Two-timescale analysis of extreme mass ratio inspirals in Kerr spacetime: Orbital motion

Tanja Hinderer[1] and Éanna É. Flanagan[1,2]

[1]*Center for Radiophysics and Space Research, Cornell University, Ithaca, New York 14853, USA*
[2]*Laboratory for Elementary Particle Physics, Cornell University, Ithaca, New York 14853, USA*

(Received 18 June 2008; published 11 September 2008)

Inspirals of stellar-mass compact objects into massive black holes are an important source for future gravitational wave detectors such as Advanced LIGO and LISA. The detection and analysis of these signals rely on accurate theoretical models of the binary dynamics. We cast the equations describing binary inspiral in the extreme mass ratio limit in terms of action-angle variables, and derive properties of general solutions using a two-timescale expansion. This provides a rigorous derivation of the prescription for computing the leading order orbital motion. As shown by Mino, this leading order or adiabatic motion requires only knowledge of the orbit-averaged, dissipative piece of the self-force. The two-timescale method also gives a framework for calculating the post-adiabatic corrections. For circular and for equatorial orbits, the leading order corrections are suppressed by one power of the mass ratio, and give rise to phase errors of order unity over a complete inspiral through the relativistic regime. These post-1-adiabatic corrections are generated by the fluctuating, dissipative piece of the first order self-force, by the conservative piece of the first order self-force, and by the orbit-averaged, dissipative piece of the second order self-force. We also sketch a two-timescale expansion of the Einstein equation, and deduce an analytic formula for the leading order, adiabatic gravitational waveforms generated by an inspiral.

DOI: 10.1103/PhysRevD.78.064028 PACS numbers: 04.30.Db, 04.25.Nx, 95.30.Sf, 97.60.Lf

I. INTRODUCTION AND SUMMARY

A. Background and motivation

Recent years have seen great progress in our understanding of the two body problem in general relativity. Binary systems of compact bodies undergo an inspiral driven by gravitational radiation reaction until they merge. As illustrated in Fig. 1, there are three different regimes in the dynamics of these systems, depending on the values of the total and reduced masses M and μ of the system and the orbital separation r: (i) the early, weak field regime at $r \gg M$, which can be accurately modeled using post-Newtonian theory (see, for example, the review [1]); (ii) the relativistic, equal mass regime $r \sim M$, $\mu \sim M$, which must be treated using numerical relativity. Over the last few years, numerical relativists have succeeded for the first time in simulating the merger of black hole binaries (see, for example, the review [2] and references therein). The last regime is (iii) the relativistic, extreme mass ratio regime $r \sim M$, $\mu \ll M$. Over timescales short compared to the dephasing time $\sim M\sqrt{M/\mu}$, systems in this regime can be accurately modeled using black hole perturbation theory [3], with the mass ratio $\varepsilon \equiv \mu/M$ serving as the expansion parameter. The subject of this paper is the approximation methods that are necessary to treat such systems over the longer inspiral timescale $\sim M^2/\mu$ necessary for computation of complete inspirals.

This extreme mass ratio regime has direct observational relevance: Compact objects spiraling into much larger black holes are expected to be a key source for both LIGO and LISA. *Intermediate-mass-ratio inspirals* (IMRIs) are inspirals of black holes or neutron stars into intermediate mass ($50 \leq M \leq 1000 M_\odot$) black holes; these would be visible to Advanced LIGO out to distances of several hundred Mpc [4], where the event rate could be about 3–30 per year [4,5]. *Extreme mass ratio inspirals* (EMRIs) are inspirals of stellar-mass compact objects (black holes, neutron stars, or possibly white dwarfs) into massive ($10^4 \leq M \leq 10^7 M_\odot$) black holes in galactic nuclei; these will be visible to LISA out to redshifts $z \approx 1$

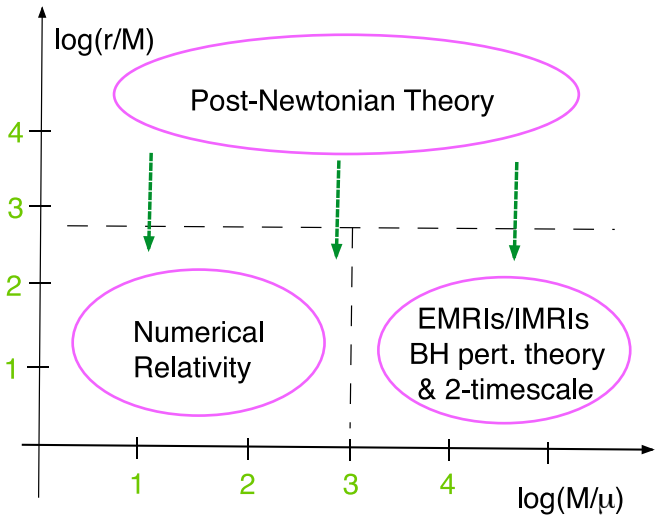

FIG. 1 (color online). The parameter space of inspiralling compact binaries in general relativity, in terms of the inverse mass ratio $M/\mu = 1/\varepsilon$ and the orbital radius r, showing the different regimes and the computational techniques necessary in each regime. Individual binaries evolve downwards in the diagram (green dashed arrows).

 © 2008 The American Physical Society

[6–8]. It has been estimated [9,10] that LISA should see about 50 such events per year, based on calculations of stellar dynamics in galaxies' central cusps [11]. Because of an IMRI's or EMRI's small mass ratio $\varepsilon = \mu/M$, the small body lingers in the large black hole's strong-curvature region for many wave cycles before merger: hundreds of cycles for LIGO's IMRIs; hundreds of thousands for LISA's EMRIs [6]. In this relativistic regime the post-Newtonian approximation has completely broken down, and full numerical relativity simulations become prohibitively difficult as ε is decreased. Modeling of these sources therefore requires a specialized approximation method.

Gravitational waves from these sources will be rich with information [7,8]:

(i) The waves carry not only the details of the evolving orbit, but also a map of the large body's spacetime geometry, or equivalently the values of all its multipole moments, as well as details of the response of the horizon to tidal forces [12,13]. Extracting the map (*bothrodesy*) is a high priority for LISA, which can achieve ultrahigh accuracy, and a moderate priority for LIGO, which will have a lower (but still interesting) accuracy [4]. Measurements of the black hole's quadrupole (with a fractional accuracy of about 10^{-3} for LISA [14,15], about 1 for Advanced LIGO [4]) will enable tests of the black hole's *no hair property*, that all of the mass and current multipole moments are uniquely determined in terms of the first two, the mass and spin. Potentially, these measurements could lead to the discovery of non-black-hole central objects such as boson stars [16,17] or naked singularities.

(ii) One can measure the mass and spin of the central black hole with fractional accuracies of order 10^{-4} for LISA [18,19] and about 10^{-2}–10^{-1} for Advanced LIGO [4]. Observing many events will therefore provide a census of the masses and spins of the massive central black holes in nonactive galactic nuclei like M31 and M32. The spin can provide useful information about the hole's growth history (mergers versus accretion) [20].

(iii) For LISA, one can measure the inspiralling objects' masses with precision about 10^{-4}, teaching us about the stellar population in the central parsec of galactic nuclei.

(iv) If the LISA event rate is large enough, one can measure the Hubble constant H_0 to about 1% [21], which would indirectly aid dark energy studies [22]. The idea is to combine the measured luminosity distance of cosmological ($z \sim 1/2$) EMRIs with a statistical analysis of the redshifts of candidate host galaxies located within the error box on the sky.

To realize the science goals for these sources requires accurate theoretical models of the waveforms for matched filtering. The accuracy requirement is roughly that the theoretical template's phase must remain accurate to ~ 1 cycle over the $\sim \varepsilon^{-1}$ cycles of waveform in the highly relativistic regime ($\sim 10^2$ cycles for LIGO, $\sim 10^5$ for LISA). For signal detection, the requirement is slightly less stringent than this, while for parameter extraction the requirement is slightly more stringent: The waveforms are characterized by 14 parameters, which makes a fully coherent search of the entire data train computationally impossible. Therefore, detection templates for LISA will use short segments of the signal and require phase coherence for $\sim 10^4$ cycles [10]. Once the presence of a signal has been established, the source parameters will be extracted using measurement templates that require a fractional phase accuracy of order the reciprocal of the signal-to-noise ratio [10], in order to keep systematic errors as small as the statistical errors.

B. Methods of computing orbital motion and waveforms

A variety of approaches to computing waveforms have been pursued in the community. We now review these approaches in order to place the present paper in context. The foundation for all approaches is the fact that, since $\varepsilon = \mu/M \ll 1$, the field of the compact object can be treated as a small perturbation to the large black hole's gravitational field. On short timescales $\sim M$, the compact object moves on a geodesic of the Kerr geometry, characterized by its conserved energy E, z component of angular momentum L_z, and Carter constant Q. Over longer timescales $\sim M/\varepsilon$, radiation reaction causes the parameters E, L_z, and Q to evolve adiabatically and the orbit to shrink. The effect of the internal structure of the object is negligible,[1] so it can be treated as a point particle. At the end of the inspiral, the particle passes through an innermost stable orbit where adiabaticity breaks down, and it transitions onto a geodesic plunge orbit [27–30]. In this paper we restrict attention to the adiabatic portion of the motion.

Numerical relativity.—Numerical relativity has not yet been applied to the extreme mass ratio regime. However, given the recent successful simulations in the equal mass regime $\varepsilon \sim 1$, one could contemplate trying to perform simulations with smaller mass ratios. There are a number of difficulties that arise as ε gets small: (i) The orbital timescale and the radiation-reaction timescale are separated by the large factor $\sim 1/\varepsilon$. The huge number of

[1]There are two exceptions, where corrections to the point-particle model can be important: (i) white-dwarf EMRIs, where tidal interactions can play a role [23]; (ii) the effect due to the spin, if any, of the inspiralling object, whose importance has been emphasized by Burko [24,25]. While this effect is at most marginally relevant for signal detection [26], it is likely quite important for information extraction. We neglect the spin effect in the present paper, since it can be computed and included in the waveforms relatively easily.

wave cycles implies an impractically large computation time. (ii) There is a separation of length scales: the compact object is smaller than the central black hole by a factor ε. (iii) Most importantly, in the strong field region near the small object, the piece of the metric perturbation responsible for radiation reaction is of order ε, and since one requires errors in the radiation reaction to be of order ε, the accuracy requirement on the metric perturbation is of order ε^2. These difficulties imply that numerical simulations will likely not be possible in the extreme mass ratio regime in the foreseeable future, unless major new techniques are devised to speed up computations.

Use of post-Newtonian methods.—Approximate waveforms which are qualitatively similar to real waveforms can be obtained using post-Newtonian methods or using hybrid schemes containing some post-Newtonian elements [26,31,32]. Although these waveforms are insufficiently accurate for the eventual detection and data analysis of real signals, they have been very useful for approximately scoping out the detectability of inspiral events and the accuracy of parameter measurement, both for LIGO [4] and LISA [10,26]. They have the advantage that they can be computed relatively quickly.

Black hole perturbation theory: first order.—There is a long history of using first order perturbation theory [3] to compute gravitational waveforms from particles in geodesic orbits around black holes [33–36]. These computations have recently been extended to fully generic orbits [37–39]. However, first order perturbation theory is limited to producing "snapshot" waveforms that neglect radiation reaction.[2] Such waveforms fall out of phase with true waveforms after a dephasing time $\sim M/\sqrt{\varepsilon}$, the geometric mean of the orbital and radiation-reaction timescales, and so are of limited utility.[3]

Black hole perturbation theory: second order—One can, in principle, go to second order in perturbation theory [41–43]. At this order, the particle's geodesic motion must be corrected by self-force effects describing its interaction with its own spacetime distortion. This gravitational self-force is analogous to the electromagnetic Abraham-Lorentz-Dirac force. Although a formal expression for the self-force is known [44,45], it has proved difficult to translate this expression into a practical computational scheme for Kerr black holes because of the mathematical complexity of the self-field regularization which is required. Research into this topic is ongoing; see, for example, the review [46] and Refs. [43,47–54] for various approaches and recent progress.

When the self-force is finally successfully computed, second order perturbation theory will provide a self-consistent framework for computing the orbital motion and the waveform, but only over short timescales. The second order waveforms will fall out of phase with the true waveforms after only a dephasing time $\sim M/\sqrt{\varepsilon}$[4] [55,56]. Computing accurate waveforms describing a full inspiral therefore requires going beyond black hole perturbation theory.

Use of conservation laws.—This well-explored method allows tracking an entire inspiral for certain special classes of orbits. Perturbation theory is used to compute the fluxes of E and L_z to infinity and down the horizon for geodesic orbits, and imposing global conservation laws, one infers the rates of change of the orbital energy and angular momentum. For circular orbits and equatorial orbits these determine the rate of change of the Carter constant Q, and thus the inspiralling trajectory. The computation can either be done in the frequency domain [33–36], or in the time domain by numerically integrating the Teukolsky equation as a $2+1$ partial differential equation with a suitable numerical model of the point-particle source [57–66]. However, this method fails for generic orbits since there is no known global conservation law associated with the Carter constant Q.

Adiabatic approximation: leading order.—Over the last few years, it has been discovered how to compute inspirals to leading order for generic orbits. The method is based on Mino's realization [67] that, in the adiabatic limit, one needs only the time-averaged, dissipative piece of the first order self-force, which can be straightforwardly computed from the half retarded minus half advanced prescription. This sidesteps the difficulties associated with regularization that impede computations of the full, first order self-force. From the half advanced minus half retarded prescription, one can derive an explicit formula for a time average of $\dot{Q}$ in terms of mode expansion [37,68–71]. Using this formula it will be straightforward to compute inspirals to the leading order.

We now recap and assess the status of these various approaches. All of the approaches described above have shortcomings and limitations [56]. Suppose that one computes the inspiral motion, either from conservation laws, or from the time-averaged dissipative piece of the first order self-force, or from the exact first order self-force when it becomes available. It is then necessary to compute the radiation generated by this inspiral. One might be tempted

[2]The source for the linearized Einstein equation must be a conserved stress-energy tensor, which for a point particle requires a geodesic orbit.

[3]Drasco has argued that snapshot waveforms may still be useful for signal detections in certain limited parts of the IMRI/EMRI parameter space, since the phase coherence time is actually $\sim 100M/\sqrt{\varepsilon}$ [40].

[4]The reason is as follows. Geodesic orbits and true orbits become out of phase by ~ 1 cycle after a dephasing time. Therefore, since the linear metric perturbation is sourced by a geodesic orbit, fractional errors in the linear metric perturbation must be of order unity. Therefore, the second order metric perturbation must become comparable to the first order term after a dephasing time.

to use linearized perturbation theory for this purpose. However, two problems then arise:

(i) As noted above, the use of linearized perturbation theory with nongeodesic sources is mathematically inconsistent. This inconsistency has often been remarked upon, and various *ad hoc* methods of modifying the linearized theory to get around the difficulty have been suggested or implemented [45,72,73].

(ii) A related problem is that the resulting waveforms will depend on the gauge chosen for the linearized metric perturbation, whereas the exact waveforms must be gauge invariant.

It has often been suggested that these problems can be resolved by going to second order in perturbation theory [43,46]. However, as discussed above, a second order computation will be valid only for a dephasing time, and not for a full inspiral.

Of course, the above problems are not fatal, since the motion is locally very nearly geodesic, and so the inconsistencies and ambiguities are suppressed by a factor $\sim\varepsilon$ relative to the leading order waveforms.[5] Nevertheless, it is clearly desirable to have a well-defined approximation method that gives a unique, consistent result for the leading order waveform. Also, for parameter extraction, it will be necessary to compute the phase of the waveform beyond the leading order. For this purpose it will clearly be necessary to have a more fundamental computational framework.

C. The two-timescale expansion method

In this paper we describe an approximation scheme which addresses and resolves all of the theoretical difficulties discussed above. It is based on the fact that the systems evolve adiabatically: the radiation-reaction timescale $\sim M/\varepsilon$ is much longer than the orbital timescale $\sim M$ [67]. It uses two-timescale expansions, which are a systematic method for studying the cumulative effect of a small disturbance on a dynamical system that is active over a long time [74].

The essence of the method is simply an ansatz for the dependence of the metric $g_{ab}(\varepsilon)$ on ε, and an ansatz for the dependence of the orbital motion on ε, which are justified *a posteriori* order by order via substitution into Einstein's equation. The ansatz for the metric is more complex than the Taylor series ansatz which underlies standard perturbation theory. The two-timescale method has roughly the same relationship to black hole perturbation theory as post-Newtonian theory has to perturbation theory of Minkowski spacetime. The method is consistent with standard black hole perturbation theory locally in time, at each instant, but extends the domain of validity to an entire inspiral. The method provides a systematic procedure for computing the leading order waveforms, which we call the adiabatic waveforms, as well as higher order corrections. We call the $O(\varepsilon)$ corrections the post-1-adiabatic corrections, the $O(\varepsilon^2)$ corrections post-2-adiabatic, etc., paralleling the standard terminology in post-Newtonian theory.

The use of two-timescale expansions in the extreme mass ratio regime was first suggested in Refs. [55,75]. The method has already been applied to some simplified model problems: a computation of the inspiral of a point particle in Schwarzschild spacetime subject to electromagnetic radiation-reaction forces by Pound and Poisson [76], and a computation of the scalar radiation generated by a inspiralling particle coupled to a scalar field by Mino and Price [77]. We will extend and generalize these analyses, and develop a complete approximation scheme.

There are two independent parts to the approximation scheme. The first is a two-timescale analysis of the inspiralling orbital motion, which is the focus of the present paper. Our formalism enables us to give a rigorous derivation and clarification of the prescription for computing the leading order motion that is valid for all orbits, and resolves some controversies in the literature [76]. It also allows us to systematically calculate the higher order corrections. For these corrections, we restrict our attention to inspirals in Schwarzschild spacetime, and to circular and equatorial inspirals in Kerr spacetime. Fully generic inspirals in Kerr spacetime involve a qualitatively new feature—the occurrence of transient resonances—which we will discuss in the forthcoming papers [78,79].

The second part of the approximation scheme is the application of the two-timescale method to the Einstein equation, and a meshing of that expansion to the analysis of the orbital motion. This allows computation of the observable gravitational waveforms, and is described in detail in the forthcoming paper [80]. We briefly sketch this formalism in Sec. I E below, and give an analytic result for the leading order waveforms.

We note that alternative methods of attempting to overcome the problems with standard perturbation theory, and of going beyond adiabatic order, have been developed by Mino [56,72,81,82]. These methods have some overlap with the method discussed here, but differ in some crucial aspects. We do not believe that these methods provide a systematic framework for going to higher orders, unlike the two-timescale method.

D. Orbital motion

We now turn to a description of our two-timescale analysis of the orbital motion. The first step is to exploit the Hamiltonian structure of the unperturbed, geodesic motion to rewrite the governing equations in terms of generalized action-angle variables. We start from the

[5]This is true both for the instantaneous amplitude and for the accumulated phase of the waveform.

forced geodesic equation

$$\frac{d^2x^\nu}{d\tau^2} + \Gamma^\nu_{\sigma\rho}\frac{dx^\sigma}{d\tau}\frac{dx^\rho}{d\tau} = \varepsilon a^{(1)\nu} + \varepsilon^2 a^{(2)\nu} + O(\varepsilon^3). \quad (1.1)$$

Here τ is proper time and $a^{(1)\nu}$ and $a^{(2)\nu}$ are the first order and second order self-accelerations. In Sec. II we augment these equations to describe the leading order backreaction of the inspiral on the mass M and spin a of the black hole, and show they can be rewritten as [cf. Eqs. (2.47) below]

$$\frac{dq_\alpha}{d\tau} = \omega_\alpha(J_\sigma) + \varepsilon g^{(1)}_\alpha(q_r, q_\theta, J_\sigma) + \varepsilon^2 g^{(2)}_\alpha(q_r, q_\theta, J_\sigma) + O(\varepsilon^3), \quad (1.2a)$$

$$\frac{dJ_\lambda}{d\tau} = \varepsilon G^{(1)}_\lambda(q_r, q_\theta, J_\sigma) + \varepsilon^2 G^{(2)}_\lambda(q_r, q_\theta, J_\sigma) + O(\varepsilon^3). \quad (1.2b)$$

Here the variables J_λ are the three conserved quantities of geodesic motion, with the dependence on the particle mass scaled out, together with the black hole mass and spin parameters:

$$J_\lambda = (E/\mu, L_z/\mu, Q/\mu^2, M, a). \quad (1.3)$$

The variables $q_\alpha = (q_r, q_\theta, q_\phi, q_t)$ are a set of generalized angle variables associated with the r, θ, ϕ, and t motions in Boyer-Lindquist coordinates, and are defined more precisely in Sec. II D below. The variables q_r, q_θ, and q_ϕ each increase by 2π after one cycle of motion of the corresponding variable r, θ, or ϕ. The functions $\omega_\alpha(J_\sigma)$ are the fundamental frequencies of geodesic motion in the Kerr metric. The functions $g^{(1)}_\alpha$, $G^{(1)}_\lambda$ are currently not known explicitly, but their exact form will not be needed for the analysis of this paper. They are determined by the first order self-acceleration [44,45]. Similarly, the functions $g^{(2)}_\alpha$ and $G^{(2)}_\lambda$ are currently not known explicitly, and are determined in part by the second order self-acceleration [83–87]; see Sec. II F for more details.

In Secs. IV and V below we analyze the differential equations (1.2) using two-timescale expansions. In the nonresonant case, and up to post-1-adiabatic order, the results can be summarized as follows. Approximate solutions of the equations can be constructed via a series of steps: (i) We define the slow time variable $\tilde\tau = \varepsilon\tau$. (ii) We construct a set of functions $\psi^{(0)}_\alpha(\tilde\tau)$, $\mathcal{J}^{(0)}_\lambda(\tilde\tau)$, $\psi^{(1)}_\alpha(\tilde\tau)$, and $\mathcal{J}^{(1)}_\lambda(\tilde\tau)$ of the slow time. These functions are defined by a set of differential equations that involve the functions ω_α, $g^{(1)}_\alpha$, $G^{(1)}_\lambda$, $g^{(2)}_\alpha$, and $G^{(2)}_\lambda$, and which are independent of ε [Eqs. (5.26), (5.31), (5.29), (5.39), and (5.37) below]. (iii) We define a set of auxiliary phase variables ψ_α by

$$\psi_\alpha(\tau, \varepsilon) = \frac{1}{\varepsilon}\psi^{(0)}_\alpha(\varepsilon\tau) + \psi^{(1)}_\alpha(\varepsilon\tau) + O(\varepsilon), \quad (1.4)$$

where the $O(\varepsilon)$ symbol refers to the limit $\varepsilon \to 0$ at fixed $\tilde\tau = \varepsilon\tau$. (iv) Finally, the solution to post-1-adiabatic order is given by

$$q_\alpha(\tau, \varepsilon) = \psi_\alpha + O(\varepsilon), \quad (1.5a)$$

$$J_\lambda(\tau, \varepsilon) = \mathcal{J}^{(0)}_\lambda(\varepsilon\tau) + \varepsilon\mathcal{J}^{(1)}_\lambda(\varepsilon\tau) + \varepsilon H_\lambda[\psi_r, \psi_\theta, \mathcal{J}^{(0)}_\sigma(\varepsilon\tau)] + O(\varepsilon^2), \quad (1.5b)$$

where the $O(\varepsilon)$ and $O(\varepsilon^2)$ symbols refer to $\varepsilon \to 0$ at fixed $\tilde\tau$ and ψ_α. Here H_λ is a function which is periodic in its first two arguments and which can be computed from the function $G^{(1)}_\lambda$ [Eq. (7.3) below].

We now turn to a discussion of the implications of the final result (1.5). First, we emphasize that the purpose of the analysis is *not* to give a convenient, practical scheme to integrate the orbital equations of motion. Such a scheme is not needed, since once the self-acceleration is known, it is straightforward to numerically integrate the forced geodesic equations (1.1). Rather, the main benefit of the analysis is to give an analytic understanding of the dependence of the motion on ε in the limit $\varepsilon \to 0$. This serves two purposes. First, it acts as a foundation for the two-timescale expansion of the Einstein equation and the computation of waveforms (Sec. I E below and Ref. [80]). Second, it clarifies the utility of different approximations to the self-force that have been proposed, by determining which pieces of the self-force contribute to the adiabatic order and post-1-adiabatic order motions [37,68]. This second issue is discussed in detail in Sec. VII below. Here we give a brief summary.

Consider first the motion to adiabatic order, given by the functions $\psi^{(0)}_\alpha$ and $\mathcal{J}^{(0)}_\lambda$. These functions are obtained from the differential equations [Eqs. (5.26), (5.31), and (5.29) below]

$$\frac{d\psi^{(0)}_\alpha}{d\tilde\tau}(\tilde\tau) = \omega_\alpha[\mathcal{J}^{(0)}_\sigma(\tilde\tau)], \quad (1.6a)$$

$$\frac{d\mathcal{J}^{(0)}_\lambda}{d\tilde\tau}(\tilde\tau) = \langle G^{(1)}_\lambda\rangle[\mathcal{J}^{(0)}_\sigma(\tilde\tau)], \quad (1.6b)$$

where $\langle\ldots\rangle$ denotes the average[6] over the 2-torus

$$\langle G^{(1)}_\lambda\rangle(J_\sigma) \equiv \frac{1}{(2\pi)^2}\int_0^{2\pi} dq_r \int_0^{2\pi} dq_\theta G^{(1)}_\lambda(q_r, q_\theta, J_\sigma). \quad (1.7)$$

This zeroth order approximation describes the inspiralling motion of the particle. In Sec. II G below we show that only the dissipative (i.e. half retarded minus half advanced) piece of the self-force contributes to the torus average

[6]This phase space average is uniquely determined by the dynamics of the system, and resolves concerns in the literature about inherent ambiguities in the choice of averaging [76].

(1.7). Thus, the leading order motion depends only on the dissipative self-force, as argued by Mino [67]. Our result extends slightly that of Mino, since he advocated using an infinite time average on the right-hand side of Eq. (1.6b), instead of the phase space or torus average. The two averaging methods are equivalent for generic geodesics, but not for geodesics for which the ratio of radial and azimuthal periods is a rational number. The time-average prescription is therefore correct for generic geodesics, while the result (1.6) is valid for all geodesics. The effect of the nongeneric geodesics is discussed in detail in Refs. [78,79].

Consider next the subleading, post-1-adiabatic corrections to the inspiral given by the functions $\psi_\alpha^{(1)}$ and $\mathcal{J}_\lambda^{(1)}$. These corrections are important for assessing the accuracy of the adiabatic approximation, and will be needed for accurate data analysis of detected signals. The differential equations determining $\psi_\alpha^{(1)}$ and $\mathcal{J}_\lambda^{(1)}$ are Eqs. (5.37) and (5.39) below. These equations depend on (i) the oscillating (not averaged) piece of the dissipative, first order self-force; (ii) the conservative piece of the first order self-force, and (iii) the torus-averaged, dissipative piece of the second order self-force. Thus, all three of these quantities will be required to compute the inspiral to subleading order, confirming arguments made in Refs. [37,68,88,89]. In particular, knowledge of the full first order self-force will not enable computation of more accurate inspirals until the averaged, dissipative piece of the second order self-force is known.[7]

E. Two-timescale expansion of the Einstein equations and adiabatic waveforms

We now turn to a brief description of the two-timescale expansion of the Einstein equations; more details will be given in the forthcoming paper [80]. We focus our attention on a region $\mathcal{R}$ of spacetime defined by the following conditions: (i) The distance from the particle is large compared to its mass μ; (ii) the distance r from the large black hole is small compared to the inspiral time, $r \ll M^2/\mu$; and (iii) the extent of the region in time covers the entire inspiral in the relativistic regime. In this domain we make an ansatz for the form of the metric that is justified *a posteriori* order by order.

At distances $\sim\mu$ from the particle, one needs to use a different type of analysis (e.g. black hole perturbation theory for a small black hole), and to mesh that analysis with the solution in the region $\mathcal{R}$ by matching in a domain of common validity. This procedure is very well understood and is the standard method for deriving the first order self-force [44,46,90]. It is valid for our metric ansatz (1.8) below since that ansatz reduces, locally in time at each instant, to standard black hole perturbation theory. Therefore we do not focus on this aspect of the problem here. Similarly, at large distances, one needs to match the solution within $\mathcal{R}$ onto an outgoing wave solution in order to read off the asymptotic waveforms.[8]

Within the region $\mathcal{R}$, our ansatz for the form of the metric in the nonresonant case is

$$g_{\alpha\beta}(\bar{t}, \bar{x}^j; \varepsilon) = g^{(0)}_{\alpha\beta}(\bar{x}^j) + \varepsilon g^{(1)}_{\alpha\beta}(q_r, q_\theta, q_\phi, \tilde{t}, \bar{x}^j) + \varepsilon^2 g^{(1)}_{\alpha\beta}(q_r, q_\theta, q_\phi, \tilde{t}, \bar{x}^j) + O(\varepsilon^3). \quad (1.8)$$

Here $g^{(0)}_{\alpha\beta}$ is the background, Kerr metric. The coordinates $(\bar{t}, \bar{x}^j)$ can be any set of coordinates in Kerr geometry, subject only to the restriction that $\partial/\partial\bar{t}$ is the timelike Killing vector. On the right-hand side, $\tilde{t}$ is the slow time variable $\tilde{t} = \varepsilon\bar{t}$, and the quantities q_r, q_θ, and q_ϕ are the values of the orbit's angle variables at the intersection of the inspiralling orbit with the hypersurface $\bar{t}$ = constant. These are functions of $\bar{t}$ and of ε, and can be obtained from the solutions (1.4) and (1.5a) of the inspiral problem by eliminating the proper time τ. The result is of the form

$$q_i(\bar{t}, \varepsilon) = \frac{1}{\varepsilon} f_i^{(0)}(\tilde{t}) + f_i^{(1)}(\tilde{t}) + O(\varepsilon), \quad (1.9)$$

for some functions $f_i^{(0)}$, $f_i^{(1)}$. On the right-hand side of Eq. (1.8), the $O(\varepsilon^3)$ refers to an asymptotic expansion associated with the limit $\varepsilon \to 0$ at fixed q_i, $\bar{x}^k$, and $\tilde{t}$. Finally, the functions $g^{(1)}_{\alpha\beta}$ and $g^{(2)}_{\alpha\beta}$ are assumed to be multiply periodic in q_r, q_θ, and q_ϕ with period 2π in each variable.

By inserting the ansatz (1.8) into Einstein's equations, one obtains a set of equations that determines the free functions, order by order. At leading order we obtain an equation of the form

$$\mathcal{D}\, g^{(0)}_{\alpha\beta} = 0, \quad (1.10)$$

where $\mathcal{D}$ is a linear differential operator on the six-dimensional manifold with coordinates $(q_r, q_\theta, q_\phi, \bar{x}^j)$. In solving this equation, $\tilde{t}$ is treated as a constant. The solution that matches appropriately onto the worldline source can be written as

$$g^{(1)}_{\alpha\beta} = \frac{\partial g^{(0)}_{\alpha\beta}}{\partial M}\delta M(\tilde{t}) + \frac{\partial g^{(0)}_{\alpha\beta}}{\partial a}\delta a(\tilde{t}) + \ldots + \mathcal{F}_{\alpha\beta}[q_r, q_\theta, q_\phi, \bar{x}^j, E(\tilde{t}), L_z(\tilde{t}), Q(\tilde{t})]. \quad (1.11)$$

Here the terms on the first line are the secular pieces of the solution. They arise since the variable $\tilde{t}$ is treated as a constant, and so one can obtain a solution by taking the perturbation to the metric generated by allowing the parameters of the black hole (mass, spin, velocity, center of

[7] This statement remains true when one takes into account resonances [79].

[8] This matching is not necessary at the leading, adiabatic order, for certain special choices of time coordinate in the background spacetime, as argued in Ref. [77]. It is needed to higher orders.

mass location) to vary as arbitrary functions of $\tilde{t}$. For example, the mass of the black hole can be written as $M(\tilde{t}) = M + \delta M(\tilde{t})$, where $M = M(0)$ is the initial mass. The functions $\delta M(\tilde{t})$, $\delta a(\tilde{t})$ etc. are freely specifiable at this order, and will be determined at the next (post-1-adiabatic) order.

The second line of Eq. (1.11) is the oscillatory piece of the solution. Here one obtains a solution by taking the function $\mathcal{F}_{\alpha\beta}$ to be the function

$$\mathcal{F}_{\alpha\beta}(q_r, q_\theta, q_\phi, \bar{x}^j, E, L_z, Q)$$

that one obtains from standard linear perturbation theory with a geodesic source. This function is known analytically in Boyer-Lindquist coordinates (t, r, θ, ϕ) in terms of a mode expansion.[9,10]

The gauge freedom in this formalism consists of those one-parameter families of diffeomorphisms which preserve the form (1.8) of the metric ansatz. To the leading order, these consist of (i) gauge transformations of the background coordinates that are independent of ε, which preserve the timelike Killing vector, and (ii) transformations of the form

$$x^\alpha \to x^\alpha + \varepsilon \xi^\alpha(q_r, q_\theta, q_\phi, \tilde{t}, x^j) + O(\varepsilon^2). \qquad (1.12)$$

Note that this is *not* the standard gauge freedom of linear perturbation theory, since ξ^α depends on four "time variables" instead of one. This modified gauge group allows the two-timescale method to evade the two problems discussed at the end of Sec. I B above, since the gradual evolution is described entirely by the $\tilde{t}$ dependence, and, at each fixed $\tilde{t}$, the leading order dependence on the variables q_r, q_θ, q_ϕ, r, θ, and ϕ is the same as in standard perturbation theory with the same gauge transformation properties.

[9]In coordinates $\bar{t} = t - r$, r, θ, ϕ, the explicit form of the asymptotic solution can be obtained by taking Eq. (3.1) of Ref. [40], eliminating the phases χ_{lmkn} using Eq. (8.29) of Ref. [68], and making the identifications $q_r = \Omega_r[t - r - t_0 + \hat{t}_r(-\lambda_{r0}) - \hat{t}_\theta(-\lambda_{\theta 0})] - \Upsilon_r \lambda_{r0}$, $q_\theta = \Omega_\theta[t - r - t_0 + \hat{t}_r(-\lambda_{r0}) - \hat{t}_\theta(-\lambda_{\theta 0})] - \Upsilon_\theta \lambda_{\theta 0}$, and $q_\phi = \Omega_\phi[t - r - t_0 + \hat{t}_r(-\lambda_{r0}) - \hat{t}_\theta(-\lambda_{\theta 0})] + \phi_0 - \hat{\phi}_r(-\lambda_{r0}) + \hat{\phi}_\theta(-\lambda_{\theta 0})$.

[10]The function $\mathcal{F}_{\alpha\beta}$ depends on q_ϕ and ϕ only through the combination $q_\phi - \phi$. This allows us to show that the two-timescale form (1.8) of the metric reduces to a standard Taylor series expansion, locally in time near almost every value $\tilde{t}_0$ of $\tilde{t}$. For equatorial orbits there is no dependence on q_θ, and the ε dependence of the metric has the standard form up to linear order, in coordinates (t', r', θ', ϕ') defined by $t' = (\tilde{t} - \tilde{t}_0)/\varepsilon + [f_r^{(0)}(\tilde{t}_0)/\varepsilon]/\omega_{r0}$, $\phi' = \phi + \omega_{\phi 0}[f_r^{(0)}(\tilde{t}_0)/\varepsilon]/\omega_{r0} - [f_\phi^{(0)}(\tilde{t}_0)/\varepsilon]$, $r' = r$, $\theta' = \theta$, where $\omega_{r0} = f_r^{(0)\prime}(\tilde{t}_0)$, $\omega_{\phi 0} = f_\phi^{(0)\prime}(\tilde{t}_0)$, and for any number x, $[x] \equiv x + 2\pi n$ where the integer n is chosen so that $0 \le [x] < 2\pi$. A similar construction works for circular orbits for which there is no dependence on q_r. For generic orbits a slightly more involved construction works, but only if $\omega_{r0}/\omega_{\phi 0}$ is irrational [80], which occurs for almost every value of $\tilde{t}_0$.

F. Organization of this paper

The organization of this paper is as follows. In Sec. II we derive the fundamental equations describing the inspiral of a point particle into a Kerr black hole in terms of generalized action-angle variables. In Sec. III we define a class of general, weakly perturbed dynamical systems of which the inspiral motion in Kerr spacetime is a special case. We then study the solutions of this class of systems using two-timescale expansions, first for a single degree of freedom in Sec. IV, and then for the general case in Sec. V. Section VI gives an example of a numerical integration of a system of this kind, and Sec. VII gives the final discussion and conclusions.

G. Notation and conventions

Throughout this paper we use units with $G = c = 1$. Lower case Roman indices $a, b, c, \ldots$ denote abstract indices in the sense of Wald [91]. We use these indices both for tensors on spacetime and for tensors on the eight-dimensional phase space. Lower case Greek indices $\nu, \lambda, \sigma, \tau, \ldots$ from the middle of the alphabet denote components of spacetime tensors on a particular coordinate system; they thus transform under spacetime coordinate transformations. They run over 0, 1, 2, 3. Lower case Greek indices $\alpha, \beta, \gamma \ldots$ from the start of the alphabet label position or momentum coordinates on eight-dimensional phase space that are not associated with coordinates on spacetime. They run over 0, 1, 2, 3 and do not transform under spacetime coordinate transformations. In Sec. V, and just in that section, indices $\alpha, \beta, \gamma, \delta, \varepsilon, \ldots$ from the start of the Greek alphabet run over $1 \ldots N$, and indices $\lambda, \mu, \nu, \rho, \sigma, \ldots$ from the second half of the alphabet run over $1 \ldots M$. Bold faced quantities generally denote vectors, as in $\mathbf{J} = (J_1, \ldots, J_M)$, although in Sec. II the bold faced notation is used for differential forms.

II. EXTREME MASS RATIO INSPIRALS IN KERR SPACETIME FORMULATED USING ACTION-ANGLE VARIABLES

In this section we derive the form of the fundamental equations describing the inspiral of a point particle into a Kerr black hole, using action-angle-type variables. Our final result is given in Eqs. (2.47) below, and the properties of the solutions of these equations are analyzed in detail in the remaining sections of this paper.

The description of geodesic motion in Kerr geometry in terms of action-angle variables was first given by Schmidt [92], and has been reviewed by Glampedakis and Babak [93]. We follow closely Schmidt's treatment, except that we work in an eight-dimensional phase space instead of a six-dimensional phase space, thus treating the time and spatial variables on an equal footing. We also clarify the extent to which the fundamental frequencies of geodesic motion are uniquely determined and gauge invariant, as claimed by Schmidt.

We start in Sec. II A by reviewing the geometric definition of action-angle variables in Hamiltonian mechanics, which is based on the Liouville-Arnold theorem [94]. This definition does not apply to geodesic motion in Kerr spacetime, since the level surfaces defined by the conserved quantities in the eight-dimensional phase space are noncompact. In Sec. II B we discuss how generalized action-angle variables can be defined for noncompact level surfaces, and in Sec. II C we apply this to give a coordinate-independent construction of generalized action-angle variables for generic bound geodesics in Kerr geometry. Section II D specializes to Boyer-Lindquist coordinates on phase space, and describes explicitly, following Schmidt [92], the explicit canonical transformation from those coordinates to the generalized action-angle variables.

We then turn to using these variables to describe a radiation-reaction driven inspiral. In Sec. II E we derive the equations of motion in terms of the generalized action-angle variables. These equations define a flow on the eight-dimensional phase space, and do not explicitly exhibit the conservation of rest mass. In Sec. II F we therefore switch to a modified set of variables and equations in which the conservation of rest mass is explicit. We also augment the equations to describe the backreaction of gravitational radiation passing through the horizon of the black hole.

A. Review of action-angle variables in geometric Hamiltonian mechanics

We start by recalling the standard geometric framework for Hamiltonian mechanics [94]. A Hamiltonian system consists of a $2N$-dimensional differentiable manifold $\mathcal{M}$ on which there is defined a smooth function H (the Hamiltonian), and a nondegenerate 2-form Ω_{ab} which is closed, $\nabla_{[a}\Omega_{bc]} = 0$. Defining the tensor Ω^{ab} by $\Omega^{ab}\Omega_{bc} = \delta^a_c$, the Hamiltonian vector field is defined as

$$v^a = \Omega^{ab}\nabla_b H, \tag{2.1}$$

and the integral curves of this vector field give the motion of the system. The 2-form Ω_{ab} is called the symplectic structure. Coordinates (q_α, p_α) with $1 \le \alpha \le N$ are called symplectic coordinates if the symplectic structure can be written as $\mathbf{\Omega} = dp_\alpha \wedge dq_\alpha$, i.e. $\Omega_{ab} = 2\nabla_{[a}p_\alpha\nabla_{b]}q_\alpha$.

We shall be interested in systems that possess $N-1$ first integrals of motion which, together with the Hamiltonian H, form a complete set of N-independent first integrals. We denote these first integrals by P_α, $1 \le \alpha \le N$, where $P_1 = H$. These quantities are functions on $\mathcal{M}$ for which the Poisson brackets

$$\{P_\alpha, H\} \equiv \Omega^{ab}(\nabla_a P_\alpha)(\nabla_b H) \tag{2.2}$$

vanish for $1 \le \alpha \le N$. If the first integrals satisfy the stronger condition that all the Poisson brackets vanish,

$$\{P_\alpha, P_\beta\} = 0 \tag{2.3}$$

for $1 \le \alpha, \beta \le N$, then the first integrals are said to be in involution. If the 1-forms $\nabla_a P_\alpha$ for $1 \le \alpha \le N$ are linearly independent, then the first integrals are said to be independent. A system is said to be *completely integrable* in some open region $\mathcal{U}$ in $\mathcal{M}$ if there exist N first integrals which are independent and in involution at every point of $\mathcal{U}$.

For completely integrable systems, the phase space $\mathcal{M}$ is foliated by invariant level sets of the first integrals. For a given set of values $\mathbf{p} = (p_1, \ldots, p_N)$, we define the level set

$$\mathcal{M}_{\mathbf{p}} = \{x \in \mathcal{M} | P_\alpha(x) = p_\alpha, 1 \le \alpha \le N\}, \tag{2.4}$$

which is an N-dimensional submanifold of $\mathcal{M}$. The level sets are invariant under the Hamiltonian flow by Eq. (2.2). Also the pullback of the symplectic structure $\mathbf{\Omega}$ to $\mathcal{M}_{\mathbf{p}}$ vanishes, since the vector fields $\vec{v}_\alpha$ defined by

$$v^a_\alpha = \Omega^{ab}\nabla_b P_\alpha \tag{2.5}$$

for $1 \le \alpha \le N$ form a basis of the tangent space to $\mathcal{M}_{\mathbf{p}}$ at each point, and satisfy $\Omega_{ab}v^a_\alpha v^b_\beta = 0$ for $1 \le \alpha, \beta \le N$ by Eq. (2.3).

A classic theorem of mechanics, the Liouville-Arnold theorem [94], applies to systems which are completely integrable in a neighborhood of some level set $\mathcal{M}_{\mathbf{p}}$ that is connected and compact. The theorem says that (i) the level set $\mathcal{M}_{\mathbf{p}}$ is diffeomorphic to an N-torus T^N. Moreover, there is a neighborhood $\mathcal{V}$ of $\mathcal{M}_{\mathbf{p}}$ which is diffeomorphic to the product $T^N \times \mathcal{B}$ where $\mathcal{B}$ is an open ball, such that the level sets are the N-tori. (ii) There exist symplectic coordinates (q_α, J_α) for $1 \le \alpha \le N$ (action-angle variables) on $\mathcal{V}$ for which the angle variables q_α are periodic,

$$q_\alpha + 2\pi \equiv q_\alpha,$$

and for which the first integrals depend only on the action variables, $P_\alpha = P_\alpha(J_1, \ldots, J_N)$ for $1 \le \alpha \le N$.

An explicit and coordinate-invariant prescription for computing a set of action variables J_α is as follows [94]. A symplectic potential $\mathbf{\Theta}$ is a 1-form which satisfies $d\mathbf{\Theta} = \Omega$. Since the 2-form $\mathbf{\Omega}$ is closed, such 1-forms always exist locally. For example, in any local symplectic coordinate system (q_α, p_α), the 1-form $\mathbf{\Theta} = p_\alpha dq_\alpha$ is a symplectic potential. It follows from the hypotheses of the Liouville-Arnold theorem that there exist symplectic potentials that are defined in a neighborhood of $\mathcal{M}_{\mathbf{p}}$ [95]. The first homotopy group $\Pi_1(\mathcal{M}_{\mathbf{p}})$ is defined to be the set of equivalence classes of loops on $\mathcal{M}_{\mathbf{p}}$, where two loops are equivalent if one can be continuously deformed into the other. Since $\mathcal{M}_{\mathbf{p}}$ is diffeomorphic to the N-torus, this group is isomorphic to $(\mathbf{Z}^k, +)$, the group of N-tuples of integers under addition. Pick a set of generators $\gamma_1, \ldots, \gamma_N$ of $\Pi_1(\mathcal{M}_{\mathbf{p}})$, and for each loop γ_α define

$$J_\alpha = \frac{1}{2\pi}\int_{\gamma_\alpha} \mathbf{\Theta}. \tag{2.6}$$

This integral is independent of the choice of symplectic potential $\mathbf{\Theta}$.[11] It is also independent of the choice of loop γ_α in the equivalence class of the generator of $\Pi_1(\mathcal{M}_\mathbf{p})$, since if γ_α and γ'_α are two equivalent loops, we have

$$\int_{\gamma_\alpha} \mathbf{\Theta} - \int_{\gamma'_\alpha} \mathbf{\Theta} = \int_{\partial\mathcal{R}} \mathbf{\Theta} = \int_{\mathcal{R}} d\mathbf{\Theta} = \int_{\mathcal{R}} \mathbf{\Omega} = 0. \tag{2.7}$$

Here $\mathcal{R}$ is a two-dimensional surface in $\mathcal{M}_\mathbf{p}$ whose boundary is $\gamma_\alpha - \gamma'_\alpha$; we have used Stokes theorem, and in the last equality we have used the fact that the pullback of $\mathbf{\Omega}$ to the level set $\mathcal{M}_\mathbf{p}$ vanishes.

Action-angle variables for a given system are not unique [96]. There is a freedom to redefine the coordinates via

$$q_\alpha \to A_{\alpha\beta} q_\beta, \qquad J_\alpha \to B_{\alpha\beta} J_\beta, \tag{2.8}$$

where $A_{\alpha\beta}$ is a constant matrix of integers with determinant ± 1, and $A_{\alpha\beta}B_{\alpha\gamma} = \delta_{\beta\gamma}$. This is just the freedom present in choosing a set of generators of the group $\Pi_1(\mathcal{M}_\mathbf{p}) \sim (\mathbf{Z}^N, +)$. Fixing this freedom requires the specification of some additional information, such as a choice of coordinates on the torus; once the coordinates q_α are chosen, one can take the loops γ_α to be the curves $q_\beta =$ constant for $\beta \neq \alpha$. There is also a freedom to redefine the origin of the angle variables separately on each torus:

$$q_\alpha \to q_\alpha + \frac{\partial Z(J_\beta)}{\partial J_\alpha}, \qquad J_\alpha \to J_\alpha. \tag{2.9}$$

Here $Z(J_\beta)$ can be an arbitrary function of the action variables.

B. Generalized action-angle variables for noncompact level sets

One of the crucial assumptions in the Liouville-Arnold theorem is that the level set $\mathcal{M}_\mathbf{p}$ is compact. Unfortunately, this assumption is not satisfied by the dynamical system of bound orbits in Kerr spacetime which we discuss in Sec. II C below, because we will work in the eight-dimensional phase space and the motion is not bounded in the time direction. We shall therefore use instead a generalization of the Liouville-Arnold theorem to noncompact level sets, due to Fiorani, Giachetta, and Sardanashvily [95].

Consider a Hamiltonian system which is completely integrable in a neighborhood $\mathcal{U}$ of a connected level set $\mathcal{M}_\mathbf{p}$, for which the N vector fields (2.5) are complete on $\mathcal{U}$, and for which the level sets $\mathcal{M}_{\mathbf{p}'}$ foliating $\mathcal{U}$ are all diffeomorphic to one another. For such systems Fiorani *et al.* [95] prove the following:

(i) There is an integer k with $0 \le k \le N$ such that the level set $\mathcal{M}_\mathbf{p}$ is diffeomorphic to the product $T^k \times \mathbf{R}^{N-k}$, where $\mathbf{R}$ is the set of real numbers. Moreover, there is a neighborhood $\mathcal{V}$ of $\mathcal{M}_\mathbf{p}$ which is diffeomorphic to the product $T^k \times \mathbf{R}^{N-k} \times \mathcal{B}$, where $\mathcal{B}$ is an open ball.

(ii) There exist symplectic coordinates (q_α, J_α) for $1 \le \alpha \le N$ (generalized action-angle variables) on $\mathcal{V}$ for which the first k variables q_α are periodic,

$$q_\alpha + 2\pi \equiv q_\alpha, \qquad 1 \le \alpha \le k,$$

and for which the first integrals depend only on the action variables, $P_\alpha = P_\alpha(J_1, \ldots, J_N)$ for $1 \le \alpha \le N$.

Thus, there are k compact dimensions in the level sets, and $N - k$ noncompact dimensions. In our application to Kerr spacetime below, the values of these parameters will be $k = 3$ and $N - k = 1$.

The freedom in choosing generalized action-angle variables is larger than the corresponding freedom for action-angle variables discussed above. The first k action variables can be computed in the same way as before, via the integral (2.6) evaluated on a set of generators $\gamma_1, \ldots, \gamma_k$ of $\Pi_1(\mathcal{M}_\mathbf{p})$, which in this case is isomorphic to $(\mathbf{Z}^k, +)$. This prescription is unique up to a group of redefinitions of the form [cf. Eq. (2.8) above]

$$q_\alpha \to \sum_{\beta=1}^{k} A_{\alpha\beta} q_\beta, \qquad J_\alpha \to \sum_{\beta=1}^{k} B_{\alpha\beta} J_\beta, \tag{2.10}$$

for $1 \le \alpha \le k$, where the $k \times k$ matrix $A_{\alpha\beta}$ is a constant matrix of integers with determinant ± 1, and $A_{\alpha\beta}B_{\alpha\gamma} = \delta_{\beta\gamma}$. There is additional freedom present in the choice of the rest of the action variables $J_{k+1}, \ldots, J_N$. As a consequence, the remaining freedom in choosing generalized action-angle variables consists of the transformations (2.9) discussed earlier, together with transformations of the form

$$q_\alpha \to A_{\alpha\beta} q_\beta, \qquad J_\alpha \to B_{\alpha\beta} J_\beta, \tag{2.11}$$

where $A_{\alpha\beta}$ and $B_{\alpha\beta}$ are constant real $N \times N$ matrices with $A_{\alpha\beta}B_{\alpha\gamma} = \delta_{\beta\gamma}$ such that $J_1, \ldots, J_k$ are preserved.

In generalized action-angle variables, the equations of motion take the simple form

$$\dot{q}_\alpha = \frac{\partial H(\mathbf{J})}{\partial J_\alpha} \tag{2.12}$$

and

$$\dot{J}_\alpha = -\frac{\partial H(\mathbf{J})}{\partial q_\alpha} = 0. \tag{2.13}$$

We define the quantities

[11] An argument of the type used in Ref. [95] can be used to show that the pullback to $\mathcal{M}_\mathbf{p}$ of the difference between two symplectic potentials is exact since it is closed.

$$\Omega_\alpha(\mathbf{J}) \equiv \frac{\partial H(\mathbf{J})}{\partial J_\alpha}, \tag{2.14}$$

which are angular frequencies for $1 \le \alpha \le k$ but not for $k+1 \le \alpha \le N$. The solutions of the equations of motion are then

$$q_\alpha(t) = \Omega_\alpha(\mathbf{J}_0)t + q_{\alpha 0}, \tag{2.15a}$$
$$J_\alpha(t) = J_{\alpha 0}, \tag{2.15b}$$

for some constants $\mathbf{J}_0$ and $\mathbf{q}_0$.

C. Application to bound geodesic motion in Kerr spacetime

We now apply the general theory discussed above to give a coordinate-invariant definition of action-angle variables for a particle on a bound orbit in Kerr spacetime. We denote by $(\mathcal{M}_{\rm K}, g_{ab})$ the Kerr spacetime, and we denote by ξ^a and η^a the timelike and axial Killing vector fields. The cotangent bundle over $\mathcal{M}_{\rm K}$ forms an eight-dimensional phase space $\mathcal{M} = T^*\mathcal{M}_{\rm K}$. Given any coordinate system x^ν on Kerr spacetime, we can define a coordinate system (x^ν, p_ν) on $\mathcal{M}$, such that the point (x^ν, p_ν) corresponds to the covector or 1-form $p_\nu dx^\nu$ at x^ν in $\mathcal{M}_{\rm K}$. The natural symplectic structure on $\mathcal{M}$ is then defined by demanding that all such coordinate systems (x^ν, p_ν) be symplectic [94]. The Killing vector fields ξ^a and η^a on $\mathcal{M}_{\rm K}$ have natural extensions to vector fields on phase space which Lie derive the symplectic structure.

Consider now a particle of mass μ on a bound geodesic orbit. A Hamiltonian on $\mathcal{M}$ that generates geodesic motion is given by

$$H(x^\nu, p_\nu) = \tfrac{1}{2} g^{\nu\sigma}(x^\rho) p_\nu p_\sigma; \tag{2.16}$$

this definition is independent of the choice of coordinate system x^ν. If we interpret p_ν to be the 4-momentum of the particle, then the conserved value of H is $-\mu^2/2$, and the evolution parameter is the affine parameter $\lambda = \tau/\mu$ where τ is proper time.

As is well known, geodesics on Kerr geometry possess three first integrals: the energy $E = -\xi^a p_a$, the z component of angular momentum $L_z = \eta^a p_a$, and the Carter constant $Q = Q^{ab} p_a p_b$, where Q^{ab} is a Killing tensor [97]. Together with the Hamiltonian we therefore have four first integrals:

$$P_\alpha = (P_0, P_1, P_2, P_3) = (H, E, L_z, Q). \tag{2.17}$$

An explicit computation of the 4-form $dH \wedge dE \wedge dL_z \wedge dQ$ on $\mathcal{M}$ shows that it is nonvanishing for bound orbits except for the degenerate cases of circular (i.e. constant Boyer-Lindquist radial coordinate) and equatorial orbits. Also, the various Poisson brackets $\{P_\alpha, P_\beta\}$ vanish: $\{E, H\}$ and $\{L_z, H\}$ vanish since ξ^a and η^a are Killing fields, $\{E, L_z\}$ vanishes since these Killing fields commute, $\{Q, H\}$ vanishes since Q^{ab} is a Killing tensor, and finally $\{E, Q\}$ and $\{L_z, Q\}$ vanish since the Killing tensor is invariant under the flows generated by ξ^a and η^a. Therefore, for generic orbits the theorem due to Fiorani *et al.* discussed in the last subsection applies.[12] The relevant parameter values are $k = 3$ and $N = 4$, since the level sets $\mathcal{M}_{\mathbf{p}}$ are noncompact in the time direction only. Thus geodesic motion can be parametrized in terms of generalized action-angle variables.

We next discuss how to resolve in this context the nonuniqueness in the choice of generalized action-angle variables discussed in the last subsection. Consider first the freedom (2.10) associated with the choice of generators of $\Pi_1(\mathcal{M}_{\mathbf{p}})$. One of these generators can be chosen to be an integral curve of the extension to $\mathcal{M}$ of the axial Killing field η^a. The other two can be chosen as follows. Let π: $\mathcal{M} \to \mathcal{M}_{\rm K}$ be the natural projection from phase space $\mathcal{M}$ to spacetime $\mathcal{M}_{\rm K}$ that takes (x^ν, p_ν) to x^ν. A loop $(x^\nu(\lambda), p_\nu(\lambda))$ in the level set $\mathcal{M}_{\rm p}$ then projects to the curve $x^\nu(\lambda)$ in $\pi(\mathcal{M}_{\rm p})$. Requiring that this curve intersect the boundary of $\pi(\mathcal{M}_{\rm p})$ only twice determines the two other generators of $\Pi_1(\mathcal{M}_{\mathbf{p}})$.[13] The resulting three generators coincide with the generators obtained from the motions in the r, θ, and ϕ directions in Boyer-Lindquist coordinates [92]. We denote the resulting generalized action-angle variables by $(q_t, q_r, q_\theta, q_\phi, J_t, J_r, J_\theta, J_\phi)$.

The remaining ambiguity (2.11) is of the form

$$J_i \to J_i, \qquad J_t \to \gamma J_t + v^i J_i, \tag{2.18}$$

where i runs over r, θ, and ϕ, and the parameters γ and v^i are arbitrary. The corresponding transformation of the frequencies (2.14) is

$$\Omega_t \to \gamma^{-1}\Omega_t, \qquad \Omega_i \to \Omega_i - \gamma^{-1} v^i \Omega_t. \tag{2.19}$$

A portion of this ambiguity (the portion given by $\gamma = 1$, $v^r = v^\theta = 0$) is that associated with the choice of rotational frame, $\phi \to \phi + \Omega t$ where Ω is an angular velocity. It is not possible to eliminate this rotational-frame ambiguity using only the spacetime geometry in a neighborhood of the orbit. In this sense, the action-angle variables are not uniquely determined by local geometric information. However, we can resolve the ambiguity using global geometric information, by choosing

$$J_t = \frac{1}{2\pi} \int_{\gamma_t} \Theta, \tag{2.20}$$

where γ_t is an integral curve of length 2π of the extension to $\mathcal{M}$ of the timelike Killing field ξ^a.[14] The definition

[12] One can check that the two other assumptions in the theorem listed in the second paragraph of Sec. II B are satisfied.

[13] This excludes, for example, loops which wind around twice in the r direction and once in the θ direction.

[14] The Killing field ξ^a encodes global geometric information since it is defined to be timelike and of unit norm at spatial infinity.

(2.20) is independent of the choice of such a curve γ_t and of the choice of symplectic potential $\mathbf{\Theta}$.

To summarize, we have given a coordinate-invariant definition of the generalized action-angle variables $(q_t, q_r, q_\theta, q_\phi, J_t, J_r, J_\theta, J_\phi)$ for generic bound orbits in Kerr spacetime. These variables are uniquely determined up to relabeling and up to the residual ambiguity (2.9). A similar construction has been given by Schmidt [92], except that Schmidt first projects out the time direction of the level sets, and then defines three action variables (J_r, J_θ, J_ϕ) and three angle variables (q_r, q_θ, q_ϕ).

D. Explicit expressions in terms of Boyer-Lindquist coordinates

In Boyer-Lindquist coordinates (t, r, θ, ϕ), the Kerr metric is

$$ds^2 = -\left(1 - \frac{2Mr}{\Sigma}\right)dt^2 + \frac{\Sigma}{\Delta}dr^2 + \Sigma d\theta^2 + \left(r^2 + a^2 + \frac{2Ma^2r}{\Sigma}\sin^2\theta\right)\sin^2\theta d\phi^2 - \frac{4Mar}{\Sigma}\sin^2\theta dt d\phi, \tag{2.21}$$

where

$$\Sigma = r^2 + a^2\cos^2\theta, \qquad \Delta = r^2 - 2Mr + a^2, \tag{2.22}$$

and M and a are the black hole mass and spin parameters. The timelike and axial Killing fields are $\vec{\xi} = \partial/\partial t$ and $\vec{\eta} = \partial/\partial\phi$, and so the energy and z component of angular momentum are

$$E = -\vec{\xi}\cdot\vec{p} = -p_t \tag{2.23a}$$

and

$$L_z = \vec{\eta}\cdot\vec{p} = p_\phi. \tag{2.23b}$$

The Carter constant is given by [97]

$$Q = p_\theta^2 + a^2\cos^2\theta(\mu^2 - p_t^2) + \cot^2\theta p_\phi^2, \tag{2.23c}$$

and the Hamiltonian (2.16) is

$$H = \frac{\Delta}{2\Sigma}p_r^2 + \frac{1}{2\Sigma}p_\theta^2 + \frac{(p_\phi + a\sin^2\theta p_t)^2}{2\Sigma\sin^2\theta} - \frac{[(r^2+a^2)p_t + ap_\phi]^2}{2\Sigma\Delta}. \tag{2.23d}$$

Following Schmidt [92], we can obtain an invertible transformation from the Boyer-Lindquist phase space coordinates (x^ν, p_ν) to the generalized action-angle variables (q_α, J_α) as follows. Equations (2.23) can be inverted to express the momenta p_ν in terms of x^ν and the four first integrals

$$P_\alpha = (H, E, L_z, Q) = (-\tfrac{1}{2}\mu^2, E, L_z, Q) \tag{2.24}$$

up to some signs [97]:

$$p_t = -E, \qquad p_\phi = L_z, \qquad p_r = \pm\frac{\sqrt{V_r(r)}}{\Delta}, \qquad p_\theta = \pm\sqrt{V_\theta(\theta)}. \tag{2.25}$$

Here the potentials $V_r(r)$ and $V_\theta(\theta)$ are defined by

$$V_r(r) = [(r^2+a^2)E - aL_z]^2 - \Delta[\mu^2 r^2 + (L_z - aE)^2 + Q], \tag{2.26a}$$

$$V_\theta(\theta) = Q - \left[(\mu^2 - E^2)a^2 + \frac{L_z^2}{\sin^2\theta}\right]\cos^2\theta. \tag{2.26b}$$

Using these formulas together with the symplectic potential $\mathbf{\Theta} = p_\nu dx^\nu$ in the definitions (2.6) and (2.20) gives

$$J_r = \frac{1}{2\pi}\oint\frac{\sqrt{V_r}}{\Delta}dr, \tag{2.27a}$$

$$J_\theta = \frac{1}{2\pi}\oint\sqrt{V_\theta}d\theta, \tag{2.27b}$$

$$J_\phi = \frac{1}{2\pi}\oint p_\phi d\phi = L_z, \tag{2.27c}$$

$$J_t = \frac{1}{2\pi}\int_0^{2\pi} p_t dt = -E. \tag{2.27d}$$

These expressions give the action variables as functions of the first integrals, $J_\alpha = J_\alpha(P_\beta)$. The theorem discussed in Sec. II B above guarantees that these relations can be inverted to give

$$P_\alpha = P_\alpha(J_\beta). \tag{2.28}$$

Next, to obtain expressions for the corresponding generalized angle variables, we use the canonical transformation from the symplectic coordinates (x^ν, p_ν) to (q_α, J_α) associated with a general solution of the Hamilton-Jacobi equation

$$H\left[x^\nu, \frac{\partial\mathcal{S}}{\partial x^\nu}\right] + \frac{\partial\mathcal{S}}{\partial\lambda} = 0. \tag{2.29}$$

As shown by Carter [97], this equation is separable and the general solution[15] can be written in terms of the first integrals P_α,

$$\mathcal{S}(x^\nu, P_\alpha, \lambda) = -H\lambda + \mathcal{W}(x^\nu, P_\alpha) \tag{2.30}$$

where $H = -\mu^2/2$,

$$\mathcal{W}(x^\nu, P_\alpha) = -Et + L_z\phi \pm \mathcal{W}_r(r) \pm \mathcal{W}_\theta(\theta), \tag{2.31}$$

$$\mathcal{W}_r(r) = \int^r dr\frac{\sqrt{V_r}}{\Delta}, \tag{2.32}$$

[15] As indicated by the $\pm$ signs in Eq. (2.31), there are actually four different solutions, one on each of the four coordinate patches on which (x^ν, P_α) are good coordinates, namely, $\mathrm{sgn}(p_r) = \pm 1$, $\mathrm{sgn}(p_\theta) = \pm 1$.

and

$$\mathcal{W}_\theta(\theta) = \int^\theta d\theta \sqrt{V_\theta}. \tag{2.33}$$

Using the relation (2.28) the function $\mathcal{W}$ can be expressed in terms of the Boyer-Lindquist coordinates x^ν and the action variables J_α:

$$\mathcal{W} = \mathcal{W}(x^\nu, J_\alpha). \tag{2.34}$$

This is a type II generating function that generates the required canonical transformation from (x^ν, p_ν) to (q_α, J_α):

$$p_\nu = \frac{\partial \mathcal{W}}{\partial x^\nu}(x^\nu, J_\beta), \tag{2.35a}$$

$$q_\alpha = \frac{\partial \mathcal{W}}{\partial J_\alpha}(x^\nu, J_\beta). \tag{2.35b}$$

Equation (2.35a) is already satisfied by virtue of the definition (2.31) of $\mathcal{W}$ together with Eqs. (2.25). Equation (2.35b) furnishes the required formulas for the generalized angle variables q_α.[16]

Although it is possible, in principle, to express the first integrals P_α in terms of the action variables J_α using Eqs. (2.27), it is not possible to obtain explicit analytic expressions for $P_\alpha(J_\beta)$. However, as pointed out by Schmidt [92], it is possible to obtain explicit expressions for the partial derivatives $\partial P_\alpha/\partial J_\beta$, and this is sufficient to compute the frequencies Ω_α. We review this in Appendix A.

E. Application to slow inspiral motion in Kerr

The geodesic equations of motion in terms of the generalized action-angle variables (q_α, J_α) are [cf. Eqs. (2.12), (2.13), and (2.14) above]

$$\frac{dq_\alpha}{d\lambda} = \Omega_\alpha(J_\beta), \tag{2.36a}$$

$$\frac{dJ_\alpha}{d\lambda} = 0, \tag{2.36b}$$

for $0 \le \alpha \le 3$. Here $\lambda = \tau/\mu$ where τ is proper time and μ is the mass of the particle. In this section we derive the modifications to these equations required to describe the radiation-reaction driven inspiral of a particle in Kerr spacetime. Our result is of the form

$$\frac{dq_\alpha}{d\lambda} = \Omega_\alpha(J_\beta) + \mu^2 f_\alpha(q_\beta, J_\beta), \tag{2.37a}$$

$$\frac{dJ_\alpha}{d\lambda} = \mu^2 F_\alpha(q_\beta, J_\beta). \tag{2.37b}$$

[16]The freedom (2.9) to redefine the origin of the angle variables on each torus is just the freedom to add to $\mathcal{W}$ any function of P_α. We choose to resolve this freedom by demanding that $q_r = 0$ at the minimum value of r, and $q_\theta = 0$ at the minimum value of θ.

We will derive explicit expressions for the forcing terms f_α and F_α in these equations.

The equation of motion for a particle subject to a self-acceleration a^ν is

$$\frac{d^2x^\nu}{d\lambda^2} + \Gamma^\nu_{\sigma\rho}\frac{dx^\sigma}{d\lambda}\frac{dx^\rho}{d\lambda} = \mu^2 a^\nu. \tag{2.38}$$

Rewriting this second order equation as two first order equations allows us to use the Jacobian of the coordinate transformation $\{x^\nu, p_\nu\} \to \{q_\alpha, J_\alpha\}$ to relate the forcing terms for the two sets of variables:

$$\frac{dx^\nu}{d\lambda} = g^{\nu\sigma}p_\sigma, \tag{2.39a}$$

$$\frac{dp_\nu}{d\lambda} = -\frac{1}{2}g^{\sigma\rho}{}_{,\nu}p_\sigma p_\rho + \mu^2 a_\nu. \tag{2.39b}$$

We start by deriving the equation of motion for the action variables J_α. Taking a derivative with respect to λ of the relation $J_\alpha = J_\alpha(x^\nu, p_\nu)$ and using Eqs. (2.39) gives

$$\begin{aligned}\frac{dJ_\alpha}{d\lambda} &= \frac{\partial J_\alpha}{\partial x^\nu}p^\nu + \frac{\partial J_\alpha}{\partial p_\nu}\frac{dp_\nu}{d\lambda} \\ &= \left[\frac{\partial J_\alpha}{\partial x^\nu}g^{\nu\sigma}p_\sigma - \frac{1}{2}\frac{\partial J_\alpha}{\partial p_\nu}g^{\sigma\rho}{}_{,\nu}p_\sigma p_\rho\right] + \mu^2\frac{\partial J_\alpha}{\partial p_\nu}a_\nu.\end{aligned} \tag{2.40}$$

The term in square brackets must vanish identically since J_α is conserved in the absence of any acceleration a_ν. Rewriting the second term using $J_\alpha = J_\alpha(P_\beta)$ and the chain rule gives an equation of motion of the form (2.37b), where the forcing terms F_α are

$$F_\alpha = \frac{\partial J_\alpha}{\partial P_\beta}\left(\frac{\partial P_\beta}{\partial p_\nu}\right)_x a_\nu. \tag{2.41}$$

Here the subscript x on the round brackets means that the derivative is to be taken holding x^ν fixed. When the sum over β is evaluated, the contribution from $P_1 = H$ vanishes since $a_\nu p^\nu = 0$, and we obtain, using Eqs. (2.17) and (2.27),

$$F_t = a_t, \tag{2.42a}$$

$$F_r = -\frac{\partial J_r}{\partial E}a_t + \frac{\partial J_r}{\partial Q}a_Q + \frac{\partial J_r}{\partial L_z}a_\phi, \tag{2.42b}$$

$$F_\theta = -\frac{\partial J_\theta}{\partial E}a_t + \frac{\partial J_\theta}{\partial Q}a_Q + \frac{\partial J_\theta}{\partial L_z}a_\phi, \tag{2.42c}$$

$$F_\phi = a_\phi. \tag{2.42d}$$

Here we have defined $a_Q = 2Q^{\nu\sigma}p_\nu a_\sigma$, and the various coefficients $\partial J_\alpha/\partial P_\beta$ are given explicitly as functions of P_α in Appendix A.

We use a similar procedure to obtain the equation of motion (2.37a) for the generalized angle variables q_α. Differentiating the relation $q_\alpha = q_\alpha(x^\nu, p_\nu)$ with respect to λ and combining with the two first order equations of

motion (2.39) gives

$$\frac{dq_\alpha}{d\lambda} = \left[\frac{\partial q_\alpha}{\partial x^\nu} g^{\nu\sigma} p_\sigma - \frac{1}{2}\frac{\partial q_\alpha}{\partial p_\nu} g^{\sigma\rho}{}_{,\nu} p_\sigma p_\rho\right] + \mu^2 \frac{\partial q_\alpha}{\partial p_\nu} a_\nu. \tag{2.43}$$

By comparing with Eq. (2.36a) in the case of vanishing acceleration we see that the term in square brackets is $\Omega_\alpha(J_\beta)$. This gives an equation of motion of the form (2.37a), where the forcing term f_α is

$$f_\alpha = \left(\frac{\partial q_\alpha}{\partial p_\nu}\right)_x a_\nu. \tag{2.44}$$

Using the expression (2.35b) for the angle variable q_α together with $J_\alpha = J_\alpha(P_\beta)$ gives

$$\left(\frac{\partial q_\alpha}{\partial p_\nu}\right)_x = \left(\frac{\partial P_\gamma}{\partial p_\nu}\right)_x \left[\frac{\partial P_\beta}{\partial J_\alpha}\left(\frac{\partial^2 \mathcal{W}}{\partial P_\beta \partial P_\gamma}\right)_x + \left(\frac{\partial \mathcal{W}}{\partial P_\beta}\right)_x \frac{\partial}{\partial P_\gamma}\left(\frac{\partial P_\beta}{\partial J_\alpha}\right)\right]. \tag{2.45}$$

This yields for the forcing term

$$f_\alpha = a_\nu \left(\frac{\partial P_\gamma}{\partial p_\nu}\right)_x \frac{\partial P_\delta}{\partial J_\alpha}\left[\left(\frac{\partial^2 \mathcal{W}}{\partial P_\delta \partial P_\gamma}\right)_x - \left(\frac{\partial \mathcal{W}}{\partial P_\beta}\right)_x \frac{\partial P_\beta}{\partial J_\varepsilon}\frac{\partial^2 J_\varepsilon}{\partial P_\gamma \partial P_\delta}\right]. \tag{2.46}$$

In this expression the first two factors are the same as the factors which appeared in the forcing term (2.41) for the action variables. The quantities $\partial P_\delta/\partial J_\alpha$, $\partial P_\beta/\partial J_\varepsilon$, and $\partial^2 J_\varepsilon/(\partial P_\gamma \partial P_\delta)$ can be evaluated explicitly as functions of P_α using the techniques discussed in Appendix A. The remaining factors in Eq. (2.46) can be evaluated by differentiating the formula (2.31) for Hamilton's principal function $\mathcal{W}$ and using the formulas (2.26) for the potentials V_r and V_θ.

F. Rescaled variables and incorporation of back reaction on the black hole

We now augment the action-angle equations of motion (2.37) in order to describe the backreaction of the gravitational radiation on the black hole. We also modify the equations to simplify and make explicit the dependence on the mass μ of the particle. The resulting modified equations of motion, whose solutions we will analyze in the remainder of the paper, are

$$\frac{dq_\alpha}{d\tau} = \omega_\alpha(\tilde{P}_j, M_B) + \varepsilon g_\alpha^{(1)}(q_A, \tilde{P}_j, M_B) + \varepsilon^2 g_\alpha^{(2)}(q_A, \tilde{P}_j, M_B) + O(\varepsilon^3), \tag{2.47a}$$

$$\frac{d\tilde{P}_i}{d\tau} = \varepsilon G_i^{(1)}(q_A, \tilde{P}_j, M_A) + \varepsilon^2 G_i^{(2)}(q_A, \tilde{P}_j, M_B) + O(\varepsilon^3), \tag{2.47b}$$

$$\frac{dM_A}{d\tau} = \varepsilon^2 \hat{G}_A(q_A, \tilde{P}_j, M_B) + O(\varepsilon^3). \tag{2.47c}$$

Here α runs over 0, 1, 2, 3; i, j run over 1, 2, 3; A, B run over 1, 2; and $q_A = (q_r, q_\theta)$, $M_A = (M_1, M_2)$, and $\tilde{P}_i = (\tilde{P}_1, \tilde{P}_2, \tilde{P}_3)$. Also, all of the functions ω_α, $g_\alpha^{(1)}$, $g_\alpha^{(2)}$, $G_i^{(1)}$, $G_i^{(2)}$, and $\hat{G}_A$ that appear on the right-hand sides are smooth functions of their arguments whose precise form will not be needed for this paper (and are currently unknown aside from ω_α).

Our final equations (2.47) are similar in structure to the original equations (2.37), but there are a number of differences:

(i) We have switched the independent variable in the differential equations from affine parameter λ to proper time $\tau = \mu\lambda$.

(ii) We have introduced the ratio

$$\varepsilon = \frac{\mu}{M} \tag{2.48}$$

of the particle mass μ and black hole mass M, and have expanded the forcing terms as a power series in ε.

(iii) The forcing terms $g_\alpha^{(1)}$, $g_\alpha^{(2)}$, $G_i^{(1)}$, $G_i^{(2)}$, and $\hat{G}_A$ depend only on the two angle variables $q_A \equiv (q_r, q_\theta)$, and are independent of q_t and q_ϕ.

(iv) Rather than evolving the action variables J_α, we evolve two different sets of variables, $\tilde{P}_i$ and M_A. The first of these sets consists of three of the first integrals of the motion, with the dependence on the mass μ of the particle scaled out:

$$\tilde{P}_i = (\tilde{P}_1, \tilde{P}_2, \tilde{P}_3) \equiv (E/\mu, L_z/\mu, Q/\mu^2). \tag{2.49}$$

The second set consists of the mass and spin parameters of the black hole, which gradually evolve due to absorption of gravitational radiation by the black hole:

$$M_A = (M_1, M_2) = (M, a). \tag{2.50}$$

We now turn to a derivation of the modified equations of motion (2.47). The derivation consists of several steps. First, since the mapping (2.27) between the first integrals P_α and the action variables J_α is a bijection, we can use the P_α as dependent variables instead of J_α.[17] Equation (2.37a)

[17]Note that since the variables J_α are adiabatic invariants, so are the variables P_α.

is unmodified except that the right-hand side is expressed as a function of P_α instead of J_α. Equation (2.37b) is replaced by

$$\frac{dP_\alpha}{d\lambda} = \mu^2 \left(\frac{\partial P_\alpha}{\partial p_\nu}\right)_x a_\nu. \tag{2.51a}$$

Second, we switch to using modified versions $\tilde{P}_\alpha$ of the first integrals P_α with the dependence on the mass μ scaled out. These rescaled first integrals are defined by

$$\tilde{P}_\alpha = (\tilde{H}, \tilde{E}, \tilde{L}_z, \tilde{Q}) \equiv (H/\mu^2, E/\mu, L_z/\mu, Q/\mu^2). \tag{2.52}$$

We also change the independent variable from affine parameter λ to proper time $\tau = \mu\lambda$. This gives from Eqs. (2.37) and (2.44) the system of equations

$$\frac{dq_\alpha}{d\tau} = \frac{1}{\mu}\Omega_\alpha(P_\beta) + \mu\left(\frac{\partial q_\alpha}{\partial p_\nu}\right)_x a_\nu, \tag{2.53a}$$

$$\frac{d\tilde{P}_\alpha}{d\tau} = \mu^{1-n_\alpha}\left(\frac{\partial P_\alpha}{\partial p_\nu}\right)_x a_\nu, \tag{2.53b}$$

where we have defined $n_\alpha = (2, 1, 1, 2)$.

Third, we analyze the dependence on the mass μ of the right-hand sides of these equations. Under the transformation $(x^\nu, p_\nu) \to (x^\nu, sp_\nu)$ for $s > 0$, we obtain the following transformation laws for the first integrals (2.24), the action variables (2.27), and Hamilton's principal function (2.31):

$$P_\alpha \to s^{n_\alpha} P_\alpha \quad \text{with} \quad n_\alpha = (2, 1, 1, 2), \tag{2.54a}$$

$$J_\alpha \to sJ_\alpha, \tag{2.54b}$$

$$\mathcal{W} \to s\mathcal{W}. \tag{2.54c}$$

From the definitions (2.14) and (2.35b) of the angular frequencies Ω_α and the angle variables q_α, we also deduce

$$\Omega_\alpha \to s\Omega_\alpha, \tag{2.55a}$$

$$q_\alpha \to q_\alpha. \tag{2.55b}$$

If we write the angular velocity Ω_α as a function $\omega_\alpha(P_\beta)$ of the first integrals P_β, then it follows from the scalings (2.54a) and (2.55a) that the first term on the right-hand side of Eq. (2.53a) is

$$\frac{\Omega_\alpha}{\mu} = \frac{\omega_\alpha(P_\beta)}{\mu} = \frac{\omega_\alpha(\mu^{n_\beta}\tilde{P}_\beta)}{\mu} = \omega_\alpha(\tilde{P}_\beta). \tag{2.56}$$

This quantity is thus independent of μ at fixed $\tilde{P}_\beta$, as we would expect.

Similarly, if we write the angle variable q_α as a function $\bar{q}_\alpha(x^\nu, p_\nu)$ of x^ν and p_ν, then the scaling law (2.55b) implies that $\bar{q}_\alpha(x^\nu, sp_\nu) = \bar{q}_\alpha(x^\nu, p_\nu)$, and it follows that the coefficient of the 4-acceleration in Eq. (2.53a) is[18]

$$\mu\frac{\partial \bar{q}_\alpha}{\partial p_\nu}(x^\sigma, p_\sigma) = \mu\frac{\partial \bar{q}_\alpha}{\partial p_\nu}(x^\sigma, \mu u_\sigma) = \frac{\partial \bar{q}_\alpha}{\partial p_\nu}(x^\sigma, u_\sigma), \tag{2.57}$$

where u_σ is the 4-velocity. This quantity is also independent of μ at fixed $\tilde{P}_\beta$. We will denote this quantity by $f^\nu_\alpha(q_\beta, \tilde{P}_\beta)$. It can be obtained explicitly by evaluating the coefficient of a_ν in Eq. (2.46) at $P_\alpha = \tilde{P}_\alpha$, $p_\nu = u_\nu$. A similar analysis shows that the driving term on the right-hand side of Eq. (2.53b) can be written in the form

$$F^\nu_\alpha(q_\beta, \tilde{P}_\beta)a_\nu \equiv (0, -a_t, a_\phi, 2Q^{\nu\sigma}u_\nu a_\sigma). \tag{2.58}$$

The resulting rescaled equations of motion are

$$\frac{dq_\alpha}{d\tau} = \omega_\alpha(\tilde{P}_\beta) + f^\nu_\alpha(q_\beta, \tilde{P}_\beta)a_\nu, \tag{2.59a}$$

$$\frac{d\tilde{P}_\alpha}{d\tau} = F^\nu_\alpha(q_\beta, \tilde{P}_\beta)a_\nu. \tag{2.59b}$$

Note that this formulation of the equations is completely independent of the mass μ of the particle (except for the dependence on μ of the radiation-reaction acceleration a_ν which we will discuss below).

Fourth, since $P_0 = H = -\mu^2/2$, the rescaled variable is $\tilde{P}_0 = -1/2$ from Eq. (2.52). Thus we can drop the evolution equation for $\tilde{P}_0$, and retain only the equations for the remaining rescaled first integrals

$$\tilde{P}_i = (\tilde{P}_1, \tilde{P}_2, \tilde{P}_3) = (\tilde{E}, \tilde{L}_z, \tilde{Q}). \tag{2.60}$$

We can also omit the dependence on $\tilde{P}_0$ in the right-hand sides of the evolution equations (2.59), since $\tilde{P}_0$ is a constant. This yields

$$\frac{dq_\alpha}{d\tau} = \omega_\alpha(\tilde{P}_j) + f^\nu_\alpha(q_\beta, \tilde{P}_j)a_\nu, \tag{2.61a}$$

$$\frac{d\tilde{P}_i}{d\tau} = F^\nu_i(q_\beta, \tilde{P}_j)a_\nu. \tag{2.61b}$$

Fifth, the self-acceleration of the particle can be expanded in powers of the mass ratio $\varepsilon = \mu/M$ as

$$a_\nu = \varepsilon a^{(1)}_\nu + \varepsilon^2 a^{(2)}_\nu + O(\varepsilon^3). \tag{2.62}$$

Here $a^{(1)}_\nu$ is the leading order self-acceleration derived by Mino, Sasaki, and Tanaka [44] and by Quinn and Wald [45], discussed in the Introduction. The subleading self-acceleration $a^{(2)}_\nu$ has been computed in Refs. [83–87]. The accelerations $a^{(1)}_\nu$ and $a^{(2)}_\nu$ are independent of μ and thus depend only on x^ν and u_ν or, equivalently, on q_α and $\tilde{P}_i$. This yields the system of equations

[18]Note that $\mu\partial/\partial p_\nu$ cannot be simplified to $\partial/\partial u_\nu$ because we are working in the eight-dimensional phase space $\mathcal{M}$ where μ is a coordinate and not a constant.

$$\frac{dq_\alpha}{d\tau} = \omega_\alpha(\tilde{P}_j) + \varepsilon g^{(1)}_\alpha(q_\beta, \tilde{P}_j) + \varepsilon^2 g^{(2)}_\alpha(q_\beta, \tilde{P}_j) + O(\varepsilon^3), \tag{2.63a}$$

$$\frac{d\tilde{P}_i}{d\tau} = \varepsilon G^{(1)}_i(q_\beta, \tilde{P}_j) + \varepsilon^2 G^{(2)}_i(q_\beta, \tilde{P}_j) + O(\varepsilon^3). \tag{2.63b}$$

Here the forcing terms are given by

$$g^{(s)}_\alpha = f^\nu_\alpha a^{(s)}_\nu, \tag{2.64a}$$

$$G^{(s)}_i = F^\nu_i a^{(s)}_\nu, \tag{2.64b}$$

for $s = 1, 2$.

The formula (2.62) for the self-acceleration, with the explicit formula for $a^{(1)}_\nu$ from Refs. [44,45], is valid when one chooses the Lorentz gauge for the metric perturbation. The form of Eq. (2.62) is also valid in a variety of other gauges; see Ref. [98] for a discussion of the gauge transformation properties of the self-force. However, there exist gauge choices which are incompatible with Eq. (2.62), which can be obtained by making ε-dependent gauge transformations. We shall restrict our attention to classes of gauges which are consistent with our ansatz (1.8) for the metric, as discussed in Sec. I E above. This class of gauges has the properties that (i) the deviation of the metric from Kerr is $\lesssim \varepsilon$ over the entire inspiral, and (ii) the expansion (2.62) of the self-acceleration is valid. These restrictions exclude, for example, the gauge choice which makes $a^{(1)}_\nu \equiv 0$, since in that gauge the particle does not inspiral, and the metric perturbation must therefore become of order unity over an inspiral time. We note that alternative classes of gauges have been suggested and explored by Mino [56,72,81].

Sixth, from the formula (2.35b) for the generalized angle variables q_α together with Eqs. (2.27d) and (2.31), it follows that q_t can be written as

$$q_t = t + f_t(r, \theta, P_\alpha) \tag{2.65}$$

for some function f_t. All of the other angle and action variables are independent of t. Therefore the vector field $\partial/\partial t$ on phase space is just $\partial/\partial q_t$; the symmetry $t \to t + \Delta t$ with x^i, p_μ fixed is the same as the symmetry $q_t \to q_t + \Delta t$ with q_r, q_θ, q_ϕ, and J_α fixed. Since the self-acceleration as well as the background geodesic motion respect this symmetry, all of the terms on the right-hand side of Eqs. (2.63) must be independent of q_t. A similar argument shows that they are independent of q_ϕ. This gives

$$\frac{dq_\alpha}{d\tau} = \omega_\alpha(\tilde{P}_j) + \varepsilon g^{(1)}_\alpha(q_A, \tilde{P}_j) + \varepsilon^2 g^{(2)}_\alpha(q_A, \tilde{P}_j) + O(\varepsilon^3), \tag{2.66a}$$

$$\frac{d\tilde{P}_i}{d\tau} = \varepsilon G^{(1)}_i(q_A, \tilde{P}_j) + \varepsilon^2 G^{(2)}_i(q_A, \tilde{P}_j) + O(\varepsilon^3), \tag{2.66b}$$

where $q_A \equiv (q_r, q_\theta)$.

Seventh, consider the evolution of the black hole background. So far in our analysis we have assumed that the particle moves in a fixed Kerr background, and is subject to a self-force $a_\nu = \varepsilon a^{(1)}_\nu + \varepsilon^2 a^{(2)}_\nu + O(\varepsilon^3)$. In reality, the center of mass, 4-momentum, and spin angular momentum of the black hole will gradually evolve due to the gravitational radiation passing through the event horizon. The total change in the mass M of the black hole over the inspiral timescale $\sim M/\varepsilon$ is $\sim M\varepsilon$. It follows that the timescale for the black hole mass to change by a factor of order unity is $\sim M/\varepsilon^2$. The same timescale governs the evolution of the other black hole parameters.

This effect of evolution of the black hole background will alter the inspiral at the first subleading order (post-1-adiabatic order) in our two-timescale expansion. A complete calculation of the inspiral to this order requires solving simultaneously for the motion of the particle and the gradual evolution of the background. We introduce the extra variables

$$M_A = (M_1, M_2) = (M, a), \tag{2.67}$$

the mass and spin parameters of the black hole. We modify the equations of motion (2.66) by showing explicitly the dependence of the frequencies ω_α and the forcing functions $g^{(n)}_\alpha$ and $G^{(n)}_i$ on these parameters (the dependence has up to now been implicit). We also add to the system of equations the following evolution equations for the black hole parameters:

$$\frac{dM_A}{d\tau} = \varepsilon^2 \hat{G}_A(q_B, \tilde{P}_j, M_B) + O(\varepsilon^3), \tag{2.68}$$

where $A = 1, 2$. Here $\hat{G}_A$ are some functions describing the fluxes of energy and angular momentum down the horizon, whose explicit form will not be important for our analyses. They can, in principle, be computed using, for example, the techniques developed in Ref. [99].[19] The reason for the prefactor of ε^2 is that the evolution timescale for the black hole parameters is $\sim M/\varepsilon^2$, as discussed above. The functions $\hat{G}_A$ are independent of q_t and q_ϕ for the reason discussed near Eq. (2.66): the fluxes through the horizon respect the symmetries of the background spacetime. Finally, we have omitted in the set of new variables (2.67) the orientation of the total angular momentum, the location of the center of mass, and the total linear momen-

[19] These techniques naturally furnish the derivatives of M_A with respect to Boyer-Lindquist time t, not proper time τ as in Eq. (2.68). However, this difference is unimportant; one can easily convert from one variable to the other by multiplying the functions $\hat{G}_A$ by the standard expression for $dt/d\tau$ [100],

$$\frac{dt}{d\tau} = \frac{\tilde{E}}{\Sigma}\left(\frac{\varpi^4}{\Delta} - a^2\sin^2\theta\right) + \frac{a\tilde{L}_z}{\Sigma}\left(1 - \frac{\varpi^2}{\Delta}\right),$$

where $\varpi = \sqrt{r^2 + a^2}$. This expression can be written in terms of q_A, $\tilde{P}_i$, and M_A, and is valid for accelerated motion as well as geodesic motion by Eqs. (2.25) and (2.39a).

tum of the system, since these parameters are not coupled to the inspiral motion at the leading order. However, it would be possible to enlarge the set of variables M_A to include these parameters without modifying in any way the analyses in the rest of this paper.

These modifications result in the final system of equations (2.47).

Finally we note that an additional effect arises due to the fact that the action-angle variables we use are defined, at each instant, to be the action-angle variables associated with the black hole background at that time. In other words, the coordinate transformation on phase space from $(x^\nu, p_\nu) \to (q_\alpha, J_\alpha)$ acquires an additional dependence on time. Therefore, the Jacobian of this transformation, which was used in deriving the evolution equations (2.37), has an extra term. However, the corresponding correction to the evolution equations can be absorbed into a redefinition of the forcing term $g^{(2)}_\alpha$.

G. Conservative and dissipative pieces of the forcing terms

In this subsection we define a splitting of the forcing terms g_α and G_i in the equations of motion (2.47) into conservative and dissipative pieces, and review some properties of this decomposition derived by Mino [67].

We start by defining some notation. Suppose that we have a particle at a point $\mathcal{P}$ with 4-velocity u^μ, and that we are given a linearized metric perturbation $h_{\mu\nu}$ which is a solution (not necessarily the retarded solution) of the linearized Einstein equation for which the source is a delta function on the geodesic determined by $\mathcal{P}$ and u^μ. The self-acceleration of the particle is then some functional of $\mathcal{P}$, u^μ, $h_{\mu\nu}$ and of the spacetime metric $g_{\mu\nu}$, which we write as

$$a^\mu[\mathcal{P}, u^\mu, g_{\mu\nu}, h_{\mu\nu}]. \tag{2.69}$$

Note that this functional does not depend on a choice of time orientation for the manifold, and also it is invariant under $u^\mu \to -u^\mu$. The retarded self-acceleration is defined as

$$a^\mu_{\rm ret}[\mathcal{P}, u^\mu, g_{\mu\nu}] = a^\mu[\mathcal{P}, u^\mu, g_{\mu\nu}, h^{\rm ret}_{\mu\nu}], \tag{2.70}$$

where $h^{\rm ret}_{\mu\nu}$ is the retarded solution to the linearized Einstein equation obtained using the time orientation that is determined by demanding that u^μ be future directed. This is the physical self-acceleration which is denoted by a^μ throughout the rest of this paper. Similarly, the advanced self-acceleration is

$$a^\mu_{\rm adv}[\mathcal{P}, u^\mu, g_{\mu\nu}] = a^\mu[\mathcal{P}, u^\mu, g_{\mu\nu}, h^{\rm adv}_{\mu\nu}], \tag{2.71}$$

where $h^{\rm adv}_{\mu\nu}$ is the advanced solution. It follows from these definitions that

$$a^\mu_{\rm ret}[\mathcal{P}, -u^\mu, g_{\mu\nu}] = a^\mu_{\rm adv}[\mathcal{P}, u^\mu, g_{\mu\nu}]. \tag{2.72}$$

We define the conservative and dissipative self-accelerations to be

$$a^\mu_{\rm cons} = \tfrac{1}{2}(a^\mu_{\rm ret} + a^\mu_{\rm adv}) \tag{2.73}$$

and

$$a^\mu_{\rm diss} = \tfrac{1}{2}(a^\mu_{\rm ret} - a^\mu_{\rm adv}). \tag{2.74}$$

The physical self-acceleration can then be decomposed as

$$a^\mu = a^\mu_{\rm ret} = a^\mu_{\rm cons} + a^\mu_{\rm diss}. \tag{2.75}$$

A similar decomposition applies to the forcing functions (2.64):

$$g^{(s)}_\alpha = g^{(s)}_{\alpha\rm cons} + g^{(s)}_{\alpha\rm diss}, \tag{2.76a}$$

$$G^{(s)}_i = G^{(s)}_{i\rm cons} + G^{(s)}_{i\rm diss}, \tag{2.76b}$$

for $s = 1, 2$.

Next, we note that if ψ is any diffeomorphism from the spacetime to itself, then the self-acceleration satisfies the covariance relation

$$a^\nu_{\rm ret}[\psi(\mathcal{P}), \psi^* u^\nu, \psi^* g_{\mu\nu}] = \psi^* a^\nu_{\rm ret}[\mathcal{P}, u^\nu, g_{\mu\nu}]. \tag{2.77}$$

Taking the point $\mathcal{P}$ to be $(t_0, r_0, \theta_0, \phi_0)$ in Boyer-Lindquist coordinates, and choosing ψ to be $t \to 2t_0 - t$, $\phi \to 2\phi_0 - \phi$, then ψ is an isometry, $\psi^* g_{\mu\nu} = g_{\mu\nu}$. It follows that

$$a^\nu_{\rm ret}(-u_t, u_r, u_\theta, -u_\phi) = -\epsilon_\nu a^\nu_{\rm ret}(u_t, u_r, u_\theta, u_\phi), \tag{2.78}$$

where

$$\epsilon_\nu = (1, -1, -1, 1) \tag{2.79}$$

and there is no summation over ν on the right-hand side. Combining this with the identity (2.72) gives

$$a^\nu_{\rm adv}(u_t, u_r, u_\theta, u_\phi) = -\epsilon_\nu a^\nu_{\rm ret}(u_t, -u_r, -u_\theta, u_\phi). \tag{2.80}$$

Now, under the transformation $p_r \to -p_r$, $p_\theta \to -p_\theta$ with other quantities fixed, the action variables and the quantities P_α are invariant, the angle variables q_r and q_θ transform as $q_r \to 2\pi - q_r$, $q_\theta \to 2\pi - q_\theta$, while $q_t - t$ and $q_\phi - \phi$ flip sign. This can be seen from the definitions (2.31) and (2.35b). Explicitly we have

$$\bar{q}_t(x^\gamma, \epsilon_\delta p_\delta) - t = -[\bar{q}_t(x^\gamma, p_\delta) - t], \tag{2.81a}$$

$$\bar{q}_\phi(x^\gamma, \epsilon_\delta p_\delta) - \phi = -[\bar{q}_\phi(x^\gamma, p_\delta) - \phi], \tag{2.81b}$$

$$\bar{q}_A(x^\gamma, \epsilon_\delta p_\delta) = 2\pi - \bar{q}_A(x^\gamma, p_\delta), \tag{2.81c}$$

$$P_i(x^\gamma, \epsilon_\delta p_\delta) = P_i(x^\gamma, p_\delta), \tag{2.81d}$$

where we use the values (2.79) of ϵ_α, the functions $\bar{q}_\alpha$ are defined before Eq. (2.57), and $q_A = (q_r, q_\theta)$. If we now differentiate with respect to p_α holding x^α fixed and use the definitions (2.53b), (2.57), and (2.59b) of the functions f^ν_α and F^ν_i, we obtain

$$f^\nu_\alpha(x^\beta, \epsilon_\gamma u_\gamma) = -\epsilon_\nu f^\nu_\alpha(x^\beta, u_\gamma), \tag{2.82a}$$
$$F^\nu_i(x^\beta, \epsilon_\gamma u_\gamma) = \epsilon_\nu F^\nu_i(x^\beta, u_\gamma). \tag{2.82b}$$

We now compute the conservative and dissipative pieces of the forcing functions $g^{(1)}_\alpha$ and $G^{(1)}_i$, using the definitions (2.64) and (2.76). Using the results (2.80) and (2.82) we obtain

$$\begin{aligned} g^{(1)}_{\alpha\mathrm{adv}}(u_\gamma) &= f^\nu_\alpha(u_\gamma) a^{(1)}_{\nu\mathrm{adv}}(u_\gamma) \\ &= [-\epsilon_\nu f^\nu_\alpha(\epsilon_\gamma u_\gamma)][-\epsilon_\nu a^{(1)}_{\nu\mathrm{ret}}(\epsilon_\gamma u_\gamma)] \\ &= g^{(1)}_{\alpha\mathrm{ret}}(\epsilon_\gamma u_\gamma). \end{aligned} \tag{2.83}$$

A similar computation gives

$$G^{(1)}_{i\mathrm{adv}}(u_\gamma) = -G^{(1)}_{i\mathrm{ret}}(\epsilon_\gamma u_\gamma), \tag{2.84}$$

and using that the mapping $x^\nu \to x^\nu$, $u_\mu \to \epsilon_\mu u_\mu$ corresponds to $\tilde{P}_j \to \tilde{P}_j$, $q_r \to 2\pi - q_r$, $q_\theta \to 2\pi - q_\theta$ finally yields the identities

$$g^{(1)}_{\alpha\mathrm{cons}}(q_A, \tilde{P}_j) = [g^{(1)}_\alpha(q_r, q_\theta, \tilde{P}_j) + g^{(1)}_\alpha(2\pi - q_r, 2\pi - q_\theta, \tilde{P}_j)]/2, \tag{2.85a}$$
$$g^{(1)}_{\alpha\mathrm{diss}}(q_A, \tilde{P}_j) = [g^{(1)}_\alpha(q_r, q_\theta, \tilde{P}_j) - g^{(1)}_\alpha(2\pi - q_r, 2\pi - q_\theta, \tilde{P}_j)]/2, \tag{2.85b}$$

and

$$G^{(1)}_{i\mathrm{cons}}(q_A, \tilde{P}_j) = [G^{(1)}_i(q_r, q_\theta, \tilde{P}_j) - G^{(1)}_i(2\pi - q_r, 2\pi - q_\theta, \tilde{P}_j)]/2, \tag{2.86a}$$
$$G^{(1)}_{i\mathrm{diss}}(q_A, \tilde{P}_j) = [G^{(1)}_i(q_r, q_\theta, \tilde{P}_j) + G^{(1)}_i(2\pi - q_r, 2\pi - q_\theta, \tilde{P}_j)]/2. \tag{2.86b}$$

Here we have used the fact that the forcing functions are independent of q_t and q_ϕ, as discussed in the last subsection. Similar equations apply with $g^{(1)}_\alpha$ and $G^{(1)}_i$ replaced by the higher order forcing terms $g^{(s)}_\alpha$ and $G^{(s)}_i$, $s \geq 2$.

It follows from the identity (2.86a) that, for the action-variable forcing functions $G^{(1)}_i$, the average over the 2-torus parametrized by q_r and q_θ of the conservative piece vanishes. For generic orbits (for which ω_r and ω_θ are incommensurate), the torus average is equivalent to a time average, and so it follows that the time average vanishes, a result first derived by Mino [67]. Similarly, from Eqs. (2.85) it follows that the torus average of the dissipative pieces of $g^{(1)}_\alpha$ vanishes.

III. A GENERAL WEAKLY PERTURBED DYNAMICAL SYSTEM

In the remainder of this paper we will study in detail the behavior of a one-parameter family of dynamical systems parametrized by a dimensionless parameter ε. We shall be interested in the limiting behavior of the systems as $\varepsilon \to 0$. The system contains $N + M$ dynamical variables

$$\mathbf{q}(t) = (q_1(t), q_2(t), \ldots, q_N(t)), \tag{3.1a}$$
$$\mathbf{J}(t) = (J_1(t), J_2(t), \ldots, J_M(t)), \tag{3.1b}$$

and is defined by the equations

$$\frac{dq_\alpha}{dt} = \omega_\alpha(\mathbf{J}, \tilde{t}) + \varepsilon g_\alpha(\mathbf{q}, \mathbf{J}, \tilde{t}, \varepsilon), \qquad 1 \leq \alpha \leq N, \tag{3.2a}$$
$$\frac{dJ_\lambda}{dt} = \varepsilon G_\lambda(\mathbf{q}, \mathbf{J}, \tilde{t}, \varepsilon), \qquad 1 \leq \lambda \leq M. \tag{3.2b}$$

Here the variable $\tilde{t}$ is the "slow time" variable defined by

$$\tilde{t} = \varepsilon t. \tag{3.3}$$

We assume that the functions g_α and G_λ can be expanded as

$$\begin{aligned} g_\alpha(\mathbf{q}, \mathbf{J}, \tilde{t}, \varepsilon) &= \sum_{s=1}^{\infty} g^{(s)}_\alpha(\mathbf{q}, \mathbf{J}, \tilde{t}) \varepsilon^{s-1} \\ &= g^{(1)}_\alpha(\mathbf{q}, \mathbf{J}, \tilde{t}) + g^{(2)}_\alpha(\mathbf{q}, \mathbf{J}, \tilde{t})\varepsilon + O(\varepsilon^2) \end{aligned} \tag{3.4}$$

and

$$\begin{aligned} G_\lambda(\mathbf{q}, \mathbf{J}, \tilde{t}, \varepsilon) &= \sum_{s=1}^{\infty} G^{(s)}_\lambda(\mathbf{q}, \mathbf{J}, \tilde{t}) \varepsilon^{s-1} \\ &= G^{(1)}_\lambda(\mathbf{q}, \mathbf{J}, \tilde{t}) + G^{(2)}_\lambda(\mathbf{q}, \mathbf{J}, \tilde{t})\varepsilon + O(\varepsilon^2). \end{aligned} \tag{3.5}$$

These series are assumed to be asymptotic series in ε as $\varepsilon \to 0$ that are uniform in $\tilde{t}$.[20] We assume that the functions ω_α, $g^{(s)}_\alpha$, and $G^{(s)}_\lambda$ are smooth functions of their arguments, and that the frequencies ω_α are nowhere vanishing. Finally, the functions g_α and G_λ are assumed to be periodic

[20] In other words, there exists $\tilde{T} > 0$ such that for every $\mathbf{q}$, $\mathbf{J}$, every integer N, and every $\delta > 0$, there exists $\epsilon_1 = \epsilon_1(\mathbf{q}, \mathbf{J}, N, \delta)$ such that

$$\left| g_\alpha(\mathbf{q}, \mathbf{J}, \tilde{t}, \varepsilon) - \sum_{s=1}^{N} g^{(s)}_\alpha(\mathbf{q}, \mathbf{J}, \tilde{t}) \varepsilon^{s-1} \right| < \delta \varepsilon^{N-1}$$

for all $\tilde{t}$ with $0 < \tilde{t} < \tilde{T}$ and for all ε with $0 < \varepsilon < \epsilon_1$.

in each variable q_α with period 2π:

$$g_\alpha(\mathbf{q}+2\pi\mathbf{k},\mathbf{J},\tilde{t}) = g_\alpha(\mathbf{q},\mathbf{J},\tilde{t}), \qquad 1\le\alpha\le N, \tag{3.6a}$$
$$G_\lambda(\mathbf{q}+2\pi\mathbf{k},\mathbf{J},\tilde{t}) = G_\lambda(\mathbf{q},\mathbf{J},\tilde{t}), \qquad 1\le\lambda\le M, \tag{3.6b}$$

where $\mathbf{k} = (k_1,\ldots,k_N)$ is an arbitrary N-tuple of integers.

Equations (2.47) derived in the previous section describing the inspiral of a point particle into a Kerr black hole are a special case of the dynamical system (3.2). This can be seen using the identifications $t=\tau$, $\mathbf{q} = (q_t, q_r, q_\theta, q_\phi)$, $\mathbf{J} = (\tilde{P}_2, \tilde{P}_3, \tilde{P}_4, M_1, M_2)$, $G^{(1)}_\lambda = (G^{(1)}_2, G^{(1)}_3, G^{(1)}_4, 0, 0)$, and $G^{(2)}_\lambda = (G^{(2)}_2, G^{(2)}_3, G^{(2)}_4, \hat{G}_1, \hat{G}_2)$. The forcing functions $g^{(s)}_\alpha$ and $G^{(s)}_\lambda$ are periodic functions of q_α since they depend only on the variables $q_A = (q_r, q_\theta)$ which are angle variables; they do not depend on the variable q_t which is not an angle variable. Note that the system (3.2) allows the forcing functions $g^{(s)}_\alpha$, $G^{(s)}_\lambda$ and frequencies ω_α to depend in an arbitrary way on the slow time $\tilde{t}$, whereas no such dependence is seen in the Kerr inspiral system (2.47). The system studied here is thus slightly more general than is required for our specific application. We include the dependence on $\tilde{t}$ for greater generality and because it does not require any additional complexity in the analysis.

Another special case of the system (3.2) is when $N = M$ and when there exists a function $H(\mathbf{J},\tilde{t})$ such that

$$\omega_\alpha(\mathbf{J},\tilde{t}) = \frac{\partial H(\mathbf{J},\tilde{t})}{\partial J_\alpha} \tag{3.7}$$

for $1\le\alpha\le N$. In this case the system (3.2) represents a Hamiltonian system with slowly varying Hamiltonian $H(\mathbf{J},\tilde{t})$, with action-angle variables (q_α, J_α), and subject to arbitrary weak perturbing forces that vary slowly with time. The perturbed system is not necessarily Hamiltonian.

Because of the periodicity conditions (3.6), we can, without loss of generality, interpret the variables q_α to be coordinates on the N-torus T^N, and take Eqs. (3.2) to be defined on the product of this N-torus with an open set. This interpretation will be useful below.

In the next several sections we will study in detail the behavior of solutions of the system (3.2) in the limit $\varepsilon\to 0$ using a two-timescale expansion. We follow closely the exposition in the book by Kevorkian and Cole [74], except that we generalize their analysis and also correct some errors (see Appendix B). For clarity we treat first, in Sec. IV, the simple case of a single degree of freedom, $N = M = 1$. Section V treats the case of general N and M, but with the restriction that the forcing functions g_α and G_λ contain no resonant pieces (this is defined in Sec. V C). The general case with resonances is treated in the forthcoming papers [78,79]. Finally, in Sec. VI we present a numerical integration of a particular example of a dynamical system, in order to illustrate and validate the general theory of Secs. IV and V.

IV. SYSTEMS WITH A SINGLE DEGREE OF FREEDOM

A. Overview

For systems with a single degree of freedom, the general equations of motion (3.2) discussed in Sec. III reduce to

$$\dot{q}(t) = \omega(J,\tilde{t}) + \varepsilon g(q,J,\tilde{t},\varepsilon), \tag{4.1a}$$
$$\dot{J}(t) = \varepsilon G(q,J,\tilde{t},\varepsilon), \tag{4.1b}$$

for some functions G and g, where $\tilde{t} = \varepsilon t$ is the slow time variable. The asymptotic expansions (3.4) and (3.5) of the forcing functions reduce to

$$g(q,J,\tilde{t},\varepsilon) = \sum_{s=1}^{\infty} g^{(s)}(q,J,\tilde{t})\varepsilon^{s-1} = g^{(1)}(q,J,\tilde{t}) + g^{(2)}(q,J,\tilde{t})\varepsilon + O(\varepsilon^2) \tag{4.2}$$

and

$$G(q,J,\tilde{t},\varepsilon) = \sum_{s=1}^{\infty} G^{(s)}(q,J,\tilde{t})\varepsilon^{s-1} = G^{(1)}(q,J,\tilde{t}) + G^{(2)}(q,J,\tilde{t})\varepsilon + O(\varepsilon^2). \tag{4.3}$$

Also, the periodicity conditions (3.6) reduce to

$$g(q+2\pi,J,\tilde{t}) = g(q,J,\tilde{t}), \tag{4.4a}$$
$$G(q+2\pi,J,\tilde{t}) = G(q,J,\tilde{t}). \tag{4.4b}$$

In this section we apply two-timescale expansions to study classes of solutions of Eqs. (4.1) in the limit $\varepsilon\to 0$. We start in Sec. IV B by defining our conventions and notations for Fourier decompositions of the perturbing forces. The heart of the method is the ansatz we make for the form of the solutions, which is given in Sec. IV C. Section IV D summarizes the results we obtain at each order in the expansion, and the derivations are given in Sec. IV E. Although the results of this section are not directly applicable to the problem of inspirals in Kerr spacetime, the analysis of this section gives an introduction to the method of analysis, and is considerably simpler than the multivariable case treated in Sec. V below.

B. Fourier expansions of the perturbing forces

The periodicity conditions (4.4) apply at each order in the expansion in powers of ε:

$$g^{(s)}(q+2\pi,J,\tilde{t}) = g^{(s)}(q,J,\tilde{t}), \tag{4.5a}$$
$$G^{(s)}(q+2\pi,J,\tilde{t}) = G^{(s)}(q,J,\tilde{t}). \tag{4.5b}$$

It follows that these functions can be expanded as Fourier series:

$$g^{(s)}(q, J, \tilde{t}) = \sum_{k=-\infty}^{\infty} g_k^{(s)}(J, \tilde{t}) e^{ikq}, \tag{4.6a}$$

$$G^{(s)}(q, J, \tilde{t}) = \sum_{k=-\infty}^{\infty} G_k^{(s)}(J, \tilde{t}) e^{ikq}, \tag{4.6b}$$

where

$$g_k^{(s)}(J, \tilde{t}) = \frac{1}{2\pi} \int_0^{2\pi} dq e^{-ikq} g^{(s)}(q, J, \tilde{t}), \tag{4.7a}$$

$$G_k^{(s)}(J, \tilde{t}) = \frac{1}{2\pi} \int_0^{2\pi} dq e^{-ikq} G^{(s)}(q, J, \tilde{t}). \tag{4.7b}$$

For any periodic function $f = f(q)$, we introduce the notation

$$\langle f \rangle = \frac{1}{2\pi} \int_0^{2\pi} f(q) dq \tag{4.8}$$

for the average part of f, and

$$\hat{f}(q) = f(q) - \langle f \rangle \tag{4.9}$$

for the remaining part of f. It follows from these definitions that

$$\langle g^{(s)}(q, J, \tilde{t}) \rangle = g_0^{(s)}(J, \tilde{t}), \qquad \langle G^{(s)}(q, J, \tilde{t}) \rangle = G_0^{(s)}(J, \tilde{t}), \tag{4.10}$$

and that

$$\hat{g}^{(s)}(q, J, \tilde{t}) = \sum_{k \neq 0} g_k^{(s)}(J, \tilde{t}) e^{ikq}, \tag{4.11a}$$

$$\hat{G}^{(s)}(q, J, \tilde{t}) = \sum_{k \neq 0} G_k^{(s)}(J, \tilde{t}) e^{ikq}. \tag{4.11b}$$

We also have the identities

$$\langle f_{,q} \rangle = \langle \hat{f} \rangle = 0, \tag{4.12a}$$

$$\langle fg \rangle = \langle \hat{f}\, \hat{g} \rangle + \langle f \rangle \langle g \rangle \tag{4.12b}$$

for any periodic functions $f(q)$, $g(q)$.

For any periodic function f, we also define a particular antiderivative $\mathcal{I}\hat{f}$ of $\hat{f}$ by

$$(\mathcal{I}\hat{f})(q) \equiv \sum_{k \neq 0} \frac{f_k}{ik} e^{ikq}, \tag{4.13}$$

where $f_k = \int dq e^{-ikq} f(q)/(2\pi)$ are the Fourier coefficients of f. This operator satisfies the identities

$$(\mathcal{I}\hat{f})_{,q} = \hat{f}, \tag{4.14a}$$

$$\langle (\mathcal{I}\hat{f}) \hat{g} \rangle = -\langle \hat{f} (\mathcal{I}\hat{g}) \rangle, \tag{4.14b}$$

$$\langle \hat{f} (\mathcal{I}\hat{f}) \rangle = 0. \tag{4.14c}$$

C. Two-timescale ansatz for the solution

We now discuss the *ansatz* we use for the form of the solutions of the equations of motion. This ansatz will be justified *a posteriori* order by order in ε. The method used here is sometimes called the "method of strained coordinates" [74].

We assume that q and J have asymptotic expansions in ε as functions of two different variables, the slow time parameter $\tilde{t} = \varepsilon t$, and a phase variable Ψ (also called a "fast time parameter"), the dependence on which is periodic with period 2π. Thus we assume

$$q(t, \varepsilon) = \sum_{s=0}^{\infty} \varepsilon^s q^{(s)}(\Psi, \tilde{t}) = q^{(0)}(\Psi, \tilde{t}) + \varepsilon q^{(1)}(\Psi, \tilde{t}) + O(\varepsilon^2), \tag{4.15a}$$

$$J(t, \varepsilon) = \sum_{s=0}^{\infty} \varepsilon^s J^{(s)}(\Psi, \tilde{t}) = J^{(0)}(\Psi, \tilde{t}) + \varepsilon J^{(1)}(\Psi, \tilde{t}) + O(\varepsilon^2). \tag{4.15b}$$

These asymptotic expansions are assumed to be uniform in $\tilde{t}$. The expansion coefficients $J^{(s)}$ are each periodic in the phase variable Ψ with period 2π:

$$J^{(s)}(\Psi + 2\pi, \tilde{t}) = J^{(s)}(\Psi, \tilde{t}). \tag{4.16}$$

The phase variable Ψ is chosen so that angle variable q increases by 2π when Ψ increases by 2π; this implies that the expansion coefficients $q^{(s)}$ satisfy

$$q^{(0)}(\Psi + 2\pi, \tilde{t}) = q^{(0)}(\Psi, \tilde{t}) + 2\pi, \tag{4.17a}$$

$$q^{(s)}(\Psi + 2\pi, \tilde{t}) = q^{(s)}(\Psi, \tilde{t}), \qquad s \geq 1. \tag{4.17b}$$

The angular velocity $\Omega = d\Psi/dt$ associated with the phase Ψ is assumed to depend only on the slow time variable $\tilde{t}$ (so it can vary slowly with time), and on ε. We assume that it has an asymptotic expansion in ε as $\varepsilon \to 0$, which is uniform in $\tilde{t}$:

$$\frac{d\Psi}{dt} = \Omega(\tilde{t}, \varepsilon) = \sum_{s=0}^{\infty} \varepsilon^s \Omega^{(s)}(\tilde{t}) \tag{4.18}$$

$$= \Omega^{(0)}(\tilde{t}) + \varepsilon \Omega^{(1)}(\tilde{t}) + O(\varepsilon^2). \tag{4.19}$$

Equation (4.19) serves to define the phase variable Ψ in terms of the angular velocity variables $\Omega^{(s)}(\tilde{t})$, $s = 0, 1, 2 \ldots$, up to constants of integration. One constant of integration arises at each order in ε. Without loss of generality we choose these constants of integration so that

$$q^{(s)}(0, \tilde{t}) = 0 \tag{4.20}$$

for all s, $\tilde{t}$. Note that this does not restrict the final solutions $q(t, \varepsilon)$ and $J(t, \varepsilon)$, as we show explicitly below, because there are additional constants of integration that arise when solving for the functions $q^{(s)}(\Psi, \tilde{t})$ and $J^{(s)}(\Psi, \tilde{t})$.

Roughly speaking, the meaning of these assumptions is the following. The solution of the equations of motion consists of a mapping from (t, ε) to (q, J). That mapping contains dynamics on two different timescales, the dy-

namical timescale ~ 1 and the slow timescale $\sim 1/\varepsilon$. The mapping can be uniquely written as the composition of two mappings,

$$(t, \varepsilon) \rightarrow (\Psi, \tilde{t}, \varepsilon) \rightarrow (q, J), \tag{4.21}$$

such that the first mapping contains all the fast dynamics, and is characterized by the slowly evolving frequency $\Omega(\tilde{t}, \varepsilon)$, and the second mapping contains dynamics only on the slow timescale.

D. Results of the two-timescale analysis

By substituting the ansatz (4.15b), (4.16), (4.17), (4.18), (4.19), and (4.20) into the equations of motion (4.1), we find that all of the assumptions made in the ansatz can be satisfied, and that all of the expansion coefficients are uniquely determined, order by order in ε. This derivation is given in Sec. IV E below. Here we list the results obtained for the various expansion coefficients up to the leading and subleading orders.

1. Terminology for various orders of the approximation

We can combine the definitions just summarized to obtain an explicit expansion for the quantity of most interest, the angle variable q as a function of time. From the periodicity condition (4.17a) it follows that the function $q^{(0)}(\Psi, \tilde{t})$ can be written as $\Psi + \bar{q}^{(0)}(\Psi, \tilde{t})$ where $\bar{q}^{(0)}$ is a periodic function of Ψ. [We shall see that $\bar{q}^{(0)}$ in fact vanishes; cf. Eq. (4.27) below.] From the definitions (3.3) and (4.19), we can write the phase variable Ψ as

$$\Psi = \frac{1}{\varepsilon}\psi^{(0)}(\tilde{t}) + \psi^{(1)}(\tilde{t}) + \varepsilon\psi^{(2)}(\tilde{t}) + O(\varepsilon^2), \tag{4.22}$$

where the functions $\psi^{(s)}(\tilde{t})$ are defined by

$$\psi^{(s)}(\tilde{t}) = \int^{\tilde{t}} d\tilde{t}' \Omega^{(s)}(\tilde{t}'). \tag{4.23}$$

Inserting this into the expansion (4.15a) of q and using the above expression for $q^{(0)}$ gives

$$q(t, \varepsilon) = \frac{1}{\varepsilon}\psi^{(0)}(\tilde{t}) + [\psi^{(1)}(\tilde{t}) + \bar{q}^{(0)}(\Psi, \tilde{t})] + \varepsilon[\psi^{(2)}(\tilde{t}) + q^{(1)}(\Psi, \tilde{t})] + O(\varepsilon^2). \tag{4.24}$$

We will call the leading order $O(1/\varepsilon)$ term in Eq. (4.24) the *adiabatic approximation*, the subleading $O(1)$ term the *post-1-adiabatic* term, the next $O(\varepsilon)$ term the *post-2-adiabatic* term, etc. This choice of terminology is motivated by the terminology used in post-Newtonian theory.

It is important to note that the expansion in powers of ε in Eq. (4.24) is *not* a straightforward power series expansion at fixed $\tilde{t}$, since the variable Ψ depends on ε. [The precise definition of the expansion of the solution which we are using is given by Eqs. (4.15a), (4.15b), (4.16), (4.17), (4.18), (4.19), and (4.20)]. Nevertheless, the expansion (4.24) as written correctly captures the ε dependence of the secular pieces of the solution, since the functions $\bar{q}^{(0)}$ and $q^{(1)}$ are periodic functions of Ψ and so have no secular pieces.

2. Adiabatic order

First, the zeroth order action variable is given by

$$J^{(0)}(\Psi, \tilde{t}) = \mathcal{J}^{(0)}(\tilde{t}), \tag{4.25}$$

where $\mathcal{J}^{(0)}$ satisfies the differential equation

$$\frac{d\mathcal{J}^{(0)}(\tilde{t})}{d\tilde{t}} = G_0^{(1)}[\mathcal{J}^{(0)}(\tilde{t}), \tilde{t}]. \tag{4.26}$$

Here the right-hand side denotes the average over q of the forcing term $G^{(1)}[q, \mathcal{J}^{(0)}(\tilde{t}), \tilde{t}]$; cf. Eqs. (4.6) above. The zeroth order angle variable is given by

$$q^{(0)}(\Psi, \tilde{t}) = \Psi, \tag{4.27}$$

and the angular velocity Ω that defines the phase variable Ψ is given to zeroth order by

$$\Omega^{(0)}(\tilde{t}) = \omega[\mathcal{J}^{(0)}(\tilde{t}), \tilde{t}]. \tag{4.28}$$

Note that this approximation is equivalent to the following simple prescription: (i) Truncate the equations of motion (4.1) to the leading order in ε:

$$\dot{q}(t) = \omega(J, \tilde{t}) + \varepsilon g^{(1)}(q, J, \tilde{t}), \tag{4.29a}$$

$$\dot{J}(t) = \varepsilon G^{(1)}(q, J, \tilde{t}); \tag{4.29b}$$

(ii) omit the driving term $g^{(1)}$ in the equation for the angle variable; and (iii) replace the driving term $G^{(1)}$ in the equation for the action variable with its average over q.

3. Post-1-adiabatic order

Next, the first order action variable is given by

$$J^{(1)}(\Psi, \tilde{t}) = \frac{\mathcal{I}\hat{G}^{(1)}[\Psi, \mathcal{J}^{(0)}(\tilde{t}), \tilde{t}]}{\Omega^{(0)}(\tilde{t})} + \mathcal{J}^{(1)}(\tilde{t}), \tag{4.30}$$

where the symbol $\mathcal{I}$ on the right-hand side denotes the integration operator (4.13) with respect to Ψ. In Eq. (4.30) the quantity $\mathcal{J}^{(1)}(\tilde{t})$ satisfies the differential equation

$$\frac{d\mathcal{J}^{(1)}(\tilde{t})}{d\tilde{t}} - \frac{\partial G_0^{(1)}}{\partial J}[\mathcal{J}^{(0)}(\tilde{t}), \tilde{t}]\mathcal{J}^{(1)}(\tilde{t}) = \frac{\langle \frac{\partial \hat{G}^{(1)}}{\partial J}\mathcal{I}\hat{G}^{(1)}\rangle}{\Omega^{(0)}(\tilde{t})} - \frac{\langle \hat{G}^{(1)}\hat{g}^{(1)}\rangle}{\Omega^{(0)}(\tilde{t})} + G_0^{(2)}. \tag{4.31}$$

Here it is understood that the quantities on the right-hand side are evaluated at $q = q^{(0)} = \Psi$ and $J = \mathcal{J}^{(0)}(\tilde{t})$. The subleading correction to the phase variable Ψ is given by

$$\Omega^{(1)}(\tilde{t}) = \frac{\partial \omega}{\partial J}[\mathcal{J}^{(0)}(\tilde{t}), \tilde{t}]\mathcal{J}^{(1)}(\tilde{t}) + g_0^{(1)}[\mathcal{J}^{(0)}(\tilde{t}), \tilde{t}]. \quad (4.32)$$

Finally, the subleading term in the angle variable is

$$q^{(1)}(\Psi, \tilde{t}) = \hat{q}^{(1)}(\Psi, \tilde{t}) + \mathcal{Q}^{(1)}(\tilde{t}), \quad (4.33)$$

where

$$\hat{q}^{(1)}(\Psi, \tilde{t}) = \frac{1}{\Omega^{(0)}(\tilde{t})^2}\frac{\partial \omega}{\partial J}[\mathcal{J}^{(0)}(\tilde{t}), \tilde{t}]\mathcal{I}^2\hat{G}^{(1)}[\Psi, \mathcal{J}^{(0)}(\tilde{t}), \tilde{t}] + \frac{1}{\Omega^{(0)}(\tilde{t})}\mathcal{I}\hat{g}^{(1)}[\Psi, \mathcal{J}^{(0)}(\tilde{t}), \tilde{t}] \quad (4.34)$$

and

$$\mathcal{Q}^{(1)}(\tilde{t}) = -\hat{q}^{(1)}(0, \tilde{t}). \quad (4.35)$$

4. Discussion

One of the key results of the general analysis of this section is the identification of which pieces of the external forces are required to compute the adiabatic and post-1-adiabatic solutions. From Eqs. (4.24), (4.26), and (4.28), the adiabatic solution depends only on the averaged piece $G_0^{(1)}(J, \tilde{t}) = \langle G^{(1)}(q, J, \tilde{t})\rangle$ of the leading order external force $G^{(1)}$. This quantity is purely dissipative, as can be seen in the context of inspirals in Kerr spacetime from Eqs. (2.85) and (2.86). More generally, if the perturbing forces g and G arise from a perturbation $\varepsilon\Delta H = \sum_s \varepsilon^s \Delta H^{(s)}$ to the Hamiltonian, then the forcing function $G^{(s)}$ is

$$G^{(s)}(q, J, \tilde{t}) = -\frac{\partial \Delta H^{(s)}(q, J, \tilde{t})}{\partial q},$$

and it follows that the average over q of $G^{(s)}$ vanishes.

At the next order, the post-1-adiabatic term $\psi^{(1)}(\tilde{t})$ depends on the averaged piece $G_0^{(2)}(J, \tilde{t}) = \langle G^{(2)}(q, J, \tilde{t})\rangle$ of the subleading force $G^{(2)}$, again purely dissipative, as well as the remaining conservative and dissipative pieces of the leading order forces $G^{(1)}(q, J, \tilde{t})$ and $g^{(1)}(q, J, \tilde{t})$. This can be seen from Eqs. (4.31) and (4.32). These results have been previously discussed briefly in the EMRI context in Refs. [37,68]. For circular, equatorial orbits, the fact that there is a post-1-adiabatic order contribution from the second order self-force was first argued by Burko [88].

5. Initial conditions and the generality of our ansatz

We will show in the next subsection that our ansatz (4.15a), (4.15b), (4.16), (4.17), (4.18), (4.19), and (4.20) is compatible with the one-parameter family of differential equations (4.1). However, it does not necessarily follow that our ansatz is compatible with the most general one-parameter family $[q(t, \varepsilon), J(t, \varepsilon)]$ of solutions, because of the possibility of choosing arbitrary, ε-dependent initial conditions $q(0, \varepsilon)$ and $J(0, \varepsilon)$ at the initial time $t = 0$.[21] In general, the ε dependence of the solutions arises from both the ε dependence of the initial conditions and the ε dependence of the differential equations. It is possible to choose initial conditions which are incompatible with our ansatz.

To see this explicitly, we evaluate the expansions (4.24) and (4.30) at $t = \tilde{t} = 0$. This gives

$$q(0, \varepsilon) = \varepsilon^{-1}\psi^{(0)}(0) + \psi^{(1)}(0) + O(\varepsilon), \quad (4.36a)$$

$$J(0, \varepsilon) = \mathcal{J}^{(0)}(0) + \varepsilon\mathcal{J}^{(1)}(0) + \varepsilon\frac{\mathcal{I}\hat{G}^{(1)}[\varepsilon^{-1}\psi^{(0)}(0) + \psi^{(1)}(0), \mathcal{J}^{(0)}(0), 0]}{\omega[\mathcal{J}^{(0)}, 0]} + O(\varepsilon^2). \quad (4.36b)$$

Recalling that parameters $\psi^{(0)}(0)$, $\psi^{(1)}(0)$, $\mathcal{J}^{(0)}(0)$, and $\mathcal{J}^{(1)}(0)$ are assumed to be independent of ε, we see that the conditions (4.36) strongly constrain the allowed ε dependence of the initial conditions. We note, however, that the choice of constant (ε-independent) initial conditions

$$q(0, \varepsilon) = q_0, \qquad J(0, \varepsilon) = J_0 \quad (4.37)$$

can be accommodated, which is sufficient for most applications of the formalism. To achieve this one chooses

$$\psi^{(0)}(0) = 0, \qquad \psi^{(1)}(0) = q_0, \qquad \mathcal{J}^{(0)}(0) = J_0, \quad (4.38)$$

and

$$\mathcal{J}^{(1)}(0) = -\frac{\mathcal{I}\hat{G}^{(1)}[q_0, J_0, 0]}{\omega[J_0, 0]}. \quad (4.39)$$

E. Derivation

In this subsection we give the derivation of the results (4.25), (4.26), (4.27), (4.28), (4.29), (4.30), (4.31), (4.32), (4.33), (4.34), and (4.35) summarized above. At each order s we introduce the notation $\mathcal{J}^{(s)}(\tilde{t})$ for the average part of $J^{(s)}(\Psi, \tilde{t})$:

$$\mathcal{J}^{(s)}(\tilde{t}) \equiv \langle J^{(s)}(\Psi, \tilde{t})\rangle = \frac{1}{2\pi}\int_0^{2\pi} J^{(s)}(\Psi, \tilde{t})d\Psi. \quad (4.40)$$

We denote by $\hat{J}^{(s)}$ the remaining part of $J^{(s)}$, as in Eq. (4.9). This gives the decomposition

$$J^{(s)}(\Psi, \tilde{t}) = \mathcal{J}^{(s)}(\tilde{t}) + \hat{J}^{(s)}(\Psi, \tilde{t}) \quad (4.41)$$

for all $s \geq 0$. Similarly, for the angle variable we have the decomposition

$$q^{(s)}(\Psi, \tilde{t}) = \mathcal{Q}^{(s)}(\tilde{t}) + \hat{q}^{(s)}(\Psi, \tilde{t}) \quad (4.42)$$

[21] More generally, we could consider specifying initial conditions at some time $t = t_0$. In that case we would modify the definition of the rescaled time coordinate to $\tilde{t} = \varepsilon(t - t_0)$.

for all $s \geq 1$. [We do not use this notation for the $s = 0$ case for the angle variable, since $q^{(0)}$ is not a periodic function of Ψ, by Eq. (4.17a)].

Using the expansions (4.15a) and (4.15b) of q and J together with the expansion (4.19) of $d\Psi/dt$, we obtain

$$\frac{dq}{dt} = \Omega^{(0)} q^{(0)}_{,\Psi} + \varepsilon[\Omega^{(1)} q^{(0)}_{,\Psi} + \Omega^{(0)} q^{(1)}_{,\Psi} + q^{(0)}_{,\tilde{t}}] + \varepsilon^2[\Omega^{(2)} q^{(0)}_{,\Psi} + \Omega^{(0)} q^{(2)}_{,\Psi} + \Omega^{(1)} q^{(1)}_{,\Psi} + q^{(1)}_{,\tilde{t}}] + O(\varepsilon^3). \tag{4.43}$$

Here we use commas to denote partial derivatives. We now insert this expansion together with a similar expansion for dJ/dt into the equations of motion (4.1) and use the expansions (4.2) and (4.3) of the external forces g and G. Equating coefficients[22] of powers of ε then gives at zeroth order

$$\Omega^{(0)} q^{(0)}_{,\Psi} = \omega, \tag{4.44a}$$

$$\Omega^{(0)} J^{(0)}_{,\Psi} = 0, \tag{4.44b}$$

at first order

$$\Omega^{(0)} q^{(1)}_{,\Psi} - \omega_{,J} J^{(1)} = -\Omega^{(1)} q^{(0)}_{,\Psi} - q^{(0)}_{,\tilde{t}} + g^{(1)}, \tag{4.45a}$$

$$\Omega^{(1)} J^{(0)}_{,\Psi} + \Omega^{(0)} J^{(1)}_{,\Psi} = -J^{(0)}_{,\tilde{t}} + G^{(1)}, \tag{4.45b}$$

and at second order

$$\Omega^{(0)} q^{(2)}_{,\Psi} - \omega_{,J} J^{(2)} = \tfrac{1}{2}\omega_{,JJ}(J^{(1)})^2 + g^{(1)}_{,q} q^{(1)} + g^{(1)}_{,J} J^{(1)} + g^{(2)} - \Omega^{(2)} q^{(0)}_{,\Psi} - \Omega^{(1)} q^{(1)}_{,\Psi} - q^{(1)}_{,\tilde{t}}, \tag{4.46a}$$

$$\Omega^{(2)} J^{(0)}_{,\Psi} + \Omega^{(0)} J^{(2)}_{,\Psi} = G^{(1)}_{,q} q^{(1)} + G^{(1)}_{,J} J^{(1)} - \Omega^{(1)} J^{(1)}_{,\Psi} - J^{(1)}_{,\tilde{t}} + G^{(2)}. \tag{4.46b}$$

Here it is understood that all functions of q and J are evaluated at $q^{(0)}$ and $J^{(0)}$.

1. Zeroth order analysis

The zeroth order equations (44) can be written more explicitly as

$$\Omega^{(0)}(\tilde{t}) q^{(0)}_{,\Psi}(\Psi, \tilde{t}) = \omega[J^{(0)}(\Psi, \tilde{t}), \tilde{t}], \tag{4.47a}$$

$$\Omega^{(0)}(\tilde{t}) J^{(0)}_{,\Psi}(\Psi, \tilde{t}) = 0. \tag{4.47b}$$

The second of these equations implies that $J^{(0)}$ is independent of Ψ, so we obtain $J^{(0)}(\Psi, \tilde{t}) = \mathcal{J}^{(0)}(\tilde{t})$. The first equation then implies that $q^{(0)}_{,\Psi}$ is independent of Ψ, and integrating with respect to Ψ gives

$$q^{(0)}(\Psi, \tilde{t}) = \frac{\omega[\mathcal{J}^{(0)}(\tilde{t}), \tilde{t}]}{\Omega^{(0)}(\tilde{t})} \Psi + \mathcal{Q}^{(0)}(\tilde{t}), \tag{4.48}$$

where $\mathcal{Q}^{(0)}$ is some function of $\tilde{t}$. The periodicity condition (4.17a) now implies that the coefficient of Ψ in Eq. (4.48) must be unity, which gives the formula (4.28) for the angular velocity $\Omega^{(0)}(\tilde{t})$. Finally, the assumption (4.20) forces $\mathcal{Q}^{(0)}(\tilde{t})$ to vanish, and we recover the formula (4.27) for $q^{(0)}(\Psi, \tilde{t})$.

2. First order analysis

The first order equation (4.45b) can be written more explicitly as

$$\Omega^{(0)}(\tilde{t}) J^{(1)}_{,\Psi}(\Psi, \tilde{t}) = -\mathcal{J}^{(0)}_{,\tilde{t}}(\tilde{t}) + G^{(1)}[\Psi, \mathcal{J}^{(0)}(\tilde{t}), \tilde{t}], \tag{4.49}$$

where we have simplified using the zeroth order solutions (4.25) and (4.27). We now take the average with respect to Ψ of this equation. The left-hand side vanishes since it is a total derivative, and we obtain, using the definition (4.7), the differential equation (4.26) for $\mathcal{J}^{(0)}(\tilde{t})$. Next, we subtract from Eq. (4.49) its averaged part, and use the decomposition (4.41) of $J^{(1)}$. This gives

$$\Omega^{(0)}(\tilde{t}) \hat{J}^{(1)}_{,\Psi}(\Psi, \tilde{t}) = \hat{G}^{(1)}[\Psi, \mathcal{J}^{(0)}(\tilde{t}), \tilde{t}]. \tag{4.50}$$

We solve this equation using the Fourier decomposition (4.11b) of $\hat{G}^{(1)}$ to obtain

$$\hat{J}^{(1)}(\Psi, \tilde{t}) = \sum_{k \neq 0} \frac{G^{(1)}_k[\mathcal{J}^{(0)}(\tilde{t}), \tilde{t}] e^{ik\Psi}}{ik\Omega^{(0)}(\tilde{t})}. \tag{4.51}$$

This yields the first term in the result (4.30) for $J^{(1)}$ when we use the notation (4.13).

Next, we simplify the first order equation (4.45a) using the zeroth order solutions (4.25) and (4.27), to obtain

$$\Omega^{(0)}(\tilde{t}) q^{(1)}_{,\Psi}(\Psi, \tilde{t}) - \omega_{,J}[\mathcal{J}^{(0)}(\tilde{t}), \tilde{t}] J^{(1)}[\Psi, \tilde{t}] = -\Omega^{(1)}(\tilde{t}) + g^{(1)}[\Psi, \mathcal{J}^{(0)}(\tilde{t}), \tilde{t}]. \tag{4.52}$$

Averaging with respect to Ψ and using the decompositions (4.41) and (4.42) of $J^{(1)}$ and $q^{(1)}$ now gives the formula (4.32) for $\Omega^{(1)}(\tilde{t})$. Note, however, that the function $\mathcal{J}^{(1)}(\tilde{t})$ in that formula has not yet been determined; it will be necessary to go to one higher order to compute this function.

Finally, we subtract from Eq. (4.52) its average over Ψ using the decompositions (4.41) and (4.42) and then integrate with respect to Ψ using the notation (4.13). This gives

$$\hat{q}^{(1)}(\Psi, \tilde{t}) = \frac{1}{\Omega^{(0)}(\tilde{t})}\{\omega_{,J}[\mathcal{J}^{(0)}(\tilde{t}), \tilde{t}] \mathcal{I} \hat{J}^{(1)}[\Psi, \tilde{t}] + \mathcal{I} \hat{g}^{(1)}[\Psi, \mathcal{J}^{(0)}(\tilde{t}), \tilde{t}]\}. \tag{4.53}$$

Using the result for $\hat{J}^{(1)}$ given by the first term in Eq. (4.30) now yields the formula (4.34) for $\hat{q}^{(1)}(\Psi, \tilde{t})$, and the result

[22] As is well known, this procedure is valid for asymptotic series as well as normal power series.

(4.33) for $q^{(1)}$ then follows from the decomposition (4.42) together with the initial condition (4.20).

3. Second order analysis

We simplify the second order equation (4.46b) using the zeroth order solutions (4.25) and (4.27), average over Ψ, and simplify using the decompositions (4.41) and (4.42) and the identities (4.12). The result is

$$\mathcal{J}^{(1)}_{,\tilde{t}}(\tilde{t}) = G^{(1)}_{0,J}[\mathcal{J}^{(0)}(\tilde{t}), \tilde{t}]\mathcal{J}^{(1)}(\tilde{t}) + G^{(2)}_{0}[\mathcal{J}^{(0)}(\tilde{t}), \tilde{t}] + \langle \hat{q}^{(1)}(\Psi, \tilde{t})\hat{G}^{(1)}_{,q}[\Psi, \mathcal{J}^{(0)}(\tilde{t}), \tilde{t}]\rangle + \langle \hat{J}^{(1)}(\Psi, \tilde{t})\hat{G}^{(1)}_{,J}[\Psi, \mathcal{J}^{(0)}(\tilde{t}), \tilde{t}]\rangle. \tag{4.54}$$

Using the expressions (4.30) and (4.34) for $\hat{q}^{(1)}$ and $\hat{J}^{(1)}$ and simplifying using the identities (4.14) now gives the differential equation (4.31) for $\mathcal{J}^{(1)}$.

4. Extension to arbitrary order

In this subsection we prove by induction that solutions are uniquely determined at each order in ε. Our inductive hypothesis is that, given the equations up to order s, we can compute all of the expansion coefficients $q^{(u)}(\Psi, \tilde{t})$, $J^{(u)}(\Psi, \tilde{t})$, and $\Omega^{(u)}(\tilde{t})$ for $0 \le u \le s$, except for the averaged piece $\mathcal{J}^{(s)}(\tilde{t})$ of $J^{(s)}(\Psi, \tilde{t})$, and except for $\Omega^{(s)}(\tilde{t})$, which is assumed to be determined by $\mathcal{J}^{(s)}(\tilde{t})$. From the preceding subsections this hypothesis is true for $s = 0$ and for $s = 1$. We shall assume it is true at order $s - 1$ and prove it is true at order s.

The equations of motion at order s, when simplified using the zeroth order solutions (4.25) and (4.27), can be written as

$$\Omega^{(0)}q^{(s)}_{,\Psi} + \Omega^{(s)} - \omega_{,J}J^{(s)} = \omega_{,JJ}J^{(1)}J^{(s-1)} + g^{(1)}_{,q}q^{(s-1)} + g^{(1)}_{,J}J^{(s-1)} - \Omega^{(1)}q^{(s-1)}_{,\Psi} - \Omega^{(s-1)}q^{(1)}_{,\Psi} - q^{(s-1)}_{,\tilde{t}} + \mathcal{S}, \tag{4.55a}$$

$$\Omega^{(0)}J^{(s)}_{,\Psi} = G^{(1)}_{,q}q^{(s-1)} + G^{(1)}_{,J}J^{(s-1)} - \Omega^{(s-1)}J^{(1)}_{,\Psi} - \Omega^{(1)}J^{(s-1)}_{,\Psi} - J^{(s-1)}_{,\tilde{t}} + \mathcal{T}. \tag{4.55b}$$

Here $\mathcal{S} = \mathcal{S}(\Psi, \tilde{t})$ and $\mathcal{T} = \mathcal{T}(\Psi, \tilde{t})$ are expressions involving the forces $G^{(u)}$ and $g^{(u)}$ for $0 \le u \le s$ evaluated at $q = q^{(0)} = \Psi$ and $J = J^{(0)} = \mathcal{J}^{(0)}$, and involving the coefficients $q^{(u)}$, $J^{(u)}$, and $\Omega^{(u)}$ for $0 \le u \le s - 2$, which by the inductive hypothesis are known. Therefore we can treat $\mathcal{S}$ and $\mathcal{T}$ as known functions.

Averaging Eq. (4.55b) over Ψ yields the differential equation

$$\mathcal{J}^{(s-1)}_{,\tilde{t}} - G^{(1)}_{0,J}\mathcal{J}^{(s-1)} = \langle \mathcal{T}\rangle + \langle \hat{G}^{(1)}_{,q}\hat{q}^{(s-1)}\rangle + \langle \hat{G}^{(1)}_{,J}\hat{J}^{(s-1)}\rangle. \tag{4.56}$$

By the inductive hypothesis all the terms on the right-hand side are known, so we can solve this differential equation to determine $\mathcal{J}^{(s-1)}$.

Next, averaging Eq. (4.55a) yields

$$\Omega^{(s)} - \omega_{,J}\mathcal{J}^{(s)} = \omega_{,JJ}\langle \hat{J}^{(1)}\hat{J}^{(s-1)}\rangle + \omega_{,JJ}\mathcal{J}^{(1)}\mathcal{J}^{(s-1)} + \langle \hat{g}^{(1)}_{,q}\hat{q}^{(s-1)}\rangle + \langle \hat{g}^{(1)}_{,J}\hat{J}^{(s-1)}\rangle + g^{(1)}_{0,J}\mathcal{J}^{(s-1)} - \mathcal{Q}^{(s-1)}_{,\tilde{t}} + \langle \mathcal{S}\rangle. \tag{4.57}$$

Since $\mathcal{J}^{(s-1)}$ has already been determined, the right-hand side of this equation is known, and therefore the equation can be used to solve for $\Omega^{(s)}$ once $\mathcal{J}^{(s)}$ is specified, in accord with the inductive hypothesis. Next, Eq. (4.55b) with the average part subtracted can be used to solve for $\hat{J}^{(s)}$, and once $\hat{J}^{(s)}$ is known Eq. (4.55a) with the average part subtracted can be used to solve for $\hat{q}^{(s)}$. Finally, the averaged piece $\mathcal{Q}^{(s)}(\tilde{t})$ of $q^{(s)}(\Psi, \tilde{t})$ can be computed from $\hat{q}^{(s)}$ using the initial condition (4.20) and the decomposition (4.42). Thus the inductive hypothesis is true at order s if it is true at order $s - 1$.

V. SYSTEMS WITH SEVERAL DEGREES OF FREEDOM SUBJECT TO NONRESONANT FORCING

A. Overview

In this section we generalize the analysis of the preceding section to the general system of equations (3.2) with several degrees of freedom. For convenience we reproduce those equations here:

$$\frac{dq_\alpha}{dt} = \omega_\alpha(\mathbf{J}, \tilde{t}) + \varepsilon g^{(1)}_\alpha(\mathbf{q}, \mathbf{J}, \tilde{t}) + \varepsilon^2 g^{(2)}_\alpha(\mathbf{q}, \mathbf{J}, \tilde{t}) + O(\varepsilon^3), \qquad 1 \le \alpha \le N, \tag{5.1a}$$

$$\frac{dJ_\lambda}{dt} = \varepsilon G^{(1)}_\lambda(\mathbf{q}, \mathbf{J}, \tilde{t}) + \varepsilon^2 G^{(2)}_\lambda(\mathbf{q}, \mathbf{J}, \tilde{t}) + O(\varepsilon^3), \qquad 1 \le \lambda \le M. \tag{5.1b}$$

For the remainder of this paper, unless otherwise specified, indices $\alpha, \beta, \gamma, \delta, \varepsilon, \ldots$ from the start of the Greek alphabet will run over $1 \ldots N$, and indices $\lambda, \mu, \nu, \rho, \sigma, \ldots$ from the second half of the alphabet will run over $1 \ldots M$.

The generalization from one to several variables is straightforward except for the treatment of resonances [74]. The key new feature in the N variable case is that the asymptotic expansions now have additional terms pro-

portional to $\sqrt{\varepsilon}$, $\varepsilon^{3/2}, \ldots$ as well as the integer powers of ε. The coefficients of these half-integer powers of ε obey source-free differential equations, except at resonances. Therefore, in the absence of resonances, all of these coefficients can be set to zero without loss of generality. In this paper we develop the general theory with both types of terms present, but we specialize to the case where no resonances occur. Subsequent papers [78,79] will extend the treatment to include resonances, and derive the form of the source terms for the half-integer power coefficients.

We start in Sec. V B by defining our conventions and notations for Fourier decompositions of the perturbing forces. In Sec. V C we discuss the assumptions we make that prevent the occurrence of resonances in the solutions. The heart of the method is the ansatz we make for the form of the solutions, which is given in Sec. V D. Section V E summarizes the results we obtain at each order in the expansion, and the derivations are given in Sec. V F. The implications of the results are discussed in detail in Sec. VII below.

B. Fourier expansions of perturbing forces

The periodicity condition (3.6) applies at each order in the expansion in powers of ε, so we obtain

$$g_\alpha^{(s)}(\mathbf{q}+2\pi\mathbf{k},\mathbf{J},\tilde{t}) = g_\alpha^{(s)}(\mathbf{q},\mathbf{J},\tilde{t}), \tag{5.2a}$$

$$G_\lambda^{(s)}(\mathbf{q}+2\pi\mathbf{k},\mathbf{J},\tilde{t}) = G_\lambda^{(s)}(\mathbf{q},\mathbf{J},\tilde{t}), \tag{5.2b}$$

for $s \geq 1$, $1 \leq \alpha \leq N$, and $1 \leq \lambda \leq M$. Here $\mathbf{k} = (k_1, \ldots, k_N)$ can be an arbitrary N-tuple of integers. It follows from Eqs. (5.2) that these functions can be expanded as multiple Fourier series:

$$g_\alpha^{(s)}(\mathbf{q},\mathbf{J},\tilde{t}) = \sum_{\mathbf{k}} g_{\alpha\mathbf{k}}^{(s)}(\mathbf{J},\tilde{t}) e^{i\mathbf{k}\cdot\mathbf{q}}, \tag{5.3a}$$

$$G_\lambda^{(s)}(\mathbf{q},\mathbf{J},\tilde{t}) = \sum_{\mathbf{k}} G_{\lambda\mathbf{k}}^{(s)}(\mathbf{J},\tilde{t}) e^{i\mathbf{k}\cdot\mathbf{q}}, \tag{5.3b}$$

where

$$g_{\alpha\mathbf{k}}^{(s)}(\mathbf{J},\tilde{t}) = \frac{1}{(2\pi)^N}\int d^N q e^{-i\mathbf{k}\cdot\mathbf{q}} g_\alpha^{(s)}(\mathbf{q},\mathbf{J},\tilde{t}), \tag{5.4a}$$

$$G_{\lambda\mathbf{k}}^{(s)}(\mathbf{J},\tilde{t}) = \frac{1}{(2\pi)^N}\int d^N q e^{-i\mathbf{k}\cdot\mathbf{q}} G_\lambda^{(s)}(\mathbf{q},\mathbf{J},\tilde{t}). \tag{5.4b}$$

Here we adopt the usual notations

$$\sum_{\mathbf{k}} \equiv \sum_{k_1=-\infty}^{\infty} \cdots \sum_{k_N=-\infty}^{\infty}, \tag{5.5}$$

$$\int d^N q \equiv \int_0^{2\pi} dq_1 \ldots \int_0^{2\pi} dq_N, \tag{5.6}$$

and

$$\mathbf{k}\cdot\mathbf{q} \equiv \sum_{\alpha=1}^{N} k_\alpha q_\alpha. \tag{5.7}$$

For any multiply periodic function $f = f(\mathbf{q})$, we introduce the notation

$$\langle f\rangle = \frac{1}{(2\pi)^N}\int d^N q f(\mathbf{q}) \tag{5.8}$$

for the average part of f, and

$$\hat{f}(\mathbf{q}) = f(\mathbf{q}) - \langle f\rangle \tag{5.9}$$

for the remaining part of f. It follows from these definitions that

$$\langle g_\alpha^{(s)}(\mathbf{q},\mathbf{J},\tilde{t})\rangle = g_{\alpha\mathbf{0}}^{(s)}(\mathbf{J},\tilde{t}), \qquad \langle G_\lambda^{(s)}(\mathbf{q},\mathbf{J},\tilde{t})\rangle = G_{\lambda\mathbf{0}}^{(s)}(\mathbf{J},\tilde{t}), \tag{5.10}$$

and that

$$\hat{g}_\alpha^{(s)}(\mathbf{q},\mathbf{J},\tilde{t}) = \sum_{\mathbf{k}\neq\mathbf{0}} g_{\alpha\mathbf{k}}^{(s)}(\mathbf{J},\tilde{t}) e^{i\mathbf{k}\cdot\mathbf{q}}, \tag{5.11a}$$

$$\hat{G}_\lambda^{(s)}(\mathbf{q},\mathbf{J},\tilde{t}) = \sum_{\mathbf{k}\neq\mathbf{0}} G_{\lambda\mathbf{k}}^{(s)}(\mathbf{J},\tilde{t}) e^{i\mathbf{k}\cdot\mathbf{q}}. \tag{5.11b}$$

We also have the identities

$$\left\langle \frac{\partial f}{\partial q_\alpha}\right\rangle = \langle \hat{f}\rangle = 0, \tag{5.12a}$$

$$\langle fg\rangle = \langle \hat{f}\,\hat{g}\rangle + \langle f\rangle\langle g\rangle \tag{5.12b}$$

for any multiply periodic functions $f(\mathbf{q})$, $g(\mathbf{q})$.

For any multiply periodic function f and for any vector $\mathbf{v} = (v_1, \ldots, v_N)$, we also define the quantity $\mathcal{I}_\mathbf{v}\hat{f}$ by

$$(\mathcal{I}_\mathbf{v}\hat{f})(\mathbf{q}) \equiv \sum_{\mathbf{k}\neq\mathbf{0}} \frac{f_\mathbf{k}}{i\mathbf{k}\cdot\mathbf{v}} e^{i\mathbf{k}\cdot\mathbf{q}}, \tag{5.13}$$

where $f_\mathbf{k} = \int d^N q e^{-i\mathbf{k}\cdot\mathbf{q}} f(\mathbf{q})/(2\pi)^N$ are the Fourier coefficients of f. The operator $\mathcal{I}_\mathbf{v}$ satisfies the identities

$$\mathcal{I}_\mathbf{v}(\mathbf{v}\cdot\nabla\hat{f}) = \hat{f}, \tag{5.14a}$$

$$\langle(\mathcal{I}_\mathbf{v}\hat{f})\hat{g}\rangle = -\langle\hat{f}(\mathcal{I}_\mathbf{v}\hat{g})\rangle, \tag{5.14b}$$

$$\langle\hat{f}(\mathcal{I}_\mathbf{v}\hat{f})\rangle = 0. \tag{5.14c}$$

C. The no-resonance assumption

For each set of action variables $\mathbf{J}$ and for each time $\tilde{t}$, we will say that an N-tuple of integers $\mathbf{k} \neq 0$ is a *resonant N-tuple* if

$$\mathbf{k}\cdot\boldsymbol{\omega}(\mathbf{J},\tilde{t}) = 0, \tag{5.15}$$

where $\boldsymbol{\omega} = (\omega_1, \ldots, \omega_N)$ are the frequencies that appear on the right-hand side of the equation of motion (3.2a). This condition governs the occurrence of resonances in our perturbation expansion, as is well known in the context of perturbations of multiply periodic Hamiltonian systems [94]. We will assume that for a given $\mathbf{k}$, the set of values of $\tilde{t}$ at which the quantity

$$\sigma_{\mathbf{k}}(\tilde{t}) \equiv \mathbf{k}\cdot\boldsymbol{\omega}[\mathcal{J}^{(0)}(\tilde{t}),\tilde{t}] \tag{5.16}$$

vanishes (i.e. the resonant values) consists of isolated points. Here $\mathcal{J}^{(0)}(\tilde{t})$ is the leading order solution for $\mathbf{J}$ given by Eq. (5.29) below. This assumption excludes persistent resonances that last for a finite interval in $\tilde{t}$. Generically, we expect this to be true because of the time dependence of $\mathcal{J}^{(0)}(\tilde{t})$.

Our no-resonance assumption is essentially that the Fourier components of the forcing terms vanish for resonant N-tuples. More precisely, for each fixed $\mathbf{k}$ and for each time $\tilde{t}_{\rm r}$ for which $\sigma_{\mathbf{k}}(\tilde{t}_{\rm r})=0$, we assume that

$$g^{(s)}_{\alpha\mathbf{k}}[\mathcal{J}^{(0)}(\tilde{t}),\tilde{t}]=0, \tag{5.17a}$$

$$G^{(s)}_{\lambda\mathbf{k}}[\mathcal{J}^{(0)}(\tilde{t}),\tilde{t}]=0, \tag{5.17b}$$

for $s\geq 1$ and for all $\tilde{t}$ in a neighborhood of $\tilde{t}_{\rm r}$. Our no-resonance assumption will be relaxed in the forthcoming papers [78,79].

In our application to inspirals in Kerr black holes, the no-resonance condition will be automatically satisfied for two classes of orbits: circular and equatorial orbits. This is because for these orbits there is either no radial motion, or no motion in θ, and so the two-dimensional torus (q_r, q_θ) reduces to a one-dimensional circle. The resonance condition $k_r\omega_r + k_\theta\omega_\theta = 0$ reduces to $k_r\omega_r = 0$ for equatorial orbits, or $k_\theta\omega_\theta = 0$ for circular orbits, and these conditions can never be satisfied since the fundamental frequencies ω_r and ω_θ are positive.

D. Two-timescale ansatz for the solution

We now discuss the two-timescale *ansatz* we use for the form of the solutions of the equations of motion. This ansatz will be justified *a posteriori* order by order in $\sqrt{\varepsilon}$. Our ansatz essentially consists of the assumption that the mapping from (t,ε) to $(\mathbf{q},\mathbf{J})$ can be written as an asymptotic expansion in $\sqrt{\varepsilon}$, each term of which is the composition of two maps, the first from (t,ε) to an abstract N-torus with coordinates $\boldsymbol{\Psi}=(\Psi_1,\ldots,\Psi_N)$, and the second from $(\boldsymbol{\Psi},\tilde{t},\varepsilon)$ to $(\mathbf{q},\mathbf{J})$. Here $\tilde{t}=\varepsilon t$ is the slow time parameter. All the fast timescale dynamics is encapsulated in the first mapping. More precisely, we assume

$$\begin{aligned} q_\alpha(t,\varepsilon) &= \sum_{n=0}^{\infty}\varepsilon^{n/2}q^{(n/2)}_\alpha(\boldsymbol{\Psi},\tilde{t}) \\ &= q^{(0)}_\alpha(\boldsymbol{\Psi},\tilde{t}) + \sqrt{\varepsilon}q^{(1/2)}_\alpha(\boldsymbol{\Psi},\tilde{t}) + \varepsilon q^{(1)}_\alpha(\boldsymbol{\Psi},\tilde{t}) \\ &\quad + \varepsilon^{3/2}q^{(3/2)}_\alpha(\boldsymbol{\Psi},\tilde{t}) + O(\varepsilon^2), \end{aligned} \tag{5.18a}$$

$$\begin{aligned} J_\lambda(t,\varepsilon) &= \sum_{n=0}^{\infty}\varepsilon^{n/2}J^{(n/2)}_\lambda(\boldsymbol{\Psi},\tilde{t}) \\ &= J^{(0)}_\lambda(\boldsymbol{\Psi},\tilde{t}) + \sqrt{\varepsilon}J^{(1/2)}_\lambda(\boldsymbol{\Psi},\tilde{t}) + \varepsilon J^{(1)}_\lambda(\boldsymbol{\Psi},\tilde{t}) \\ &\quad + \varepsilon^{3/2}J^{(3/2)}_\lambda(\boldsymbol{\Psi},\tilde{t}) + O(\varepsilon^2). \end{aligned} \tag{5.18b}$$

These asymptotic expansions are assumed to be uniform in $\tilde{t}$. The expansion coefficients $J^{(s)}_\lambda$, where $s=0,1/2,1,\ldots$, are multiply periodic in the phase variables Ψ_α with period 2π in each variable:

$$J^{(s)}_\lambda(\boldsymbol{\Psi}+2\pi\mathbf{k},\tilde{t}) = J^{(s)}_\lambda(\boldsymbol{\Psi},\tilde{t}). \tag{5.19}$$

Here $\mathbf{k}=(k_1,\ldots,k_N)$ is an arbitrary N-tuple of integers. The mapping of the abstract N-torus with coordinates $\boldsymbol{\Psi}$ into the torus in phase space parametrized by $\mathbf{q}$ is assumed to have a trivial wrapping, so that the angle variable q_α increases by 2π when Ψ_α increases by 2π; this implies that the expansion coefficients $q^{(s)}$ satisfy

$$q^{(0)}_\alpha(\boldsymbol{\Psi}+2\pi\mathbf{k},\tilde{t}) = q^{(0)}_\alpha(\boldsymbol{\Psi},\tilde{t}) + 2\pi k_\alpha, \tag{5.20a}$$

$$q^{(s)}_\alpha(\boldsymbol{\Psi}+2\pi\mathbf{k},\tilde{t}) = q^{(s)}_\alpha(\boldsymbol{\Psi},\tilde{t}), \qquad s\geq 1/2, \tag{5.20b}$$

for arbitrary $\mathbf{k}$. The variables $\Psi_1,\ldots,\Psi_N$ are sometimes called fast time parameters.

The angular velocity

$$\Omega_\alpha = d\Psi_\alpha/dt \tag{5.21}$$

associated with the phase Ψ_α is assumed to depend only on the slow time variable $\tilde{t}$ (so it can vary slowly with time), and on ε. We assume that it has an asymptotic expansion in $\sqrt{\varepsilon}$ as $\varepsilon\to 0$ which is uniform in $\tilde{t}$:

$$\Omega_\alpha(\tilde{t},\varepsilon) = \sum_{n=0}^{\infty}\varepsilon^{n/2}\Omega^{(n/2)}_\alpha(\tilde{t}) \tag{5.22}$$

$$\begin{aligned} &= \Omega^{(0)}_\alpha(\tilde{t}) + \sqrt{\varepsilon}\Omega^{(1/2)}_\alpha(\tilde{t}) + \varepsilon\Omega^{(1)}_\alpha(\tilde{t}) + \varepsilon^{3/2}\Omega^{(3/2)}_\alpha(\tilde{t}) \\ &\quad + O(\varepsilon^2). \end{aligned} \tag{5.23}$$

Equations (5.21) and (5.23) serve to define the phase variable Ψ_α in terms of the angular velocity variables $\Omega^{(s)}_\alpha(\tilde{t})$, $s=0,1/2,1\ldots$, up to constants of integration. One constant of integration arises at each order in $\sqrt{\varepsilon}$, for each α. Without loss of generality we choose these constants of integration so that

$$q^{(s)}_\alpha(\mathbf{0},\tilde{t}) = 0 \tag{5.24}$$

for all α, s, and $\tilde{t}$. Note that this does not restrict the final solutions $q_\alpha(t,\varepsilon)$ and $J_\lambda(t,\varepsilon)$, as we show explicitly below, because there are additional constants of integration that arise when solving for the functions $q^{(s)}_\alpha(\boldsymbol{\Psi},\tilde{t})$ and $J^{(s)}_\lambda(\boldsymbol{\Psi},\tilde{t})$.

E. Results of the two-timescale analysis

By substituting the ansatz (5.18b), (5.19), (5.20), (5.21), (5.22), (5.23), and (5.24) into the equations of motion (3.2), we find that all of the assumptions made in the ansatz can be satisfied, and that all of the expansion coefficients are uniquely determined, order by order in $\sqrt{\varepsilon}$. This derivation is given in Sec. V F below. Here we list the results obtained for the various expansion coefficients up to the first three orders.

1. Terminology for various orders of the approximation

We can combine the definitions just summarized to obtain an explicit expansion for the quantity of most interest, the angle variables q_α as a function of time. From the periodicity condition (4.17a) it follows that the function $q_\alpha^{(0)}(\mathbf{\Psi}, \tilde{t})$ can be written as $\Psi_\alpha + \bar{q}_\alpha^{(0)}(\mathbf{\Psi}, \tilde{t})$ where $\bar{q}_\alpha^{(0)}$ is a multiply periodic function of $\mathbf{\Psi}$. From the definitions (3.3) and (5.23), we can write the phase variables Ψ_α as

$$\Psi_\alpha = \frac{1}{\varepsilon}\psi_\alpha^{(0)}(\tilde{t}) + \frac{1}{\sqrt{\varepsilon}}\psi_\alpha^{(1/2)}(\tilde{t}) + \psi_\alpha^{(1)}(\tilde{t}) + \sqrt{\varepsilon}\psi_\alpha^{(3/2)}(\tilde{t}) + \varepsilon\psi_\alpha^{(2)}(\tilde{t}) + O(\varepsilon^{3/2}), \tag{5.25}$$

where the functions $\psi_\alpha^{(s)}(\tilde{t})$ are defined by

$$\psi_\alpha^{(s)}(\tilde{t}) = \int^{\tilde{t}} d\tilde{t}' \Omega_\alpha^{(s)}(\tilde{t}'). \tag{5.26}$$

Inserting this into the expansion (5.18a) of q_α gives

$$q_\alpha(t, \varepsilon) = \frac{1}{\varepsilon}\psi_\alpha^{(0)}(\tilde{t}) + \frac{1}{\sqrt{\varepsilon}}\psi_\alpha^{(1/2)}(\tilde{t}) + [\psi_\alpha^{(1)}(\tilde{t}) + \bar{q}_\alpha^{(0)}(\mathbf{\Psi}, \tilde{t})] + \sqrt{\varepsilon}[\psi_\alpha^{(3/2)}(\tilde{t}) + q_\alpha^{(1/2)}(\mathbf{\Psi}, \tilde{t})] + \varepsilon[\psi_\alpha^{(2)}(\tilde{t}) + q_\alpha^{(1)}(\mathbf{\Psi}, \tilde{t})] + O(\varepsilon^{3/2}). \tag{5.27}$$

We will call the leading order $O(1/\varepsilon)$ term in Eq. (5.27) the *adiabatic approximation*, the subleading $O(1/\sqrt{\varepsilon})$ term the *post-1/2-adiabatic* term, the next $O(1)$ term the *post-1-adiabatic* term, etc. Below we will see that the functions $\bar{q}_\alpha^{(0)}$ and $q_\alpha^{(1/2)}$ in fact vanish identically, and so the oscillatory, $\mathbf{\Psi}$-dependent terms in the expansion (5.27) arise only at post-2-adiabatic and higher orders.

As before we note that the expansion in powers of ε in Eq. (5.27) is *not* a straightforward power series expansion at fixed $\tilde{t}$, since the variable Ψ depends on ε. [The precise definition of the expansion of the solution which we are using is given by Eqs. (5.18a), (5.18b), (5.19), (5.20), (5.21), (5.22), (5.23), and (5.24)]. Nevertheless, the expansion (5.27) as written correctly captures the ε dependence of the secular pieces of the solution, since the functions $\bar{q}^{(0)}$, $q_\alpha^{(1/2)}$, and $q_\alpha^{(1)}$ are multiply periodic functions of Ψ and so have no secular pieces.

2. Adiabatic order

The zeroth order action variables are given by

$$J_\lambda^{(0)}(\mathbf{\Psi}, \tilde{t}) = \mathcal{J}_\lambda^{(0)}(\tilde{t}), \tag{5.28}$$

where $\boldsymbol{\mathcal{J}}^{(0)}(\tilde{t}) = (\mathcal{J}_1^{(0)}(\tilde{t}), \ldots, \mathcal{J}_M^{(0)}(\tilde{t}))$ satisfies the set of coupled ordinary differential equations

$$\frac{d\mathcal{J}_\lambda^{(0)}(\tilde{t})}{d\tilde{t}} = G_{\lambda\mathbf{0}}^{(1)}[\boldsymbol{\mathcal{J}}^{(0)}(\tilde{t}), \tilde{t}]. \tag{5.29}$$

Here the right-hand side denotes the average over $\mathbf{q}$ of the forcing term $G_\lambda^{(1)}[\mathbf{q}, \boldsymbol{\mathcal{J}}^{(0)}(\tilde{t}), \tilde{t}]$; cf. Eqs. (5.4) above. The zeroth order angle variables are given by

$$q_\alpha^{(0)}(\mathbf{\Psi}, \tilde{t}) = \Psi_\alpha, \tag{5.30}$$

and the angular velocity Ω_α that defines the phase variable Ψ_α is given to zeroth order by

$$\Omega_\alpha^{(0)}(\tilde{t}) = \omega_\alpha[\boldsymbol{\mathcal{J}}^{(0)}(\tilde{t}), \tilde{t}]. \tag{5.31}$$

Note that this approximation is equivalent to the following simple prescription: (i) Truncate the equations of motion (5.1) to the $O(\varepsilon)$; (ii) omit the driving terms $g_\alpha^{(1)}$ in the equations for the angle variables; and (iii) replace the driving terms $G_\lambda^{(1)}$ in the equations for the action variables with their averages over $\mathbf{q}$.

3. Post-1/2-adiabatic order

Next, the $O(\sqrt{\varepsilon})$ action variables are given by

$$J_\lambda^{(1/2)}(\mathbf{\Psi}, \tilde{t}) = \mathcal{J}_\lambda^{(1/2)}(\tilde{t}), \tag{5.32}$$

where $\boldsymbol{\mathcal{J}}^{(1/2)}(\tilde{t}) = (\mathcal{J}_1^{(1/2)}(\tilde{t}), \ldots, \mathcal{J}_M^{(1/2)}(\tilde{t}))$ satisfies the set of coupled, source-free ordinary differential equations

$$\frac{d\mathcal{J}_\lambda^{(1/2)}(\tilde{t})}{d\tilde{t}} - \frac{\partial G_{\lambda\mathbf{0}}^{(1)}}{\partial J_\mu}[\boldsymbol{\mathcal{J}}^{(0)}(\tilde{t}), \tilde{t}]\mathcal{J}_\mu^{(1/2)}(\tilde{t}) = 0. \tag{5.33}$$

Equation (5.33) will acquire a source term in Ref. [79] where we include the effects of resonances. The $O(\sqrt{\varepsilon})$ angle variables are given by

$$q_\alpha^{(1/2)}(\mathbf{\Psi}, \tilde{t}) = 0, \tag{5.34}$$

and the angular velocity Ω_α that defines the phase variable Ψ_α is given to $O(\sqrt{\varepsilon})$ by

$$\Omega_\alpha^{(1/2)}(\tilde{t}) = \frac{\partial \omega_\alpha}{\partial J_\lambda}[\boldsymbol{\mathcal{J}}^{(0)}(\tilde{t}), \tilde{t}]\mathcal{J}_\lambda^{(1/2)}(\tilde{t}). \tag{5.35}$$

Note that Eqs. (5.33) and (5.35) can be obtained simply by linearizing Eqs. (5.29) and (5.31) about the zeroth order solution. That is, $\boldsymbol{\mathcal{J}}^{(0)} + \sqrt{\varepsilon}\boldsymbol{\mathcal{J}}^{(1/2)}$ and $\mathbf{\Omega}^{(0)} + \sqrt{\varepsilon}\mathbf{\Omega}^{(1/2)}$ satisfy the zeroth order equations (5.29) and (5.31) to $O(\sqrt{\varepsilon})$. This means that setting $\boldsymbol{\mathcal{J}}^{(1/2)}$ and $\mathbf{\Omega}^{(1/2)}$ to zero does not cause any loss of generality in the solutions (under the no-resonance assumption of this paper), as long as we allow initial conditions to have sufficiently general dependence on ε.

4. Post-1-adiabatic order

The first order action variable is given by

$$J_\lambda^{(1)}(\mathbf{\Psi}, \tilde{t}) = \mathcal{I}_{\mathbf{\Omega}^{(0)}(\tilde{t})}\hat{G}_\lambda^{(1)}[\mathbf{\Psi}, \boldsymbol{\mathcal{J}}^{(0)}(\tilde{t}), \tilde{t}] + \mathcal{J}_\lambda^{(1)}(\tilde{t}), \tag{5.36}$$

where the symbol $\mathcal{I}$ on the right-hand side denotes the integration operator (5.13) with respect to $\mathbf{\Psi}$, $\hat{G}_\lambda^{(1)}$ is the nonconstant piece of $G_\lambda^{(1)}$ as defined in Eq. (5.9), and $\mathbf{\Omega}^{(0)}$ is given by Eq. (5.31). In Eq. (5.36) the quantity $\boldsymbol{\mathcal{J}}^{(1)}(\tilde{t})$ satisfies the differential equation

$$\frac{d\mathcal{J}^{(1)}_\lambda(\tilde{t})}{d\tilde{t}} - \frac{\partial G^{(1)}_{\lambda\mathbf{0}}}{\partial J_\mu}[\mathcal{J}^{(0)}(\tilde{t},\tilde{t}]\mathcal{J}^{(1)}_\mu(\tilde{t}) = G^{(2)}_{\lambda\mathbf{0}} + \frac{1}{2}\frac{\partial^2 G^{(1)}_{\lambda\mathbf{0}}}{\partial J_\mu \partial J_\sigma}\mathcal{J}^{(1/2)}_\mu \mathcal{J}^{(1/2)}_\sigma + \left\langle \frac{\partial \hat{G}^{(1)}_\lambda}{\partial J_\mu} I_{\mathbf{\Omega}^{(0)}} \hat{G}^{(1)}_\mu \right\rangle + \left\langle \frac{\partial \hat{G}^{(1)}_\lambda}{\partial q_\alpha} I_{\mathbf{\Omega}^{(0)}} \hat{g}^{(1)}_\alpha \right\rangle + \frac{\partial \omega_\alpha}{\partial J_\mu}\left\langle \frac{\partial \hat{G}^{(1)}_\lambda}{\partial q_\alpha} I_{\mathbf{\Omega}^{(0)}} I_{(0)} \hat{G}^{(1)}_\mu \right\rangle. \tag{5.37}$$

Here it is understood that the quantities on the right-hand side are evaluated at $\mathbf{J} = \mathcal{J}^{(0)}(\tilde{t})$ and $\mathbf{q} = \mathbf{q}^{(0)} = \mathbf{\Psi}$. The last three terms on the right-hand side of Eq. (5.37) can be written more explicitly using the definition (5.13) of I and the definition (5.8) of the averaging $\langle \ldots \rangle$ as

$$\sum_{\mathbf{k}\neq\mathbf{0}} \frac{1}{\mathbf{\Omega}^{(0)}\cdot\mathbf{k}}\left\{ ik_\alpha \frac{\partial \omega_\alpha}{\partial J_\mu}\frac{G^{(1)*}_{\lambda\mathbf{k}} G^{(1)}_{\mu\mathbf{k}}}{\mathbf{\Omega}^{(0)}\cdot\mathbf{k}} - k_\alpha G^{(1)*}_{\lambda\mathbf{k}} g^{(1)}_{\alpha\mathbf{k}} - iG^{(1)}_{\mu\mathbf{k}}\frac{\partial G^{(1)*}_{\lambda\mathbf{k}}}{\partial J_\mu}\right\}. \tag{5.38}$$

The $O(\varepsilon)$ correction to the angular velocity Ω_α is given by

$$\Omega^{(1)}_\alpha(\tilde{t}) = g^{(1)}_{\alpha\mathbf{0}}[\mathcal{J}^{(0)}(\tilde{t},\tilde{t}] + \frac{\partial \omega_\alpha}{\partial J_\lambda}[\mathcal{J}^{(0)}(\tilde{t},\tilde{t}]\mathcal{J}^{(1)}_\lambda(\tilde{t}) + \frac{1}{2}\frac{\partial^2 \omega_\alpha}{\partial J_\lambda \partial J_\mu}[\mathcal{J}^{(0)}(\tilde{t},\tilde{t}]\mathcal{J}^{(1/2)}_\lambda(\tilde{t})\mathcal{J}^{(1/2)}_\mu(\tilde{t}). \tag{5.39}$$

Finally, the subleading term in the angle variable is

$$q^{(1)}_\alpha(\mathbf{\Psi},\tilde{t}) = \hat{q}^{(1)}_\alpha(\mathbf{\Psi},\tilde{t}) + \mathcal{Q}^{(1)}_\alpha(\tilde{t}), \tag{5.40}$$

where

$$\hat{q}^{(1)}_\alpha(\mathbf{\Psi},\tilde{t}) = \frac{\partial \omega_\alpha}{\partial J_\lambda}[\mathcal{J}^{(0)}(\tilde{t},\tilde{t}]I_{\mathbf{\Omega}^{(0)}(\tilde{t})} I_{\mathbf{\Omega}^{(0)}(\tilde{t})}\hat{G}^{(1)}_\lambda[\mathbf{\Psi},\mathcal{J}^{(0)}(\tilde{t},\tilde{t}] + I_{\mathbf{\Omega}^{(0)}(\tilde{t})}\hat{g}^{(1)}_\alpha[\mathbf{\Psi},\mathcal{J}^{(0)}(\tilde{t},\tilde{t}] \tag{5.41}$$

and

$$\mathcal{Q}^{(1)}_\alpha(\tilde{t}) = -\hat{q}^{(1)}_\alpha(\mathbf{0},\tilde{t}). \tag{5.42}$$

5. Discussion

One of the key results of the general analysis of this section is the identification of which pieces of the external forces are required to compute the adiabatic, post-1/2-adiabatic, and post-1-adiabatic solutions. From Eqs. (5.27), (5.29), and (5.31), the adiabatic solution depends only on the averaged piece $G^{(1)}_{\lambda\mathbf{0}}(\mathbf{J},\tilde{t}) = \langle G^{(1)}_\lambda(\mathbf{q},\mathbf{J},\tilde{t})\rangle$ of the leading order external force $G^{(1)}_\lambda$. Only the dissipative piece of the force $G^{(1)}_\lambda$ normally contributes to this average. For our application to inspirals in Kerr spacetime, this follows from the identity (2.86a) which shows that the average of the conservative piece of $G^{(1)}_\lambda$ vanishes. For a Hamiltonian system with $N = M$, if the perturbing forces g_α and G_β arise from a perturbation $\varepsilon\Delta H = \sum_s \varepsilon^s \Delta H^{(s)}$ to the Hamiltonian, then the forcing function $G^{(s)}_\beta$ is

$$G^{(s)}_\beta(\mathbf{q},\mathbf{J},\tilde{t}) = -\frac{\partial \Delta H^{(s)}(\mathbf{q},\mathbf{J},\tilde{t})}{\partial q_\beta},$$

and it follows that the average over $\mathbf{q}$ of $G^{(s)}_\beta$ vanishes.

At the next, post-1/2-adiabatic order, it follows from Eqs. (5.33) and (5.35) that the term $\psi^{(1/2)}_\alpha(\tilde{t})$ depends again only on the averaged, dissipative piece $G^{(1)}_{\lambda\mathbf{0}}$ of the leading order force. However, we shall see in the forthcoming paper [79] that when the effects of resonances are included, additional dependencies on the remaining (nonaveraged) pieces of the first order self-forces will arise.

At the next, post-1-adiabatic order, the term $\psi^{(1)}_\alpha(\tilde{t})$ in Eq. (5.27) depends on the averaged piece $G^{(2)}_{\lambda\mathbf{0}}(\mathbf{J},\tilde{t}) = \langle G^{(2)}_\lambda(\mathbf{q},\mathbf{J},\tilde{t})\rangle$ of the subleading force $G^{(2)}_\lambda$, again normally purely dissipative, as well as the remaining conservative and dissipative pieces of the leading order forces $G^{(1)}_\lambda(\mathbf{q},\mathbf{J},\tilde{t})$ and $g^{(1)}_\alpha(\mathbf{q},\mathbf{J},\tilde{t})$. This can be seen from Eqs. (5.37) and (5.39). These results have been previously discussed briefly in the EMRI context in Refs. [37,68]. For circular, equatorial orbits, the fact that there is a post-1-adiabatic order contribution from the second order self-force was first argued by Burko [88].

Finally, we consider the choice of initial conditions for the approximate differential equations we have derived. The discussion and conclusions here parallel those in the single variable case, given in Sec. IV D 5 above, and the results are summarized in Sec. VII C below.

F. Derivation

We will denote by $\mathcal{R}(\tilde{t})$ the set of resonant N-tuples $\mathbf{k}$ at time $\tilde{t}$, and by $\mathcal{R}^c(\tilde{t})$ the remaining nonresonant nonzero N-tuples. The set of all N-tuples can therefore be written as the disjoint union

$$\mathbf{Z}^N = \{\mathbf{0}\}\dot{\cup}\mathcal{R}(\tilde{t})\dot{\cup}\mathcal{R}^c(\tilde{t}). \tag{5.43}$$

At each order s we introduce the notation $\mathcal{J}^{(s)}_\lambda(\tilde{t})$ for the average part of $J^{(s)}_\lambda(\mathbf{\Psi},\tilde{t})$:

$$\mathcal{J}^{(s)}_\lambda(\tilde{t}) \equiv \langle J^{(s)}_\lambda(\mathbf{\Psi},\tilde{t})\rangle = \frac{1}{(2\pi)^N}\int_0^{2\pi} d\Psi_1 \ldots \int_0^{2\pi} d\Psi_N J^{(s)}_\lambda(\mathbf{\Psi},\tilde{t}). \tag{5.44}$$

We denote by $\hat{J}^{(s)}_\beta$ the remaining part of $J^{(s)}_\beta$, as in Eq. (5.9). This gives the decomposition

$$J^{(s)}_\lambda(\mathbf{\Psi},\tilde{t}) = \mathcal{J}^{(s)}_\lambda(\tilde{t}) + \hat{J}^{(s)}_\lambda(\mathbf{\Psi},\tilde{t}) \tag{5.45}$$

for all $s \geq 0$. Similarly, for the angle variable we have the decomposition

$$q_\alpha^{(s)}(\mathbf{\Psi},\tilde{t}) = \mathcal{Q}_\alpha^{(s)}(\tilde{t}) + \hat{q}_\alpha^{(s)}(\mathbf{\Psi},\tilde{t}) \tag{5.46}$$

for all $s \geq 1/2$. For the case $s = 0$ we use the fact that $q_\alpha^{(0)}(\mathbf{\Psi},\tilde{t}) - \Psi_\alpha$ is a multiply periodic function of $\mathbf{\Psi}$, by Eq. (5.20a), to obtain the decomposition

$$q_\alpha^{(0)}(\mathbf{\Psi},\tilde{t}) = \Psi_\alpha + \mathcal{Q}_\alpha^{(0)}(\tilde{t}) + \hat{q}_\alpha^{(0)}(\mathbf{\Psi},\tilde{t}), \tag{5.47}$$

where $\hat{q}_\alpha^{(0)}(\mathbf{\Psi},\tilde{t})$ is multiply periodic in $\mathbf{\Psi}$ with zero average.

Using the expansions (5.18a) and (5.18b) of q_α and J_β together with the expansion (5.23) of $d\Psi_\alpha/dt$, we obtain

$$\begin{aligned}\frac{dq_\alpha}{dt} &= \Omega_\beta^{(0)} q^{(0)}_{\alpha,\Psi_\beta} + \sqrt{\varepsilon}[\Omega_\beta^{(1/2)} q^{(0)}_{\alpha,\Psi_\beta} + \Omega_\beta^{(0)} q^{(1/2)}_{\alpha,\Psi_\beta}] + \varepsilon[\Omega_\beta^{(1)} q^{(0)}_{\alpha,\Psi_\beta} + \Omega_\beta^{(1/2)} q^{(1/2)}_{\alpha,\Psi_\beta} + \Omega_\beta^{(0)} q^{(1)}_{\alpha,\Psi_\beta} + q^{(0)}_{\alpha,\tilde{t}}] \\ &\quad + \varepsilon^{3/2}[\Omega_\beta^{(3/2)} q^{(0)}_{\alpha,\Psi_\beta} + \Omega_\beta^{(1)} q^{(1/2)}_{\alpha,\Psi_\beta} + \Omega_\beta^{(1/2)} q^{(1)}_{\alpha,\Psi_\beta} + \Omega_\beta^{(0)} q^{(3/2)}_{\alpha,\Psi_\beta} + q^{(1/2)}_{\alpha,\tilde{t}}] \\ &\quad + \varepsilon^2[\Omega_\beta^{(2)} q^{(0)}_{\alpha,\Psi_\beta} + \Omega_\beta^{(3/2)} q^{(1/2)}_{\alpha,\Psi_\beta} + \Omega_\beta^{(1)} q^{(1)}_{\alpha,\Psi_\beta} + \Omega_\beta^{(1/2)} q^{(3/2)}_{\alpha,\Psi_\beta} + \Omega_\beta^{(0)} q^{(2)}_{\alpha,\Psi_\beta} + q^{(1)}_{\alpha,\tilde{t}}] + O(\varepsilon^{5/2}).\end{aligned} \tag{5.48}$$

We now insert this expansion together with a similar expansion for dJ_λ/dt into the equations of motion (3.2) and use the expansions (3.4) and (3.5) of the external forces g_α and G_λ. Equating coefficients of powers[23] of $\sqrt{\varepsilon}$ then gives at zeroth order

$$\Omega_\beta^{(0)} q^{(0)}_{\alpha,\Psi_\beta} = \omega_\alpha, \tag{5.49a}$$

$$\Omega_\beta^{(0)} J^{(0)}_{\lambda,\Psi_\beta} = 0, \tag{5.49b}$$

at order $O(\sqrt{\varepsilon})$

$$\Omega_\beta^{(0)} q^{(1/2)}_{\alpha,\Psi_\beta} = -\Omega_\beta^{(1/2)} q^{(0)}_{\alpha,\Psi_\beta} + \omega_{\alpha,J_\lambda} J_\lambda^{(1/2)}, \tag{5.50a}$$

$$\Omega_\beta^{(0)} J^{(1/2)}_{\lambda,\Psi_\beta} = -\Omega_\beta^{(1/2)} J^{(0)}_{\lambda,\Psi_\beta}, \tag{5.50b}$$

at order $O(\varepsilon)$

$$\Omega_\beta^{(0)} q^{(1)}_{\alpha,\Psi_\beta} = -\Omega_\beta^{(1/2)} q^{(1/2)}_{\alpha,\Psi_\beta} - \Omega_\beta^{(1)} q^{(0)}_{\alpha,\Psi_\beta} - q^{(0)}_{\alpha,\tilde{t}} + g_\alpha^{(1)} + \omega_{\alpha,J_\lambda} J_\lambda^{(1)} + \tfrac{1}{2}\omega_{\alpha,J_\lambda J_\mu} J_\lambda^{(1/2)} J_\mu^{(1/2)}, \tag{5.51a}$$

$$\Omega_\beta^{(0)} J^{(1)}_{\lambda,\Psi_\beta} = -\Omega_\beta^{(1/2)} J^{(1/2)}_{\lambda,\Psi_\beta} - \Omega_\beta^{(1)} J^{(0)}_{\lambda,\Psi_\beta} - J^{(0)}_{\lambda,\tilde{t}} + G_\lambda^{(1)}, \tag{5.51b}$$

at order $O(\varepsilon^{3/2})$

$$\begin{aligned}\Omega_\beta^{(0)} q^{(3/2)}_{\alpha,\Psi_\beta} &= -\Omega_\beta^{(1/2)} q^{(1)}_{\alpha,\Psi_\beta} - \Omega_\beta^{(1)} q^{(1/2)}_{\alpha,\Psi_\beta} - \Omega_\beta^{(3/2)} q^{(0)}_{\alpha,\Psi_\beta} - q^{(1/2)}_{\alpha,\tilde{t}} + g^{(1)}_{\alpha,q_\beta} q_\beta^{(1/2)} + g^{(1)}_{\alpha,J_\lambda} J_\lambda^{(1/2)} + \omega_{\alpha,J_\lambda} J_\lambda^{(3/2)} \\ &\quad + \omega_{\alpha,J_\lambda J_\mu} J_\lambda^{(1/2)} J_\mu^{(1)} + \tfrac{1}{6}\omega_{\alpha,J_\lambda J_\mu J_\sigma} J_\lambda^{(1/2)} J_\mu^{(1/2)} J_\sigma^{(1/2)},\end{aligned} \tag{5.52a}$$

$$\Omega_\beta^{(0)} J^{(3/2)}_{\lambda,\Psi_\beta} = -\Omega_\beta^{(1/2)} J^{(1)}_{\lambda,\Psi_\beta} - \Omega_\beta^{(1)} J^{(1/2)}_{\lambda,\Psi_\beta} - \Omega_\beta^{(3/2)} J^{(0)}_{\lambda,\Psi_\beta} - J^{(1/2)}_{\lambda,\tilde{t}} + G^{(1)}_{\lambda,q_\beta} q_\beta^{(1/2)} + G^{(1)}_{\lambda,J_\mu} J_\mu^{(1/2)}, \tag{5.52b}$$

and at order $O(\varepsilon^2)$

$$\begin{aligned}\Omega_\beta^{(0)} q^{(2)}_{\alpha,\Psi_\beta} &= -\Omega_\beta^{(1/2)} q^{(3/2)}_{\alpha,\Psi_\beta} - \Omega_\beta^{(1)} q^{(1)}_{\alpha,\Psi_\beta} - \Omega_\beta^{(3/2)} q^{(1/2)}_{\alpha,\Psi_\beta} - \Omega_\beta^{(2)} q^{(0)}_{\alpha,\Psi_\beta} - q^{(1)}_{\alpha,\tilde{t}} + g_\alpha^{(2)} + g^{(1)}_{\alpha,q_\beta} q_\beta^{(1)} + g^{(1)}_{\alpha,J_\lambda} J_\lambda^{(1)} \\ &\quad + \tfrac{1}{2} g^{(1)}_{\alpha,q_\beta q_\gamma} q_\beta^{(1/2)} q_\gamma^{(1/2)} + \tfrac{1}{2} g^{(1)}_{\alpha,J_\lambda J_\mu} J_\lambda^{(1/2)} J_\mu^{(1/2)} + g^{(1)}_{\alpha,q_\beta J_\lambda} q_\beta^{(1/2)} J_\lambda^{(1/2)} + \omega_{\alpha,J_\lambda} J_\lambda^{(2)} + \tfrac{1}{2}\omega_{\alpha,J_\lambda J_\mu J_\sigma} J_\lambda^{(1)} J_\mu^{(1/2)} J_\sigma^{(1/2)} \\ &\quad + \tfrac{1}{2}\omega_{\alpha,J_\lambda J_\mu} J_\lambda^{(1)} J_\mu^{(1)} + \omega_{\alpha,J_\lambda J_\mu} J_\lambda^{(1/2)} J_\mu^{(3/2)} + \tfrac{1}{24}\omega_{\alpha,J_\lambda J_\mu J_\sigma J_\tau} J_\lambda^{(1/2)} J_\mu^{(1/2)} J_\sigma^{(1/2)} J_\tau^{(1/2)},\end{aligned} \tag{5.53a}$$

$$\begin{aligned}\Omega_\beta^{(0)} J^{(2)}_{\lambda,\Psi_\beta} &= -\Omega_\beta^{(1/2)} J^{(3/2)}_{\lambda,\Psi_\beta} - \Omega_\beta^{(1)} J^{(1)}_{\lambda,\Psi_\beta} - \Omega_\beta^{(3/2)} J^{(1/2)}_{\lambda,\Psi_\beta} - \Omega_\beta^{(2)} J^{(0)}_{\lambda,\Psi_\beta} - J^{(1)}_{\lambda,\tilde{t}} + G_\lambda^{(2)} + G^{(1)}_{\lambda,q_\beta} q_\beta^{(1)} + G^{(1)}_{\lambda,J_\mu} J_\mu^{(1)} \\ &\quad + \tfrac{1}{2} G^{(1)}_{\lambda,q_\beta q_\gamma} q_\beta^{(1/2)} q_\gamma^{(1/2)} + \tfrac{1}{2} G^{(1)}_{\lambda,J_\mu J_\sigma} J_\mu^{(1/2)} J_\sigma^{(1/2)} + G^{(1)}_{\lambda,q_\beta J_\mu} q_\beta^{(1/2)} J_\mu^{(1/2)}.\end{aligned} \tag{5.53b}$$

[23] This is justified since both sides are asymptotic expansions in powers of $\sqrt{\varepsilon}$ at fixed $\mathbf{\Psi}$, $\tilde{t}$.

Here it is understood that all functions of $\mathbf{q}$ and $\mathbf{J}$ are evaluated at $\mathbf{q}^{(0)}$ and $\mathbf{J}^{(0)}$.

1. Zeroth order analysis

The zeroth order equations (5.49) can be written more explicitly as

$$\Omega^{(0)}_\beta(\tilde{t}) q^{(0)}_{\alpha,\Psi_\beta}(\mathbf{\Psi},\tilde{t}) = \omega_\alpha[\mathbf{J}^{(0)}(\mathbf{\Psi},\tilde{t}),\tilde{t}], \tag{5.54a}$$

$$\Omega^{(0)}_\beta(\tilde{t}) J^{(0)}_{\lambda,\Psi_\beta}(\mathbf{\Psi},\tilde{t}) = 0. \tag{5.54b}$$

Since $\mathbf{J}^{(0)}$ is a multiply periodic function of $\mathbf{\Psi}$ by Eq. (5.19), we can rewrite Eq. (5.54b) in terms of the Fourier components $J^{(0)}_{\lambda\mathbf{k}}(\tilde{t})$ of $J^{(0)}_\lambda$ as

$$\sum_{\mathbf{k}} [i\mathbf{\Omega}^{(0)}(\tilde{t})\cdot\mathbf{k}] J^{(0)}_{\lambda\mathbf{k}}(\tilde{t}) e^{i\mathbf{k}\cdot\mathbf{\Psi}} = 0. \tag{5.55}$$

For nonresonant N-tuples $\mathbf{k}$ we have

$$\mathbf{\Omega}^{(0)}(\tilde{t})\cdot\mathbf{k} \neq 0 \tag{5.56}$$

by Eqs. (5.15) and (5.31) unless $\mathbf{k} = \mathbf{0}$. This implies that $J^{(0)}_{\lambda\mathbf{k}}(\tilde{t}) = 0$ for all nonzero nonresonant $\mathbf{k}$.

It follows that, for a given $\mathbf{k}$, $J^{(0)}_{\lambda\mathbf{k}}(\tilde{t})$ must vanish except at those values of $\tilde{t}$ at which $\mathbf{k}$ is resonant. Since we assume that $J^{(0)}_{\lambda\mathbf{k}}(\tilde{t})$ is a continuous function of $\tilde{t}$, and since the set of resonant values of $\tilde{t}$ for a given $\mathbf{k}$ consists of isolated points (cf. Sec. V C above), it follows that $J^{(0)}_{\lambda\mathbf{k}}(\tilde{t})$ vanishes for all nonzero $\mathbf{k}$. The formula (5.28) now follows from the decomposition (5.45).

Next, substituting the formula (5.28) for $\mathbf{J}^{(0)}$ and the decomposition (5.47) of $\mathbf{q}^{(0)}$ into Eq. (5.54a) gives

$$\Omega^{(0)}_\alpha(\tilde{t}) + \sum_{\mathbf{k}} [i\mathbf{\Omega}^{(0)}(\tilde{t})\cdot\mathbf{k}] \hat{q}^{(0)}_{\alpha\mathbf{k}}(\tilde{t}) e^{i\mathbf{k}\cdot\mathbf{\Psi}} = \omega_\alpha[\mathcal{J}^{(0)}(\tilde{t}),\tilde{t}], \tag{5.57}$$

where $\hat{q}^{(0)}_{\alpha\mathbf{k}}(\tilde{t})$ are the Fourier components of $\hat{q}^{(0)}_\alpha(\mathbf{\Psi},\tilde{t})$. The $\mathbf{k} = 0$ Fourier component of this equation gives the formula (5.31) for the zeroth order angular velocity $\mathbf{\Omega}^{(0)}$. The $\mathbf{k} \neq 0$ Fourier components imply, using an argument similar to that just given for Eq. (5.54b), that $\hat{q}^{(0)}_{\alpha\mathbf{k}}(\tilde{t}) = 0$ for all nonzero $\mathbf{k}$. The decomposition (5.47) then gives

$$q^{(0)}_\alpha(\mathbf{\Psi},\tilde{t}) = \Psi_\alpha + \mathcal{Q}^{(0)}_\alpha(\tilde{t}). \tag{5.58}$$

Finally, the assumption (5.24) forces $\mathcal{Q}^{(0)}_\alpha(\tilde{t})$ to vanish, and we recover the formula (5.30) for $q^{(0)}_\alpha(\mathbf{\Psi},\tilde{t})$.

2. Order $O(\sqrt{\varepsilon})$ analysis

The $O(\sqrt{\varepsilon})$ equation (5.50b) can be written more explicitly as

$$\Omega^{(0)}_\beta(\tilde{t}) J^{(1/2)}_{\lambda,\Psi_\beta}(\mathbf{\Psi},\tilde{t}) = 0, \tag{5.59}$$

where we have simplified using the zeroth order solution (5.28). An argument similar to that given in Sec. V F 1 now forces the $\mathbf{\Psi}$-dependent piece of $\mathbf{J}^{(1/2)}$ to vanish, and so we obtain the formula (5.32).

Next, we simplify the order $O(\sqrt{\varepsilon})$ equation (5.50a) using the solutions (5.28), (5.30), and (5.32) to obtain

$$\Omega^{(0)}_\beta(\tilde{t}) q^{(1/2)}_{\alpha,\Psi_\beta}(\mathbf{\Psi},\tilde{t}) = \omega_{\alpha,J_\lambda}[\mathcal{J}^{(0)}(\tilde{t}),\tilde{t}]\mathcal{J}^{(1/2)}_\lambda(\tilde{t}) - \Omega^{(1/2)}_\alpha(\tilde{t}). \tag{5.60}$$

After averaging with respect to $\mathbf{\Psi}$, the term on the left-hand side vanishes since it is a total derivative, and we obtain the formula (5.35) for $\Omega^{(1/2)}(\tilde{t})$. Note, however, that the function $\mathcal{J}^{(1/2)}(\tilde{t})$ in that formula has not yet been determined; it will be necessary to go to two higher orders in $\sqrt{\varepsilon}$ to compute this function.

Next, we subtract from Eq. (5.60) its averaged part and use the decomposition (5.46) of $q^{(1/2)}_\alpha$ to obtain

$$\Omega^{(0)}_\beta(\tilde{t}) \hat{q}^{(1/2)}_{\alpha,\Psi_\beta}(\mathbf{\Psi},\tilde{t}) = 0. \tag{5.61}$$

An argument similar to that given in Sec. V F 1 now shows that $\hat{\mathbf{q}}^{(1/2)} = 0$, and the result (5.34) then follows from the decomposition (5.46) together with the initial condition (5.24).

3. Order $O(\varepsilon)$ analysis

The first order equation (5.51b) can be written more explicitly as

$$\Omega^{(0)}_\beta(\tilde{t}) J^{(1)}_{\lambda,\Psi_\beta}(\mathbf{\Psi},\tilde{t}) = -\mathcal{J}^{(0)}_{\lambda,\tilde{t}}(\tilde{t}) + G^{(1)}_\lambda[\mathbf{\Psi},\mathcal{J}^{(0)}(\tilde{t}),\tilde{t}], \tag{5.62}$$

where we have simplified using the zeroth order solutions (5.28) and (5.30) and the $O(\sqrt{\varepsilon})$ solution (5.32). We now take the average with respect to $\mathbf{\Psi}$ of this equation. The left-hand side vanishes since it is a derivative, and we obtain, using the definition (5.4), the differential equation (5.29) for $\mathcal{J}^{(0)}(\tilde{t})$. Next, we subtract from Eq. (5.62) its averaged part, and use the decomposition (5.45) of $\mathbf{J}^{(1)}$. This gives

$$\Omega^{(0)}_\beta(\tilde{t}) \hat{J}^{(1)}_{\lambda,\Psi_\beta}(\mathbf{\Psi},\tilde{t}) = \hat{G}^{(1)}_\lambda[\mathbf{\Psi},\mathcal{J}^{(0)}(\tilde{t}),\tilde{t}]. \tag{5.63}$$

We solve this equation using the Fourier decomposition (5.11b) of $\hat{G}^{(1)}_\lambda$ to obtain

$$\hat{J}^{(1)}_{\lambda}(\mathbf{\Psi}, \tilde{t}) = \sum_{\mathbf{k}\in\mathcal{R}^c(\tilde{t})} \frac{G^{(1)}_{\lambda\mathbf{k}}[\mathcal{J}^{(0)}(\tilde{t}), \tilde{t}]}{i\mathbf{k}\cdot\mathbf{\Omega}^{(0)}(\tilde{t})} e^{i\mathbf{k}\cdot\mathbf{\Psi}} + \sum_{\mathbf{k}\in\mathcal{R}(\tilde{t})} J^{(1)}_{\lambda\mathbf{k}}(\tilde{t}) e^{i\mathbf{k}\cdot\mathbf{\Psi}}. \tag{5.64}$$

Here the first term is a sum over nonresonant N-tuples, and the second term is a sum over resonant N-tuples, for which the coefficients are unconstrained by Eq. (5.63). However, for each fixed $\mathbf{k}$, the values of $\tilde{t}$ that correspond to resonances are isolated, and furthermore, by the no-resonance assumption (5.56) we have $G^{(1)}_{\beta\mathbf{k}}[\mathcal{J}^{(0)}(\tilde{t}), \tilde{t}] = 0$ in the vicinity of those values of $\tilde{t}$. Therefore, using the assumed continuity of $J^{(1)}_{\lambda\mathbf{k}}(\tilde{t})$ in $\tilde{t}$, we can simplify Eq. (5.64) to

$$\hat{J}^{(1)}_{\lambda}(\mathbf{\Psi}, \tilde{t}) = \sum_{\mathbf{k}\neq\mathbf{0}} \frac{G^{(1)}_{\lambda\mathbf{k}}[\mathcal{J}^{(0)}(\tilde{t}), \tilde{t}]}{i\mathbf{k}\cdot\mathbf{\Omega}^{(0)}(\tilde{t})} e^{i\mathbf{k}\cdot\mathbf{\Psi}}, \tag{5.65}$$

where any terms of the form $0/0$ that appear in the coefficients are interpreted to be 0. This yields the first term in the result (5.36) for $\mathbf{J}^{(1)}$ when we use the notation (5.13).

Next, we simplify the first order equation (5.51a) using the zeroth order solutions (5.28) and (5.30) and the $O(\sqrt{\varepsilon})$ solutions (5.32) and (5.34), to obtain

$$\begin{aligned}\Omega^{(0)}_{\beta}(\tilde{t}) q^{(1)}_{\alpha,\Psi_\beta}(\mathbf{\Psi}, \tilde{t}) &= g^{(1)}_{\alpha}[\mathbf{\Psi}, \mathcal{J}^{(0)}(\tilde{t}), \tilde{t}] - \Omega^{(1)}_{\alpha}(\tilde{t}) + \omega_{\alpha,J_\lambda}[\mathcal{J}^{(0)}(\tilde{t}), \tilde{t}] J^{(1)}_{\lambda}[\mathbf{\Psi}, \tilde{t}] \\ &\quad + \tfrac{1}{2}\omega_{\alpha,J_\lambda J_\mu}[\mathcal{J}^{(0)}(\tilde{t}), \tilde{t}] \mathcal{J}^{(1/2)}_{\lambda}(\tilde{t}) \mathcal{J}^{(1/2)}_{\mu}(\tilde{t}).\end{aligned} \tag{5.66}$$

Averaging with respect to $\mathbf{\Psi}$ and using the decompositions (5.45) and (5.46) of $\mathbf{J}^{(1)}$ and $\mathbf{q}^{(1)}$ now gives the formula (5.39) for $\Omega^{(1)}(\tilde{t})$. Note, however, that the function $\mathcal{J}^{(1)}(\tilde{t})$ in that formula has not yet been determined; it will be necessary to go to two higher orders in $\sqrt{\varepsilon}$ to compute this function.

Finally, we subtract from Eq. (5.66) its average over $\mathbf{\Psi}$ using the decompositions (5.45) and (5.46), and then solve the resulting partial differential equation using the notation (5.13) and the convention described after Eq. (5.65). This gives

$$\hat{q}^{(1)}_{\alpha}(\mathbf{\Psi}, \tilde{t}) = \frac{\partial\omega_\alpha}{\partial J_\lambda}[\mathcal{J}^{(0)}(\tilde{t}), \tilde{t}] I_{\mathbf{\Omega}^{(0)}(\tilde{t})} \hat{J}^{(1)}_{\lambda}[\mathbf{\Psi}, \tilde{t}] + I_{\mathbf{\Omega}^{(0)}(\tilde{t})} \hat{g}^{(1)}_{\alpha}[\mathbf{\Psi}, \mathcal{J}^{(0)}(\tilde{t}), \tilde{t}]. \tag{5.67}$$

Using the result for $\hat{J}^{(1)}_{\beta}$ given by the first term in Eq. (5.36) now yields the formula (5.41) for $\hat{q}^{(1)}_{\alpha}(\mathbf{\Psi}, \tilde{t})$, and the result (5.40) for $q^{(1)}_{\alpha}$ then follows from the decomposition (5.46) together with the initial condition (5.24).

4. Order $O(\varepsilon^{3/2})$ analysis

The $O(\varepsilon^{3/2})$ equation (5.52b) can be written more explicitly as

$$\Omega^{(0)}_{\beta}(\tilde{t}) J^{(3/2)}_{\lambda,\Psi_\beta}(\mathbf{\Psi}, \tilde{t}) = -\Omega^{(1/2)}_{\beta}(\tilde{t}) J^{(1)}_{\lambda,\Psi_\beta}(\mathbf{\Psi}, \tilde{t}) - \mathcal{J}^{(1/2)}_{\lambda,\tilde{t}}(\tilde{t}) + G^{(1)}_{\lambda,J_\mu}[\mathbf{\Psi}, \mathcal{J}^{(0)}(\tilde{t}), \tilde{t}] \mathcal{J}^{(1/2)}_{\mu}(\tilde{t}), \tag{5.68}$$

where we have simplified using the lower order solutions (5.28), (5.30), (5.32), and (5.34). We now take the average with respect to $\mathbf{\Psi}$ of this equation. Two terms vanish since they are total derivatives, and we obtain, using the definition (5.4), the differential equation (5.33) for $\mathcal{J}^{(1/2)}(\tilde{t})$. The remaining nonzero Fourier components of Eq. (5.68) can be used to solve for $\hat{\mathbf{J}}^{(3/2)}$, which we will not need in what follows.

Next, we simplify the $O(\varepsilon^{3/2})$ equation (5.52a) using the lower order solutions (5.28), (5.30), (5.32), and (5.34) to obtain

$$\begin{aligned}&\Omega^{(0)}_{\beta}(\tilde{t}) q^{(3/2)}_{\alpha,\Psi_\beta}(\mathbf{\Psi}, \tilde{t}) \\ &= g^{(1)}_{\alpha,J_\lambda}[\mathbf{\Psi}, \mathcal{J}^{(0)}(\tilde{t}), \tilde{t}] \mathcal{J}^{(1/2)}_{\lambda}(\tilde{t}) - \Omega^{(3/2)}_{\alpha}(\tilde{t}) \\ &\quad - \Omega^{(1/2)}_{\beta}(\tilde{t}) q^{(1)}_{\alpha,\Psi_\beta}(\mathbf{\Psi}, \tilde{t}) + \omega_{\alpha,J_\lambda}[\mathcal{J}^{(0)}(\tilde{t}), \tilde{t}] J^{(3/2)}_{\lambda}[\mathbf{\Psi}, \tilde{t}] \\ &\quad + \omega_{\alpha,J_\lambda J_\mu}[\mathcal{J}^{(0)}(\tilde{t}), \tilde{t}] J^{(1)}_{\lambda}[\mathbf{\Psi}, \tilde{t}] \mathcal{J}^{(1/2)}_{\mu}(\tilde{t}) \\ &\quad + \tfrac{1}{2}\omega_{\alpha,J_\lambda J_\mu J_\sigma}[\mathcal{J}^{(0)}(\tilde{t}), \tilde{t}] \mathcal{J}^{(1/2)}_{\lambda}(\tilde{t}) \mathcal{J}^{(1/2)}_{\mu}(\tilde{t}) \mathcal{J}^{(1/2)}_{\sigma}(\tilde{t}).\end{aligned} \tag{5.69}$$

The $\mathbf{k} = 0$ component of this equation yields a formula for $\mathbf{\Omega}^{(3/2)}(\tilde{t})$ in terms of $\mathcal{J}^{(1/2)}(\tilde{t})$ and $\mathcal{J}^{(3/2)}(\tilde{t})$, and the Fourier components with $\mathbf{k} \neq \mathbf{0}$ yield a formula for $\hat{\mathbf{q}}^{(3/2)}$ which we shall not need.

5. Order $O(\varepsilon^2)$ analysis

We simplify the second order equation (5.53b) using the lower order solutions (5.28), (5.30), (5.32), and (5.34), average over $\mathbf{\Psi}$, and simplify using the decompositions (5.45) and (5.46) and the identities (5.12). The result is

$$\begin{aligned}\frac{d}{d\tilde{t}}\mathcal{J}^{(1)}_{\lambda}(\tilde{t}) &= \frac{\partial G^{(1)}_{\lambda\mathbf{0}}}{\partial J_\mu}[\mathcal{J}^{(0)}(\tilde{t}), \tilde{t}] \mathcal{J}^{(1)}_{\mu}(\tilde{t}) + G^{(2)}_{\lambda\mathbf{0}}[\mathcal{J}^{(0)}(\tilde{t}), \tilde{t}] \\ &\quad + \frac{1}{2}\frac{\partial^2 G^{(1)}_{\lambda\mathbf{0}}}{\partial J_\mu \partial J_\sigma}[\mathcal{J}^{(0)}(\tilde{t}), \tilde{t}] \mathcal{J}^{(1/2)}_{\mu}(\tilde{t}) \mathcal{J}^{(1/2)}_{\sigma}(\tilde{t}) \\ &\quad + \left\langle \hat{q}^{(1)}_{\alpha}(\mathbf{\Psi}, \tilde{t}) \frac{\partial \hat{G}^{(1)}_{\lambda}}{\partial \Psi_\alpha}[\mathbf{\Psi}, \mathcal{J}^{(0)}(\tilde{t}), \tilde{t}] \right\rangle \\ &\quad + \left\langle \hat{J}^{(1)}_{\mu}(\mathbf{\Psi}, \tilde{t}) \frac{\partial \hat{G}^{(1)}_{\lambda}}{\partial J_\mu}[\mathbf{\Psi}, \mathcal{J}^{(0)}(\tilde{t}), \tilde{t}] \right\rangle.\end{aligned} \tag{5.70}$$

Using the expressions (5.36) and (5.41) for $\hat{q}^{(1)}_\alpha$ and $\hat{J}^{(1)}_\alpha$ now gives the differential equations (5.37) for $\mathcal{J}^{(1)}$.[24]

VI. NUMERICAL INTEGRATION OF AN ILLUSTRATIVE EXAMPLE

In this section we present a numerical integration of a particular example of a dynamical system, in order to illustrate and validate the general theory of Secs. IV and V.

Consider the system of equations

$$\dot{q} = \omega(J) + \varepsilon g^{(1)}(q, J), \tag{6.1a}$$

$$\dot{J} = \varepsilon G^{(1)}(q, J), \tag{6.1b}$$

where

$$\omega(J) = 1 + J - J^2/4, \qquad g^{(1)}(q, J) = \sin(q)/J,$$
$$G^{(1)}(q, J) = -J - J^2/4 - J\cos(q) - J^2\sin(q), \tag{6.2}$$

together with the initial conditions $q(0) = 1, J(0) = 1$, and with $\varepsilon = 10^{-3}$. The exact solution of this system is shown in Fig. 2.

Consider now the adiabatic approximation to this system. From Eqs. (4.23), (4.24), (4.25), (4.26), (4.27), and (4.28) the adiabatic approximation is given by the system

$$\frac{d\psi^{(0)}}{d\tilde{t}} = \omega(\mathcal{J}^{(0)}), \tag{6.3a}$$

$$\frac{d\mathcal{J}^{(0)}}{d\tilde{t}} = -\mathcal{J}^{(0)} - \mathcal{J}^{(0)2}/4, \tag{6.3b}$$

where $\tilde{t} = \varepsilon t$. The adiabatic solution $(q_{\rm ad}, J_{\rm ad})$ is given in terms of the functions $\psi^{(0)}(\tilde{t})$ and $\mathcal{J}^{(0)}(\tilde{t})$ by

$$q_{\rm ad}(t, \varepsilon) = \varepsilon^{-1}\psi^{(0)}(\varepsilon t), \qquad J_{\rm ad}(t, \varepsilon) = \mathcal{J}^{(0)}(\varepsilon t). \tag{6.4}$$

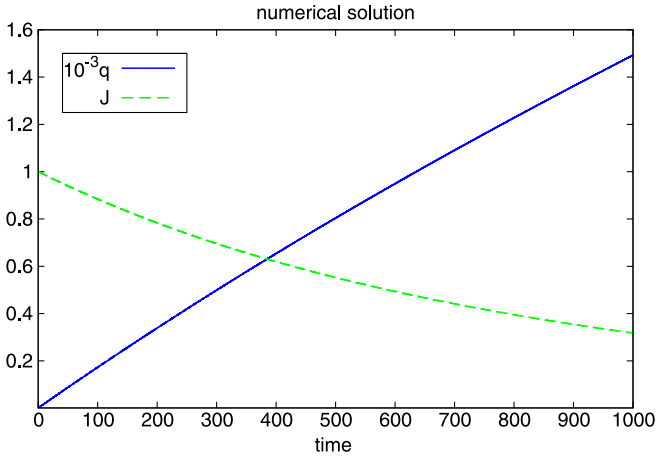

FIG. 2 (color online). The exact numerical solution of the system of equations (6.1). After a time $\sim 1/\varepsilon$, the action variable J is $O(1)$, while the angle variable q is $O(1/\varepsilon)$.

To this order, the initial conditions on $(q_{\rm ad}, J_{\rm ad})$ are the same as those for (q, J), which gives $\psi^{(0)}(0) = \varepsilon$[25] and $\mathcal{J}^{(0)}(0) = 1$. We expect to find that after a time $t \sim 1/\varepsilon$, the errors are of order ~ 1 for $q_{\rm ad}(t)$, and of order $\sim\varepsilon$ for $J_{\rm ad}(t)$. This is confirmed by the two upper panels in Fig. 3, which show the differences $q - q_{\rm ad}$ and $J - J_{\rm ad}$.

Consider next the post-1-adiabatic approximation. From Eqs. (4.31) and (4.32) this approximation is given by the system of equations

$$\frac{d\psi^{(1)}}{d\tilde{t}} = \omega_{,J}(\mathcal{J}^{(0)})\mathcal{J}^{(1)}, \tag{6.5a}$$

$$\frac{d\mathcal{J}^{(1)}}{d\tilde{t}} = -(1 + \mathcal{J}^{(0)}/2)\mathcal{J}^{(1)} + \frac{\mathcal{J}^{(0)}(\mathcal{J}^{(0)} + 1)}{2\omega(\mathcal{J}^{(0)})}, \tag{6.5b}$$

together with the adiabatic system (6.3). From Eqs. (4.24) and (4.30) the post-1-adiabatic solution $(q_{\rm p1a}, J_{\rm p1a})$ is given by

$$q_{\rm p1a}(t, \varepsilon) = \varepsilon^{-1}\psi^{(0)}(\varepsilon t) + \psi^{(1)}(\varepsilon t), \tag{6.6a}$$

$$J_{\rm p1a}(t, \varepsilon) = \mathcal{J}^{(0)}(\varepsilon t) + \varepsilon\mathcal{J}^{(1)}(\varepsilon t) + \varepsilon H[\mathcal{J}^{(0)}(\varepsilon t), q_{\rm p1a}(t, \varepsilon)], \tag{6.6b}$$

where the function H is given by

$$H(\mathcal{J}, q) = \frac{\mathcal{J}^2\cos q - \mathcal{J}\sin q}{\omega(\mathcal{J})}. \tag{6.7}$$

Consider next the choice of initial conditions $\psi^{(0)}(0)$, $\psi^{(1)}(0)$, $\mathcal{J}^{(0)}(0)$, and $\mathcal{J}^{(1)}(0)$ for the system of equa-

[24] We remark that a slight inconsistency arises in our solution ansatz (5.18) at this order, $O(\varepsilon^2)$. Consider the $\mathbf{k} \neq 0$ Fourier components of the second order equations (5.53). For resonant n-tuples $\mathbf{k}$, the left-hand sides of these two equations vanish by definition, but the right-hand sides are generically nonzero, due to the effects of subleading resonances. A similar inconsistency would arise in the $O(\varepsilon)$ equations (5.51), but for the fact that our no-resonance assumption (5.17) forces the right-hand sides of those equations to vanish for resonant n-tuples. However, the no-resonance assumption (5.17) is insufficient to make the right-hand sides of the $O(\varepsilon^2)$ equations (5.53) vanish, because of the occurrence of quadratic cross terms such as

$$g^{(1)}_{\alpha\mathbf{k}} g^{(1)}_{\beta\mathbf{k}'} e^{i(\mathbf{k}+\mathbf{k}')\cdot\boldsymbol{\Psi}}.$$

It can be shown, by an analysis similar to that given in Ref. [79], that the effect of these subleading resonances is to (i) restrict the domain of validity of the expansion (5.18) to exclude times $\tilde{t}$ at which subleading resonances occur, and (ii) add source terms to the differential equation for $\mathcal{J}^{(3/2)}$ which encode the effect of passing through a subleading resonance. These modifications do not affect any of the conclusions in the present paper.

[25] Strictly speaking, our derivations assumed that $\psi^{(0)}(\tilde{t})$ is independent of ε, and so it is inconsistent to use this initial condition for $\psi^{(0)}(0)$. Instead we should set $\psi^{(0)}(0) = 0$, and take account of the nonzero initial phase $q(0)$ at the next order, in the variable $\psi^{(1)}(0)$. However, moving a constant from $\psi^{(1)}(\tilde{t})$ to $\varepsilon^{-1}\psi^{(0)}(\tilde{t})$ does not affect the solution, and so we are free to choose the initial data as done here.

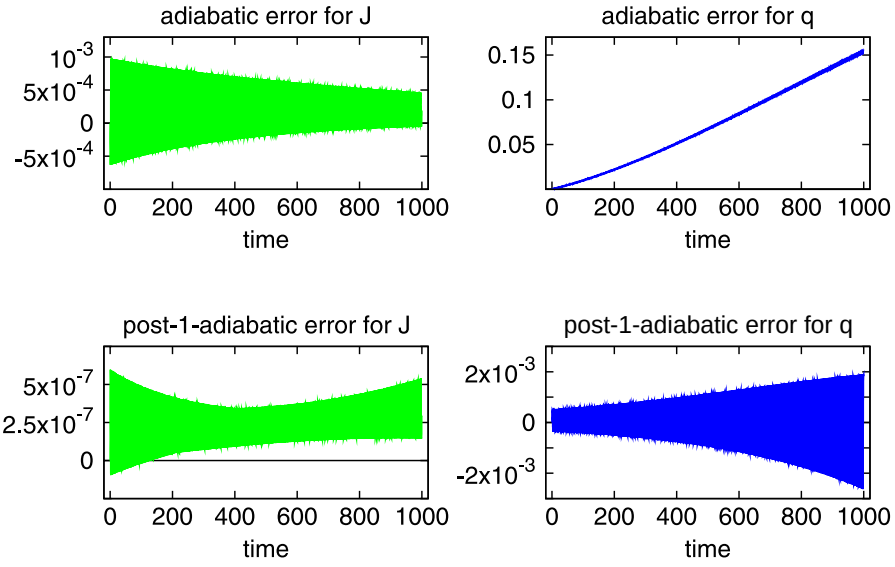

FIG. 3 (color online). *Upper panels*: The difference between the solution of the exact dynamical system (6.1) and the adiabatic approximation given by Eqs. (6.3) and (6.4). For the action variable J, this difference is $O(\varepsilon)$, while for the angle variable q, this difference is $O(1)$, as expected. *Lower panels*: The difference between the exact solution and the post-1-adiabatic approximation given by Eqs. (6.3), (6.5), and (6.6). Again the magnitudes of these errors are as expected: $O(\varepsilon^2)$ for J and $O(\varepsilon)$ for q.

tions (6.3) and (6.5). From Eqs. (6.6) these choices are constrained by, to $O(\varepsilon^2)$,

$$q(0) = \varepsilon^{-1}\psi^{(0)}(0) + \psi^{(1)}(0), \tag{6.8a}$$

$$J(0) = \mathcal{J}^{(0)}(0) + \varepsilon\mathcal{J}^{(1)}(0) + \varepsilon H[J(0), q(0)]. \tag{6.8b}$$

We solve these equations by taking $\psi^{(0)}(0) = 0$, $\psi^{(1)}(0) = q(0) = 1$, $\mathcal{J}^{(0)}(0) = J(0) = 1$, and $\mathcal{J}^{(1)}(0) = -H[J(0), q(0)]$. We expect to find that after a time $t \sim 1/\varepsilon$, the errors are of order $\sim\varepsilon$ for $q_{\rm 1pa}(t)$, and of order $\sim\varepsilon^2$ for $J_{\rm 1pa}(t)$. This is confirmed by the two lower panels in Fig. 3, which show the differences $q - q_{\rm 1pa}$ and $J - J_{\rm 1pa}$.

VII. DISCUSSION

In Sec. II above we derived the set of equations (2.47) describing the radiation-reaction driven inspiral of a particle into a spinning black hole, in terms of generalized action-angle variables. Although those equations contain some functions which are currently unknown, it is possible to give a general analysis of the dependence of the solutions on the mass ratio $\varepsilon = \mu/M$ as $\varepsilon \to 0$, using two-timescale expansions. That analysis was presented in Secs. III, IV, V, and VI above, for the general class of equation systems (3.2) of which the Kerr inspiral example (2.47) is a special case. In this final section we combine these various results and discuss the implications for our understanding of inspirals into black holes.

A. Consistency and uniqueness of approximation scheme

Our analysis has demonstrated that the adiabatic approximation method gives a simple and unique prescription for computing successive approximations to the exact solution, order by order, which is free of ambiguities. In this sense it is similar to the post-Newtonian approximation method.[26] This is shown explicitly in Sec. IV E 4, which shows that the adiabatic method can be extended to all orders for the case of a single degree of freedom, and in Sec. VI, which shows how the method works in practice in a numerical example. In particular, there is no ambiguity in the assignment of initial conditions when computing adiabatic or post-1-adiabatic approximations.

This conclusion appears to be at odds with a recent analysis of Pound and Poisson (PP) [76]. These authors conclude that "An adiabatic approximation to the exact differential equations and initial conditions, designed to capture the secular changes in the orbital elements and to discard the oscillations, would be very difficult to formulate without prior knowledge of the exact solution." The reason for the disagreement is, in part, a matter of terminology: PP's definition of "adiabatic approximation" is different from ours.[27] They take it to mean an approximation which (i) discards all the pieces of the true solutions that vary on the rapid timescale ~ 1 and retains the pieces

[26]The analogy is closer when the two-timescale method is extended to include the field equations and wave generation as well as the inspiral motion [80].

[27]In a later version of their paper they call it instead a "secular approximation."

that vary on the slow timescale $\sim 1/\varepsilon$, and (ii) is globally accurate to some specified order in ε over an inspiral time—throughout their paper they work to the first subleading order, i.e. post-1-adiabatic order. In our terminology, their approximation would consist of the adiabatic approximation, plus the secular piece of the post-1-adiabatic approximation [given by omitting the first term in Eq. (5.36)].

The difference in the terminology used here and in PP is not the only reason for the different conclusions. Our formalism shows that PP's "adiabatic approximation" is actually straightforward to formulate, and that prior knowledge of the exact solution is not required. The reason for the different conclusions is as follows. By "exact solution" PP in fact meant any approximation which includes the rapidly oscillating pieces at post-1-adiabatic order. Their intended meaning was that, since the secular and rapidly oscillating pieces are coupled together at post-1-adiabatic order, any approximation which completely neglects the oscillations cannot be accurate to post-1-adiabatic order [101]. We agree with this conclusion.

On the other hand, we disagree with the overall pessimism of PP's conclusion, because we disagree with their premise. Since the qualitative arguments that were originally presented for the radiative approximation involved discarding oscillatory effects [37,67], PP chose to examine general approximation schemes that neglect oscillatory effects[28] and correctly concluded that such schemes cannot be accurate beyond the leading order. However, our viewpoint is that there is no need to restrict attention to schemes that neglect all oscillatory effects. The two-timescale scheme presented here yields leading order solutions which are not influenced by oscillatory effects, and higher order solutions whose secular pieces are. The development of a systematic approximation scheme that exploits the disparity in orbital and radiation-reaction timescales need not be synonymous with neglecting all oscillatory effects.

B. Effects of conservative and dissipative pieces of the self-force

As we have discussed in Secs. IV D 4 and V E 5 above, our analysis shows rigorously that the dissipative piece of the self-force contributes to the leading order, adiabatic motion, while the conservative piece does not, as argued in Refs. [37,67]. It is possible to understand this fundamental difference in a simple way as follows. We use units where the orbital timescale is ~ 1 and the inspiral timescale is $\sim 1/\varepsilon$. Then the total phase accumulated during the inspiral is $\sim 1/\varepsilon$, and this accumulated phase is driven by the dissipative piece of the self-force.

Consider now the effect of the conservative piece of the self-force. As a helpful thought experiment, imagine setting to zero the dissipative piece of the first order self-force. What then is the effect of the conservative first order self-force on the dynamics? We believe that the perturbed motion is likely to still be integrable; arguments for this will be presented elsewhere [78,79]. However, even if the perturbed motion is not integrable, the Kolmogorov-Arnold-Moser (KAM) theorem [94] implies that the perturbed motion will generically be confined to a torus in phase space for sufficiently small ε. The effect of the conservative self-force is therefore roughly to give an $O(\varepsilon)$ distortion to this torus, and to give $O(\varepsilon)$ corrections to the fundamental frequencies.[29] If one now adds the effects of dissipation, we see that after the inspiral time $\sim 1/\varepsilon$, the corrections due to the conservative force will give a fractional phase correction of order $\sim \varepsilon$, corresponding to a total phase correction ~ 1. This correction therefore comes in at post-1-adiabatic order.

Another way of describing the difference is that the dissipative self-force produces secular changes in the orbital elements, while the conservative self-force does not at the leading order in ε. In Ref. [37] this difference was overstated: it was claimed that the conservative self-force does not produce any secular effects. However, once one goes beyond the leading order, adiabatic approximation, there are in fact conservative secular effects. At post-1-adiabatic order these are described by the first term on the right-hand side of Eq. (5.39). This error was pointed out by Pound and Poisson [76,102].

C. The radiative approximation

So far in this paper we have treated the self-force as fixed, and have focused on how to compute successive approximations to the inspiralling motion. However, as explained in the Introduction, the first order self-force is currently not yet known explicitly. The time-averaged, dissipative[30] piece is known from the work of Mino and others [37,67–70]. The remaining, fluctuating piece of the dissipative first order self-force has not been computed but will be straightforward to compute.[31] The conservative piece of the first order self-force will be much more difficult to compute, and is the subject of much current research [46,49–52].

It is natural therefore to consider the *radiative approximation* obtained by using only the currently available, radiative piece of the first order self-force, as suggested by Mino [67], and by integrating the orbital equations

[28] In the strong sense of neglecting the influence of the oscillatory pieces of the solution on the secular pieces, as well as neglecting the oscillatory pieces themselves.

[29] This corresponds to adding to the frequency ω_α in Eq. (5.1a) the average over $\mathbf{q}$ of the term $\varepsilon g^{(1)}_\alpha$.

[30] We use the terms radiative and dissipative interchangeably; both denote the time-odd piece of the self-force, as defined by Eq. (2.74) above.

[31] For example, by evaluating $J_{\omega lmkn}$ from Eq. (8.21) of Ref. [68] at $\omega = \omega_{mk'n'}$ instead of $\omega = \omega_{mkn}$.

exactly (e.g. numerically). How well will this approximation perform?

From our analysis it follows that the motion as computed in this approximation will agree with the true motion to adiabatic order, and will differ at post-1-adiabatic order. At post-1-adiabatic order, it will omit effects due to the conservative first order force, and also effects due to the dissipative second order self-force. It will include post-1-adiabatic effects due to the fluctuating pieces of the first order, dissipative self-force, and so would be expected to be more accurate than the adiabatic approximation.[32] EMRI waveforms computed using this approximation will likely be the state of the art for quite some time.

Our conclusions about the radiative approximation appear to differ from those of PP [76], who argue that " The radiative approximation does not achieve the goals of an adiabatic approximation." Here, however, the different conclusions arise entirely from a difference in terminology, since PP define "adiabatic approximation" to include slowly varying pieces of the solution to at least post-1-adiabatic order. The radiative approximation does produce solutions that are accurate to adiabatic order, as we have defined it.

We now discuss in more detail the errors that arise in the radiative approximation. These errors occur at post-1-adiabatic order. For discussing these errors, we will neglect post-2-adiabatic effects, and so it is sufficient to use our post-1-adiabatic dynamical equations (5.37) and (5.39). These equations have the structure

$$\mathcal{D}\begin{bmatrix} \psi_\alpha^{(1)}(\tilde{t}) \\ \mathcal{J}_\lambda^{(1)}(\tilde{t}) \end{bmatrix} = \mathcal{S}, \tag{7.1}$$

where $\mathcal{D}$ is a linear differential operator and $\mathcal{S}$ is a source term. The appropriate initial conditions are [see Sec. IV D 5]

$$\psi_\alpha^{(0)} = 0, \qquad \mathcal{J}_\lambda^{(0)}(0) = J_\lambda(0), \tag{7.2a}$$

$$\psi_\alpha^{(1)} = q_\alpha(0), \qquad \mathcal{J}_\lambda^{(1)}(0) = -H_\lambda[\mathbf{q}(0), \mathbf{J}(0)], \tag{7.2b}$$

where $\mathbf{q}(0)$ and $\mathbf{J}(0)$ are the exact initial conditions and the function H_λ is given by, from Eq. (5.36),

$$H_\lambda(\mathbf{q}, \mathbf{J}) = \mathcal{I}_{\mathbf{\Omega}^{(0)}(0)} \hat{G}_\lambda^{(1)}[\mathbf{q}, \mathbf{J}, 0]. \tag{7.3}$$

In terms of these quantities, the radiative approximation is equivalent to making the replacements

$$g_\alpha^{(1)}(\mathbf{q}, \mathbf{J}) \to g_{\alpha\mathrm{diss}}^{(1)}(\mathbf{q}, \mathbf{J}), \tag{7.4a}$$

$$G_i^{(1)}(\mathbf{q}, \mathbf{J}) \to G_{i\mathrm{diss}}^{(1)}(\mathbf{q}, \mathbf{J}), \tag{7.4b}$$

$$G_i^{(2)}(\mathbf{q}, \mathbf{J}) \to 0. \tag{7.4c}$$

These replacements have two effects: (i) They give rise to an error in the source term $\mathcal{S}$ in Eq. (7.1), and (ii) they give rise to an error in the function H_λ and hence in the initial conditions (7.2). There are thus two distinct types of errors that occur in the radiative approximation.[33]

The second type of error could, in principle, be removed by adjusting the initial conditions appropriately. For fixed initial conditions $\mathbf{q}(0)$ and $\mathbf{J}(0)$, such an adjustment would require knowledge of the conservative piece of the self-force, and so is not currently feasible. However, in the context of searches for gravitational wave signals, matched filtering searches will automatically vary over a wide range of initial conditions. Therefore the second type of error will not be an impediment to detecting gravitational wave signals. It will, however, cause errors in parameter extraction.

This fact that the error in the radiative approximation can be reduced by adjusting the initial conditions was discovered by Pound and Poisson [103], who numerically integrated inspirals in Schwarzschild spacetime using post-Newtonian self-force expressions. Their "time-averaged" initial conditions, which they found to give the highest accuracy, correspond to removing the second type of error discussed above, that is, using the initial conditions (7.2) with the exact function H_λ rather than the radiative approximation to H_λ.

Finally, we note that, given the radiative approximation to the self-force, one can compute waveforms using the radiative approximation as described above, or compute waveforms in the adiabatic approximation by solving Eqs. (5.26), (5.29), and (5.31) using the replacement (7.4b). This second option would be easier although somewhat less accurate.

D. Utility of adiabatic approximation for detection of gravitational wave signals

The key motivation for accurate computations of waveforms from inspiral events is of course their use for detecting and analyzing gravitational wave signals. How well will the adiabatic and radiative approximations perform in practice? In this section, we review the studies that have been made of this question. These studies are largely consistent with one another, despite differences in emphasis and interpretation that can be found in the literature. We restrict our attention to inspirals in Schwarzschild spacetime, and to circular or equatorial inspirals in Kerr space-

[32] It is of course possible that, due to an accidental near-cancellation of different post-1-adiabatic terms, the adiabatic approximation may be closer to the true solution than the radiative approximation.

[33] These two errors are both secular, varying on long timescales. There is in addition a rapidly oscillating error caused by the correction to the first term in the expression (5.36) for $J_\lambda^{(1)}$.

time; fully general orbits present additional features that will be discussed elsewhere [78,79].

First, we note that in this paper we have focused on how the post-1-adiabatic error in phase scales with the mass ratio $\varepsilon = \mu/M$. However, one can also ask how the error scales with the post-Newtonian expansion parameter $v/c \sim \sqrt{M/r}$. From Eq. (A10) of Ref. [68] it follows that the post-1-adiabatic phase errors scale as

$$\sim \left(\frac{\mu}{M}\right)^0 \left(\frac{v}{c}\right)^{-3};$$

this scaling is consistent with the more recent analysis of Ref. [103]. This scaling does imply that the error gets larger in the weak field regime, as correctly argued in Ref. [103]. However, it does not necessarily imply large errors in the relativistic regime $v/c \sim 1$ relevant to LISA observations.

The first, order of magnitude estimates of the effects of the conservative piece of the self-force were made by Burko in Refs. [104,105]. References [37,68] computed the post-1-adiabatic phase error within the post-Newtonian approximation for circular orbits, minimized over some of the template parameters, and evaluated at frequencies relevant for LISA. The results indicated a total phase error of order one cycle, not enough to impede detection given that maximum coherent integration times are computationally limited to ~3 weeks [10]. This result was extended to eccentric orbits with eccentricities $\lesssim 0.4$ in Refs. [106,107], with similar results. Similar computations were performed by Burko in Refs. [25,108], although without minimization over template parameters.

These analyses all focused on extreme mass ratio inspirals for LISA. For intermediate-mass-ratio inspirals, potential sources for LIGO, the post-1-adiabatic corrections were studied within the post-Newtonian approximation in Refs. [4,109]. Reference [4] computed fitting factors in addition to phase errors, found that the associated loss of signal-to-noise ratio would be less than 10% in all but the most rapidly spinning cases, and concluded that it would be ''worthwhile but not essential'' to go beyond adiabatic order for detection templates.

The most definitive study to date of post-adiabatic errors for LISA in the case of inspirals in Schwarzschild spacetime was performed by Pound and Poisson (PP1) [103]. PP1 numerically integrated the geodesic equations with post-Newtonian expressions for the self-force, with and without conservative terms. PP1 found large phase errors, $\delta\phi \gtrsim 100$, in the weak field regime. However, the regime relevant to LISA observations is $p \lesssim 30$ [6],[34] where p is the dimensionless semilatus rectum parameter defined by PP1, and PP1's results are focused mostly on values of p larger than this.[35] It is therefore difficult to compare the results of PP1 with earlier estimates or to use them directly to make inferences about signal detection with LISA. PP1's results do show clearly that the errors increase rapidly with increasing eccentricity.

We have repeated PP1's calculations, reproducing the results of their Fig. 6, and extended their calculations to more relativistic systems at lower values of p. More specifically, we performed the following computation: (i) Select values of the mass parameters M and μ, and the initial eccentricity e; (ii) choose the initial value of semilatus rectum p to correspond to one year before the last stable orbit, which occurs on the separatrix $p = 6 + 2e$ [110]; (iii) choose the radiative evolution and the exact evolution to line up at some matching time $t_{\rm m}$ during the last year of inspiral; (iv) start the radiative and exact evolutions with slightly different initial conditions in order that the secular pieces of the evolutions initially coincide—this is the time-averaged initial data prescription of PP1; (v) compute the maximum of the absolute value of the phase error $\delta\phi$ incurred during the last year; (vi) minimize over the matching time $t_{\rm m}$; and (vii) repeat for different values of M, μ, and e. As an example, for $M = 10^6 M_\odot$ and $\mu = 10 M_\odot$, an inspiral starting at $(p, e) = (10.77, 0.300)$ ends up at (6.31, 0.153) after one year. We match the two evolutions at 0.2427 years before plunge, with the exact evolution starting at $(p, e) = (8.819\,33, 0.210\,700)$ and the radiative evolution starting at $(p, e) = (8.819\,28, 0.210\,681)$. The maximum phase error incurred in the last year is then 0.91 cycles.

The phase error incurred during an inspiral from some initial values of e and p to the plunge is independent of the masses M and μ in the small mass ratio limit. However, the phase error incurred during the last year of inspiral is not, since the initial value of p depends on the inspiral timescale $\sim M^2/\mu$. The result is that the phase error depends only on the combination of masses M^2/μ to a good approximation.

Our results are shown in Fig. 4. This figure shows, first, that the computational method of PP1 gives results for low eccentricity systems that are roughly consistent with earlier, cruder estimates, with total phase errors of less than one cycle over most of the parameter space. It also shows that for large eccentricity systems the total phase error can be as large as two or three cycles.

How much will the phase errors shown in Fig. 4 impede the use of the radiative approximation to detect signals? There are two factors which will help. First, Fig. 4 shows

[34]It is true that there will be some binaries visible to LISA at higher values of p that do not merge within the LISA mission lifetime. However, post-Newtonian templates should be sufficient for the detection of these systems.

[35]The second panel of their Fig. 6 does show phase shifts for smaller values of p, but these are all for a mass ratio of $\varepsilon = 0.1$, too large to be a good model of LISA observations; although the phase shift becomes independent of ε as $\varepsilon \to 0$, their Fig. 6 shows that it can vary by factors of up to ~10 as ε varies between 0.1 and 0.001.

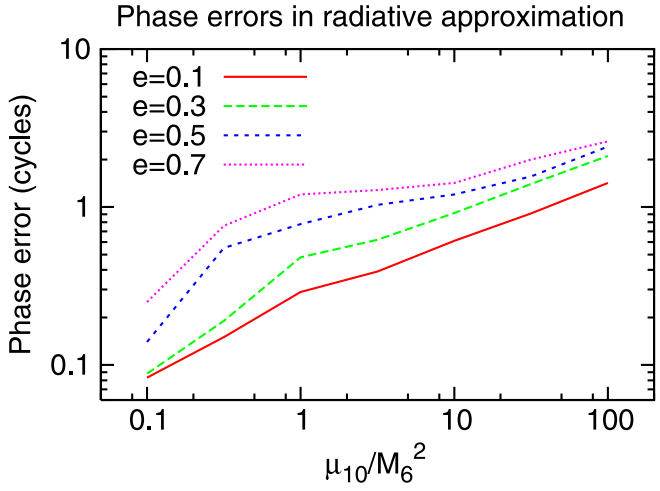

FIG. 4 (color online). The maximum orbital phase error in cycles, $\delta N = \delta\phi/(2\pi)$, incurred in the radiative approximation during the last year of inspiral, as a function of the mass M_6 of the central black hole in units of $10^6 M_\odot$, the mass μ_{10} of the small object in units of $10M_\odot$, and the eccentricity e of the system at the start of the final year of inspiral. The exact and radiative inspirals are chosen to line up at some time $t_{\rm m}$ during the final year, and the value of $t_{\rm m}$ is chosen to minimize the phase error. The initial data at time $t_{\rm m}$ for the radiative evolution are slightly different from those used for the exact evolution, in order that the secular pieces of the two evolutions initially coincide; this is the time-averaged initial data prescription of Pound and Poisson [103]. All evolutions are computed using the hybrid equations of motion of Kidder, Will, and Wiseman [118] in the osculating-element form given by Pound and Poisson.

the maximum phase error during the last year of inspiral, while for detection phase coherence is needed only for periods of $\sim$3 weeks [10]. Second, the matched filtering search process will automatically select parameter values to maximize the overlap between the template and true signal, and parameter mismatches will therefore be likely to reduce the effect of the phase error.[36] On the other hand, for large eccentricities, the phase error $\delta\phi(t)$ is typically a rapidly oscillating function, rather than a smooth function, which may counteract the helpful effects of smaller time windows or parameter mismatches. Also, we note that a sign flip will occur in the integrand of an overlap integral once the gravitational wave phase error $2\delta\phi$ exceeds π, corresponding to the number of cycles plotted in Fig. 4 exceeding $1/4$. This occurs in a large part of the parameter space.

[36]We note that there are already two minimizations over parameters included in the phase errors shown in Fig. 4: a minimization over $t_{\rm m}$ as discussed above, and the replacement $m_1 \to m_1 + m_2$ used by PP1 in the derivation of their self-force expressions in order to eliminate the leading order piece of the self-force.

Thus, there is a considerable amount of uncertainty as to whether the radiative approximation will be sufficiently accurate for signal detection. A detailed study would require computation of fitting factors and optimizing over all template parameters, and modeling the hierarchical detection algorithm discussed in Ref. [10]. Such a study is beyond the scope of this paper. Based on the results shown in Fig. 4, we agree with the conclusions of PP1 that the early estimates based on circular orbits [37,68] were too optimistic, and that it is not clear that the radiative approximation is sufficiently accurate. (Moreover, parameter extraction will clearly require going beyond the radiative approximation.)

For gravitational wave searches, it might therefore be advisable to use *hybrid* waveforms, computed using the fully relativistic dissipative piece of the self-force, and using post-Newtonian expressions for the conservative piece. Although the post-Newtonian expressions are not expected to be very accurate in the relativistic regime, improved versions have been obtained recently based on comparisons between post-Newtonian and fully numerical waveforms from binary black hole mergers; see, for example, the effective one body approximation of Refs. [111–116]. It seems likely that hybrid EMRI waveforms incorporating such improved post-Newtonian expressions for the conservative self-force will be more accurate than radiative waveforms. Hybrid waveforms may be the best that can be done until the fully relativistic conservative self-force is computed.

VIII. CONCLUSIONS

In this paper we have developed a systematic two-timescale approximation method for computing the inspirals of particles into spinning black holes. Future papers in this series will deal with the effects of transient resonances [78,79], and will give more details of the two-timescale expansion of the Einstein equations [80] that meshes consistently with the approximation method for the orbital motion discussed here.

ACKNOWLEDGMENTS

We thank Steve Drasco, Marc Favata, John Friedman, Scott Hughes, Yasushi Mino, Eric Poisson, Adam Pound, and Eran Rosenthal for helpful conversations. This research was supported in part by NSF Grant No. PHY-0457200 and NASA Grant No. NAGW-12906..

APPENDIX A: EXPLICIT EXPRESSIONS FOR THE COEFFICIENTS IN THE ACTION-ANGLE EQUATIONS OF MOTION

From the formulas (2.27) for the action variables together with the definitions (2.26) of the potentials V_r and V_θ, we can compute the partial derivatives $\partial J_\alpha/\partial P_\beta$. The nontrivial derivatives are

$$\frac{\partial J_r}{\partial H} = \frac{Y}{\pi}, \tag{A1a}$$

$$\frac{\partial J_r}{\partial E} = \frac{W}{\pi}, \tag{A1b}$$

$$\frac{\partial J_r}{\partial L_z} = -\frac{Z}{\pi}, \tag{A1c}$$

$$\frac{\partial J_r}{\partial Q} = -\frac{X}{2\pi}, \tag{A1d}$$

$$\frac{\partial J_\theta}{\partial H} = \frac{2\sqrt{z_+}a^2}{\pi\beta}[K(k) - E(k)], \tag{A1e}$$

$$\frac{\partial J_\theta}{\partial E} = \frac{2\sqrt{z_+}Ea^2}{\pi\beta}[K(k) - E(k)], \tag{A1f}$$

$$\frac{\partial J_\theta}{\partial L_z} = \frac{2L_z}{\pi\beta\sqrt{z_+}}[K(k) - \Pi(\pi/2, z_-, k)], \tag{A1g}$$

$$\frac{\partial J_\theta}{\partial Q} = \frac{1}{\pi\beta\sqrt{z_+}}K(k). \tag{A1h}$$

Here the quantities W, X, Y, and Z are the radial integrals defined by Schmidt[37] as [92]

$$W = \int_{r_1}^{r_2} \frac{r^2E(r^2 + a^2) - 2Mra(L_z - aE)}{\Delta\sqrt{V_r}}dr, \tag{A2a}$$

$$X = \int_{r_1}^{r_2} \frac{dr}{\sqrt{V_r}}, \tag{A2b}$$

$$Y = \int_{r_1}^{r_2} \frac{r^2}{\sqrt{V_r}}dr, \tag{A2c}$$

$$Z = \int_{r_1}^{r_2} \frac{r[L_z r - 2M(L_z - aE)]}{\Delta\sqrt{V_r}}dr, \tag{A2d}$$

where r_1 and r_2 are the turning points of the radial motion, i.e. the two largest roots of $V_r(r) = 0$. In these equations $K(k)$ is the complete elliptic integral of the first kind, $E(k)$ is the complete elliptic integral of the second kind, and $\Pi(\phi, n, k)$ is the Legendre elliptic integral of the third kind [117]:

$$K(k) = \int_0^{\pi/2} \frac{d\theta}{\sqrt{1 - k^2\sin^2\theta}}, \tag{A3}$$

$$E(k) = \int_0^{\pi/2} d\theta\sqrt{1 - k^2\sin^2\theta}, \tag{A4}$$

$$\Pi(\phi, n, k) = \int_0^{\phi} \frac{d\theta}{(1 - n\sin^2\theta)\sqrt{1 - k^2\sin^2\theta}}. \tag{A5}$$

Also, we have defined $\beta^2 = a^2(\mu^2 - E^2)$ and $k = \sqrt{z_-/z_+}$, where $z = \cos^2\theta$[38] and z_- and z_+ are the two roots of $V_\theta(z) = 0$ with $0 < z_- < 1 < z_+$.

Combining the derivatives (A1) with the chain rule in the form

$$\frac{\partial P_\alpha}{\partial J_\beta}\frac{\partial J_\beta}{\partial P_\gamma} = \delta^\alpha_\gamma \tag{A6}$$

yields the following expression for the frequencies (2.14) as functions of P_α:

$$\Omega_t = \frac{K(k)W + a^2z_+E[K(k) - E(k)]X}{K(k)Y + a^2z_+[K(k) - E(k)]X}, \tag{A7a}$$

$$\Omega_r = \frac{\pi K(k)}{K(k)Y + a^2z_+[K(k) - E(k)]X}, \tag{A7b}$$

$$\Omega_\theta = \frac{\pi\beta\sqrt{z_+}X/2}{K(k)Y + a^2z_+[K(k) - E(k)]X}, \tag{A7c}$$

$$\Omega_\phi = \frac{K(k)Z + L_z[\Pi(\pi/2, z_-, k) - K(k)]X}{K(k)Y + a^2z_+[K(k) - E(k)]X}. \tag{A7d}$$

APPENDIX B: COMPARISON WITH TREATMENT OF KEVORKIAN AND COLE

As explained in Sec. III above, our two-timescale analysis of the general system of equations (3.2) closely follows that of the textbook [74] by Kevorkian and Cole (KC), which is a standard reference on asymptotic methods. In this appendix we explain the minor ways in which our treatment in Secs. IV and V extends and corrects that of KC. Section 4.4 of KC covers the one variable case. We simplify this treatment by using action-angle variables, and also extend it by showing that the method works to all orders in ε. Our general system of equations (3.2) is studied by KC in their Sec. 4.5. We generalize this analysis by including the half-integer powers of ε, which are required for the treatment of resonances. A minor correction is that their solution (4.5.54a) is not generally valid, since it requires Ω_i and τ_i to be collinear, which will not always be the case. However, it is easy to repair this error by using the continuity argument given in Sec. V F 1 above, which in effect replaces the expression (4.5.54a) with one constructed using Fourier methods; cf. Eq. (5.65) above. This argument does require a more precise version of the no-resonance assumption than given by KC. Finally, our treatment of resonances [78,79] will closely follow KC's Sec. 5.4, except that our analysis will apply to the general system (3.2), generalizing KC's treatment of special cases.

[37]There is a typo in the definition of W given in Eq. (44) of Schmidt [92].

[38]Here we follow Drasco and Hughes [38] rather than Schmidt who defines $z = \cos\theta$.

[1] L. Blanchet, Living Rev. Relativity **5**, 3 (2002).
[2] F. Pretorius, arXiv:0710.1338.
[3] S. A. Teukolsky, Astrophys. J. **185**, 635 (1973).
[4] D. A. Brown *et al.*, Phys. Rev. Lett. **99**, 201102 (2007).
[5] I. Mandel, D. A. Brown, J. R. Gair, and M. C. Miller, arXiv:0705.0285.
[6] L. S. Finn and K. S. Thorne, Phys. Rev. D **62**, 124021 (2000).
[7] C. Cutler and K. S. Thorne, in *Proceedings of General Relativity and Gravitation XVI*, edited by N. T. Bishop and S. D. Maharaj (World Scientific, Singapore, 2002).
[8] P. Amaro-Seoane *et al.*, Classical Quantum Gravity **24**, R113 (2007).
[9] L. Barack *et al.*, Estimates of Detection Rates for LISA Capture Sources, http://www.tapir.caltech.edu/listwg1/EMRI/LISTEMRIreport.pdf.
[10] J. R. Gair *et al.*, Classical Quantum Gravity **21**, S1595 (2004).
[11] C. Hopman and T. Alexander, Astrophys. J. Lett. **645**, L133 (2006).
[12] F. D. Ryan, Phys. Rev. D **52**, 5707 (1995).
[13] C. Li and G. Lovelace, Phys. Rev. D **77**, 064022 (2008).
[14] F. D. Ryan, Phys. Rev. D **56**, 1845 (1997).
[15] L. Barack and C. Cutler, Phys. Rev. D **75**, 042003 (2007).
[16] M. Colpi, S. L. Shapiro, and I. Wasserman, Phys. Rev. Lett. **57**, 2485 (1986).
[17] F. D. Ryan, Phys. Rev. D **55**, 6081 (1997).
[18] E. Poisson, Phys. Rev. D **54**, 5939 (1996).
[19] L. Barack and C. Cutler, Phys. Rev. D **69**, 082005 (2004).
[20] S. A. Hughes and R. D. Blandford, Astrophys. J. **585**, L101 (2003).
[21] C. L. MacLeod and C. J. Hogan, Phys. Rev. D **77**, 043512 (2008).
[22] W. Hu, ASP Conf. Ser. **339**, 215 (2005).
[23] Y. Rathore, R. D. Blandford, and A. E. Broderick, Mon. Not. R. Astron. Soc. **357**, 834 (2005).
[24] L. M. Burko, Phys. Rev. D **69**, 044011 (2004).
[25] L. M. Burko, Classical Quantum Gravity **23**, 4281 (2006).
[26] L. Barack and C. Cutler, Phys. Rev. D **69**, 082005 (2004).
[27] A. Ori and K. S. Thorne, Phys. Rev. D **62**, 124022 (2000).
[28] A. Buonanno and T. Damour, Phys. Rev. D **62**, 064015 (2000).
[29] R. O'Shaughnessy, Phys. Rev. D **67**, 044004 (2003).
[30] P. A. Sundararajan, Phys. Rev. D **77**, 124050 (2008).
[31] K. Glampedakis, S. A. Hughes, and D. Kennefick, Phys. Rev. D **66**, 064005 (2002).
[32] S. Babak *et al.*, Phys. Rev. D **75**, 024005 (2007).
[33] C. Cutler, D. Kennefick, and E. Poisson, Phys. Rev. D **50**, 3816 (1994).
[34] M. Shibata, Phys. Rev. D **50**, 6297 (1994).
[35] K. Glampedakis and D. Kennefick, Phys. Rev. D **66**, 044002 (2002).
[36] S. A. Hughes, Phys. Rev. D **61**, 084004 (2000).
[37] S. A. Hughes, S. Drasco, E. E. Flanagan, and J. Franklin, Phys. Rev. Lett. **94**, 221101 (2005).
[38] S. Drasco and S. A. Hughes, Phys. Rev. D **73**, 024027 (2006).
[39] S. Drasco, arXiv:0711.4644.
[40] S. Drasco, Classical Quantum Gravity **23**, S769 (2006).
[41] M. Campanelli and C. O. Lousto, Phys. Rev. D **59**, 124022 (1999).
[42] R. Gleiser, C. Nicasio, R. Price, and J. Pullin, Phys. Rep. **325**, 41 (2000).
[43] C. O. Lousto and H. Nakano, arXiv:0804.3824.
[44] Y. Mino, M. Sasaki, and T. Tanaka, Phys. Rev. D **55**, 3457 (1997).
[45] T. C. Quinn and R. M. Wald, Phys. Rev. D **56**, 3381 (1997).
[46] E. Poisson, Living Rev. Relativity **7**, 6 (2004).
[47] L. Barack and A. Ori, Phys. Rev. D **61**, 061502 (2000).
[48] L. Barack and C. O. Lousto, Phys. Rev. D **72**, 104026 (2005).
[49] L. Barack and A. Ori, Phys. Rev. Lett. **90**, 111101 (2003).
[50] S. E. Gralla, J. L. Friedman, and A. G. Wiseman, arXiv:gr-qc/0502123.
[51] T. S. Keidl, J. L. Friedman, and A. G. Wiseman, Phys. Rev. D **75**, 124009 (2007).
[52] L. Barack and N. Sago, Phys. Rev. D **75**, 064021 (2007).
[53] I. Vega and S. Detweiler, Phys. Rev. D **77**, 084008 (2008).
[54] L. Barack, D. A. Golbourn, and N. Sago, Phys. Rev. D **76**, 124036 (2007).
[55] E. E. Flanagan, Workshop on Radiation Reaction in General Relativity, Penn State (2002).
[56] Y. Mino, Prog. Theor. Phys. **115**, 43 (2006).
[57] W. Krivan, P. Laguna, P. Papadopoulos, and N. Andersson, Phys. Rev. D **56**, 3395 (1997).
[58] L. M. Burko and G. Khanna, Phys. Rev. D **67**, 081502 (2003).
[59] K. Martel, Ph.D. thesis, University of Guelph, Guelph, Ontario, Canada, 2003.
[60] R. Lopez-Aleman, G. Khanna, and J. Pullin, Classical Quantum Gravity **20**, 3259 (2003).
[61] G. Khanna, Phys. Rev. D **69**, 024016 (2004).
[62] E. Pazos-Avalos and C. O. Lousto, Phys. Rev. D **72**, 084022 (2005).
[63] C. F. Sopuerta, P. Sun, P. Laguna, and J. Xu, Classical Quantum Gravity **23**, 251 (2006).
[64] C. F. Sopuerta and P. Laguna, Phys. Rev. D **73**, 044028 (2006).
[65] P. A. Sundararajan, G. Khanna, and S. A. Hughes, Phys. Rev. D **76**, 104005 (2007).
[66] J. L. Barton *et al.*, arXiv:0804.1075.
[67] Y. Mino, Phys. Rev. D **67**, 084027 (2003).
[68] S. Drasco, E. E. Flanagan, and S. A. Hughes, Classical Quantum Gravity **22**, S801 (2005).
[69] N. Sago, T. Tanaka, W. Hikida, and H. Nakano, Prog. Theor. Phys. **114**, 509 (2005).
[70] N. Sago *et al.*, Prog. Theor. Phys. **115**, 873 (2006).
[71] S. Drasco and N. Sago (unpublished).
[72] Y. Mino, Prog. Theor. Phys. **113**, 733 (2005).
[73] P. A. Sundararajan, G. Khanna, S. A. Hughes, and S. Drasco, Phys. Rev. D **78**, 024022 (2008).
[74] J. Kevorkian and J. D. Cole, *Multiple Scale and Singular Perturbation Methods* (Springer, New York, 1996).
[75] T. Hinderer and E. Flanagan, APS Meeting Abstracts 11005 (2006).
[76] A. Pound and E. Poisson, Phys. Rev. D **77**, 044012 (2008).
[77] Y. Mino and R. Price, Phys. Rev. D **77**, 064001 (2008).
[78] E. E. Flanagan and T. Hinderer, "Transient resonances in the inspirals of point particles into black holes" (unpublished).

[79] E. E. Flanagan and T. Hinderer, ''Two-timescale analysis of extreme mass ratio inspirals in Kerr spacetime. II. Transient resonances'' (unpublished).
[80] E. E. Flanagan and T. Hinderer, ''Two-timescale analysis of extreme mass ratio inspirals in Kerr spacetime. III. Wave generation'' (unpublished).
[81] Y. Mino, Classical Quantum Gravity **22**, S717 (2005).
[82] Y. Mino, Classical Quantum Gravity **22**, S375 (2005).
[83] E. Rosenthal, Phys. Rev. D **73**, 044034 (2006).
[84] E. Rosenthal, Phys. Rev. D **74**, 084018 (2006).
[85] C. R. Galley, Ph.D. thesis, University of Maryland, 2007.
[86] C. R. Galley and B. L. Hu, arXiv:0801.0900.
[87] C. R. Galley, 11th Eastern Gravity Meeting, Penn State (2008).
[88] L. M. Burko, Phys. Rev. D **67**, 084001 (2003).
[89] T. Tanaka, Prog. Theor. Phys. Suppl. **163**, 120 (2006).
[90] S. Detweiler, Phys. Rev. Lett. **86**, 1931 (2001).
[91] R. M. Wald, *General Relativity* (University of Chicago Press, Chicago, 1984), p. 504.
[92] W. Schmidt, Classical Quantum Gravity **19**, 2743 (2002).
[93] K. Glampedakis and S. Babak, Classical Quantum Gravity **23**, 4167 (2006).
[94] V. Arnold, *Mathematical Methods of Classical Mechanics*, Graduate Texts in Mathematics Vol. 60 (Springer, Berlin, Germany; New York, USA, 1995), 2nd ed.
[95] E. Fiorani, G. Giachetta, and G. Sardanashvily, J. Phys. A **36**, L101 (2003).
[96] M. Antonowicz, J. Phys. A **14**, 1099 (1981).
[97] B. Carter, Phys. Rev. **174**, 1559 (1968).
[98] L. Barack and A. Ori, Phys. Rev. D **64**, 124003 (2001).
[99] E. Poisson, Phys. Rev. D **70**, 084044 (2004).
[100] C. W. Misner, K. S. Thorne, and J. A. Wheeler, *Gravitation* (W. H. Freeman and Co., San Francisco, 1973).
[101] A. Pound and E. Poisson (private communication).
[102] A. Pound, E. Poisson, and B. G. Nickel, Phys. Rev. D **72**, 124001 (2005).
[103] A. Pound and E. Poisson, Phys. Rev. D **77**, 044013 (2008).
[104] L. M. Burko, Int. J. Mod. Phys. A **16**, 1471 (2001).
[105] L. M. Burko, Phys. Rev. D **67**, 084001 (2003).
[106] M. Favata, Ph.D. thesis, Cornell, 2006.
[107] M. Favata and E. Flanagan, APS Meeting Abstracts 11006 (2006).
[108] L. M. Burko, in *Laser Interferometer Space Antenna: 6th International LISA Symposium*, edited by S. M. Merkovitz and J. C. Livas, AIP Conf. Proc. No. 873 (AIP, New York, 2006), pp. 269–273.
[109] P. Ajith, B. R. Iyer, C. A. K. Robinson, and B. S. Sathyaprakash, Classical Quantum Gravity **22**, S1179 (2005).
[110] C. Cutler, D. Kennefick, and E. Poisson, Phys. Rev. D **50**, 3816 (1994).
[111] A. Buonanno and T. Damour, Phys. Rev. D **59**, 084006 (1999).
[112] T. Damour, Phys. Rev. D **64**, 124013 (2001).
[113] A. Buonanno, Classical Quantum Gravity **19**, 1267 (2002).
[114] T. Damour, B. R. Iyer, P. Jaranowski, and B. S. Sathyaprakash, Phys. Rev. D **67**, 064028 (2003).
[115] T. Damour and A. Nagar, Phys. Rev. D **76**, 064028 (2007).
[116] A. Buonanno *et al.*, Phys. Rev. D **76**, 104049 (2007).
[117] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in C. The Art of Scientific Computing* (Cambridge University Press, Cambridge, 1992), 2nd ed.
[118] L. E. Kidder, C. M. Will, and A. G. Wiseman, Phys. Rev. D **47**, 3281 (1993).