A rigorous derivation of gravitational self-force

2008 Class. Quantum Grav. 25 205009

(http://iopscience.iop.org/0264-9381/25/20/205009)

IOP PUBLISHING CLASSICAL AND QUANTUM GRAVITY

Class. Quantum Grav. **25** (2008) 205009 (33pp) doi:10.1088/0264-9381/25/20/205009

A rigorous derivation of gravitational self-force

Samuel E Gralla and Robert M Wald

Enrico Fermi Institute and Department of Physics, University of Chicago, 5640 S Ellis Avenue, Chicago, IL 60637, USA

Received 21 July 2008, in final form 18 August 2008
Published 30 September 2008
Online at stacks.iop.org/CQG/25/205009

Abstract

There is general agreement that the MiSaTaQuWa equations should describe the motion of a 'small body' in general relativity, taking into account the leading order self-force effects. However, previous derivations of these equations have made a number of ad hoc assumptions and/or contain a number of unsatisfactory features. For example, all previous derivations have invoked, without proper justification, the step of 'Lorenz gauge relaxation', wherein the linearized Einstein equation is written in the form appropriate to the Lorenz gauge, but the Lorenz gauge condition is then not imposed—thereby making the resulting equations for the metric perturbation inequivalent to the linearized Einstein equations. (Such a 'relaxation' of the linearized Einstein equations is essential in order to avoid the conclusion that 'point particles' move on geodesics.) In this paper, we analyze the issue of 'particle motion' in general relativity in a systematic and rigorous way by considering a one-parameter family of metrics, $g_{ab}(\lambda)$, corresponding to having a body (or black hole) that is 'scaled down' to zero size and mass in an appropriate manner. We prove that the limiting worldline of such a one-parameter family must be a geodesic of the background metric, $g_{ab}(\lambda = 0)$. Gravitational self-force—as well as the force due to coupling of the spin of the body to curvature—then arises as a first-order perturbative correction in λ to this worldline. No assumptions are made in our analysis apart from the smoothness and limit properties of the one-parameter family of metrics, $g_{ab}(\lambda)$. Our approach should provide a framework for systematically calculating higher order corrections to gravitational self-force, including higher multipole effects, although we do not attempt to go beyond first-order calculations here. The status of the MiSaTaQuWa equations is explained.

PACS number: 04.25.−g

1. Introduction

The physical content of general relativity is contained in Einstein's equation, which has a well-posed initial-value formulation (see, e.g., [1]). In principle, therefore, to determine the motion of bodies in general relativity—such as binary neutron stars or black holes—one simply needs to provide appropriate initial data (satisfying the constraint equations) on a spacelike slice and then evolve these data via Einstein's equation. However, in practice, it is generally impossible to find exact solutions of physical interest describing the motion of bodies by analytic methods. Although it is now possible to find solutions numerically in many cases of interest, it is difficult and cumbersome to do so, and one may overlook subtle effects and/or remain unenlightened about some basic general features of the solutions. Therefore, it is of considerable interest to develop methods that yield approximate descriptions of motion in some cases of interest.

In general, the motion of a body of finite size will depend on the details of its composition as well as the details of its internal states of motion. Therefore, one can expect to get a simple description of motion only in some kind of 'point particle limit'. However, Einstein's equation is nonlinear, and a straightforward analysis [2] shows that it does not make any mathematical sense to consider solutions of Einstein's equation with a distributional stress–energy tensor supported on a worldline[1]. Physically, if one tried to shrink a body down to zero radius at fixed mass, collapse to a black hole would occur before the point particle limit could be reached.

Distributional stress–energy tensors supported on a worldline *do* make mathematical sense in the context of the linearized Einstein equation. Therefore, one might begin a treatment of gravitational self-force by considering a metric perturbation, h_{ab}, in a background metric, g_{ab}, sourced by the stress–energy tensor of a 'point particle' of mass M, given in coordinates (t, x^i) by

$$G^{(1)}_{ab}[h](t, x^i) = 8\pi M u_a(t) u_b(t) \frac{\delta^{(3)}(x^i - z^i(t))}{\sqrt{-g}} \frac{\mathrm{d}\tau}{\mathrm{d}t}, \tag{1}$$

where u^a is the unit tangent (i.e., 4-velocity) of the worldline γ defined by $x^i(t) = z^i(t)$ and τ is the proper time along γ. (Here $\delta^{(3)}(x^i - z^i(t))$ is the 'coordinate delta function', i.e., $\int \delta^{(3)}(x^i - z^i(t))\, \mathrm{d}^3 x^i = 1$. The right-hand side could also be written covariantly as $8\pi M \int_\gamma \delta_4(x, z(\tau)) u_a(\tau) u_b(\tau)\, \mathrm{d}\tau$, where δ_4 is the covariant four-dimensional delta function and τ denotes the proper time along γ.) However, this approach presents two major difficulties.

First, the linearized Bianchi identity implies that the point particle stress energy must be conserved. However, as we shall see explicitly in section 4, this requires that the worldline γ of the particle be a geodesic of the background spacetime. Therefore, there are no solutions to equation (1) for non-geodesic source curves, making it hopeless to use this equation to derive corrections to geodesic motion. This difficulty has been circumvented in [5–8] and other references by modifying (1) as follows. Choose the Lorenz gauge condition, so that equation (1) takes the form

$$\nabla^c \nabla_c \tilde{h}_{ab} - 2R^c{}_{ab}{}^d \tilde{h}_{cd} = -16\pi M u_a(t) u_b(t) \frac{\delta^{(3)}(x^i - z^i(t))}{\sqrt{-g}} \frac{\mathrm{d}\tau}{\mathrm{d}t}, \tag{2}$$

$$\nabla^b \tilde{h}_{ab} = 0, \tag{3}$$

where $\tilde{h}_{ab} \equiv h_{ab} - \frac{1}{2} h g_{ab}$ with $h = h_{ab} g^{ab}$. Equation (2) by itself has solutions for any source curve γ; it is only when the Lorenz gauge condition (3) is adjoined that the equations

[1] Nevertheless, action principles corresponding to general relativity with point particle sources are commonly written (see, e.g., equations (12.1.6) and (12.4.1)–(12.4.2) of [3]). There are no solutions to the equations of motion resulting from such action principles. By contrast, distributional solutions of Einstein's equation with support on a timelike hypersurface (i.e., 'shells') do make mathematical sense [2, 4].

are equivalent to the linearized Einstein equation and geodesic motion is enforced. Therefore, if one solves the Lorenz gauge form (2) of the linearized Einstein equation while simply *ignoring* the Lorenz gauge condition[2] that was used to derive (2), one allows for the possibility non-geodesic motion. Of course, this 'gauge relaxation' of the linearized Einstein equation produces an equation inequivalent to the original. However, because deviations from geodesic motion are expected to be small, the Lorenz gauge violation should likewise be small, and it thus has been argued [6] that solutions to the two systems should agree to sufficient accuracy.

The second difficulty is that the solutions to equation (2) are singular on the worldine of the particle. Therefore, naive attempts to compute corrections to the motion due to h_{ab}—such as demanding that the particle move on a geodesic of $g_{ab}+h_{ab}$—are virtually certain to encounter severe mathematical difficulties, analogous to the difficulties encountered in treatments of the electromagnetic self-force problem.

Despite these difficulties, there is a general consensus that in the approximation that spin and higher multipole moments may be neglected, the motion of a sufficiently small body (with no 'incoming radiation') should be described by self-consistently solving equation (2) via the retarded solution together with

$$u^b\nabla_b u^a = -\tfrac{1}{2}(g^{ab}+u^a u^b)\big(2\nabla_d h^{\mathrm{tail}}_{bc} - \nabla_b h^{\mathrm{tail}}_{cd}\big)\big|_{z(\tau)} u^c u^d, \qquad (4)$$

where

$$h^{\mathrm{tail}}_{ab}(x) = M\int_{-\infty}^{\tau^-_{\mathrm{ret}}}\left(G^{+}_{aba'b'} - \frac{1}{2}g_{ab}G^{+c}{}_{ca'b'}\right)\big(x, z(\tau')\big)\, u^{a'}u^{b'}\,\mathrm{d}\tau', \qquad (5)$$

with $G^{+}_{aba'b'}$ the retarded Green's function for equation (2), normalized with a factor of -16π, following [6]. The symbol τ^-_{ret} indicates that the range of the integral extends just short of the retarded time τ_{ret}, so that only the 'tail' (i.e., interior of the light cone) portion of Green's function is used (see, e.g., [8] for details). Equations (2) and (4) are known as the MiSaTaQuWa equations, and have been derived by a variety of approaches. However, there are difficulties with all of these approaches.

One approach [5] that has been taken is to parallel the analysis of [9, 10] in the electromagnetic case and use conservation of effective gravitational stress energy to determine the motion. However, this use of distributional sources at second order in perturbation theory results in infinities that must be 'regularized'. Although these regularization procedures are relatively natural looking, the mathematical status of such a derivation is unclear.

Another approach [6] is to postulate certain properties that equations of gravitational self-force should satisfy. This yields a mathematically clean derivation of the self-force corrected equations of motion. However, as the authors of [6] emphasized, the motion of bodies in general relativity is fully described by Einstein's equation together with the field equations of the matter sources, so no additional postulates should be needed to obtain an equation of motion, beyond the 'small-body' assumption and other such approximations. The analysis given by [6] shows that equation (4) follows from certain plausible assumptions. However, their derivation is thus only a plausibility argument for equation (4). Similar remarks apply to a later derivation by Poisson [8] that uses a Green's function decomposition developed by Detweiler and Whiting [7].

A third approach, taken by Mino, Sasaki and Tanaka [5] and later Poisson [8]—building on previous work of Burke [11], d'Eath [12], Kates [13], Thorne and Hartle [14], and others—involves the use of matched astymptotic expansions. Here one assumes a metric form in the 'near zone'—where the metric is assumed to be that of the body, with a small correction due

[2] In some references, the failure to satisfy equation (3) truly is ignored in the sense that it is not even pointed out that one has modified the linearized Einstein equation, and no attempt is made to justify this modification.

to the background spacetime—and in the 'far zone'—where the metric is assumed to be that of the background spacetime, with a small correction due to the body. One then assumes that there is an overlap region of the body where both metric forms apply, and matches the expressions. The equations of motion of the body then arise from the matching conditions. However, as we shall indicate below, in addition to the 'Lorenz gauge relaxation', there are a number of assumptions and steps in these derivations that have not been adequately justified.

A more rigorous approach to deriving gravitational self-force is suggested by the work of Geroch and Jang [15] and later Geroch and Ehlers [16] on geodesic motion of small bodies (see also [17]). In [15], one considers a fixed spacetime background metric g_{ab} and considers a smooth one-parameter family of stress–energy smooth tensors $T_{ab}(\lambda)$ that satisfy the dominant energy condition and have support on a world tube. As the parameter goes to zero, the world tube shrinks to a timelike curve. It is then proven that this timelike curve must be a geodesic. This result was generalized in [16] to allow g_{ab} to also vary with λ so that Einstein's equation is satisfied. Within the framework of [16], it therefore should be possible to derive perturbative corrections to geodesic motion, including gravitational self-force. However, the conditions imposed in [16] in effect require the mass of the body to go to zero faster than λ^2. Consequently, in this approach, a self-force correction like (4) to the motion of the body would arise at higher order than finite size effects and possibly other effects that would depend on the composition of the body. Thus, while the work of [16] provides a rigorous derivation of geodesic motion of a 'small body' to lowest order, it is not a promising approach to derive gravitational self-force corrections to geodesic motion.

In this paper, we shall take an approach similar in spirit to that of [16], but we will consider a different smooth, one-parameter family of metrics $g_{ab}(\lambda)$, wherein, in effect, we have a body (or black hole) present that scales to zero size in a self-similar manner, with both the size and the mass of the body being proportional to λ. In the limit as $\lambda \to 0$, the body (or black hole) shrinks down to a worldline, γ. As in [15, 16], we prove that γ must be a geodesic of the 'background spacetime' $g_{ab}(\lambda = 0)$—although our method of proving this differs significantly from [15, 16]. To first order in λ, the correction to the motion is described by a vector field, Z^i, on γ, which gives the 'infinitesimal displacement' to the new worldline. We will show that, for any such one-parameter family $g_{ab}(\lambda)$, in the Lorenz gauge $Z^i(\tau)$ satisfies

$$\frac{\mathrm{d}^2 Z^i}{\mathrm{d}t^2} = \frac{1}{2M} S^{kl} R_{kl0}{}^{i} - R_{0j0}{}^{i} Z^j - \left(h^{\mathrm{tail}\,i}{}_{0,0} - \frac{1}{2} h^{\mathrm{tail}}{}_{00}{}^{,i} \right). \tag{6}$$

Here M and S_{ij} are, respectively, the mass and spin of the body. The terms in parentheses on the right-hand side of this equation correspond exactly to the gravitational self-force term in equation (4); the first term is the Papapetrou spin force [18]; the second term is simply the usual right-hand side of the geodesic deviation equation. Equation (6) is 'universal' in the sense that it holds for any one-parameter family satisfying our assumptions, so it holds equally well for a (sufficiently small) blob of ordinary matter or a (sufficiently small) black hole.

Our derivation of (6) is closely related to the matched asymptotic expansions analyses. However, our derivation is a rigorous, perturbative result. In addition, we eliminate a number of ad hoc, unjustified and/or unnecessary assumptions made in previous approaches, including assumptions about the form of the 'body metric' and its perturbations, assumptions about rate of change of these quantities with time, the imposition of certain gauge conditions, the imposition of boundary conditions at the body and, most importantly, the step of Lorenz gauge relaxation. Furthermore, in our approach, the notion of a 'point particle' is a concept that is *derived* rather than assumed. It will also be manifest in our approach that the results hold for all bodies (or black holes) and that the physical results do not depend on a choice of gauge (although $Z^i(\tau)$ itself is a gauge-dependent quantity, i.e., the description of the corrections to particle motion depend on how the spacetimes at different λ are identified—see section 5 and the appendix). In particular, because the Lorenz gauge plays no preferred role in our derivation (aside from being a computationally convenient choice), our gauge transformation law is not, as in previous work [19], restricted to gauges continuously related to the Lorenz gauge. Our approach also holds out the possibility of being extended so as to systematically take higher order corrections into account. However, we shall not attempt to undertake such an extension in this paper.

Although (6) holds rigorously as a first-order perturbative correction to geodesic motion, this equation would not be expected to provide a good global-in-time approximation to motion, since the small local corrections to the motion given by (6) may accumulate with time, and equation (6) would not be expected to provide a good description of the perturbed motion when Z^i becomes large. We will argue in this paper that the MiSaTaQuWa equations, equations (2) and (4), should provide a much better global-in-time approximation to motion than equation (6), and they therefore should be used for self-consistent calculations of the motion of a small body, such as for calculations of extreme-mass-ratio inspiral[3].

We note in passing that, in contrast to Einstein's equation, Maxwell's equations are linear, and it makes perfectly good mathematical sense to consider distributional solutions to Maxwell's equations with point particle sources. However, if the charge–current sources are not specified in advance but rather are determined by solving the matter equations of motion—which are assumed to be such that the total stress energy of the matter and electromagnetic field is conserved—then the full, coupled system of Maxwell's equations together with the equations of motion of the sources becomes nonlinear in the electromagnetic field. Point particle sources do not make mathematical sense in this context either. It is for this reason that—despite more than a century of work on this problem—there is no mathematically legitimate derivation of the electromagnetic self-force on a charged particle. The methods of this paper can be used to rigorously derive a perturbative expression for electromagnetic self-force by considering suitable one-parameter families of coupled electromagnetic-matter systems, and we shall carry out this analysis elsewhere [20]. However, we shall restrict consideration in this paper to the gravitational case.

This paper is organized as follows. In section 2, we motivate the kind of one-parameter family of metrics, $g_{ab}(\lambda)$, that we will consider by examining the one-parameter family of Schwarzschild–de Sitter spacetimes with black hole mass equal to λ. One way of taking the limit as $\lambda \to 0$ yields de Sitter spacetime. We refer to this way of taking the limit as the 'ordinary limit'. But we show that if we take the limit as $\lambda \to 0$ in another way and also rescale the metric by λ^{-2}, we obtain Schwarzschild spacetime. We refer to this second way of taking the limit as $\lambda \to 0$ as the 'scaled limit'. The scaled limit can be taken at any time t_0 on the worldline γ. The simultaneous existence of both types of limits defines the kind of one-parameter family of metrics we seek, wherein a body (or black hole) is shrinking down to a worldline γ in an asymptotically self-similar manner. The precise, general conditions we impose on $g_{ab}(\lambda)$ are formulated in section 2. Some basic properties of $g_{ab}(\lambda)$ that follow directly from our assumptions are given in section 3. In particular, we obtain general 'far-zone' and 'near-zone' expansions and we show that, at any t_0, the scaled limit always yields a stationary, asymptotically flat spacetime at $\lambda = 0$. In section 4, we prove that γ must be a geodesic of the 'background spacetime' (i.e., the spacetime at $\lambda = 0$ resulting from taking the ordinary limit). In other words, to zeroth order in λ a small body or black hole always moves on a geodesic. We also show that, to first order in λ, the metric perturbation associated

[3] This viewpoint is in contrast to that of [8], where it is argued that second-order perturbations are needed for self-consistent evolution.

with such a body or black hole corresponds to that of a structureless 'point particle'. In section 5, we define the motion of the body (or black hole) to first order in λ by finding a coordinate displacement that makes the mass dipole moment of the stationary, asymptotically flat spacetime appearing in the scaled limit vanish. (This can be interpreted as a displacement to the 'center of mass' of the body.) In section 6, we then derive equation (6) as the first order in λ correction to γ in the Lorenz gauge. (An appendix provides the transformation to other gauges.) Finally, in section 7 we explain the status of the MiSaTaQuWa equation (4). Our spacetime conventions are those of Wald [1], and we work in units where $G = c = 1$. Lower case Latin indices early in the alphabet $(a, b, c, \ldots)$ will denote abstract spacetime indices; Greek indices will denote coordinate components of these tensors; Latin indices from mid-alphabet $(i, j, k, \ldots)$ will denote spatial coordinate components.

2. Motivating example and assumptions

As discussed in the introduction, we seek a one-parameter family of metrics $g_{ab}(\lambda)$ that describes a material body or black hole that 'shrinks down to zero size' in an asymptotically self-similar manner. In order to motivate the general conditions on $g_{ab}(\lambda)$ that we shall impose, we consider here an extremely simple example of the type of one-parameter family that we seek. This example will provide an illustration of the two types of limits that we shall use to characterize $g_{ab}(\lambda)$.

Our example is built from Schwarzschild–de Sitter spacetime,

$$\mathrm{d}s^2 = -\left(1 - \frac{2M_0}{r} - C_0 r^2\right)\mathrm{d}t^2 + \left(1 - \frac{2M_0}{r} - C_0 r^2\right)^{-1}\mathrm{d}r^2 + r^2\,\mathrm{d}\Omega^2. \tag{7}$$

(This metric, of course, is a solution to the vacuum Einstein's equation with a cosmological constant $\Lambda = \frac{2}{3}C_0$ rather than a solution with $\Lambda = 0$, but the field equations will not play any role in the considerations of this section; we prefer to use this example because of its simplicity and familiarity.) The desired one-parameter family is

$$\mathrm{d}s^2(\lambda) = -\left(1 - \frac{2M_0\lambda}{r} - C_0 r^2\right)\mathrm{d}t^2 + \left(1 - \frac{2M_0\lambda}{r} - C_0 r^2\right)^{-1}\mathrm{d}r^2 + r^2\,\mathrm{d}\Omega^2, \tag{8}$$

where we consider only the portion of the spacetime with $r > \lambda R_0$ for some $R_0 > 2M_0$. For each λ, this spacetime describes the exterior gravitational field of a spherical body or black hole of mass λM_0 in an asymptotically de Sitter spacetime. As $\lambda \to 0$, the body/black hole shrinks to zero size and mass. For $\lambda = 0$, the spacetime is de Sitter spacetime, which extends smoothly to the worldline $r = 0$, corresponding to where the body/black hole was just before it 'disappeared'.

As explained clearly in [21], the limit of a one-parameter family of metrics $g_{ab}(\lambda)$ depends on how the spacetime manifolds at different values of λ are identified. This identification of spacetime manifolds at different λ can be specified by choosing coordinates x^μ for each λ and identifying points labeled by the same value of the coordinates x^μ. If we use the coordinates (t, r, θ, ϕ) in which the one-parameter family of metrics (8) was presented to do the identification, then it is obvious that the limit as $\lambda \to 0$ of $g_{ab}(\lambda)$ is the de Sitter metric,

$$\mathrm{d}s^2(\lambda = 0) = -(1 - C_0 r^2)\,\mathrm{d}t^2 + (1 - C_0 r^2)^{-1}\,\mathrm{d}r^2 + r^2\,\mathrm{d}\Omega^2. \tag{9}$$

This corresponds to the view that the body/black hole shrinks to zero size and mass and 'disappears' as $\lambda \to 0$.

However, there is another way of taking the limit of $g_{ab}(\lambda)$ as $\lambda \to 0$; the existence of this second limit is one of the key ideas in this paper. Choose an arbitrary time t_0 and, for

$\lambda > 0$, introduce new time and radial coordinates by $\bar{t} \equiv (t - t_0)/\lambda$ and $\bar{r} \equiv r/\lambda$. In the new coordinates, the one-parameter family of metrics becomes

$$\mathrm{d}s^2(\lambda) = -\lambda^2 \left(1 - \frac{2M_0}{\bar{r}} - C_0\lambda^2\bar{r}^2\right) \mathrm{d}\bar{t}^2 + \lambda^2 \left(1 - \frac{2M_0}{\bar{r}} - C_0\lambda^2\bar{r}^2\right)^{-1} \mathrm{d}\bar{r}^2 + \lambda^2\bar{r}^2\,\mathrm{d}\Omega^2, \qquad \bar{r} > R_0. \tag{10}$$

We now consider the limit as $\lambda \to 0$ by identifying the spacetimes with different λ at the same values of the barred coordinates. It is clear by inspection of equation (10) that the limit of $g_{ab}(\lambda)$ as $\lambda \to 0$ at fixed $\bar{x}^\mu$ exists, but is zero. In essence, the spacetime interval between any two events labeled by $\bar{x}_1^\mu$ and $\bar{x}_2^\mu$ goes to zero as $\lambda \to 0$ because these events are converging to the same point on γ. Thus, this limit of $g_{ab}(\lambda)$ exists but is not very interesting. However, an interesting limit can be taken by considering a new one-parameter family of metrics $\bar{g}_{ab}(\lambda)$ by[4]

$$\bar{g}_{\bar{\mu}\bar{\nu}} \equiv \lambda^{-2} g_{\bar{\mu}\bar{\nu}}, \tag{11}$$

so that

$$\mathrm{d}\bar{s}^2(\lambda) = \left(1 - \frac{2M_0}{\bar{r}} - C_0\lambda^2\bar{r}^2\right) \mathrm{d}\bar{t}^2 + \left(1 - \frac{2M_0}{\bar{r}} - C_0\lambda^2\bar{r}^2\right)^{-1} \mathrm{d}\bar{r}^2 + \bar{r}^2\,\mathrm{d}\Omega^2, \qquad \bar{r} > R_0. \tag{12}$$

By inspection, the limit of this family of metrics is simply

$$\mathrm{d}\bar{s}^2|_{\lambda=0} = -\left(1 - \frac{2M_0}{\bar{r}}\right) \mathrm{d}\bar{t}^2 + \left(1 - \frac{2M_0}{\bar{r}}\right)^{-1} \mathrm{d}\bar{r}^2 + \bar{r}^2\,\mathrm{d}\Omega^2, \qquad \bar{r} > R_0, \tag{13}$$

which is just the Schwarzschild metric with mass M_0.

The meaning of this second limit can be understood as follows. As already discussed above, the one-parameter family of metrics (8) describes the exterior gravitational field of a spherical body or black hole that shrinks to zero size and mass as $\lambda \to 0$. The second limit corresponds to how this family of spacetimes looks to the family of observers placed at the events labeled by fixed values of $\bar{x}^\mu$. In going from, say, the $\lambda = 1$ to the $\lambda = 1/100$ spacetime, each observer will see that the body/black hole has shrunk in size and mass by a factor of 100 and each observer also will find himself 'closer to the origin' by this same factor of 100. Suppose now that this family of observers also 'changes units' by this same factor of 100, i.e., they use centimeters rather than meters to measure distances. Then, except for small effects due to the de Sitter background, the family of observers for the $\lambda = 1/100$ spacetime will report the same results (expressed in centimeters) as the observers for the $\lambda = 1$ spacetime had reported (in meters). In the limit as $\lambda \to 0$, these observers simply see a Schwarzschild black hole of mass M_0, since the effects of the de Sitter background on what these observers will report disappear entirely in this limit.

We will refer to the first type of limit (i.e., the limit as $\lambda \to 0$ of $g_{ab}(\lambda)$ taken at fixed x^μ) as the *ordinary limit* of $g_{ab}(\lambda)$, and we will refer to the second limit (i.e., the limit as $\lambda \to 0$ of $\lambda^{-2}g_{ab}(\lambda)$ taken at fixed $\bar{x}^\mu$) as the *scaled limit* of $g_{ab}(\lambda)$. The simultaneous existence of both types of the above limits contains a great deal of relevant information about the one-parameter family of spacetimes (8). In essence, the existence of the first type of limit is telling us that the

[4] A scaling of space (but not time) has previously been considered by Futamase [22] in the post-Newtonian context. Scaled coordinates also appear in the work of D'Eath [12] and Kates [13].

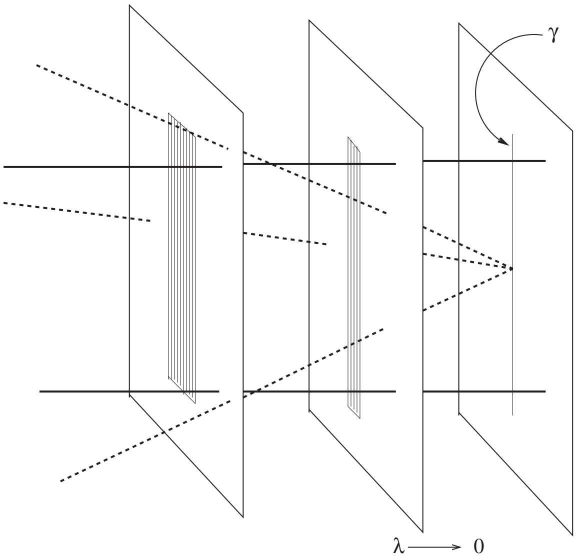

Figure 1. A spacetime diagram illustrating the type of one-parameter family we wish to consider, as well as the two limits we define. As $\lambda \to 0$, the body shrinks and finally disappears, leaving behind a smooth background spacetime with a preferred worldline, γ, picked out. The solid lines illustrate this 'ordinary limit' of $\lambda \to 0$ at fixed r, which is taken along paths that terminate away from γ (i.e., $r > 0$). By contrast, the 'scaled limit' as $\lambda \to 0$, shown in dashed lines, is taken along paths at fixed $\bar{r}$ that converge to a point on γ.

body/black hole is shrinking down to a worldline γ, with its mass going to zero (at least) as rapidly as its radius. The existence of the second type of limit is telling us that the body/black hole is shrinking to zero size in an asymptotically self-similar manner: in particular, its mass is proportional to its size, its shape is not changing and it is not undergoing any (high-frequency) oscillations in time. Figure 1 illustrates the two limits we consider.

We wish now to abstract from the above example the general conditions to impose upon one-parameter families that express in a simple and precise way the condition that we have an arbitrary body (or black hole) that is shrinking to zero size—in an asymptotically self-similar way—in an arbitrary background spacetime. Most of the remainder of this paper will be devoted to obtaining 'equations of motion' for these bodies that are accurate enough to include gravitational self-force effects. A first attempt at specifying the type of one-parameter families $g_{ab}(\lambda)$ that we seek is as follows:

- (i) *Existence of the 'ordinary limit'*: $g_{ab}(\lambda)$ is such that there exists coordinates x^α such that $g_{\mu\nu}(\lambda, x^\alpha)$ is jointly smooth in (λ, x^α), at least for $r > \bar{R}\lambda$ for some constant $\bar{R}$, where $r \equiv \sqrt{\sum (x^i)^2}$ $(i = 1, 2, 3)$. For all λ and for $r > \bar{R}\lambda$, $g_{ab}(\lambda)$ is a vacuum solution of Einstein's equation. Furthermore, $g_{\mu\nu}(\lambda = 0, x^\alpha)$ is smooth in x^α, including at $r = 0$, and, for $\lambda = 0$, the curve γ defined by $r = 0$ is timelike.
- (ii) *Existence of the 'scaled limit'*: for each t_0, we define $\bar{t} \equiv (t - t_0)/\lambda, \bar{x}^i \equiv x^i/\lambda$. Then the metric $\bar{g}_{\bar{\mu}\bar{\nu}}(\lambda; t_0; \bar{x}^\alpha) \equiv \lambda^{-2} g_{\bar{\mu}\bar{\nu}}(\lambda; t_0; \bar{x}^\alpha)$ is jointly smooth in $(\lambda, t_0; \bar{x}^\alpha)$ for $\bar{r} \equiv r/\lambda > \bar{R}$.

Here we have used the notation $g_{\bar{\mu}\bar{\nu}}$ to denote the components of g_{ab} in the $\bar{x}^\alpha$ coordinates, whereas $\bar{g}_{\bar{\mu}\bar{\nu}}$ denotes the components of $\bar{g}_{ab}$ in the $\bar{x}^\alpha$ coordinates. It should be noted that, since the barred coordinates differ only by scale (and shift of time origin) from the corresponding unbarred coordinates, we have[5]

$$g_{\bar{\mu}\bar{\nu}} = \lambda^2 g_{\mu\nu}. \tag{14}$$

Consequently, we have

$$\bar{g}_{\bar{\mu}\bar{\nu}}(\lambda; t_0; \bar{t}, \bar{x}^i) = g_{\mu\nu}(\lambda; t_0 + \lambda\bar{t}, \lambda\bar{x}^i), \tag{15}$$

since there is a cancelation of the factors of λ resulting from the definition of $\bar{g}_{ab}$ and the coordinate rescalings. It should also be noted that there is a redundancy in our description of the one-parameter family of metrics when taking the scaled limit: we define a scaled limit for all values of t_0, but these arise from a single one-parameter family of metrics $g_{ab}(\lambda)$. Indeed, it is not difficult to see that we have

$$\bar{g}_{\bar{\mu}\bar{\nu}}(\lambda; t_0; \bar{t} + \bar{s}, \bar{x}^i) = \bar{g}_{\bar{\mu}\bar{\nu}}(\lambda; t_0 + \lambda\bar{s}; \bar{t}, \bar{x}^i). \tag{16}$$

In fact, our requirements on g_{ab} of the existence of both an 'ordinary limit' and a 'scaled limit' are not quite strong enough to properly specify the one-parameter families we seek. To explain this and obtain the desired strengthened condition, it is convenient to define the new variables

$$\alpha \equiv r, \qquad \beta \equiv \lambda/r, \tag{17}$$

where the range of β is $0 \leqslant \beta < 1/\bar{R}$. Let f denote a component of $g_{ab}(\lambda)$ in the coordinates x^α. However, instead of viewing f as a function of (λ, x^α), we view f as a function of $(\alpha, \beta, t, \theta, \phi)$, where θ and ϕ are defined in terms of x^i by the usual formula for spherical polar angles. In terms of these new variables, taking the 'ordinary limit' corresponds to letting $\beta \to 0$ at any fixed $\alpha > 0$, whereas taking the 'scaled limit' corresponds to letting $\alpha \to 0$ at any fixed $\beta > 0$ (see figure 2(*b*)). Now, our assumptions concerning the ordinary limit imply that, at fixed (t, θ, ϕ) and at fixed $\alpha > 0$, f depends smoothly on β, including at $\beta = 0$. On the other hand, our assumptions concerning the scaled limit imply that at fixed (t, θ, ϕ) and at fixed $\beta > 0$, f is smooth in α. Furthermore, the last condition in the ordinary limit implies that for $\beta = 0$ and fixed (t, θ, ϕ), f is smooth in α, including at $\alpha = 0$.

Thus, at fixed (t, θ, ϕ), our previously stated assumptions imply that f is well defined at the 'origin' $(\alpha, \beta) = (0, 0)$, and is smooth in α along the α-axis (i.e., $\beta = 0$). However, our previously stated assumptions do not say anything about the continuity or smoothness of f as $(\alpha, \beta) \to (0, 0)$ from directions other than along the α-axis. Such limiting behavior would correspond to letting $r \to 0$ as $\lambda \to 0$ but at a rate *slower* than λ, i.e., such that $r/\lambda \to \infty$. To see the meaning and relevance of this limiting behavior, let us return to our original motivating example, equation (8) and take f to be the time–time component of this metric[6]. In terms of our new variables (17), we have

$$f = -(1 - 2M_0\beta - C_0\alpha^2), \tag{18}$$

which is jointly smooth in (α, β) at $(0, 0)$. By contrast, suppose we had a one-parameter family of metrics $\tilde{g}_{ab}(\lambda)$ that satisfies our above assumptions about the ordinary and scaled limits, but fails to be jointly smooth in (α, β) at $(0, 0)$. For example, suppose the time–time component of such a one-parameter family varied as

$$\tilde{f} = -\left(1 - \frac{\alpha\beta}{\alpha^2 + \beta^2}\right). \tag{19}$$

[5] Note that in this equation and in other equations occurring later in this paper, we relate components of tensors in the barred coordinates to the corresponding components of tensors in unbarred coordinates. Thus, a bar appears over the indices on the left-hand side of this equation, but not over the indices appearing on the right-hand side of this equation.

[6] Note that if we wished to consider other components of this metric, we would have to transform back from the coordinates (r, θ, ϕ) to 'Cartesian-like' coordinates x^i that are well behaved at the origin $x^i = 0$ when $\lambda = 0$.

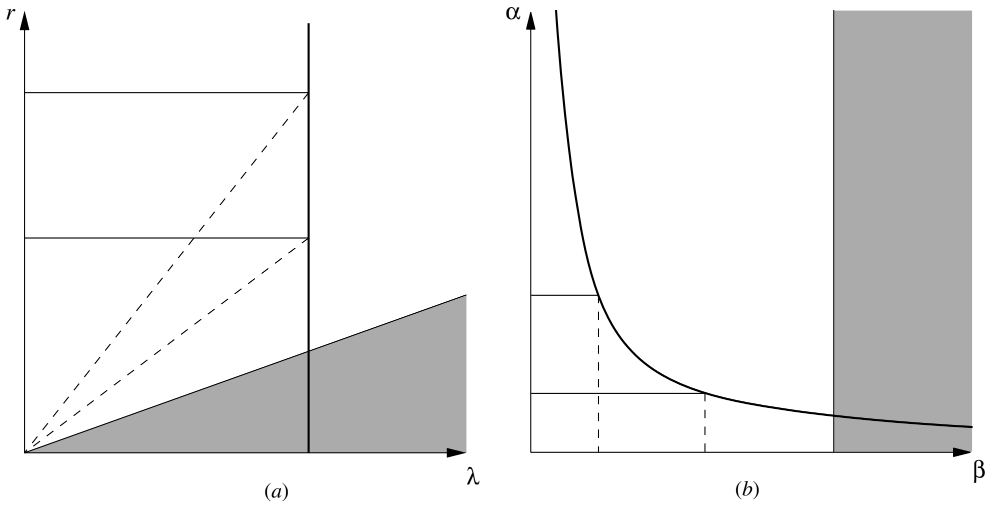

Figure 2. The two limits. (*a*) The two limits in terms of r and λ. A constant-λ spacetime is shown as a thick line. The shaded region corresponds to the interior of the (shrinking) body, about which we make no assumptions. The ordinary limit is represented by solid lines and the scaled limit is represented by dashed lines. While the ordinary background metric is on the r-axis, the scaled background metric is contained in the discontinuous behavior of the metric family at $r = \lambda = 0$. (*b*) The two limits in terms of α and β. In the new variables, the two types of limit appear on a more equal footing, with the ordinary and scaled background metrics placed on either axis. The body is 'pushed out' to finite β, so that assumptions made in a neighborhood of $\alpha = \beta = 0$ make no reference to the body.

In terms of the original variables (λ, r), this corresponds to behavior of the form

$$\tilde{f} = -\left(1 - \frac{\lambda r^2}{\lambda^2 + r^4}\right). \tag{20}$$

If we take the 'ordinary limit' ($\lambda \to 0$ at fixed $r > 0$) of $\tilde{f}$, we find that $\tilde{f}$ smoothly goes to -1. Similarly, if we take the 'scaled limit' ($\lambda \to 0$ at fixed $\bar{r} = r/\lambda > 0$), we also find that $\tilde{f}$ smoothly goes to -1. However, suppose we let $\lambda \to 0$ but let r go to zero as $r = c\sqrt{\lambda}$. Then $\tilde{f}$ will approach a different limit, namely $c^2/(1 + c^4)$. In essence, $\tilde{g}_{ab}(\lambda)$ corresponds to a one-parameter family in which there is a 'bump of curvature' at $r \propto \sqrt{\lambda}$. Although this 'bump of curvature' does not register when one takes the ordinary or scaled limits, it is present in the one-parameter family of spacetimes and represents unacceptable limiting behavior as $\lambda \to 0$ of this one-parameter family.

In order to eliminate this kind of non-uniform behavior in λ and r, we now impose the following additional condition:

- (iii) *Uniformity condition*: each component of $g_{ab}(\lambda)$ in the coordinates x^μ is a jointly smooth function of the variables (α, β) at $(0, 0)$ (at fixed (t, θ, ϕ)), where α and β are defined by equation (17).

Assumptions (i)–(iii) constitute all of the conditions that we shall impose on $g_{ab}(\lambda)$. No additional assumptions will be made in this paper.

We note that the coordinate freedom allowed by our conditions are precisely all coordinate transformations

$$x^\mu \to x'^\mu(\lambda, x^\nu), \tag{21}$$

such that $x'^{\mu}(\lambda, x^{\nu})$ is jointly smooth in (λ, x^{ν}) for all $r > C\lambda$ for some constant C, and such that the Jacobian matrix $\partial x'^{\mu}/\partial x^{\nu}$ is jointly smooth in (α, β) at $(0, 0)$ at fixed (t, θ, ϕ).

It should be emphasized that our assumptions place absolutely no restrictions on the one-parameter family of spacetimes for $r < \lambda\bar{R}$, i.e., this portion of these spacetimes could equally well be 'filled in' with ordinary matter or a black hole[7]. It should also be noted that the 'large r' region of the spacetime will not be relevant to any of our considerations, so it is only necessary that our conditions hold for $r < K$ for some constant K.

Finally, although it may not be obvious upon first reading, we note that our assumptions concerning $g_{ab}(\lambda)$ are closely related to the assumptions made in matched asymptotic expansion analyses. As we shall see in the following section, in essence, our assumption about the existence of an ordinary limit of $g_{ab}(\lambda)$ corresponds to assuming the existence of a 'far-zone' expansion; our assumption about the existence of a scaled limit of $g_{ab}(\lambda)$ corresponds to assuming the existence of a 'near-zone' expansion; and our uniformity assumption corresponds closely to the assumption of the existence of a 'buffer zone' where both expansions are valid.

3. Consequences of our assumptions

In this section, we derive some immediate consequences of the assumptions of the previous section that will play a key role in our analysis. These results will follow directly from the 'uniformity condition' and the consistency relation (16).

Since, by the uniformity assumption, the coordinate components of the one-parameter family of metrics $g_{ab}(\lambda)$ are jointly smooth in the variables (α, β) at $(0, 0)$, we may approximate $g_{\mu\nu}$ by a joint Taylor series in α and β to any finite orders N and M by

$$g_{\mu\nu}(\lambda; t, r, \theta, \phi) = \sum_{n=0}^{N}\sum_{m=0}^{M} \alpha^n \beta^m (a_{\mu\nu})_{nm}(t, \theta, \phi) + O(\alpha^{N+1}) + O(\beta^{M+1}). \tag{22}$$

Substituting for α and β from equation (17), we have

$$g_{\mu\nu}(\lambda; t, r, \theta, \phi) = \sum_{n=0}^{N}\sum_{m=0}^{M} r^n \left(\frac{\lambda}{r}\right)^m (a_{\mu\nu})_{nm}(t, \theta, \phi), \tag{23}$$

where here and in the following, we drop the error term. We can rewrite this equation as a perturbation expansion in λ,

$$g_{\mu\nu}(\lambda; t, r, \theta, \phi) = \sum_{m=0}^{M} \lambda^m \sum_{n=0}^{N} r^{n-m} (a_{\mu\nu})_{nm}(t, \theta, \phi). \tag{24}$$

We will refer to equation (24) as the *far-zone expansion* of $g_{ab}(\lambda)$. It should be noted that the mth-order term in λ in the far-zone perturbation expansion has leading order behavior of $1/r^m$ at small r. However, arbitrarily high positive powers of r may occur at each order in λ. Finally, we note that the angular dependence of $(a_{\mu\nu})_{n0}(t, \theta, \phi)$ is further restricted by the requirement that the metric components $g_{\mu\nu}(\lambda = 0)$ must be smooth at $r = 0$ when re-expressed as functions of x^i. In particular, this implies that $(a_{\mu\nu})_{00}$ cannot have any angular dependence.

[7] Indeed, it could also be 'filled in' with 'exotic matter' (failing to satisfy, say, the dominant energy condition) or a naked singularity (of positive or negative mass). However, in cases where there fails to be a well-posed initial-value formulation (as would occur with certain types of exotic matter and with naked singularities) and/or there exist instabilities (as would occur with other types of exotic matter), if is far from clear that one should expect there to exist a one-parameter family of solutions $g_{ab}(\lambda)$ satisfying our assumptions.

Equivalently, in view of equation (15), we can expand $\bar{g}_{\bar{\mu}\bar{\nu}}$ as

$$\bar{g}_{\bar{\mu}\bar{\nu}}(\lambda; t_0; \bar{t}, \bar{r}, \theta, \phi) = \sum_{n=0}^{N}\sum_{m=0}^{M}(\lambda\bar{r})^n\left(\frac{1}{\bar{r}}\right)^m (a_{\mu\nu})_{nm}(t_0 + \lambda\bar{t}, \theta, \phi)$$
$$= \sum_{n=0}^{N}\sum_{m=0}^{M}\lambda^n\left(\frac{1}{\bar{r}}\right)^{m-n} (a_{\mu\nu})_{nm}(t_0 + \lambda\bar{t}, \theta, \phi). \tag{25}$$

By further expanding $(a_{\mu\nu})_{nm}$ in $\bar{t}$ about $\bar{t} = 0$, we obtain

$$\bar{g}_{\bar{\mu}\bar{\nu}}(\lambda; t_0; \bar{t}, \bar{r}, \theta, \phi) = \sum_{n=0}^{N}\sum_{m=0}^{M}\sum_{p=0}^{P}\lambda^{n+p}\bar{t}^p\left(\frac{1}{\bar{r}}\right)^{m-n}(b_{\mu\nu})_{nmp}(t_0, \theta, \phi), \tag{26}$$

where

$$(b_{\mu\nu})_{nmp} \equiv \frac{1}{p!}\frac{\partial^p}{\partial t^p}(a_{\mu\nu})_{nm}\Big|_{t=t_0}. \tag{27}$$

We can rewrite this as a perturbation series expansion in λ,

$$\bar{g}_{\bar{\mu}\bar{\nu}}(\lambda; t_0; \bar{t}, \bar{r}, \theta, \phi) = \sum_{q=0}^{N+P}\lambda^q\sum_{p=0}^{\min(q,P)}\sum_{m=0}^{M}\bar{t}^p\left(\frac{1}{\bar{r}}\right)^{m-q+p}(b_{\mu\nu})_{(q-p)mp}(t_0; \theta, \phi). \tag{28}$$

We will refer to equation (28) as the *near-zone expansion* of $g_{ab}(\lambda)$. We see from this formula that the scaled metric, viewed as a perturbation series in λ, follows the rule that the combined powers of $\bar{t}$ and $\bar{r}$ are allowed to be only as positive as the order in perturbation theory. By contrast, inverse powers of $\bar{r}$ of arbitrarily high order are always allowed. Of course, only non-negative powers of $\bar{t}$ can occur.

By inspection of equation (28), we see that the 'background' ($\lambda = 0$) scaled metric is given by

$$\bar{g}_{\bar{\mu}\bar{\nu}}(\lambda = 0; t_0; \bar{t}, \bar{r}, \theta, \phi) = \sum_{m=0}^{M}\left(\frac{1}{\bar{r}}\right)^m (a_{\mu\nu})_{0m}(t_0; \theta, \phi), \tag{29}$$

where we have used the fact that $(b_{\mu\nu})_{0m0} = (a_{\mu\nu})_{0m}$. Thus, we see that there is no dependence of $\bar{g}_{\bar{\mu}\bar{\nu}}(\lambda = 0; t_0)$ on $\bar{t}$ and only non-positive powers of $\bar{r}$ occur. *Thus, we see that* $\bar{g}_{ab}(\lambda = 0)$ *is a stationary, asymptotically flat spacetime.* However, the limiting, stationary, asymptotically flat spacetime that we obtain may depend on the choice of the time, t_0, on the worldline, γ, about which the scaling is done.

Our 'far-zone expansion', equation (24), appears to correspond closely to the far-zone expansion used in matched asymptotic expansion analyses [5, 8]. However, our 'near-zone expansion' differs in that we define a separate expansion for each time t_0 rather than attempting a uniform in time approximation with a single expansion. Such expansions require an additional 'quasi-static' or slow-time variation assumption for the evolution of the metric perturbations. A further difference is that the conclusion that the background ($\lambda = 0$) metric is stationary and asymptotically flat has been derived here rather than assumed. Indeed, in other analyses, a particular form of the background metric (such as the Schwarzschild metric) is assumed, and the possibility that this metric form might change with time (i.e., depend upon t_0) is not considered. In addition, in other analyses boundary conditions at small $\bar{r}$ (such as regularity on the Schwarzschild horizon) are imposed. In our analysis, we make no assumptions other than the assumptions (i)–(iii) stated in the previous section. In particular, since we make no assumptions about the form of the metric for $\bar{r} < \bar{R}$, we do not impose any boundary conditions at small $\bar{r}$.

Finally, it is also useful to express the consistency relation (16) in a simple, differential form. We define

$$(K_{\mu\nu})_{npm}(\lambda; t_0; \bar{t}, \bar{x}^i) \equiv \left(\frac{\partial}{\partial\lambda}\right)^n \left(\frac{\partial}{\partial t_0}\right)^p \left(\frac{\partial}{\partial\bar{t}}\right)^m \bar{g}(\lambda; t_0; \bar{t}, \bar{x}^i)|_{\lambda=0, t_0, \bar{t}=0}. \tag{30}$$

Then, a short calculation shows that

$$K_{(n+1)p(m+1)} = (n+1)K_{n(p+1)m} \tag{31}$$

as well as

$$K_{npm} = 0 \qquad \text{if} \quad n < m. \tag{32}$$

Setting $n = 0$, we see that the last relation implies that $\bar{g}(\lambda = 0; t_0; \bar{t}, \bar{x}^i)$ is stationary, as we have already noted.

4. Geodesic motion

In this section, we will prove that the worldline γ appearing in assumption (i) of section 2 must, in fact, be a geodesic of the background metric $g_{ab}(\lambda = 0)$. This can be interpreted as establishing that, to zeroth order in λ, any body (or black hole) moves on a geodesic of the background spacetime. In fact, we will show considerably more than this: we will show that, to first order in λ, the far-zone description of $g_{ab}(\lambda)$ is that of a 'point particle'. As previously mentioned in the introduction, our derivation of geodesic motion is similar in spirit to that of [16] in that we consider one-parameter families of solutions with a worldtube that shrinks down to a curve γ, but the nature of the one-parameter families that we consider here are very different from those considered by [16], and our proof of geodesic motion is very different as well. Our derivation of geodesic motion also appears to differ significantly from previous derivations using matched asymptotic expansions [5, 8, 12, 13].

We begin by writing the lowest order terms in the far-zone expansion, equation (24), as follows:

$$\begin{aligned} g_{\alpha\beta}(\lambda) &= (a_{\alpha\beta})_{00}(t) + (a_{\alpha\beta})_{10}(t,\theta,\phi)r + O(r^2) \\ &\quad + \lambda\left[(a_{\alpha\beta})_{01}(t,\theta,\phi)\frac{1}{r} + (a_{\alpha\beta})_{11}(t,\theta,\phi) + O(r)\right] + O(\lambda^2), \qquad r > 0, \end{aligned} \tag{33}$$

where we have used the fact that $(a_{\alpha\beta})_{00}$ can depend only on t, as noted in the previous section. Since the worldline γ, given by $x^i = 0$, was assumed to be timelike[8] in the spacetime $g_{ab}(\lambda = 0)$, without loss of generality, we may choose our coordinates x^α so that $g_{\alpha\beta}(\lambda = 0, x^i = 0) = (a_{\alpha\beta})_{00}(t) = \eta_{\alpha\beta}$. (One such possible choice of coordinates would be Fermi normal coordinates with respect to γ in the metric $g_{ab}(\lambda = 0)$. We emphasize that we make the coordinate choice $(a_{\alpha\beta})_{00} = \eta_{\alpha\beta}$ purely for convenience—so that, e.g., coordinate time coincides with proper time on γ—but it plays no essential role in our arguments.) Choosing these coordinates, and letting $h_{\alpha\beta}$ denote the $O(\lambda)$ piece of the metric, we see that $g_{\alpha\beta}(\lambda)$ takes the form

$$g_{\alpha\beta} = \eta_{\alpha\beta} + O(r) + \lambda h_{\alpha\beta} + O(\lambda^2), \tag{34}$$

where

$$h_{\alpha\beta} = \frac{c_{\alpha\beta}(t,\theta,\phi)}{r} + O(1), \tag{35}$$

[8] We made this assumption explicitly in condition (i) of section 2. However, if, instead, we had assumed that the 'interior region' $r \leqslant \lambda\bar{R}$ was 'filled in' with matter satisfying the dominant energy condition, then it should be possible to prove that γ must be timelike.

where in equation (35), the term '$O(1)$' denotes a term that, when multiplied by r^ϵ for any $\epsilon > 0$, vanishes as $r \to 0$.

Now, by assumption (i) of section 2, for each λ, $g_{ab}(\lambda)$ is a vacuum solution of Einstein's equation for $r > \lambda\bar{R}$ and is jointly smooth in (λ, x^α) in this coordinate range. It follows that for all $r > 0$, h_{ab} is a solution to the linearized Einstein equation off of $g_{ab}(\lambda = 0)$, i.e.,

$$G^{(1)}_{ab}[h_{cd}] = -\tfrac{1}{2}\nabla_a\nabla_b h^c_c - \tfrac{1}{2}\nabla^c\nabla_c h_{ab} + \nabla^c\nabla_{(b}h_{a)c} = 0, \qquad r > 0, \tag{36}$$

where here and in the following, ∇_a denotes the derivative operator associated with $g_{ab}(\lambda = 0)$, and indices are raised and lowered with $g_{ab}(\lambda = 0)$. Equation (36) holds only for $r > 0$, and, indeed, if $c_{\alpha\beta} \neq 0$, h_{ab} is singular at $r = 0$. However, even if $c_{\alpha\beta} \neq 0$, the singularity of each component of h_{ab} is locally integrable with respect to the volume element associated with $g_{ab}(\lambda = 0)$, i.e., each component, $h_{\alpha\beta}$, is a locally L^1 function on the entire spacetime manifold, including $r = 0$. *Thus, h_{ab} is well defined as a distribution on all of spacetime.* The quantity $T_{ab} \equiv G^{(1)}_{ab}[h_{cd}]/8\pi$ is therefore automatically well defined as a distribution. This quantity has the interpretation of being the 'source' for the metric perturbation (35)—even though all of our spacetimes $g_{ab}(\lambda)$ for $\lambda > 0$ have excluded the 'source region' $r \leqslant \lambda\bar{R}$. It follows immediately from equation (36) that, as a distribution, T_{ab} must have support on γ in the sense that it must vanish when acting on any test tensor field f^{ab} whose support does not intersect γ. We now compute T_{ab}.

By definition, $T_{ab} \equiv G^{(1)}_{ab}[h_{cd}]/8\pi$ is the distribution on spacetime whose action on an arbitrary smooth, compact support, symmetric tensor field $f_{cd} = f_{dc}$ is given by

$$8\pi T(f) = \int_M G^{(1)}_{ab}[f_{cd}]h^{ab}\sqrt{-g}\,\mathrm{d}^4x = 0, \tag{37}$$

where $\sqrt{-g}\,\mathrm{d}^4x$ denotes the volume element associated with $g_{ab}(\lambda = 0)$ and we have used the fact that the operator $G^{(1)}_{ab}$ is self-adjoint[9]. Note that the right-hand side of this equation is well defined since $G^{(1)}_{ab}[f_{cd}]$ is a smooth tensor field of compact support and h^{ab} is locally L^1. We can evaluate the right-hand side of equation (37) by integrating over the region $r > \epsilon > 0$ and then taking the limit as $\epsilon \to 0$. In the region $r > \epsilon$, h^{ab} is smooth, and a straightforward 'integration by parts' calculation shows that

$$G^{(1)}_{ab}[f_{cd}]h^{ab} - G^{(1)}_{ab}[h_{cd}]f^{ab} = \nabla_c s^c, \tag{38}$$

where

$$\begin{aligned} s^c = {}& h^{ab}\nabla^c f_{ab} - \nabla^c h^{ab} f_{ab} + h^{bc}\nabla_b f - \nabla_b h^{bc} + 2\nabla^a h^{bc} f_{ab} - 2h_{ab}\nabla^a f^{bc} \\ & + \nabla^c h f - h\nabla^c f + h\nabla_b f^{bc} - \nabla_b h f^{bc}, \end{aligned} \tag{39}$$

where $f = f_{ab}g^{ab}(\lambda = 0)$. Since $G^{(1)}_{ab}[h_{cd}] = 0$ for $r > 0$, it follows immediately that

$$T(f) = \frac{1}{8\pi}\lim_{\epsilon\to 0}\int_{r>\epsilon} G^{(1)}_{ab}[f_{cd}]h^{ab} = \frac{1}{8\pi}\lim_{\epsilon\to 0}\int_{r=\epsilon} s^a n_a\,\mathrm{d}S. \tag{40}$$

Using equations (35) and (39), we find that $T(f)$ takes the form

$$T(f) = \int \mathrm{d}t\, N_{ab}(t) f^{ab}(t, r = 0), \tag{41}$$

[9] See [23] for the definition of adjoint being used here. If $G^{(1)}_{ab}$ were not self-adjoint, then the adjoint of $G^{(1)}_{ab}$ would have appeared in equation (37).

where $N_{ab}(t)$ is a smooth, symmetric ($N_{ab} = N_{ba}$) tensor field on γ whose components are given in terms of suitable angular averages of $c_{\alpha\beta}$ and its first angular derivatives. In other words, the distribution T_{ab} is given by[10]

$$T_{ab} = N_{ab}(t)\frac{\delta^{(3)}(x^i)}{\sqrt{-g}}\frac{\mathrm{d}\tau}{\mathrm{d}t}, \tag{42}$$

where $\delta^{(3)}(x^i)$ is the 'coordinate delta function' (i.e., $\int \delta^{(3)}(x^i)\,\mathrm{d}^3x^i = 1$).

We now use the fact that, since the differential operator $G^{(1)}_{ab}$ satisfies the linearized Bianchi identity $\nabla^a G^{(1)}_{ab} = 0$, the distribution T_{ab} must satisfy $\nabla^a T_{ab} = 0$ in the distributional sense. This means that the action of T_{ab} must vanish on any test tensor field of the form $f_{ab} = \nabla_{(a} f_{b)}$ where f_a is smooth and of compact support. In other words, by equation (41), the tensor field N_{ab} on γ must be such that for an arbitrary smooth vector field f^a on spacetime, we have

$$\int \mathrm{d}t\, N_{ab}(t)\nabla^a f^b(t, r = 0) = 0. \tag{43}$$

Now for any $i = 1, 2, 3$, choose f^a to have components of the form $f^\mu = x^i F(x^1, x^2, x^3)c^\mu(t)$, where each c^μ ($\mu = 0, 1, 2, 3$) is an arbitrary smooth function of compact support in t and F is a smooth function of compact spatial support, with $F = 1$ in a neighborhood of γ. Then equation (43) yields

$$\int \mathrm{d}t\, N_{i\mu}(t)c^\mu(t) = 0 \tag{44}$$

for all $c^\mu(t)$, which immediately implies that $N_{i\mu} = N_{\mu i} = 0$ for all $i = 1, 2, 3$ and all $\mu = 0, 1, 2, 3$. In other words, we have shown that $N_{ab}(t)$ must take the form

$$N_{ab} = M(t)u_a u_b, \tag{45}$$

where u^a denotes the unit tangent to γ, i.e., u^a is the 4-velocity of γ. Now choose f^a to be an arbitrary smooth vector field of compact support. Then equations (43) and (45) yield

$$0 = \int \mathrm{d}t\, M(t)u_b(u_a\nabla^a f^b) = -\int \mathrm{d}t\, u^a\nabla_a(M(t)u_b)f^b, \tag{46}$$

where we integrated by parts in t to obtain the last equality. Since f^a is arbitrary, this immediately implies that

$$u^a\nabla_a(M(t)u_b) = 0. \tag{47}$$

This, in turn, implies that

$$\mathrm{d}M/\mathrm{d}t = 0, \tag{48}$$

i.e., M is a constant along γ, and, if $M \neq 0$,

$$u^a\nabla_a u^b = 0, \tag{49}$$

i.e., in the case where $M \neq 0$, γ is a geodesic of $g_{ab}(\lambda = 0)$, as we desired to show[11].

In summary, we have shown that for any one-parameter family of metrics $g_{ab}(\lambda)$ satisfying assumptions (i)–(iii) of section 2, to first order in λ, the far-zone metric perturbation h_{ab} corresponds to a solution to the linearized Einstein equation with a point particle source

$$T_{ab} = M u_a u_b\frac{\delta^{(3)}(x^i)}{\sqrt{-g}}\frac{\mathrm{d}\tau}{\mathrm{d}t}, \tag{50}$$

[10] In fact, by our coordinate choice, we have $\sqrt{-g} = 1$ on γ and $\frac{\mathrm{d}\tau}{\mathrm{d}t} = 1$, but we prefer to leave in these factors so that this formula holds for an arbitrary choice of coordinates.

[11] Some previous derivations [5, 8, 12] of geodesic motion do not appear to make explicit use of the fact that $M \neq 0$. It is critical that this assumption be used in any valid derivation of geodesic motion, since a derivation that holds for $M = 0$ effectively would show that all curves are geodesics.

where M is a constant and u^a is the 4-velocity of γ, which must be a geodesic if $M \neq 0$. We refer to M as the *mass* of the particle. It is rather remarkable that the point particle source (50) is an *output* of our analysis rather than an input. Indeed, we maintain that the result we have just derived is what provides the justification for the notion of 'point particles'—a notion that has played a dominant role in classical physics for more than three centuries. In fact, the notion of point particles makes no mathematical sense in the context of nonlinear field theories like general relativity. Nevertheless, we have just shown that the notion of a (structureless) 'point particle' arises naturally as an *approximate* description of sufficiently small bodies—namely, a description that is valid to first order in λ in the far zone for arbitrary one-parameter families of metrics $g_{ab}(\lambda)$ satisfying the assumptions of section 2. This description is valid independently of the nature of the 'body', e.g., it holds with equal validity for a small black hole as for a small blob of ordinary matter.

5. Description of motion to first order in λ

In the previous section, we established that, to zeroth order in λ, any body (or black hole) of nonvanishing mass moves on a geodesic of the background spacetime. Much of the remainder of this paper will be devoted to finding the corrections to this motion, valid to first order in λ in the far zone. In this section, we address the issue of what is meant by the 'motion of the body' to first order in λ.

The first point that should be clearly recognized is that it is far from obvious how to describe 'motion' in terms of a worldline for $\lambda > 0$. Indeed, the metric $g_{ab}(\lambda)$ is defined only for $r > \lambda\bar{R}$, so at finite λ the spacetime metric may not even be defined in a neighborhood of γ. If we were to assume that $\bar{R} \gg M$ and that the region $r < \lambda\bar{R}$ were 'filled in' with sufficiently 'weak field matter'—so that $\bar{\mathcal{R}}\bar{R}^2 \ll 1$, where $\bar{\mathcal{R}}$ denotes the supremum of the components of the Riemann curvature tensor of $\bar{g}_{ab}(\lambda)$ in the 'filled-in' region—then it should be possible to define a 'center-of-mass' worldline at finite λ, and we could use this worldline to characterize the motion of the body [24]. However, we do not wish to make any 'weak field' assumptions here, since we wish to describe to motion of small black holes and other 'strong field' objects. Since it is not clear how to associate a worldline to the body at finite λ—and, indeed, the 'body' is excluded from the spacetime region we consider at finite λ—it is not clear what one would mean by a 'perturbative correction' to γ to first or higher order in λ.

A second point that should be understood is that if we have succeeded in defining the worldlines describing the motion the body at finite λ,

$$x^i(\lambda, t) = z^i(\lambda, t) = \lambda Z^i(t) + O(\lambda^2), \tag{51}$$

then the 'first order in λ perturbative correction', Z^i, to the zeroth-order motion γ (given by $x^i(t) = 0$) is most properly viewed as the spatial components of a vector field, Z^a, defined along γ. This vector field describes the 'infinitesimal displacement' to the corrected motion to first order in λ. The time component, Z^0, of Z^a depends on how we identify the time parameter of the worldlines at different values of λ and is not physically relevant; we will set $Z^0 = 0$ so that Z^a is orthogonal to the tangent, u^a, to γ in the background metric $g_{ab}(\lambda = 0)$. Thus, when we seek equations of motion describing the first-order perturbative corrections to geodesic motion, we are seeking equations satisfied by the vector field $Z^a(t)$ on γ.

A third point that should be clearly recognized is that Z^a and any equations of motion satisfied by Z^a will depend on our choice of gauge for h_{ab}. To see this explicitly, suppose that we perform a smooth[12] gauge transformation of the form

$$x^\mu \to \hat{x}^\mu = x^\mu - \lambda A^\mu(x^\nu) + O(\lambda^2). \tag{52}$$

Under this gauge transformation, we have

$$h_{\mu\nu} \to \hat{h}_{\mu\nu} = h_{\mu\nu} - 2\nabla_{(\mu}A_{\nu)}. \tag{53}$$

However, clearly, the new description of motion will be of the form [19]

$$\hat{x}^i(\lambda, t) = \hat{z}^i(\lambda, t), \tag{54}$$

where

$$\hat{z}^i(t) = z^i(t) - \lambda A^i(t, x^j = 0) + O(\lambda^2). \tag{55}$$

Thus, we see that Z^a transforms as

$$Z^i(t) \to \hat{Z}^i(t) = Z^i(t) - A^i(t, x^j = 0) \tag{56}$$

in order that it describe the same perturbed motion. Thus, the first-order correction, $Z^a(t)$, to the background geodesic motion contains no meaningful information by itself and, indeed, it can always be set to zero by a smooth gauge transformation. Only the pair $(h_{ab}, Z^a(t))$ has gauge-invariant meaning.

We turn now to the definition of the first-order perturbed motion. Our definition relies on the fact, proven in section 3, that for each t_0, $\bar{g}_{\bar{\mu}\bar{\nu}}(\lambda = 0; t_0; \bar{x}^\alpha)$ is a stationary, asymptotically flat spacetime. Therefore, $\bar{g}_{ab}(\lambda = 0)$ has well-defined sets of mass ('electric parity') and angular momentum ('magnetic parity') multipole moments [25, 26] and, indeed, the spacetime is characterized by the values of these two sets of multipole moments [27, 28]. The multipole moments (other than the lowest nonvanishing multipoles of each type) depend upon a choice of conformal factor [25, 26], which, roughly speaking, corresponds to a choice of 'origin'. We choose the conformal factor $\Omega = 1/\bar{r}^2$ to define all of the multipoles, corresponding to choosing the origin at $\bar{r} = 0$. For a metric of the form equation (29)—with $(a_{\mu\nu})_{00} = \eta_{\mu\nu}$ by our coordinate choice imposed in the previous section that $g_{\mu\nu}(\lambda = 0) = \eta_{\mu\nu}$ on γ—the mass will be simply the $l = 0$ part of the coefficient of $1/\bar{r}$ in the large $\bar{r}$ expansion of $\frac{1}{2}\bar{g}_{\bar{t}\bar{t}}(\lambda = 0; t_0)$. Similarly, the mass dipole moment will be the coefficient of the $l = 1$ part of this quantity at order $1/\bar{r}^2$.

It is well known that if the mass is nonzero, the mass dipole moment is 'pure gauge' and can be set to zero by choice of conformal factor/'origin'. We now explicitly show that, with our choice of conformal factor $\Omega = 1/\bar{r}^2$, the mass dipole moment can be set to zero by a smooth gauge transformation of the form (52). It follows from the linearized Einstein equation (36) with source (50) applied to h_{ab}, equation (35), that the time–time component of h_{ab} takes the form

$$h_{tt} = \frac{2M}{r} + O(1), \tag{57}$$

i.e., in the notation of equation (24), we have $(a_{tt})_{01} = 2M$. Comparing with equation (28) (and also using the fact that $(a_{tt})_{00} = -1$), we see that at each t_0

$$\bar{g}_{\bar{t}\bar{t}}(\lambda = 0; t_0) = -\left(1 - \frac{2M}{\bar{r}}\right) + O(1/\bar{r}^2). \tag{58}$$

[12] Gauge transformations where A^μ is not smooth at $x^i = 0$ are also permitted under the coordinate freedom stated at the end of section 2. However, it suffices to consider smooth A^ν for our considerations here. The change in the description of motion under non-smooth gauge transformations will be treated in the appendix.

From this equation, we see that the 'particle mass', M, of the 'source' of the far-zone metric perturbation (see equation (50)) is also the Komar/ADM mass of the stationary, asymptotically flat spacetime $\bar{g}_{ab}(\lambda = 0; t_0)$. We now calculate the effect of the coordinate transformation (52) on $\bar{g}_{\bar{t}\bar{t}}(\lambda = 0; t_0)$. The transformation (52) corresponds to changing the barred coordinates by

$$\bar{x}^\mu \to \hat{\bar{x}}^\mu = \bar{x}^\mu - A^\mu(t_0, x^i = 0) + O(\lambda), \tag{59}$$

i.e., to zeroth order in λ, it corresponds to a 'constant displacement' of coordinates. Since

$$\frac{1}{\bar{r}} = \frac{1}{|\hat{\bar{x}}^i + A^i(t_0, 0)|} = \frac{1}{\hat{\bar{r}}} - \frac{A_i x^i}{\hat{\bar{r}}^3} + O(1/\hat{\bar{r}}^3), \tag{60}$$

it can be seen that this transformation has the effect of changing the mass dipole moment by $-MA^i$. In particular, we can always choose A^i so as to set the mass dipole moment to zero.

Now, the 'near-zone' coordinates $\hat{\bar{x}}^i$ for which the mass dipole moment vanishes have the interpretation of being 'body centered' coordinates to zeroth order in λ. The origin $\hat{x}^i = 0$ of the corresponding 'far-zone' coordinates $\hat{x}^i$ therefore has the interpretation of representing the 'position' of the center of mass of the body to first order in λ. We shall use this to define the correction to geodesic motion to first order in λ by proceeding as follows:

First, we shall choose our coordinates, x^μ, to zeroth order in λ by choosing convenient coordinates for the 'background spacetime' $g_{ab}(\lambda = 0)$. (We will use Fermi normal coordinates based on γ.) Next, we will define our coordinates, x^μ, to first order in λ by choosing a convenient gauge for h_{ab}, equation (35). (We will choose the Lorenz gauge $\nabla^a(h_{ab} - \frac{1}{2}hg_{ab}) = 0$.) Then we will introduce the smooth coordinate transformation (52), and impose the requirement that A^μ be such that the mass dipole moment of $\bar{g}_{ab}(\lambda = 0; t_0)$ vanish for all t_0. Since the 'location' of the body in the new coordinates is $\hat{z}^i(t) = 0$, the first-order perturbative correction $Z^a(t)$ to the motion of the body in our original coordinates x^μ will be given by

$$Z^i(t) = A^i(t, x^j = 0). \tag{61}$$

Of course, the particular $Z^a(t)$ that we obtain in any given case will depend upon the particular one-parameter family $g_{ab}(\lambda)$ that we consider. What is of interest is any 'universal relations' satisfied by $Z^a(t)$ that are independent of the choice of one-parameter family satisfying assumptions (i)–(iii) of section 2. Such universal relations would provide us with 'laws of motion' for point particles that take self-force effects into account. In the following section, we will show (via a lengthy calculation) that such a universal relation exists for $\mathrm{d}^2Z^i/\mathrm{d}t^2$, thus providing us with general 'equations of motion' for all 'point particles', valid to first order in λ.

Finally, we note that if we wish to describe motion beyond first order in λ, it will be necessary to define a 'representative world line' in the far zone to at least second order in λ. We shall not attempt to do so in this paper. The definition of a suitable representative worldline is probably the greatest obstacle to extending the results of this paper to higher order in perturbation theory.

6. Computation of perturbed motion

In the section 4, we found that first-order far-zone perturbations of the background spacetime $g_{ab}(\lambda = 0)$ are sourced by a point particle stress energy, equation (50). For the remainder of this paper, we will assume that $M \neq 0$, so that, as shown in section 4, the lowest order motion is described by a geodesic, γ, of the background spacetime. We will need expressions for the components of the far-zone metric, $g_{\mu\nu}|_{\lambda=0}$, its first-order perturbation, $h_{\mu\nu} \equiv \partial g_{\mu\nu}/\partial\lambda|_{\lambda=0}$, and its second-order perturbation $j_{\mu\nu} \equiv \frac{1}{2}\partial^2 g_{\mu\nu}/\partial\lambda^2|_{\lambda=0}$. It is convenient to choose our coordinates x^μ to zeroth order in λ to be Fermi normal coordinates with respect to the background geodesic γ, and choose these coordinates to first order in λ so that $h_{\mu\nu}$ satisfies the Lorenz gauge condition $\nabla^b\tilde{h}_{ab} = 0$, where $\tilde{h}_{ab} \equiv h_{ab} - \frac{1}{2}hg_{ab}|_{\lambda=0}$ with $h \equiv h_{ab}g^{ab}|_{\lambda=0}$. Then the linearized Einstein equation reads

$$\nabla^c\nabla_c\tilde{h}_{ab} - 2R^c{}_{ab}{}^d\tilde{h}_{cd} = -16\pi M\int_\gamma \delta_4(x, z(\tau))u_a(\tau)u_b(\tau)\,\mathrm{d}\tau, \tag{62}$$

$$\nabla^b\tilde{h}_{ab} = 0. \tag{63}$$

This system of equations can be solved using the Hadamard expansion techniques of DeWitt and Brehme [5, 6, 9]. Since this technology has been used in all previous derivations of gravitational self-force, we do not review it here but simply present results. Equation (2.27) of [5] provides a covariant expression for the perturbations in terms of parallel propagators and Synge's world function on the background metric (see, e.g., [8] for definitions of these quantities). The Fermi normal coordinate components of these tensors are easily calculated with the aid of expressions from section 8 of Poisson [8]. Combining this with the form of the background metric in Fermi normal coordinates, we obtain

$$\begin{aligned} g_{\alpha\beta}(\lambda; t, x^i) = \eta_{\alpha\beta} + B_{\alpha i\beta j}(t)x^ix^j + O(r^3) + \lambda\bigg(\frac{2M}{r}\delta_{\alpha\beta} + h^{\text{tail}}_{\alpha\beta}(t, 0) \\ + h^{\text{tail}}_{\alpha\beta i}(t, 0)x^i + M\mathcal{R}_{\alpha\beta}(t, x^i) + O(r^2)\bigg) + O(\lambda^2), \end{aligned} \tag{64}$$

where the quantities $B_{\alpha\beta\gamma\delta}$ and $\mathcal{R}_{\alpha\beta}$ are defined by the following expressions in terms of the Fermi normal coordinate components of the Riemann tensor of the background metric

$$B_{0k0l} = -R_{0k0l} \qquad \mathcal{R}_{00} = 7R_{0k0l}\frac{x^kx^l}{r} \tag{65}$$

$$B_{ik0l} = -\frac{2}{3}R_{ik0l} \qquad \mathcal{R}_{i0} = \frac{2}{3}R_{ik0l}\frac{x^kx^l}{r} - 2R_{i0k0}x^k \tag{66}$$

$$B_{ikjl} = -\frac{1}{3}R_{ikjl} \qquad \mathcal{R}_{ij} = -\frac{13}{3}R_{ikjl}\frac{x^kx^l}{r} - 4rR_{i0j0}, \tag{67}$$

and $h^{\text{tail}}_{\alpha\beta}$ and $h^{\text{tail}}_{\alpha\beta\gamma}$ are given by

$$h^{\text{tail}}_{\alpha\beta}(x) \equiv M\int_{-\infty}^{\tau^-_{\text{ret}}}\left(G_{+\alpha\beta\alpha'\beta'} - \frac{1}{2}g_{\alpha\beta}G_+{}^\gamma{}_{\gamma\alpha'\beta'}\right)(x, z(\tau'))u^{\alpha'}u^{\beta'}\,\mathrm{d}\tau', \tag{68}$$

$$h^{\text{tail}}_{\alpha\beta\gamma}(x) \equiv M\int_{-\infty}^{\tau^-_{\text{ret}}}\nabla_\gamma\left(G_{+\alpha\beta\alpha'\beta'} - \frac{1}{2}g_{\alpha\beta}G_+{}^\delta{}_{\delta\alpha'\beta'}\right)(x, z(\tau'))u^{\alpha'}u^{\beta'}\,\mathrm{d}\tau'. \tag{69}$$

In these expressions, G_+ is the Lorenz gauge retarded Green's function, normalized with a factor of -16π, following [6]. As previously mentioned, the symbol τ^-_{ret} indicates that the range of the integral extends just short of the retarded time τ_{ret}, so that only the 'tail' (i.e., interior of the light cone) portion of Green's function is used (see, e.g., [8] for details). We define $h^{\text{tail}}_{\alpha\beta\gamma}$, rather than working with derivatives of $h^{\text{tail}}_{\alpha\beta}$, because $h^{\text{tail}}_{\alpha\beta}$ is not differentiable on the worldline. (However, this non-differentiability is limited only to spatial derivatives of spatial components of $h^{\text{tail}}_{\alpha\beta}$,[13] so that expressions like (4) are well defined.) A choice of retarded solution (corresponding to 'no incoming radiation') was made in writing these equations. This choice is not necessary, and one could add an arbitrary smooth solution $h_{\alpha\beta}$ of the linearized Einstein equation to the first order in λ term on the right-hand side of equation (64), which could then be carried through all of our calculations straightforwardly. However, for simplicity, we will not consider the addition of such a term.

Our derivation of gravitational self-force to first order in λ will require consideration of second-order metric perturbations, so we will have to carry the expansion of $g_{ab}(\lambda)$ somewhat beyond equation (64). (This should not be surprising in view of fact that our above derivation in section 4 of geodesic motion at zeroth order in λ required consideration of first-order metric perturbations.) In particular, we will need an explicit expression for the quantity $(a_{\mu\nu})_{02}$ appearing in the far-zone expansion equation (24), i.e., the term of order λ^2 that has the most singular behavior in $1/r$ (namely, $1/r^2$).

The second-order perturbation j_{ab} satisfies the second-order Einstein equation, which takes the form

$$G^{(1)}_{ab}[j] = -G^{(2)}_{ab}[h, h], \tag{70}$$

where $G^{(2)}_{ab}$ denotes the second-order Einstein tensor about the background metric $g^{ab}|_{\lambda=0}$. Since the $O(1/r)$ part of h_{ab} corresponds to the linearized Schwarzschild metric in isotropic coordinates (see equation (64)), it is clear that there is a particular solution to equation (70) of the form

$$j^I_{\alpha\beta} = \frac{M^2}{r^2}(-2t_\alpha t_\beta + 3n_\alpha n_\beta) + O(r^{-1}) \tag{71}$$

as $r \to 0$, where

$$n^i \equiv x^i/r \tag{72}$$

and $n^0 = 0$, whereas $t_\alpha \equiv \delta_{\alpha 0}$. (The explicitly written term on the right-hand side of equation (71) is just the $O(M^2)$ part of the Schwarzschild metric in isotropic coordinates.) The general solution to equation (70) can then be written as

$$j_{ab} = j^I_{ab} + j^H_{ab}, \tag{73}$$

where j^H_{ab} is a homogeneous solution of the linearized Einstein equation. We wish to compute the $O(1/r^2)$ part of j^H_{ab}, i.e., writing

$$j^H_{ab} = \frac{C_{ab}(t, \theta, \phi)}{r^2} + O(r^{-1}), \tag{74}$$

we wish to compute C_{ab}. Now, although the equations of motion to first order in λ depend upon a choice of gauge to first order in λ (see section 5), they cannot depend upon a choice of gauge to second order in λ, since a second-order gauge transformation cannot affect the mass dipole moment of the background-scaled metric $\bar{g}_{\bar{\mu}\bar{\nu}}(\lambda = 0)$. (We have also verified by a direct, lengthy computation that second-order gauge transformations do not produce changes in the equations of motion to first order in λ.) Therefore, we are free to impose any (admissible) second-order gauge condition on j^H_{ab}. It will be convenient to require that the Lorenz gauge condition $\nabla^a\left(j^H_{ab} - \frac{1}{2}j^H g_{ab}\right) = 0$ be satisfied to order $1/r^3$. The $O(1/r^4)$ part of the linearized Einstein equation together with the $O(1/r^3)$ part of the Lorenz gauge condition then yields

[13] This can be seen from the fact that differentiation of $h^{\text{tail}}_{\alpha\beta}$ on the worldline γ yields $h^{\text{tail}}_{\alpha\beta\gamma}$ plus the coincidence limit of the integrand of (68), which is proportional to $R_{\alpha 0\beta 0}$ times the gradient of τ_{ret}.

$$\partial^i \partial_i \left(\frac{1}{r^2} \tilde{C}_{\mu\nu}(t, \theta, \phi) \right) = 0 \tag{75}$$

$$\partial^i \left(\frac{1}{r^2} \tilde{C}_{i\mu}(t, \theta, \phi) \right) = 0, \tag{76}$$

where $\tilde{C}_{ab} = C_{ab} - \frac{1}{2} C \eta_{ab}$. This system of equations for $\frac{1}{r^2} C_{\mu\nu}$ is the same system of equations as is satisfied by stationary solutions of the flat spacetime linearized Einstein equation (but our $C_{\mu\nu}$ may depend on time). The general solution of these equations is $\tilde{C}_{ij} = 0$, $\tilde{C}_{i0} = F(t) n_i + S_{ij}(t) n^j$, and $\tilde{C}_{00} = 4 P_i(t) n^i$, where S_{ij} is antisymmetric, $S_{ij} = -S_{ji}$, where F, S_{ij} and P_i have no spatial dependence, and where n^i was defined by equation (72). By a further second-order gauge transformation (of the form $\xi_\mu = \delta_{\mu 0} F(t)/r$), we can set $F(t) = 0$. We thus obtain

$$C_{00}(t, \theta, \phi) = 2 P_i(t) n^i(\theta, \phi) \tag{77}$$

$$C_{i0}(t, \theta, \phi) = S_{ij}(t) n^j(\theta, \phi) \tag{78}$$

$$C_{ij}(t, \theta, \phi) = 2 \delta_{ij} P_k(t) n^k(\theta, \phi), \tag{79}$$

which is of the same form as the general stationary $l = 1$ perturbation of Minkowski spacetime (see, e.g., [29]), except that time dependence is allowed for S_{ij} and P_i. As we shall see shortly, S_{ij} and P_i correspond, respectively, to the spin and mass dipole moment of the body.

We now may write for the metric through $O(\lambda^2)$,

$$\begin{aligned} g_{\alpha\beta}(\lambda; t, x^i) = {} & \eta_{\alpha\beta} + B_{\alpha i \beta j}(t) x^i x^j + O(r^3) \\ & + \lambda \left(\frac{2M}{r} \delta_{\alpha\beta} + h^{\text{tail}}_{\alpha\beta}(t, 0) + h^{\text{tail}}_{\alpha\beta i}(t, 0) x^i + M \mathcal{R}_{\alpha\beta}(t) + O(r^2) \right) \\ & + \lambda^2 \left(\frac{M^2}{r^2} (-2 t_\alpha t_\beta + 3 n_\alpha n_\beta) + \frac{2}{r^2} P_i(t) n^i \delta_{\alpha\beta} + \frac{1}{r^2} t_{(\alpha} S_{\beta) j}(t) n^j \right. \\ & \left. + \frac{1}{r} K_{\alpha\beta}(t, \theta, \phi) + H_{\alpha\beta}(t, \theta, \phi) + O(r) \right) + O(\lambda^3), \end{aligned} \tag{80}$$

where we have introduced the unknown tensors K and H, and $S_{\alpha\beta}$ is the antisymmetric tensor whose spatial components are S_{ij} and whose time components vanish, i.e.,

$$S_{0i} = 0. \tag{81}$$

We now follow the strategy outlined in section 5. We consider a *smooth* coordinate shift of the form (52),

$$\hat{x}^\mu = x^\mu - \lambda A^\mu(x^\nu) + O(\lambda^2), \tag{82}$$

and choose A^μ so as to make the mass dipole moment of $\bar{g}_{\hat{\alpha}\hat{\beta}}(\lambda, t_0)$ vanish for all t_0. A straightforward application of the coordinate transformation (82) to the metric (80) yields

$$\begin{aligned} g_{\hat{\alpha}\hat{\beta}} = {} & \eta_{\alpha\beta} + B_{\alpha i \beta j}(\hat{t}) \hat{x}^i \hat{x}^j + O(r^3) \\ & + \lambda \left(\frac{2M}{\hat{r}} \delta_{\alpha\beta} + h^{\text{tail}}_{\alpha\beta}(\hat{t}, 0) + h^{\text{tail}}_{\alpha\beta i}(\hat{t}, 0) \hat{x}^i + M \mathcal{R}_{\alpha\beta}(\hat{t}, \hat{x}^i) + 2 A_{\alpha,\beta}(\hat{t}, \hat{x}^i) \right. \\ & \left. + 2 B_{\alpha i \beta j}(\hat{t}) \hat{x}^i A^j(\hat{t}, \hat{x}^i) + O(r^2) \right) + \lambda^2 \left(\frac{M^2}{\hat{r}^2} (-2 t_\alpha t_\beta + 3 n_\alpha n_\beta) \right. \\ & + \frac{2}{\hat{r}^2} [P_i(\hat{t}) - M A_i(\hat{t}, 0)] n^i \delta_{\alpha\beta} + \frac{1}{\hat{r}^2} t_{(\alpha} S_{\beta) j}(\hat{t}) n^j \\ & \left. + \frac{1}{\hat{r}} K_{\alpha\beta}(\hat{t}, \theta, \phi) + H_{\alpha\beta}(\hat{t}, \theta, \phi) + O(r) \right) + O(\lambda^3), \end{aligned} \tag{83}$$

where we have 'absorbed' the effects of the gauge transformation at orders $\lambda^2 r^{-1}$ and $\lambda^2 r^0$ into the unknown tensors H, K. The corresponding 'near-zone expansion' (see equations (24), (27) and (28)) is

$$
\begin{aligned}
\bar{g}_{\hat{\bar{\alpha}}\hat{\bar{\beta}}}(\hat{t}_0) = {} & \eta_{\alpha\beta} + \frac{2M}{\hat{\bar{r}}}\delta_{\alpha\beta} + \frac{M^2}{\hat{\bar{r}}^2}(-2t_\alpha t_\beta + 3n_\alpha n_\beta) + \frac{1}{\hat{\bar{r}}^2}t_{(\alpha}S_{\beta)j}n^j + \frac{2}{\hat{\bar{r}}^2}[P_i - MA_i]n^i\delta_{\alpha\beta} \\
& + O\left(\frac{1}{\hat{\bar{r}}^3}\right) + \lambda\left[h^{\text{tail}}_{\alpha\beta} + 2A_{(\alpha,\beta)} + \frac{1}{\hat{\bar{r}}}K_{\alpha\beta} + \frac{\hat{\bar{t}}}{\hat{\bar{r}}^2}(t_{(\alpha}S_{\beta)j,0}n^j\right. \\
& \left. + 2[P_{i,0} - MA_{i,0}]n^i\delta_{\alpha\beta}) + O\left(\frac{1}{\hat{\bar{r}}^2}\right) + \hat{\bar{t}}\,O\left(\frac{1}{\hat{\bar{r}}^3}\right)\right] + \lambda^2\left[B_{\alpha i\beta j}\hat{\bar{x}}^i\hat{\bar{x}}^j + h^{\text{tail}}_{\alpha\beta\gamma}\hat{\bar{x}}^\gamma\right. \\
& + M\mathcal{R}_{\alpha\beta}(\hat{\bar{x}}^i) + 2B_{\alpha i\beta j}A^i\hat{\bar{x}}^j + 2A_{(\alpha,\beta)\gamma}\hat{\bar{x}}^\gamma + H_{\alpha\beta} + \frac{\hat{\bar{t}}}{\hat{\bar{r}}}K_{\alpha\beta,0} + \frac{\hat{\bar{t}}^2}{\hat{\bar{r}}^2}(t_{(\alpha}S_{\beta)j,00}n^j \\
& \left. + 2[P_{i,00} - MA_{i,00}]n^i\delta_{\alpha\beta}) + O\left(\frac{1}{\hat{\bar{r}}}\right) + \hat{\bar{t}}\,O\left(\frac{1}{\hat{\bar{r}}^3}\right) + \hat{\bar{t}}^2 O\left(\frac{1}{\hat{\bar{r}}^3}\right)\right] + O(\lambda^3).
\end{aligned}
\tag{84}
$$

Note that the indices on the left-hand side of this equation have both a 'hat' and 'bar' on them to denote that they are components of $\bar{g}_{ab}$ in the scaled coordinates associated with our new coordinates $\hat{x}^\mu$. By contrast, the indices on the right-hand side have neither a 'hat' nor a 'bar', since they denote the corresponding components in the unscaled, original coordinates x^μ. Thus, for example, $A_{\alpha,\beta}$ denotes the matrix of first partial derivatives of the x^μ-components of A_a with respect to the x^μ coordinates[14]. It also should be understood that all tensor components appearing on the right-hand side of equation (84) are evaluated at time $\hat{t}_0$, and that A_α and its derivatives, as well as $h^{\text{tail}}_{\alpha\beta}$ and $h^{\text{tail}}_{\alpha\beta\gamma}$, are evaluated at $\hat{x}^i = 0$ (i.e., on the worldline γ). Finally, the 'reversals' in the roles of various terms in going from the far-zone expansion of the metric equation (80) to the near-zone expansion equation (84) should be noted. For example, the spin term $\frac{1}{r^2}t_{(\alpha}S_{\beta)j}n^j$ originated as a second-order perturbation in the far zone, but it now appears as part of the background-scaled metric in the near-zone expansion. By contrast, the term $B_{\alpha i\beta j}x^i x^j$ originated as part of the background metric in the far zone, but it now appears as a second-order perturbation in the near-zone expansion.

It is easy to see from equation (84) that $P^i - MA^i$ is the mass dipole moment of $\bar{g}_{\hat{\bar{\alpha}}\hat{\bar{\beta}}}$ at time t_0. We therefore set

$$
A^i(t) = P^i(t)/M \tag{85}
$$

for all t. Consequently, no mass dipole term will appear in our expressions below.

Although we have 'solved' for A^i in equation (85), we have not learned anything useful about the motion[15]. All useful information about A^i will come from demanding that the metrics $g_{ab}(\lambda)$—or, equivalently, $\bar{g}_{ab}(\lambda)$—be solutions of Einstein's equation. We may apply Einstein's equation perturbatively either via the far-zone expansion or the near-zone expansion. The resulting systems of equations are entirely equivalent, but the terms are organized very

[14] Note that the term $A_{(\alpha,\beta)\gamma}$ arises from Taylor expanding $A_{(\alpha,\beta)}$ with respect to the $\hat{x}^\mu$ coordinates, so, in principle, the second partial derivative in this expression should be with respect to $\hat{x}^\gamma$ rather than x^γ. However, since $\hat{x}^\gamma$ coincides with x^γ at zeroth order in λ and the $A_{(\alpha,\beta)\gamma}$ appears at second order in λ, we may replace the partial derivative with respect to $\hat{x}^\gamma$ by the partial derivative with respect to x^γ.

[15] However, equation (85) indicates clearly that solving for the displacement to center-of-mass coordinates A^i is equivalent to simply determining the mass dipole moment P^i in the original coordinates. The computations of this section may therefore be recast as simply solving enough of the second-order perturbation equations for the mass dipole moment—and hence the motion—to be determined.

differently. We find it more convenient to work with the near-zone expansion, and will do so below. We emphasize, however, that we could equally well have used the far-zone perturbation expansion. We also emphasize that no new information whatsoever can be generated by matching the near- and far-zone expansions, since these expansions have already been fully 'matched' via equations (24), (27) and (28).

In the following, in order to make the notation less cumbersome, we will drop the 'hat' on the near-zone coordinates $\hat{\bar{x}}^\mu$ and on the components $\bar{g}_{\hat{\bar{\alpha}}\hat{\bar{\beta}}}$. No confusion should arise from this, since we will not have occasion to use the original, scaled coordinates $\bar{x}^\mu$ below. On the other hand, we will maintain the 'hat' on the coordinates $\hat{x}^\mu$, since we will have occasion to use both $\hat{x}^\mu$ and x^μ below. Using this notation and setting the mass dipole terms to zero, equation (84) becomes

$$\begin{aligned}\bar{g}_{\bar{\alpha}\bar{\beta}}(\hat{t}_0) = {} & \eta_{\alpha\beta} + \frac{2M}{\bar{r}}\delta_{\alpha\beta} + \frac{M^2}{\bar{r}^2}(-2t_\alpha t_\beta + 3n_\alpha n_\beta) + \frac{1}{\bar{r}^2}t_{(\alpha}S_{\beta)j}n^j + O\left(\frac{1}{\bar{r}^3}\right) \\ & + \lambda\left[h^{\text{tail}}_{\alpha\beta} + 2A_{(\alpha,\beta)} + \frac{1}{\bar{r}}K_{\alpha\beta} + \frac{\bar{t}}{\bar{r}^2}t_{(\alpha}\dot{S}_{\beta)j}n^j + O\left(\frac{1}{\bar{r}^2}\right) + \bar{t}\,O\left(\frac{1}{\bar{r}^3}\right)\right] \\ & + \lambda^2\Bigg[B_{\alpha i\beta j}\bar{x}^i\bar{x}^j + h^{\text{tail}}_{\alpha\beta\gamma}\bar{x}^\gamma + M\mathcal{R}_{\alpha\beta}(\bar{x}^i) + 2B_{\alpha i\beta j}A^i\bar{x}^j + 2A_{(\alpha,\beta)\gamma}\bar{x}^\gamma \\ & + H_{\alpha\beta} + \frac{\bar{t}}{\bar{r}}\dot{K}_{\alpha\beta} + \frac{\bar{t}^2}{\bar{r}^2}t_{(\alpha}\ddot{S}_{\beta)j}n^j + O\left(\frac{1}{\bar{r}}\right) + \bar{t}\,O\left(\frac{1}{\bar{r}^2}\right) + \bar{t}^2 O\left(\frac{1}{\bar{r}^3}\right)\Bigg] + O(\lambda^3),\end{aligned} \tag{86}$$

where the 'dots' denote derivatives with respect to t.

We now apply the vacuum linearized Einstein equation—up to leading order, $1/\bar{r}^3$, in $1/\bar{r}$ as $\bar{r} \to \infty$—to the first-order term in λ appearing in equation (86), namely

$$\bar{g}^{(1)}_{\bar{\alpha}\bar{\beta}} = h^{\text{tail}}_{\alpha\beta} + 2A_{(\alpha,\beta)} + \frac{1}{\bar{r}}K_{\alpha\beta}(\theta,\phi) + \frac{\bar{t}}{\bar{r}^2}t_{(\alpha}\dot{S}_{\beta)j}n^j + O\left(\frac{1}{\bar{r}^2}\right) + \bar{t}\,O\left(\frac{1}{\bar{r}^3}\right). \tag{87}$$

It is clear that the terms of order $1/\bar{r}^2$ and $\bar{t}/\bar{r}^3$ in $\bar{g}^{(1)}_{\bar{\alpha}\bar{\beta}}$ cannot contribute to the linearized Ricci tensor to order $1/\bar{r}^3$. Similarly, it is clear that the terms of order $1/\bar{r}^2$ and higher in the background-scaled metric cannot contribute to the linearized Ricci tensor to order $1/\bar{r}^3$, so, to order $1/\bar{r}^3$, we see that $\bar{g}^{(1)}_{\bar{\alpha}\bar{\beta}}$ satisfies the linearized Einstein equation about the Schwarzschild metric. It is therefore useful to expand $\bar{g}^{(1)}_{\bar{\alpha}\bar{\beta}}$ in tensor spherical harmonics.

We obtain one very useful piece of information by extracting the $\ell = 1$, magnetic parity part of the linearized Ricci tensor that is even under time reversal. On account of the symmetries of the background Schwarzschild metric, only the $\ell = 1$, magnetic parity, even under time reversal part of the metric perturbation can contribute. Now, a general $\ell = 1$, symmetric (but not necessarily trace free) tensor field $Q_{\alpha\beta}(t, r, \theta, \phi)$ can be expanded in tensor spherical harmonics as (see, e.g., [30] or [31] equations (A.16)–(A.18))

$$\begin{aligned} Q_{00} &= Q^A_i n^i \\ Q_{i0} &= Q^B_j n^j n_i + Q^C_i + Q^M_k \epsilon_{ij}{}^k n^j \\ Q_{ij} &= Q^D_k n^k n_i n_j + Q^E_{(i} n_{j)} + Q^F_k \delta_{ij} n^k + Q^N_k \epsilon^k{}_{l(i} n_{j)} n^l, \end{aligned} \tag{88}$$

where the expansion coefficients $Q^A_i, Q^B_i, Q^C_i, Q^D_i, Q^E_i, Q^F_i, Q^M_i, Q^N_i$ are functions of (t, r). The 3-vector index on these coefficients corresponds to the three different 'm-values' for each $\ell = 1$ harmonic. Thus, we see that there are a grand total of eight types of $\ell = 1$ tensor spherical harmonics. The six harmonics associated with labeling indices $A - F$ are of electric parity, whereas the two harmonics associated with M, N are of magnetic parity.

For the metric perturbation (87), the 'constant tensors' $h^{\rm tail}_{\alpha\beta}$ and $A_{(\alpha,\beta)}$ are purely electric parity and cannot contribute. It turns out that $\frac{1}{\bar{r}}K_{\alpha\beta}(\theta,\phi)$ also does not contribute to the $\ell = 1$, magnetic parity part of the linearized Ricci tensor that is even under time reversal: since $K_{\alpha\beta}$ is independent of $\bar{t}$ the 'M' part of $K_{\alpha\beta}$ is odd under time reversal, whereas the 'N' part of $\frac{1}{\bar{r}}K_{\alpha\beta}(\theta,\phi)$ is pure gauge. Thus, the only term that contributes to order $1/\bar{r}^3$ to the $\ell = 1$, magnetic parity part of the linearized Ricci tensor that is even under time reversal is $\frac{\bar{t}}{\bar{r}^2}t_{(\alpha}\dot{S}_{\beta)j}n^j$. The satisfaction of the vacuum linearized Einstein equation at order $1/\bar{r}^3$ requires that this term vanish. We thereby learn that

$$\frac{\mathrm{d}S_{ij}}{\mathrm{d}t} = 0, \tag{89}$$

i.e., to lowest order, the spin is parallelly propagated with respect to the background metric along the worldline γ.

Having set the spin term to zero in equation (87), we may now substitute the remaining terms in equation (87) into the linearized Einstein equation and set the $1/\bar{r}^3$ part equal to zero. It is clear that we will thereby obtain relations between $h^{\rm tail}_{\alpha\beta}$, $A_{(\alpha,\beta)}$, and $K_{\alpha\beta}$. However, these relations will not be of direct interest for obtaining 'equations of motion'—i.e., equations relating A^i and its time derivatives to known quantities—because the quantity of interest $A_{i,0}$ always appears in combination with the quantity $A_{0,i}$, which is unrelated to the motion. Therefore, we shall not explicitly compute the relations arising from the linearized Einstein equation here.

We now consider the information on A^i that can be obtained from the near-zone second-order Einstein equation

$$G^{(1)}_{ab}[\bar{g}^{(2)}] = -G^{(2)}_{ab}[\bar{g}^{(1)}, \bar{g}^{(1)}], \tag{90}$$

where, from equation (86), we see that

$$\begin{aligned}\bar{g}^{(2)}_{\bar{\alpha}\bar{\beta}} = {} & B_{\alpha i\beta j}\bar{x}^i\bar{x}^j + D_{\alpha\beta\gamma}\bar{x}^\gamma + M\mathcal{R}_{\alpha\beta}(\bar{x}^\mu) + H_{\alpha\beta}(\theta,\phi) + \frac{\bar{t}}{\bar{r}}\dot{K}_{\alpha\beta}(\theta,\phi) \\ & + O\left(\frac{1}{\bar{r}}\right) + \bar{t}\,O\left(\frac{1}{\bar{r}^2}\right) + \bar{t}^2 O\left(\frac{1}{\bar{r}^3}\right),\end{aligned} \tag{91}$$

where we have defined

$$D_{\alpha\beta 0} \equiv h^{\rm tail}_{\alpha\beta 0} + 2A_{(\alpha,\beta)0} \tag{92}$$

$$D_{\alpha\beta i} \equiv h^{\rm tail}_{\alpha\beta i} + 2A_{(\alpha,\beta)i} + 2B_{\alpha i\beta j}A^j. \tag{93}$$

We wish to impose the second-order Einstein equation to orders $1/\bar{r}^2$ and $\bar{t}/\bar{r}^3$, which, as we shall see below, are the lowest nontrivial orders in $1/\bar{r}$ as $\bar{r} \to \infty$ that occur. First, we consider $G^{(2)}_{ab}[\bar{g}^{(1)}, \bar{g}^{(1)}]$. The terms appearing in this quantity can be organized into terms of the following general forms (i) $\bar{g}^{(1)}\partial\partial\bar{g}^{(1)}$; (ii) $\partial\bar{g}^{(1)}\partial\bar{g}^{(1)}$; (iii) $\Gamma\bar{g}^{(1)}\partial\bar{g}^{(1)}$ where Γ denotes a Christoffel symbol of the background-scaled metric; (iv) $\Gamma\Gamma\bar{g}^{(1)}\bar{g}^{(1)}$ and (v) $(\partial\Gamma)\bar{g}^{(1)}\bar{g}^{(1)}$. From the form of $\bar{g}^{(1)}$ together with the fact that $\Gamma = O(1/\bar{r}^2)$ and $\partial\Gamma = O(1/\bar{r}^3)$, it is clear that none of these terms can contribute to $G^{(2)}_{ab}[\bar{g}^{(1)}, \bar{g}^{(1)}]$ to order $1/\bar{r}^2$ or $\bar{t}/\bar{r}^3$. Therefore, we may treat $\bar{g}^{(2)}$ as satisfying the homogeneous, vacuum linearized Einstein equation.

We now consider the linearized Ricci tensor of the perturbation $\bar{g}^{(2)}$. By inspection of equation (91), it might appear that terms that are $O(1)$ (from two partial derivatives acting on the 'B' term) and $O(1/\bar{r})$ (from various terms) will arise. However, it is not difficult to show that the total contribution to the $O(1)$ and $O(1/\bar{r})$ terms will vanish by virtue of the fact that the metric $g_{ab}(\lambda = 0)$ is a solution to Einstein's equation and the term proportional to

λ in equation (80) satisfies the far-zone linearized Einstein equation (which has already been imposed). It also is clear that there is no contribution of $\bar{g}^{(2)}$ to the linearized Ricci tensor that is of order $\bar{t}/\bar{r}^2$. Thus, the lowest nontrivial orders that arise in the second-order Einstein equation are indeed $1/\bar{r}^2$ and $\bar{t}/\bar{r}^3$, as claimed above.

The computation of the linearized Ricci tensor to orders $1/\bar{r}^2$ and $\bar{t}/\bar{r}^3$ for the metric perturbation $\bar{g}^{(2)}$ is quite complicated, so we will save considerable labor by focusing on the relevant part of the linearized Einstein equation to these orders. Our hope/expectation (which will be borne out by our calculation) is to obtain an equation for $A^i{}_{,00}$. Since this quantity is of $\ell = 1$, electric parity type and is even under time reversal, we shall focus on the $\ell = 1$, electric parity, even under time reversal part of the linearized Ricci tensor of $\bar{g}^{(2)}$ at orders $1/\bar{r}^2$ and $\bar{t}/\bar{r}^3$. From equation (88), we see that the $\ell = 1$ electric parity part of the Ricci tensor that is $O(1/\bar{r}^2)$ and even under time reversal can be written as

$$R^{(1)}_{00}\big|_{\ell=1,+,\frac{1}{\bar{r}^2}} = \frac{1}{\bar{r}^2} R^A_i n^i \tag{94}$$

$$R^{(1)}_{ij}\big|_{\ell=1,+,\frac{1}{\bar{r}^2}} = \frac{1}{\bar{r}^2}\left(R^D_k n^k n_i n_j + R^E_{(i} n_{j)} + R^F_k n^k \delta_{ij}\right), \tag{95}$$

whereas the $\ell = 1$ part of the Ricci tensor that is $O(\bar{t}/\bar{r}^3)$ and even under time reversal can be written as

$$R^{(1)}_{i0}\big|_{\ell=1,+,\frac{\bar{t}}{\bar{r}^3}} = \frac{\bar{t}}{\bar{r}^3}\left(R^B_j n^j n_i + R^C_i\right). \tag{96}$$

Here, in contrast to the usage of (88), R^A_i, R^B_i, R^C_i, R^D_i, R^E_i, R^F_i are 'constants', i.e., they have no dependence on $(\bar{t}, \bar{r})$.

We now consider the terms in $\bar{g}^{(2)}$ that can contribute to these Ricci terms. The term $B_{\alpha i\beta j}\bar{x}^i\bar{x}^j$ has no $\ell = 1$ part. Nevertheless, the $\ell = 2$ magnetic parity part of this term can, in effect, combine with the $\ell = 1$ magnetic parity 'spin term' $\frac{1}{\bar{r}^2} t_{(\alpha} S_{\beta) j} n^j$ in the background-scaled metric to produce a contribution to the linearized Ricci tensor of the correct type. This contribution will be proportional to

$$F_i \equiv S^{kl} R_{kl0i}. \tag{97}$$

For the remaining terms in $\bar{g}^{(2)}$, the 'spin term' $\frac{1}{\bar{r}^2} t_{(\alpha} S_{\beta) j} n^j$ in the background-scaled metric will not contribute to the relevant parts of the linearized Ricci tensor, so we may treat the remaining terms in $\bar{g}^{(2)}$ as though they were perturbations of Schwarzschild. Thus, only the $\ell = 1$, electric parity, even under time reversal part of these terms can contribute. The remaining contributors to R^A_i, R^B_i, R^C_i, R^D_i, R^E_i and R^F_i are

$$\begin{aligned} D_{00k}\bar{x}^k &= \bar{r} D^A_i n^i \\ D_{i00}\bar{t} &= \bar{t} D^C_i \\ D_{ijk}\bar{x}^k|_{\ell=1,+} &= \bar{r}\left(n_{(i} D^E_{j)} + \delta_{ij} n^k D^F_k\right), \end{aligned} \tag{98}$$

and

$$H_{00}|_{\ell=1,+} = H^A_i n^i \tag{99}$$

$$\dot{K}_{i0}|_{\ell=1,+} = \dot{K}^B_j n^j n_i + \dot{K}^C_i \tag{100}$$

$$H_{ij}|_{\ell=1,+} = H^D_k n^k n_i n_j + H^E_{(i} n_{j)} + H^F_k n^k \delta_{ij}, \tag{101}$$

where, in equation (98), we have

$$D_i^A = D_{00i} \tag{102}$$

$$D_i^C = D_{i00} \tag{103}$$

$$D_i^E = \tfrac{1}{5}\big(3D_{ik}^k - D_{ki}^k\big) \tag{104}$$

$$D_i^F = \tfrac{1}{5}\big(-D_{i\,k}^{\,k} + 2D^k{}_{ki}\big). \tag{105}$$

(The curvature term $\mathcal{R}_{\alpha\beta}$ has not appeared in the above equations because it has no $\ell = 1$ part.) The D_i^A, D_i^C, D_i^E, D_i^F, H_i^A, H_i^D, H_i^E, H_i^F, $\dot{K}_i^B$, $\dot{K}_i^C$ are also 'constants' in these expressions. A lengthy calculation now yields

$$\begin{pmatrix} R_i^A \\ R_i^B \\ R_i^C \\ R_i^D \\ R_i^E \\ R_i^F \end{pmatrix} = -\frac{1}{2}\begin{pmatrix} \frac{-16}{5} & -3M & 0 & -M & -M & -2 & 0 & 0 & 0 & -2 & 2 \\ 0 & 0 & -6M & 0 & 0 & 0 & 0 & 0 & 0 & -3 & -3 \\ 0 & 0 & 2M & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ \frac{-6}{5} & -6M & 0 & 0 & -6M & -3 & 3 & 0 & 3 & -6 & 0 \\ \frac{-16}{5} & 0 & 4M & 0 & 4M & 2 & -2 & 0 & -2 & 2 & -2 \\ 2 & 0 & 3M & -3M & 9M & 1 & -3 & 0 & -3 & 2 & 0 \end{pmatrix}\begin{pmatrix} F_i \\ D_i^A \\ D_i^C \\ D_i^E \\ D_i^F \\ H_i^A \\ H_i^D \\ H_i^E \\ H_i^F \\ \dot{K}_i^B \\ \dot{K}_i^C \end{pmatrix}. \tag{106}$$

Using the vacuum linearized Einstein equation $R_{ab}^{(1)} = 0$, we thus obtain six linear equations for our 11 unknowns. However, in order to find 'universal' behavior, we are interested in relations that do not involve $H_{\alpha\beta}$ and $\dot{K}_{\alpha\beta}$. It can be shown that there are two such relations[16], namely

$$-4F_i - 3MD_i^A + 2MD_i^C - 2MD_i^E + 4MD_i^F = 0 \tag{107}$$

and

$$-F_i - MD_i^A + 2MD_i^C = 0. \tag{108}$$

The first equation involves A^0 and spatial derivatives of A^i, and does not yield restrictions on the motion. However, the second equation gives desired equations of motion. Plugging in the definitions of the quantities appearing in equation (108), we obtain

$$-S^{kl}R_{kl0i} - M\big(h_{00,i}^{\mathrm{tail}} + 2R_{0j0i}A^j + 2A_{0,0i}\big) + 2M\big(h_{i0,0}^{\mathrm{tail}} + A_{i,00} + A_{0,i0}\big) = 0, \tag{109}$$

[16] There will be three such relations in total, because the vanishing of the mass dipole moment for all time implies through $O(\lambda^2)$ in near-zone perturbation theory the vanishing of the value, time derivative and second time derivative of the mass dipole at time t_0. The third condition should follow from the first-order near-zone Einstein equation, which we did not fully use. In fact, it should only be necessary to impose that the mass dipole have no second time derivative in order to define the motion.

where we have taken advantage of the fact (noted above) that for 00 and $0i$ components we have $h^{\rm tail}_{\alpha\beta\gamma} = h^{\rm tail}_{\alpha\beta,\gamma}$. Using the equality of mixed partials $A_{0,i0} = A_{0,0i}$, we obtain

$$A_{i,00} = \frac{1}{2M} S^{kl} R_{kl0i} - R_{0j0i} A^j - \left(h^{\rm tail}_{i0,0} - \frac{1}{2} h^{\rm tail}_{00,i} \right). \tag{110}$$

Thus, according to the interpretation provided in section 5 above, the first-order perturbative correction, $Z^i(t)$, to the geodesic γ of the background spacetime satisfies

$$\frac{d^2 Z^i}{dt^2} = \frac{1}{2M} S^{kl} R_{kl0}{}^i - R_{0j0}{}^i Z^j - \left(h^{{\rm tail}\,i}{}_{0,0} - \frac{1}{2} h^{\rm tail}{}_{00}{}^{,i} \right). \tag{111}$$

In addition, we have previously found that M and S_{ij} are constant along γ. Taking account of the fact that this equation is written in Fermi normal coordinates of γ and that $Z^0 = 0$, we can rewrite this equation in a more manifestly covariant looking form as

$$u^c \nabla_c (u^b \nabla_b Z^a) = \frac{1}{2M} R_{bcd}{}^a S^{bc} u^d - R_{bcd}{}^a u^b Z^c u^d - (g^{ab} + u^a u^b) \left(\nabla_d h^{\rm tail}_{bc} - \frac{1}{2} \nabla_b h^{\rm tail}_{cd} \right) u^c u^d, \tag{112}$$

where u^a is the tangent to γ and $u^c \nabla_c S_{ab} = 0$. However, it should be emphasized that this equation describes the perturbed motion only when the metric perturbation is in the Lorenz gauge (see the appendix).

The first term in equation (111) (or, equivalently, in equation (112)) is the 'spin force' first obtained by Papapetrou [18]. Contributions from higher multipole moments do not appear in our equation because they scale to zero faster than the spin dipole moment, and thus would arise at higher order in λ in our perturbation scheme. The second term corresponds to the right-hand side of the geodesic deviation equation, and appears because the perturbed worldline is not (except at special points) coincident with the background worldline[17]. The final term is the 'gravitational self-force', which is seen to take the form of a (regularized) gravitational force from the particle's own field. Our derivation has thus provided a rigorous justification of the regularization schemes that have been proposed elsewhere.

Finally, we note that, although our analysis has many points of contact with previous analyses using matched asymptotic expansions, there are a number of significant differences. We have already noted in section 4 that our derivation of geodesic motion at zeroth order in λ appears to differ from some other derivations [5, 8, 12], which do not appear to impose the requirement that $M \neq 0$. We also have already noted that in other approaches to self-force [5, 8], what corresponds to our scaled metric at $\lambda = 0$ is *assumed* to be of Schwarzschild form. In these other approaches, first-order perturbations in the near-zone expansion are treated as time independent, and are required to be regular on the Schwarzschild horizon. By contrast, we make no assumptions about the time dependence of the perturbations of the scaled metric beyond those that follow from our fundamental assumptions (i)–(iii) of section 2. Thus, our first-order perturbations are allowed to have linear dependence on $\bar{t}$, and our second-order perturbations can depend quadratically on $\bar{t}$. We also make no assumptions about the spacetime at $\bar{r} < \bar{R}$ and therefore impose no boundary conditions at small $\bar{r}$. Finally, there is a significant difference in the manner in which the gauge conditions used to define the motion are imposed. In [5, 8], the entire $\ell = 1$ electric parity part of what corresponds to our second-order near-zone perturbation is set to zero without proper justification[18]. By contrast, our 'no

[17] Consider a one-parameter family wherein the initial position for a body is 'moved over' smoothly with λ. In the limit $M \to 0$, the body then moves on a family of geodesics of the background metric parametrized by λ, and the perturbative description of motion should indeed be the geodesic deviation equation.

[18] Note that the part of the $\ell = 1$ electric parity perturbation that is relevant for obtaining equations of motion in [5, 8] is of 'acceleration type' (with linear growth in $\bar{r}$) and does not have an obvious interpretation in terms of a shift in the center of mass.

mass dipole' condition applies to the *background* near-zone metric and has been justified as providing 'center-of-mass' coordinates.

7. Beyond perturbation theory

As already mentioned near the beginning of section 5, the quantity Z^i in equation (111) is a 'deviation vector' defined on the background geodesic γ that describes the first order in λ perturbation to the motion. For any one-parameter family of spacetimes $g_{ab}(\lambda)$ satisfying the assumptions stated in section 2, equation (111) is therefore guaranteed to give a good approximation to the deviation from the background geodesic motion γ as $\lambda \to 0$. In other words, if γ is described by $x^i(t) = 0$, then the new worldline obtained defined by $x^i(t) = \lambda Z^i(t)$ is the correct description of motion to first order in λ (when the metric perturbation is in Lorenz gauge) and is therefore guaranteed to be accurate at small λ. However, this guarantee is of the form that if one wants to describe the motion accurately up to time t, then it always will be possible to choose λ sufficiently small that $x^i(t) = \lambda Z^i(t)$ is a good approximation up to time t. The guarantee is *not* of the form that if λ is chosen to be sufficiently small, then $x^i(t) = \lambda Z^i(t)$ will accurately describe the motion for all time. Indeed, for any fixed $\lambda > 0$, it is to be expected that $Z^i(t)$ will grow large at sufficiently late times, and it is clear that the approximate description of motion $x^i(t) = \lambda Z^i(t)$ cannot be expected to be good when $Z^i(t)$ is large, since by the time the motion has deviated significantly from the original background geodesic γ, the motion clearly cannot be accurately described in the framework of being a 'small correction' to γ. However, the main intended application of the first-order corrected equations of motion is to compute motion in cases, such as inspiral, where the deviations from the original geodesic motion become large at late times. It is therefore clear that equation (111), as it stands, is useless for computing long-term effects, such as inspiral.

One possible response to the above difficulty would be to go to higher order in perturbation theory. However, it seems clear that this will not help. Although the equations of motion obtained from nth-order perturbation theory will be more accurate than the first-order equations, they will not have a domain of validity that is significantly larger than the first-order equations. The perturbative description at any finite order will continue to treat the motion as a 'small deviation' from γ, and cannot be expected to describe motion accurately when the deviations are, in fact, large. In essence, by the time that the deviation from γ has become sufficiently large to invalidate first-order perturbation theory—so that, e.g., the second-order corrections are comparable in magnitude to the first-order corrections—then one would expect that the $(n + 1)$th-order corrections will also be comparable to the nth-order corrections, so nth-order perturbation theory will not be accurate either. Only by going to all orders in perturbation theory can one expect to get an accurate, global in time, description of motion via perturbation theory. Of course, if one goes to all orders in perturbation theory, then there is little point in having done perturbation theory at all.

Nevertheless, for a sufficiently small body of sufficiently small mass, it seems clear that the corrections to geodesic motion should be *locally* small and should be locally described by equation (111). By the time these small corrections have built up and the body has deviated significantly from the original geodesic approximating its motion, it should then be close to a *new* geodesic, perturbing off of which should give a better approximation to the motion for that portion of time. One could then attempt to 'patch together' such solutions to construct a world-line that accurately describes the motion of the particle for a longer time. In the limit of many such patches with small times between them, one expects the resulting worldline to be described by a single 'self-consistent' differential equation, which should then well approximate the motion as long as it remains *locally* close to geodesic motion.

A simple, familiar example will help illustrate all of the above points. Consider the cooling of a 'black body'. To choose a definite problem that can be put in a framework similar to that considered in this paper, let us consider a body (such as a lump of hot coal) that is put in a box with perfect reflecting walls, but a hole of area A is cut in this wall. We are interested in determining how the energy, E, of the body changes with time. At finite A, this is a very difficult problem, since the body will not remain in exact thermal equilibrium as it radiates energy through the hole. However, let us consider a one-parameter family of cavities where $A(\lambda)$ smoothly goes to zero as $\lambda \to 0$. When $\lambda = 0$, we find that the energy, $E_0 \equiv E(\lambda = 0)$, does not change with time, and the body will remain in thermal equilibrium at temperature T_0 for all time. When we do first-order perturbation theory in λ, we will find that the first order in λ correction, $E^{(1)}$, to the energy satisfies[19]

$$\frac{\mathrm{d}E^{(1)}}{\mathrm{d}t} = -\sigma A^{(1)} T_0^4, \tag{113}$$

where σ is the Stefan–Boltzmann constant and $A^{(1)} \equiv \mathrm{d}A/\mathrm{d}\lambda|_{\lambda=0}$. Note that only the zeroth-order temperature, T_0, enters the right-hand side of this equation because the quantity $A^{(1)}$ is already first order in λ, so the effect of any changes in temperature would appear only to higher order in λ. Since T_0 is a constant, it is easy to integrate equation (113) to obtain

$$E^{(1)}(t) = -\sigma A^{(1)} T_0^4 t. \tag{114}$$

Thus, first-order perturbation theory approximates the behavior of $E(\lambda, t)$ as

$$E(\lambda, t) = E_0 - \lambda \sigma A^{(1)} T_0^4 t. \tag{115}$$

Although this is a good approximation at early times, it is a horrible approximation at late times, as it predicts that the energy will go negative. If one went to second order in perturbation theory, one would obtain corrections to equation (113) that would take into account the first-order energy loss as well as various non-equilibrium effects. However, one would still be perturbing off of the non-radiating background, and the late time predictions using second (or any finite higher order) perturbation theory would still be very poor.

However, there is an obvious major improvement that can be obtained by noting that if A is sufficiently small, then the body should remain nearly in thermal equilibrium as it loses energy. Therefore, although perturbation theory off of the zeroth-order solution may give poor results at late times, first-order perturbation theory off of *some* thermal equilibrium solution should give locally accurate results at all times. This suggests that if A is sufficiently small, the cooling of the body should be described by

$$\frac{\mathrm{d}E}{\mathrm{d}t} = -\sigma A T^4(t). \tag{116}$$

When supplemented with the formula, $E = E(T)$, that relates energy to temperature when the body is in thermal equilibrium, this equation should provide an excellent description of the cooling of the body that is valid at all times. In effect, equation (116) takes into account the higher order perturbative effects (to all orders in λ) associated with the cooling of the body, but it neglects various perturbative effects associated with the body failing to remain in thermal equilibrium as it cools. Equation (116) is *not* an exact equation (since it does not take various non-equilibrium effects into account) and it is *not* an equation that arises directly from perturbation theory. Rather, it is an equation that corresponds to applying

[19] Of course, when A becomes small compared to the typical wavelengths of the radiation (as it must as we let $A \to 0$), we enter a physical optics regime where our formulae are no longer valid. We ignore such effects here, just as in our above analysis of the motion of bodies in general relativity we ignored quantum gravity effects even though they should be important when the size of the body is smaller than the Planck scale.

first-order perturbation theory to a background that itself undergoes changes resulting from the perturbation. We will refer to such an equation as a 'self-consistent perturbative equation'. Such equations are commonly written for systems that can be described *locally in time* by a small deviation from a simple solution.

How does one find a 'self-consistent perturbative equation' for a given system for which one has derived first-order perturbative equations? We do not believe that there is any general method for deriving a self-consistent perturbative equation. However, the following appear to be appropriate criteria to impose on a self-consistent perturbative equation: (1) it should have a well-posed initial-value formulation. (2) It should have the same number of degrees of freedom as the first-order perturbative system, so that a correspondence can be made between initial data for the self-consistent perturbative equation and the first-order perturbative system. (3) For corresponding initial data, the solutions to the self-consistent perturbative equation should be close to the corresponding solutions of the first-order perturbative system over the time interval for which the first-order perturbative description should be accurate. We do not know of any reason why, for any given system, there need exist a self-consistent perturbative equation satisfying these criteria. In cases where a self-consistent perturbative equation satisfying these criteria does exist, we would not expect it to be unique. For example, we could modify equation (116) by adding suitable terms proportional to A^2 to the right-hand side of this equation.

The first-order perturbative equations for the motion of a small body are that the first-order metric perturbation satisfies

$$\nabla^c\nabla_c\tilde{h}_{ab} - 2R^c{}_{ab}{}^d\tilde{h}_{cd} = -16\pi M u_a u_b \frac{\delta^{(3)}(x^i)}{\sqrt{-g}}\frac{\mathrm{d}\tau}{\mathrm{d}t}, \tag{117}$$

where $x^i = 0$ corresponds to a geodesic, γ of the background spacetime and u^a is the tangent to γ. If we consider the retarded solution to this equation (which automatically satisfies the Lorenz gauge condition), we have proven rigorously in this paper that the first order in λ deviation of the motion from γ satisfies

$$u^c\nabla_c(u^b\nabla_b Z^a) = -R_{bcd}{}^a u^b Z^c u^d - (g^{ab} + u^a u^b)\left(\nabla_d h^{\text{tail}}_{bc} - \tfrac{1}{2}\nabla_b h^{\text{tail}}_{cd}\right)u^c u^d, \tag{118}$$

with

$$h^{\text{tail}}_{ab}(x) = M\int_{-\infty}^{\tau^-_{\text{ret}}}\left(G^+_{aba'b'} - \frac{1}{2}g_{ab}G^{+\ c}_{c\ a'b'}\right)(x, z(\tau'))u^{a'}u^{b'}\,\mathrm{d}\tau', \tag{119}$$

where, for simplicity, we have dropped the spin term. The MiSaTaQuWa equations

$$\nabla^c\nabla_c\tilde{h}_{ab} - 2R^c{}_{ab}{}^d\tilde{h}_{cd} = -16\pi M u_a(t)u_b(t)\frac{\delta^{(3)}(x^i - z^i(t))}{\sqrt{-g}}\frac{\mathrm{d}\tau}{\mathrm{d}t}, \tag{120}$$

$$u^b\nabla_b u^a = -(g^{ab} + u^a u^b)\left(\nabla_d h^{\text{tail}}_{bc} - \frac{1}{2}\nabla_b h^{\text{tail}}_{cd}\right)u^c u^d, \tag{121}$$

$$h^{\text{tail}}_{ab}(x) = M\int_{-\infty}^{\tau^-_{\text{ret}}}\left(G^+_{aba'b'} - \frac{1}{2}g_{ab}G^{+\ c}_{c\ a'b'}\right)(x, z(\tau'))u^{a'}u^{b'}\,\mathrm{d}\tau', \tag{122}$$

(where one chooses the retarded solution to equation (120)) are an excellent candidate for self-consistent perturbative equations corresponding to the above first-order perturbative system[20]. Here, $u^a(\tau)$ (normalized in the background metric) refers to the self-consistent motion $z(\tau)$,

[20] The Riemann tensor term does not appear on the right-hand side of equation (121), since in the self-consistent perturbative equation, the deviation from the self-consistent worldline should vanish.

rather than to a background geodesic as before. Although a proper mathematical analysis of this integro-differential system has not been carried out, it appears plausible that our above criteria (1)–(3) will be satisfied by the MiSaTaQuWa equations. If so, they should provide a good, global in time, description of motion for problems like extreme mass ratio inspiral.

Acknowledgments

We wish to thank Abraham Harte and Eric Poisson for helpful discussions. This research was supported in part by NSF grant PHY04-56619 to the University of Chicago and a National Science Foundation Graduate Research Fellowship to SG.

Appendix. Self-force in an arbitrary allowed gauge

As discussed in section 5, the description of motion will change under first-order changes of gauge. Indeed, in that section, we noted that under a smooth gauge transformation, the description of motion changes by equation (56). However, as previously stated near the end of section 2 (see equation (21)), the allowed coordinate freedom includes transformations that are not smooth at $r = 0$. Since such gauges may arise in practice[21], we provide here the expression for the first-order perturbative equation of motion in an arbitrary gauge allowed by our assumptions. We also present the corresponding self-consistent perturbative equations of motion.

As previously noted in section 6 (see the remark below equation (74)), the equations of motion to first order in λ depend only upon the first-order gauge transformation ξ^a. As we have seen, the mass dipole moment appears at second order in (far-zone) perturbation theory, so we must consider the effects of first-order gauge transformations on second-order perturbations. This is given by $g^{(2)} \to g^{(2)} + \delta g^{(2)}$, with

$$\delta g^{(2)}_{ab} = (\mathcal{L}_\xi g^{(1)})_{ab} + \left(\mathcal{L}^2_\xi g^{(0)}\right)_{ab}, \tag{A.1}$$

where $\mathcal{L}$ denotes the Lie derivative. Equivalently, we have

$$\delta g^{(2)}_{ab} = \xi^c \nabla_c g^{(1)}_{ab} + 2\nabla^c \xi_{(a} g^{(1)}_{b)c} + \xi^c \nabla_c \nabla_{(a}\xi_{b)} + \nabla_c \xi_{(a} \nabla^c \xi_{b)} + \nabla_c \xi_{(a} \nabla_{b)} \xi^c, \tag{A.2}$$

where ∇_a is the derivative operator associated with the background metric $g_{ab}(\lambda = 0)$. In order to satisfy the criteria on allowed gauge transformations (see equation (21)), the components of ξ^a must be of the form

$$\xi^\mu = F^\mu(t, \theta, \phi) + O(r), \tag{A.3}$$

i.e., ξ^a cannot 'blow up' at $r = 0$ but it can be singular in the sense that its components can have direction-dependent limits.

The mass dipole moment, P^i, is one-half of the coefficient of the $\ell = 1$ part of the leading order, $1/r^2$, part of the second-order metric perturbation, $g^{(2)}_{00}$. Therefore, P^i may be extracted from the formula

$$P^i = \frac{3}{8\pi} \lim_{R\to 0} \int_{r=R} g^{(2)}_{00} n^i \, \mathrm{d}S, \tag{A.4}$$

[21] For example, the Regge–Wheeler gauge (used for perturbations of the Schwarzschild metric) is not smoothly related to the Lorenz gauge [19]. However, it is possible that the gauge vector is bounded [19], in which case perturbations in the Regge–Wheeler gauge would satisfy our assumptions (see equation (A.3)), and equations of motion could be defined. On the other hand, point particle perturbations expressed in radiation gauges (used for perturbations of the Kerr metric) contain a log singularity along a string [19], and therefore do not satisfy our assumptions.

where dS is the area element on the sphere of radius R. Under the gauge transformation generated by ξ^a, we have

$$\delta g^{(2)}_{00} = \xi^c \nabla_c g^{(1)}_{00} + 2\nabla^c \xi_0 g^{(1)}_{0c} + \xi^c \nabla_c \nabla_0 \xi_0 + \nabla_c \xi_0 \nabla^c \xi_0 + \nabla_c \xi_0 \nabla_0 \xi^c. \tag{A.5}$$

As previously noted, for an arbitrary first-order perturbation satisfying our assumptions, we have

$$g^{(1)}_{00} = -1 + \frac{2M}{r} + O(1), \tag{A.6}$$

where M is the mass of the body. From equations (A.3), (A.5) and (A.6), we see that the change in $g^{(2)}_{00}$ induced by our gauge transformation is

$$\delta g^{(2)}_{00} = -\frac{2M}{r^2}\xi^i n_i + \chi + O\left(\frac{1}{r}\right), \tag{A.7}$$

$$\chi \equiv 2\partial^i \xi_0 g^{(1)}_{0i} + \partial_i \xi_0 \partial^i \xi_0. \tag{A.8}$$

Therefore, by equation (A.4), the induced change in the mass dipole moment is

$$\delta P^i = \frac{3}{8\pi} \lim_{r\to 0} \int (-2M\xi^j n_j + r^2 \chi) n^i \, d\Omega, \tag{A.9}$$

where $d\Omega$ is the area element on the unit sphere.

Equation (A.9) gives the change in the mass dipole moment induced by the possibly non-smooth gauge transformation generated by ξ^a. The corresponding change in the first-order perturbative equation of motion is determined by the change in the *smooth* vector field A^a required to eliminate the mass dipole. Writing $A^a \to A^a + \delta A^a$, this change is given by

$$\delta A^i = \delta P^i / M \tag{A.10}$$

(see equation (85)). Thus, the change $Z^i \to \hat{Z}^i = Z^i + \delta Z^i$ induced in the deviation vector describing the perturbed worldline is

$$\delta Z^i = \frac{3}{8\pi} \lim_{r\to 0} \int (-2\xi^j n_j + M^{-1} r^2 \chi) n^i \, d\Omega, \tag{A.11}$$

with χ given by equation (A.8). In the case where our original gauge was the Lorenz gauge, it follows immediately from equation (111) that the new equation of motion for $\hat{Z}^i$ is

$$\frac{d^2 \hat{Z}^i}{dt^2} = -R_{0j0}{}^i Z^j - \left({h^{\text{tail}}}^i{}_{0,0} - \frac{1}{2} {h^{\text{tail}}}_{00}{}^{,i} \right) + \delta \ddot{Z}^i, \tag{A.12}$$

where δZ^i is given by equation (A.11), and where, for simplicity, we have dropped the spin term. We may rewrite equation (A.12) as

$$\frac{d^2 \hat{Z}^i}{dt^2} = -R_{0j0}{}^i \hat{Z}^j - \left({h^{\text{tail}}}^i{}_{0,0} - \frac{1}{2} h^{\text{tail},i}_{00} \right) + \delta \ddot{Z}^i + R_{0j0}{}^i \delta Z^j. \tag{A.13}$$

Note that although equation (A.13) provides us with the desired equation of motion in an arbitrary allowed gauge, the terms involving components of h^{tail} must still be computed in the Lorenz gauge.

Now suppose one wishes to pass to a self-consistent perturbative equation associated with the new choice of gauge. It is not obvious how one might wish to modify the evolution equations for the metric perturbations in the new gauge. (One possibility would be to simply use equation (120) and then modify the result by the addition of $2\nabla_{(a}\xi_{b)}$ but it might be preferable to find a new equation based on a suitable 'relaxed' version of the linearized Einstein equation for the new gauge.) However, it appears that a natural choice of self-consistent perturbative equation associated with equation (A.13) would be

$$u^b\nabla_b u^a = -(g^{ab} + u^a u^b)\left(\nabla_d h^{\text{tail}}_{bc} - \tfrac{1}{2}\nabla_b h^{\text{tail}}_{cd}\right) u^c u^d + \delta\ddot{Z}^a + R_{cbd}{}^a u^c u^d \delta Z^b. \tag{A.14}$$

In the case where ξ^a is smooth (so that, by equation (A.11), we have $\delta Z^i = -\xi^i$) this agrees with the proposal of Barack and Ori [19].

References

[1] Wald R M 1984 *General Relativity* (Chicago, IL: University of Chicago Press)
[2] Geroch R and Traschen J 1987 *Phys. Rev.* D **36** 1017
[3] Weinberg S 1972 *Gravitation and Cosmology* (New York: Wiley)
[4] Israel W 1966 *Nuovo Cimento* B **44** 1
[5] Mino Y, Sasaki M and Tanaka T 1997 *Phys. Rev.* D **55** 3457–76
[6] Quinn T C and Wald R M 1997 *Phys. Rev.* D **56** 3381–94
[7] Detweiler S and Whiting B F 2003 *Phys. Rev.* D **67** 024025
[8] Poisson E 2004 *Living Rev. Rel.* **7** 6
[9] DeWitt B S and Brehme R W 1960 *Ann. Phys., NY* **9** 220–59
[10] Dirac P A M 1938 *Proc. R. Soc. Lond.* A **167** 148
[11] Burke W L 1971 *J. Math. Phys., NY* **12** 401
[12] D'Eath P D 1975 *Phys. Rev.* D **11** 1387
[13] Kates R 1980 *Phys. Rev.* D **22** 1853
[14] Thorne K S and Hartle J B 1985 *Phys. Rev.* D **31** 1815
[15] Geroch R and Jang P S 1975 *J. Math. Phys.* **16** 65–7
[16] Ehlers J and Geroch R 2004 *Ann. Phys.* **309** 232
[17] Stuart D A M 2004 *J. Math. Pures Appl.* **83** 541
[18] Papapetrou A 1951 *Proc. R. Soc. Lond.* A **209** 248–58
[19] Barack L and Ori A 2001 *Phys. Rev.* D **64** 124003, 1–13
[20] Gralla S, Harte A and Wald R in preparation
[21] Geroch R 1969 *Commun. Math. Phys.* **13** 180
[22] Futamase T and Itoh Y 2007 *Living Rev. Rel.* **10** 2
[23] Wald R M 1978 *Phys. Rev. Lett.* **41** 203
[24] Beiglbock W 1967 *Commun. Math. Phys.* **5** 106
[25] Geroch R 1970 *J. Math. Phys.* **11** 2580
[26] Hansen R O 1974 *J. Math. Phys.* **15** 46
[27] Beig R and Simon W 1980 *Gen. Rel. Grav.* **12** 1003
[28] Kundu P 1981 *J. Math. Phys.* **22** 1236
[29] Zhang X H 1986 *Phys. Rev.* D **34** 991
[30] Thorne K S 1980 *Rev. Mod. Phys.* **52** 299
[31] Blanchet L and Damour T 1986 *Phil. Trans. R. Soc. Lond.* A **320** 379