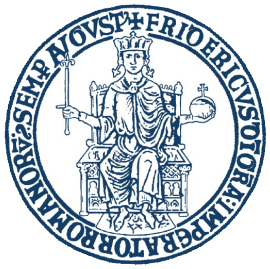

UNIVERSITÀ DEGLI STUDI
DI NAPOLI FEDERICO II

Scuola Politecnica e delle Scienze di Base

Area Didattica di Scienze Matematiche, Fisiche e Naturali

Dipartimento di Fisica "Ettore Pancini"

Laurea Magistrale in Fisica

Gravitational self-force in extreme-mass-ratio inspirals

Conservative effects as a testing ground for theories of gravity

Relatore
Prof. Mariafelicia De Laurentis

Candidato
Vincenzo Ventriglia
Matr. N94000440

Anno Accademico 2019/2020

There's so many different worlds

So many different suns

And we have just one world

But we live in different ones

Contents

List of Figures

Introduction

Homo sum, humani nihil a me alienum puto.
(Publius Terentius Afer, *Heautontimorùmenos*, 165 BC)

The Nobel-prize laureate Chandrasekhar expressed the simplicity of the mathematical description of a rotating black hole saying that

> "Rotating black holes are the most perfect macroscopic objects in the Universe. And, as the theory of General Relativity provides a single, unique, two-parameter solution for their description, they are also the simplest objects."

Nevertheless, two such simple black holes together are already too difficult to handle, as they form a remarkable complex dynamical system for which only approximate methods allow us to grasp the physics behind it. Even in scenarios where our approximation and idealisation can make sense in some effective way, the orbital dynamics remains truly complicated. The detailed description of the entire process will involve inspiral and merger phases, and it has been a key theme in gravitational research for almost 50 years, driven by the ambition to predict the pattern of gravitational waves through which such systems can be observationally studied.

The landmark observation of merging black holes by the LIGO-Virgo Collaboration in 2015 [1] has conclusively brought black holes, inspiraling black-hole binaries and gravitational waves in the realm of physics rather than mathematical speculation, marking the birth of *gravitational-wave astronomy*. The LIGO-Virgo discoveries would not have been possible without an accurate model of the inspiral and merger; actually their analysis [2] concluded that the quality of science extractable from future observations may well be limited not by experimental precision, but by the accuracy of available theoretical models.

One of the next steps in the gravitational astronomy programme is the observation of an inspiral scenario where one of the black holes is much lighter than the other one – the so-called "extreme-mass-ratio inspiral", or EMRI. Nature seems to abound with EMRIs, which are expected to emit gravitational waves in millihertz frequencies, impossible to be detected by Earth-borne detectors. Indeed, EMRIs are prime targets for the planned LISA mission, whose peak sensitivity will be exactly in the millihertz band. The will to hunt these sources of gravitational waves is explained by their nature as an extraordinary laboratory for strong-gravity physics.

Our journey will begin in chapter 1 with an introduction to the astrophysical background and the primary scientific motivations that bring our attention to EMRIs. We will explore the detectability of these sources at LISA and illustrate how they bring an immediate problem in the study of their dynamics, making the development of new approaches to the Einstein equations absolutely essential. The final focus of the chapter will be on the interplay of the gravitational self-force technique with the other methods developed so far for tackling the evolution problem of binaries.

Once we have defined the physical background in which we will move, in chapter 2 we will go on setting out the technical basis and developments of the self-force method. The discussion will start from the early works at the end of the 1990s, passing through the more modern reformulations and introducing the issue concerning the implementation of numerical methods. We will thus have set the stage to discuss the orbital evolution of EMRIs, capitalising on the adiabatic nature of the inspiral process. Finally, we turn to a discussion of the radiative effects of the gravitational self-force.

In chapter 3 we will pause for a moment the main discussion in order to motivate and introduce the concept of Extended Theories of Gravity. What is the point of considering them and what theories should we consider in the vast panorama of possible extensions of a theory so praised as General Relativity? We will first try to answer these queries and then deal with the analysis of the self-force

in scalar-tensor gravity, finally specialising our discussion to the case where both objects in question are black holes.

The analysis of the conservative effects of the self-force will then engage us throughout chapter 4. The reason why we delayed the discussion at this point is that we will be able to examine these effects first in General Relativity, and then move on to an "effective" theory – relying on a theory-agnostic parametrisation – that takes into account the corrections induced by extended theories, as discussed in the previous chapter.

Throughout this work, the use of geometrized units – in which $G = c = 1$ – is implied, unless otherwise specified. In this system all units are expressed in terms of powers of length, as it is customary in the literature of General Relativity.

1. Scientific background and motivations

How many bodies are required before we have a problem?

In 18th-century Newtonian Mechanics, the 3-body problem was insoluble. With the birth of General Relativity around 1910 and Quantum Electrodynamics in 1930, the 2- and 1-body problem was insoluble. And, within modern Quantum Field Theory, the problem of zero bodies (vacuum) is insoluble. So, if we are out after exact solutions, no bodies at all is already too many!

(Richard Mattuck, 1976)

When two black holes (BHs) are in orbit around each other, they form a strikingly complex dynamical system. No closed-form solutions are known, and even numerical solutions have been forbiddingly hard to obtain until well into the 21st century. It is true beyond a shadow of a doubt that the complexity of the gravitational two-body problem in General Relativity (GR) stands in stark contrast to its elementary nature in the context of point-particle Newtonian gravity, where all possible orbital configurations are simple conical sections. First, point-particle idealisation is problematic in GR; second, a gravitationally bound system of two masses admits no stationary configurations, as gravitational waves (GWs) constantly carry orbital energy away from the system and back-reaction from that radiation gradually drives the two objects closer together, and, given enough time, the two bodies will eventually merge. If these bodies are Kerr black holes, a single, larger Kerr black hole eventually forms.[1]

Observational evidence indicates that the centres of most galaxies contain supermassive black holes (SMBHs), with masses between a few tens of thousands and a few billion solar masses. Mergers involving SMBHs are powerful sources of GWs, but these can only be observed from space: that is because of the seismic noise limiting the sensitivity of ground-based detectors to frequencies above approx. 1 Hz, and hence with total mass no greater than a few hundred solar masses.

The *Laser Interferometer Space Antenna* (LISA) will have sensitivity to GWs in the millihertz band generated by merging systems having total mass in the range $10^4 - 10^7\ M_\odot$, and is expected to observe a great variety of sources. However, the primary source for LISA are the SMBHs in the centres of galaxies, that generate GWs either when they merge with other SMBHs – which is expected to occur following mergers between their host galaxies – or when they merge with much smaller, stellar-mass compact objects – such as white dwarfs, neutron stars or stellar-mass BHs.[2] The latter systems are called *extreme-mass-ratio inspirals* (EMRIs), because of the large difference in mass between the two objects involved.

EMRIs describe the long-lasting inspiral (from months to a few years) and plunge of stellar-origin BHs, with mass of $10 - 60\ M_\odot$, into SMBHs of $10^5 - 10^6\ M_\odot$ in the centre of galaxies. The orbits of EMRIs are expected to be generic and highly relativistic. These events, which constitute one of the eight science objectives of LISA [3], are tremendously interesting because, depending on the chirp mass of the system, the signal can stay in the LISA bandwidth for years; the small object may then spend many cycles in close vicinity of the SMBH, with its orbit displaying extreme forms of periastron and orbital plane precession. The intricate GW signal cleanly encodes within it a detailed mapping of the spacetime around the SMBH, allowing us to test its geometry and providing accurate measurements

[1] Recent direct observations of the gravitational wave emission from a binary black hole merger indicate that the resulting black hole remnant of a binary black hole merger is in general boosted along a particular direction with respect to the asymptotic Lorentz frame at null infinity where such emissions have been detected. In this general case an additional parameter – the boost parameter – has to be added, so that the solution has three parameters (mass, spin and boost) and corresponds to the most general configuration that an astrophysical black hole can have, reducing to the standard Kerr solution when the boost parameter is zero. [7]

[2] A normal star would be tidally disrupted by the central BH.

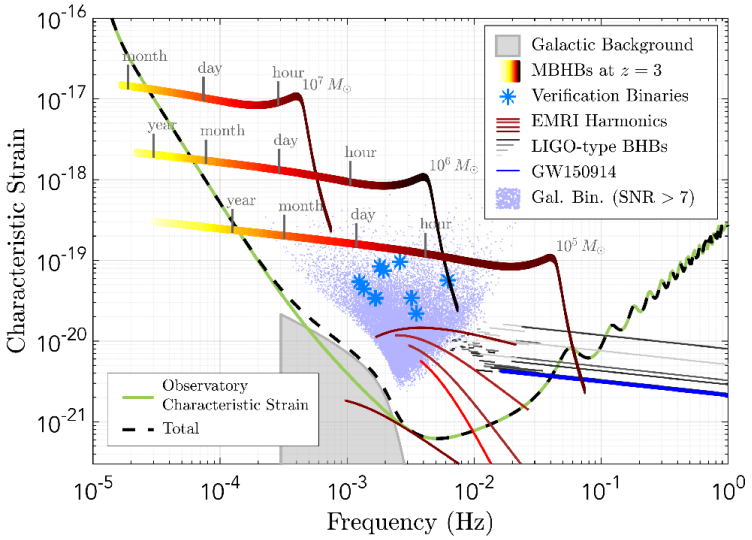

Figure 1.1.: Examples of GW sources in the frequency range of LISA, compared with its sensitivity for a 3-arm configuration. [3]

of its mass and spin, confirm whether it is a BH as GR predicts and eliminate or tightly constrain a heap of proposed alternatives to GR: an "ideal experiment" of Nature in strong gravity.

1.1. LISA mission

Space-based GW interferometry will open the low-frequency window (mHz to Hz) in the GW landscape, which is complementary to *Earth-based* detectors (Hz to kHz) and *pulsar timing array* experiments (nHz).

The proposed LISA mission [3] enables the detection of GWs from SMBHs coalescences within a vast cosmic volume encompassing all ages, from cosmic dawn to the present, across the epochs of the earliest quasars and of the rise of galaxy structure. LISA will provide the finest possibility to map the spacetime around the SMBHs which populate the centres of galaxies, using compact objects as particle-like probes; this great science objective will be achieved using three arms and three identical spacecrafts in a triangular formation (constellation) in a heliocentric orbit trailing the Earth by about 20° – between 50 and 65 million km from Earth. The expected sensitivity and some potential signals, including EMRIs, are shown in fig. 1.1. An observatory that is expected to deliver fine-detailed science is described by a sensitivity curve which, below 3 mHz, will be limited by acceleration noise at the level demonstrated by LISA Pathfinder; interferometry noise instead dominates above 3 mHz. Such a sensitivity can be achieved with a mean 2.5 million km arm-length constellation (the inter-spacecraft separation) with 30 cm telescopes and 2 W laser systems.

GWs change the optical path-length along the three sides of a triangular configuration defined by free-falling test masses. The test masses will follow their geodesic trajectories with sub-femto $g/\sqrt{\text{Hz}}$ spurious acceleration and will be located inside the three identical co-orbiting drag-free spacecrafts. Laser interferometers, all-sky monitors of GWS, will measure the pm $= 10^{-12}$ m to nm $= 10^{-9}$ m path-length variations caused by GWs. These distance changes are small compared to the variations caused by solar system celestial dynamics (some 10^4 km), but they can be distinguished because the former are at mHz frequencies (timescale of 10^3 s), whereas the latter are quiet at mHz frequencies (with periods of many months).

The constellation is fully symmetric, with similar measurements taking place in both direction along each of the three arms. Three independent interferometric combinations of the light travel time measurements between the test masses are possible, allowing the synthesis of two virtual Michelson interferometers (allowing the simultaneous measurement of the two polarisations of the GW) plus a third, "Sagnac" configuration (insensitive to GWs, used to characterise the instrumental noise background).

The yearly rotation of the constellation about itself and its orbit around the Sun allows to reconstruct the source direction for sources that can be observed for, at least, several weeks. Noisy non-gravitational interactions acting on the spacecrafts require the use of test masses as geodesic reference: these test particles are actually shielded by the containing spacecrafts. Two test masses per spacecraft are used, each one dedicated to a single interferometry arm.

1.1.1. EMRIs@LISA

When the orbit as a function of time is not a simple harmonic motion, we do not expect the radiated power to be monochromatic. In fact, computing the frequency spectrum of the radiated power, one finds that the circular orbit value $P_\circ$ gets replaced by [4]

$$P_\circ = \frac{32G^4\mu^2m^3}{5c^5a^5} \longrightarrow P = \sum_{n=1}^{+\infty} P_n = \frac{32G^4\mu^2m^3}{5c^5a^5}\sum_{n=1}^{+\infty} g_n\left(e\right),$$

where $g_n\left(e\right)$ depends on the eccentricity e through Bessel functions in such a way that, when $e=0$, one gets back the result for circular motion, *i.e.* $g_n\left(0\right)=\delta_{n2}$. For a generic value of $0<e<1$ all harmonics contribute, and radiation at all frequencies $\omega_n = n\omega_0$ appears for all integer values $n \geq 1$. Increasing the ellipticity, increase also the value of $n=\bar{n}$ where $g_n\left(e\right)$, as well as P_n, is maximum and the total radiated power. Furthermore, for a generic orbit one has three fundamental frequencies – corresponding to radial, azimuthal and polar motion[3] – with the corresponding harmonics. The signal then has a set of harmonics, which in an inspiraling binary evolve simultaneously in time, giving in principle a very specific signature.

A raw EMRI signal will have an instantaneous amplitude an order of magnitude below the instrumental noise of LISA and (at low frequencies) several orders of magnitude below the GW foreground from galactic compact binaries [5], making detection a rather difficult problem to address. However, the signals are very long lived and will be observed over at least 10^4 cycles, which in principle allows the signal-to-noise ratio (SNR) to be built up over time using matched filtering. One can estimate, for instance, that a binary with

$$M = 10^6\ M_\odot \qquad \text{and} \qquad m = 10\ M_\odot,$$

at $z \simeq 1.2$, spends about 10^5 inspiral cycles in the LISA bandwidth. This large number of cycles gives the possibility of extracting a small signal from the noise, and allows in principle extremely precise measurements of the parameters of the system. Hence, EMRIs can bring a wealth of information. However, the computational problem becomes formidable: we need templates able to track the signal accurately enough for $10^4 - 10^5$ cycles. This not only requires very accurate theoretical waveforms, but also the scanning of a very fine grid in the parameter space of the binary. Furthermore, their orbits are quite relativistic, with generic inclination with respect to the spin of the central BH, and display extreme forms of periastron and orbital precession, therefore providing a real "map" of the (Kerr) geometry produced by the central SMBH. The important lesson we are taught is that, in order to gain access to this opulence, accurate templates for the sources must be developed.

The theoretical uncertainty in the rate of formation of EMRIs is relatively large – typical estimates are of order $10^{-8}-10^{-6}$ yr^{-1} for a central SMBH with mass $M=10^6\ M_\odot$, corresponding to a detection rate for LISA ranging from a few to hundreds of events per year (see [4, 5] and references therein). Captures occur when two objects in the dense stellar cusp[4] surrounding a galactic SMBH undergo a close encounter, sending one of them into an orbit tight enough that the orbital decay through emission of gravitational radiation dominates the subsequent orbital evolution. The latter may be qualitatively divided into three stages: in the first and longest stage, the orbit is extremely eccentric and GWs

[3] Stated differently, there are (*i*) the orbital frequency, (*ii*) the perihelion precession frequency and (*iii*) the frequency of precession of the orbital plane.

[4] The evolution of a star cluster containing a SMBH has been (robustly, but only theoretically) proven to admit a quasi-steady state solution, where the stellar density takes a power-law form $\rho \propto r^{-\gamma}$, the so-called stellar density *cusp*, fully developed after a *relaxation time*, necessary for the randomisation of the cluster phase space via close encounters between stars.

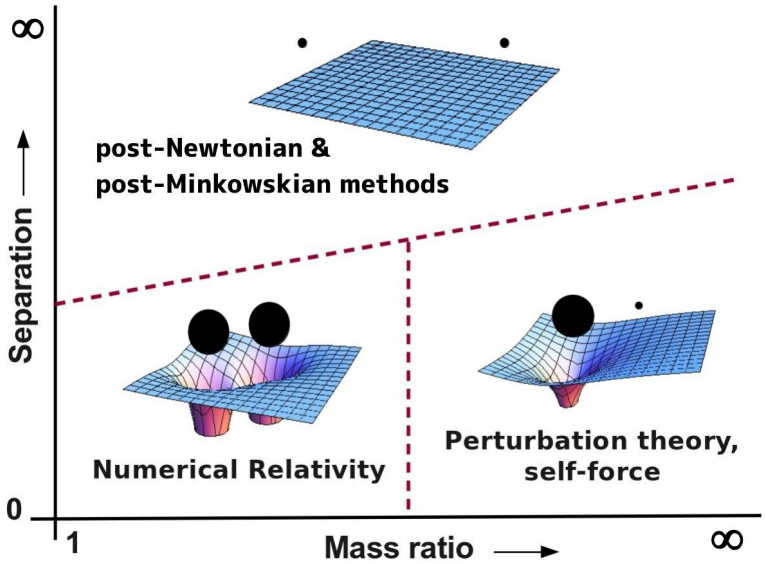

Figure 1.2.: Domains of the two-body problem in GR. Different principle lead to different methods: PN expands about flat spacetime, GSF expands about the exact field geometry of the central BH, while NR digs into the full non-linear dynamics. Synergistic work seeks to interface between these results [21].

are emitted in pulses which slowly remove energy and angular momentum from the system, and the orbit gradually shrinks and circularises; after $10^3 - 10^8$ years (depending on the two masses and the initial eccentricity), the evolution enters its second stage, when the orbit is sufficiently circular so that the emission becomes approximately continuous; finally, the adiabatic inspiral transitions to a direct plunge, as the object reaches the last stable orbit, quickly plunges through the SMBH horizon and the GW signal cuts off. We will talk more extensively about these stages in subsection 4.2.1. While individually-resolvable captures will mostly be detectable during the last $1 - 100$ years of the second stage, radiation emitted during the first stage (mostly in the form of short bursts near periastron passages) will contribute significantly to the confusion background [6]. The detection will be further complicated by the fact that about 10^7 galactic white-dwarf binaries generate a stochastic background that is above the LISA noise floor for frequencies below $2 - 3$ mHz. One also expects that there will be a stronger signal form the merger of a few SMBH binaries, and all these signals will be superimposed in time and frequency. Therefore, one needs techniques for separating the various signals, removing the stronger source before being able to see the signals due to EMRIs.

1.2. Waveforms and synergies

The GW strain signal, $h(t)$, called *waveform*, together with its frequency domain representation, $\tilde{h}(\nu)$, encodes exquisite information about both intrinsic (*e.g.*, mass and spin) and extrinsic (*e.g.*, inclination, luminosity distance, location) parameters of the source. The desire to maximise the science return from GW experiments drives the theory programme to improve waveform models across the full parameter space relevant to observation. However, the theoretical computation of templates able to follow the actual waveforms with high accuracy over $10^4 - 10^5$ cycles is a highly non-trivial issue.

For EMRIs, the post-Newtonian (PN) expansion is not adequate, since EMRIs are quite relativistic and it would converge too slowly. Numerical relativity (NR) is also not suitable, since relatively long numerical simulations are currently only possible for mass ratios up to about 5-10, due to the numerical complications introduced by the different length-scales associated with the two BHs – and, even in that case, they cover "only" $O\left(10^2\right)$ inspiral cycles. One can however make use of the fact that in EMRIs the mass ratio is small, *e.g.* of order $10^{-4} - 10^{-6}$, so one can perform an expansion in this parameter.

A technique that is particularly suitable in the case of EMRI systems is the self-force approach, which we will be dealing with throughout this work. As will become more clear in the following,

a comparison of results from different techniques provide a powerful overall check and mutual tests for different methods, as each relies on subtle procedures and computational manipulations. But such comparison can achieve more than just mutual tests: the gravitational self-force (GSF) and PN methods are both systematic approximation approaches to the two-body problem in GR, each based on a perturbative expansion in a different limiting domain of the problem. In fact, the PN method expands about the limit of large separation keeping the mass ratio η arbitrary, while the GSF method expands about the limit $\eta \to 0$ keeping the separation arbitrary. Joining forces is useful in different ways, as the results of one method can become a benchmark for the other, or can help predicting higher order terms that are difficult to determine. But what is perhaps best is the chance to enhance the science return by exploiting synergies and "sew" predictions in intermediate domains, which may not be accessible to either of the approximation methods separately.

The synergistic approach shall be expanded to include NR, which enables us to directly solve the fully non-linear problem describing the inspiral and merger dynamics of BH binaries. NR simulations result in the best tool to model the two-body dynamics of two comparable-mass BHs close to their merger, but unfortunately become computationally prohibitive when the separation is large or the mass ratio too small ($\lesssim 1:10$, in practice).

The idea of a three-fold synergy is depicted in fig. 1.2, showing the respective domain of applicability of each of the three methods in the essential parameter space of the two-body problem, *i.e.* mass ratio *vs.* separation. Exploring the interfaces between the approaches – and reaching across them – is often done within the framework of the effective-one-body (EOB) approach, which aims to provide a universal, semi-analytical model of the two-body dynamics across its entire parameter space. For a full review on the overlap between the various approaches to the binary BH problem, refer to [29].

2. Gravitational self-force

Es ist nicht das Ziel der Wissenschaft, der unendlichen Weisheit eine Tür zu öffnen, sondern eine Grenze zu setzen dem unendlichen Irrtum.

(Bertolt Brecht, *Leben des Galilei*, 1938)

The first encounter with the concept of "self-force", through which an object does "feel its own field", usually happens in the case of a Galilean-relativistic accelerating charge in flat spacetime. In fact, when actually taking into account the change in the field due to the acceleration of the particle, one finds that the motion obeys the Abraham-Lorentz equation [21, 8]

$$m\frac{\mathrm{d}^2\boldsymbol{z}}{\mathrm{d}t^2} = \boldsymbol{f}_{\mathrm{ext}} + \frac{2e^2}{3m}\frac{\mathrm{d}\boldsymbol{f}_{\mathrm{ext}}}{\mathrm{d}t}.$$

The second term is a self-force; more specifically, it is a radiation-reaction. Unlike the Coulomb potential of the static charge, the Liénard-Wiechert potentials of the accelerating charge contain an unbound piece [9], which carries energy-momentum out to infinity in the form of radiation (dissipation). That emission causes a recoil, pushing the particle in the opposite direction.

The Abraham-Lorentz equation does not evince a particularly direct relationship between the self-force and the field of the particle. In order to obtain a more physically compelling picture and to understand GSF physics, it is instructive to start from the back-reaction problem for a special-relativistic electric charge, moving in a Minkowski spacetime under the influence of some external force.

2.1. Electromagnetic self-force in flat spacetime

A typical situation would consist of an electron moving in a bound orbit in the field created by a positive charge, taken to be very massive. The motion of the electron generates an electric current j^μ, which sources the electromagnetic (EM) field. In the Lorenz gauge, $\partial_\mu A^\mu = 0$, the EM field satisfies the equation

$$\Box A^\mu = -4\pi j^\mu. \tag{2.1}$$

If we assume that the electron is point-like, the current j^μ has a support only on the electron's worldline. Given the trajectory of the electron, j^μ induces radiative solutions at infinity for A^μ in eq. (2.1). These out-going EM waves carry away energy and angular momentum; the electron will therefore lose energy and angular momentum, and will inspiral onto the central massive charge, at least as long as the classical description is still appropriate. In other words, the EM field generated by the electron acts back on the electron itself, producing a back-reaction force, which should be included in the equations of motion of the electron, at least iteratively.

When solving eq. (2.1) to compute the electromagnetic waves produced at infinity, one uses the retarded Green's function, leading to the solution

$$A^\mu_{\mathrm{ret}}(x) = \int \mathrm{d}^4x' G_{\mathrm{ret}}(x, x')\, j^\mu(x').$$

A problem that immediately arises in the computation of the self-force on the electron is that the field $A^\mu_{\mathrm{ret}}(x)$ generated by the electron is singular at the position of the electron itself, so one might fear that the self-force is also divergent. But that's not the case. The situation considered here, in which the radiation is propagating outward and the charge is spiraling inward, breaks the *time-reversal invariance* of Maxwell's theory: a specific time direction was adopted when, among all possible solutions to the wave equation, we chose the retarded one as the physically relevant solution. Specifically, observation

of the motion of a single particle could not give an indication as to the direction of time – only the statistical behaviour of systems with a large number of degrees of freedom single out a direction. The asymmetry is therefore introduced by the, somewhat arbitrary, rejection of the advanced solution. Reflecting the time-symmetry of the laws, any field can be described equivalently by an initial *or* final value problem (with arbitrary boundary conditions). Hence one can always decompose the field as

$$A^\mu_{\rm ret}(x) = \frac{1}{2}\left[A^\mu_{\rm ret}(x) + A^\mu_{\rm adv}(x)\right] + \frac{1}{2}\left[A^\mu_{\rm ret}(x) - A^\mu_{\rm adv}(x)\right]. \tag{2.2}$$

This decomposition tells us something very important. The linear superposition

$$A^\mu_{\rm S} \equiv \frac{1}{2}\left[A^\mu_{\rm ret} + A^\mu_{\rm adv}\right]$$

restores time-reversal invariance: outgoing and incoming radiation would be present in equal amounts, there would be no net loss nor gain of energy and momentum by the system, and the charge would undergo no radiation reaction. While $A^\mu_{\rm S}$ does not exert a force on the charged particle, it is just as singular as the retarded field in the vicinity of the world line. This follows from the fact that $A^\mu_{\rm ret}$, $A^\mu_{\rm adv}$ and $A^\mu_{\rm S}$ all satisfy eq. (2.1), whose source term is infinite on the worldline: despite different wave-zone behaviours, they share the same singular behaviour near the worldline (dominated by the particle's Coulomb field). So, the subscript "S" stands for "symmetric" as well as "singular".

Because $A^\mu_{\rm S}$ is just as singular as $A^\mu_{\rm ret}$, removing it from the retarded solution gives rise to a well behaved field in a neighbourhood of the worldline; and, because $A^\mu_{\rm S}$ is known not to affect the motion of the charged particle, this new field must be entirely responsible for the radiation reaction. We therefore introduce the field

$$A^\mu_{\rm R} \equiv A^\mu_{\rm ret} - A^\mu_{\rm S} = \frac{1}{2}\left[A^\mu_{\rm ret} - A^\mu_{\rm adv}\right],$$

which coincides with the second term in the decomposition (2.2).[1] This field satisfies the homogeneous version of eq. (2.1), $\Box A^\mu_{\rm R} = 0$, so there is no singular source to produce a singular behaviour on the worldline. Furthermore, satisfying the homogeneous wave equation, $A^\mu_{\rm R}$ can be thought of as a free radiation field, and the subscript "R" stands for "radiative" as well as "regular".

The self-action of the charge own field can then be understood as follows: a singular field $A^\mu_{\rm S}$ can be removed from the retarded field and shown not to affect the motion of the particle; what remains is a well behaved field $A^\mu_{\rm R}$ that must be solely responsible for the radiation reaction. From the regular field we form an electromagnetic field tensor, $F^{\rm R}_{\mu\nu} = \partial_\mu A^{\rm R}_\nu - \partial_\nu A^{\rm R}_\mu$, and we take the equations of motion (EoMs) of the particle to be

$$m a_\mu = f^{\rm ext}_\mu + e F^{\rm R}_{\mu\nu} u^\nu,$$

where $u^\mu \equiv {\rm d}z^\mu/{\rm d}\tau$ is the charge four-velocity, $a^\mu = {\rm d}u^\mu/{\rm d}\tau$ its acceleration, m its (renormalised[2]) mass and $f^{\rm ext}_\mu$ an external force potentially acting on the particle. Computing the regular field on the worldline yields the Abraham–Lorentz–Dirac equation[3]

$$m a^\mu = f^\mu_{\rm ext} + \frac{2e^2}{3m}\left(\delta^\mu_\nu + u^\mu u_\nu\right)\frac{{\rm d}f^\nu_{\rm ext}}{{\rm d}\tau},$$

where the second, $O\left(e^2\right)$ term is the self-force responsible for the radiation reaction, which is orthogonal to the four-velocity and depends on the rate of change of the external force.

[1] Sources actually determine only the difference $A^\mu_{\rm ret} - A^\mu_{\rm adv}$, similar to $-iT = S - \mathbb{1}$ in the interaction picture of the S-matrix in QFT.

[2] $A^\mu_{\rm ret}$ is singular on the worldline. It can be shown that this does not affect the EoM, but gives a divergence that can be reabsorbed into a renormalisation of the mass of the particle.

[3] The original equation, established in 1938, actually involved the rate of change of the acceleration, resulting into the well-known (unphysical) runaway solutions. In order to eliminate this problem, one can replace the term ${\rm d}a^\nu/{\rm d}\tau$ with $m^{-1}{\rm d}f^\nu_{\rm ext}/{\rm d}\tau$ on the right-hand side. The order-reduced equation delivers a description of the motion whose expected accuracy is just as high as that of the original one, nevertheless suffering from none of the problems of the original equation. (See [20] and references therein for a discussion)

2.2. Notion of *force* in curved spacetime and gravitational self-force

One would now like to repeat a similar analysis for a light mass moving in a gravitational field, which is the typical situation of EMRIs. But the question arises: *is there any place for the notion of* (gravitational) *"force" in General Relativity* (GR)? In GR, test masses simply move on geodesics and no concept of force is required. But a useful notion of gravitational force can be introduced as follows. Consider a particle moving in a metric decomposable as

$$g'_{\mu\nu} = g_{\mu\nu} + h_{\mu\nu}, \qquad |h_{\mu\nu}| \ll 1. \tag{2.3}$$

At this level, $h_{\mu\nu}$ is a smooth, weak gravitational perturbation (*e.g.* an incident GW) of the background spacetime. The geodesic equation in the full metric is

$$\frac{\mathrm{d}^2 x^\mu}{\mathrm{d}\tau'^2} + \Gamma'^{\mu}_{\nu\rho}\frac{\mathrm{d}x^\nu}{\mathrm{d}\tau'}\frac{\mathrm{d}x^\rho}{\mathrm{d}\tau'} = 0, \tag{2.4}$$

where $\Gamma'^{\mu}_{\nu\rho}$ is the connection in the metric $g'_{\mu\nu}$ and τ' is an affine parameter along the trajectory. Adopting the decomposition (2.3) means that it is possible to reinterpret the particle's motion in terms of a trajectory in the background spacetime. Under this interpretation, the trajectory (in the background $g_{\mu\nu}$) is no longer geodesic; rather, the particle experiences an "external gravitational force", which is of course fictitious. This leads to

$$\frac{\mathrm{d}^2 x^\mu}{\mathrm{d}\tau^2} + \Gamma^{\mu}_{\nu\rho}\frac{\mathrm{d}x^\nu}{\mathrm{d}\tau}\frac{\mathrm{d}x^\rho}{\mathrm{d}\tau} = a^\mu_{\rm grav}, \tag{2.5}$$

where $\Gamma^{\mu}_{\nu\rho}$ is the connection in the background metric $g_{\mu\nu}$. Note that, in this non-covariant description, $h_{\mu\nu}$ ($\Gamma^{\mu}_{\nu\rho}$) is treated as a tensor (pseudo-tensor) field in $g_{\mu\nu}$, and similarly $a^\mu_{\rm grav}$ and $u^\mu \equiv \mathrm{d}x^\mu/\mathrm{d}\tau$ are regarded as vectors in the full metric. To determine $a^\mu_{\rm grav}$ in terms of the perturbation $h_{\mu\nu}$, one can make use of

$$\frac{\mathrm{d}}{\mathrm{d}\tau} = \left(\frac{\mathrm{d}\tau'}{\mathrm{d}\tau}\right)\frac{\mathrm{d}}{\mathrm{d}\tau'}$$

in eq. (2.5) and introduce the quantity $\Delta\Gamma^{\mu}_{\nu\rho} \equiv \Gamma'^{\mu}_{\nu\rho} - \Gamma^{\mu}_{\nu\rho}$. In light of eq. (2.4), this gives

$$a^\mu_{\rm grav} = -\Delta\Gamma^{\mu}_{\nu\rho}u^\nu u^\rho + \zeta u^\mu,$$

with $\zeta \equiv (\mathrm{d}\tau/\mathrm{d}\tau')\left(\mathrm{d}^2\tau'/\mathrm{d}\tau^2\right)$. From its definition in eq. (2.5), $a^\mu_{\rm grav}$ must be perpendicular to u^μ; hence, projecting $a^\mu_{\rm grav}$ orthogonally to u^μ keeps it unchanged, thus giving

$$a^\mu_{\rm grav} = -\left(\delta^\mu_\lambda + u^\mu u_\lambda\right)\Delta\Gamma^{\lambda}_{\nu\rho}u^\nu u^\rho.$$

Expanding now $\Delta\Gamma^{\lambda}_{\nu\rho}$ to linear order in $h_{\mu\nu}$, one obtains

$$a^\mu_{\rm grav} = -\frac{1}{2}\left(g^{\mu\lambda} + u^\mu u^\lambda\right)\left(\nabla_\sigma h_{\lambda\rho} + \nabla_\rho h_{\lambda\sigma} - \nabla_\lambda h_{\rho\sigma}\right)u^\rho u^\sigma \equiv \nabla^{\mu\beta\gamma}h_{\beta\gamma}, \tag{2.6}$$

where ∇_μ is the covariant derivative with respect to the background metric. The particle moves along the geodesic of the full metric $g'_{\mu\nu}$; however, to first order in the perturbation $h_{\mu\nu}$, its EoM can be *formally* written as a geodesic equation with respect to the background, supplemented by a gravitational "force" $F^\mu_{\rm grav} = ma^\mu_{\rm grav}$, m being the mass of the particle. The differential operator $\nabla^{\mu\beta\gamma}$, explicitly given by

$$\nabla^{\mu\beta\gamma} = \frac{1}{2}\left(g^{\mu\delta}u^\beta - 2g^{\mu\beta}u^\delta - u^\mu u^\beta u^\delta\right)u^\gamma\nabla_\delta, \tag{2.7}$$

determines the gravitational force exerted by any given external perturbation.

We must readily observe that $F^\mu_{\rm grav}$ – just like the metric perturbation itself – is gauge-dependent, as can be checked by considering a small gauge displacement,

$$x^\mu \longrightarrow x'^\mu = x^\mu + \xi^\mu(x),$$

(with ξ^μ assumed to scale like the external perturbation) under which the perturbation transforms as

$$h_{\mu\nu}(x) \longrightarrow h'_{\mu\nu}(x') = h_{\mu\nu}(x) - (\nabla_\mu \xi_\nu + \nabla_\nu \xi_\mu), \qquad |\nabla_\mu \xi_\nu| \lesssim |h_{\mu\nu}|, \tag{2.8}$$

inducing a change in the gravitational force given by[4]

$$\delta_\xi F^\mu_{\text{grav}} = m \nabla^{\mu\beta\gamma} (\nabla_\beta \xi_\gamma + \nabla_\beta \xi_\gamma). \tag{2.9}$$

This has to be expected since, because of the equivalence principle, there can be no diffeomorphism-invariant notion of gravitational force in GR.

One might be tempted to simply interpret the *gravitational self-force* (GSF, or sometimes just SF) as an example of a gravitational force of the type just discussed, with the source of the metric perturbation now being the particle itself. This naïve interpretation would be problematic.

The physical perturbation due to the particle – a retarded solution of the linearised Einstein equations, $h^{\text{ret}}_{\mu\nu}$ – is singular at the location of the particle, and the statement that the particle follows a geodesic of

$$g'_{\mu\nu} = g_{\mu\nu} + h^{\text{ret}}_{\mu\nu}$$

is therefore physically meaningless. Obviously, trying to apply eq. (2.6) with the perturbation replaced by the self-perturbation $h^{\text{ret}}_{\mu\nu}$ would yield a singular, hence meaningless, result. Relatedly, since we are now considering the self-gravity of the particle (it is no longer a test particle), we must make a mathematical sense of its being "point-like". This is not a trivial matter to address in curved spacetime, since, mathematically, the usual delta-function representation of a point particle stress-energy is inconsistent with the non-linearity of the full Einstein equations (see appendix B). The mathematical consistency of a delta-function source can be restored in the linear theory, but it remains a challenging task to understand how the notion of a point particle might emerge – rather than be pre-assumed – from a suitable limiting procedure.

A more profound consideration is as follows. The GSF is conceptually different from the external forces discussed above, in that the latter are, in truth, just fictitious forces resulting from our insistence to artificially split the physical spacetime into a background plus a perturbation. The GSF, in contrast, must be viewed as a genuine physical effect – even if a delicate one, as it is gauge dependent. There indeed exists an interpretation of the motion wherein the particle moves freely on a geodesic of a certain smooth, perturbed spacetime, subject to no GSF. However, in this description the smooth geometry will not be the physical spacetime of the background plus particle system – *i.e.*, the metric of this geometry is not a retarded solution of the linearised Einstein equations.

In 1997 two independent groups – Mino, Sasaki, Tanaka [12], and Quinn, Wald [13] – published three independent derivations of the GSF. This initial work was crucially inspired by the classical analyses of the electromagnetic SF problem, both in flat (Dirac, 1938 [10]) and curved (DeWitt and Brehme, 1960 [11]) spacetimes. Both groups pre-assumed a notion of point particle without seeking to make a consistent sense of this notion; however, in [12] an implementation of the idea of matched asymptotic expansions was pursued as well. This approach relies on the assumption that there can be identified two separate lenghtscales in the problem: one associated with the particle's mass m, and another, much larger, associated with the typical radius of curvature of the geometry in which the particle is moving (that, in the EMRI problem, is provided by the mass of the central BH, $M \gg m$). The two separate scales in the setup define a "near zone", $r \ll M$, and a "far zone", $r \gg m$ (r is a suitable measure of distance from the small BH). In the near zone, the geometry is approximately that of the small Schwarzschild BH with small tidal-type corrections from the background geometry. As we zoom away from the small object and enter the far zone, the effect of the small object's detailed structure becomes gradually less important, and at the far zone limit the geometry becomes that of the background spacetime, weakly perturbed by what is now a distant "point particle" – it is indeed the far zone limit through which a notion of point mass can be defined in a consistent way. When

[4] Terms arising from the gauge transformation of $\nabla^{\alpha\beta\gamma}$ are quadratic in the magnitude of ξ and are neglected.

$m \ll r \ll M$, one has a "buffer region": both near zone and far zone descriptions of the geometry are expected to be valid. Matching the near and far zone metrics (expressed as asymptotic expansions in r/M and m/r, respectively) constrains the motion of the particle (from a far zone point of view) and thus yields an expression for the GSF. This independent approach offered for the first time a fully GR-consistent treatment of the problem.

More recently, Gralla and Wald [14] developed a new procedure for deriving the GSF, offering improved mathematical rigour as well as a generalisation of the MiSaTaQuWa formula. Rather than relying on two separate asymptotic expansions of the metric, they introduced a single one-parameter family of metrics, which, through two different limiting procedures, can produce both near and far zone metrics in a natural way. The analysis proves that in the far-zone limit the particle is described precisely by the usual delta-function distribution, and that at the very limit $m \to 0$ this particle moves on a geodesic of the background. Furthermore, the analysis relaxes all assumptions about the nature of the small object: it no longer need to be a Schwarzschild BH, but can assume the form of any sufficiently small BH or even a blob of ordinary matter. This allows, in particular, for a spin-force term to appear in the resulting, generalised version of the MiSaTaQuWa formula.

The main end product of these theoretical developments is a firmly established general formula for the SF in a class of background spacetimes including Kerr. It is worth nothing that the GSF formula stems from nothing else than the Einstein equations with the usual conservation laws. It does not – and should not, as a matter of principle[5] – rely on any form or regularisation or subtraction of infinities.

2.3. MiSaTaQuWa equation

Stating the formula, we ignore spin-force terms and focus on the self-interaction part.

Consider a timelike geodesic Γ in a background spacetime with metric $g_{\mu\nu}$. For concreteness, let us think of Γ as a test particle orbit outside a Kerr BH, so $g_{\mu\nu}$ is the Kerr metric. Let τ be the proper time along Γ, and let $x^\mu = z^\mu(\tau)$ describe Γ in some smooth coordinate system and $u^\mu \equiv \mathrm{d}z^\mu/\mathrm{d}\tau$ be the four velocity of the test particle. Denote by $h^{\mathrm{ret}}_{\mu\nu}$ the physical, retarded metric perturbation from a particle of mass m whose worldline is Γ. Assume $h^{\mathrm{ret}}_{\mu\nu}$ is given in the Lorenz gauge, *i.e.*

$$\nabla^\mu \bar{h}^{\mathrm{ret}}_{\mu\nu} = 0, \qquad \bar{h}^{\mathrm{ret}}_{\mu\nu} \equiv h^{\mathrm{ret}}_{\mu\nu} - \frac{1}{2} g_{\mu\nu} g^{\alpha\beta} h_{\alpha\beta}, \tag{2.10}$$

where we introduced the trace-reversed perturbation $\bar{h}^{\mathrm{ret}}_{\mu\nu}$. Throughout the discussion indices are raised and lowered by the background metric $g_{\mu\nu}$, and covariant derivatives are taken with respect to this metric.

At any spacetime point x, the retarded perturbation can be written as a sum of two pieces,

$$\bar{h}^{\mathrm{ret}}_{\mu\nu} = \bar{h}^{\mathrm{dir}}_{\mu\nu} + \bar{h}^{\mathrm{tail}}_{\mu\nu},$$

the former being the "direct" contribution from the intersection of the past light-cone of x with Γ, and the latter being the "tail" contribution arising from the part of Γ inside the light-cone (see fig. 2.1). The occurrence of a tail term is a well-known feature of the wave equation in 3+1D curved spacetime: Huygens' principle no longer holds and the tail term can be interpreted physically as arising from the effect of waves being scattered off spacetime curvature.[6] Both $\bar{h}^{\mathrm{ret}}_{\mu\nu}$ and

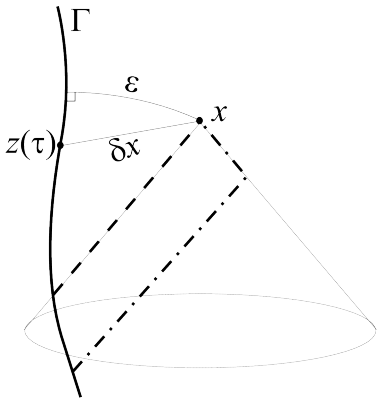

Figure 2.1.: The direct and tail contributions.

[5] At a fundamental level, there should be no need for regularisation in GR, because, apart from curvature singularities inside BHs, everything in the problem should be finite.

[6] The move to curved spacetime brings a major change to the physics of the problem. In flat spacetime, waves propagate at the speed of light, along null rays, while in curved spacetime, waves scatter off the spacetime curvature, causing solutions to propagate not just on light-cones, but also within them. Because of this, the retarded field depends not only on the state of the particle at the retarded point $z^\mu(\tau_{\mathrm{ret}})$, but on its state at all prior points $z^\mu(\tau < \tau_{\mathrm{ret}})$.

$\bar{h}^{\rm dir}_{\mu\nu}$ obviously diverge when evaluated on Γ; however, $\bar{h}^{\rm tail}_{\mu\nu}$ is continuous and differentiable everywhere, including on the worldline. Notably, though, the tail field is not a smooth (*i.e.*, analytic) function on the worldline, and is *not* a vacuum solution of the linearised Einstein equations.

The MiSaTaQuWa formula states that the GSF at a given point z along Γ results simply from the back reaction of the tail field,

$$F^{\mu}_{\rm self}(z) = m\bar{\nabla}^{\mu\beta\gamma}\bar{h}^{\rm tail}_{\mu\nu}(z)\,. \tag{2.11}$$

Here

$$\bar{\nabla}^{\mu\beta\gamma} \equiv \frac{1}{4}\left(2g^{\mu\delta}u^{\beta}u^{\gamma} - 4g^{\mu\beta}u^{\gamma}u^{\delta} - 2u^{\mu}u^{\beta}u^{\gamma}u^{\delta} + u^{\mu}g^{\beta\gamma}u^{\delta} + g^{\mu\delta}g^{\beta\gamma}\right)\nabla_{\delta}$$

is the "force operator" in terms of the trace-reversed metric perturbation defined along the worldline of the oarticle, as it depends upon the four-velocity u^{μ} and the background metric $g_{\mu\nu}$ at point z.[7]

Remark In the original MiSaTaQuWa formulation the SF is not expressed directly in terms of this "gradient". The EoM was found to be

$$\frac{D^2 z^{\mu}}{{\rm d}\tau^2} = -\frac{1}{2}\epsilon P^{\mu\nu}\left(2h^{\rm tail}_{\nu\alpha\beta} - h^{\rm tail}_{\alpha\beta\nu}\right)u^{\alpha}u^{\beta} + O\left(\epsilon^2\right), \tag{2.12}$$

where $D/{\rm d}\tau \equiv u^{\mu}\nabla_{\mu}$ is the directional covariant derivative, $P^{\mu\nu} \equiv g^{\mu\nu} + u^{\mu}u^{\nu}$ projects orthogonally to the worldline, and the tail term is given by

$$h^{\rm tail}_{\mu\nu\rho}(z(\tau)) = 4m\int_{-\infty}^{\tau^-} \nabla_{\rho}\bar{G}_{\mu\nu\mu'\nu'}u^{\mu'}u^{\nu'}\,{\rm d}\tau'; \tag{2.13}$$

the integral covers all of the worldline earlier than the point $z^{\mu}(\tau)$ at which the force is evaluated. The bar atop $G_{\mu\nu\mu'\nu'}$ again denotes a trace reversal. We must note as well that $G_{\mu\nu\mu'\nu'}$ is a bitensor, an object which lives in the tangent spaces of two different points x and x'.

The intuitive picture to glean from the MiSaTaQuWa result is that the direct piece of the field is analogous to a Coulomb field, moving with the particle and exerting no force on it, in the same way the self-field exerts no force on a body in Newtonian gravity. Loosely speaking, from the perspective of the particle, the tail, consisting as it does of backscattered radiation, is *indistinguishable from any other incoming radiation*. In other words, it is effectively an *external field*, and like an external field *it exerts a force*.

2.3.1. Detweiler-Whiting reformulation

In 2003 Detweiler and Whiting [15] replaced the direct/tail decomposition of the retarded perturbation

$$\bar{h}^{\rm ret}_{\mu\nu} = \bar{h}^{\rm S}_{\mu\nu} + \bar{h}^{\rm R}_{\mu\nu}.$$

The R(egular)-field includes all the backscattered radiation in the tail, but, unlike the tail field, is a smooth, vacuum solution of the perturbation equations, which, nonetheless, gives rise to the same physical SF as the tail field; *i.e.*,

$$F^{\mu}_{\rm self}(z) = m\bar{\nabla}^{\mu\beta\gamma}\bar{h}^{\rm R}_{\beta\gamma}(z)\,. \tag{2.14}$$

The S(ingular)-field, which mimics the singular behaviour of the retarded field near the particle, exerts no SF and does not affect the motion of the particle.

The breakthrough of the discovery that the SF can be expressed as the back-reaction force from a smooth vacuum perturbation leads to an interesting re-interpretation of the GSF effect: *the particle effectively moves freely along a geodesic of a smooth perturbed spacetime with metric* $g_{\mu\nu} + h^{\rm R}_{\mu\nu} \equiv \tilde{g}_{\mu\nu}$. That is very welcome, because in this alternative picture, which is more in the spirit of *GR equivalence principle*, the notion of a SF becomes artificial (and obsolete) in much the same way that the notion of an external gravitational force is artificial.

[7] A crucial point to have in mind is that the MiSaTaQuWa formula (2.11) is guaranteed to hold true only if $\bar{h}^{\rm ret}_{\mu\nu}$ satisfies the Lorenz gauge condition (2.10).

Anyway, the R-field does not represent the actual physical perturbation from the particle, which is of course $\bar{h}^{\rm ret}_{\mu\nu}$. The R-field has peculiar causal properties, since its value at an event x depends not only on events in the causal past of x, but also on events outside the light-cone of x. Rather than an entity of physical substance, it should be viewed as an effective field that allows us to describe the dynamics in terms of geodesic motion.

The two descriptions of the perturbed motion – self-accelerated motion in $g_{\mu\nu}$ *vs.* geodesic motion in $g_{\mu\nu}+h^{\rm R}_{\mu\nu}$ – are alternative, equivalent interpretations of the same, genuine physical effect. The two points of view are not contradictory, but rather they are complementary in their perspective on the problem. Therefore we can say that eqs. (2.11), (2.14) prescribe the correct regularisation of the GSF and form the fundamental basis for all modern calculations.

Remark In the original formulation, this translates into

$$\frac{D^2 z^\mu}{\mathrm{d}\tau^2} = -\frac{1}{2}\epsilon P^{\mu\nu}\left(2\nabla_\beta h^{\rm R}_{\nu\alpha} - \nabla_\nu h^{\rm R}_{\alpha\beta}\right) u^\alpha u^\beta + O\left(\epsilon^2\right), \tag{2.15}$$

or explicitly as the geodesic equation in the metric $\tilde{g}_{\mu\nu}$,

$$\frac{\tilde{D}^2 z^\mu}{\mathrm{d}\tilde{\tau}^2} = O\left(\epsilon^2\right), \tag{2.16}$$

where $\tilde{D}/\mathrm{d}\tilde{\tau} \equiv \tilde{u}^\mu \tilde{\nabla}_\mu$ is a directional covariant derivative compatible with $\tilde{g}_{\mu\nu}$, $\tilde{u}^\mu = \mathrm{d}z^\mu/\mathrm{d}\tilde{\tau}$ is the four-velocity normalised in $\tilde{g}_{\mu\nu}$ and $\tilde{\tau}$ is the proper time along z^μ as measured in $\tilde{g}_{\mu\nu}$. Eq. (2.15) is equivalent to eq. (2.12), because on the worldline $h^{\rm R}_{\mu\nu}$ differs from $h^{\rm tail}_{\mu\nu}$ only by (*i*) Riemann terms that cancel in eq. (2.15) and (*ii*) terms proportional to the worldline acceleration, which can be treated as effectively higher order because the acceleration is already $O(\epsilon)$.

2.3.2. Generalised Equivalence Principle

It is worth to dwell longer on the interpretation of the R-field. Because $\tilde{g}_{\mu\nu} = g_{\mu\nu} + h^{\rm R}_{\mu\nu}$ is a smooth vacuum solution, at the position of the particle an observer cannot distinguish it from $g_{\mu\nu}$. Although a part of $\tilde{g}_{\mu\nu}$ comes from the retarded field sourced by the particle, to the observer on the worldline it appears just as would any metric sourced by a distant object. However, this interpretation of the effective metric as an effectively external metric is delicate. To realise a split into S- and R-fields, both of them must be made a-causal when evaluated off the worldline. More precisely, in addition to depending on the causal past of the particle, $h^{\rm R}_{\mu\nu}(x)$ depends on the particle at spatially related points x'; so in this sense its interpretability as a physical external field is limited. Yet when evaluated *on* the worldline, $h^{\rm R}_{\mu\nu}$ and its derivatives *are* causal, and this is the sense in which $\tilde{g}_{\mu\nu}$ appears as a *physical metric* on the worldline. These properties of $h^{\rm R}_{\mu\nu}$ are what makes eq. (2.16) a meaningful result. Although it may not be an obvious fact at first glance, any equation of motion can be written as the geodesic equation in some smooth piece of the metric.

Remark This is most easily seen by writing the EoM in a frame that comoves with the particle. In locally Cartesian coordinates (t, x^i) adapted to that frame, such as Fermi-Walker coordinates,[8] the particle's EoM reads $a_i = F_i$, where $a^\mu \equiv D^2 z^\mu/\mathrm{d}\tau^2$ is the covariant acceleration of the particle and F^μ is the force (per unit mass) acting on it. Now suppose we were to define some smooth field $h^r_{\mu\nu}$ and a corresponding singular field $h^s_{\mu\nu} \equiv h_{\mu\nu} - h^r_{\mu\nu}$. In the comoving coordinates, the linearised geodesic equation in the regular metric $g_{\mu\nu} + h^r_{\mu\nu}$, in the form analogous to eq. (2.15), reads

$$a_i = -\partial_t h^r_{ti} + \frac{1}{2}\partial_i h^r_{tt}.$$

No matter what force F^μ acts on the particle, the EoM $a_i = F_i$ could be written as the geodesic equation in $g_{\mu\nu} + h^r_{\mu\nu}$ simply by choosing

$$\partial_t h^r_{ti}|_\gamma = 0 \quad \text{and} \quad \partial_i h^r_{tt}|_\gamma = 2F_i.$$

[8] See section A.1.

Besides those two conditions, the regular field could be entirely freely specified. Regardless of the specification we made, we would have defined a split $h_{\mu\nu} = h^s_{\mu\nu} + h^r_{\mu\nu}$ in which $h^s_{\mu\nu}$ is singular and exerts no force, and $g_{\mu\nu} + h^r_{\mu\nu}$ is a regular metric in which the motion is geodesic.

Given this, it is of *no special significance* that the MiSaTaQuWa equation is equivalent to geodesic motion in *some* effective metric. But it *is* significant that this equation is equivalent to geodesic motion in the *particular* effective metric $\tilde{g}_{\mu\nu} = g_{\mu\nu} + h^{\mathrm{R}}_{\mu\nu}$ identified by Detweiler and Whiting, because of the particular properties of this metric: $\tilde{g}_{\mu\nu}$ is a smooth vacuum solution that is causal on the particle's worldline. And, because of these properties, we may think of the MiSaTaQuWa equation as a generalised equivalence principle: *any object, if it is sufficiently compact and slowly spinning, regardless of its internal composition, falls freely in a gravitational field* $\tilde{g}_{\mu\nu}$ that can be thought of as the physical "external" gravitational field at its "position".

2.4. Gauge, motion and long-term dynamics

In spite of differences between the concepts of a general (external) gravitational force and that of the SF, the two notions share a common feature: they are both defined via a mapping of the physical trajectory from a perturbed spacetime to a background spacetime. In both cases, such a procedure gives rise to a gauge ambiguity. Consider a small[9] gauge displacement,

$$x^\mu \longrightarrow x^\mu - \epsilon\xi^\mu(x)\,.$$

The change this induces on the physical SF can be computed and found to have exactly the same form as the gauge transformation law for the external gravitational force [17], which is not surprising, given the similar geometrical origin of the gauge freedom in both cases.

Two important facts must be understood: in perturbation theory, motion is intimately related to gauge freedom; and, in problems of astrophysical interest, the most important dynamical effects occur on the very long time scale $\sim 1/\epsilon$.

At leading order, the object's motion is geodesic in the background metric $g_{\mu\nu}$ and all deviation from this motion is driven by an order-ϵ force. Suppose the self-accelerated worldline γ is a smooth function of ϵ; then one can write its parametric relation as an expansion [16]

$$z^\mu(s,\epsilon) = z^\mu_0(s) + \epsilon z^\mu_1(s) + O\left(\epsilon^2\right), \tag{2.17}$$

where s is an affine parameter on the worldline, and the zeroth-order term z^μ_0 is a geodesic of $g_{\mu\nu}$. Considering the effect of a gauge transformation generated by a vector $\epsilon\xi^\mu$, z^μ is shifted to a curve $z'^\mu = z^\mu - \epsilon\xi^\mu + O\left(\epsilon^2\right)$. Since the vector ξ^μ is generic, nothing prevents us from choosing $\xi^\mu = z^\mu_1$, which leaves us with $z'^\mu = z^\mu_0(s) + O\left(\epsilon^2\right)$, entirely eliminating the first-order deviation from z^μ_0. This same reasoning can be carried to arbitrary order, meaning we can precisely set $z'^\mu = z^\mu_0$: the effect of the SF appears to be pure gauge.

The MiSaTaQuWa equation is not gauge invariant and cannot by themself produce a meaningful answer to a well-posed physical question; to obtain such an answer it is necessary to combine the EoMs with the metric perturbation so as to form gauge-invariant quantities that will correspond to direct observables. This point is very important and cannot be over-emphasized.

But let us take a closer look at the issue. If we examine any finite region of spacetime and consider the limit $\epsilon \to 0$ in this region, the deviation from a background geodesic is, indeed, pure gauge. This does not mean it is irrelevant: in any given gauge, it must be accounted for to obtain the correct metric in that gauge. But it need not be accounted for in the linearised metric, since we can always substitute the expansion (2.17) into $h^1_{\mu\nu}(x;z)$ to obtain

$$\epsilon h^1_{\mu\nu}(x;z) = \epsilon h^1_{\mu\nu}(x;z_0) + \epsilon^2 \delta h^1_{\mu\nu}(x;z_0,z_1) + O\left(\epsilon^3\right), \tag{2.18}$$

[9] Small compared to the mass m of the particle; recall that ϵ counts powers of m. This contrasts with the argument in section 2.2, where the gauge displacement should be assumed to scale like the external perturbation $h_{\mu\nu}$.

so we can transfer the term $\delta h^1_{\mu\nu}$ into the second-order perturbation, $\epsilon^2 h^2_{\mu\nu}$.

However, this analysis assumes that one works in a fixed, finite domain – and one does not typically work in such a domain in problems of interest. Consider an EMRI: GWs carry away orbital energy from the system at a rate $\dot{E}/E \sim \epsilon$, then the inspiral occurs on the time scale $t_{\rm rr} \sim 1/\epsilon$ (the so-called radiation-reaction time). So, in practice, we are not looking at the limit $\epsilon \to 0$ on a finite time interval $[0, T]$, where T is ϵ-independent; instead, we are looking at that limit on a time interval $[0, T/\epsilon]$ that blows up.

This consideration forces to adjust the thinking about motion and gauge. Loosely speaking, the deviation from geodesic motion, ϵz^μ_1, is governed by an equation of the form $\mathrm{d}^2 z^\mu_1/\mathrm{d}t^2 \sim F^\mu_1$. On the radiation-reaction time scale, it therefore behaves as

$$\epsilon z^\mu_1 \sim \epsilon F^\mu_1 t^2_{\rm rr} \sim F^\mu_1/\epsilon,$$

and it blows up in the limit $\epsilon \to 0$. So, on this domain, one cannot rightly write the worldline as a geodesic plus a self-forced correction, and one cannot use a small gauge transformation to shift the perturbed worldline onto a background geodesic – the gauge transformation would have to blow up in the limit $\epsilon \to 0$. But there is a way out of this situation, when we observe that these arguments about time scales translate into arguments about spatial scales. So, if one's accuracy is limited to a time span Δt, then it is also limited to a spatial region of similar size.

Let us denote by $\mathcal{D}_{\rho(\epsilon)}$ a spacetime region of size, both temporal and spatial, $\rho(\epsilon)$. Call an asymptotic solution to the Einstein equations a "good solution" in $\mathcal{D}_{\rho(\epsilon)}$ if it is uniform in that region, that is, the asymptotic expansion

$$\mathfrak{g}_{\mu\nu} = g_{\mu\nu} + \sum_{n=1}^{+\infty} \epsilon^n h^n_{\mu\nu}$$

must satisfy

$$\lim_{\epsilon\to 0} \frac{\epsilon h^1_{\mu\nu}}{g_{\mu\nu}} = 0 \quad \text{and} \quad \lim_{\epsilon\to 0} \frac{\epsilon^n h^n_{\mu\nu}}{\epsilon^{n-1} h^{n-1}_{\mu\nu}} = 0$$

uniformly (*e.g.*, in a sup-norm).

For the EMRI problem, we are interested in obtaining a good solution in a domain of size $\rho = t_{\rm rr} \sim 1/\epsilon$. Suppose we use an asymptotic expansion of the form (2.18) and incorporate $\delta h^1_{\mu\nu}$ into $h^2_{\mu\nu}$. In a gauge such as the Lorenz gauge, z^μ_1 grows as $\sim F^\mu_1 t^2_{\rm rr}$, and so $\delta h^1_{\mu\nu}$ likewise grows as $t^2_{\rm rr}$. Hence, on $\mathcal{D}_{1/\epsilon}$ its contribution to $\epsilon^2 h^2_{\mu\nu}$ behaves at best as ϵ^0, comparable to $g_{\mu\nu}$ – clearly, this is not a good approximation. Suppose we instead eliminated z^μ_1 using a gauge transformation generated by $\xi^\mu = z^\mu_1$. This removes the offending growth in $h^2_{\mu\nu}$, but it commits a worse offence: it alters $\epsilon h^1_{\mu\nu}$ by an amount $2\epsilon\xi_{(\mu;\nu)}$, which behaves at best as $\epsilon t_{\rm rr}$, or as ϵ^0 on $\mathcal{D}_{1/\epsilon}$. Hence: if we are in a gauge where the SF is non-vanishing, $h^2_{\mu\nu}$ behaves poorly; if we are in a gauge where the SF is vanishing, even $h^1_{\mu\nu}$ behaves poorly.

Let us chase the consequences of this. To obtain a good approximation in $\mathcal{D}_{1/\epsilon}$, we need to work in a class of gauges compatible with uniformity in $\mathcal{D}_{1/\epsilon}$. This means, in particular, that if we obtain a good approximation in a particular gauge – call it a "good gauge" – we must confine ourselves to a class of gauges related to the good one by uniformly small transformations. In turn, this means that the effects of the SF are *not* pure gauge on $\mathcal{D}_{1/\epsilon}$. In more physical words, due to dissipation, z^μ will deviate from any given geodesic z^μ_0 by a very large amount in $\mathcal{D}_{1/\epsilon}$, but, by using an allowed gauge transformation, we may shift it only by a very small, order-ϵ amount on that domain.

At the end of the day, we learn that, although the self-forced deviation from z^μ_0 is pure gauge on a domain like $\mathcal{D}_{\epsilon^0}$, it is no longer pure gauge in the domain $\mathcal{D}_{1/\epsilon}$.

The gauge dependence of the SF by no means implies that there is something "unphysical" about it – the SF is as *physical* as the (gauge-dependent) metric perturbation itself. The gauge dependence does mean, however, that one needs to exercise *some care* in decoding the physical content of the SF, if based on the value of the SF alone.

2.5. Equations of motion

Given the GSF, the EoM of the particle becomes

$$m\frac{D^2 z^\alpha}{\mathrm{d}\tau^2} = mu^\beta \nabla_\beta u^\alpha := F^\alpha_{\mathrm{self}}. \tag{2.19}$$

This equation, along with the MiSaTaQuWa equation (2.11), describes the dynamics of the particle given the metric perturbation (and assuming one has a way of extracting the tail piece out of the full perturbation). To close the system of equations, we need to know how the metric perturbation is determined from the trajectory of the particle; this is provided by the linearised Einstein equation, which takes the Lorenz-gauge form

$$\nabla^\sigma \nabla_\sigma \bar{h}_{\mu\nu} + 2R^\alpha{}_\mu{}^\beta{}_\nu \bar{h}_{\alpha\beta} = -16\pi m \int_\gamma \delta_4\left(x, z\left(\tau\right)\right) u_\mu u_\nu \mathrm{d}\tau, \tag{2.20}$$

where $g \equiv \det g_{\mu\nu}$ and $\delta_4\left(x,z\right) \equiv \delta^4\left(x^\mu - z^\mu\right)/\sqrt{-g}$ is the invariant four-dimensional delta-function. The field equation (2.20) is to be supplemented by the gauge condition (2.10) and by suitable boundary conditions.

The set of equations (2.11), (2.19), (2.20) and (2.10) (together with a method to obtain $\bar{h}^{\mathrm{tail}}_{\mu\nu}$ out of the retarded solution $\bar{h}_{\mu\nu}$) should in principle determine the dynamics of the orbit at linear order in perturbation theory. But there is a slight problem: the field equation (2.20) is only consistent with the Lorenz gauge condition (2.10) if the particle is moving strictly along a geodesic – which would then be inconsistent with the EoM (2.19). To resolve this inconsistency while allowing for orbital evolution, Gralla and Wald suggested a "Lorenz gauge relaxation" approach, wherein one relaxes the gauge condition and consider solutions of the set (2.11), (2.19), (2.20). One then expects that, in situations where the orbit is very nearly geodesic (as is usually the case with LISA-relevant astrophysical inspirals), such solutions would give a faithful, albeit approximate description of the actual orbit [13, 14].

Other proposals were put forward, like to use self-consistent solutions of the set (2.11), (2.19), (2.20) for modelling the slow orbital evolution at linear order in perturbation theory; or a different mathematical framework, based on techniques from multi-scale perturbation theory.[10]

2.6. Survey of computational methods

Using the differential operator (2.7), the EoM can be cast in the form

$$m\frac{D^2 z^\alpha}{\mathrm{d}\tau^2} = m \lim_{x\to z} \nabla^{\alpha\beta\gamma} h^{\mathrm{R}}_{\beta\gamma}\left(x\right) := F^\alpha\left(z\right). \tag{2.21}$$

The argument x represents a field point in a neighbourhood of the worldline and z is the worldline point where the GSF is evaluated. Writing the R-field as the difference between the physical (retarded) metric perturbation sourced by the particle and the S-field,[11] the EoM can be written as

$$F^\alpha\left(z\right) = m \lim_{x\to z} \nabla^{\alpha\beta\gamma}\left[h^{\mathrm{ret}}_{\beta\gamma}\left(x\right) - h^{\mathrm{S}}_{\beta\gamma}\left(x\right)\right]. \tag{2.22}$$

The limit procedure is necessary here because the individual fields, unlike their difference, are each singular at $x \to z$, and so are their derivatives. Suppose that the full perturbation satisfies the Lorenz-gauge form of the linearised Einstein equations, that we already saw in eq. (2.20); then, choosing a point-particle source, we can write it schematically as

$$\tilde{\Box} h_{\alpha\beta} = S_{\alpha\beta}. \tag{2.23}$$

[10] More on this in subsection 2.7.2.

[11] Choosing between the direct+tail and the S+R decomposition of the retarded field does not alter the discussion; the only exception is that statements referring to the smoothness of the R-field would need to be formulated more carefully to reflect the irregularity in the higher derivatives of the tail field.

In fact, recall again that the linearised Einstein tensor (B.3) in the Lorenz gauge takes the simple, hyperbolic form

$$\delta G_{\alpha\beta}[h] = -\frac{1}{2}\Box\bar{h}_{\alpha\beta} - R_{\mu\ \nu}^{\ \alpha\ \ \beta}\bar{h}_{\alpha\beta},$$

where $R_{\mu\ \nu}^{\ \alpha\ \ \beta}$ is the Riemann tensor of the background metric. This gives the explicit form of the wave-like operator $\tilde{\Box}$ in eq. (2.23), and $S_{\alpha\beta}$ is the linearised energy-momentum tensor associated with the particle, with support (a delta-function) confined to the worldline of the particle. Eq. (2.23) has to be solved with physical, retarded boundary conditions, which are most conveniently imposed at infinity and on the BH event horizon. The conditions are that no radiation should be coming in from (past null) infinity, and that no radiation should be coming out from inside the BH.

From a computational point of view, solving (2.23) and evaluating (2.22) pose two main difficulties:

1. the Lorenz-gauge field equations (2.23) constitute a complicated set of coupled partial differential equations (PDEs). Even though these equations are linear and manifestly hyperbolic, solving them numerically is computationally expensive and technically challenging – due, in particular, to the need to resolve the diverging field near the particle with sufficient accuracy and also to the occurrence of certain mode instabilities;

2. there is a "subtraction problem": to implement (2.22) and obtain the GSF, one has to subtract one divergent quantity from another (which is usually given only numerically), before taking the regular limit to the particle. This is obviously problematic in practice.

2.6.1. Mode-sum and puncture methods

Let us move on now to describe two methods for tackling the second technical difficulty, the subtraction problem.

Mode-sum This method is a general procedure addressing the subtraction problem. The basic idea is simple: instead of directly subtracting the divergent field $\nabla^{\alpha\beta\gamma}h^{\rm S}_{\beta\gamma}(x)$ from the other divergent one $\nabla^{\alpha\beta\gamma}h^{\rm ret}_{\beta\gamma}(x)$, (*i*) decompose each of these fields into multipolar-mode components (using a basis of angular harmonics defined on spheres around the large BH), then (*ii*) perform the subtraction mode-by-mode; finally (*iii*) add up all the "regularised" modal contributions.

The benefit of such an approach is twofold. *First*, due to the particular Coulomb-like form of the singularity of the field near the particle, the individual multipole modes of $\nabla^{\alpha\beta\gamma}h^{\rm S}_{\beta\gamma}(x)$ (and of $\nabla^{\alpha\beta\gamma}h^{\rm ret}_{\beta\gamma}(x)$) have finite values even at the location of the particle, so one only ever subtracts finite quantities. *Second*, the perturbation field $h^{\rm ret}_{\beta\gamma}(x)$ is typically solved for mode-by-mode anyway, so the necessary input for the mode-sum procedure is readily available without any extra work.

To describe this more precisely (while avoiding technical details that can be found, for example in [18, 19]) define the fields

$$F^{\alpha}_{\rm ret}(x) \equiv m\nabla^{\alpha\beta\gamma}h^{\rm ret}_{\beta\gamma}(x) \qquad \text{and} \qquad F^{\alpha}_{\rm S}(x) \equiv m\nabla^{\alpha\beta\gamma}h^{\rm S}_{\beta\gamma}(x) \tag{2.24}$$

in the (Kerr) BH geometry, introducing the standard Boyer-Lindquist coordinates (t, r, θ, ϕ) covering the exterior of the BH. Then consider the decomposition of these fields into spherical-harmonic modes, defined on spheres of constant t and r around the BH, *i.e.*

$$F^{\alpha}_{\rm ret} = \sum_{\ell=0}^{+\infty} F^{\alpha\ell}_{\rm ret}, \qquad F^{\alpha\ell}_{\rm ret} \equiv \sum_{m=-\ell}^{\ell} F_{\ell m}(t,r)\, Y_{\ell m}(\theta,\phi),$$

with $Y_{\ell m}$ being the usual spherical harmonics; similarly for $F^{\alpha}_{\rm S}$. Eq. (2.22) thus becomes

$$F^{\alpha}(z) = \lim_{x\to z}\sum_{\ell=0}^{+\infty}\left[F^{\alpha\ell}_{\rm ret}(x) - F^{\alpha\ell}_{\rm S}(x)\right].$$

The individual mode-sums of $F^{\alpha\ell}_{\text{ret}}$ and $F^{\alpha\ell}_{\text{S}}$ both diverge at the particle: in the multipolar space, the Coulomb-like particle singularity has turned into a large-ℓ (*i.e.* ultraviolet) divergence. However, the mode-sum of the difference $F^{\alpha\ell}_{\text{ret}}(x) - F^{\alpha\ell}_{\text{S}}(x)$ converges faster than any power of $1/\ell$ everywhere, even at the particle, since $h^{\text{R}}_{\beta\gamma}(x)$ in eq. (2.22) has to be a smooth field. This implies that $F^{\alpha\ell}_{\text{ret}}$ and $F^{\alpha\ell}_{\text{S}}$ share the same ultraviolet singularity structure. In fact, based on the detailed form of the singular field, one can show that[12]

$$F^{\alpha\ell}_{\text{S}}(z) \sim A^\alpha \ell + B^\alpha + C^\alpha \ell^{-1}, \qquad \ell \gg 1, \tag{2.25}$$

where A^α, B^α and C^α are ℓ-independent expansion coefficients encoding the local ultraviolet structure, whose values depend on the background geometry as well on the the particle's location and velocity. We can thus write

$$F^\alpha(z) = \sum_{\ell=0}^{+\infty} \left[F^{\alpha\ell}_{\text{ret}}(x) - \left(A^\alpha \ell + B^\alpha + C^\alpha \ell^{-1}\right)\right] - D^\alpha,$$

$$D^\alpha \equiv \sum_{\ell=0}^{+\infty} \left[F^{\alpha\ell}_{\text{S}} - \left(A^\alpha \ell + B^\alpha + C^\alpha \ell^{-1}\right)\right],$$

where the two individual sums are convergent (at least as $\sim \ell^{-1}$). Neglecting the acceleration caused by the force, the GSF has historically been calculated by approximating the source orbit as a geodesic; to this end, these four parameters have been derived analytically for arbitrary geodesic orbits in Kerr spacetime. In particular, it has been shown that, when the acceleration is neglected, C^α and D^α always vanish identically. One thus arrives at the working form of the mode-sum formula, *i.e.*

$$F^\alpha(z) = \sum_{\ell=0}^{+\infty} \left[F^{\alpha\ell}_{\text{ret}}(x) - (A^\alpha \ell + B^\alpha)\right]. \tag{2.26}$$

This provides a practical means of evaluating the GSF at any point along a given (geodesic) orbit: first obtain the multipole modes of the physical field $h^{\text{ret}}_{\alpha\beta}$ by solving a suitable version of the linearised Einstein equations mode-by-mode (usually done numerically); from each mode then subtract the analytically given quantity $A^\alpha \ell + B^\alpha$, and finally add up all the modal contributions.

The above schematic description suppresses important details and one should also mention that the mode-sum in eq. (2.26) converges only slowly, since the summand typically falls off only as $\sim \ell^{-2}$ – this means that one normally has to compute a large number of ℓ-modes ($\ell_{\text{max}} \sim 50$ is typical), which can become computationally expensive. Anyway, the mode-sum scheme, in its many variants, has been the primary framework for GSF calculations.

Puncture (or "effective source") This other method begins by recognizing that an approximation to the exact singular potential can be used to regularise the delta-function source term of the original field equation. The puncture method addresses the subtraction problem differently from mode-sum, since the "regularisation" is performed already at the level of the field equation (2.23). Instead of solving for the physical field $h_{\alpha\beta}$ and then subtracting the singular field, one solves directly for a local approximation to the regular field $h^{\text{R}}_{\alpha\beta}$ [21]. Specifically, it is designed an analytic function $h^{\mathcal{P}}_{\alpha\beta}(x)$ that approximates the singular field $h^{\text{S}}_{\alpha\beta}(x)$ near the particle sufficiently well that

$$\lim_{x\to z}\left(h^{\mathcal{P}}_{\alpha\beta} - h^{\text{S}}_{\alpha\beta}\right) = 0 \quad \text{and} \quad \lim_{x\to z}\left(\nabla^{\alpha\beta\gamma} h^{\mathcal{P}}_{\beta\gamma} - \nabla^{\alpha\beta\gamma} h^{\text{S}}_{\beta\gamma}\right) = 0.$$

Then it is perfectly allowable to replace the true singular field in eq. (2.22) with its puncture-field approximant

$$F^\alpha(z) = m \lim_{x\to z} \nabla^{\alpha\beta\gamma} h^{\mathcal{R}}_{\beta\gamma}(x), \tag{2.27}$$

where we have introduced the residual field $h^{\mathcal{R}}_{\beta\gamma} \equiv h_{\beta\gamma} - h^{\mathcal{P}}_{\beta\gamma}$. We then make $h^{\mathcal{R}}_{\beta\gamma}$ the subject of the field equation (2.23), *i.e.*

$$\tilde{\Box} h^{\mathcal{R}}_{\alpha\beta} = S_{\alpha\beta} - \tilde{\Box} h^{\mathcal{P}}_{\alpha\beta} \equiv S^{\text{eff}}_{\alpha\beta}. \tag{2.28}$$

[12] See [19]; as a simple example to understand this general behaviour, see appendix C.

The effective source $S^{\text{eff}}_{\alpha\beta}$ contains no delta function on the particle worldline; its residual non-smoothness there is determined by how well the puncture field approximates the S-field. The field $h^{\mathcal{R}}_{\alpha\beta}$ is at least once differentiable (unlike the physical field $h_{\alpha\beta}$, which is divergent), and it directly yields the GSF, via eq. (2.27).

In practice, one can restrict the support of $S^{\text{eff}}_{\alpha\beta}$ to within a small region around the particle worldline, so as to avoid having to control the behaviour of the puncture field and effective source far from the particle. This can be achieved with a suitable window function or by introducing a "worldtube" around the worldline, such that one solves for $h^{\mathcal{R}}_{\alpha\beta}$ inside the worldtube and for the original perturbation $h_{\alpha\beta}$ outside it, with the analytically known value of the puncture field used to communicate between the two variables across the boundary of the tube. This scheme can be implemented numerically without any multipole decomposition, directly evolving the hyperbolic PDE (2.28) from initial conditions in 3+1 dimensions. The relevance of this method is that it has also proved useful when applied in conjunction with a mode decomposition: one can separate the field into azimuthal modes, $\sim e^{im\phi}$, and evolve each of the m-modes separately in 2+1 dimensions (this is possible and useful even on a Kerr background, thanks to its axial symmetry).

The utility and significance of this idea becomes fully manifest when coming to solve the second-order field equation (B.4), where applying the mode-sum method becomes impossible in general. Recall that the source term appearing in that second-order equation is sufficiently singular that the equation does not actually admit a globally valid solution. Even restricting to $x^\mu \neq z^\mu$, the singularity in the second-order solution is strong enough (a consequence of the distributionally ill-defined source) that its individual ℓ modes diverge at the particle: this means that, even if one were given the modes of the retarded field, one could not apply mode-sum regularisation to extract the regular field. For these reasons, the puncture idea takes on a more fundamental status in second-order GSF calculations. Finally, we mention that the basic idea can also be applied to control the behaviour of the second-order solution near the horizon and at large distances, in circumstances where one would otherwise encounter infrared-type divergences.

Note on "regularisation" In the GSF literature, one often speaks of "regularising" the field or the SF. This can mistakenly give the impression that one has introduced infinities into the problem, and that one must regularise them to recover the physical result. But in the formalism we described, one only ever deals with finite quantities. The various "regularisation" methods that have been used to remove the "singular part" of the field arise only as a practical necessity: we cannot easily determine the physical metric inside the object, nor are we interested in doing so, which prompts us to replace it with the fiction of a singular field solely *as a means* of calculating the physical metric outside the object. Computational techniques such as puncture schemes and mode-sum "regularisation" are not methods of removing singularities; they are simply methods of calculating the particular, finite quantities in question. In mode-sum regularisation, for example, one rigorously writes a spherical-harmonic mode decomposition of $h_{\mu\nu}(z)$ by decomposing $h^{\text{R}}_{\mu\nu}(z) = \lim_{x\to z}\left(h_{\mu\nu} - h^{\text{S}}_{\mu\nu}\right)$, with $h_{\mu\nu}$ being the field of a point mass. Every quantity in the calculation is finite every step of the way.

2.6.2. Alternative choices of gauge

We turn now to a brief discussion of the first technical difficulty. It has often been referred to as the "gauge problem", since, when it was first derived, the correct GSF was only known in the Lorenz gauge, meaning one had to calculate the perturbation in a gauge which, despite being convenient for describing the local singularity near the particle, does not sit very well with the global symmetries of the BH background.

The application of both the mode-sum and puncture methods involves, in some form, solving linear field equations for the metric perturbation (or for its multipole modes), which must usually be done numerically. In the above discussion we have referred specifically to the Lorenz-gauge form of these equations, (2.23) or (2.28). This form is convenient for a number of reasons: first and foremost, it yields the perturbation field in a gauge consistent with that assumed in the original GSF formulation, and so ready to be used in calculations; related to this, the singularity of $h^{\text{S}}_{\mu\nu}$ in the Lorenz gauge has

an intuitive, Coulomb-like form; finally, the field equations themselves are hyperbolic and form part of a mathematically well-posed initial value problem.

However, the direct Lorenz-gauge approach has several serious weakness [21]. Prime among these is the fact that the Lorenz-gauge field equations (2.23) cannot be decomposed into individual, decoupled multipole modes on a Kerr background, in any known form; this restricts one to time-domain numerical evolutions in 3+1 or 2+1 dimensions, which are computationally expensive and cumbersome. Second, even on a Schwarzschild background where the equations are separable into (tensorial-type) spherical harmonics, they still constitute a complicated set of 10 equations that couple between the various tensorial components. Third, time-domain evolutions of the Lorenz-gauge equations appear to suffer from linear instabilities associated with certain non-physical gauge modes, whose removal is still an open issue.

These complications have motivated the development of methods of calculating the GSF in alternative gauges, facilitated by theoretical work to extend the GSF formalism beyond the Lorenz gauge.[13] These developments have focused on the most traditionally useful gauges in BH perturbation theory: the Regge-Wheeler-Zerilli gauge in the Schwarzschild case, and the so-called "radiation gauge" in the Kerr case. In these gauges, the first-order metric perturbation can be obtained from one or more scalar quantities that satisfy fully separable field equations, reducing the numerical calculation of $h_{\mu\nu}$ to solving a set of ordinary differential equations. For example, in the case of radiation gauge, the scalar quantity is the linear perturbation of one of the Newman-Penrose curvature scalars – either Ψ_0 or Ψ_4 – constructed from the Riemann tensor; the separable field equation is then the well-known Teukolsky equation.

Unfortunately, these alternative gauges become poorly behaved in the presence of a point-particle source, introducing pathological singularities into the metric perturbation. However, by analysing the local form of the transformation to the Lorenz gauge, it has been showed that the GSF and related quantities can still be rigorously calculated from a mode-sum formula in the radiation gauge. This has effectively resolved the "gauge problem".

2.7. Orbital evolution in EMRIs – perturbative approach

The calculation of the local GSF acting on the small object in an EMRI system is a first, crucial step in the programme to model the long-term orbital dynamics and emitted GWs. One must also devise a method that uses the GSF information to construct a sufficiently accurate description of the evolving orbit and emitted radiation. Such a method is based on a systematic perturbative expansion that exploits the adiabatic nature of the inspiral process [21].

2.7.1. Bound geodesic orbits in Kerr geometry

The geometry of a Kerr spacetime is stationary – so orbital energy is conserved. However, this space is not maximally symmetric, since there is no longer spherical symmetry – so total angular momentum is not conserved and orbits are generally non-planar. Nonetheless, the geometry is axially symmetric, meaning that the projection L_z of the angular momentum along the symmetry axis (the direction of the BH spin) is conserved. As a result, the orbital plane performs a simple precession motion about the direction of the BH spin. Bound geodesic orbits around a Kerr BH are thus, generally, tri-periodic. Each orbit has: an epicyclic period T_r, equal to the time between two successive periapsis passages; a "rotational" period T_ϕ, associated with the mean azimuthal motion; a "longitudinal" period T_z, equal to the time interval between two successive minima of the longitudinal angle of the object. The combination of two precessional (libration-type) motions traces out a complicated trajectory, which is generically *ergodic*.

The conserved quantities E and L_z constitute first integrals of the geodesic EoM. The conserved mass of the particle, m, is a third such integral. Carter identified also a fourth integral, Q, that is

[13] Recall, *e.g.*, footnote 7 on page 13.

associated with a more subtle symmetry of the Kerr geometry.[14] Orbits that are initially equatorial remain equatorial (due to the symmetry of the geometry under reflection across the "equatorial plane, *i.e.* the plane orthogonal to the BH spin direction) and have $Q = 0$. The set of constants $\{E, L_z, Q\}$ completely and uniquely parametrises all geodesic orbits in Kerr spacetime, up to initial phases.[15]

The existence of four first integrals of motion – the above mentioned trio, in addition to the mass – allows to write the geodesic EoMs in a convenient form. Moreover, as noted by Carter [23], one can choose the affine parameter along the orbit so that the radial and longitudinal libration motions become manifestly decoupled from one another. Let $x^\alpha_{\rm p}(\lambda) = (t_{\rm p}(\lambda), r_{\rm p}(\lambda), \theta_{\rm p}(\lambda), \phi_{\rm p}(\lambda))$ represent the particle geodesic trajectory in Boyer-Lindquist coordinates, where λ is the Carter's parameter (also known as "Mino time"). The equations of geodesic motion in Kerr geometry take the remarkably simple form

$$\begin{aligned}
\frac{{\rm d}r_{\rm p}}{{\rm d}\lambda} &= \pm\sqrt{R(r_{\rm p})}, \\
\frac{{\rm d}\theta_{\rm p}}{{\rm d}\lambda} &= \pm\sqrt{\Theta(\cos\theta_{\rm p})}, \\
\frac{{\rm d}\phi_{\rm p}}{{\rm d}\lambda} &= \Phi_r(r_{\rm p}) + \Phi_\theta(\theta_{\rm p}), \\
\frac{{\rm d}t_{\rm p}}{{\rm d}\lambda} &= \mathcal{T}_r(r_{\rm p}) + \mathcal{T}_\theta(\theta_{\rm p}),
\end{aligned}$$

where each of the right-hand-side functions depends only on its indicated argument (as well as on the constant parameters E/m, L_z/m, Q/m and on the BH mass and spin). Note that, given initial conditions – say, $x^\alpha_{\rm p}(0)$ – the radial and longitudinal motions can be independently determined from the first two of these equations. Then, supplied with $r_{\rm p}(\lambda)$ and $\theta_{\rm p}(\lambda)$, one can solve for the azimuthal motion using the third equation. The fourth of these equations relates the affine parameter λ to the standard time coordinate t.

Geodesic motion in Kerr is thus *manifestly integrable*. As such, it proves convenient to formulate the problem in terms of action-angle variables. Let $J_\alpha = \{E, L_z, Q, m\}$ be the action variables and $q^\alpha = \{q^t, q^r, q^\theta, q^\phi\}$ be the generalised angle variables, associated with the t, r, θ and ϕ motions. The equations of geodesic motion in Kerr then take the form

$$\dot{J}_\alpha = 0, \qquad \dot{q}^\alpha = \omega^\alpha(J_\mu). \tag{2.29}$$

Here the over-dots denote differentiation with respect to any suitable parameter along the orbit – like λ, t or the proper time, with suitable redefinitions of ω^α and q^α. The four quantities ω^α are generalised frequencies associated with q^α; the parameters J_α are the principal elements of the orbit, which describe the "shape" of the orbit and determine physical attributes such as the orbital eccentricity and semimajor axis; the parameters q^α are the positional elements of the orbit, which contain the phase information of the orbit and determine physical attributes such as the (time-dependent) direction of the periapsis and orientation of the orbital plane.

In general, the radial and longitudinal libration motions lead to ergodic behaviour; however, there is a special class of geodesic orbits for which ω^r/ω^θ is a rational number. Such resonant orbits are non-ergodic and precisely periodic: the orbit completes a certain integer number of radial cycles at the same time it completes a certain integer number of longitudinal cycles. The unusual periodic nature of resonant orbits manifests itself more profoundly the smaller those integers are: that is why resonances with small integers are sometimes called "strong". During the gradual radiative inspiral of an EMRI system, the orbit will become tangent to numerous such resonances at any moment, including, generically, ones that are strong. Such resonant crossings have interesting dynamical consequences [21].

[14] Q does not have a simple physical interpretation or a Newtonian analogue, except in the weak-field or Schwarzschild limits, where it roughly corresponds to $L_x^2 + L_y^2$.

[15] One could think that the trio of periods $\{T_r, T_\phi, T_\theta\}$ is another parametrisation; actually, as pointed out in [22], it is not, as there are (infinitely many) pairs of physically distinct orbits exhibiting the same three periods.

A distinctive property of orbits around BHs is the presence of an innermost stable circular orbit (ISCO).[16] For rotating BHs, the ISCO location depends on the spin: a higher spin rate gives rise to a smaller ISCO radius for objects that co-rotate with the BH, and to a larger ISCO radius for counter-rotation motion.

2.7.2. Self-consistent and two-timescale descriptions of the orbital evolution

The full GSF obtained from the MiSaTaQuWa equation has both a *dissipative* part – which describes the change in the particle trajectory due to the emission of radiation at infinity and down the central BH event horizon, causing the small object to spiral into it – and a *conservative* part – which, for instance, modifies the precession rate of the periastron. Because the GSF is small, we know that the spiraling process must be slow: over a single radial period the worldline traced out by the object must be very nearly a geodesic of the background spacetime. Hence, the inspiral is an adiabatic process, slowly evolving through a sequence of background geodesics.

As we already noticed in section 2.4, thinking of the inspiral as an evolution through a sequence of geodesics, we can describe it as a slow change of the "constants" of motion, E, L_z and Q, happening at the rate $\dot{E}/E \propto m/M^2$; this introduces the large time scale $t_{\rm rr} \sim M/m$ into the system. Just as the small size of the object, expressed by m, led to a failure of ordinary perturbation theory (motivating the use of matched asymptotic expansions), the presence of the large scale $t_{\rm rr}$ does likewise. To clarify this point, write the full metric in the explicit form of an *ordinary Taylor expansion*,

$$\mathfrak{g}_{\mu\nu}(x,\eta) = g_{\mu\nu}(x) + \eta h^1_{\mu\nu}(x) + \eta^2 h^2_{\mu\nu}(x) + O\left(\eta^3\right), \tag{2.30}$$

where x^μ is any suitable set of background coordinates, such as the Boyer-Lindquist coordinates associated with the central BH, and $\eta \equiv m/M$ is the (small) mass ratio.[17] Clearly, the metric perturbation produced by the object will depend on the worldline z^μ, which may lead us to expect that each of the $h^n_{\mu\nu}$ depends on z^μ. But the worldline satisfying the EoM plainly depends on η, while the coefficients $h^n_{\mu\nu}$ in the above expansion are η-independent. It follows that $h^n_{\mu\nu}$ cannot depend on the whole of z^μ. Instead, we can only utilise an expansion of the form (2.30) if we also expand z^μ in the same way, *i.e.*

$$z^\mu(\tau,\eta) = z^\mu_0(\tau) + \eta z^\mu_1(\tau) + \eta^2 z^\mu_2(\tau) + O\left(\eta^3\right). \tag{2.31}$$

The zeroth-order worldline z^μ_0 is a geodesic of the background spacetime $g_{\mu\nu}$, and the GSF introduces small corrections to that worldline. One obtains evolution equations for the individual terms in z^μ_n. $h^1_{\mu\nu}$ depends only on z^μ_0 and creates the GSF that drives z^μ_1; $h^2_{\mu\nu}$ depends on z^μ_1 in addition to z^μ_0, and it (together with $h^1_{\mu\nu}$) drives z^μ_2; and so on. This approach, used systematically by Gralla and Wald [14], is the only consistent way to apply ordinary perturbation theory to the problem. Unfortunately, it is not suitable for treating long-term effects. Suppose the small object initially moves tangentially to a geodesic z^μ_0; as the inspiral progresses, the object moves further and further from z^μ_0, until the expansion (2.31) breaks down.

So, ordinary perturbation theory fails miserably in accounting for the long-term changes in the worldline. We then seek an asymptotic expansion[18] that allows the metric perturbation to depend on the full, unexpanded z^μ; we may write this as

$$\mathfrak{g}_{\mu\nu}(x,\eta) = g_{\mu\nu}(x) + \eta h^1_{\mu\nu}(x;z) + \eta^2 h^2_{\mu\nu}(x;z) + O\left(\eta^3\right),$$

with each coefficient containing an implicit functional dependence on z^μ. This is the type of expansion implicitly use in the preceding sections. It is called the *self-consistent approximation*, indicating

[16] Stable in the sense that a small perturbation applied to the orbit (*e.g.*, one that takes it away from being circular) remains small over time.

[17] In the EMRI case, an expansion in the point-particle limit roughly corresponds to an expansion of the metric in powers of the mass ratio η. In fact, even if in principle η and ϵ are different parameters, here $\eta \sim m/\mathcal{R} \equiv \epsilon$ (see appendix B for the definition of $\mathcal{R}$). In the present subsection we have set $\epsilon = m/\mathcal{R} \equiv 1$ and for clarity redefined each of the perturbations $h^n_{\mu\nu}$ with a factor η^n pulled out.

[18] Asymptotic series need not be convergent, and a convergent series need not be asymptotic. Denoting the partial sums of a series by $S_N(x)$, convergence is concerned with the behaviour of $S_N(x)$ as $N \to +\infty$ with x fixed, whereas asymptoticity (at $x = 0$) is concerned with the behaviour of $S_N(x)$ as $x \to 0$ with N fixed.

that in it, the trajectory z^μ is obtained by solving the coupled field equations and EoM together, self-consistently. This approximation successfully eliminates the growing errors associated with the expansion (2.31), and it has the advantage of being formulated in a generic (vacuum) spacetime. However it is not quite ideal for the EMRI problem, since it does not capitalise on the particular, adiabatic character of the orbital inspiral, which is only slowing evolving and hence very nearly tri-periodic. Furthermore, it is not designed to accurately incorporate a second type of slow change in the system: the slow evolution of the large BH. Over time, the BH absorbs energy and angular momentum in the form of GWs, causing its mass and spin to slowly change and leading again to growing errors.

An approximation more specifically tailored to EMRIs is offered by a *multi-scale expansion* (also known as a "*two-timescale*" expansion [24, 25]), a method of singular perturbation theory that, in the case of an EMRI, expresses the evolving worldline as a function of both "slow time" and "fast time" variables. By likewise writing the metric perturbation in terms of these variables, we can split the EoM and field equations into corresponding slow- and fast-time equations. The cleverness now becomes evident: at fixed values of slow time, the fast-time equations have the same tri-periodicity as a background geodesic, providing "snapshots" of the inspiral on the orbital time scale; the slow-time equations then govern the smooth evolution from one snapshot to the next.

This approximation scheme is most easily described in terms of the action-angle variables (J_α, q^α). As a slow time, we may use $\tilde{t} \equiv \eta t$; when Boyer-Lindquist time t is comparable to the radiation-reaction time, $t \sim M/\eta$, the slow time is of order 1, $\tilde{t} \sim M$. As fast-time variables, we can use the angle variables q^α. The expansion of the metric then becomes

$$\mathfrak{g}_{\mu\nu}\left(t, x^a, \eta\right) = g_{\mu\nu}\left(x^a\right) + \eta h^1_{\mu\nu}\left(x^a, \tilde{t}, q^\alpha\right) + \eta^2 h^2_{\mu\nu}\left(x^a, \tilde{t}, q^\alpha\right) + O\left(\eta^3\right), \tag{2.32}$$

where x^a can be any set of coordinates on spacetime slices of constant t, such as the Boyer-Lindquist coordinates $x^a = (r, \theta, \phi)$. The coefficients in the expansion are required to be bounded functions of $\tilde{t}$ and periodic functions of q^α, with Fourier expansions

$$h^n_{\mu\nu} = \sum_{k_\alpha} h^{n,k_\alpha}_{\mu\nu}\left(x^a, \tilde{t}\right) e^{-ik_\alpha q^\alpha}, \tag{2.33}$$

where the sum runs over sets of integer constants k_α. From the coefficients $h^{n,k_\alpha}_{\mu\nu}$ one can obtain the slow evolution of the "constants" J_α, which, instead of eq. (2.29), now satisfy an equation of the form

$$\dot{J}_\alpha = \sum_{k_A} \left[\eta G^{1,k_A}_\alpha\left(\tilde{t}\right) + \eta^2 G^{2,k_A}_\alpha\left(\tilde{t}\right) + O\left(\eta^3\right)\right] e^{-ik_A q^A}. \tag{2.34}$$

Here the over-dot denotes differentiation with respect to t. The Fourier coefficients G^{n,k_A}_α are constructed from those of the GSF (therefore, from the coefficients $h^{n,k_\alpha}_{\mu\nu}$). Here the sum is over pairs of integers $k_A = (k_r, k_\theta)$; because the background is stationary and axially symmetric, the "forces" in eq. (2.34) are independent of q^t and q^ϕ. Once the evolution of the action variables is determined, one can obtain the evolution of the angle variables simply by integrating

$$\dot{q}^\alpha = \omega^\alpha\left(J_A\right) = \omega^\alpha_0\left(\tilde{t}\right) + \eta\omega^\alpha_1\left(\tilde{t}\right) + O\left(\eta^2\right); \tag{2.35}$$

this gives

$$q^\alpha = \frac{1}{\eta}\int \omega^\alpha \mathrm{d}\tilde{t} = \frac{1}{\eta}\left[q^\alpha_0\left(\tilde{t}\right) + \eta q^\alpha_1\left(\tilde{t}\right) + O\left(\eta^2\right)\right]. \tag{2.36}$$

Finally, with $J_\alpha(t, \eta)$ and $q^\alpha(t, \eta)$ in hand, one obtains the metric perturbations coefficients (2.33) as ordinary functions of t, x^a and η [26].

One is generally interested in the phase evolution, *i.e.* the scalar $\psi \equiv k_A q^A$, since accurately tracking the phase of the waveform is the critical requirement for matched filtering (see subsection 1.1.1). It has been found [26] that in order to obtain the leading term in the phase evolution (2.36), one requires only the time-averaged dissipative piece of the first-order force (sometimes referred to as an "adiabatic approximation"). In order to obtain both the leading and sub-leading term in eq. (2.36), one requires the entire first-order force, along with the time-averaged dissipative piece of the second-order force

(sometimes referred to as a "post-adiabatic approximation"). Because the sub-leading term is of order η^0, a post-adiabatic approximation is absolutely necessary for accurate modelling; hence one concludes that EMRI science requires at least part of the second-order metric perturbation. This has been the primary motivation for developing second-order GSF theory.

Given its clear advantages, the multi-scale expansion is the most promising method of tackling long-term evolution; yet this method has failings of its own. One limitation is that it breaks down on large distances, where length scales become comparable to the radiation-reaction time scale. This failure manifests itself as an infrared divergence in the retarded solution. A similar failure can also occur near the large BH event horizon, which – as regards wave propagation – plays a role similar to that of infinity. Overcoming these failures calls for the introduction of additional, complementary expansions near infinity and the horizon, which can be combined with multi-scale expansion by once again appealing to the method of matched asymptotic expansions.

2.7.3. Transient resonances

Generically, all the $k_A \neq 0$ modes in eq. (2.34) are oscillatory, averaging out to zero on the radiation-reaction time. The long-term, average evolution is then driven by the (approximately constant) $k_A = 0$ modes, giving rise to an equation of the form

$$\langle \dot{J}_\alpha \rangle = \eta G_\alpha^{1,0}\left(\tilde{t}\right) + \eta^2 G_\alpha^{2,0}\left(\tilde{t}\right) + O\left(\eta^3\right), \tag{2.37}$$

where $\langle\cdot\rangle$ denotes an average over the torus with coordinates $\left(q^r, q^\theta\right)$. However, this situation changes near a resonance, where for some period of time one of the $k_A \neq 0$ modes becomes approximately stationary. In order to understand how this occurs, and its dynamical consequences, let us examine how the phase $\psi = k_A q^A$ evolves near a resonance. Suppose a resonance occurs at a time $t_{\rm res}$, meaning that the ratio $\omega^r\left(t_{\rm res}\right)/\omega^\theta\left(t_{\rm res}\right)$ is rational. For some integers $k_A = k_A^{\rm res}$, then the combination

$$k_A^{\rm res}\omega^A\left(t_{\rm res}\right) = k_r^{\rm res}\omega^r + k_\theta^{\rm res}\omega^\theta$$

vanishes. Now consider the resonant phase $\psi^{\rm res} \equiv k_A^{\rm res} q^A$ near $t_{\rm res}$. Expanding $q^A\left(t\right)$ around $q^A\left(t_{\rm res}\right)$, recalling that $\dot{q}^A = \omega^A\left(\tilde{t}\right)$, we get

$$\psi^{\rm res}\left(t\right) = \psi^{\rm res}\left(\tilde{t}_{\rm res}\right) + k_A^{\rm res}\omega^A\left(\tilde{t}_{\rm res}\right)\left(t - t_{\rm res}\right) + \frac{\eta}{2}k_A^{\rm res}\left.\frac{{\rm d}\omega^A}{{\rm d}\tilde{t}}\right|_{\tilde{t}_{\rm res}}\left(t - t_{\rm res}\right)^2 + O\left(\left(t - t_{\rm res}\right)^3\right).$$

In usual circumstances, away from a resonance, the second term on the right-hand side dominates the evolution, and the phase varies on the orbital timescale $\sim 1/k_A\omega^A \sim M$, causing the exponential term $e^{-ik_A q^A}$ to oscillate and average to zero. But because $k_A^{\rm res}\omega^A\left(\tilde{t}_{\rm res}\right) = 0$, near the resonance the third term dominates the evolution. The phase then varies slowly, on the long timescale of order

$$\sqrt{\eta k_A^{\rm res}\frac{{\rm d}\omega^A}{{\rm d}\tilde{t}}}^{-1} \sim \frac{M}{\sqrt{\eta}}.$$

Over periods shorter than this, the terms $G_A^{n,k_A^{\rm res}} e^{-i\psi^{\rm res}}$ in eq. (2.34) become approximately stationary. They then appear as additional, driving terms on the right-hand side of eq. (2.37), altering the average rate of change. After a time of order $m/\sqrt{\eta}$, when the orbit has completed its passage through resonance, the additional driving term will have shifted the action variables J_α by an amount $\sim \dot{J}_\alpha/\sqrt{\eta} \sim \sqrt{\eta}$, that induces a corresponding order-$\sqrt{\eta}$ shift to the orbital frequencies, inducing an order-$\sqrt{\eta}$ term in eq. (2.35). Since the remainder of the inspiral, after the resonance, lasts a time of order M/η, the shift in the frequencies leads to a dramatic, $\sim 1/\sqrt{\eta}$ cumulative shift in the orbital phases.

The emergence of the dynamical time scale invalidate the assumed form (2.32) of the metric, causing the multi-scale approximation to break down. The significant issue is that the resonance leads to an overall loss of accuracy, even if the passage through it is accurately modelled. And that is a subtle consequence of the sensitive dependence of the dynamics on the resonant phase $\psi^{\rm res}$ – resonant

phenomena are harbingers of the onset of dynamical chaos. The details of the radiative dynamics across the resonance depend sensitively on $\psi^{\rm res}$; thus, to model the radiative transition across a resonance one has to know the exact phase of the orbit as it enters the resonance.

Studies [27, 28] have shown that essentially all astrophysically relevant systems encounter one – possibly even more – strong resonance while in the LISA band, and it has been estimated that resonances are expected to reduce the number of EMRI detections with LISA by no more than $\sim 4\%$.

2.8. Dissipative effects and orbital evolution

As we already mentioned, the GSF is made up of a conservative and a dissipative piece. By this we mean that it can be written as a sum of a time-symmetric and a time-antisymmetric piece. Recall the decomposition (2.2) and think at the gravitational analogue, adapting that discussion to a curved spacetime; then we have a definition of $F_\alpha^{\rm cons}$ and $F_\alpha^{\rm diss}$, alluding to the retarded and advanced metric perturbations: $F_\alpha^{\rm cons}$ is the force exerted by the perturbation $\left(h^{\rm adv}_{\alpha\beta}+h^{\rm ret}_{\alpha\beta}\right)/2-h^{\rm S}_{\alpha\beta}$, while $F_\alpha^{\rm diss}$ is exerted by $\left(h^{\rm adv}_{\alpha\beta}-h^{\rm ret}_{\alpha\beta}\right)/2-h^{\rm S}_{\alpha\beta}$.

We now shift our attention to a set of concrete results from numerical GSF calculations. This series of results concerns conservative and dissipative effects: the latter will be analysed in the next sections, while the former will be the focus of chapter 4. So, we start now with a brief discussion of the dissipative (or radiative) effects, analysing in particular their impact on the orbital evolution.

2.8.1. Balance laws and adiabatic evolution

What is perhaps the most intuitive aspect of GSF physics is the back-reaction from emission of GWs. The dissipative piece of the GSF "does work"[19] on the particle, dissipating its orbital energy and angular momentum and thereby driving its gradual inspiral deeper and deeper into the potential well of the central BH. The lost energy and angular momentum of the particle is transferred into the gravitational field, which then carries them away to infinity in the form of GWs. In this intuitive picture, the loss of orbital energy and angular momentum is "balanced" by the emitted radiation. The only slight problem is that – due to the fundamental absence of a notion of local energy in GR – such a balance cannot usually be established in a momentary sense[20]: there is no meaningful way to relate the momentary rate of, say, orbital-energy dissipation to the momentary flux of energy in the GWs.

In certain circumstances, however, even in GR it is possible to formulate balance relations holding in a certain time-averaged sense. It has been shown [30, 31] for bound geodesic orbits in Kerr that "balance laws" for energy and angular momentum, when a suitable orbital averaging is applied, take the simple form

$$\begin{aligned} \left\langle F_t^{\rm diss}/u^t\right\rangle &= \left\langle \dot{\mathcal{E}}_\infty\right\rangle+\left\langle \dot{\mathcal{E}}_{\rm H}\right\rangle, \\ -\left\langle F_\phi^{\rm diss}/u^t\right\rangle &= \left\langle \dot{\mathcal{L}}_\infty\right\rangle+\left\langle \dot{\mathcal{L}}_{\rm H}\right\rangle, \end{aligned} \tag{2.38}$$

where on the right-hand side are time-averaged asymptotic fluxes of energy and angular momentum out to infinity and down the event horizon of the BH, and throughout the rest of this subsection an over-dot denotes $\mathrm{d}/\mathrm{d}t$. Some clarifying remarks are necessary. Following the logic of the adiabatic approximation, quantities on both sides are calculated while approximating the orbit as a fixed geodesic. We use $\langle\cdot\rangle$ to denote t-averaging; for intrinsically periodic geodesic orbits it suffices to average over one period of the motion, but in general for ergodic geodesic one must average over an infinite amount of time, while still treating the orbit as a fixed geodesic – or, equivalently, over the orbital $\left(q^r,q^\theta\right)$ torus. On the left-hand side of the balance equations (2.38) are the averaged t and ϕ components of the local GSF, normalised by the time component of the local four-velocity.[21] Notice that the left-hand sides here are always positive, and so are the fluxes $\left\langle \dot{\mathcal{E}}_\infty\right\rangle$ and $\left\langle \dot{\mathcal{L}}_\infty\right\rangle$. However, the fluxes $\left\langle \dot{\mathcal{E}}_{\rm H}\right\rangle$ and $\left\langle \dot{\mathcal{L}}_{\rm H}\right\rangle$ of radiation absorbed by the BH can be either positive or – for certain orbits in Kerr – negative. Negative

[19] *Ich entschuldige mich feierlich, Onkel Albert.*

[20] Unlike in the analogous flat-space relativistic electrodynamics problem that inspires it.

[21] u^t simply translates from the local proper time used in the definition of the GSF to the usual coordinate time used in defining the asymptotic fluxes.

horizon fluxes mark a super-radiant behaviour, in which some of the BH rotational energy and angular momentum are, in effect, transferred to the orbit [32].

One may also write the balance laws in the more enlightening form

$$\begin{aligned} \langle \dot{E} \rangle &= -\langle \dot{\mathcal{E}}_\infty \rangle - \langle \dot{\mathcal{E}}_{\rm H} \rangle, \\ \langle \dot{L}_z \rangle &= -\langle \dot{\mathcal{L}}_\infty \rangle - \langle \dot{\mathcal{L}}_{\rm H} \rangle. \end{aligned}$$

Hence, if we compute the fluxes on the right-hand side for a given geodesic with parameters E and L_z, these laws allow us to evolve to a new geodesic with parameters $E + \langle \dot{E} \rangle \delta t$ and $L_z + \langle \dot{L}_z \rangle \delta t$. This gives us a way of grasping the dominant, adiabatic evolution without actually resorting to a calculation of the local GSF, and that is because computational methods for asymptotic fluxes in BH perturbation theory have been well developed since the early '70s. Such calculations are much less computationally expensive than local GSF calculations and can be done in the convenient framework of Teukolsky's perturbation formalism, working with Ψ_0 or Ψ_4 instead of the full metric perturbation [33].

To get a full description of the adiabatic evolution for a generic orbit, one must also be able to calculate the evolution of the third constant of motion, Q. However, there is no way to relate $\langle \dot{Q} \rangle$ to some asymptotic fluxes of radiation, as the formula for $\langle \dot{Q} \rangle$ involves both quantities encoded in the asymptotic radiation (readily calculable within the Teukolsky formalism) and quantities locally defined as integrals along the orbit (nonetheless also calculable) [31]. With this in hand, it is completely feasible to calculate the evolution of generic EMRI orbits in Kerr – at leading, adiabatic order – without performing an actual GSF calculation.

What is possibly even more important, is that the balance laws (2.38) provide an important benchmark for GSF calculations. GSF codes all involve both a calculation of the local GSF and a calculation of the global perturbation field, from which the time-averaged asymptotic fluxes may readily be extracted. Thus, GSF codes offer the opportunity to test the balance relations (2.38) – or, conversely, depending on one's point of view, the balance relations offer the opportunity to test codes. Explicit numerical calculations demonstrating the validity of (2.38) have been actually carried out for a variety of cases [34, 35].

2.8.2. Inspiral orbits with the full (first-order) GSF

The adiabatic, flux-based calculations described above capture the main, average dissipative effect of the (first-order) GSF, but they completely neglect the conservative piece, as well as subleading effects of the dissipative first-order force that average out at leading order in the two-timescale approximation [21]. These neglected effect of the first-order GSF contribute at the first post-adiabatic order, together with that of the averaged dissipative second-order GSF. These post-adiabatic terms, including the conservative piece of the GSF, do have an important secular effect on the phase evolution in EMRI systems and it is important to include them in EMRI models.

The most direct way of incorporating these effects would be to directly integrate the EoM in a self-consistent manner, but this is a computationally challenging task. In a time-domain implementation, one would need to compute the GSF time-step by time-step, and at each step accelerate the orbit that sources the field by a suitable amount. Computing entire EMRI inspirals using this method does not seem to be a realistic prospect without a major improvement in numerical methodology. Frequency-domain implementations seem to be even less viable for such direct integrations of the EoM, as, at least in their current form, they all assume a fixed geodesic source, with a given, unevolving frequency spectrum.

Orbital-evolution calculations so far have been based on the idea of osculating geodesics. In this approach, the inspiral orbit is reconstructed as a smooth sequence of geodesics, each lying tangent to (or "osculating") the true orbit at a particular moment. This amounts to modelling the true orbit as an evolving geodesic with dynamical orbital elements. At each instant t_0, one then approximates the GSF by computing it as if, for all $t < t_0$, the particle had been moving on the geodesic that is instantaneously tangential to the evolving orbit at the instant t_0. This GSF is then used to calculate the momentary rate of change in each of the orbital elements – principal as well as positional; this is a key difference from the adiabatic approximation, which likewise models the orbit as a smooth sequence

of geodesics, but does not account for the GSF conservative effect on the evolution of the positional elements. The osculating-geodesics equations are precisely equivalent to the original EoM; the only non-numerical source of error introduced in this procedure is in the use of the GSF calculated while treating the orbit as a fixed geodesic (rather than as the true, evolving orbit). The resulting error in the momentary self-acceleration is a priori comparable to the error introduced by neglecting the second-order GSF. The problem is that assessment of the actual magnitude of error from using the "geodesic" GSF would require comparison with calculations based on a direct time-domain evolution.

From the implementations of the osculating-geodesics method one can actually see how radiation reaction drives the orbit to become more and more circular until very near the final plunge, where the eccentricity can briefly increase. Important information can be also provided in the form of the number of radians by which the periastron position will rotate under the effect of the conservative piece of the GSF from a given point until plunge – indicating the total phase error that one would be making by neglecting the effect of the conservative GSF. Notice that GSF acts to decrease the rate of periastron advance.

The full inspiral trajectories of the osculating-geodesics method are still rather slow to evaluate in practice.[22] While this evolution method represents significant progress compared to the leading-order adiabatic model, it still misses several important pieces, all formally second order in the GSF approximation, but all expected to contribute at the first post-adiabatic order just like the first-order conservative GSF. Missing is the dissipative piece of the second-order GSF, and missing too are the aforementioned corrections due to the use of the "geodesic" GSF approximation. If the small BH is spinning, the Papapetrou force must also be included, along with dissipative forcing terms associated with the spin. Eventually, work on orbital evolution has so far concentrated on the Schwarzschild problem as the main hurdle for the programme to extend its reach to the Kerr case is, again, computational.

[22] That is because the forcing terms in the evolution equations depend explicitly on the orbital phases, so one must resolve the inspiral trajectory over the short orbital timescale. This problem can be partly circumvented by means of what is known in the theory of dynamical systems as a near-identity transformation: apply a "small" transformation to the phase-space variables, such that the resulting forcing terms no longer depend on the orbital phase, while the solution to the modified problem remains uniformly close to the original solution. This procedure leads to a substantial gain in computational efficiency [36].

3. Extended theories of gravity

Causas rerum naturalium non plures admitti debere, quàm quæ et veræ sint et earum phænomenis explicandis sufficiunt.

(Isaac Newton, *Philosophiæ Naturalis Principia Mathematica*, 1687)

Gravity is allegedly the most poorly understood among the fundamental interactions. Seeking desperately the guidance of experimental facts, GWs promise to turn research into a data-driven field, letting those facts take the lead in the development of new concepts and potentially providing invaluable and much needed insights into fundamental physics.

As we already mentioned, extracting useful information from GW observations is never an easy task, for the signal is buried inside noise and extracting it requires precise modelling. Doing so in GR is already a formidable feat – it only gets harder when one tries to add new ingredients. Nonetheless, the motivation is strong, since the forthcoming experimental perspectives will provide us with an unprecedented opportunity to unveil the gravitational Universe.

3.1. Beyond Einstein's gravity

Over the last few decades several shortcomings of Einstein's theory have prompted scientists to wonder whether GR is the only theory capable of explaining the gravitational interaction. Its groundbreaking predictions – such as the light deflection, the Shapiro time delay, the precession of perihelia and the Nordtvedt effect – passed experimental tests with flying colours. Nevertheless it has received little, direct experimental verification in the strong-field regime, leaving room for alternative or extended theories of gravity which reduce to GR in a weak-field limit.

On the other hand, GR is a classical theory which is not expected to be a fundamental one; so, scientists' concern is naturally about seeking a quantum description of spacetime and gravity. When you put it that way, it may seem like quantum gravity has more to do with philosophy than science, because of that sort of "human discomfort" with the fact that we have:

- on one side, our present description of electro-weak and strong interactions, unified within the Standard Model of particle physics, which is a quantum field theory;
- on the other side, gravity, described by GR, which is governed by the laws of classical mechanics.

What we should deal with is the well-defined, scientific problem concerning the objective of getting quantitative predictions for measurements in which both gravity and Standard-Model effects cannot be neglected. We do not already have scientific data in these situations, but we already expect our current theories to fail in describing those data. Some logical and mathematical inconsistencies are encountered even before getting to the point of a numerical prediction: like two tesserae of different mosaics, GR and the Standard Model cannot be put together without modifications. The very fact that our theories fail to generate consistent predictions, albeit in some hard-to-produce contexts, should make us concerned about the general robustness of these theories. While these problems typically involve energy scales higher than the Planck energy scale

$$E_{\rm P} \equiv \sqrt{\frac{\hbar c^5}{G}} \sim 10^{19}\ {\rm GeV},$$

it is legitimate and appropriate to wonder whether the necessary new elements might affect – in some however indirect way – also some processes which do not involve higher-than-$E_{\rm P}$ energies. Even at low

energies we should, at the very least, expect small quantum-gravity corrections to the predictions of our current theories. Stated differently, the GR-picture of spacetime as a Riemannian manifold breaks down around the Planck length scale

$$\ell_{\rm P} \equiv \sqrt{\frac{\hbar G}{c^3}} \sim 10^{-33}\ {\rm cm},$$

where quantum fluctuations of the gravitational field become important, but deviations in the strong-field regime may be exhibited even at macroscopic length scales.

Proposing a tentative semi-classical description of gravity may turn out to be a crucial attempt in order to shed light on the full problem of quantum gravity. In this direction, one of the most fruitful approaches comes from the so-called *Extended Theories of Gravity* (ETGs), a broad class of theories which enlarge Einstein's theory by adding higher-order curvature invariant, drawing their motivation from effective quantum gravity actions [37], and scalar fields that can be non-minimally coupled to gravity in the action.

There are desirable advantages coming from this approach. In fact, ETGs can:

- act as *effective field theories* for describing certain effects and phenomena;
- provide a framework for obtaining *predictions* for binary evolutions and waveforms;
- combine *constraints* coming from the strong-gravity regime with other bounds from, say, weak-field regime, cosmology, astrophysics, laboratory tests, etc.

Nevertheless, there are also unpleasant drawbacks not to be underestimated, since ETGs:

- require a *theory-dependent modelling*, which can be really demanding;
- can lead to ill-posed *initial value problem* (IVP) for the metric, altering dramatically the character of the theory and potentially undermining its physical viability.

Modelling the evolution of a binary system for given initial data is a type of IVP. Recall that an IVP is said to be *well-posed* (in the sense of Hadamard) if a solution (*i*) exists, (*ii*) it is unique and (*iii*) exhibits a continuous dependence on the initial data. A theory with an ill-posed IVP cannot make predictions, so one may be tempted to use well-posedness as a selection criterion for ETGs – as, if there is to be science, a theory certainly needs to be predictive.[1] Regarding the candidate theories, well-posed formulations are generally not known, or presently available only in the form of a continuous limit to GR or linearisation around some background (see table 1 in [40]). Noticeable exceptions are scalar-tensor (ST) theories of gravity and, through mathematical equivalence, a subset of $f(R)$ theories, inheriting the well-posedness of GR through the Einstein-frame formulation.

Overcoming the technical barrier of the theory-dependent modelling may initially be achieved using *theory-agnostic*, strong-field *parametrisations*. This rather simple idea underlines the fact that, for a given theory, it proves very difficult at present to distinguish between GR and non-GR spacetimes, motivating an approach independent of the particular details of any specific theory. This brings a crystal-clear advantage, as it drastically simplifies the modelling in a theory-independent way. Obviously, all of the consequent constraints for such *phenomenological parameters* need then to be interpreted within the framework of a specific theory.

3.2. Scalar-tensor gravity

Among the copious possibilities, one of the simplest ways to modify GR is to introduce a scalar field that is non-minimally coupled to gravity. The emergence of a scalar field is a well-known generic

[1] However, as we mentioned, these theories can be thought of as effective theories, *i.e.* intrinsically limited in their range of validity and often containing spurious degrees of freedom (*e.g.* ghosts) that lead to pathological dynamics. In linearised theories it is possible to remove such pathologies, but there is no unique prescription of doing so in general. Hence, instead of setting aside theories that appear to be ill-posed, perhaps one should look for a way to "cure" them and convey predictability at non-linear level.

property of dimensionally-reduced higher-dimensional models, such as string theory.[2]

ST theories benefit from a well-posed Cauchy formulation [40], a necessary feature in order to be amenable for numerical relativity simulations. But some post-Newtonian results indicate that binary BHs in ST theory are indistinguishable from binaries in GR [41, 42], just like the end-state of collapse, isolated BHs in the general proof of Faraoni and Sotiriou [43]. Howbeit, if the scalar field is made time-dependent or is given inhomogeneities through an external mechanism like a potential, the binary also emits dipolar radiation: a dramatic effect on the binary BH dynamics – due to accretion of the scalar field by the merging BHs – and noticeable differences on the GW emission are found [44].

Furthermore, ST theories would present one of the most conspicuous strong-field deviation from GR, the spontaneous scalarization of isolated neutron stars [45], a phenomenon occurring when a non-zero value of the scalar field inside the neutron star becomes energetically favourable over the zero-field configuration, hence developing a sudden scalar charge when previously having none.

A primary motivation for studying ST gravity in the EMRI regime is to provide an additional and possibly more appropriate framework to constrain these theories by means of the effects of GSF. This task has been recently addressed by Zimmerman [47]. As we will see, even when the scalar field configuration is the trivial one, as in the BH scenario, there is room for promising constraints due to the existence of a scalar component of the GSF in addition to the gravitational one.

The leading-order EoMs for small compact objects in ST theories were provided by Gralla [48]. Building on the "Bianchi identity" for the theory, he showed that the linearised equations for a small extended body of mass m and (scalar) charge q in the Einstein frame reduces to the point-particle form

$$\begin{aligned} \delta G_{\alpha\beta} - 8\pi\delta T_{\alpha\beta} &= 8\pi \int_\gamma m(\tau)\, u_\alpha u_\beta \delta_4(x, z(\tau))\, \mathrm{d}\tau, \\ \delta\Box\phi &= -8\pi \int_\gamma q(\tau)\, \delta_4(x, z(\tau))\, \mathrm{d}\tau, \end{aligned}$$

where γ is a timelike worldline, $G_{\alpha\beta}$ is the Einstein tensor, $T_{\alpha\beta}$ is the stress-energy tensor of the bulk scalar field, $\Box \equiv g^{\alpha\beta}\nabla_\alpha\nabla_\beta$ is the covariant wave operator in the background spacetime and $\delta_4(x, z) \equiv \delta^4(x^\mu - z^\mu)/\sqrt{-g}$ is the invariant four-dimensional delta-function. Note that the charge is not constrained by any evolution equation inherent to ST theory; a separate postulate describing the internal structure of the object is required.

3.2.1. Basic concepts and field equations

Following [47], the action for a generic ST theory in the Jordan frame can be written in full generality as

$$\mathcal{S} = \frac{1}{16\pi}\int \mathrm{d}^4x\sqrt{-\bar{g}}\left[a(\bar{\phi})\,\bar{R} - b(\bar{\phi})\,\bar{g}^{\mu\nu}\bar{\nabla}_\mu\bar{\phi}\bar{\nabla}_\nu\bar{\phi} - 2c(\bar{\phi})\right] + \mathcal{S}_{\mathrm{M}}(\bar{g}^{\mu\nu}, \Psi), \tag{3.1}$$

where Ψ collectively denotes the matter fields, a, b and c are field-dependent ST parameters and the over-bar indicates that we are dealing with Jordan frame. Nothing prevents us from redefining the scalar fields in such a way that $a(\bar{\phi}) \mapsto \bar{\phi}$, leaving the theory with two free functions: the coupling function, which precedes the kinetic term (more typically written as ω), and a cosmological function. The coupling is responsible for the scalarization phenomenon, while the cosmological function provides the scalar field with mass and plays the role of the cosmological constant Λ.

In ST gravity the inertial mass and the structural properties of the small body are generally influenced by the scalar field, due to the variability of the Newton's gravitational constant – more on that later. This introduces a dependence incorporated into the point-particle model by allowing the mass to vary with the scalar field. The point-particle action for the theory is thus given by

$$\mathcal{S}_{\mathrm{M}}^{\mathrm{pp}} = -\int_\gamma m(\bar{\phi})\,\mathrm{d}\bar{\tau}.$$

[2] Considering a $(4+d)$-dimensional spacetime, one can write a higher dimensional Hilbert action, which can be dimensionally reduced by performing the integral over the extra, compactified dimensions d. This turns the action into that of a scalar field (known as the *dilaton*) coupled to gravity in the Einstein frame and characterising the size of the extra-dimensional manifold [38].

Performing the conformal transformation

$$\begin{aligned} g_{\mu\nu} &= a\left(\bar{\phi}\right)\bar{g}_{\mu\nu}, \\ \chi\left(\bar{\phi}\right) &= \int\left(\frac{3}{4}\left(\frac{a'\left(\phi\right)}{a\left(\phi\right)}\right)^2+\frac{b\left(\phi\right)}{2a\left(\phi\right)}\right)\mathrm{d}\phi, \\ A\left(\chi\right) &= a^{-1/2}\left(\bar{\phi}\right), \\ F\left(\chi\right) &= \frac{c\left(\bar{\phi}\right)}{a^2\left(\bar{\phi}\right)}, \end{aligned}$$

one gets the Einstein-frame action

$$\mathcal{S}=\frac{1}{16\pi}\int\mathrm{d}^4x\sqrt{-g}\left[R-g^{\mu\nu}\nabla_\mu\chi\nabla_\nu\chi-2F\left(\chi\right)\right]-\int_\gamma A\left(\chi\right)m\left(\chi\right)\mathrm{d}\tau. \tag{3.2}$$

A variation with respect to the metric yields, as usual, the field equations in the 22tein frame

$$G_{\alpha\beta}=8\pi\left(T^{\mathrm{bulk}}_{\alpha\beta}+T^{\mathrm{pp}}_{\alpha\beta}\right), \tag{3.3}$$

having defined

$$T^{\mathrm{bulk}}_{\mu\nu}\equiv\frac{1}{8\pi}\left[\nabla_\mu\chi\nabla_\nu\chi-\frac{1}{2}g_{\mu\nu}\left(\nabla_\sigma\chi\nabla^\sigma\chi+2F\right)\right]$$

as the stress-energy tensor associated with the bulk scalar field, and

$$T^{\mathrm{pp}}_{\alpha\beta}\equiv\int_\gamma A\left(\chi\right)m\left(\chi\right)u_\mu u_\nu\delta_4\left(x,z\right)\mathrm{d}\tau$$

as the stress-energy tensor associated with the point particle. Varying the action with respect to the scalar field produces the scalar wave equation

$$\Box\chi-F'\left(\chi\right)=8\pi\int_\gamma\frac{\mathrm{d}Am}{\mathrm{d}\chi}\delta_4\left(x,z\right)\mathrm{d}\tau, \tag{3.4}$$

which governs the evolution of χ; here again $\Box\equiv g^{\mu\nu}\nabla_\mu\nabla_\nu$ and $F'\left(\chi\right)\equiv\mathrm{d}F/\mathrm{d}\chi$.

The point-particle gives rise to perturbations of the fields around their background values. We can make this explicit by adopting the decompositions

$$g_{\mu\nu}=g^0_{\mu\nu}+h_{\mu\nu},\qquad\chi=\Phi+f, \tag{3.5}$$

for the metric and scalar fields; $g^0_{\mu\nu}\equiv g_{\mu\nu}\left(0\right)$ and $\Phi\equiv\chi\left(0\right)$ denote the background fields, *i.e.* full fields taken at $m=0$. For the sake of notation, we drop the superscript 0 as we always work with either the perturbation or the background metric, and we conveniently define the trace-reversed metric perturbation as

$$\bar{h}_{\alpha\beta}\equiv h_{\alpha\beta}-\frac{1}{2}g_{\alpha\beta}h,$$

with $h=g^{\rho\sigma}h_{\rho\sigma}$, since it has a vanishing divergence in the Lorenz gauge. We can consider a slight generalisation of the Lorenz gauge as a one-parameter family of gauge conditions, reading

$$\nabla_\alpha\bar{h}^{\alpha\beta}=2\lambda f\nabla^\beta\Phi, \tag{3.6}$$

where λ is a free dimensionless parameter, which proves useful in order to put the field equations in a weakly hyperbolic form.[3]

From eq. (3.4) stems that the perturbed scalar field then according to the linearised equation

$$\Box f+N^{\cdot}{}_{|\alpha\beta}\bar{h}^{\alpha\beta}+N^{\cdot}{}_{|\cdot}f=-4\pi\rho, \tag{3.7}$$

[3]The advantage of this gauge over the standard Lorenz gauge is that it eliminates the derivative coupling when $\lambda=1$.

with[4]

$$\begin{aligned} N^{\cdot}{}_{|\alpha\beta} &= -\left(\nabla_\alpha\nabla_\beta\Phi - \frac{1}{2}F' g_{\alpha\beta}\right), \\ N^{\cdot}{}_{|\cdot} &= -\left(2\lambda\nabla_\sigma\Phi\nabla^\sigma\Phi + F''\right), \end{aligned}$$

playing the role of external potentials, and the scalar source given by

$$\begin{aligned} \rho &= -2\int_\gamma m(\Phi) A(\Phi)\left(\frac{A'(\Phi)}{A(\Phi)} + \frac{m'(\Phi)}{m(\Phi)}\right)\delta_4(x,z)\,\mathrm{d}\tau \\ &\equiv -2\int_\gamma m(\Phi) A(\Phi)\,\alpha(\Phi)\,\delta_4(x,z)\,\mathrm{d}\tau. \end{aligned}$$

The perturbed Einstein equation coming from (3.3) takes the form

$$\delta G_{\alpha\beta} = 8\pi\left(\delta T^{\text{bulk}}_{\alpha\beta} + t_{\alpha\beta}\right),$$

where the perturbed Einstein tensor is given by

$$\begin{aligned} 2\delta G_{\alpha\beta} &= -\Box\bar{h}_{\alpha\beta} + 2\nabla_{(\alpha|}\nabla_\sigma\bar{h}^\sigma{}_{|\beta)} - g_{\alpha\beta}\nabla_\gamma\nabla_\delta\bar{h}^{\gamma\delta} - 2R^{\gamma\ \ \delta}_{\ \alpha\ \ \beta}\bar{h}_{\gamma\delta} \\ &\quad + 2R_{\rho(\alpha}\bar{h}^\rho_{\ \beta)} + g_{\alpha\beta}R^{\gamma\delta}\bar{h}_{\gamma\delta} - R\bar{h}_{\alpha\beta}, \end{aligned}$$

which is sourced by the perturbed stress-energy tensor of the bulk scalar field,

$$\begin{aligned} 8\pi\delta T^{\text{bulk}}_{\alpha\beta} &= 2\nabla_{(\alpha} f\nabla_{\beta)}\Phi - g_{\alpha\beta}\nabla_\sigma\Phi\nabla^\sigma f - g_{\alpha\beta}F' f + \frac{1}{2}g_{\alpha\beta}\bar{h}^{\gamma\delta}\nabla_\gamma\Phi\nabla_\delta\Phi \\ &\quad - \left(\frac{1}{2}\nabla_\rho\Phi\nabla^\rho\Phi + F\right)\bar{h}_{\alpha\beta} + \frac{1}{2}g_{\alpha\beta}Fg^{\sigma\nu}\bar{h}_{\sigma\nu}, \end{aligned}$$

and by the stress-energy of the point particle,

$$t_{\alpha\beta} = \int_\gamma m(\Phi) A(\Phi) u_\alpha u_\beta \delta_4(x,z)\,\mathrm{d}\tau.$$

Putting all the nightmarish pieces together, the perturbed Einstein field equation reads

$$\Box\bar{h}^{\alpha\beta} + M^{\alpha\beta}{}_{|\cdot\sigma}\nabla^\sigma f + N^{\alpha\beta}{}_{|\gamma\delta}\bar{h}^{\gamma\delta} + N^{\alpha\beta}{}_{|\cdot} f = -16\pi t^{\alpha\beta}, \tag{3.8}$$

where

$$\begin{aligned} M^{\alpha\beta}{}_{|\cdot\sigma} &= 2(1-\lambda)\left(\delta^\alpha_\sigma\nabla^\beta\Phi + \delta^\beta_\sigma\nabla^\alpha\Phi - g^{\alpha\beta}\nabla_\sigma\Phi\right), \\ N^{\alpha\beta}{}_{|\gamma\delta} &= 2R^{\alpha\ \ \beta}_{\ (\gamma\ \ \delta)} - \delta^\alpha_{(\gamma}\nabla^\beta\Phi\nabla_{\delta)}\Phi - \delta^\beta_{(\gamma}\nabla^\alpha\Phi\nabla_{\delta)}\Phi, \\ N^{\alpha\beta}{}_{|\cdot} &= -2\left[(1-\lambda)g^{\alpha\beta}F' + 2\lambda\nabla^\alpha\nabla^\beta\Phi\right]. \end{aligned}$$

Note that selecting the gauge parameter to be $\lambda = 1$, which implies $M^{\alpha\beta}{}_{|\cdot\sigma} = 0$, removes the derivate coupling from the field equations (3.8).

3.2.2. Field decomposition and regular field

The two coupled field equations, (3.7) and (3.8), can be jointly handled all at once using a convenient notation introduced in [46], called condensed index notation.

This notation introduces a meta-index A, which can stand for different types of tensorial indices: a pair of indices, $A = \alpha\beta$, which is understood to be symmetrised; or a single index, $A = \alpha$; or no index, $A = \cdot$. This enable us to collect tensor, vector and scalar fields into a single meta-object. In the

[4] The dot-and-bar notation to be briefly discussed below is borrowed from [46]. For what concerns the present discussion, one could safely replace $N^{\cdot}{}_{|\alpha\beta} \mapsto N_{\alpha\beta}$ and $N^{\cdot}{}_{|\cdot} \mapsto N$.

present case, we are concerned with a 2-tensor, $\bar{h}^{\alpha\beta}$, and a 0-tensor, f, which are then combined into the field "doublet" $\psi_A = \left\{f, \bar{h}^{\alpha\beta}\right\}$. In a similar fashion, the source terms $t_{\alpha\beta}$ and ρ are collected into the meta-object

$$\mu^A = \int_\gamma g^A_{\ M}\left(x,z\right) q^M\left(\tau\right) \delta_4\left(x,z\right) \mathrm{d}\tau,$$

where

$$q^A \equiv \begin{cases} 4m\left(\Phi\right) A\left(\Phi\right) u^\alpha u^\beta, & A = \alpha\beta \\ -2m\left(\Phi\right) A\left(\Phi\right) \alpha\left(\Phi\right) & A = \cdot \end{cases},$$

and $g^A_{\ M}$ defines the parallel propagator[5] which transports tensors at x to x' [20],

$$g^A_{\ B'}\left(x,x'\right) \equiv \begin{cases} g^{(\alpha}_{\ \ \gamma'}\left(x,x'\right) g^{\ \beta)}_{\delta'}\left(x,x'\right), & A = \alpha\beta,\ B = \gamma'\delta' \\ 1, & A = B = \cdot \end{cases}.$$

In the more explicit notation, a vertical bar has been used to separate the indices collected in A from those collected in B; *e.g.*, $N^A_{\ B}$ is denoted $N^{\alpha\beta}_{\ \ \ |\cdot}$ when $A = \alpha\beta$ and $B = \cdot$.

In this condensed notation, both perturbation equations are combined into the single equation

$$\Box\psi^A + M^A_{\ B\sigma}\nabla^\sigma\psi^B + N^A_{\ B}\psi^B = -4\pi\mu^A.$$

As it is a linear wave equation, its solution can be represented as

$$\psi_A = \int \mathrm{d}^4x' \sqrt{-\bar{g}} G_{AB'}\mu^{B'},$$

where $G_{AB'}$ is a Green's function, *i.e.* a solution of the equation

$$\Box G^A_{\ B'}\left(x,x'\right) + M^A_{\ B\sigma}\nabla^\sigma G^B_{\ B'}\left(x,x'\right) + N^A_{\ B} G^B_{\ B'}\left(x,x'\right) = -4\pi g^A_{\ B'}\delta_4\left(x,x'\right),$$

chosen to satisfy retarded boundary conditions, as all radiation is purely outgoing. The principal parts of the off-diagonal Green's functions are solutions to homogeneous equations, *e.g.* $\Box G^{\cdot}_{\ \ |\alpha\beta} = 0$; whereas the principal parts of the diagonal Green's functions have distributional sources.

It can be shown that, in a normal neighbourhood of the worldline γ described by relations $z^\mu\left(\tau\right)$, the retarded solution takes the form

$$\psi^A\left(x\right) = \frac{1}{r} U^A_{\ B'}\left(x,x'\right) q^{B'}\left(u\right) + \int_{\tau_<}^{u} V^A_{\ M}\left(x,z\right) q^M\left(\tau\right) \mathrm{d}\tau + \int_{-\infty}^{\tau_<} G^A_{\ M}\left(x,z\right) q^M\left(\tau\right) \mathrm{d}\tau, \tag{3.9}$$

where $U^A_{\ B'}$ and $V^A_{\ B'}$ are smooth bi-tensors, u is the retarded time at the point $z\left(u\right)$ where a past-directed null ray starting from x intersects the worldline, v is the advanced time of a point connecting x to the worldline by a future-directed null ray, r is the retarded distance from x to $z\left(u\right)$ and $\tau_<$ is the proper time where the worldline intersects the convex normal neighbourhood of x [20]. This solution takes the form of a local leading-order $1/r$ piece plus tail integrals.

The near-zone behaviour of the full, retarded solution (3.9) exhibits the Coulomb-type $1/r$ behaviour leading to a singularity at the location of the particle, $r = 0$. As it is now well established, the Detweiler-Whiting prescription involves a singular field $\psi_{\mathrm{S}}$to be subtracted from the retarded field, yielding its regular part $\psi_{\mathrm{R}} \equiv \psi - \psi_{\mathrm{S}}$. As in GR-GSF theory, the S-field is built from a "singular" Green's function which (*i*) vanishes in the causal future and past of the point, (*ii*) is symmetric in its arguments and (*iii*) solves the inhomogeneous wave equation [20]. The R-field is then found to be

$$\begin{aligned} \psi^A_{\mathrm{R}}\left(x\right) \ = \ & \frac{1}{2r} U^A_{\ B'}\left(x,x'\right) q^{B'}\left(u\right) - \frac{1}{2r_{\mathrm{adv}}} U^A_{\ B''}\left(x,x''\right) q^{B''}\left(v\right) \\ & \int_{\tau_<}^{u} V^A_{\ M}\left(x,z\right) q^M\left(\tau\right) \mathrm{d}\tau + \frac{1}{2}\int_u^v V^A_{\ M}\left(x,z\right) q^M\left(\tau\right) \mathrm{d}\tau \\ & + \int_{-\infty}^{\tau_<} G^A_{\ M}\left(x,z\right) q^M\left(\tau\right) \mathrm{d}\tau, \end{aligned} \tag{3.10}$$

[5] For the notion of parallel propagator, see section A.2.

where $r_{\rm adv}$ is the advanced distance from x to the worldline and v is the advanced time at that point.

To evaluate the R-fields, it proves useful to choose local coordinates; in particular, one can adopt Fermi normal coordinates[6] (FNC) adapted to the worldline of the body. To that end, the retarded/advanced time dependencies in eq. (3.10) must be translated into dependencies on $\bar{x}$ (details can be found in [20, 46]). Recalling that $s^2 = 2\sigma(x, \bar{x})$ is the perturbatively small geodesic distance between x and $\bar{x}$, the R-field in FNC through $O(s)$ can be found to be

$$\begin{aligned}\psi^A_{\rm R}(t, x^a) &= -(1 - a_c x^c)\dot{U}^A(t) + \frac{1}{3}U^A(t)\dot{a}_c x^c + \psi^A_{\rm tail} + O(s^2)\\ &= -g^A_{\ \bar{A}}\left(\dot{\hat{q}}^{\bar{A}} + \hat{q}^{\bar{B}}U^{\bar{A}}_{\ \bar{B}t}\right)(1 - a_c x^c) + g^A_{\ \bar{A}}\left(\frac{1}{3}\hat{q}^{\bar{A}}\dot{a}_c + \hat{q}^{\bar{B}}\dot{U}^{\bar{A}}_{\ \bar{B}c} + \dot{\hat{q}}^{\bar{B}}U^{\bar{A}}_{\ \bar{B}c} + \hat{q}^{\bar{B}}U^{\bar{A}}_{\ \bar{B}tc}\right)x^c\\ &\quad + \frac{1}{2}\hat{q}^{\bar{B}}R^A_{\ \bar{B}tc}x^c + \psi^A_{\rm tail} + O(s^2),\end{aligned}$$

where $a_a = a_{\bar{\alpha}}e^{\bar{\alpha}}_a$ and similar are components of tensors in FNC (evaluated at $\bar{x}$), $\hat{q}^A = q^A$, $\hat{q} = q/2$, and $\psi^A_{\rm tail}$ denotes the contribution from the chronological past,

$$\psi^A_{\rm tail}(x) = \int_{-\infty}^{t_-} G_{AB}(x, z)\, q^B(\tau)\, {\rm d}\tau.$$

This tail integral is cut short at $t_- \equiv t - 0^+$ to avoid the singular behaviour of the retarded Green's function at the coincidence point $\bar{x}$. Explicitly, the R-fields to next-to-leading-order in the geodesic displacement parameter s are given by

$$f_{\rm R} = -(1 - a_c x^c)\dot{\hat{q}} + \frac{1}{6}\hat{q}(2\dot{a}_c R_{tc})x^c + f_{\rm tail} + O(s^2) \tag{3.11}$$

and

$$\begin{aligned}\bar{h}^{\alpha\beta}_{\rm R} &= -\dot{q}^{tt}e^\alpha_{\ t}e^\beta_{\ t} - 2\dot{q}^{tb}e^{(\alpha}_{\ t}e^{\beta)}_{\ b}\\ &\quad + \left[e^\alpha_{\ t}e^\beta_{\ t}\left(3a_c\dot{q}^{tt} + \frac{1}{3}q^{tt}\dot{a}_c + \frac{1}{6}q^{tt}R_{tc}\right) + e^\alpha_{\ t}e^\beta_{\ b}\left(4\dot{q}^{tb}a_c - q^{tt}R^b_{\ tct}\right)\right]x^c\\ &\quad + \bar{h}^{\alpha\beta}_{\rm tail} + O(s^2).\end{aligned} \tag{3.12}$$

In eqs. (3.11) and (3.12), the tail terms are defined by

$$\begin{aligned}\bar{h}^{\alpha\beta}_{\rm tail} &\equiv \int_{-\infty}^{t_-} G^{\alpha\beta}_{\ \ \ |\gamma\delta}(x, z)\, q^{\gamma\delta}(\tau)\, {\rm d}\tau + \int_{-\infty}^{t_-} G^{\alpha\beta}_{\ \ \ |\cdot}(x, z)\, \hat{q}(\tau)\, {\rm d}\tau,\\ f_{\rm tail} &\equiv \int_{-\infty}^{t_-} G^{\cdot}_{\ \ |\gamma\delta}(x, z)\, q^{\gamma\delta}(\tau)\, {\rm d}\tau + \int_{-\infty}^{t_-} G^{\cdot}_{\ \ |\cdot}(x, z)\, \hat{q}(\tau)\, {\rm d}\tau.\end{aligned}$$

Since the EoMs are written in terms of the non-trace-reversed metric perturbation, the latter can be computed using FNC. It also turns out that one gets an additional local contribution from the time derivatives of the tail terms in the normal neighbourhood at the present time t [47].

3.2.3. Equations of motion

Recalling the Einstein-frame action (3.2) and the definition (3.5) of the perturbed scalar field, one can write the point-particle action in the background geometry as

$$\mathcal{S}^{0,{\rm pp}}_{\rm M} = -\int_\gamma A(\Phi)\, m(\Phi)\, {\rm d}\tau.$$

Just as we did for the field equations, we find the EoMs in the background from the stationarity of the action, which yields

$$(mA)\, a^\mu = -(mA\alpha)\, w^{\mu\nu}\nabla_\nu\Phi \equiv \mathsf{q}w^{\mu\nu}\nabla_\nu\Phi, \tag{3.13}$$

[6] See section A.3.

where $w^{\mu\nu} \equiv g^{\mu\nu} + u^\mu u^\nu$ projects along the direction orthogonal to u^μ, and $\mathsf{q} \equiv -mA\alpha$ can be interpreted as the charge of the body. The inertial mass parameter satisfies the evolution equation

$$\frac{\mathrm{d}\left(mA\right)}{\mathrm{d}\tau} = -\mathsf{q}u^\mu \nabla_\mu \Phi.$$

In the same way, the EoMs in the perturbed geometry can be found by varying the first-order perturbed point-particle action

$$\mathcal{S}_{\mathrm{M}}^{1,\mathrm{pp}} = \frac{1}{8}\int_\gamma q^{\mu\nu}h_{\mu\nu}\mathrm{d}\tau + \int_\gamma \hat{q}f\mathrm{d}\tau$$

with respect to the worldline coordinate z^μ. Working first in FNC, employing the background EoM (3.13), using the expression $R_{ta} = \nabla_t\Phi\nabla_a\Phi + O\left(s\right)$ for the Ricci tensor in terms of the background scalar field, and having defined $\mathsf{m} \equiv m\left(\Phi\right)A\left(\Phi\right)$, $\mathsf{q} \equiv \hat{q}\left(\Phi\right)$, one can show that the local SF term is given by [47]

$$\begin{aligned} F_\alpha^{\mathrm{loc}} &= \mathsf{q}w_\alpha^\beta\nabla_\beta\Phi + \left[\frac{23}{6}\mathsf{q}^2 - \frac{11}{6}\mathsf{m}^2 + \frac{1}{3}\left(\frac{\mathsf{q}}{\mathsf{m}}\right)^2\left(2\mathsf{q}^2 + \mathsf{m}\left(\alpha\mathsf{q} - \mathsf{m}\alpha'\right)\right)\right. \\ &\quad \left. -\frac{11}{3}\mathsf{m}\left(\alpha\mathsf{q} - \mathsf{m}\alpha'\right) + \left(\alpha\mathsf{q} - \mathsf{m}\alpha'\right)^2\right]w_\alpha^\beta u^\gamma\nabla_\gamma\Phi\nabla_\beta\Phi \\ &\quad + \frac{1}{3}\left(\frac{\mathsf{q}^3}{\mathsf{m}} - 5\mathsf{q}\mathsf{m}\right)w_\alpha^\beta u^\gamma\nabla_\gamma\nabla_\beta\Phi \end{aligned} \tag{3.14}$$

in covariant form.

In addition to the local term acting at the present time, one also finds a non-local contribution coming from the tail part of the SF, which depends on the chronological past. Within the past light-cone of the present, the order-reduced tail contribution is

$$\begin{aligned} F_\alpha^{\mathrm{tail}} &= w_{\alpha\beta}\left[-\alpha\mathsf{q}\nabla^\beta\Phi f_{\mathrm{tail}} + \left(\alpha\mathsf{q} - \mathsf{m}\alpha'\right)\nabla^\beta\Phi f_{\mathrm{tail}} + \mathsf{q}\nabla^\beta f_{\mathrm{tail}} - \mathsf{q}\nabla^\gamma\Phi h_\gamma^{\mathrm{tail},\beta}\right. \\ &\quad \left. -\mathsf{q}\nabla^\beta\Phi h_{\gamma\delta}^{\mathrm{tail}}u^\gamma u^\delta + \frac{1}{2}\mathsf{m}\left(\nabla^\beta h_{\gamma\delta}^{\mathrm{tail}} - 2\nabla_\gamma h_\delta^{\mathrm{tail},\beta}\right)u^\gamma u^\delta\right]. \end{aligned} \tag{3.15}$$

What is absolutely new with respect to GR is that the perturbed R-field also contributes to the evolution equation for the particle's mass, *i.e.*

$$\begin{aligned} \frac{D\mathsf{m}}{\mathrm{d}\tau} &= -\mathsf{q}u^\alpha\nabla_\alpha\Phi + 2\mathsf{q}\mathsf{m}\left(\nabla_\gamma\nabla_\delta\Phi - \frac{1}{2}F'g_{\gamma\delta}\right)u^\gamma u^\delta + \mathsf{q}^2\left(\nabla_\alpha\Phi\nabla^\alpha\Phi + \frac{1}{2}F''\right) \\ &\quad -\frac{1}{12}\mathsf{q}^2R + \mathsf{q}'^2\left(u^\alpha\nabla_\alpha\Phi\right)^2 - \mathsf{q}'\left(u^\alpha\nabla_\alpha\Phi\right)f_{\mathrm{tail}} - \mathsf{q}u^\alpha\nabla_\alpha f_{\mathrm{tail}}, \end{aligned} \tag{3.16}$$

as well as to the evolution equation for its charge,

$$\begin{aligned} \frac{D\mathsf{q}}{\mathrm{d}\tau} &= \mathsf{q}'u^\alpha\nabla_\alpha\Phi - \mathsf{q}'\mathsf{q}''\left(u^\alpha\nabla_\alpha\Phi\right)^2 - 2\mathsf{m}\mathsf{q}'\left(\nabla_\gamma\nabla_\delta\Phi + \frac{1}{2}F'g_{\gamma\delta}\right)u^\gamma u^\delta \\ &\quad -\mathsf{q}\mathsf{q}'\left(\nabla_\alpha\Phi\nabla^\alpha\Phi + \frac{1}{2}F''\right) + \frac{1}{12}\mathsf{q}\mathsf{q}'R + \mathsf{q}''u^\alpha\nabla_\alpha\Phi f_{\mathrm{tail}} + \mathsf{q}'u^\alpha\nabla_\alpha f_{\mathrm{tail}}, \end{aligned} \tag{3.17}$$

where $\mathsf{q}' \equiv \alpha\mathsf{q} - \mathsf{m}\alpha'$ and $\mathsf{q}'' \equiv 2\alpha'\mathsf{q} + \alpha\mathsf{q}' - \mathsf{m}\alpha''$. Despite the great complication of the equations, a phenomenon emerges: starting with the initial condition $\mathsf{q}\left(\Phi\left(0,\vec{x}\right)\right) = 0$, eq. (3.17) allows for a non-zero charge at some later time. In other words, the background scalar field can dynamically *induce* a charge on the worldline of the body, suggesting that scalarization could be present in the EMRI context as well [47].

3.2.4. Stationary black-hole background

The Kerr metric features a distinguished role in GR, its relevance coming mainly from the uniqueness theorem which states that – under rather general conditions – the Kerr spacetime is the only asymptotically flat, stationary, vacuum BH [39]. Therefore, the Kerr metric describes the exterior end-state of any sufficiently massive, collapsing, isolated system which obeys the cosmic censorship conjecture (assuming that some equilibrium state is reached).

In light of the general results obtained by Sotiriou and Faraoni [43] – proving that stationary and axisymmetric BHs are solution of the class of ST theories if and only if they are solutions of GR – the Kerr-Newman metric supplemented by an external, constant scalar field represents the most general stationary, axisymmetric, and vacuum solution for ST gravity.

Field equations

The condition that the BH be isolated can be translated into the asymptotic flatness requirement: the metric should approximate Minkowski and the scalar field should go over to a constant value. Resorting to the field equation that one derives from the action (3.1) by varying with respect to $\bar{g}_{\mu\nu}$ (see [43]), one gets the condition $c\left(\bar{\phi}_0\right)=0$, where $\bar{\phi}_0$ denotes a constant configuration for the scalar field; by varying with respect to $\bar{\phi}$, one also gets

$$c'\left(\bar{\phi}\right)\big|_{\bar{\phi}_0}=0,$$

In the Einstein frame, these conditions read

$$\Phi=\Phi_0, \qquad F\left(\Phi_0\right)=F'\left(\Phi\right)\big|_{\Phi_0}=0. \tag{3.18}$$

Consequently, the background Ricci tensor vanishes, implying that the stress-energy tensor of the scalar is also zero; the background Einstein field equations are thus given by the vacuum field equation, $R_{\mu\nu}=0$.

Considering a non-spinning, massive point-like object moving in the BH spacetime of a ST theory, its stress-energy perturbs the background geometry. Within the Lorenz gauge, corresponding to $\lambda=0$ in the one-parameter family of gauge conditions (3.6), the field equations (3.7), (3.8) governing the BH perturbations under the conditions in (3.18) take the completely decoupled form

$$\begin{aligned}\Box\bar{h}_{\alpha\beta}+2R_{\alpha\ \beta}^{\ \mu\ \nu}\bar{h}_{\mu\nu} &= -16\pi\mathsf{m}\int_\gamma u_\alpha u_\beta\delta_4\left(x,z\right)\mathrm{d}\tau,\\ \Box f-\mu^2 f &= -8\pi\mathsf{q}\int_\gamma\delta_4\left(x,z\right)\mathrm{d}\tau,\end{aligned}$$

where $\mathsf{m}\equiv m\left(\Phi_0\right)A\left(\Phi_0\right)$ and $\mathsf{q}\equiv -m\left(\Phi_0\right)A\left(\Phi_0\right)\alpha\left(\Phi_0\right)=-\mathsf{m}\alpha\left(\Phi_0\right)$ are the constant mass and charge of the point-like object, respectively, and $\mu^2\equiv F''\left(\Phi\right)\big|_{\Phi_0}$ is the mass associated with the scalar field. The field equation governing the metric perturbations is nothing else that the well-known, Lorenz-gauge wave equation that one gets in GR (see eq 2.15 astro2) and the scalar field equation is the curved-spacetime, massive wave equation.

The retarded solutions are given in terms of the diagonal Green's functions,

$$\begin{aligned}\bar{h}_{\alpha\beta}\left(x\right) &= 4\mathsf{m}\int_\gamma G_{\alpha\beta\gamma\delta}\left(x,z\left(\tau\right)\right)u^\gamma u^\delta\mathrm{d}\tau,\\ f\left(x\right) &= \mathsf{q}\int_\gamma G\left(x,z\left(\tau\right)\right)\mathrm{d}\tau.\end{aligned}$$

Each solution, again, has to be separately regularised, according to the Detweiler-Whiting's prescription, in order to obtain the R-fields responsible for the SF [20].

Equations of motion

For a stationary BH spacetime with a constant scalar field the leading order motion of the particle is geodesic in the background geometry, *i.e.* $a^\alpha = 0$.

The next-to-leading order motion is determined by the first-order SF which, recalling the conditions (3.18) and the fact that $R = 0$, turns out to be the sum of the tail parts of the vacuum gravitational and scalar SF in eq. (3.15), *i.e.*

$$\begin{aligned} F^\alpha_{\rm self} &= F^\alpha_{\rm tail} = F^\alpha_{\rm g,tail} + F^\alpha_{\rm s,tail} \\ &= \frac{1}{2}\mathsf{m} w^\alpha_\beta \left(\nabla^\beta h^{\rm tail}_{\gamma\delta} - 2\nabla_\gamma h^{{\rm tail},\beta}_{\delta}\right) u^\gamma u^\delta + \mathsf{q} w^\alpha_\beta \nabla^\beta f_{\rm tail}. \end{aligned} \tag{3.19}$$

The absence of coupling terms in the linearised field equations manifests itself as the lack of local terms in the equation of motion, which are seen to vanish from eq. (3.14).

The mass evolves according to (3.16), which reduces to

$$\frac{D\mathsf{m}}{{\rm d}\tau} = \frac{1}{2}\mathsf{q}^2\mu^2 - \mathsf{q}u^\alpha\nabla_\alpha f_{\rm tail},$$

while the charge, which follows eq. (3.17), evolves according to

$$\frac{D\mathsf{q}}{{\rm d}\tau} = -\frac{1}{2}\mathsf{q}\mathsf{q}'\mu^2 + \mathsf{q}' u^\alpha\nabla_\alpha f_{\rm tail}.$$

What might have been less obvious is that the mass and the charge of the particle as well experience a local self-field correction due to the mass μ associated with the scalar field, the charge $\mathsf{q} \equiv -\alpha(\Phi_0)\,\mathsf{m}(\Phi_0) = -\,\mathsf{m}'(\Phi)|_{\Phi_0}$ and the profile of the charge $\mathsf{q}' = -\,\mathsf{m}''(\Phi)|_{\Phi_0}$ at the present time, along with a history-dependent term from the scalar field.[7]

At the end of the day, the result to take home is that the EoM decompose into the form

$$\mathsf{m}a^\mu = F^\mu_{\rm back} + F^\mu_{\rm loc} + F^\mu_{\rm tail}, \tag{3.20}$$

where $\mathsf{m} \equiv m(\Phi)\,A(\Phi)$ is the ST mass parameter and with the SF splitting into:

- a term exerted by the background field in the background spacetime, $F^\mu_{\rm back}$;
- a local contribution, $F^\mu_{\rm loc}$, built from background quantities evaluated on the worldline at the present position of the particle;
- and a tail contribution, $F^\mu_{\rm tail}$, depending on the past history of the particle.

However, the absence of coupling in the asymptotically flat, constant scalar field BH scenario leads to a SF which is simply given by the tail part, in turn given by the sum of the gravitational and scalar SF; *i.e.*

$$\mathsf{m}a^\mu = F^\mu_{\rm tail} = F^\mu_{\rm g,tail} + F^\mu_{\rm s,tail}.$$

3.3. Effective gravitational constant

In the framework of ETGs the strength of gravity, given by the local value of the gravitational coupling, may depend on time and location. The variation of the gravitational constant $G_{\rm eff}$ (which is, in general, distinct from the standard Newton's constant G) implies that ETGs need *not* necessarily satisfy the strong equivalence principle (SEP). Talking about ST gravity for definiteness – or, through mathematical equivalence, $f(R)$ – this amounts to say that local gravitational physics depends on the scalar field strength.

[7] Cosmological constraints imply bounds as low as $\mu \sim 10^{-33}$ eV; so local changes in mass and charge of the particle would be far from relevant with respect to the tail pieces.

Besides the variability of the gravitational coupling, another effect that one expects on general grounds is that the standard Newtonian potential is modified by a Yukawa-like correction. This occurrence was first noticed by Stelle [49], who added terms proportional to $R_{\mu\nu}R^{\mu\nu}$ and R^2 to the Lagrangian; a recent review with an eye on astronomy and cosmology can be found in [50].

Yukawa-like corrections come into the story when one starts considering a rather general class of higher-order-ST theories (in four spacetime dimensions), whose action in the Jordan frame can be written as

$$\mathcal{S} = \int \mathrm{d}^4x\sqrt{-g}\left[F\left(R, \Box R, \ldots, \Box^k R, \phi\right) - \frac{\varepsilon}{2}g^{\mu\nu}\nabla_\mu\phi\nabla_\nu\phi\right] + \mathcal{S}_\mathrm{M}, \tag{3.21}$$

where F is a generic function of curvature invariants and a scalar field ϕ, $\varepsilon = \mp 1, 0$ is a constant which merely specifies the nature and the dynamics of the scalar field (whether it a standard, a phantom or a dynamics-free scalar field, respectively), and $\mathcal{S}_\mathrm{M}$ represents the minimally coupled ordinary matter contribution (considered as a perfect fluid). This action encompasses a great variety of theories, as lots of information is packed in F, which by definition is a Lagrangian density containing also the contribution of a potential $V(\phi)$; from time to time, one can consider to deal with functions of fields without a potential, and this will be made clear by the lower case letter f. Formalities aside, by varying the action (3.21) with respect to ϕ, one gets the Klein-Gordon-like equation

$$\varepsilon\Box\phi = -\frac{\partial F}{\partial\phi}.$$

Considering as an example $F = f(\phi)R - V(\phi)$ and $\varepsilon = -1$, with the generic functions $f(\phi)$ and $V(\phi)$ describing the coupling and the potential of the scalar field, one gets

$$\Box\phi - R\frac{\mathrm{d}f}{\mathrm{d}\phi} + \frac{\mathrm{d}V}{\mathrm{d}\phi} = 0;$$

for a constant coupling and a time-independent field, this reduces to

$$\left(\triangle - m^2\right)\phi = 0,$$

where the effective mass m is given by the minimum of the potential $V(\phi)$. It should be clear now what one should expect: the mass term smears out the $1/r$ behaviour, introducing a Yukawa-like correction to the Newtonian potential, which of course disappears at spatial infinity and therefore allows to recover the Newtonian limit and the cherished Minkowski spacetime.

Quite in general, a great deal of ETGs admit a weak field limit that, momentarily restoring SI units, can be expressed as (see again [50])

$$\Phi(r) = -\frac{GM}{r}\left(1 + \sum_{k=1}^{n}\alpha_k e^{-r/r_k}\right) \equiv -\frac{G_\mathrm{eff}M}{r},$$

where r_k is a characteristic length-scale for the interaction introduced by the k-th component of the non-Newtonian corrections, and G_eff is the effective gravitational constant. The amplitude α_k of each component is appropriately normalised to the standard Newtonian term and its sign indicate whether the correction is attractive or repulsive. As the Yukawa-like terms are negligible for $r \gg \max r_k$, G stands for the gravitational constant as measured at the spatial infinity. Also note that in this picture the inverse-square law holds, as the changes with respect to GR are encoded in the coupling G_eff. In general, any correction will introduce a characteristic range, acting at a certain length-scale, which can be translated into the mass m_k of a pseudo-particle, whose Compton's length is given by

$$r_k = \frac{\hbar}{m_k c}.$$

From a quantum-field point of view, this suggests that, in the low-energy limit, effective theories attempting to unify gravity with the other interactions introduce massive particles which carry the gravitational interaction along with the massless graviton.

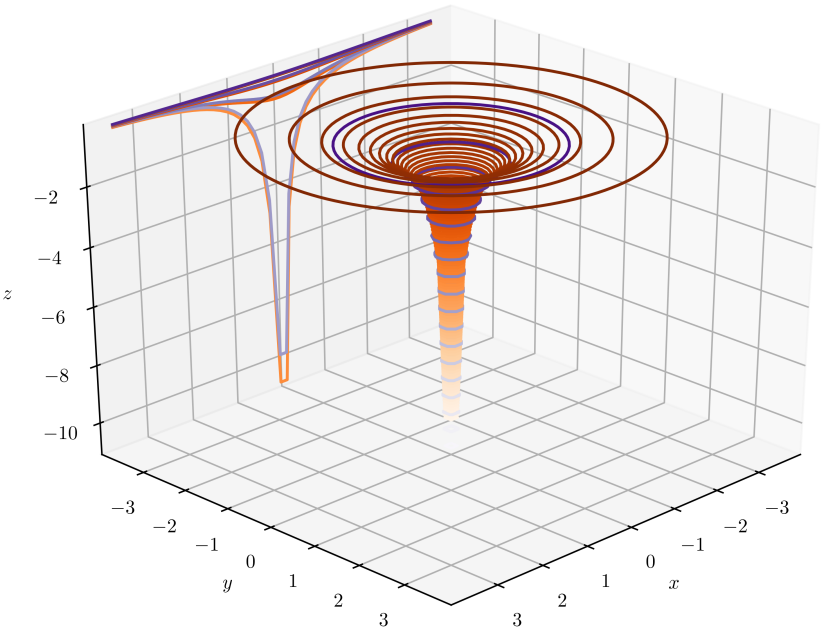

Figure 3.1.: Newtonian (purple) *vs.* non-Newtonian (orange); contour lines are projected on the coordinate plane.

Taking into account just one component – or, alternatively, the leading-order term of the summation – one has

$$\Phi(r) = -\frac{GM}{r}\left(1 + \alpha e^{-r/\lambda}\right).$$

This potential is depicted in fig. 3.1 together with the Newtonian case ($\alpha = 0$), showing how the former amounts to a tiny modification of the latter. In a nutshell, the effect of a non-Newtonian term can be conveniently parametrised by (α, λ). For distances $r \gg \lambda$, the exponential vanishes and the gravitational coupling is G; when $r \ll \lambda$, one can linearise the last equation and get

$$-\frac{GM}{r}\left(1 + \alpha\left(1 - \frac{r}{\lambda}\right)\right) \overset{r\ll\lambda}{\longrightarrow} -\frac{GM}{r}(1+\alpha),$$

which, compared to the gravitational force measured in the laboratory, implies

$$G_{\text{lab}} \simeq G(1+\alpha).$$

Here $G_{\text{lab}} = 6.674\,30\,(15) \times 10^{-11}\ \text{m}^3\text{kg}^{-1}\text{s}^{-2}$ denotes the usual Newton's constant as measured by the Cavendish-like experiments. Of course, $G \equiv G_{\text{lab}}$ in standard GR gravity.

The variability of the effective gravitational constant, as given by

$$G_{\text{eff}}(r) \equiv G\left(1 + \alpha e^{-r/\lambda}\right),$$

implies that a measured value of the ratio $G(r_1)/G(r_2) \equiv \beta$ at some distances r_1 and r_2 constrains α and λ to lie on a curve in the (α, λ) plane, *i.e.*

$$\alpha(\lambda) = \frac{\beta - 1}{e^{-r_2/\lambda} - \beta e^{-r_1/\lambda}}. \tag{3.22}$$

Varying β within the limits of experimental errors causes the curve to sweep out an allowed region. For example, in figure 3.2 we plotted eq. (3.22) using four values of β, corresponding to variations of $\pm 1\%$ and $\pm 2\%$ of $G_{\text{lab}} = 6.674\,30\,(15) \times 10^{-11}\ \text{m}^3\text{kg}^{-1}\text{s}^{-2}$. In this plot, α is expected to lie between the curves, in the shaded area. Note that we chose to plot $|\alpha|$ against λ in order to have positive values only and therefore adopt a log-log scale. Also note that, due to the singularity in eq. (3.22), neither

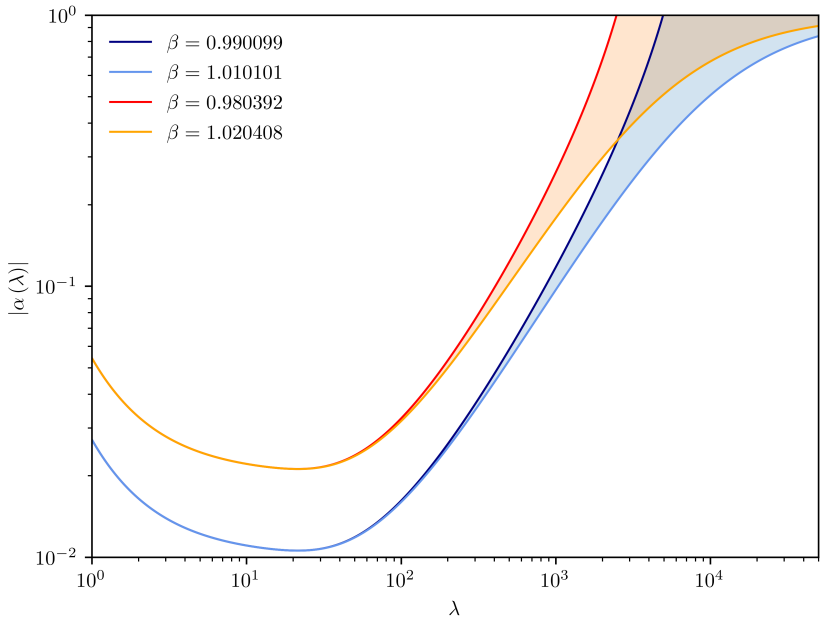

Figure 3.2.: $|\alpha|$ plotted against λ in a log-log scale for eq. (3.22).

curve can be extended indefinitely for small and large λ – of course, one should be comfortable with that, as these parameters are effective.

In conclusion, there is room for experimentally constraining λ and α, and this has been actually done in the range 1 cm $< r < 10^8$ cm with a great deal of different techniques, giving the estimates

$$|\alpha| \lesssim 10^{-2}, \qquad \lambda \gtrsim 10^4 \text{ cm}$$

for the parameters. Very-long-baseline interferometry imposes a limit of $|\alpha| \sim 1.4 \times 10^{-2}$, while binary-pulsar experiments provide a limit in the range $10^{-2} \lesssim |\alpha| \lesssim 10^{-4}$.

4. Conservative effects

Wir müssen wissen – wir werden wissen!

(David Hilbert, address to the Society of German Scientists and Physicians, 1930)

The crucial point of our analysis of the actual predictions of the GSF theory was the acknowledgement that different pieces in the MiSaTaQuWa equation contribute to different effects. As we mentioned in section 2.8, a second set of concrete results from numerical calculations comprises various specific conservative physical effects of the GSF.

At least at first order, the conservative effects can be cleanly disentangled from those of dissipation by writing the GSF as a sum of a time-symmetric piece, $F_\alpha^{\rm cons}$, and a time-antisymmetric one, $F_\alpha^{\rm diss}$, and then considering the EoM with the full GSF replaced with either $F_\alpha^{\rm cons}$ or $F_\alpha^{\rm diss}$. As it is equivalent, we can define the conservative and dissipative components as the parts of the GSF which are (correspondingly) symmetric and antisymmetric under ret $\leftrightarrow$ adv,

$$F^\alpha \left(\equiv F^\alpha_{\rm ret}\right) = F^\alpha_{\rm cons} + F^\alpha_{\rm diss},$$

where

$$F^\alpha_{\rm cons} \equiv \frac{1}{2}\left(F^\alpha_{\rm ret} + F^\alpha_{\rm adv}\right), \qquad F^\alpha_{\rm diss} \equiv \frac{1}{2}\left(F^\alpha_{\rm ret} - F^\alpha_{\rm adv}\right). \tag{4.1}$$

What we may call the "conservative dynamics" is thus described by the solution of the EoM

$$m\frac{D^2 z^\alpha}{{\rm d}\tau^2} = F^\alpha_{\rm cons}\left(z\right),$$

obtained from (2.21) by setting $F^\alpha_{\rm diss}$ to zero. The extraction of the conservative and dissipative pieces through, say, the mode-sum formula (2.26) entails a calculation of both retarded and advanced perturbations. This would normally double the computation time, as it requires one to solve the perturbation equations twice, changing the boundary conditions in order to obtain the retarded and advanced solutions. Fortunately, in the case of a Kerr (and consequently Schwarzschild) background one can avoid this extra computational burden. For an eccentric geodesic, following [52, 26], we think of the SF as a function of τ along the orbit. Without loss of generality, we can take $\tau = 0$ to correspond to a certain periapsis passage, $r_{\rm p}\left(0\right) = r_{\rm min}$. Then one gets the symmetry relation

$$F^\alpha_{\rm adv}\left(\tau\right) = \epsilon_{(\alpha)} F^\alpha_{\rm ret}\left(-\tau\right),$$

with no summation over α where $\epsilon_{(\alpha)} = (-1, 1, 1, -1)$ in Schwarzschild coordinates. This trick enables us to write

$$F^\alpha_{\rm cons}\left(\tau\right) = \frac{1}{2}\left[F^\alpha_{\rm ret}\left(\tau\right) + \epsilon_{(\alpha)} F^\alpha_{\rm ret}\left(-\tau\right)\right], \tag{4.2}$$

$$F^\alpha_{\rm diss}\left(\tau\right) = \frac{1}{2}\left[F^\alpha_{\rm ret}\left(\tau\right) - \epsilon_{(\alpha)} F^\alpha_{\rm ret}\left(-\tau\right)\right], \tag{4.3}$$

i.e. eq. (4.1) in terms of the retarded SF alone. This welcome feature allows to extract the conservative and dissipative pieces without resorting to a calculation of the advanced perturbation.

In section 4.1 we set the stage for analysing the shift in location and frequency of the ISCO in a Schwarzschild background (section 4.2) and addressing the GSF-induced modification of the spin-precession rate (section 4.3). It is worth a mention that the conservative GSF modifies also the standard relativistic periastron advance [62]. Finally, in section 4.4 we build on our previous discussions to consider these effects within the framework of ETGs. As we will occasionally remark, conservative

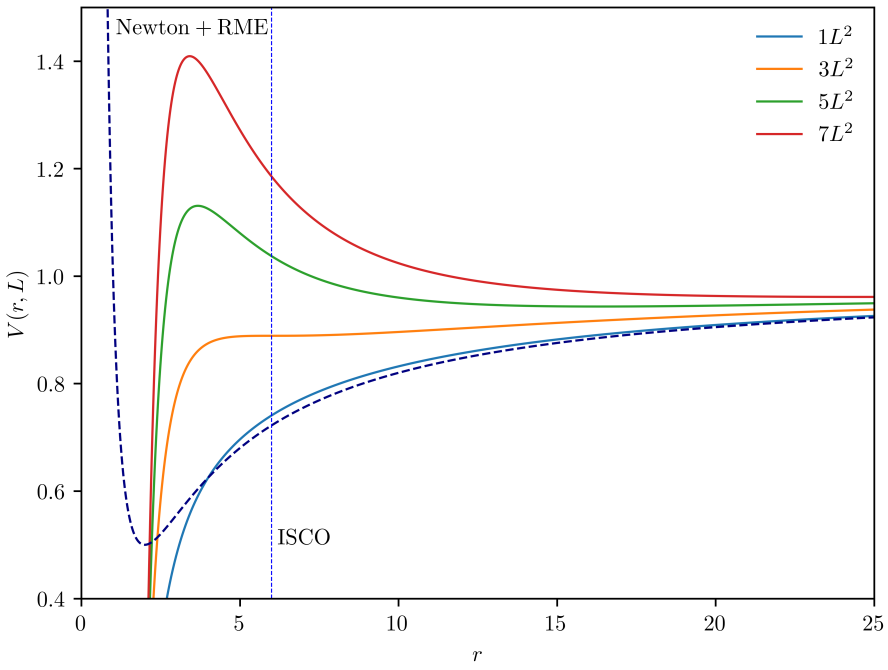

Figure 4.1.: Effective radial potential for geodesics around a Schwarzschild BH. The dashed line represents the Newtonian effective potential with rest mass energy (RME) added to match the relativistic case. In this plot, $M = 1$ and $L = 2$.

effects may happen to be buried underneath dissipative effects – as exemplified by eq. (4.26). However, this by no means implies that our analysis becomes less useful, since the conservative part of the GSF does influence the positional elements of the orbit, *i.e.* parameters containing phase information of the orbit and determining physical attributes such as the (time-dependent) direction of the periapsis and orientation of the orbital plane.

4.1. Perturbed geodesics in Schwarzschild

We consider a bound orbit of a non-spinning point-like particle with mass m around a Schwarzschild BH of mass $M \gg m$. In the limit $m \to 0$ the trajectory becomes a timelike geodesic of the background spacetime, which, in Schwarzschild coordinates, we parametrise by the proper time τ in the form

$$z^\mu_{\rm p}(\tau) = \left(t_{\rm p}(\tau), r_{\rm p}(\tau), \theta_{\rm p}(\tau), \phi_{\rm p}(\tau)\right),$$

with four-velocity $u^\mu = {\rm d}z^\mu_{\rm p}/{\rm d}\tau$. Without loss of generality we can set $\theta_{\rm p}(\tau) = \pi/2$,[1] so that the geodesic equations are given by

$$\frac{{\rm d}t_{\rm p}}{{\rm d}\tau} = E\left(1 - \frac{2M}{r_{\rm p}}\right), \qquad \frac{{\rm d}\phi_{\rm p}}{{\rm d}\tau} = \frac{L}{r_{\rm p}^2},$$

$$\left(\frac{{\rm d}r_{\rm p}}{{\rm d}\tau}\right)^2 = E^2 - V(r_{\rm p}, L), \qquad \text{with} \qquad V(r, L) \equiv \left(1 - \frac{2M}{r}\right)\left(1 + \frac{L^2}{r^2}\right), \tag{4.4}$$

where $E \equiv -u_t$ and $L \equiv u_\phi$ are the integrals of motion corresponding to the (specific, *i.e.* per unit mass) energy and angular momentum of the particle.

When $L^2 > 12M^2$, the effective potential $V(r, L)$ of the well-known radial motion develops a maximum and a minimum, and hence eccentric (bound) orbits exist – see also fig. 4.1. These orbits can

[1] Recall that the Killing vectors of the Schwarzschild metric that lead to conservation of the direction of angular momentum imply that the particle will move in a plane. We can choose this to be the equatorial plane of our coordinate system; if the particle is not in this plane, we can rotate coordinates until it is.

be parametrised by the values of $r_{\rm p}$ at the turning points, $r_{\rm min}$ and $r_{\rm max}$ ("periastron" and "apastron", respectively) or, alternatively, by the semilatus rectum p and the eccentricity e, defined through

$$p \equiv \frac{2r_{\rm min}r_{\rm max}}{M\left(r_{\rm min}+r_{\rm max}\right)}, \qquad e \equiv \frac{r_{\rm max}-r_{\rm min}}{r_{\rm max}+r_{\rm min}}. \tag{4.5}$$

From the conditions $V\left(r_{\rm min}\right)=V\left(r_{\rm max}\right)=E^2$, one obtains

$$\begin{aligned} E^2 &= \frac{(p-2-2e)\,(p-2+2e)}{p\,(p-3-e^2)}, \\ L^2 &= \frac{p^2M^2}{p-3-e^2}. \end{aligned}$$

Bound geodesics have $0\leq e<1$ and $p>6+2e$; stable circular orbits have $e=0$ and $p\geq 6$, for which E^2 equals the minimum of $V\left(r,L\right)$; the point $(p,e)=(6,0)$ is referred to as the *innermost stable circular orbit* (ISCO) [51].

The notion of ISCO can actually be looked at from another point of view in the unperturbed, exactly geodesic, $O\left(m^0\right)$ case. The radial equation that one obtains by differentiating (4.4) with respect to τ reads

$$\frac{{\rm d}^2r_{\rm p}}{{\rm d}\tau^2}=\mathcal{F}_{\rm eff}\left(r_{\rm p},L\right), \qquad \text{with} \qquad \mathcal{F}_{\rm eff}\left(r,L\right)\equiv-\frac{1}{2}\frac{\partial V\left(r,L\right)}{\partial r}. \tag{4.6}$$

When dealing with a slightly eccentric orbit that represents an e-perturbation of a circular orbit with radius $r_\circ$, one can write

$$r_{\rm p}\left(\tau\right)=r_\circ+er_1\left(\tau\right)+O\left(e^2\right), \tag{4.7}$$

assuming that $r_1\left(\tau\right)$ is e-independent. Substituting in eq. (4.6) and reading the $O\left(e\right)$ term gives

$$e\frac{{\rm d}^2r_1}{{\rm d}\tau^2}=\left.\frac{\partial\mathcal{F}_{\rm eff}\left(r_{\rm p},L\right)}{\partial r_{\rm p}}\right|_{e=0}\delta_e r_{\rm p}+\left.\frac{\partial\mathcal{F}_{\rm eff}\left(r_{\rm p},L\right)}{\partial\left(L^2\right)}\right|_{e=0}\delta_e L^2,$$

where δ_e denotes a linear variation with respect to e (holding $r_\circ$ fixed). Since

$$\delta_e L^2\equiv\left.\frac{\rm d}{{\rm d}\alpha}L^2\left(e+\alpha\delta e\right)\right|_0=\frac{2p^2M^2\delta e}{\left(p-3-e^2\right)^2}$$

vanishes for a geodesic ($\delta e=0$) and $\delta_e r_{\rm p}=er_1$, we obtain

$$\frac{{\rm d}^2r_1}{{\rm d}\tau^2}=-\omega_r^2r_1, \tag{4.8}$$

with

$$\omega_r^2\equiv-\left.\frac{\partial\mathcal{F}_{\rm eff}\left(r_{\rm p},L\right)}{\partial r_{\rm p}}\right|_{e=0}=\frac{1}{2}\left.\frac{\partial^2V\left(r_{\rm p},L\right)}{\partial r_{\rm p}^2}\right|_{e=0}=\frac{M\left(r_\circ-6M\right)}{r_\circ^3\left(r_\circ-3M\right)}.$$

Thus $O\left(e\right)$ radial motion is simple-harmonic in τ with frequency ω_r and the orbit is stable under small-e perturbation when $\omega_r^2>0$, namely for $r_\circ>6M$, and is perturbatively unstable when $\omega_r^2<0$. The ISCO is identified by the condition $\omega_r^2=0$, giving $r_{\rm isco}=6M$. Integrating eq. (4.8) with the initial condition of a periapsis passage at $\tau=0$, we obtain $r_1=-r_\circ\cos\omega_r\tau$, and hence

$$r_{\rm p}\left(\tau\right)=r_\circ\left(1-e\cos\omega_r\tau\right)+O\left(e^2\right).$$

Dealing with the finiteness of m (yet still much smaller than M), we consider the $O\left(m\right)$ correction to the orbit caused by the conservative piece of the GSF. The EoMs become [34]

$$m\frac{{\rm d}\tilde{E}}{{\rm d}\tilde{\tau}}=-F_t^{\rm cons}, \qquad m\frac{{\rm d}\tilde{L}}{{\rm d}\tilde{\tau}}=F_\phi^{\rm cons}, \tag{4.9}$$

$$m\frac{\mathrm{d}^2\tilde{r}_\mathrm{p}}{\mathrm{d}\tilde{\tau}^2} = m\mathcal{F}_\mathrm{eff}\left(\tilde{r}_\mathrm{p}, \tilde{L}\right) + F^r_\mathrm{cons}, \tag{4.10}$$

where an over-tilde indicates quantities associated with the SF-corrected orbit (*i.e.* no longer geodesic). The SF-corrected specific energy and angular-momentum $\tilde{E}(\tilde{\tau})$ and $\tilde{L}(\tilde{\tau})$ – not expected to be integrals of motion any longer – are defined through

$$\frac{\mathrm{d}\tilde{t}_\mathrm{p}}{\mathrm{d}\tilde{\tau}} = \tilde{E}\left(1 - \frac{2M}{\tilde{r}_\mathrm{p}}\right), \qquad \frac{\mathrm{d}\tilde{\phi}_\mathrm{p}}{\mathrm{d}\tilde{\tau}} = \frac{\tilde{L}}{\tilde{r}^2_\mathrm{p}}. \tag{4.11}$$

Assuming, as it is reasonable, that the orbit remains bounded under the effect of the conservative SF, we write $\tilde{r}_\mathrm{min} \leq \tilde{r}_\mathrm{p}(\tilde{\tau}) \leq \tilde{r}_\mathrm{max}$ and define p and e as in eq. (4.5), replacing $r_\mathrm{min} \to \tilde{r}_\mathrm{min}$ and $r_\mathrm{max} \to \tilde{r}_\mathrm{max}$ (we leave e and p untilded for notational clarity). Without loss of generality, we take $\tilde{r}(\tilde{\tau} = 0) = \tilde{r}_\mathrm{min}$.

The radial component F^r_cons, recalling eq. (4.2), is an even and periodic function of τ along the geodesic $z_\mathrm{p}(\tau)$; hence also, at leading order in m, an even and periodic function of $\tilde{\tau}$ along the perturbed orbit $\tilde{z}_\mathrm{p}(\tilde{\tau})$.[2] Since $\tilde{r}_\mathrm{p}(\tilde{\tau})$ too is even and periodic in $\tilde{\tau}$ – and monotonically increasing between $\tilde{r}_\mathrm{min}$ and $\tilde{r}_\mathrm{max}$ – one may express F^r_cons as a function of $\tilde{r}_\mathrm{p}$ only (for given p and e), and we shall indicate this by $F^r_\mathrm{cons} = F^r_\mathrm{cons}(\tilde{r}_\mathrm{p}; p, e)$. In eq. (4.10) the quantities $\tilde{r}_\mathrm{p}(\tau)$, $\mathrm{d}^2\tilde{r}_\mathrm{p}/\mathrm{d}\tilde{\tau}^2$ and F^r_cons are all periodic and even in τ, and we conclude that $\tilde{L}$ too is periodic and even in τ. Hence, we write $\tilde{L} = \tilde{L}(\tilde{r}_\mathrm{p}; p, e)$.

Specialising to a slightly eccentric and GSF-perturbed orbit, through $O(e)$ we can write

$$\tilde{r}_\mathrm{min} = r_\circ(1 - e) \qquad \text{and} \qquad \tilde{r}_\mathrm{max} = r_\circ(1 + e), \tag{4.12}$$

where $r_\circ = pM + O\left(e^2\right)$ is the radius of the unperturbed circular orbit. In analogy to the eq. (4.7), we write

$$\tilde{r}_\mathrm{p}(\tau) = r_\circ + e\tilde{r}_1(\tau) + O\left(e^2\right),$$

and similarly we find

$$F^r_\mathrm{cons} = F^r_\mathrm{cons}(\tilde{r}_\mathrm{p}; r_\circ, e) \qquad \text{and} \qquad \tilde{L} = \tilde{L}(\tilde{r}_\mathrm{p}; r_\circ, e).$$

Through $O\left(e^0\right)$, *i.e.* at the circular-orbit limit, $\tilde{L}$ is constant along the orbit and solving eq. (4.10) with $\mathrm{d}^2\tilde{r}_\mathrm{p}/\mathrm{d}\tilde{\tau}^2 = 0$ gives

$$\tilde{L}^2_\circ = \frac{Mr^2_\circ}{r_\circ - 3M}\left(1 - \frac{r^2_\circ}{mM}F^r_\circ\right), \tag{4.13}$$

where the subscript "$\circ$" denotes circular-orbit values; in particular, $F^r_\circ$ denotes the circular-orbit value of F^r_cons, where we omitted the label "cons" for brevity. Then, through $O(e)$, eq. (4.10) gives

$$e\frac{\mathrm{d}^2\tilde{r}_1}{\mathrm{d}\tilde{\tau}^2} = \left.\frac{\partial\mathcal{F}_\mathrm{eff}\left(\tilde{r}_\mathrm{p}, \tilde{L}_\circ\right)}{\partial\tilde{r}_\mathrm{p}}\right|_{\tilde{r}_\mathrm{p}=r_\circ}\delta_e\tilde{r}_\mathrm{p} + \left.\frac{\partial\mathcal{F}_\mathrm{eff}\left(r_\circ, \tilde{L}\right)}{\partial\left(\tilde{L}^2\right)}\right|_{\tilde{L}=\tilde{L}_\circ}\delta_e\tilde{L}^2 + m^{-1}\delta_e F^r_\mathrm{cons}. \tag{4.14}$$

The first linear variation is trivially given by $\delta_e\tilde{r}_\mathrm{p} = e\tilde{r}_1$; to evaluate $\delta_e\tilde{L}^2$, one can exploit the fact that, through $O(e)$, $\tilde{L}$ depends on e both explicitly and implicitly, through $\tilde{r}_\mathrm{p}(\tilde{\tau}; r_\circ, e)$. This can be seen expanding

$$\tilde{L}(\tilde{r}_\mathrm{p}; r_\circ, e) = \tilde{L}^{(0)}(\tilde{r}_\mathrm{p}; r_\circ) + e\tilde{L}^{(1)}(\tilde{r}_\mathrm{p}; r_\circ) + O\left(e^2\right),$$

where the coefficients $L^{(n)}$ depend on e only implicitly through $\tilde{r}_\mathrm{p}$. Now notice that sending $e \to -e$ in eq. (4.12) amounts to replacing $\tilde{r}_\mathrm{min} \leftrightarrow \tilde{r}_\mathrm{max}$; hence $\tilde{L}(\tilde{r}_\mathrm{min}; r_\circ, \mp e) = \tilde{L}(\tilde{r}_\mathrm{max}; r_\circ, \pm e)$, which can hold only if $\tilde{L}^{(1)}(\tilde{r}_\mathrm{min}; r_\circ) = -\tilde{L}^{(1)}(\tilde{r}_\mathrm{max}; r_\circ)$. Assuming that $\tilde{L}$ is a continuous function of $\tilde{r}_\mathrm{p}$, the last statement implies that $\tilde{L}^{(1)} = O(e)$ for all $\tilde{r}_\mathrm{p}$; we thus conclude that

$$\tilde{L}(\tilde{r}_\mathrm{p}; r_\circ, e) = \tilde{L}^{(0)}(\tilde{r}_\mathrm{p}; r_\circ) + O\left(e^2\right),$$

[2] This is because $\tilde{z}_\mathrm{p}(\tilde{\tau}) - z_\mathrm{p}(\tau) \sim O(m)$, while F^r_cons is already $O\left(m^2\right)$.

i.e., working at $O(e)$, $\tilde{L}$ may depend on e only implicitly through $\tilde{r}_{\rm p}$. The same argument, with the same conclusion, applies to $F^r_{\rm cons}$. Hence, through $O(e)$, we may write

$$\delta_e \tilde{L}^2 = \tilde{r}_1 \delta_{\tilde{r}_{\rm p}} \tilde{L}^2 = e\tilde{r}_1 \left.\frac{{\rm d}\tilde{L}^2}{{\rm d}\tilde{r}_{\rm p}}\right|_{e=0} \qquad \text{and} \qquad \delta_e F^r_{\rm cons} = \tilde{r}_1 \delta_{\tilde{r}_{\rm p}} F^r_{\rm cons} = e\tilde{r}_1 \left.\frac{{\rm d}F^r_{\rm cons}}{{\rm d}\tilde{r}_{\rm p}}\right|_{e=0},$$

casting eq. (4.14) into the simple-harmonic form

$$\frac{{\rm d}^2\tilde{r}_1}{{\rm d}\tilde{\tau}^2} = -\tilde{\omega}_r^2 \tilde{r}_1, \tag{4.15}$$

with

$$\tilde{\omega}_r^2 \equiv -\left.\frac{\rm d}{{\rm d}\tilde{r}_{\rm p}}\left[\mathcal{F}_{\rm eff}\left(\tilde{r}_{\rm p}, \tilde{L}\left(\tilde{r}_{\rm p}\right)\right) + m^{-1} F^r_{\rm cons}\left(\tilde{r}_{\rm p}\right)\right]\right|_{\tilde{r}_{\rm p}=r_\circ}. \tag{4.16}$$

One can finally work for a more explicit expression for the GSF-shifted radial frequency $\tilde{\omega}_r$ expanding $\tilde{L}$ and $F^r_{\rm cons}$ through $O(e)$. Like before, solving eq. (4.15) with the initial condition $\tilde{r}_{\rm p} = \tilde{r}_{\rm min}$ one finds $\tilde{r}_1 = r_\circ \cos\tilde{\omega}_r\tilde{\tau}$. Then we expand

$$F^r_{\rm cons} = F^r_\circ + e\tilde{r}_1 \left.\frac{{\rm d}F^r_{\rm cons}}{{\rm d}\tilde{r}_{\rm p}}\right|_{\tilde{r}_{\rm p}=r_\circ} + O\left(e^2\right) = F^r_\circ + eF^r_1 \cos\tilde{\omega}_r\tilde{\tau} + O\left(e^2\right), \tag{4.17}$$

where we put

$$F^r_1 \equiv -r_\circ \left.\frac{{\rm d}F^r_{\rm cons}}{{\rm d}\tilde{r}_{\rm p}}\right|_{\tilde{r}_{\rm p}=r_\circ}.$$

Similarly, by also using eq. (4.9), one finds

$$F^{\rm cons}_\phi = e\tilde{\omega}_r F^1_\phi \sin\tilde{\omega}_r\tilde{\tau} + O\left(e^2\right), \tag{4.18}$$

where we put

$$F^1_\phi \equiv mr_\circ \left.\frac{{\rm d}\tilde{L}}{{\rm d}\tilde{r}_{\rm p}}\right|_{\tilde{r}_{\rm p}=r_\circ}.$$

Finally, with eqs. (4.17), (4.18) and (4.13) in hand, the GSF-shifted radial frequency (4.16) becomes

$$\begin{aligned}
\tilde{\omega}_r^2 &= \omega_r^2 - \frac{3(r_\circ - 4M)}{r_\circ(r_\circ - 3M)} m^{-1}F^r_\circ + \frac{1}{r_\circ} m^{-1}F^r_1 - \frac{2}{r_\circ^4}\sqrt{M(r_\circ - 3M)}\, m^{-1}F^1_\phi \\
&= \frac{M}{r_\circ^3(r_\circ - 3M)}\left[r_\circ - 6M - \frac{3r_\circ^2(r_\circ - 4M)}{mM}F^r_\circ\right. \\
&\quad \left. + \frac{r_\circ^2(r_\circ - 3M)}{mM}F^r_1 - \frac{2(r_\circ - 3M)\sqrt{M(r_\circ - 3M)}}{mMr_\circ}F^1_\phi\right],
\end{aligned} \tag{4.19}$$

describing the $O(m)$ conservative shift in the radial frequency off its geodesic value. Notice that it requires the knowledge of the GSF through $O(e)$, as the circular-orbit GSF does not suffice.

4.2. ISCO shift

The GSF shifts the ISCO both in frequency and location. To see this, we resort to the defining condition

$$\tilde{\omega}_r^2\left(r_\circ = \tilde{r}_{\rm ISCO}\right) \stackrel{!}{=} 0$$

through $O(m)$. Notice that in eq. (4.19) we are allowed to substitute $r_\circ \simeq r_{\rm ISCO} = 6M$ in all the three GSF terms, since such terms are already $O(m)$ and this substitution introduces error only at $O\left(m^2\right)$. With this in mind, we readily obtain

$$\begin{aligned}
\Delta r_{\rm isco} &\equiv \tilde{r}_{\rm isco} - r_{\rm isco} = \tilde{r}_{\rm isco} - 6M \\
&= \frac{M^2}{m}\left(216F^r_{\circ{\rm isco}} - 108F^r_{1{\rm isco}} + \frac{\sqrt{3}}{M^2}F^1_{\phi{\rm isco}}\right)
\end{aligned} \tag{4.20}$$

through $O(m)$, where we have put

$$F^r_{\circ\mathrm{isco}} \equiv F^r_\circ\left(r_\circ = 6M\right), \qquad F^r_{1\mathrm{isco}} \equiv F^r_1\left(r_\circ = 6M\right), \qquad F^1_{\phi\mathrm{isco}} \equiv F^1_\phi\left(r_\circ = 6M\right).$$

This is a nice result obtained by [34], but one has to bear in mind that, since the shift Δr_{isco} in the position is gauge dependent (just like the GSF itself), it is not very useful as a benchmark for comparisons among different methods. That is why one may want to consider also another quantity, the (GSF-corrected) circular-orbit azimuthal frequency,

$$\tilde{\Omega} \equiv \frac{\mathrm{d}\tilde{\phi}_{\mathrm{p}}}{\mathrm{d}\tilde{t}_{\mathrm{p}}} = \left(\frac{\mathrm{d}\tilde{t}_{\mathrm{p}}}{\mathrm{d}\tilde{\tau}}\right)^{-1}\frac{\mathrm{d}\tilde{\phi}_{\mathrm{p}}}{\mathrm{d}\tilde{\tau}} = \frac{\tilde{u}^\phi}{\tilde{u}^t}.$$

This quantity, as discussed in subsection 4.2.2, is invariant under all $O(m)$ gauge transformations whose generators respect the helical symmetry of the circular-orbit configuration. Using eqs. (4.11) and (4.13), we get

$$\tilde{\Omega} = \Omega\left[1 - \frac{r_\circ\left(r_\circ - 3M\right)}{2mM\left(1 - \frac{2M}{r_\circ}\right)}F^r_\circ\right], \tag{4.21}$$

where

$$\Omega \equiv \sqrt{\frac{M}{r^3_\circ}} \tag{4.22}$$

is the geodesic (not GSF-corrected) value. When evaluated at $r_\circ = \tilde{r}_{\mathrm{isco}}$, the GSF-induced frequency shift $\Delta\Omega \equiv \tilde{\Omega} - \Omega$ through $O(m)$ reads [34]

$$\begin{aligned}\Delta\Omega_{\mathrm{isco}} &\equiv \tilde{\Omega}\left(\tilde{r}_{\mathrm{isco}}\right) - \Omega\left(r_{\mathrm{isco}}\right)\\ &= -\frac{1}{6^{3/2}M}\left[\frac{\Delta r_{\mathrm{isco}}}{4M} + \frac{27M}{2m}F^r_{\circ\mathrm{isco}}\right],\end{aligned} \tag{4.23}$$

where $\left(6^{3/2}M\right)^{-1} = \Omega\left(r_{\mathrm{isco}}\right)$; notice the minus sign, as the conservative piece of the GSF acts to decrease the azimuthal frequency at the ISCO.

Despite its gauge-invariant character (in the sense to be specified in subsection 4.2.2), some care must be exercised while interpreting the quantity expressed in (4.23). In fact, as pointed out in [55], the Lorenz-gauge metric perturbation has the undesirable feature that its tt component (in Schwarzschild coordinates) does not fall to zero as $r \to +\infty$, but instead to the constant value[3]

$$h_{tt} \overset{r\to\infty}{\longrightarrow} -2\alpha, \qquad \alpha \equiv \frac{m}{\sqrt{r_\circ\left(r_\circ - 3M\right)}}.$$

A gauge transformation away from Lorenz into an asymptotically flat gauge is obtained by rescaling the Lorenz-gauge time coordinate as

$$t \longrightarrow \hat{t} \equiv \sqrt{1+2\alpha}t = \left(1+\alpha\right)t + O\left(m^2\right).$$

This leads to a "rescaled" frequency, given through $O(m)$ by

$$\hat{\tilde{\Omega}} = \left(1-\alpha\right)\tilde{\Omega},$$

as this quantity has an explicit reference to t in its definition. It is a standard practice in relativistic gravity to resort to asymptotically flat coordinates, because they relate coordinates to physical units. This rescaled version now refers to an asymptotically-flat, Poincaré-Minkowski coordinate system – as the one employed in PN and EOB methods – and it therefore provides a useful reference point

[3] This peculiarity merely relates to the choice of gauge; the underlying perturbed geometry is, of course, asymptotically flat.

for comparison between GSF and PN/EOB calculations, together with a handle on the strong-field conservative dynamics. Hence, it is useful to give also the "t-rescaled" version of eq. (4.23), *i.e.*

$$\Delta\hat{\Omega}_{\rm isco} = -\frac{1}{6^{3/2}M}\left[\frac{\Delta r_{\rm isco}}{4M} + \frac{27}{2}\frac{M}{m}F^r_{\circ{\rm isco}} + \frac{1}{\sqrt{18}}\frac{m}{M}\right], \tag{4.24}$$

where we used $\alpha\,(r_\circ = 6M) = (m/M)/\sqrt{18}$.

In view of eqs. (4.24) and (4.20), the task of calculating the orbital frequency amounts to obtaining numerical values for $F^r_{\circ{\rm isco}}$, $F^r_{1{\rm isco}}$ and $F^1_{\phi{\rm isco}}$. The first one is the GSF along a strictly circular geodesic with radius $r_\circ = 6M$, while the computation of the remaining is much more delicate, as it requires to resolve numerically the small variation in SF under a small-e perturbation of a circular orbit (recall eqs. (4.17), (4.18)). The numerical values, obtained by [34] through mode-sum regularisation, are

$$\begin{aligned} F^r_{\circ{\rm isco}} &= 0.024\ 466\ 5\,(1)\,\eta/M, \\ F^r_{1{\rm isco}} &= 0.062\ 095\,(1)\,\eta/M, \\ F^1_{\phi{\rm isco}} &= -1.066\ 5\,(8)\,m, \end{aligned}$$

and

$$\Delta\hat{\Omega}_{\rm isco} = \Omega_{\rm isco} \times 0.251\ 2\,(4)\,\eta, \tag{4.25}$$

where, recall, $\Omega_{\rm isco} \equiv \left(6^{3/2}M\right)^{-1}$ is the geodesic value of the orbital frequency at $r = 6M$.

4.2.1. Transition regime

Regardless of the shape and orientation of the orbit, the orbital evolution of an EMRI system can be divided into three regimes:

- the *adiabatic inspiral regime*, during which the small body gradually descends through a sequence of geodesic orbits with slowly-varying "constants" of the motion E, L_z and Q (in the general case of a Kerr spacetime);
- the *transition regime*, during which the character of the orbit gradually changes from inspiral to plunge;
- the *plunge regime*, during which the small body plunges into the horizon along a geodesic with (nearly) fixed E, L_z and Q.

When the inspiraling body is far away from to the ISCO, it moves on a circular geodesic orbit with angular velocity Ω. As it moves, the body radiates energy into GWs at a rate $\dot{E}_{\rm GW} = -\dot{E}$; this energy loss causes the orbit to shrink adiabatically at a rate given by

$$\frac{{\rm d}r}{{\rm d}t} = -\dot{E}_{\rm GW}\frac{{\rm d}r}{{\rm d}E}.$$

The inspiral continues adiabatically until the body nears the ISCO, where its inspiral gradually ceases to be adiabatic and it enters the transition regime. Radiation reaction, as controlled by $\dot{E}_{\rm GW}$, continues to drive the orbital evolution throughout the transition regime, but gradually becomes unimportant as the transition ends and pure plunge takes over.

The breakdown of the adiabatic approximation can be understood in terms of the effective potential (4.4) governing the 1D radial motion. Throughout the inspiral and transition regimes, the body moves along a nearly circular orbit – its change of radius during each revolution around the BH is $\Delta r \ll r$, that becomes comparable to r only after the body is well into its final plunge. This near-circular motion guarantees that the radiated energy is proportional to the radiated angular momentum via the orbital angular velocity of the body [56],

$$\frac{{\rm d}E}{{\rm d}\tau} = \Omega\frac{{\rm d}L}{{\rm d}\tau}.$$

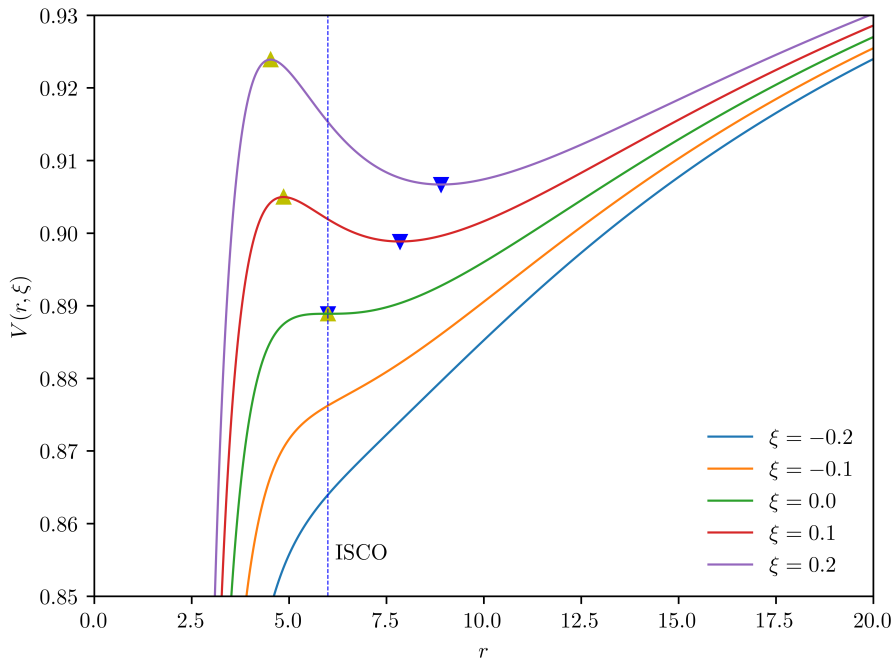

Figure 4.2.: The gradually changing effective potential for radial geodesic motion, plotted for different values of $\xi \equiv L - L_{\rm isco}$. As ξ decreases due to radiation reaction, the body at first sits at the minimum of the potential (*inspiral*); as ξ nears zero, the body cannot keep up with the inward motion of the minimum (*transition*); as ξ becomes negative, the potential has become so steep that radiation reaction is no longer important and the body plunges toward the central BH with nearly constant energy and angular momentum (*plunge*). In this plot, $M = 1$ and $L_{\rm isco} = \sqrt{12}$. Notice that the scales are different from those adopted in fig. 4.1; $\xi = 0$ corresponds to the $3L^2$ case.

Correspondingly, during transition regime, *i.e.*0 a narrow range of radii around $r_{\rm isco}$, the energy and angular momentum of the body are related by

$$E = E_{\rm isco} + \Omega\xi, \qquad \xi \equiv L - L_{\rm isco}.$$

At the end of the day, this means that we could regard the effective potential (4.4) as a function of r and the difference ξ of the angular momentum of the body from that at the ISCO. Fig. 4.2 shows

$$V\left(r,\xi\right) \equiv \left(1 - \frac{2M}{r}\right)\left(1 + \frac{\left(L_{\rm isco} + \xi\right)^2}{r^2}\right)$$

for a sequence of angular momenta around $\xi = 0$, where

$$L_{\rm isco} \equiv \sqrt{12}M$$

is a textbook-result for the value of the angular momentum at the ISCO in a Schwarzschild spacetime. As ξ decreases to zero, the maximum of the potential (denoted by ▲) migrates toward larger radii, until it merges with the minimum (▼) at the ISCO, at which point both extrema disappear; in fact, no extrema occur for $\xi < 0$.

This situation, neglecting the conservative part of the SF, has been studied in [57], showing that the width (in terms of the azimuthal frequency) of the transition regime scales with a power of the mass ratio,

$$\frac{\Delta\Omega_{\rm diss}}{\Omega_{\rm isco}} \simeq 4.387\eta^{2/5}.$$

Hence,

$$\frac{\Delta\Omega_{\rm diss}}{\Delta\hat{\Omega}_{\rm isco}} \simeq 17\eta^{-3/5}, \tag{4.26}$$

giving, for example, $\simeq 67\,700$, $17\,000$ and $4\,300$ for mass ratios $\eta = 10^{-6}$, 10^{-5} and 10^{-4}, respectively. Thus, for η in the astrophysically relevant range, the dissipative effect which drives the inspiral and the eventual transition to the plunge is dominant; nevertheless, also the conservative part of the GSF shifts the location of the ISCO away from $r_{\rm isco} = 6M$ by an $O(m)$ amount, quantified by (4.20).

What we should bring home from this discussion is that (*i*) for a true inspiral process under the full GSF the notion of a well localized ISCO gets replaced with that of a "smeared out" transition regime, and (*ii*) eq. (4.26) provides a firm justification for omitting the conservative part of the SF in the analysis. Still, there is a fundamental practical value steaming from (4.25), in that it provides an accurate, strong-field benchmark to inform the development of approximate methods, such as PN, and make useful comparisons.

4.2.2. More on gauge invariance

We have referred to the azimuthal frequency as to a gauge-invariant feature of circular orbits. But, in truth, that is not a gauge invariant in the usual sense of perturbation theory.

Let $h_{\alpha\beta} \sim O(m)$ be the metric perturbation due to the particle, given in a specific gauge, and consider the $O(\eta)$ gauge transformation

$$x^\alpha \longrightarrow x'^\alpha = x^\alpha + \xi^\alpha(x)\,.$$

As we already saw in section 2.2, eq. (2.8), this induces the gauge displacement

$$\delta_\xi h_{\alpha\beta} \equiv -2\nabla_{(\alpha}\xi_{\beta)},$$

as well as a change in the SF given by eq. (2.9), which we restate here for convenience:

$$\delta_\xi F^\alpha = 2m\nabla^{\alpha\beta\gamma}\nabla_{(\beta}\xi_{\gamma)}.$$

Restricting the discussion to gauge transformations within the family of the "physically reasonable" ones – meaning that $h_{\alpha\beta}$ has to retain its symmetries – it is possible to show that [53]

$$\begin{aligned}\delta_\xi F_t = \delta_\xi F_\theta = \delta_\xi F_\phi &= 0,\\ \delta_\xi F_r &= -3m\frac{L_\circ^2}{r_\circ^4}\xi^r.\end{aligned} \tag{4.27}$$

Hence, the "physically reasonable" gauge transformations may affect only the radial component of the GSF. Using eq. (4.27) along with

$$\delta_\xi r_\circ = \xi^r, \tag{4.28}$$

it can be immediately verified that

$$\delta_\xi\tilde{\Omega} = \delta_\xi\Omega = 0$$

within the class

$$\left\{\xi^\alpha : \xi^\phi = 0 \text{ and } (\partial_t + \Omega\partial_\phi)\,\xi^\alpha = 0\right\}$$

of "physically reasonable" transformations [54], whose generators respect the helical symmetry of the circular-orbit configuration. The first condition prevents ξ^α from generating a rotation with respect to the original coordinate system;[4] the second condition implies that, along the orbit, $\mathrm{d}\xi^\alpha/\mathrm{d}\tau = 0$ through $O(m)$, implying that $\delta_\xi u^\alpha = 0$, u^α being the four-velocity of the particle. This discussion demonstrates that the azimuthal frequency (4.21) is a gauge-invariant quantity under a suitable class of transformations, while the radial coordinate of the ISCO is not, as it is trivially seen from eq. (4.28).

The practical problem arises: in perturbation theory gauge conditions are not usually given as explicit coordinate conditions, being more typically formulated in terms of conditions on the metric

[4] Take, for instance, $\xi^\alpha = \varpi t\delta^\alpha_\phi$ (where ϖ is some $O(\eta)$ constant), under which $\phi \to \phi - \varpi t$ and thus $\Omega \to \Omega - \varpi$. This (mathematically legitimate) transformation has clearly modified the frequency, as ξ^α generated a rotation with respect to the original coordinate system. It is not really surprising to find that the frequency defined with respect to the rotating system differs from that defined in the original coordinate system.

perturbation itself; hence, given a metric perturbation in a particular gauge, how do we tell if the gauge is "physically reasonable"? This is an absolutely important question, because it impacts on our ability to assign a proper physical interpretation to the GSF results and on the ability to compare results with those obtained in other gauges or using other methods. Therefore, to address this question, we must accompany the definition of each "invariant" quantity with a precise statement about the class of "physically reasonable" gauges within which it remains invariant. That statement should be in the form of conditions on the metric perturbation, and these conditions must refer to some genuine invariant attributes of the spacetime in question. The formulation of conditions to define a class of suitable gauges is a subtle matter, which ultimately depends on the invariant quantity being calculated. If one finds that the gauges most convenient to work with – say, the Lorenz or radiation gauges – fall outside the class of gauges deemed suitable for a particular calculation, one has to introduce, *a posteriori*, a correction to the quantity being calculated, which accounts for the effect of a gauge transformation from the working gauge onto the class of physically suitable ones.

4.3. Spin precession and self-torque

In 1916, Einstein proposed three test of his theory that were supposed to exploit observations of (*i*) the precession of Mercury's perihelion, (*ii*) the deflection of the light by the Sun and (*iii*) the gravitational redshift. In the same year, de Sitter outlined a fourth test based on the precession of the spin of a system [58]. His prediction was that the rotation axis of the Earth-Moon system as it moves around the Sun experiences a non-Keplerian precession of ~ 1.9 arcsec/century, an effect confirmed by the Lunar Laser Ranging experiment [59].

De Sitter precession, also known as "geodesic precession", is a familiar effect of GR dynamics that is associated with the failure of a spin vector to return to itself after being parallel-transported along a closed curve in a curved spacetime. More extreme examples of relativistic precession are found outside the Solar system: the spin of one member of the (only known) double-pulsar system PSR J0737-3039 appears to be precessing at a rate of $\Omega_s \sim 4.8°/\mathrm{yr}$ [60]; given the orbital period of the binary of only 2.45 hours, this translates to $\sim 3.7 \times 10^{-6}$ rad of precession angle per radian of orbital revolution – a rather meagre effect, nevertheless consistent with GR. An extreme manifestation of spin precession can be found in the late inspiral of compact-object binaries, including EMRIs, where its rate can reach $O(1)$ radians per radian of orbital revolution.

This effect was first quantified in the context of GSF by [61], who considered geodesic precession for a slowly spinning compact body of mass m in a circular orbit around a Schwarzschild BH of mass $M \gg m$, calculating the $O(\eta)$ shift in the precession rate due to the back-reaction of the conservative piece of the GSF (which may be viewed as a "self-torque"). Starting with a test-particle in the limit $m \to 0$ (back-reaction effects are negligible) and assuming that its spin s_a is non-zero, but sufficiently small so as not to affect its motion,[5] one can state that such a particle follows a timelike geodesics γ,[6]

$$\frac{Du^a}{\mathrm{d}\lambda} = u^b \nabla_b u^a = 0, \tag{4.29}$$

with its spin parallel-transported along the path,

$$\frac{Ds^a}{\mathrm{d}\lambda} = u^b \nabla_b s^a = 0. \tag{4.30}$$

Here, u^a is the four-velocity of the particle, $D/\mathrm{d}\lambda \equiv u^b \nabla_b$ is the directional covariant derivative with ∇_b the covariant derivative compatible with g_{ab}. From eqs. (4.29) and (4.30) it follows that the magnitudes $g_{ab}u^a u^b = -1$ and $g^{ab}u_a u_b \equiv s^2$ are conserved along γ, like the product $u^a s_a = 0$ as well (s_a is spatial in the rest frame of the object).

[5] This amounts to the possibility of neglecting the Mathisson-Papapetrou torque, which is a factor of s/m^2 smaller than the self-torque.

[6] Latin indices $a, b, \ldots$ from the beginning of the alphabet are abstract, while the letters $i, j, \ldots$ refer to spatial components in a particular frame.

Although the magnitude of the spin is conserved, its direction can precess. Considering an orthonormal triad e_i^a along γ with legs orthogonal to u^a, eq. (4.30) can be written as

$$\frac{\mathrm{d}\boldsymbol{s}}{\mathrm{d}\tau} = \boldsymbol{\omega}_s \times \boldsymbol{s},$$

where $(\boldsymbol{s})_i = e_i^a s_a$ are the frame components of the spin vector, τ is the proper time γ and $\boldsymbol{\omega}_s$ depends on the choice of the triad. One can single out a natural class of locally defined triad by considering circular orbits and thus exploiting the existence of a time-like Killing vector field k^a satisfying $k^a|_\gamma = u^a$; one can thus choose the frames that are "comoving" with the particle in the sense that the Lie derivative along k^a vanishes, *i.e.*

$$\mathcal{L}_k e_i^a = k^b \nabla_b e_i^a - e_i^b \nabla_b k^a = 0.$$

For any frame within this class, it can be shown that both $\boldsymbol{\omega}_s$ and $\boldsymbol{\omega}_s \cdot \boldsymbol{s}$ are constant along γ, and

$$(\boldsymbol{\omega}_s)_i = \frac{1}{2} e_i^a \varepsilon_{abcd} u^b K^{cd}, \qquad K_a{}^b \equiv \nabla_a k^b\Big|_\gamma,$$

ε_{abcd} being the natural volume element associated with g_{ab}. Hence, $\boldsymbol{s}$ undergoes a precession about the fixed direction of $\boldsymbol{\omega}_s$ with a proper-time frequency $\omega_s \equiv |\boldsymbol{\omega}_s|$ satisfying

$$\omega_s^2 = -\frac{1}{2} K_a{}^b K_b{}^a, \tag{4.31}$$

where the minus sign comes from the normalisation condition for u^a. Notice that this frequency is of course independent both of the particular choice of triad within the aforementioned class and of the angle between $\boldsymbol{s}$ and $\boldsymbol{\omega}_s$. We specialise now to the Schwarzschild geometry and introduce the standard coordinates (t, r, θ, ϕ); then we choose a circular geodesic at $\theta = \pi/2$ with radius (see eq. (4.22))

$$r_\Omega \equiv \sqrt[3]{\frac{M}{\Omega^2}},$$

where $\Omega \equiv u^\phi/u^t$ is the orbital (azimuthal) frequency as seen by a distant stationary observer – as we learnt from the discussion at the end of section 4.2. The (unique) Killing field that coincides with u^a on the geodesic is given by

$$k^a = u^t \left[(\partial_t)^a + \Omega\,(\partial_\phi)^a\right],$$

where $u^t = (1 - 3M/r_\Omega)^{-1/2}$. One can also show that $\boldsymbol{\omega}_s$ is aligned with the orbital angular momentum – thus $\boldsymbol{s}$ precesses in the same sense as the orbital motion – and $\omega_s = \Omega$. Eventually, one can introduce a convenient measure of the precession effect as given by the angle of spin precession per radian of orbital motion, which for a particle on a circular orbit around a Schwarzschild BH takes the form

$$\psi \equiv 1 - \frac{\omega_s}{u^\phi} = 1 - \sqrt{1 - \frac{3M}{r_\Omega}} = \psi\,(r_\Omega)\,. \tag{4.32}$$

Before moving further, we should notice that ψ can be as large as ~ 0.3 (at the ISCO, $r_\Omega = 6M$): this is some 100° rotation of the spin axis over a single orbital period.

This geodesic effect gets perturbed when one endows the particle with a mass $m \ll M$, with the GSF altering the spin-precession rate at $O\,(m)$. As discussed in subsections 2.3.1 and 2.3.2 (see in particular eq. (2.16)), eqs. (4.29), (4.30) remain valid through $O\,(m)$ if one replaces the background metric with the smooth effective metric $\tilde{g}_{ab} = g_{ab} + h^{\mathrm{R}}_{ab}$: the perturbed orbit $\tilde{\gamma}$ is a geodesic of $\tilde{g}_{ab}$. One can thus write

$$\tilde{u}^b \tilde{\nabla}_b \tilde{u}^a = 0, \qquad \tilde{u}^b \tilde{\nabla}_b \tilde{s}^a = 0,$$

where the over-tilde denotes that we are now referring to quantities evaluated with respect to $\tilde{g}_{ab}$. In the spirit of GR equivalence principle, one can say that, just as geodesic motion in $\tilde{g}_{ab}$ corresponds to self-acceleration in g_{ab}, the parallel transport of $\tilde{s}_a$ in $\tilde{g}_{\mu\nu}$ corresponds to a self-torque in $\tilde{g}_{\mu\nu}$. Thus, an effective torque is exerted on the spinning particle by its own perturbation, modifying the rate of precession ψ away from its geodesic value (4.32). Since the dynamics we are dealing with is conservative,

there still exists a Killing vector field $\tilde{k}^a\big|_{\tilde{\gamma}} = \tilde{u}^a$ with respect to $\tilde{g}_{ab}$, which in Schwarzschild coordinates, given the constants $\tilde{u}^t$ and $\tilde{\Omega}$, can be written as

$$\tilde{k}^a = \tilde{u}^t \left[(\partial_t)^a + \tilde{\Omega} (\partial_\phi)^a \right].$$

Again, $\tilde{\Omega}$ represents an orbital frequency and the spin precession is recovered through a "tilded" version of eq. (4.31). The perturbed precession rate is defined as

$$\tilde{\psi} \equiv 1 - \frac{\tilde{\omega}_s}{\tilde{u}^\phi};$$

letting γ be a circular geodesic of the Schwarzschild background with the same orbital frequency as $\tilde{\gamma}$, $\Omega = \tilde{\Omega}$, we simply have

$$k^a = \frac{u^t}{\tilde{u}^t} \tilde{k}^a.$$

This leads us to consider $\delta\psi \equiv \tilde{\psi} - \psi$ as a function of the invariant quantity Ω, or equivalently the "gauge-invariant" radius r_Ω. Using now eq. (4.31) together with its tilded analog, one can find

$$\delta\psi = -\frac{K^{ab}\Lambda_{ab}}{2u^\phi \omega_s},$$

where

$$\Lambda_{ab} \equiv u^c \left(\nabla_{[a} h^{\mathrm{R}}_{b]c} + R_{abcd} \delta\gamma^d \right).$$

Here, R_{abcd} is the background Riemann tensor and $\delta\gamma^d$ is a deviation vector[7] between $\tilde{\gamma}$ and γ. For circular orbits in a Schwarzschild background, this reduces to

$$\delta\psi(\Omega) = r_\Omega \left(\Lambda^{rt} - \frac{\Lambda^{r\phi}}{\Omega} \right) \tag{4.33}$$

in Schwarzschild coordinates. Evaluating this expression requires knowledge of the deviation vector, which can be cast in the form (see [61] and references therein)

$$\delta\gamma^b = -\frac{1}{6} \frac{u^a u^c \left(\nabla^b h^{\mathrm{R}}_{ac} - 2\nabla_a h^{\mathrm{R}b}_{c} \right)}{\left(1 - \frac{2M}{r_\Omega} \right) (u^\phi)^2},$$

allowing the gauge-invariant[8] function $\delta\psi(\Omega)$ to be computed directly from the knowledge of $\nabla_a h^{\mathrm{R}}_{bc}$ on the worldline.

4.4. ISCO shift and spin precession in ETGs

The asymptotic flatness of the Schwarzschild spacetime guarantees that the GSF in ST and – through mathematical equivalence – $f(R)$ theories, eq. (3.20), takes the simpler form of its tail part only, eq. (3.19), that we restate here:

$$\begin{aligned} F^\alpha_{\mathrm{ETG}} &= \mathsf{m} \frac{D^2 z^\alpha}{\mathrm{d}\tau^2} = F^\alpha_{\mathrm{g,tail}} + F^\alpha_{\mathrm{s,tail}} \\ &= \frac{1}{2} \mathsf{m} w^\alpha_\beta \left(\nabla^\beta h^{\mathrm{tail}}_{\gamma\delta} - 2\nabla_\gamma h^{\mathrm{tail},\beta}_{\delta} \right) u^\gamma u^\delta + \mathsf{q} w^\alpha_\beta \nabla^\beta f_{\mathrm{tail}}. \end{aligned} \tag{4.34}$$

[7] The one-parameter family $\gamma_s(t)$ ($s \in \mathbb{R}$) of geodesics parametrised by the affine parameter t defines a smooth two-dimensional surface embedded in a manifold $\mathcal{M}$, *i.e.* the set of points $x^\mu(s,t) \in \mathcal{M}$. The coordinates on this surface may be chosen to be s and t; one thus finds two natural vector fields: the tangent vector to the geodesics, $t^\mu \equiv \partial_t x^\mu$, and the deviation vector, $s^\mu \equiv \partial_s x^\mu$. For a small displacement between neighbouring geodesics, we denote the deviation vector by $\delta\gamma^\mu$ – or $\delta\gamma^a$ in coordinate-free, abstract notation.

[8] In the same sense that we specified in subsection 4.2.2.

Recall that $\mathsf{m} \equiv m(\Phi_0) A(\Phi_0)$ and $\mathsf{q} \equiv -\mathsf{m}(\Phi_0)\alpha(\Phi_0) = -\mathsf{m}'(\Phi)|_{\Phi_0}$ are the mass and charge of the point-like object, depending on the fixed scalar field configuration Φ_0, whose evolution is approximately given by

$$\frac{D\mathsf{m}}{\mathrm{d}\tau} = -\mathsf{q}u^\alpha \nabla_\alpha f_{\rm tail}, \qquad \frac{D\mathsf{q}}{\mathrm{d}\tau} = \mathsf{q}' u^\alpha \nabla_\alpha f_{\rm tail},$$

showing how both mass and charge of the particle experience a history-dependent scalar self-field correction. This has to be compared to the analogous GR result, given by eq. (2.12), that we restate here:

$$F^\alpha_{\rm GR} = m\frac{D^2 z^\alpha}{\mathrm{d}\tau^2} = \frac{1}{2} m w^{\alpha\beta}\left(h^{\rm tail}_{\gamma\delta\beta} - 2h^{\rm tail}_{\beta\gamma\delta}\right) u^\gamma u^\delta. \tag{4.35}$$

The tail term is defined by eq. (2.13). For both expressions, $w^{\alpha\beta} \equiv g^{\alpha\beta} + u^\alpha u^\beta$ projects orthogonally to the worldline.

Some remarks are necessary at this point. *First*, the tail field is not a nice one, as it does not satisfy any particular field equation, nor is it differentiable on the worldline [20]. As we discussed in subsection 2.3.1, Detweiler and Whiting provided a more compelling form of the MiSaTaQuWa equation, replacing the direct+tail decomposition of the retarded field with the S+R decomposition. The advantage is in bolstering the *interpretation* of the fields as *self*-fields. In fact, the S-field, as the direct piece, exhibits a $1/r$, Coulomb-like divergence at the particle; but, unlike the direct piece, it satisfies the inhomogeneous linearised Einstein equation $G_{\alpha\beta}\left[h^{\rm S1}\right] = -16\pi T^1_{\alpha\beta}$, eq. (2.20). Similarly, the R-field includes the backscattered radiation in the tail, but it is also smooth on the worldline and is a vacuum solution of the homogeneous wave equation $G_{\alpha\beta}\left[h^{\rm R1}\right]$. This is what drove our discussion about a generalised equivalence principle in subsection 2.3.2. But, as already stressed in subsection 2.3.1, it should be understood that neither of these two fields represents the actual physical perturbation from the particle, which is of course the retarded field; hence both the descriptions of the perturbed motion are equally valid interpretation of the same physical effect – as such, eqs. (2.12) and (2.15) are "interchangeable". In writing eq. (2.22), we adopted the S+R decomposition; nevertheless, one can perform that "regularisation" procedure using either $h^{\rm dir}_{\alpha\beta}$ or $h^{\rm S}_{\alpha\beta}$, because both produce the same final value for the GSF ([18] and references therein; see also fn. 11 on page 17).

Second, eq. (4.34) lives in the Einstein frame, while eq. (4.35) lives in the Jordan frame. This might seem a serious, physical concern, suggesting that the Einstein-frame action (3.2) may not describe the same physics as the Jordan-frame action (3.1) does. It should be noticed, however, that those are just different *representations* of the same theory. A misconception is that sometimes one speaks about violation of the Einstein equivalence principle or the weak equivalence principle in the Einstein frame simply implying that the Einstein-frame metric is not the metric whose geodesics coincide with free-fall trajectories. As discussed in detail in [63], this last point – even though it is correct – by no means imply a violation of the equivalence principles, simply because all that is required is that there exist *some* metric whose geodesics coincide with free-fall trajectories. And indeed there is one: the metric of the Jordan frame.[9] It should be clear that whether or not we choose to represent a theory with respect to this metric is, after all, simply irrelevant.

Coming back to the point of our discussion, in ETGs, even with a constant background scalar field, the scalar perturbation is also expected to influence the motion of a point-particle orbiting a ST-BH, as it would experience both the metric and scalar tail pieces of the GSF. At the end of the day, these deviations from GR would likely have observational consequences. In order to establish a full comparison between the theories, one can imagine to set up a numerical calculation of the GSF on a point-particle around a BH in ST theory using, *e.g.*, the mode-sum method; but doing so is, in general, a quite complicated task – we discussed this tremendous technical problem in some detail in section 2.6. Hence, we will rather follow the path outlined in section 3.3.

In order to account for the variability of Newton's constant, we must first restore G and c in our equations; this can be easily accomplished through dimensional analysis. Considering first the orbital

[9] Actually, the popular formulation of the Einstein equivalence principle in terms of the metric postulates (metric theories) is representation-dependent: in this sense, the Jordan-frame metric has a distinguished status with respect to any other conformal metric. However, it could be misleading to refer to a representation as "physical" or not; the fact that it is better highlighted in the Jordan frame that the theory satisfies the Einstein equivalence principle does not make this frame preferable, in much the same way that the local Lorentz frame is not a preferred one.

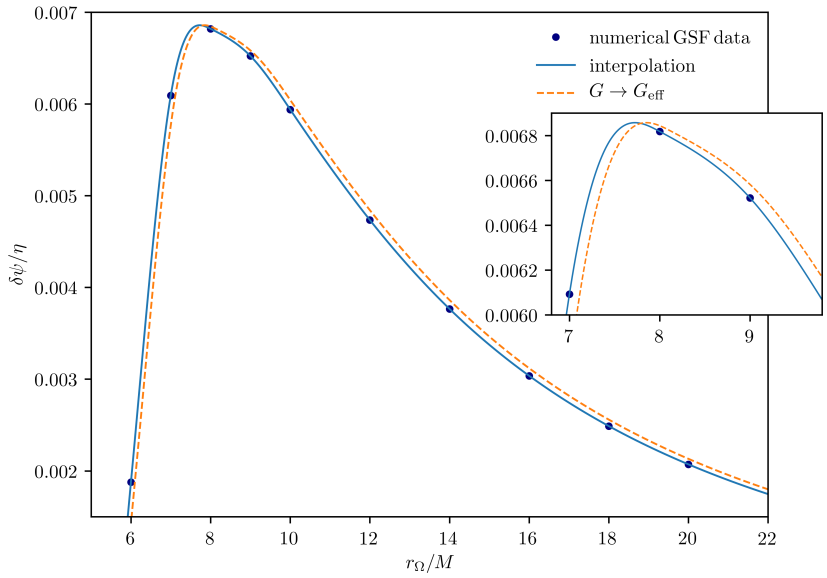

Figure 4.3.: The conservative correction to ψ as a function of the orbital radius r_Ω. The solid line interpolates the numerical GSF data [61] and the dashed line shows the estimate of deviations from GR for $\alpha = 5 \times 10^{-2}$. The inset zooms on the peak.

frequency at the ISCO, this means that

$$[\Omega_{\rm isco}] \stackrel{!}{=} \frac{1}{\mathrm{T}},$$

where square brackets indicate that we consider the physical dimensions of the quantity, which are denoted by capital Roman types. Hence, we require that

$$\mathrm{T}^{-1} \stackrel{!}{=} \left[\frac{G^a c^b}{6^{3/2} M}\right] = \frac{1}{\mathrm{M}} \left(\frac{\mathrm{L}^3}{\mathrm{MT}^2}\right)^a \left(\frac{\mathrm{L}}{\mathrm{T}}\right)^b = \mathrm{L}^{3a+b}\mathrm{M}^{-(1+a)}\mathrm{T}^{-(2a+b)},$$

which is solved by $a = -1$ and $b = 3$, therefore yielding

$$\Omega_{\rm isco} = \frac{c^3}{6^{3/2} GM}$$

in SI units. We are now in a position to incorporate the correction discussed in section 3.3 by simply replacing $G \to G\,(1+\alpha)$. In order to compare with eq. (4.24), we set again $G = c = 1$, so that our estimate of the frequency at the ISCO becomes

$$\Omega^{\rm ETG}_{\rm isco} = \frac{1}{6^{3/2} M\,(1+\alpha)} = \Omega^{\rm GR}_{\rm isco}\,(1+\alpha)^{-1}\,.$$

Our *ansatz* is that the shift in the orbital frequency retains its functional form in passing from GR to ETGs, *i.e.* $\Delta\Omega^{\rm ETG}_{\rm isco} \equiv \Delta\Omega^{\rm GR}_{\rm isco}\,(G_{\rm eff})$; recalling the numerical result given by eq. (4.25), this yields

$$\frac{\Delta\Omega^{\rm ETG}_{\rm isco}}{\Omega^{\rm GR}_{\rm isco}} = \frac{0.251\ 2\,(4)\,\eta}{1+\alpha}.$$

In a similar fashion we can deal with the spin-precession shift given by eq. (4.33). Following [61], we plot $\delta\psi/\eta$ against

$$\frac{r_\Omega}{M} = \sqrt[3]{\frac{G}{\Omega^2 M^2}},$$

in SI units. Repeating the previous reasoning, we get

$$r^{\rm ETG}_\Omega = r^{\rm GR}_\Omega \sqrt[3]{1+\alpha} \tag{4.36}$$

for the "gauge-invariant" radius. In fig. 4.3 we plot the numerical results for $\delta\psi$ calculated in [61] and compare with the shifted result given by eq. (4.36); the solid line is the interpolating cubic spline for the GSF data, intended only as a guide to the eye as no model is assumed for fitting.

Discussion and conclusion

The present work moved between two areas of current interest in theoretical physics, namely gravitational self-force and extended theories of gravity. In particular, it has been analysed the still largely unexplored observational potential that would descend from the latter when we could observe strong-gravity phenomena such as those that EMRIs will offer us. The full problem is still unapproachable due to the technical difficulties that inevitably brings with it. However, the present can be considered a preliminary work highlighting a wide margin of investigation, which could provide essential indications to be tested when LISA will be launched into orbit (presumably) in 2034.

We have followed an already firmly established indication of modifications to the equations of motion in scalar-tensor gravity. Then we used a phenomenological approach to incorporate these modifications in a simple and direct way. Clearly, the rather simple concept of a theory-agnostic parametrisation relies on the fact that it is very difficult, at present, to discern between GR and non-GR spacetimes. Hence, a theory-dependent modelling would prove to be an unnecessary titanic enterprise. This motivates the theory-independent, phenomenological approach that we pursued, being satisfactory as long as one is content to capture some preliminary expectations as to the modifications one may reasonably expect. Indeed, we have identified a class of effects as a promising test-bed for gravitation theories, and the results that we derived are summarised in section 4.4.

These results motivate the scientific interest in the promising field of gravitational self-force technique, not only for its considerable astrophysical impact, but also for the importance of the additional constraints it could provide on the plethora of extended theories that have been developed so far. In conclusion, we note again how in all our discussion the crucial role that future space-borne experiments will play is prominently highlighted, since they will open for the first time a complete window on the whole *gravitational universe*.

A. Useful definitions

A.1. Fermi-Walker transport

Let γ be a timelike worldline described by parametric relations $z^\mu(\tau)$, with τ the proper time; let $u^\mu = \mathrm{d}z^\mu/\mathrm{d}\tau$ be the normalised tangent vector of the curve and $a^\mu = Du^\mu/\mathrm{d}\tau$ is its acceleration. A vector field is said to be Fermi-Walker (FW) transported along γ if it is a solution to the differential equation

$$\frac{Dv^\mu}{\mathrm{d}\tau} = (v_\nu a^\nu)\, u^\mu - (v_\nu u^\nu)\, a^\mu.$$

Note that this reduces to parallel transport when $a^\mu = 0$ and γ is a geodesic. FW transport has two important properties that follow from the definition. First, u^μ is automatically FW transported along γ (recall that u^μ is orthogonal to a^μ); second, if the vectors v^μ and w^μ are both FW transported along γ, then their inner product is constant along γ, *i.e.* $D(v_\mu w^\mu)/\mathrm{d}\tau = 0$.

The natural moving frame associated with an accelerated observer consists of four orthonormal vectors, each of which is Fermi-Walker transported along the worldline and one of which is the four-velocity of the observer, u^μ.

A.2. Parallel propagator

Any vector field $A^\mu(z)$ on a geodesic linking x to x' can be decomposed in the *vierbein* e^μ_a as $A^\mu = A^a e^\mu_a$, and the vector frame components are $A^a = A^\mu e^a_\mu$. If A^μ is parallel transported on the geodesic, then the coefficients A^a are constants. The vector at x can then be expressed as

$$A^\alpha(x) = \left(A^{\alpha'}(x')\, e^a_{\alpha'}(x')\right) e^\alpha_a(x) = g^\alpha{}_{\alpha'}(x,x')\, A^{\alpha'}(x'), \quad g^\alpha{}_{\alpha'}(x,x') \equiv e^\alpha_a(x)\, e^a_{\alpha'}(x').$$

The object $g^\alpha{}_{\alpha'}$ is the *parallel propagator*: it takes a vector at x' and parallel-transports it to x along the unique geodesic that links these points. One can show that the orderings of indices and arguments are arbitrary. The action on tensors of arbitrary rank is easy to figure out: there are as many occurrences of the parallel propagator as indexes.

A.3. Fermi normal coordinates

Constructing the FNC of a point x in the normal convex neighbourhood of a timelike worldline γ requires considering the unique *spacelike* geodesic β passing through x and intersecting γ orthogonally. Denoting the intersection point by $\bar{x} \equiv z(t)$ – with t denoting the value of the proper time at that point – the FNC of x are defined by

$$\hat{x}^0 = t, \quad \hat{x}^a = -e^a_{\bar{\alpha}}(\bar{x})\, \sigma^{\bar{\alpha}}(x,\bar{x}), \quad \sigma_{\bar{\alpha}}(x,\bar{x})\, u^{\bar{\alpha}}(\bar{x}) = 0,$$

where barred indices indicates that the quantity transforms tensorially at the barred point; $e^a_{\bar{\alpha}}$ is a spatial triad (a runs from 1 to 3) that is FW transported along the worldline; the last equation determines $\bar{x}$ from the requirement that the vector$-\sigma^{\bar{\alpha}}$, tangent to β at $\bar{x}$, be orthogonal to $u^{\bar{\alpha}}$, the vector tangent to γ. The distance from $\bar{x}$ to x along the geodesic β is given by

$$s^2 \equiv \delta_{ab}\hat{x}^a\hat{x}^b = 2\sigma(x,\bar{x}),$$

which is perturbatively small relative to the curvature scale. This gives an immediate meaning to $\hat{x}^a$, the spatial FNC, and the time coordinate $\hat{x}^0$ is simply the proper time at the intersection point $\bar{x}$. In the text, hats are removed when no confusion can arise.

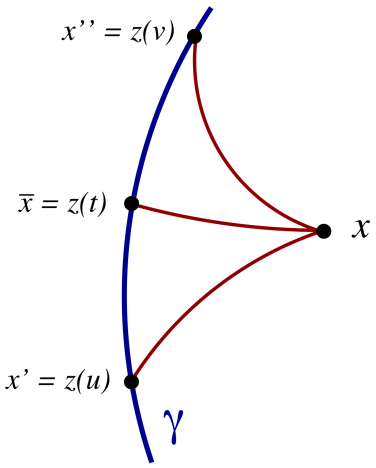

Figure A.1.: The retarded point $x' = z(u)$ is linked to x by a future-directed null geodesic; the simultaneous point $\bar{x} = z(t)$ is linked to x by a spacelike geodesic that intersects γ orthogonally; the advanced point $x'' = z(v)$ is linked to x by a past-directed null geodesic. [20]

A.4. Retarded coordinates, retarded/advanced point

FNC are constructed on the basis of a spacelike geodesic β connecting a field point x to the worldline timelike γ. The retarded coordinates are based, instead, on a *null* geodesic that links the field point to γ. Hence, we let x be within the normal convex neighbourhood of γ, β be the unique future-directed null geodesic going from γ to x, and $x' \equiv z(u)$ be the (retarded) point at which β intersects γ – with u denoting the value of the proper time at this point. The retarded coordinates are defined by

$$\hat{x}^0 = u, \qquad \hat{x}^a = -e^a_{\alpha'}\left(x'\right)\sigma^{\alpha'}\left(x, x'\right), \qquad \sigma\left(x, x'\right) = 0;$$

the second equality gives the frame components of the separation vector and the last one states that x' and x are linked by a null geodesic. From the fact that $\sigma^{\alpha'}$, tangent to β at x', is null, and that $u^{\alpha'}$ is the tangent to γ at x', one obtains

$$r \equiv \sqrt{\delta_{ab}\hat{x}^a\hat{x}^b} = u_{\alpha'}\sigma^{\alpha'},$$

and r is a positive-definite quantity by virtue of the fact that β is a future-directed null geodesic (making $\sigma^{\alpha'}$ past-directed). To fathom its meaning, consider a flat spacetime, where $\sigma^{\alpha'} = -(x - x')^{\alpha}$; in a Lorentz frame momentarily comoving with the worldline, $r = t - t' > 0$ is the spatial distance between x and x' as measured in this frame (recall that $c = 1$ here). In curved spacetime, $r = u_{\alpha'}\sigma^{\alpha'}$ can still be called the retarded distance between the point x and the worldline γ. So, the invariant quantity r is an affine parameter on the null geodesic β that links x to x', which can be loosely interpreted as the time delay between x and x' as measured by an observer comoving with the particle.

It proves convenient to introduce on the world line, along with the *retarded* and the *simultaneous* points, an *advanced* point associated with the field point x. The latter will be denoted $x'' \equiv z(v)$ – with v denoting the value of the proper time at x''. The advanced point is linked to x by a past-directed null geodesic and can be located by solving $\sigma\left(x, x''\right) = 0$ together with the requirement that $\sigma^{\alpha''}\left(x, x''\right)$ be a future-directed null vector. The affine-parameter distance between x and x'' along the null geodesic β is given by the positive-definite quantity

$$r_{\text{adv}} = -u_{\alpha''}\sigma^{\alpha''}$$

and it is called the advanced distance between x and the worldline γ.

B. Failure of point-like approximation in full General Relativity

The failure stems from the non-linearity of the Einstein field equation. From a physical perspective, we know the Einstein equations imply that a sufficiently dense mass distribution will collapse to form a BH, not a point particle. From a mathematical perspective, we know that the Einstein equations with a point-particle source do not have a well-defined solution within any suitable class of functions.

We consider an object of mass m moving in a spacetime with a much larger external length scale $\mathcal{R} \gg m$. In an EMRI system, $\mathcal{R}$ can be the mass M of the large BH (in geometrical units, where mass has dimensions of length). Wishing to take advantage of the separation of scales, we start expanding the exact metric $\mathfrak{g}_{\mu\nu}$ of the system in the limit $m/\mathcal{R} \to 0$, *i.e.*

$$\mathfrak{g}_{\mu\nu} = g_{\mu\nu} + h_{\mu\nu} = g_{\mu\nu} + \sum_{n=1}^{+\infty} \epsilon^n h^n_{\mu\nu} = g_{\mu\nu} + \epsilon h^1_{\mu\nu} + \epsilon^2 h^2_{\mu\nu} + O\left(\epsilon^3\right). \tag{B.1}$$

Here ϵ is introduced as a formal expansion parameter to count powers of $m/\mathcal{R}$. The zeroth-order term is referred to as the background metric – in the case of an EMRI, it is the metric of the large BH. The corrections $h^n_{\mu\nu}$ describe the gravitational perturbations due to the small object. The full metric must clearly satisfy the Einstein equation

$$G_{\mu\nu}\left[\mathfrak{g}\right] = 8\pi T_{\mu\nu}, \tag{B.2}$$

where $G_{\mu\nu}\left[\mathfrak{g}\right]$ is the Einstein tensor of the spacetime and $T_{\mu\nu}$ is the stress-energy tensor of the matter content of the system. For simplicity, suppose that the small object represents the only matter, such that $T_{\mu\nu}$ is the stress-energy tensor of the small object itself. The expansion (B.1) yields

$$G_{\mu\nu}\left[\mathfrak{g}\right] = G_{\mu\nu}\left[g\right] + \epsilon\delta G_{\mu\nu}\left[h^1\right] + \epsilon^2\left(\delta G_{\mu\nu}\left[h^2\right] + \delta^2 G_{\mu\nu}\left[h^1\right]\right) + O\left(\epsilon^3\right),$$

where

$$\delta G_{\mu\nu}\left[h^n\right] \equiv \left.\frac{\mathrm{d}G_{\mu\nu}\left(g + \lambda h^n\right)}{\mathrm{d}\lambda}\right|_{\lambda=0} \tag{B.3}$$

is the linearised Einstein tensor and is therefore linear in $h^n_{\mu\nu}$, while

$$\delta^2 G_{\mu\nu}\left[h^1\right] \equiv \frac{1}{2}\left.\frac{\mathrm{d}^2 G_{\mu\nu}\left(g + \lambda h^1\right)}{\mathrm{d}\lambda^2}\right|_{\lambda=0}$$

has the schematic form $\partial h^1 \partial h^1 + h^1 \partial^2 h^1$. Let us also suppose now that, in this limit, $T_{\mu\nu}$ is approximately that of a point particle, such that

$$T_{\mu\nu} = \epsilon T^1_{\mu\nu} + \epsilon^2 T^2_{\mu\nu} + O\left(\epsilon^3\right),$$

where $T^0_{\mu\nu} = 0$ since there is only the large BH and $T^1_{\mu\nu}$ is the stress-energy of a point mass moving in the background metric $g_{\mu\nu}$.

Through first order in ϵ, no fundamental problem arises: we just get the linearised Einstein equation with a point-particle source,

$$\delta G_{\mu\nu}\left[h^1\right] = 8\pi T^1_{\mu\nu}.$$

This is analogous to $\Box A^\mu - R^\mu{}_\nu A^\nu = -4\pi j^\mu$ for the electromagnetic potential and its solutions can be expressed in terms of Green's functions. Like in the electromagnetic case, the retarded field splits into singular and regular fields. The former behaves as $h^{\mathrm{S1}}_{\mu\nu} \sim m/r$ near the particle, r being a measure of the distance to the particle's worldline; the latter, $h^{\mathrm{R1}}_{\mu\nu}$, is a smooth vacuum solution which contains

the backscattered waves that arise from propagation within – not just on – the light cones of the background spacetime.

We move now to the second-order term in the Einstein equation, which reads

$$\delta G_{\mu\nu}\big[h^2\big] = 8\pi T^2_{\mu\nu} - \delta^2 G_{\mu\nu}\big[h^1\big]. \tag{B.4}$$

The second-order perturbation $h^2_{\mu\nu}$ is sourced by the quadratic combinations of $h^1_{\mu\nu}$ in $\delta^2 G_{\mu\nu}\big[h^1\big]$, which generally behave like $\sim 1/r^4$ near the particle: this singularity is too strong to be integrated. And there is no great hope of curing it. Since this term is constructed from a quadratic operation on an integrable function (as opposed to a linear one), it is not even well defined as a distribution. Furthermore, the source term $8\pi T^2_{\mu\nu}$, if it is well defined at all, must be a distribution solely supported on the particle's worldline – hence, it cannot cure the non-distributional divergence of $\delta^2 G_{\mu\nu}$. Hence, the field equation is ill defined and the point-particle treatment fails spectacularly. This means that in gravity we must face the extended size of the small object head on, as point particles pose increasingly worsening difficulties at non-linear orders in perturbation theory. It is actually well known that, in any well-behaved space of functions there exists *no* solution to the fully non-linear eq. (B.2) with a point-particle source [64]. Despite these obstacles, one might suppose that the point-particle idealisation could still work as the divergences may be appropriately regularised; evidences for the viability of this route have been provided by the use of dimensional regularisation in post-Newtonian theory and by effective field theories. However, at a fundamental level, there should be *no need* for such a regularisation in GR.

A principal goal of the SF theory is to generalise the point-particle approximation: to reduce the object to a few "bulk" properties (such as mass and spin) supported on a worldline, without representing it as a delta-function stress-energy tensor. In order to achieve this result, a change of perspective can be useful. In fact, the key idea in generalising the notion of a point-particle is to focus not on the small object itself, but on the gravitational field in its immediate neighbourhood, so that, rather than thinking of the point-particle approximation as a statement about the stress-energy tensor of the object, we can take it to be a statement about its field.

C. Mode-sum method: an elementary example

Consider a point-like particle of mass m at rest in flat space. The location of the particle is $\boldsymbol{x}_{\rm p}$ in a given Cartesian system. In this simple static configuration, the perturbed Einstein equations (2.20) read

$$\triangle \bar{h}_{tt} = -16\pi m \delta^3 \left(\boldsymbol{x} - \boldsymbol{x}_{\rm p}\right), \tag{C.1}$$

where $\triangle$ is the 3D Laplacian and with all other components of $\bar{h}_{\alpha\beta}$ vanishing. The static perturbation automatically satisfies the Lorenz-gauge condition (2.10). Of course, in this simple case we can immediately write down the exact physical (Coulomb-like) static solution,

$$\bar{h}^{\rm ret}_{tt} = \frac{4m}{|\boldsymbol{x} - \boldsymbol{x}_{\rm p}|},$$

and we also trivially have $F^{\alpha}_{\rm self} = 0$. However, for the sake of the discussion, let us proceed by considering the multipole expansion of the perturbation.

To this end, introduce polar coordinates (r, θ, ϕ), such that the particle is located at $\boldsymbol{x}_{\rm p} = (r_0, \theta_0, \phi_0)$, with $r_0 \neq 0$; then, expand the solution $\bar{h}^{\rm ret}_{tt}$ in spherical harmonics on the spheres with $r = {\rm const}$, in the form

$$\bar{h}^{\rm ret}_{tt} = \sum_{\ell=0}^{+\infty} \bar{h}^{\ell}_{tt}\left(r, \theta\right), \qquad \text{where} \qquad \bar{h}^{\ell}_{tt}\left(r, \theta\right) \equiv \sum_{m=-\ell}^{\ell} \tilde{h}^{\ell m}_{tt}\left(r\right) Y_{\ell m}\left(\theta, \phi\right).$$

This expansion separates the field equation (C.1) into radial and angular parts, the former reading (for each ℓ, m)

$$\partial_r^2 \tilde{h}^{\ell m}_{tt} + \frac{2}{r}\partial_r \tilde{h}^{\ell m}_{tt} - \frac{\ell\left(\ell+1\right)}{r^2}\tilde{h}^{\ell m}_{tt} = -\frac{16\pi m}{r_0^2} Y^*_{\ell m}\left(\theta_0, \phi_0\right) \delta\left(r - r_0\right),$$

where an asterisk denotes complex conjugation. The unique physical ℓ, m-mode solution, continuous everywhere and regular at both $r = 0$ and $r \to +\infty$, reads

$$\tilde{h}^{\ell m}_{tt}\left(r\right) = \frac{16\pi m}{\left(2\ell+1\right) r_0} Y^*_{\ell m}\left(\theta_0, \phi_0\right) \cdot \begin{cases} \left(r/r_0\right)^{-(\ell+1)} & r \geq r_0 \\ \left(r/r_0\right)^{\ell} & r \leq r_0 \end{cases},$$

thus giving

$$\bar{h}^{\ell}_{tt}\left(r, \theta\right) = \frac{4m}{r_0} P_{\ell}\left(\cos\Theta\right) \cdot \begin{cases} \left(r/r_0\right)^{-(\ell+1)} & r \geq r_0 \\ \left(r/r_0\right)^{\ell} & r \leq r_0 \end{cases}, \tag{C.2}$$

where P_ℓ is a Legendre polynomial and Θ is the angle subtended by the two radius vectors to $\boldsymbol{x}$ and $\boldsymbol{x}_{\rm p}$.

Now construct the force field $F^{\alpha}_{\rm ret}$ as it is defined in eq. (2.24). We find $F^{t}_{\rm ret} = 0$, and the spatial components are $F^{i}_{\rm ret} = m\bar{\nabla}^{itt}\bar{h}_{tt} = \left(m/4\right)\partial^i \bar{h}_{tt}$. Focus now on the $i = r$ component. The ℓ-mode is given as $F^{r\ell}_{\rm ret} = \left(m/4\right)\partial_r \bar{h}^{\ell}_{tt}$; using eq. (C.2) and evaluating $F^{r\ell}_{\rm ret}$ at the particle (taking $\Theta \to 0$ followed by $r \to r_0^{\pm}$), we obtain

$$F^{r\ell}_{{\rm ret}\pm}\left(\vec{x}_{\rm p}\right) = \mp L \frac{m^2}{r_0^2} - \frac{m^2}{2r_0^2}, \qquad L \equiv \ell + \frac{1}{2}.$$

Here the subscripts $\pm$ indicate the two (different) values obtained by taking the particle limit from "outside" and "inside".

Let us note the following features manifest in the above simple analysis:

- the individual ℓ-modes of the metric perturbation, $\bar{h}^{\ell}_{\alpha\beta}$, are each continuous at the location of the particle, although their derivatives are discontinuous there;

- the individual ℓ-modes $F^{\alpha\ell}_{\rm ret}$ have finite one-sided values at the particle;
- at large ℓ, each of this one-sided values of $F^{\alpha\ell}_{\rm ret}$ at the particle is dominated by a term $\propto \ell$ (the mode-sum obviously diverges at the particle, reflecting the divergence of the full force $F^{\alpha}_{\rm ret}$ there).

It turns out that all these features are quite generic and they carry over intact to the much more general problem of a particle moving in Kerr geometry [18]. Specifically, one finds that, at any point along the trajectory of the particle, the (one-sided values of) modes $F^{\alpha\ell}_{\rm ret}$ (and $F^{\alpha\ell}_{\rm S}$ as well) always admit the large-ℓ form (2.25). In this elementary problem the power series in $1/L \sim 1/\ell$ stops at the L^0 term, but in general the series can be infinite.

Bibliography

[1] LIGO Scientific Collaboration, Virgo Collaboration, *Phys. Rev. Lett.* **116**, 061102 (2016) [arXiv:1602.03837 [gr-qc]]

[2] LIGO Scientific Collaboration, Virgo Collaboration, *Class. Quantum Grav.* **34**, 104002 (2017) [arXiv:1611.07531 [gr-qc]]

[3] LISA L3 Mission Proposal [arXiv:1702.00786 [astro-ph.IM]]

[4] M. Maggiore, *Gravitational Waves*, vol. 1, *Oxford University Press* (2008)

[5] P. Amaro-Seoane, J.R. Gair, M. Freitag, M.C. Miller, I. Mandel, C.J. Cutler, S. Babak, *Class. Quant. Grav.* **24**, R113-R169 (2007) [arXiv:astro-ph/0703495]

[6] L. Barack, C. Cutler, *Phys. Rev. D* **70**, 122002 (2004)

[7] I.D. Soares, *General Relativity and Gravitation* **49**, 77 (2017) [arXiv:1608.01174 [gr-qc]]

[8] W.K.H. Panofsky, M. Phillips, *Classical Electricity and Magnetism (second edition)*, *Addison-Wesley* (1962)

[9] K. Lechner, *Classical Electrodynamics*, *Springer* (2018)

[10] P. A. M. Dirac, *Proc. R. Soc. London*, Ser. A**167**, 148 (1938)

[11] B. S. DeWitt, R. W. Brehme, *Annals of Phys.* **9**, 220 (1960)

[12] Y. Mino, M. Sasaki, T. Tanaka, *Phys. Rev. D* **55**, 3457 (1997) [arXiv:gr-qc/9606018]

[13] T. C. Quinn, R. M. Wald, *Phys. Rev. D* **56**, 3381 (1997) [arXiv:gr-qc/9610053]

[14] S. E. Gralla, R. M. Wald, *Class. Quant. Grav.* **25**, 205009 (2008) [arXiv:0806.3293 [gr-qc]]

[15] S. Detweiler, B. F. Whiting, *Phys. Rev. D* **67**, 024025 (2003) [arXiv:gr-qc/0202086]

[16] A. Pound, *Equations of Motion in Relativistic Gravity. Fundamental Theories of Physics*, vol. 179, pp 399-486, *Springer* (2015)

[17] L. Barack, A. Ori, *Phys. Rev. D* **64**, 124003 (2001) [arXiv:gr-qc/0107056]

[18] L. Barack, *Class. Quant. Grav.* **26**, 213001 (2009) [arXiv:0908.1664 [gr-qc]]

[19] L. Barack, *Phys. Rev. D* **64**, 084021 (2001) [arXiv:gr-qc/0105040]

[20] E. Poisson, A. Pound, I. Vega [arXiv:1102.0529 [gr-qc]]

[21] L. Barack, A. Pound [arXiv:1805.10385 [gr-qc]]

[22] N. Warburton, L. Barack, N. Sago, *Phys. Rev. D* **87**, 084012 (2013)

[23] B. Carter, *Phys. Rev.* **174**, 1559-1571 (1968)

[24] J. Kevorkian, J.D. Cole, *Multiple scale and singular perturbation methods*, *Springer New York* (1996)

[25] W. Eckhaus, *Asymptotic Analysis of Singular Perturbations*, *Elsevier North-Holland* (1979)

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

[27] U. Ruangsri, S.A. Hughes, *Phys. Rev. D* **89**, 084036 (2014) [arXiv:1307.6483 [gr-qc]]

[28] C.P.L. Berry, R.H. Cole, P. Caizares, J.R. Gair, *Phys. Rev. D* **94**, 124042 (2016) [arXiv:1608.08951 [gr-qc]]

[29] A. Le Tiec, *Int. J. Mod. Phys. D* **23**, 1430022 (2014) [arXiv:1408.5505 [gr-qc]]

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

[31] K. Ganz, W. Hikida, H. Nakano, N. Sago, T. Tanaka, *Prog. Theor. Phys.* **117**, 1041-1066 (2007) [arXiv:gr-qc/0702054]

[32] S.A. Teukolsky, W.H. Press, *Astrophys. J.* **193**, 443-461 (1974)

[33] S.A. Hughes, S. Drasco, E.E. Flanagan, J. Franklin, *Phys. Rev. Lett.* **94**, 221101 (2005) [arXiv:gr-qc/0504015]

[34] L. Barack, N. Sago, *Phys. Rev. D* **81**, 084021 (2010) [arXiv:1002.2386 [gr-qc]]

[35] M. van de Meent, *Phys. Rev. D* **97**, 104033 (2018) [arXiv:1711.09607 [gr-qc]]

[36] M. Van De Meent, N. Warburton, *Class. Quant. Grav.* **35**, 144003 (2018) [arXiv:1802.05281 [gr-qc]]

[37] N. Birrell, P. Davies, *Quantum Fields in Curved Space*, *Cambridge University Press* (1982)

[38] S.M. Carroll, J. Geddes, M.B. Hoffman, R.M. Wald, *Phys. Rev. D* **66**, 024036 (2002) [arXiv:hep-th/0110149]

[39] D.C. Robinson, *Phys. Rev. Lett.* **34**, 905 (1975)

[40] E. Berti *et al.*, *Class. Quantum Grav.* **32**, 243001 (2015)

[41] N. Yunes, P. Pani, V. Cardoso, *Phys. Rev. D* **85**, 102003 (2012) [arXiv:1112.3351 [gr-qc]]

[42] S. Mirshekari, C.M. Will, *Phys. Rev. D* **87**, 084070 (2013) [arXiv:1301.4680 [gr-qc]]

[43] T.P. Sotiriou, V. Faraoni, *Phys. Rev. Lett.* **108**, 081103 (2012) [arXiv:1109.6324 [gr-qc]]

[44] J. Healy, T. Bode, R. Haas, E. Pazos, P. Laguna, D.M. Shoemaker, N. Yunes, *Class. Quant. Grav.* **29**, 232002 (2012) [arXiv:1112.3928 [gr-qc]]

[45] T. Damour, G. Esposito-Farèse, *Phys. Rev. Lett.* **70**, 2220 (1993)

[46] P. Zimmerman, E. Poisson [arXiv:1406.5111 [gr-qc]]

[47] P. Zimmerman, *Phys. Rev. D* **92**, 064051 (2015) [arXiv:1507.04076 [gr-qc]]

[48] S.E. Gralla, *Phys. Rev. D* **81**, 084060 (2010) [arXiv:1002.5045 [gr-qc]]

[49] K.S. Stelle (1978) *General Relativity and Gravitation* **9**, 4 (1978)

[50] S. Capozziello, M. De Laurentis, *Ann. Phys.* **524**, 9-10 (2012) 545-578

[51] C. Cutler, D. Kennefick, E. Poisson, *Phys. Rev. D* **50**, 3816 (1994)

[52] Y. Mino, *Phys. Rev. D* **67**, 084027 (2003)

[53] L. Barack, N. Sago, *Phys. Rev. D* **75**, 064021 (2007)

[54] N. Sago, L. Barack, S. Detweiler, *Phys. Rev. D* **78**, 124024 (2008)

[55] L. Barack, C. O. Lousto, *Phys. Rev. D* **72**, 104026 (2005)

[56] Ya. B. Zel'dovich, I. D. Novikov, *Relativistic Astrophysics (vol. I), Stars and Relativity, U. Chicago Press* (1971)

[57] A. Ori, K. S. Thorne, *Phys. Rev. D* **62**, 124022 (2000) [arXiv:gr-qc/0003032]

[58] W. de Sitter, *Mon. Not. R. Astron. Soc.* **77**, 155 (1916)

[59] I. I. Shapiro, R. D. Reasenberg, J. F. Chandler, R.W. Babcock, *Phys. Rev. Lett.* **61**, 2643 (1988)

[60] R. P. Breton *et al.*, *Science* **321**, 104 (2008)

[61] S.R. Dolan, N. Warburton, A.I. Harte, A. Le Tiec, B. Wardell, L. Barack, *Phys. Rev. D* **89**, 064011 (2014) [arXiv:1312.0775 [gr-qc]]

[62] L. Barack, T. Damour, N. Sago, *Phys. Rev. D* **82**, 084036 (2010) [arXiv:1008.0935 [gr-qc]]

[63] T.P. Sotiriou, V. Faraoni, S. Liberati, *Int. J. Mod. Phys. D* **17**, 399 (2008) [arXiv:0707.2748 [gr-qc]]

[64] R. Geroch, J. Traschen, *Phys. Rev. D* **36**, 1017 (1987)