{N-1} \left[y_i - F\left(q_i, a_i; \mathbf{b}^*\right)\right]^2 /(N - M) \,, \tag{5.39}$$

where N is the number of data points and M is the number of free parameters (which is 3 in our case). We now want the value of $\delta\mathbf{b}^*$ for which the probability,

$$P[-\Delta\chi_0^2 \leq \Delta\chi^2 \leq \Delta\chi_0^2] = p_0 \,. \tag{5.40}$$

Here, p_0 is the desired confidence level and $\Delta\chi_0^2$ can be obtained from the chi-squared distribution. Further, $\delta\mathbf{b}^*$ is a measure of our error with a confidence of p_0.

In general, $\Delta\chi^2$ is approximated by,

$$\Delta\chi^2 = \sum_{kj} \delta\mathbf{b}_{\mathbf{k}}^* \, \alpha_{kj} \, \delta\mathbf{b}_{\mathbf{j}}^* \tag{5.41}$$

168

where α is a matrix whose components are

$$\alpha_{kj} = \frac{1}{\sigma^2} \sum_{i=0}^{N-1} \left[\frac{\partial F(q_i, a_i; \mathbf{b})}{\partial \mathbf{b_k}} \frac{\partial F(q_i, a_i; \mathbf{b})}{\partial \mathbf{b_j}} \right]_{\mathbf{b}^*} . \tag{5.42}$$

Finally, this leads to the error in our estimate,

$$\delta \mathbf{b_k} = \sqrt{\Delta \chi_0^2 (\alpha^{-1})_{kk}} \tag{5.43}$$

There is an important caveat in this analysis. Our $\{y_i\}$ are *deterministic*, they are *not* the result of randomness in an experiment. Thus, it is incorrect to quote confidence intervals. Nevertheless, minimizing the dimensionless χ^2 remains a valid method to estimate the best fit parameters. Also,

$$
\begin{aligned}
\chi^2(\mathbf{b}^* + \delta \mathbf{b}^*) &\simeq \chi^2(\mathbf{b}^*) \\
&+ \frac{1}{2} \sum_{kl} \delta \mathbf{b_k} \left. \frac{\partial^2 \chi^2}{\partial \mathbf{b_k^2}} \right|_{b_k^*} \left. \frac{\partial^2 \chi^2}{\partial \mathbf{b_l^2}} \right|_{b_l^*} \delta \mathbf{b_l} , \\
&\simeq \chi^2(\mathbf{b}^*) + \sum_{kj} \delta \mathbf{b_k^*} \, \alpha_{kj} \, \delta \mathbf{b_j^*} , \\
\Rightarrow \Delta \chi^2 &\simeq \sum_{kj} \delta \mathbf{b_k^*} \, \alpha_{kj} \, \delta \mathbf{b_j^*} .
\end{aligned}
$$

$$\tag{5.44}$$
$$\tag{5.45}$$
$$\tag{5.46}$$

In essence, α measures the curvature of χ^2. The errors we quote are

$$\delta \mathbf{b_i^*} = \sqrt{(\alpha^{-1})_{kk}} . \tag{5.47}$$

This is just a measure of the change in $\mathbf{b_i^*}$ needed to increase χ^2 by unity.

Bibliography

[1] B. Abbott et al. Upper limits on a stochastic background of gravitational waves. *Phys. Rev. Lett.*, 95:221101, 2005. LIGO Scientific Collaboration.

[2] K. A. Arnaud, S. Babak, J. G. Baker, M. J. Benacquista, N. J. Cornish, C. Cutler, L. S. Finn, S. L. Larson, T. Littenberg, E. K. Porter, M. Vallisneri, A. Vecchio, J.-Y. Vinet, and T. M. L. Data Challenge Task Force. An overview of the second round of the Mock LISA Data Challenges. *Classical and Quantum Gravity*, 24:551, October 2007.

[3] 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. A How-To for the Mock LISA Data Challenges. In S. M. Merkovitz and J. C. Livas, editors, *Laser Interferometer Space Antenna: 6th International LISA Symposium*, volume 873 of *American Institute of Physics Conference Series*, page 625, November 2006.

[4] 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 S. M. Merkovitz and J. C. Livas, editors, *Laser Interferometer Space Antenna: 6th International LISA Symposium*, volume 873 of *American Institute of Physics Conference Series*, pages 619–624, November 2006.

[5] K. G. Arun, B. R. Iyer, B. S. Sathyaprakash, and P. A. Sundararajan. Parameter estimation of inspiralling compact binaries using 3.5 post-Newtonian gravitational wave phasing: The nonspinning case. *Phys. Rev. D*, 71(8):084008, April 2005.

[6] S. Babak et al. Report on the second Mock LISA data challenge. *Classical and Quantum Gravity*, 25(11):114037, June 2008.

[7] S. Babak et al. The Mock LISA Data Challenges: from Challenge 1B to Challenge 3. *Classical and Quantum Gravity*, 25(18):184026, September 2008.

[8] 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*, 75(2):024005, January 2007.

[9] J. G. Baker, W. D. Boggs, J. Centrella, B. J. Kelly, S. T. McWilliams, M. C. Miller, and J. R. van Meter. Modeling Kicks from the Merger of Generic Black Hole Binaries. *Astrophys. J.*, 682:L29–L32, July 2008.

[10] J. G. Baker, J. Centrella, D.-I. Choi, M. Koppitz, J. R. van Meter, and M. C. Miller. Getting a Kick Out of Numerical Relativity. *Astrophys. J.*, 653:L93–L96, December 2006.

[11] L. Barack and C. Cutler. LISA capture sources: Approximate waveforms, signal-to-noise ratios, and parameter estimation accuracy. *Phys. Rev. D*, 69(8):082005, April 2004.

[12] L. Barack and C. Cutler. Using LISA extreme-mass-ratio inspiral sources to test off-Kerr deviations in the geometry of massive black holes. *Phys. Rev. D*, 75(4):042003, February 2007.

[13] L. Barack and N. Sago. Gravitational self-force on a particle in circular orbit around a Schwarzschild black hole. *Phys. Rev. D*, 75(6):064021, March 2007.

[14] J. M. Bardeen, W. H. Press, and S. A. Teukolsky. Rotating Black Holes: Locally Nonrotating Frames, Energy Extraction, and Scalar Synchrotron Radiation. *Astrophys. J.*, 178:347–370, December 1972.

[15] M. C. Begelman, R. D. Blandford, and M. J. Rees. Massive black hole binaries in active galactic nuclei. *Nature*, 287:307–309, September 1980.

[16] J. D. Bekenstein. Gravitational-Radiation Recoil and Runaway Black Holes. *Astrophys. J.*, 183:657–664, July 1973.

[17] L. Blanchet. Gravitational radiation from post-newtonian sources and inspiralling compact binaries. *Living Reviews in Relativity*, 9(4), 2006. http://www.livingreviews.org/lrr-2006-4.

[18] L. Blanchet, M. S. S. Qusailah, and C. M. Will. Gravitational Recoil of Inspiraling Black Hole Binaries to Second Post-Newtonian Order. *Astrophys. J.*, 635:508–515, December 2005.

[19] W. B. Bonnor and M. A. Rotenberg. Transport of Momentum by Gravitational Waves: The Linear Approximation. *Royal Society of London Proceedings Series A*, 265:109–116, December 1961.

[20] A. Buonanno and T. Damour. Transition from inspiral to plunge in binary black hole coalescences. *Phys. Rev. D*, 62(6):064015, September 2000.

[21] L. M. Burko. Private communication.

[22] L. M. Burko, S. A. Hughes, and P. A. Sundararajan. In preparation.

[23] L. M. Burko and G. Khanna. Accurate time-domain gravitational waveforms for extreme-mass-ratio binaries. *Europhysics Letters*, 78:60005, June 2007.

[24] M. Campanelli and C. O. Lousto. Second order gauge invariant gravitational perturbations of a Kerr black hole. *Phys. Rev. D*, 59(12):124022, June 1999.

[25] M. Campanelli, C. O. Lousto, Y. Zlochower, and D. Merritt. Maximum Gravitational Recoil. *Physical Review Letters*, 98(23):231102, June 2007.

[26] S. M. Carroll. *Spacetime and geometry*. Addison-Wesley, San Francisco, 2004.

[27] S. Chandrasekhar. *The mathematical theory of black holes*. Research supported by NSF. Oxford/New York, Clarendon Press/Oxford University Press (International Series of Monographs on Physics. Volume 69), 1983, 663 p., 1983.

[28] N. A. Collins and S. A. Hughes. Towards a formalism for mapping the spacetimes of massive compact objects: Bumpy black holes and their orbits. *Phys. Rev. D*, 69(12):124022, June 2004.

[29] N. J. Cornish and E. K. Porter. Searching for massive black hole binaries in the first Mock LISA Data Challenge. *Classical and Quantum Gravity*, 24:501, October 2007.

[30] C. Cutler and K. S. Thorne. An overview of gravitational-wave sources. In N. Bishop and S. Maharaj, editors, *Proceedings of GR16*, Durban, South Africa, 2002. World Scientific.

[31] C. Cutler and M. Vallisneri. LISA detections of massive black hole inspirals: Parameter extraction errors due to inaccurate template waveforms. *Phys. Rev. D*, 76(10):104018, November 2007.

[32] T. Damour. Introductory Lectures on the Effective One Body Formalism. *International Journal of Modern Physics A*, 23:1130–1148, 2008.

[33] S. Drasco. Private communication.

[34] S. Drasco, É. É. Flanagan, and S. A. Hughes. Computing inspirals in Kerr in the adiabatic regime: I. The scalar case. *Classical and Quantum Gravity*, 22:801, August 2005.

[35] S. Drasco and S. A. Hughes. Rotating black hole orbit functionals in the frequency domain. *Phys. Rev. D*, 69(4):044015, February 2004.

[36] S. Drasco and S. A. Hughes. Gravitational wave snapshots of generic extreme mass ratio inspirals. *Phys. Rev. D*, 73:024027, 2006.

[37] B. Engquist, A.-K. Tornberg, and R. Tsai. Discretization of Dirac delta functions in level set methods. *Journal of Computational Physics*, 207:28–51, July 2005.

[38] M. Favata, S. A. Hughes, and D. E. Holz. How Black Holes Get Their Kicks: Gravitational Radiation Recoil Revisited. *Astrophys. J.*, 607:L5–L8, May 2004.

[39] M. J. Fitchett. The influence of gravitational wave momentum losses on the centre of mass motion of a Newtonian binary system. *Mon. Not. R. Astron. Soc.*, 203:1049–1062, June 1983.

[40] E. E. Flanagan and S. A. Hughes. The basics of gravitational wave theory. *New Journal of Physics*, 7:204, 2005.

[41] J. L. Friedman. Private communication.

[42] 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. *Classical and Quantum Gravity*, 21:1595, October 2004.

[43] J. R. Gair and K. Glampedakis. Improved approximate inspirals of test bodies into Kerr black holes. *Phys. Rev. D*, 73(6):064037, March 2006.

[44] J. R. Gair, C. Li, and I. Mandel. Observable properties of orbits in exact bumpy spacetimes. *Phys. Rev. D*, 77(2):024035, January 2008.

[45] N. Gehrels et al. A short gamma-ray burst apparently associated with an elliptical galaxy at redshift z = 0.225. *Nature*, 437:851–854, 2005.

[46] GEO. http://geo600.aei.mpg.de/.

[47] B. F. Gerke, J. A. Newman, J. Lotz, R. Yan, P. Barmby, A. L. Coil, C. J. Conselice, R. J. Ivison, L. Lin, D. C. Koo, K. Nandra, S. Salim, T. Small, B. J. Weiner, M. C. Cooper, M. Davis, S. M. Faber, and P. Guhathakurta. The DEEP2 Galaxy Redshift Survey: AEGIS Observations of a Dual AGN at z = 0.7. *Astrophys. J.*, 660:L23–L26, May 2007.

[48] K. Glampedakis and S. Babak. Mapping spacetimes with LISA: inspiral of a test body in a 'quasi-Kerr' field. *Classical and Quantum Gravity*, 23:4167–4188, June 2006.

[49] H. Goldstein. *Classical mechanics*. Addison-Wesley, Reading, 1980.

[50] J. A. González, M. Hannam, U. Sperhake, B. Brügmann, and S. Husa. Supermassive Recoil Velocities for Binary Black-Hole Mergers with Antialigned Spins. *Physical Review Letters*, 98(23):231101, June 2007.

[51] F. Herrmann, I. Hinder, D. M. Shoemaker, P. Laguna, and R. A. Matzner. Binary black holes: Spin dynamics and gravitational recoil. *Phys. Rev. D*, 76(8):084032, October 2007.

[52] C. Hopman. Astrophysics of extreme mass ratio inspiral sources . In S. M. Merkovitz and J. C. Livas, editors, *Laser Interferometer Space Antenna: 6th International LISA Symposium*, volume 873 of *American Institute of Physics Conference Series*, page 241, November 2006.

[53] S. A. Hughes. Evolution of circular, nonequatorial orbits of Kerr black holes due to gravitational-wave emission. *Phys. Rev. D*, 61(8):084004, April 2000.

[54] S. A. Hughes, S. Drasco, É. É. Flanagan, and J. Franklin. Gravitational Radiation Reaction and Inspiral Waveforms in the Adiabatic Limit. *Physical Review Letters*, 94(22):221101, June 2005.

[55] R. A. Hulse and J. H. Taylor. Discovery of a pulsar in a binary system. *Astrophys. J.*, 195:L51–L53, January 1975.

[56] R. A. Isaacson. Gravitational Radiation in the Limit of High Frequency. II. Nonlinear Terms and the Effective Stress Tensor. *Physical Review*, 166:1272–1279, February 1968.

[57] J. D. Jackson. *Classical Electrodynamics*. Wiley, New York, 1999.

[58] D. Kennefick and A. Ori. Radiation-reaction-induced evolution of circular orbits of particles around Kerr black holes. *Phys. Rev. D*, 53:4319–4326, April 1996.

[59] G. Khanna. Teukolsky evolution of particle orbits around Kerr black holes in the time domain: Elliptic and inclined orbits. *Phys. Rev. D*, 69(2):024016, January 2004.

[60] G. Khanna, R. Gleiser, R. Price, and J. Pullin. Close limit of grazing black hole collisions: non-spinning holes. *New Journal of Physics*, 2:3, March 2000.

[61] K. D. Kokkotas. Normal modes of the Kerr black hole. *Classical and Quantum Gravity*, 8:2217–2224, December 1991.

[62] S. Komossa, V. Burwitz, G. Hasinger, P. Predehl, J. S. Kaastra, and Y. Ikebe. Discovery of a Binary Active Galactic Nucleus in the Ultraluminous Infrared Galaxy NGC 6240 Using Chandra. *Astrophys. J.*, 582:L15–L19, January 2003.

[63] M. Koppitz, D. Pollney, C. Reisswig, L. Rezzolla, J. Thornburg, P. Diener, and E. Schnetter. Recoil Velocities from Equal-Mass Binary-Black-Hole Mergers. *Physical Review Letters*, 99(4):041102, July 2007.

[64] W. Krivan, P. Laguna, P. Papadopoulos, and N. Andersson. Dynamics of perturbations of rotating black holes. *Phys. Rev. D*, 56:3395–3404, September 1997.

[65] R. N. Lang and S. A. Hughes. Localizing coalescing massive black hole binaries with gravitational waves. *Astrophys. J.*, 677:1184–1200, 2008.

[66] LIGO. http://www.ligo.caltech.edu/.

[67] LISA. http://lisa.nasa.gov.

[68] R. López-Alemán, G. Khanna, and J. Pullin. Perturbative evolution of particle orbits around Kerr black holes: time-domain calculation. *Classical and Quantum Gravity*, 20:3259–3268, July 2003.

[69] C. O. Lousto and Y. Zlochower. Modeling gravitational recoil from precessing highly spinning unequal-mass black-hole binaries. *Phys. Rev. D*, 79(6):064018, March 2009.

[70] H. L. Maness, G. B. Taylor, R. T. Zavala, A. B. Peck, and L. K. Pollack. Breaking All the Rules: The Compact Symmetric Object 0402+379. *Astrophys. J.*, 602:123–134, February 2004.

[71] Y. Mino. Regularization of Gravitational Radiation Reaction Force. *Progress of Theoretical Physics*, 99:79–95, January 1998.

[72] C. W. Misner, K. S. Thorne, and J. A. Wheeler. *Gravitation*. San Francisco: W.H. Freeman and Co., 1973.

[73] E. T. Newman and T. W. J. Unti. Behavior of Asymptotically Flat Empty Spaces. *Journal of Mathematical Physics*, 3:891–901, September 1962.

[74] A. Ori and 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*, 62(12):124022, December 2000.

[75] R. O'Shaughnessy. Transition from inspiral to plunge for eccentric equatorial Kerr orbits. *Phys. Rev. D*, 67(4):044004, February 2003.

[76] B. J. Owen. Search templates for gravitational waves from inspiraling binaries: Choice of template spacing. *Phys. Rev. D*, 53:6749–6761, June 1996.

[77] E. Pazos-Avalos and C. O. Lousto. Numerical integration of the Teukolsky Equation in the time domain. *Phys. Rev. D*, page 084022, 2005.

[78] A. Peres. Classical Radiation Recoil. *Physical Review*, 128:2471–2475, December 1962.

[79] P. C. Peters. Gravitational Radiation and the Motion of Two Point Masses. *Physical Review*, 136:1224–1232, November 1964.

[80] E. Poisson. The Motion of Point Particles in Curved Spacetime. *Living Reviews in Relativity*, 7:6, May 2004.

[81] E. Poisson and C. M. Will. Gravitational waves from inspiraling compact binaries: Parameter estimation using second-post-Newtonian waveforms. *Phys. Rev. D*, 52:848–855, July 1995.

[82] A. Pound and E. Poisson. Osculating orbits in Schwarzschild spacetime, with an application to extreme mass-ratio inspirals. *Phys. Rev. D*, 77(4):044013, February 2008.

[83] A. Pound, E. Poisson, and B. G. Nickel. Limitations of the adiabatic approximation to the gravitational self-force. *Phys. Rev. D*, 72(12):124001, December 2005.

[84] W. H. Press, B. P. Flannery, and S. A. Teukolsky. *Numerical recipes. The art of scientific computing.* Cambridge: University Press, 1986, 1986.

[85] F. Pretorius. Binary Black Hole Coalescence. *ArXiv e-prints 0710.1338*, October 2007.

[86] C. Rodriguez, G. B. Taylor, R. T. Zavala, A. B. Peck, L. K. Pollack, and R. W. Romani. A Compact Supermassive Binary Black Hole System. *Astrophys. J.*, 646:49–60, July 2006.

[87] F. D. Ryan. Effect of gravitational radiation reaction on nonequatorial orbits around a Kerr black hole. *Phys. Rev. D*, 53:3064–3069, March 1996.

[88] N. Sago, T. Tanaka, W. Hikida, K. Ganz, and H. Nakano. Adiabatic Evolution of Orbital Parameters in Kerr Spacetime. *Progress of Theoretical Physics*, 115:873–907, May 2006.

[89] N. Sago, T. Tanaka, W. Hikida, and H. Nakano. Adiabatic Radiation Reaction to Orbits in Kerr Spacetime. *Progress of Theoretical Physics*, 114:509–514, August 2005.

[90] W. Schmidt. Celestial mechanics in Kerr spacetime. *Classical and Quantum Gravity*, 19:2743–2764, May 2002.

[91] J. D. Schnittman and A. Buonanno. The Distribution of Recoil Velocities from Merging Black Holes. *Astrophys. J.*, 662:L63–L66, June 2007.

[92] J. D. Schnittman, A. Buonanno, J. R. van Meter, J. G. Baker, W. D. Boggs, J. Centrella, B. J. Kelly, and S. T. McWilliams. Anatomy of the binary black hole recoil: A multipolar analysis. *Phys. Rev. D*, 77(4):044031, February 2008.

[93] C. F. Sopuerta. Time Domain Simulations of EMRIs using Finite Element Methods. In S. M. Merkovitz and J. C. Livas, editors, *Laser Interferometer Space Antenna: 6th International LISA Symposium*, volume 873 of *American Institute of Physics Conference Series*, pages 274–278, November 2006.

[94] C. F. Sopuerta and P. Laguna. Finite element computation of the gravitational radiation emitted by a pointlike object orbiting a nonrotating black hole. *Phys. Rev. D*, 73(4):044028, February 2006.

[95] C. F. Sopuerta, N. Yunes, and P. Laguna. Gravitational recoil from binary black hole mergers: The close-limit approximation. *Phys. Rev. D*, 74(12):124010, December 2006.

[96] 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*, 77(12):124050, June 2008.

[97] P. A. Sundararajan, G. Khanna, and S. A. Hughes. Towards adiabatic waveforms for inspiral into kerr black holes: A new model of the source for the time domain perturbation equation. *Phys. Rev. D*, 76:104005, 2007.

[98] P. A. Sundararajan, G. Khanna, S. A. Hughes, and S. Drasco. Towards adiabatic waveforms for inspiral into Kerr black holes: II. Dynamical sources and generic orbits . *Phys. Rev. D*, 78:024022, 2008.

[99] S. A. Teukolsky. Rotating black holes: Separable wave equations for gravitational and electromagnetic perturbations. *Physical Review Letters*, 29:1114–1118, 1972.

[100] S. A. Teukolsky. Perturbations of a Rotating Black Hole. I. Fundamental Equations for Gravitational, Electromagnetic, and Neutrino-Field Perturbations. *Astrophys. J.*, 185:635–648, October 1973.

[101] W. Tichy and É. É. Flanagan. Angular momentum ambiguities in asymptotically flat spacetimes which are perturbations of stationary spacetimes. In V. G. Gurzadyan, R. T. Jantzen, and R. Ruffini, editors, *The Ninth Marcel Grossmann Meeting*, pages 1622–1623, 2002.

[102] A.-K. Tornberg and B. Engquist. Numerical approximations of singular source terms in differential equations. *Journal of Computational Physics*, 200:462–488, November 2004.

[103] M. J. Valtonen, K. Nilsson, H. Lehto, A. Sillanpää, L. Takalo, M. Kidger, G. Poyner, T. Pursimo, J. Heidt, K. Sadakane, J. Wu, A. Liakos, V. Hentunen, and M. Nissinen. Confirmation of the Gravitational Wave Energy Loss in the Binary Black Hole System OJ287. In *Bulletin of the American Astronomical Society*, volume 38 of *Bulletin of the American Astronomical Society*, page 942, December 2007.

[104] VIRGO. http://wwwcascina.virgo.infn.it/.

[105] J.M. Weisberg and J.H. Taylor. The relativistic binary pulsar b1913+16: Thirty years of observations and analysis. In F.A. Rasio and I.H. Stairs, editors, *Binary Radio Pulsars*, volume 328 of *ASP Conference Series*, pages 25–32, San Francisco, 2005. Astronomical Society of the Pacific.

[106] Wikipedia. http://en.wikipedia.org/wiki/Crab_Nebula.