Conditions for sustained orbital resonances in extreme mass ratio inspirals

Maarten van de Meent

Mathematical Sciences, University of Southampton, Southampton, SO17 1BJ, United Kingdom

(Dated: July 7, 2021)

We investigate the possibility of sustained orbital resonances in extreme mass ratio inspirals. Using a near-identity averaging transformation, we reduce the equations of motion for a particle moving in Kerr spacetime with self-force corrections in the neighbourhood of a resonant geodesic to a one dimensional equation for a particle moving in an effective potential. From this effective equation we obtain the necessary and sufficient conditions that the self-force needs to satisfy to allow inspiralling orbits to be captured in sustained resonance. Along the way we also obtain the full non-linear expression for the jump in the adiabatic constants of motion incurred as an inspiral transiently evolves through a strong resonance to first-order in the mass ratio. Finally, we find that if the resonance is strong enough to allow capture in sustained resonance, only a small fraction (order of the square root of mass-ratio) of all inspirals will indeed be captured. This makes observation of sustained resonances in extreme mass ratio inspirals—if they exist—very unlikely for space based observatories like eLisa.

I. INTRODUCTION

An extreme mass ratio inspiral (EMRI) is a binary system consisting of a central supermassive black hole ($\sim 10^8 M_\odot$) with a compact object of several solar masses in orbit. As the system emits gravitational radiation, the orbit gradually decays, causing the compact object to slowly spiral towards the central black hole. EMRIs are of great astrophysical interest as a source of gravitational waves for future space-based observatories \[1\]. The gradual evolution of EMRIs encodes a very detailed map of the spacetime surrounding the central supermassive black hole. This allows for precise tests of the predictions of general relativity in the strong field regime \[2\]. Furthermore, it allows for precise determination of the physical parameters (mass and spin) of the central supermassive black hole \[3\].

The dynamics of EMRIs can be studied by using perturbation theory in the small mass-ratio $\epsilon = m/M \sim 10^{-6}$. At lowest order, the motion of the compact object is that of a test particle following a geodesic in the Kerr spacetime generated by the central massive (and usually rotating) black hole. Any deviations from geodesic motion due to $\epsilon$ taking a finite value, can be included by adding a force term—known as the gravitational self-force—to the geodesic equation. In the last decade a lot of progress has been made in formulating and calculating this self-force (see \[4, 5\] for a comprehensive review).

A common phenomenon in celestial mechanics is the occurrence of resonance, well-known examples include the phase-locking of the rotational and orbital motions of the Moon, and the 3:2 spin-orbit resonance of Mercury. In general, resonances occur when a system has two interacting degrees of freedom oscillating at different frequencies $\omega_1$ and $\omega_2$. If the ratio of these frequencies is rational $\omega_1/\omega_2 = n_1/n_2$, then the linear combination $n_2 \omega_1 - n_1 \omega_2$ vanishes. Physically this means that the timescale of the associated interaction terms of the Fourier expansion diverges, leading to behaviour that is qualitatively different from the oscillatory behaviour expected for non-resonant systems.

Geodesic motion in a Kerr background can be characterized by three periods: the time between two successive periastron passes $T_r$, the time between two passes of maximum inclination $T_\theta$, and the time of one sidereal period $T_\phi$. At the geodesic level these oscillations are completely independent. The self-force corrections provide interaction between the oscillations creating the possibility of resonant effects when the periods become commensurate. Flanagan and Hinderer first worked out the effect explicitly in \[6\], although this possibility had been noted in the past. Based on a post-Newtonian approximation for the self-force (not necessarily valid in the relativistic strong field regime) they found that as an inspiral system evolves through a situation where $T_r/T_\theta$ is rational, its adiabatic constants of motion (the energy $E$, axial angular momentum $L_z$, and Carter constant $Q$) acquire a jump of order $\epsilon^{1/2}$.

Over the last few years orbital resonance in EMRIs have gained quite some attention. Grossman, Levin and Perez-Giz \[9\] have studied resonant geodesics in some detail without studying the resonant self-force effects. Gair, Yunes, and Bender \[10\] studied two different toy models for resonance and their effect on the constants of motion. By numerically calculating the “fluxes” of the constants of motion to infinity and down the central black hole horizon, Flanagan, Hughes, and Ruangsri \[11\] obtained the first true measure of the strength of the resonant effects for realistic relativistic orbits in Kerr spacetime. This was complemented by Ruangsri and Hughes \[12\] with a post-Newtonian estimate of the length of each resonant episode, and the time remaining after each episode be-
fore the orbiting object plunges into the central black hole. Isoyama et al. [13] studied the adiabatic evolution of the Carter constant when an EMRI system crosses a resonance. Brink, Geyer, and Hinderer [14] have explored the “phase space” of resonant orbits in Kerr spacetime, and noted the relevance of its structure for resonance caused by other astrophysical perturbations that couple the independent oscillations of the geodesic orbits.

Previous works have focused primarily on so-called transient resonances, situations where the EMRI evolves through the resonant situation, with the resonant condition being satisfied only at one point in time and the duration of the resonance is \( O(\epsilon^{-1/2}) \). Generally, this is what happens if the terms of the self-force coupling the independent oscillations is weak compared to the terms driving the overall evolution. However, in the general theory of resonant dynamical systems there exists the possibility that the system lingers near the resonance for a prolonged time of order of the inspiral timescale \( \sim a \). A prolonged time of order of the inspiral timescale possesses the resonant dynamics to be “strong” in some suitable sense. If a sustained resonance occurred in an EMRI, it would cause significant qualitative deviation from the normal evolution of the system, which should leave a distinctive imprint on the gravitational waveform.

In this paper, we explore exactly what conditions the gravitational self-force needs to satisfy to allow the occurrence of sustained resonances in EMRIs.

### A. Background

Geodesic motion in Kerr spacetime with mass \( M = 1 \) and spin parameter \( a \) has four constants of motion, the invariant mass \( m \), the energy per unit mass \( E \), the axial angular momentum per unit mass \( L_z \), and Carter’s constant [15]

\[
Q = Q^{\mu\nu}u_\mu u_\nu = a^2 \cos^2 \theta (1 - u_t^2) + u_\theta^2 + \cot^2 \theta u_\phi^2, \tag{1}
\]

where \( Q^{\mu\nu} \) is a Killing tensor defined by the second line, \( u_\mu \) is the 4-velocity along the geodesic, and \( r, \theta, \phi, \) and \( t \) are standard Boyer-Lindquist coordinates. These can be used to reduce the geodesic equations from four second-order differential equations to four first-order equations.

\[
\left( \frac{dr}{d\tau} \right)^2 = \frac{R(r)}{\Sigma(r, \theta)^2}, \tag{2a}
\]
\[
\left( \frac{d\theta}{d\tau} \right)^2 = \frac{\Theta(\cos \theta)}{\Sigma(r, \theta)^2}, \tag{2b}
\]
\[
\frac{d\phi}{d\tau} = \frac{\Phi_r(r) + \Phi_\theta(\theta)}{\Sigma(r, \theta)}, \tag{2c}
\]
\[
\frac{dr}{d\tau} = \frac{T_r(r) + T_\theta(\theta)}{\Sigma(r, \theta)}, \tag{2d}
\]

where \( \tau \) is proper time, \( R, \Theta, \Phi_r, \) and \( T_j \) are functions of \( r \) or \( \theta \), which can be obtained explicitly in terms of \( E, L_z \), and \( Q \), and \( \Sigma = r^2 + a^2 \cos^2 \theta \). As first noted by Carter [15], and more recently emphasized by Mino [16], these equations can be simplified by choosing an alternative time parametrization \( \lambda \) defined by

\[
\frac{d\tau}{d\lambda} = \Sigma = r^2 + a^2 \cos^2 \theta. \tag{3}
\]

This time parameter is commonly referred to as “Mino time”. The effect of this choice is that all \( \Sigma \)'s disappear from (2), and the equations for \( r \) and \( \theta \) decouple.

Since the geodesic equations in Kerr have as many constants of motion as equations (and their mutual Poisson brackets vanish), the system of equations is integrable [17]. One interesting property (out of many) of integrable systems is that they can be rewritten in action-angle variables \((q^\mu, J^\mu)\) as

\[
\frac{dq_\mu}{d\lambda} = \Upsilon_\mu(J), \tag{4a}
\]
\[
\frac{dJ_\mu}{d\lambda} = 0. \tag{4b}
\]

Schmidt [18] used this property to obtain the frequencies of Kerr orbits with respect to Boyer-Lindquist coordinate time \( \Omega_\mu \), while analytic expressions for the frequencies with respect to Mino time \( \Upsilon_\mu \) were obtained by Fujita and Hikida [19].

Figure 1 shows the parameter space of orbits in Kerr in terms of the frequencies with respect to Mino time. There are some features that are worth commenting on here. First, due to the equivalence principle the orbital dynamics cannot depend on the mass of the orbiting object \( m \), hence we only need three parameters (e.g. \( \Upsilon_r, \Upsilon_\theta \), and \( \Upsilon_\phi \)) to describe each orbit. Second, the Mino frequencies all diverge as the radius of the orbit increases—in sharp contrast to the frequencies with respect to coordinate time or proper time which vanish at infinity. This is a result of the fact that Mino time is rescaled by a factor \( \Sigma = r^2 + a^2 \cos^2 \theta \) with respect to proper time. Third,
as the radius increases the three Mino time frequencies converge to a common (diverging) value. Fourth, at the separatrix dividing bound orbits and orbits plunging into the central black hole $\Upsilon_r$ is zero.

A resonant orbit is a geodesic for which the ratio $\Upsilon_r/\Upsilon_\theta$ is a rational number, or equivalently, for which there exist coprime integers $n_r$ and $n_\theta$ such that

$$n_r\Upsilon_r + n_\theta\Upsilon_\theta = 0.$$ (5)

The sum $|n_r| + |n_\theta|$ is called the order of the resonance. As a rule of thumb for resonant dynamical systems low order resonances have greater effect than high order resonances. Hence, we will generally focus on the low order resonances. The resonant condition (5) defines a two-dimensional subspace of the parameter space, which we refer to as a resonant surface. In figure 1 we have plotted the resonant surface for the 3:2 resonance.

Note that at infinity the ratio $\Upsilon_r/\Upsilon_\theta$ is unity, while at the separatrix this ratio is zero. Hence, as an inspiral evolves from infinity to a plunge into the central black hole it passes through all resonant surfaces with $|n_r| > |n_\theta|$. Resonances are therefore a generic feature of any EMRI.

B. Overview of this paper

In this paper we will derive the exact conditions that the self-force needs to satisfy to allow sustained resonances to occur in EMRIs. In section II we first review how the self-force can be re-expressed as a correction to the equations of motion for $\Upsilon_i$ and $q_\mu$, then we apply a near identity transformation to simplify these equations of motion by absorbing all non-resonant interaction terms in $\Upsilon_i$ and $q_\mu$.

In section III we then discuss the effect of resonant interactions. We start in section III A with deriving a set of effective equations of motion in the neighbourhood of a resonant surface. At leading-order this equation describes the conservative dynamics of a particle moving in a one-dimensional effective potential. In section III B we then use this effective potential to derive the jump in the constants of motion induced by a transient crossing of a resonant surface. Section III C then continues to derive a necessary and sufficient condition for the existence of sustained resonance solutions to the equation of motion, which is related to the existence of local minima in the effective potential. Existence of such solutions however is not sufficient to conclude that they also occur in EMRI. This requires that an inspiral starting far away from the black hole is captured in a local minimum of the potential. Section III D derives the necessary and sufficient condition for this to happen as a restriction on the initial conditions. Finally, section III E discusses the fate of an EMRI after it is captured in sustained resonance.

The final section IV then discusses how the above conditions can be tested using numerical calculations of the self-force, and what can be learned from already existing numerical results.

C. Notations and Conventions

We employ units such that $c$ (speed of light) $G$ (Newton’s constant), and $M$ (the mass of the central supermassive black hole) are all unity. Consequently, all quantities appearing in this paper are dimensionless. Furthermore the constants of motion $E$, $L_z$, and $Q$ are normalized to be independent of the invariant particle mass $m$. Metrics have signature $(-+++)$.

In section III A we absorb all non-resonant interaction terms in $\Upsilon_i$ and $q_\mu$ to get

In the limit $\epsilon = 0$ the orbiting object follows a geodesic of the background spacetime. The self-force program sum-
marizes corrections order-by-order in $\epsilon$ as a term on the right-hand side (RHS) of the geodesic equation,
\[
\frac{d^2 x^\nu}{dr^2} + \Gamma^\nu_{\sigma\rho} \frac{dx^\sigma}{dr} \frac{dx^\rho}{dr} = \epsilon a^\nu(x),
\]
where $ea^\nu$ is the self-acceleration. Hinderer and Flanagan show in [12] how this second-order equation can be rewritten as two first-order equations,
\begin{align}
\frac{dP_i}{d\lambda} &= \epsilon \tilde{G}_i(\vec{P},q_r,q_{\theta}) + O(\epsilon^2), \\
\frac{dq_r}{d\lambda} &= \nu_i(\vec{P}) + \epsilon \tilde{g}_{r,i}(\vec{P},q_r,q_{\theta}) + O(\epsilon^2),
\end{align}

where the $q_r$ are generalized angle variables, and $\vec{P} = (P_1, P_2, P_3)$ are the slowly varying constants of motion. In particular they show how the forcing terms $\tilde{G}_i$ and $\tilde{g}_{r,i}$ follow from the self-acceleration $a^\nu$ (see appendix [A] for an explicit formula).

The frequencies $\nu_i$ themselves are functions of the $P_i$, and therefore constants of motion of the zeroth order system. We can therefore simplify the analysis of the system by using $\vec{\nu} = (\nu_r, \nu_{\theta}, \nu_{\phi})$ as the slow variables instead of $\vec{P}$. One might worry that the map $\nu_i : \vec{P} \mapsto \vec{\nu}$ is not invertible due to the existence of isofrequency pairs. It is known that this is the case for the frequencies $\Omega_i$, with respect to Boyer-Lindquist coordinate time $[20]$. In appendix [B] we present evidence that no such pairing occurs with respect to Mino time. In the rest of this paper we will assume that this map is indeed invertible, and use $\vec{\nu}$ as slow variables. However, the results do not crucially depend on this assumption.

Note that the RHS of equations (7) does not depend on $q_r$ and $q_{\theta}$. The equations for these two generalized angles thus decouple, and we can focus on just $q_r$ and $q_{\theta}$ as the fast variables, since $q_r$ and $q_{\theta}$ can later be recovered by a simple integration. This yields the reduced system,
\begin{align}
\frac{d\nu_i}{d\lambda} &= \epsilon G_i(\vec{\nu},\vec{q}) + O(\epsilon^2), \\
\frac{d\nu_r}{d\lambda} &= \nu_j + \epsilon g_{j,i}(\vec{\nu},\vec{q}) + O(\epsilon^2),
\end{align}

where $\vec{q} = (q_r, q_{\theta})$, $i$ runs over $(r, \theta)$, and $j$ runs over $(r, \theta)$. Explicit formula for obtaining $G_i$ and $g_{j,i}$ can found in [12] or appendix [A].

B. Averaged equations near resonance

Since the inspiral timescale of the system [8], $T_{\text{insp}} = O(\epsilon^{-1})$, is much larger than the orbital timescale $T_{\text{orb}} = O(1)$, one intuitively expects the solutions to be largely independent of the rapid oscillations of the forcing functions $G_i$ and $g_{j,i}$, and to mainly depend on their averages. This intuition falls near a resonance, where the period of some of the oscillating terms diverges. However, one still expects the other oscillating terms whose periods stay small with respect to the inspiral timescale, to play only a subdominant role. In this section we describe how to make this notion precise by employing a so-called near-identity averaging transformation following [21].

Suppose that the system encounters a resonance at $\lambda = 0$, i.e. for some coprime integers $n_r$ and $n_\theta$ the linear combination $n_r \nu_r + n_\theta \nu_\theta \equiv \nu_\parallel$ vanishes at $\lambda = 0$. The functions $G_i$ and $g_{j,i}$ can be decomposed as Fourier series,
\begin{align}
G_i(\vec{\nu},\vec{q}) &= \Omega_i(\vec{\nu}) + \sum_{N \neq 0} G_i,N(\vec{\nu}) \epsilon^{in_rq_r} \\
&\quad + \sum_{(n,k) \in R} G_i,nk(\vec{\nu}) \epsilon^{in_rq_r + ikq_\theta}, \\
g_i(\vec{\nu},\vec{q}) &= \nu_j(\vec{\nu}) + \sum_{N \neq 0} g_{j,i,N}(\vec{\nu}) \epsilon^{in_rq_r + ikq_\theta} \\
&\quad + \sum_{(n,k) \in R} g_{j,i,nk}(\vec{\nu}) \epsilon^{in_rq_r + ikq_\theta},
\end{align}

where we introduced the resonant phase combination $\nu_\parallel = n_r q_r + n_\theta q_\theta$, the sum over $N$ runs from $-\infty$ to $\infty$, and $R$ is the set $\{(n,k) \in \mathbb{Z} | (n,k) \neq N(n_r,n_\theta), \forall N \in \mathbb{Z}\}$ of all non-resonant 2-tuples. The first term in the expansion of (9) is the average adiabatic term, the first sum contains all the non-resonant oscillating terms, i.e. all oscillating terms whose period diverges at the resonance, and the final sum contains all the non-resonant oscillating terms.

We employ an appropriate near-identity averaging transformation (see [21] sec 5.1.3),
\begin{align}
\tilde{\nu}_i &= \nu_i + \epsilon T_i(\vec{\nu},\vec{q}) + O(\epsilon^2), \\
\tilde{q}_j &= q_j + \epsilon L_j(\vec{\nu},\vec{q}) + O(\epsilon^2),
\end{align}

where $T$ and $L$ are functions of $\vec{\nu}$ and $\vec{q}$ defining the transformation. The goal is to choose $T$ and $L$ in such a way as to eliminate the dependence of $\vec{\nu}$ on the rapidly oscillating non-resonant phase can be eliminated to order $\epsilon$. The details of this transformation are described in appendix [C]. The result is,
\begin{align}
\frac{d\tilde{\nu}_i}{d\lambda} &= \epsilon \tilde{G}_i(\vec{\nu},\tilde{\nu}_\parallel) + O(\epsilon^2), \\
\frac{d\tilde{q}_j}{d\lambda} &= \tilde{\nu}_j + \epsilon \tilde{g}_{j,i}(\vec{\nu},\tilde{\nu}_\parallel) + O(\epsilon^2),
\end{align}

2 Note that there is a sign ambiguity in the definition of $\nu_\parallel$. We will adopt the convention that $n_r > 0$ and $n_\theta < 0$. This has the consequence that in an inspiral $\nu_\parallel$ starts out positive and becomes negative as the infalling object crosses the resonance.

3 In the rest of this paper whenever we have a quantity $X_i$ with $i$ running over $(r, \theta)$ or $(r, \theta)$, we denote the linear combination $n_r X_r + n_\theta X_\theta$ by $X_\parallel$. 

with

\[
G_i(\tilde{T}, \hat{q}_\perp) = G_i(\tilde{T}) + \sum_{N \neq 0} G_{i,N}(\tilde{T}) e^{iN\hat{q}_\perp},
\]

\[
g_j(\tilde{T}, \hat{q}_\perp) = g_j(\tilde{T}) + \sum_{N \neq 0} g_{j,N}(\tilde{T}) e^{iN\hat{q}_\perp}.
\]

An important consequence of this result is that the forcing terms now only depend on the resonant phase \( \hat{q}_\perp \). The equations of motion for the other phases thus decouple from the system, and we can first focus on solving the system for \( q_\perp \) and \( \tilde{T} \). The other phases can then be recovered by direct integration.

### III. Resonances

#### A. Effective potential

In order to solve the averaged equations of motion (11) near the resonance at \( \lambda = 0 \) we introduce a rescaled boundary layer timescale

\[
\hat{\lambda} = \tilde{\epsilon}\lambda,
\]

with \( \tilde{\epsilon} \equiv \epsilon^\xi \) for some power \( \xi > 0 \). By expanding \( \tilde{T}_i \) and \( \tilde{q}_\perp \) in \( \tilde{\epsilon} \),

\[
\tilde{T}_i(\lambda, \epsilon) = \tilde{T}_i^0(\hat{\lambda}) + \epsilon\tilde{T}_i^1(\hat{\lambda}) + O(\epsilon^2),
\]

\[
\tilde{q}_\perp(\lambda, \epsilon) = \tilde{q}_\perp^0(\hat{\lambda}) + \epsilon\tilde{q}_\perp^1(\hat{\lambda}) + O(\epsilon^2).
\]

and plugging these into the averaged equations of motion (11) we obtain at leading-order

\[
\frac{d\tilde{T}_i^1}{d\lambda} = \epsilon^{1-\xi} G_i(\tilde{T}_i^0, \tilde{q}_\perp^0),
\]

\[
\frac{d\tilde{q}_\perp^0}{d\lambda} = \tilde{T}_i^1.
\]

Obtaining a non-singular and non-trivial limit as \( \tilde{\epsilon} \to 0 \) requires that we set \( \xi = 1/2 \).

Differentiating (15b) with respect to \( \hat{\lambda} \) and plugging in (15a) gives the second order equation,

\[
\frac{d^2\tilde{q}_\perp^0}{d\lambda^2} = G_\perp(\tilde{T}_i^0, \tilde{q}_\perp^0).
\]

By multiplying both sides of this equation with \( \frac{d\tilde{q}_\perp^0}{d\lambda} \) and integrating once with respect to \( \hat{\lambda} \), we obtain an equivalent first-order equation,

\[
\frac{1}{2} \left( \frac{d\tilde{q}_\perp^0}{d\lambda} \right)^2 = K - V(\tilde{q}_\perp^0),
\]

where

\[
V(q) = -G_\perp(\tilde{T}_i^0)q + i \sum_{N \neq 0} \frac{G_{\perp,N}(\tilde{T}_i^0)}{N} e^{iNq},
\]

and \( K \) is an integration constant.

This is the phase resonance equation studied in [10] as a toy model. It can be interpreted as the equation for a 1D particle moving in an effective potential, \( V \) with energy \( K \).

\[\text{Figure 2.}\]

Figure 2. The top figure shows the phase portrait of solutions of (17), with \( V = \tilde{q}_\perp - \frac{1}{2} \cos \tilde{q}_\perp + \frac{1}{2} \sin \tilde{q}_\perp \). The bottom figure show the potential of the same system. The levels plotted in the lower figure correspond to the solutions plotted in the top figure with the same shade.

#### B. Transient resonance

As discussed in [6], the effect of a resonance compared to the adiabatic approximation ignoring the oscillating terms is an order \( \tilde{\epsilon} \) jump in the slow variables \( \tilde{T}_i \). We now provide an expression for the size of these jumps.
The initial condition that the system crosses the resonance at \( \lambda = \hat{\lambda} = 0 \) corresponds to setting,

\[ K = V(q_0), \tag{19} \]

where \( q_0 \) is the value of \( \hat{q}_\perp(\hat{\lambda}) \) at \( \hat{\lambda} = 0 \). We can then invert \( [17] \) to find \( \hat{\lambda} \) as a function of \( \hat{q}_\perp \).

\[
\hat{\lambda}(\hat{q}_\perp) = \int_{q_0}^{\hat{q}_\perp} \frac{1}{\pm 2\sqrt{2(K - V(q))}} \, dq, \tag{20}\n\]

where the \( \pm \) sign corresponds to the branch before and after resonance.

The effective potential \( V \) can be split into an adiabatic contribution \( \overline{V} \) and a (resonant) oscillatory contribution \( \Delta V \),

\[
\begin{align}
V(\hat{q}_\perp) & = \overline{V}(\hat{q}_\perp) + \Delta V(\hat{q}_\perp), \tag{21a} \\
\overline{V}(\hat{q}_\perp) & = -G_{\perp}(\overline{\hat{\lambda}})\hat{q}_\perp, \tag{21b} \\
\Delta V(\hat{q}_\perp) & = i \sum_{N \neq 0} G_{\perp,N}(\overline{\hat{\lambda}}) e^{iN\hat{q}_\perp}. \tag{21c}
\end{align}
\]

In the adiabatic approximation (i.e. ignoring \( \Delta V \)) equation \( (20) \) can be integrated explicitly, yielding

\[
\hat{\lambda}(\hat{q}_\perp) = \pm 2\sqrt{\frac{q_0 - \hat{q}_\perp}{G_{\perp}}} \tag{22}\n\]

If we denote the full integral \( (20) \) by \( \hat{\lambda} \) and its adiabatic approximation \( \hat{\lambda}_0 \), then as the system “evolves” from \( \hat{q}_\perp = -\infty \) to \( q_0 \) and back again the full integral “accumulates” an additional amount of time over the adiabatic approximation given by,

\[
\Delta \hat{\lambda} = \lim_{q \to -\infty} \hat{\lambda}(q) - \hat{\lambda}_0(q) + \lim_{q \to \infty} \hat{\lambda}(q) - \hat{\lambda}_0(q) = \int_{-\infty}^{q_0} \left( \frac{\sqrt{2}}{\sqrt{V(q_0) - V(q)}} - \frac{\sqrt{2}}{\sqrt{V(q_0) - V(q)}} \right) \, dq. \tag{23}\n\]

In the asymptotic regime \( 1/\epsilon \gg |\hat{\lambda}| \gg 1 \), far away from the resonance the resonant terms \( \Delta V \) are oscillating rapidly and can be ignored at lowest order. Equation \( [16] \) then has solutions of the form

\[
\hat{q}_\perp(\hat{\lambda}) = \frac{G_{\perp}}{2} \left( \hat{\lambda} + \text{sign}(\hat{\lambda}) \frac{\Delta \hat{\lambda}}{2} \right)^2 + C, \tag{24}\n\]

where \( C \) is some constant that is not relevant here. From equation \( [11] \), we then find the total jump in the frequencies \( \hat{\gamma} \),

\[
\Delta \hat{\gamma}_i = \epsilon G_{\perp} \Delta \hat{\lambda} + O(\epsilon^2). \tag{25}\n\]

Together, equations \( (23) \) and \( (25) \) provide a full expression for this jump (to leading-order in \( \epsilon \)) in previous works \([6,10]\), the size of this jump was only calculated in the limit that \( |\Delta V| \ll |\overline{V}| \) (although a similar result is supposed to appear in \([22]\)). To check our result with theirs, we expand \( (23) \) in \( \Delta V \),

\[
\begin{align}
\Delta \hat{\lambda} & \approx -\frac{1}{\sqrt{2}} \int_{-\infty}^{q_0} \frac{\Delta V(q_0) - \Delta V(q)}{(V(q_0) - V(q))^{1/2}} \, dq, \tag{26a} \\
& = -i \sum_{N \neq 0} G_{\perp,N} e^{iNq_0 - \epsilon iNq} \int_{-\infty}^{q_0} (q_0 - q)^{3/2} \, dq. \tag{26b}
\end{align}\n\]

This integral can be computed explicitly,

\[
\Delta \hat{\lambda} = \sum_{N \neq 0} \frac{\sqrt{2\pi}}{G_{\perp,N}^2} G_{\perp,N} e^{iNq_0 - \epsilon iNq}, \tag{27}\n\]

where \( \sigma \) is the sign of \( N \). Plugging this result into equation \( (25) \) yields,

\[
\Delta \hat{\gamma}_i = \epsilon G_{\perp} \sum_{N \neq 0} \frac{\sqrt{2\pi}}{G_{\perp,N}^2} G_{\perp,N} e^{iNq_0 - \epsilon iNq}, \tag{28}\n\]

which agrees with the linear result in \([6]\).
where summation over the repeated index $j$ is implicit and the functions $G$ and $g$ and their derivatives are to be evaluated at $\vec{T}$. Bosley and Kevorkian in [23] describe how to solve such equations to find the evolution of a system captured in sustained resonance. For our current purpose it is enough to note such sustained resonance solutions can exist in principle and stay near the resonance for a prolonged period resulting in a qualitatively different behaviour from the adiabatic approximation of an inspiral.

A necessary and sufficient condition for sustained resonance solutions to exist is that the effective potential for the resonant phase $V$ has a local minimum. i.e. there must be values of the resonant phase $\dot{q}_\perp$ for which $\frac{dV}{dq_\perp} = 0$. For the Fourier components of the self-force this means

$$G_1(\vec{T}^0) = -\sum_{N\neq 0} G_{\perp,N}(\vec{T}^0)e^{IN\dot{q}_\perp}. \quad (31)$$

$$\Delta G_\perp(\vec{T}^0) \equiv \frac{G^{\mathrm{max}}_\perp(\vec{T}^0) - G^{\mathrm{min}}_\perp(\vec{T}^0)}{|G^{\mathrm{max}}_\perp(\vec{T}^0) + G^{\mathrm{min}}_\perp(\vec{T}^0)|} \geq 1, \quad (32)$$

where

$$G^{\mathrm{max}}_\perp(\vec{T}) = \max_{q_\perp \in [0,2\pi]} G_\perp(\vec{T}, q_\perp), \quad (33)$$

$$G^{\mathrm{min}}_\perp(\vec{T}) = \min_{q_\perp \in [0,2\pi]} G_\perp(\vec{T}, q_\perp). \quad (34)$$

In terms of the Fourier components of the self-force this implies the necessary (but not sufficient) condition,

$$|G_\perp(\vec{T}^0)| < \sum_{N\neq 0} |G_{\perp,N}(\vec{T}^0)|, \quad (35)$$

for the existence of sustained resonances. The negation of this condition is therefore sufficient (but not necessary) to show that there are no sustained resonances.

### D. Capture condition

However, the existence of sustained resonance solutions to equation (11), does not guarantee that they will also occur for inspiralling solutions. In fact, in the lowest order approximation of this equation given by (17) $K$ is a constant of motion; consequently, if a solution starts at $q_\perp = -\infty$ it cannot get stuck in a local potential well (see figure 3), and will in general return to $q_\perp = -\infty$ after reflecting off the potential. The only exceptions are the isolated solutions that asymptote to the unstable equilibrium at a local maximum of the potential.

In this section we derive the conditions under which EMRI may be captured in sustained resonance. In [24] Haberman derived these conditions for a Hamiltonian system with one degree of freedom, where the resonance is caused by a perturbation that is a predetermined function of time. We generalize his method to apply to orbital resonances in Kerr spacetime, where there are three slowly evolving frequencies and the perturbation causing the resonance (the self-force) is a function of the frequencies.

From the expansion (14), we can find the next order contribution to the second order equation of motion for $\dot{q}_\perp$,

$$\frac{d^2\dot{q}_\perp}{d\lambda^2} = \frac{dV}{dq_\perp}(\dot{q}_\perp) + \ddot{\vec{Y}}^\perp_h(\dot{q}_\perp) + O(\dot{\vec{q}}_\perp^2), \quad (36)$$

where \(\ddot{\vec{Y}}^\perp_h\) denotes the time derivative of \(\vec{Y}^\perp_h\) and \(O(\dot{\vec{q}}_\perp^2)\) represents terms of order $O(\dot{\vec{q}}_\perp^2)$.
To make this notion precise, recall that the leading-order of the potential (see figure 4).

...coming in from ˆq effective equations of motion for ˆ⃗(14). The to that of the slowly evolving frequencies ⃗ure 3, but with a constant dissipative term ˆ that is still captured, while the thick line labelled “min” corresponds to the solution with the minimal effective energy Kmin that is still captured in the potential.

with

\[ h_j^1(\hat{q}_\perp) = \frac{\partial G_i}{\partial \hat{T}_j}(\hat{T}^0, \hat{q}^0_\perp) + \frac{\partial g_l}{\partial \hat{q}_\perp}(\hat{T}^0, \hat{q}_\perp) \delta_{ij} \tag{37} \]

\[ = \frac{dG_i}{d\hat{T}_j}(\hat{T}^0) + \sum_{N \neq 0} \left( \frac{dG_{i,N}}{d\hat{T}_j}(\hat{T}^0) + iN\delta_{ij}g\perp,N(\hat{T}^0) \right)e^{iN\hat{q}_\perp}. \tag{38} \]

The first-order $\dot{\epsilon}$ correction couples the evolution of $\hat{q}_\perp$ to that of the slowly evolving frequencies $\hat{T}^1$ defined in (14). The $\hat{T}^1$ correction acts as a friction term in the effective equations of motion for $\hat{q}_\perp$, allowing solutions coming in from $\hat{q}_\perp = -\infty$ to sink into a local minimum of the potential (see figure 3).

To make this notion precise, recall that the leading-order system had a constant of motion,

\[ K = \frac{1}{2} \left( \frac{d\hat{q}_\perp}{d\lambda} \right)^2 + V(\hat{q}_\perp). \tag{39} \]

When we include the higher-order corrections, this quantity is no longer constant. Taking a derivative and plugging in (36) gives,

\[ \frac{dK}{d\lambda} = \frac{d\hat{q}_\perp}{d\lambda} \left( \frac{d^2\hat{q}_\perp}{d\lambda^2} + \frac{dV}{d\hat{q}_\perp} \right), \tag{40} \]

\[ = \dot{\epsilon} \frac{d\hat{q}_\perp}{d\lambda} \hat{T}^1_j h_j^1(\hat{q}_\perp), \tag{41} \]

which can be integrated to obtain $K$ as a function of $\hat{\lambda}$,

\[ K(\hat{\lambda}) = K_0 + \dot{\epsilon} \int_0^{\hat{\lambda}} \frac{d\hat{q}_\perp}{d\lambda} \hat{T}^1_j h_j^1(\hat{q}_\perp) d\lambda, \tag{42} \]

where $K_0$ is the value of $K$ at $\hat{\lambda} = 0$. (Recall that $\hat{\lambda} = 0$ is the time that the system (first) passes through resonance, i.e. $\hat{T}_\perp = 0$ at $\hat{\lambda} = 0$.)

In order for the system to be captured in sustained resonance, $K$ must decrease sufficiently as $\hat{q}_\perp$ “passes over” the local minimum of the potential. In general, there will be a window of values $K_0 \in [K_{\text{min}}, K_{\text{max}}]$ for which the solution is captured.

The lowest extreme, $K_{\text{min}}$, corresponds to the solution that comes in from negative infinity, barely scrapes the top of the local maximum of the potential, and then reflects of the potential at $V = K_{\text{min}}$. The solution corresponding to the highest extreme $K_{\text{max}}$, first reflects of the potential at $V = K_{\text{max}}$ and then asymptotically approaches the local maximum of the potential from the right.

To calculate the values of $K_{\text{min}}$ and $K_{\text{max}}$, we use the fact that these extremal solutions are close to the critical solutions of the lowest order system, i.e. the solutions of (17) with $K$ equal to the local maximum of the potential $K_c$. We can use this to find an approximation of $\frac{d\hat{q}_\perp}{d\lambda}$ as a function of $\hat{q}_\perp$,

\[ \frac{d\hat{q}_\perp}{d\lambda} = \pm \sqrt{2(K_c - V(\hat{q}_\perp))} + O(\dot{\epsilon}). \tag{43} \]

We can then use (15a) to obtain an approximation for $\hat{T}^1$. Plugging (43) into (15a) yields,

\[ \frac{d\hat{T}^1_j}{d\lambda} \frac{d\hat{q}_\perp}{d\lambda} = \frac{\hat{G}_j + \sum_{N \neq 0} G_{i,N} e^{iN\hat{q}_\perp}}{\pm \sqrt{2(K_c - V(\hat{q}_\perp))}} + O(\dot{\epsilon}), \tag{44} \]

where here and in the rest of the equations in this section the functions $\hat{G}_j$ and $G_{i,N}$ (and their derivatives) are understood to be evaluated at the initial $\hat{T}^0$. This equation can be integrated to obtain,

\[ \hat{T}^1_j(\hat{q}_\perp) = \int_{\hat{q}_0}^{\hat{q}_\perp} \frac{\hat{G}_j + \sum_{N \neq 0} G_{i,N} e^{iN\hat{s}}}{\sqrt{2(K_c - V(s))}} ds + O(\dot{\epsilon}). \tag{45} \]
We can then use the approximations for \( \frac{d\hat{q}}{d\lambda} \) and \( \hat{T}^1 \) to obtain leading-order approximations for \( K_{\min} \) and \( K_{\max} \) from (42)

\[
K_e = K_{\min} - \hat{\epsilon} \int_{-\infty}^{0} \frac{d\hat{q}}{d\lambda} \hat{T}^1_j(\hat{q}_\perp) \, d\lambda + O(\hat{\epsilon}^2) \tag{46}
\]

\[
= K_{\min} - \hat{\epsilon} \int_{q_0}^{q_0} \hat{T}^1_j(\hat{q}) h_j^1(\hat{q}) \, d\hat{q} + O(\hat{\epsilon}^2), \tag{47}
\]

and

\[
K_e = K_{\max} + \hat{\epsilon} \int_{0}^{\infty} \frac{d\hat{q}}{d\lambda} \hat{T}^1_j(\hat{q}_\perp) \, d\lambda + O(\hat{\epsilon}^2) \tag{48}
\]

\[
= K_{\max} - \hat{\epsilon} \int_{q_0}^{q_0} \hat{T}^1_j(\hat{q}) h_j^1(\hat{q}) \, d\hat{q} + O(\hat{\epsilon}^2). \tag{49}
\]

Together these give the size of the window \( \Delta K = K_{\max} - K_{\min} \),

\[
\Delta K = 2\hat{\epsilon} \int_{q_0}^{q_0} \hat{T}^1_j(\hat{q}) h_j^1(\hat{q}) \, d\hat{q} + O(\hat{\epsilon}^2) \tag{50}
\]

\[
= 2\hat{\epsilon} \sum_j \int_{q_0}^{q_0} d\hat{q} \int_{q_0}^{q_0} d\hat{s} \sqrt{2(K_e - V(s))} \frac{Q_j + \sum_{N \neq 0} G_{j,N} e^{iNj}}{d\hat{T}^1_j} \tag{51}
\]

\[
+ iN\delta_{j,N} g_{j,N} e^{iNj}) \right) + O(\hat{\epsilon}^2). \]

The first thing we notice is that the size of the window in \( K \) is very small (of order \( \hat{\epsilon}^2 \)). So, if we have an EMRI system with \( \epsilon = 10^{-6} \) that has a resonance strong enough to have sustained resonance solutions, only about 1 in 1000 inspirals will be captured. The rest will simply experience a transient resonance and obtain a kick to their constants of motion given by (23) and (25).

Since the window \( \Delta K \) is small we may obtain the window for the initial resonant phase \( \Delta\hat{q} \) as,

\[
\Delta\hat{q} = V'(q_0) \Delta K. \tag{52}
\]

Since \( V' = O(1) \), this implies that the window \( \Delta\hat{q} \) is also small.

Given the existence of sustained resonance solutions, a necessary and sufficient condition for the existence of captured inspiralling solutions is that \( K_{\max} \geq K_{\min} \). The intuitive interpretation of this condition is that the friction term in (36) must act to decrease \( K \) most of the time, i.e. if we heuristically think of \( \hat{T}^1_j \) as \( \frac{d\hat{q}}{d\lambda} \), then \( h_j^1 \) must be mostly negative. Given that along an inspiral we heuristically expect the Mino frequencies to decrease and the self-force to increase, we do indeed in general expect \( h_j^1 \) to be negative. This is by no means guaranteed, but we are given a general indication of the sign of \( \Delta K \). However, since the existence of sustained resonance solutions may require non-generic conditions—if they exist in EMRIs at all (see section IV)—non-generic behaviour of \( h_j^1 \) certainly is not excluded.

E. Escape from sustained resonance

Once an inspiral has been captured in resonance it will continue to evolve according to the equations of motion (30a), until it manages to escape from the resonance. In general, there are two ways in which the system can escape from sustained resonance. The first is the reverse of the capture condition, i.e. the friction term in (36) should have the “wrong” (i.e. negative) sign for an extended period of time. As we argued above, this is not what we heuristically expect to happen.

The other possibility is that the local minimum of the effective potential \( V \) disappears as the system continues to evolve. As can be gathered from the geometry of the phase diagram in figure 1, evolution of the system along a resonant surface while still decreasing the energy implies that it becomes more circular. We know that for circular orbits the self-force can only depend on the \( q_0 \) phase. Consequently the resonant oscillating terms of self-force must vanish as the orbit approaches circularity, and the local minimum of the effective potential \( V \) must disappear before the orbit becomes fully circular.

However, without a more detailed knowledge of the functional form of the self-force this is all we can say about the evolution of an inspiral captured in sustained resonance.

IV. EVALUATING THE RESONANCE CONDITIONS

In the previous sections we have derived two necessary and together sufficient conditions for the existence of EMRI solutions that are captured in sustained resonance at some frequencies \( \hat{T} \),

\[
\Delta G_{\perp}(\hat{T}) \geq 1, \tag{53a}
\]

\[
\Delta K(\hat{T}) \geq 0. \tag{53b}
\]

Note that since the \( (n,k) \)-modes of the self-force are expected to become small for large values of \( n \) and \( k \), one expects \( \Delta G_{\perp} \) to become small for resonances with large \( n_r \) and \( n_\theta \). Consequently, one can at best expect (53a) to be satisfied for low order resonances (i.e. \( |n_r/n_\theta| = 2/3, 2/4, 2/5, \ldots \)).
A useful fact for testing (53a) is that, as a result of the resonance condition, \( q_{\perp} \) is constant on any resonant orbit, and thus forms an addition constant of motion. This allows us to study the long term average value of \( G_{\perp} \) which is related to \( E, L_z, \) and \( Q \) through equation (3a)—on geodesic orbits with different \( q_{\perp} \) to find the dependence of \( G_{\perp} \) on \( q_{\perp} \). This can then be used to obtain \( \Delta G_{\perp} \). It is worth emphasizing that this procedure will provide all Fourier modes of \( G_{\perp} \) that depend only on \( q_{\perp} \), including those that are odd under the transformation \( q_{\perp} \to 2\pi - q_{\perp} \), despite the fact that for non-resonant orbits these later modes are associated with the conservative self-force.

In previous works, the additional constant of motion present for resonant orbits is defined in different ways. In [13], they define \( \Delta \lambda \) as Mino time difference between the minima of the \( r \) and \( \theta \) oscillations of the orbits. If we define the generalized angles \( q_r \) and \( q_\theta \) to be zero at their minima, then this implies

\[
\begin{align*}
q_r(\lambda) &= (\lambda - \lambda_0)T_r, \quad (54a) \\
q_\theta(\lambda) &= (\lambda - \lambda_0 + \Delta \lambda)T_\theta, \quad (54b)
\end{align*}
\]

where \( \lambda_0 \) is the time at which the \( r \) oscillations reach their minimum. Plugging this in the definition of \( q_{\perp} \) gives

\[
\begin{align*}
q_{\perp}(\lambda) &= n_r w_r(\lambda) + n_\theta w_\theta(\lambda) \quad (55a) \\
&= n_r(\lambda - \lambda_0)T_r + n_\theta(\lambda - \lambda_0 + \Delta \lambda)T_\theta \quad (55b) \\
&= n_\theta T_\theta \Delta \lambda, \quad (55c)
\end{align*}
\]

where \( n_r T_r + n_\theta T_\theta = 0 \). Hence \( \Delta \lambda = q_{\perp}/(n_\theta T_\theta) \).

In [11], \( \chi_0 \) is defined as,

\[
\chi_0 = \arccos \left( \frac{z(\lambda_0)}{z_{\max}} \right), \quad (56)
\]

where \( z = \cos \theta \). Although rather complicated, the relation between \( \chi_0 \) and \( q_{\perp} \) can be obtained explicitly from the analytic solutions of the geodesic equations in Kerr found in [19].

Since \( \Delta G_{\perp} \) only depends on the minimum and maximum of \( G_{\perp} \), it does not matter whether we obtain \( G_{\perp} \) as a function of \( q_{\perp}, \Delta \lambda, \) or \( \chi_0 \). The extrema are always the same.

In [11], \( \tilde{P} = (\tilde{E}, \tilde{L}_z, \tilde{Q}) \) was calculated by solving the Teukolsky equation for resonant orbits, and obtaining the “fluxes” at infinity and the horizon. They find that \( \Delta \tilde{P}_r \) (defined analogous to the definition in (32) of \( \Delta G_{\perp} \)) is at most of the order of a few percent. Since they only give \( \Delta \tilde{P}_r \), and not the values of \( \tilde{P}_r \), we cannot calculate \( \Delta G_{\perp} \) directly from their results, but unless \( \tilde{P}_r \) is almost tangent to the resonant surface \( \Delta G_{\perp} \) should be of similar order of magnitude as \( \Delta \tilde{P}_r \). This result is consistent with the conclusions of Flanagan and Hinderer [6] based on a post-Newtonian approximation of the self-force (even though the resonant orbits are in the strong field regime where the PN approximations lose their validity). Such small values of \( \Delta G_{\perp} \) indicate that there are no sustained resonance solutions (let alone captured sustained resonances) for the resonant orbits probed in [11].

However [11] probes only a few points in the parameter space. Consequently, we cannot quite exclude the possibility of sustained resonances.

Testing the second condition for capture, (53b), requires not only knowledge of \( G_{\perp} \), but also of its derivatives with respect to \( T_r \). Consequently, we cannot rely on the shortcut provided by calculating the changes of the constants of motion on resonant geodesics. To test it, a complete survey of the self-force in a neighbourhood of the resonant surface would be needed. Such a survey, is currently beyond the state-of-the-art.

V. DISCUSSION AND CONCLUSIONS

In this paper we derived a set of necessary and sufficient conditions (53) that the self-force needs to satisfy for sustained resonances to occur in extreme mass ratio inspirals. Along the way we obtained an expression (equations (23) and (25)) for the jump made by the constants of motion when an EMRI encounters a transient resonance. This expression—valid to lowest order in the mass ratio \( \epsilon \)—applies to resonances of any strength, including transient crossings of resonances strong enough to allow the existence of sustained resonances.

Current numerical evidence provides no indication that condition (53a) is satisfied for any resonant orbits in Kerr spacetime. However, these results are not sufficient to completely rule out that the condition may be satisfied for some particular resonant orbits. Efforts to provide a more comprehensive sweep of the parameter space of resonant surfaces are currently under way. Numerical testing of the second condition (53b) would require almost full knowledge of the self-force for generic orbits in Kerr spacetime, which at this point is not (yet) available.

The initial conditions needed to allow capture into sustained resonance however imply that even if the conditions (53) are satisfied for some orbits, the probability that an EMRI crossing that orbit is indeed captured is only of order \( \sim \epsilon^{1/2} \). This makes the chance of actually observing such an event in an astrophysical context fairly remote, unless very large numbers of EMRIs are observed. However, if an EMRI stuck in sustained resonance is ever observed, it would provide a unique glance into the resonant structure of the space of orbits in Kerr, which as pointed out in [11], is sensitive to the geometry of the spacetime.
ACKNOWLEDGMENTS

The author would like to thank Leor Barack for many helpful discussions. He also likes to thank Leor Barack and Priscilla Cañizares for a careful reading of this paper. This work was supported by a NWO Rubicon grant.

Appendix A: Expressions for the forcing terms

In [17] Hinderer and Flanagan derive the expression for the forcing terms \( \tilde{g}_i \) and \( G_i \) in terms of the self-acceleration \( a_\nu \),

\[
\tilde{g}_i(P, \eta) = \frac{d\eta}{d\lambda} \left( \frac{\partial q_i}{\partial p_\nu} \right)_x a_\nu, \quad (A1a)
\]

\[
G_i(P, \eta) = \frac{d\eta}{d\lambda} \left( \frac{\partial P_i}{\partial p_\nu} \right)_x a_\nu, \quad (A1b)
\]

where \( p_\nu \) is the four-momentum, and the parentheses with subscript \( x \) mean that the partial derivatives are performed keeping the position \( x^\nu \) fixed. They also show how to calculate the partial derivatives more explicitly. The result for \( G \) is rather simple,

\[
\tilde{G}(P, \eta) = \frac{d\eta}{d\lambda} (-a_i, a_\phi, 2Q^{\mu
u} a_\mu a_\nu). \quad (A2)
\]

The expression for \( \tilde{g} \) is much more involved, and since \( \tilde{g} \) is not needed to evaluate the capture conditions derived in this paper, we will not repeat it here.

The expressions for the forcing terms for the equation of motion in terms of the frequencies, \( G \) and \( g \), can be easily derived from \( \tilde{g} \) and \( \tilde{G} \),

\[
G_i(\tilde{T}, \eta) = \frac{\partial \tilde{Y}_i}{\partial P_j} \tilde{G}_j(\tilde{T}(P), \eta), \quad (A3a)
\]

\[
g_j(\tilde{T}, \eta) = \tilde{g}_j(\tilde{T}(P), \eta). \quad (A3b)
\]

Evaluating this requires knowledge of \( \tilde{T}(P) \). Unfortunately, no analytic form of \( \tilde{T}(P) \) is known. However, it is possible to explicitly get \( \tilde{T} \) and \( P \) as functions of \( \tilde{s} = (p, e, z_{\text{max}}) \) (see [18] [19] for an explicit formula).

We can therefore obtain \( \tilde{T}(P) \) numerically by numerically inverting the relation \( P(\tilde{s}) \). Similarly we obtain \( \frac{\partial \tilde{Y}_i}{\partial s_j} \) and \( \frac{\partial \tilde{P}_i}{\partial s_j} \) from \( \frac{\partial \tilde{Y}_i}{\partial P_j} \) and \( \frac{\partial \tilde{P}_i}{\partial P_j} \),

\[
\frac{\partial \tilde{Y}_i}{\partial P_j} = \frac{\partial \tilde{Y}_i}{\partial s_k} \frac{\partial s_k}{\partial P_j}. \quad (A4)
\]

Appendix B: Carter-Mino time frequencies

In this paper we identify (invariant tori of) bound orbits by their frequencies with respect to Mino time, \( \tilde{T} = (\tilde{T}_r, \tilde{T}_\theta, \tilde{T}_\phi) \), instead of more conventional sets of invariants such as \( (E, L_z, Q) \) or \( (p, e, z_{\text{max}}) \). This choice is convenient because it avoids the appearance of the function \( \tilde{T}(E, L_z, Q) \) and its derivatives in many of the equations. However, this choice is not crucial for any of the conclusions of the paper.

For this choice to be valid, we need the map \( F: (p, e, z_{\text{max}}) \mapsto (\tilde{T}_r, \tilde{T}_\theta, \tilde{T}_\phi) \) to be invertible, or at least for it be invertible in a neighbourhood of the resonant surfaces. It is known that this is not the case for the related map \( (p, e, z_{\text{max}}) \mapsto (\Omega_r, \Omega_\phi, \Omega_\theta) \) to the frequencies with respect to Boyer-Lindquist coordinate time [20]. In that case, there exist distinct invariant tori of bound orbits with the same triple of coordinate frequencies \( (\Omega_r, \Omega_\phi, \Omega_\theta) \), and the region of parameter space where these isofrequency pairs appear intersects at least some of the low integer resonant surfaces. This might be cause to worry that the map to Mino time frequency triples is equally degenerate.

We know of no formal proof that the map \( F \) is invertible, but our investigations indicate that this seems to be the case. We first present a formal proof that the reduced map \( (p, e) \mapsto (\tilde{T}_r, \tilde{T}_\theta = \tilde{T}_\phi) \) for orbits in Schwarzschild spacetimes is invertible, unlike the analogue map for \( \Omega \).

We then present numerical plots of the parameter space for bound orbits in Kerr spacetime presented in Mino frequencies which show the map to be regular in the plotted region, which includes the low integer resonant surfaces.

1. Schwarzschild case

Due to the spherical symmetry of Schwarzschild spacetime the orbital dynamics are independent of the inclination of the orbit. Consequently, there are only two relevant invariant parameters for each orbit. The same symmetry also implies that the oscillations about the equatorial plane must match the azimuthal period of the orbit, i.e. \( \tilde{T}_\phi = \tilde{T}_\theta \). This leaves a reduced map \( F: (p, e) \mapsto (\tilde{T}_r, \tilde{T}_\theta = \tilde{T}_\phi) \) to consider. The expressions for \( \tilde{T}_r, \tilde{T}_\phi \) in terms of \( (p, e) \) are readily obtained from the general expressions in [18] [19].

\[
\tilde{T}_r = \frac{\pi \sqrt{\left(\frac{p(p-6+2e)}{p-e^2-3}\right)}}{2K\left(\frac{4e}{p-6+2e}\right)}, \quad (B1a)
\]

\[
\tilde{T}_\phi = \frac{p}{\sqrt{p-e^2-3}}, \quad (B1b)
\]

\footnote{Here, \( p \) is the semi-latus rectum, \( e \) the eccentricity, and \( z_{\text{max}} \) is the maximal value of \( z = \cos \theta \).}
where

\[ K(x) = \int_0^{\pi/2} \frac{1}{\sqrt{1 - x \sin^2 \sigma}} \, d\sigma \]  

(B2)

is the complete elliptic integral of the first kind.

For convenience, we shift the semilatus rectum \( p \) by the \( 6 + 2e \), the value of \( p \) at the separatrix, i.e. \( \tilde{p} = \frac{6 + 2e}{4e} \). As a consequence the space of bound orbits in Schwarzschild spacetime is given by \( \tilde{p} \geq 1 \) and \( 0 \leq e < 1 \). In these variables the frequencies are

\[ \Upsilon_r = \frac{\pi}{4e} \frac{\sqrt{\tilde{p}(4e\tilde{p} - 2e + 6)}}{K(\frac{1}{\tilde{p}})}, \]

(B3a)

\[ \Upsilon_\phi = \frac{4e\tilde{p} - 2e + 6}{\sqrt{4e\tilde{p} + (1 - e)(3 + e)}} \]  

(B3b)

Note that the map \((p, e) \mapsto (\tilde{p}, e)\) is invertible. Hence \( F \) is invertible if and only if \( \tilde{F}: (\tilde{p}, e) \mapsto (\Upsilon_r, \Upsilon_\phi) \) is invertible. By the inverse function theorem it is enough to show the Jacobian of \( \tilde{F} \) is non-zero everywhere.

The Jacobian \( J \) of the map \( \tilde{F} \) is given by

\[ J = 2e\pi \sqrt{\tilde{p}(4e\tilde{p} - 2e + 6)} \times \]

\[ \frac{2(\tilde{p}^2 - \tilde{p} + 1)E(\frac{1}{\tilde{p}}) - (2\tilde{p}^2 - 3\tilde{p} + 1)K(\frac{1}{\tilde{p}})}{\tilde{p}(4e\tilde{p} - (1 - e)(3 + e))^2 K(\frac{1}{\tilde{p}})^2} \]  

(B4)

where

\[ E(x) = \int_0^{\pi/2} d\sigma \sqrt{1 - x \sin^2 \sigma} \]  

(B5)

is the complete elliptic integral of the second kind.

Since the square and square root factors are manifestly positive, \( J > 0 \) is equivalent to

\[ f(\tilde{p}) = 2(\tilde{p}^2 - \tilde{p} + 1)E\left(\frac{1}{\tilde{p}}\right) - (2\tilde{p}^2 - 3\tilde{p} + 1)K\left(\frac{1}{\tilde{p}}\right) > 0. \]  

(B6)

Note that \( f \) is only a function of \( \tilde{p} \) and not of \( e \). We will show that \( f \) is a monotonically increasing function \( \tilde{p} \). Since \( f(1) = 2 \), this would prove that \( f \) is strictly positive. We first calculate the third derivative of \( f \),

\[ f''' = \frac{(2 - \tilde{p})E\left(\frac{1}{\tilde{p}}\right) + (\tilde{p} - 1)K\left(\frac{1}{\tilde{p}}\right)}{8/15(\tilde{p} - 1)\tilde{p}^3}. \]  

(B7)

If we expand the elliptic integral in the numerator we find,

\[ \int_0^{\pi/2} d\theta (2 - \tilde{p})(1 - \frac{\sin^2 \theta}{1 + \tilde{p}})^{1/2} + \frac{\tilde{p} - 1}{(1 - \frac{\sin^2 \theta}{1 + \tilde{p}})^{1/2}}. \]  

(B8)

The integrand can be rewritten to,

\[ \frac{\cos^2 \theta + (1 + \sin^2 \theta)(\tilde{p} - 1)}{\tilde{p}(1 - \frac{\sin^2 \theta}{\tilde{p}})^{1/2}} > 0, \]  

(B9)

which is manifestly positive for all \( \tilde{p} \geq 1 \) and \( 0 < \theta < \pi/2 \). Consequently, \( f''' > 0 \) for all \( \tilde{p} \geq 0 \).

The first and second derivative of \( f \) are given by

\[ f' = \frac{(8\tilde{p}^2 - 3\tilde{p} + 2)E\left(\frac{1}{\tilde{p}}\right) - (8\tilde{p}^2 - 7\tilde{p} - 1)K\left(\frac{1}{\tilde{p}}\right)}{2\tilde{p}}, \]  

(B10)

\[ f'' = \frac{(16\tilde{p}^2 + 4\tilde{p} + 6)E\left(\frac{1}{\tilde{p}}\right) - (16\tilde{p}^2 - 4\tilde{p} + 3)K\left(\frac{1}{\tilde{p}}\right)}{4\tilde{p}^2}. \]  

(B11)

If we use that

\[ E\left(\frac{1}{\tilde{p}}\right) = \frac{\pi}{2} \left[ 1 - \frac{1}{4\tilde{p}} - \frac{3}{64\tilde{p}^2} + O\left(\frac{1}{\tilde{p}^3}\right) \right], \]  

(B12)

\[ K\left(\frac{1}{\tilde{p}}\right) = \frac{\pi}{2} \left[ 1 + \frac{1}{4\tilde{p}} + \frac{9}{64\tilde{p}^2} + O(\tilde{p}^3) \right], \]  

(B13)

we find both vanish at infinity,

\[ \lim_{\tilde{p} \to \infty} f' = 0, \]  

(B14)

\[ \lim_{\tilde{p} \to \infty} f'' = 0. \]  

(B15)

Consequently, since \( f''' \) is positive and \( \lim_{\tilde{p} \to \infty} f'' = 0 \), \( f'' \) must be strictly negative. So, \( f' \) is a monotonically decreasing function of \( \tilde{p} \). Since this also vanishes at infinity, it must be strictly positive. Hence \( f \) is a monotonically increasing function. As a consequence, \( f \) and by extension \( J \) are strictly positive, and the inverse function theorem therefore implies that \( \tilde{F} \) (and thus \( F \)) is invertible.

2. Plots for Kerr

The general expressions for \((\Upsilon_r, \Upsilon_\theta, \Upsilon_\phi)\) in terms of \((p, e, z_{max})\) in a Kerr spacetime are known \[18, 19\]. However, they are complicated enough that a direct computation of the Jacobian seems intractable, and we are not going to attempt it here. Instead we use the explicit expressions to numerically plot parameter space of bound orbits in Kerr spacetime as parametrized by the Mino frequencies.

Figures [5] and [6] show this parameter space for respectively prograde and retrograde orbits in a Kerr spacetime with spin \( a = \frac{9}{10} M \). The separatrix dividing bound and unbound orbits is given by the surface \( \Upsilon_r = 0 \). The plotted lines in the plot keep two of the original parameters \((p, e, z_{max})\) constant. A failure of invertibility of \( F \), would manifest itself as this grid degenerating. A similar plot for the frequencies with respect to Boyer-Lindquist coordinate time, easily reveals the isofrequency region where the grid folds back on itself.

A particularly strong clue that helps us think that there may be no Mino time isofrequency pairs, is that parabolic \((e = 1)\) orbits have a finite radial frequency. Hence, the
Figure 5. Parameter space of prograde bound orbits in a Kerr space time with $a = 0.9$.

Figure 6. The same parameter space as in figure 5, but for retrograde orbits.

separatrix and the plane of parabolic orbits cannot get (partly) mapped into each other as happens for coordinate and proper time frequencies. This more or less restricts to possibility of Mino time isofrequency pairings to the interior of the plots in figures 5 and 6 (where it also does not appear to occur).

The absence of any visible degeneration in these plots, is a strong indication that the Mino time frequencies are a good set of parameters for bound orbits in Kerr. At least, in the plotted region which contains the low integer ratios of $\Upsilon_r$ and $\Upsilon_\theta$. This, by no means, is a proof that there are no Mino isofrequency pairs of bound orbits in Kerr spacetimes, but it is good enough for us to assume this as a conjecture.

Appendix C: Near identity averaging transformation

In this appendix we describe the details of the near-identity averaging transformation needed to remove the non-resonant oscillatory terms from the equations of motion (8). It closely follows the procedure described in section 5.1 of [21]. Recall the equations of motion (8) for an EMRI system,

\[
\frac{d\Upsilon_i}{d\lambda} = \epsilon G_i(\vec{\Upsilon}, \vec{q}) + O(\epsilon^2), \quad (C1a)
\]

\[
\frac{dq_j}{d\lambda} = \Upsilon_j + \epsilon g_j(\vec{\Upsilon}, \vec{q}) + O(\epsilon^2), \quad (C1b)
\]

with

\[
G_i(\vec{\Upsilon}, \vec{q}) = G_i(\vec{\Upsilon}) + \sum_{N \neq 0} G_i, N(\vec{\Upsilon}) e^{iNq_z} + \sum_{(n, k) \in \mathbb{R}} g_{i, nk}(\vec{\Upsilon}) e^{inr + ikq} \quad (C2)
\]

\[
g_j(\vec{\Upsilon}, \vec{q}) = g_j(\vec{\Upsilon}) + \sum_{N \neq 0} g_{j, N}(\vec{\Upsilon}) e^{iNq_z} + \sum_{(n, k) \in \mathbb{R}} g_{j, nk}(\vec{\Upsilon}) e^{inr + ikq}. \quad (C3)
\]

We introduce a change of variables that is an identity transformation at leading-order (hence the term “near-identity”),

\[
\vec{\Upsilon}_i(\vec{T}, \vec{q}) = \Upsilon_i + \epsilon T_i(\vec{T}, \vec{q}) + O(\epsilon^2), \quad (C4a)
\]

\[
\vec{q}_j(\vec{T}, \vec{q}) = q_j + \epsilon L_j(\vec{T}, \vec{q}) + O(\epsilon^2). \quad (C4b)
\]

The inverse transformation is given by

\[
\Upsilon_i(\vec{T}, \vec{q}) = \vec{T}_i - \epsilon T_i(\vec{T}, \vec{q}) + O(\epsilon^2), \quad (C5a)
\]

\[
q_j(\vec{T}, \vec{q}) = \vec{q}_j - \epsilon L_j(\vec{T}, \vec{q}) + O(\epsilon^2). \quad (C5b)
\]

To obtain the equations of motion for the new variables $\vec{T}$ and $\vec{q}$ we first differentiate $[C4]$ with respect to $\lambda$,

\[
\frac{d\vec{T}_i}{d\lambda} = \frac{dT_i}{d\lambda} + \epsilon \left\{ \frac{dT_i}{d\lambda} \frac{d\Upsilon_k}{d\lambda} + \frac{dT_k}{d\lambda} \frac{dq_k}{d\lambda} \right\} + O(\epsilon^2), \quad (C6a)
\]

\[
\frac{d\vec{q}_j}{d\lambda} = \frac{dq_j}{d\lambda} + \epsilon \left\{ \frac{dL_j}{d\lambda} \frac{d\Upsilon_k}{d\lambda} + \frac{dL_k}{d\lambda} \frac{dq_k}{d\lambda} \right\} + O(\epsilon^2). \quad (C6b)
\]
If we then substitute the equations of motion \([C1]\), and use the inverse transformation \([C5]\) to eliminate the dependence on \(\bar{T}\) and \(\bar{q}\), we obtain,

\[
\frac{d\tilde{T}_i}{d\lambda} = \epsilon \{ G_i(\bar{T}, \bar{q}) + \frac{dT_i}{dq_k} \gamma_k \} + O(\epsilon^2), \\
\frac{d\tilde{q}_j}{d\lambda} = \tilde{T}_j + \epsilon \{ g_j(\bar{T}, \bar{q}) + \frac{dL_j}{dq_k} \gamma_k - T_j \} + O(\epsilon^2).
\]

The idea is to use the freedom in the functions \(T\) and \(L\) to eliminate the lowest order non-resonant oscillatory terms from the equations of motion. That is we want to set,

\[
\frac{dT_i}{dq_k} \gamma_k = - \sum_{(n,k) \in R} G_{i,nk}(\bar{T}) e^{in\varphi - ik\varphi}, \\
\frac{dL_j}{dq_k} \gamma_k = \tilde{T}_j - \sum_{(n,k) \in R} g_{j,nk}(\bar{T}) e^{in\varphi - ik\varphi},
\]

where \(\tilde{T}_i\) denotes the terms of \(T_i\) that depend on \(\bar{q}\). This set of first-order ODEs can easily be solved by direct integration, yielding

\[
T_i(\bar{T}, \bar{q}) = \tilde{T}_i(\bar{T}) + i \sum_{(n,k) \in R} \frac{G_{i,nk}(\bar{T})}{n\tilde{T}_r + k\tilde{T}_\theta} e^{in\varphi - ik\varphi}, \\
L_j(\bar{T}, \bar{q}) = \tilde{L}_j(\bar{T}) + i \sum_{(n,k) \in R} \frac{g_{j,nk}(\bar{T})}{n\tilde{T}_r + k\tilde{T}_\theta} e^{in\varphi - ik\varphi} + \sum_{(n,k) \in R} \frac{G_{j,nk}(\bar{T})}{(n\tilde{T}_r + k\tilde{T}_\theta)^2} e^{in\varphi - ik\varphi},
\]

and consequently

\[
\tilde{T}_i(\bar{T}, \bar{q}) = i \sum_{(n,k) \in R} \frac{G_{i,nk}(\bar{T})}{n\bar{T}_r + k\bar{T}_\theta} e^{in\varphi - ik\varphi},
\]

The appearance of the combination \(n\bar{T}_r + k\bar{T}_\theta\) in the denominator of the solution, immediately tells why this procedure cannot be used to remove the resonant oscillatory terms, since these would become singular at resonance.

The solution \([C9]\) contains the arbitrary functions \(\tilde{T}_i\) and \(\tilde{L}_j\) of \(\bar{T}\). In principle, we could set them to zero, since our objective of removing the non-resonant oscillatory terms has been achieved. However, as explained in section 5.1 of [21], this freedom can be used to make further simplifications to the equations of motion. In [21] this freedom is used to eliminate the \(O(\epsilon^2)\) averaged terms. We make a slightly different choice, we set \(\tilde{T}_i(\bar{T}) = \bar{g}_i(\bar{T})\), which eliminates the \(O(\epsilon)\) averaged term from the \(\bar{q}\) equations of motion, while we use \(\tilde{L}_j(\bar{T})\) as in [21] to eliminate the \(O(\epsilon^2)\) averaged terms from those same equations.

The equations of motion to \(O(\epsilon^2)\) resulting from these manipulations are

\[
\frac{d\tilde{T}_i}{d\lambda} = \epsilon \bar{G}_i(\bar{T}) + \epsilon \sum_{N \neq 0} \frac{G_{i,N}(\bar{T})}{n\tilde{T}_r + k\tilde{T}_\theta} e^{iN\varphi} + O(\epsilon^2); \\
\frac{d\tilde{q}_j}{d\lambda} = \tilde{T}_j + \epsilon \sum_{N \neq 0} \frac{g_{j,N}(\bar{T})}{n\tilde{T}_r + k\tilde{T}_\theta} e^{iN\varphi} + O(\epsilon^2).
\]

References:

[1] J. R. Gair, L. Barack, T. Creighton, C. Cutler, S. L. Larson, et al., Classical Quant. Grav. 21, S1595 (2004), arXiv:gr-qc/0405137 [gr-qc].

[2] J. R. Gair, M. Vallisneri, S. L. Larson, and J. G. Baker, Living Rev. Rel. 16, 7 (2013), arXiv:1212.5575 [gr-qc].

[3] L. Barack and C. Cutler, Phys. Rev. D 69, 082005 (2004), arXiv:gr-qc/0310125 [gr-qc].

[4] L. Barack, Classical Quant. Grav. 26, 213001 (2009), arXiv:0908.1664 [gr-qc].

[5] E. Poisson, A. Pound, and I. Vega, Living Rev. Rel. 14, 7 (2011), arXiv:1102.0529 [gr-qc].

[6] E. E. Flanagan and T. Hinderer, Phys. Rev. Lett. 109, 071102 (2012), arXiv:1009.4293 [gr-qc].

[7] Y. Mino, Prog. Theor. Phys. 113, 733 (2005), arXiv:gr-qc/0506003 [gr-qc].

[8] T. Tanaka, Prog. Theor. Phys. Suppl. 163, 120 (2006), arXiv:gr-qc/0508114 [gr-qc].

[9] K. Grossman, J. Levin, and G. Perez-Giz, Phys. Rev. D 85, 023012 (2012), arXiv:1105.5811 [gr-qc].

[10] J. Gair, N. Yunes, and C. M. Bender, J. Math. Phys. 53, 032903 (2012), arXiv:1111.3605 [gr-qc].

[11] E. E. Flanagan, S. A. Hughes, and U. Rangarsri, Phys. Rev. D 89, 084028 (2014), arXiv:1208.3906 [gr-qc].

[12] U. Rangarsri and S. A. Hughes, Phys. Rev. D 89, 084036 (2014), arXiv:1307.6483 [gr-qc].

[13] S. Isoyama, R. Fujita, H. Nakano, N. Sago, and T. Tanaka, Prog. Theor. Exp. Phys. 2013, 063E01 (2013), arXiv:1302.4035 [gr-qc].

[14] J. Brink, M. Geyer, and T. Hinderer, (2013), arXiv:1304.0330 [gr-qc].

[15] B. Carter, Phys. Rev. 174, 1559 (1968).

[16] Y. Mino, Phys. Rev. D 67, 084027 (2003), arXiv:gr-qc/0302075 [gr-qc].

[17] T. Hinderer and E. E. Flanagan, Phys. Rev. D 78, 064028 (2008), arXiv:0805.3337 [gr-qc].

[18] W. Schmidt, Classical Quant. Grav. 19, 2743 (2002), arXiv:gr-qc/0202090 [gr-qc].

[19] R. Fujita and W. Hikida, Classical Quant. Grav. 26, 135002 (2009), arXiv:0906.1420 [gr-qc].

[20] N. Warburton, L. Barack, and N. Sago, Phys. Rev. D 87, 084012 (2013), arXiv:1301.3918 [gr-qc].

[21] J. Kevorkian and J. D. Cole, Multiple scale and singu-
lar perturbation methods, Vol. 114 (Springer New York, 1996).

[22] T. Hinderer and E. E. Flanagan, “Two timescale analysis of extreme mass ratio inspirals in Kerr. II. Transient Resonances,” (unpublished).

[23] D. L. Bosley and J. Kevorkian, Stud. Appl. Math. 94, 83 (1995).

[24] R. Haberman, SIAM J. Appl. Math. 43, 244 (1983).