

Proving the Principle: Taking Geodesic Dynamics too Seriously in Einstein's Theory_{v2.1}*

Michael Tamir

In his initial formulation of the general theory of relativity, Einstein's proposal that *freely falling gravitating massive bodies follow geodesic paths* was submitted as an independent fundamental principle. By adopting this "geodesic principle" to supply the theory's law of motion, Einstein was immediately able to recover both the free-fall motion of bodies in non-relativistic regimes and the previously anomalous precession of the perihelion of Mercury. Over the last century numerous ostensible proofs claiming to have derived the geodesic principle from Einstein's field equations have been developed. As a result physicists and philosophers of science alike frequently herald Einstein's theory for having the unique distinction of being able to derive its dynamical "law of motion" from its own field equations.

In this paper I critically survey the multiple attempts to derive the geodesic principle in the context of Einstein's theory. Grouping these results into three major families, which I refer to as (1) *limit operation proofs*, (2) *0th-order proofs*, and (3) *singularity proofs*, I argue that none of these strategies successfully demonstrates the geodesic principle, canonically interpreted as a dynamical law that massive bodies must actually follow geodesic paths in Einstein's theory.

Specifically, I argue for the following three claims: First, limit operation proofs fail to demonstrate that massive bodies are ever guaranteed to follow geodesic paths. Second, on the contrary 0th-order proofs demonstrate that extended massive bodies generically deviate from uniformly geodesic paths. Moreover, the only potentially extended distributions of matter and energy that fail to avoid a uniform geodesic evolution are highly unstable, deviating from such motion under arbitrary perturbations of their angular momentum (or higher order moments). Third, thanks to certain mathematical theorems concerning distribution theory, alternative representations of massive bodies as unextended "point" particles must result either in precluding the possibility of coupling the particle to the spacetime metric in a way that is coherent with Einstein's field equations or in having to excise the particle (and its would-be path) from spacetime entirely. This three pronged argument reveals that not only does the geodesic "law of motion" fail to be a deductive consequence of the field equations, but also any attempt to canonically interpret the geodesic principle in such a way requires

*In the following, \mathcal{M} will be taken to be a smooth, orientable, four-dimensional manifold, and (\mathcal{M}, g_{ab}) will be referred to as a Lorentzian spacetime if g_{ab} is a smooth metric of signature $(+, -, -, -)$ defined on \mathcal{M} . Excepting quoted material all further notational conventions follow that of (Wald, 1984).

that either the gravitating body is not massive, its existence violates Einstein’s field equations, or it does not exist within the spacetime manifold at all (let alone along a geodesic).

Having rejected the canonical interpretation of the geodesic principle as providing a fundamental law of motion or dynamical equation, I briefly conclude with an alternative proposal that the geodesic principle be instead interpreted as a *universality thesis* analogous to the use of the term ‘universality’ to classify the group behavior exhibited across thermal systems during phase transitions. My suggestion is that if we weaken our interpretation of the geodesic principle, instead reading it as a stability claim about the (four-dimensional) clustering of free-fall gravitating bodies in the appropriate neighborhood of geodesic following models, we will be able to recover the confirmation results classically provided by the geodesic principle (such as the perihelion of Mercury confirmation), while still consistently applying the principle (now as a universality thesis) to actual massive and extended bodies that can coherently couple to the spacetime metric in accordance with Einstein’s theory. Moreover, under such a proposed universality interpretation of the geodesic principle, limit operation proofs (previously argued to be feckless under the canonical interpretation) may be used to provide deep justification for geodesic stability results detected through the observation of generic free-fall bodies.

1 Einstein and The Canonical Account

1.1 Geodesic Dynamics

Einstein’s adoption of the geodesic principle was originally thought to be an independent postulate establishing the dynamics of the theory. Not long after the debut of his general theory, however, numerous special-case results and plausibility arguments were developed suggesting that in fact the principle was not logically independent (given certain assumptions about free-fall bodies) from Einstein’s field equations themselves.¹ In the appendix to the third edition of *The Meaning of Relativity* (1946), Einstein notes these developments concerning what he still refers to as “the law of motion” as follows:

In the initial formulation of the theory the law of motion for a gravitating particle was introduced as an independent fundamental assumption in addition to the field law of gravitation ... which asserts that a gravitating particle moves in a geodesic line. This constitutes a hypothetical

¹Some of the earliest cited proofs and plausibility arguments include (Weyl, 1922, Eddington, 1923, Pauli, 1921, Einstein and Grommer, 1927, Mathisson, 1937, 1940). Though (Einstein and Grommer, 1927) has often been cited as the earliest result, the results by Eddington, Weyl, and Pauli clearly predate it. Lesser known variations of these results were also offered in (Kopff, 1923), (von Laue, 1921), and (Becquerel, 1922), which were popular as texts on the new theory at the time (see (Havas, 1989, 1993) for further discussion of Einstein’s evident oversight in recognizing this early work). Despite his comments on the apparent redundancy of the geodesic principle (see below), Kennefick (2005) has argued that Einstein was very likely aware of the possibility of such special-case deductions prior to (Einstein and Grommer, 1927) as evidenced by possible fragments of an unadopted manuscript for (Einstein, 1922). Moreover, Einstein was clearly aware of the possibility of a special-case deduction, which carries over to the general theory, from his *Entwurf* predecessor to the debut of the full theory in (Einstein, 1913) (see note 13 below).

translation of Galileo’s law of inertia to the case of the existence of ‘genuine’ gravitational fields. It has been shown that this law of motion - generalized to the case of arbitrarily large gravitating masses - can be derived from the field-equations of empty space alone.(Einstein, 1922, p113)

Beyond crediting the apparent redundancy of postulating the geodesic principle as an independent assumption, note that Einstein explicitly characterizes the derivation result as pertaining not to some kind of test particle of either vanishing or arbitrarily small relative mass, but to *arbitrarily large gravitating masses*. The referenced result is no doubt that of (Einstein and Grommer, 1927) (and its successors), frequently considered a *locus classicus* of early demonstrations. As we shall see in section 1.3, since (Einstein and Grommer, 1927) was considered, at least by Einstein, to be a derivation of the geodesic principle, this work serves as an invaluable guide to how he expected the principle to be interpreted. In particular, it offers significant illumination into what Einstein came to believe was the content of his geodesic principle.²

With these results it seemed that general relativity differed remarkably from other classical field theories such as classical electrodynamics or Newtonian gravitation. In (Einstein and Grommer, 1927), the authors highlight an apparent matter-field duality found in these classical field theories. Echoing this dichotomy Leopold Infeld and Alfred Schild later characterize this equation duality in classical field theories as follows:

Classical physics is dominated by a characteristic duality of field and matter. In Newton’s theory of gravitation as well as in the Maxwell-Lorentz theory of electromagnetism the physical laws fall naturally into two independent classes. The first class consists of the partial differential equations which (with suitable boundary conditions at infinity) determine the field in terms of the distribution and motion of the matter which “generates” it. The second class consists of the dynamical equations governing the motion of matter under the forces “exerted” by the field. (Infeld and Schild, 1949, pp408-9)

They then proceed to explain how the equations of general relativity (viz. Einstein’s field equations plus the geodesic equation) fit into this picture, observing that just as in cases like classical electrodynamics, where there are two sets of equations, one set for how the fields couple with source charges (Maxwell’s equations), and another for dynamics of how “passive” charged bodies behave in those fields (the Lorentz force law), so too is there a duality corresponding to the two sets of equations in Einstein’s theory. His field equations govern how the field couples with the gravitational sources, while the geodesic equation provides the “law” for how gravitating bodies then surf the resulting metric field. In contrast to other classical field theories, however, for Einstein’s theory

²The idea behind what Einstein and Grommer identify as their preferred result is to squeeze the bodies into singular curves that are then excised from the spacetime entirely. At that point the source terms in the field equations of such a spacetime vanish, which is why in the long quote above Einstein notes that it can be derived from the equations for “empty space alone.” (The bodies have been fit entirely into the excised curves making them technically “outside of” the manifold and so not source terms of the field equations.) The demonstration is supposedly completed by their argument that, if we *were* to “replace” the excised curves, they *would* be geodesics of the vacuum solutions to the re-patched spacetimes (cf. (Infeld and Schild, 1949, p410)). This vacuum-cum-singularities technique was further developed by Einstein, along with Infeld and Hoffman, in (Einstein et al., 1938, Einstein and Infeld, 1940, 1949) as Einstein became increasingly opposed to representations of matter by means of continuous fields (see section 1.3).

it now seemed that the field equations for coupling the metric to energy-momentum sources also entailed the geodesic equation for how free-fall massive bodies behave in a given geometric field. Unlike with electrodynamics and Newtonian gravitation, the dynamical equations appeared not to be logically independent.³

This “duality of field and matter” and the dichotomy of their corresponding equations, endorsed by Einstein himself, is significant for two reasons. First, it emphasizes the apparent boon for Einstein’s theory: general relativity is “special” in comparison with other classical theories because the dynamical equations of the theory appear to logically follow from its own field equations.⁴ Second, in order to even claim that general relativity has such a special status among classical field theories, one must subscribe to a key presumption about the role of the geodesic principle active in the early decades of the theory (and still endorsed frequently today), namely, that (analogous to the role of the Lorentz force law in electrodynamics) *the geodesic principle plays the role of providing the dynamics of material bodies in the general theory of relativity*. In the following, we shall refer to this account of the role of the geodesic principle as providing the dynamics of general relativity as the *canonical account*.

In his early comments on the geodesic principle, Einstein frequently endorses this canonical account. In (Einstein, 1916) as well as his (Einstein, 1922) lectures on the theory, Einstein refers to the geodesic equation or the principle as the “equation of motion” or “law of motion” over a dozen times, characterizing them in this way not only for application to “particles,” but also in describing “planetary motion” (most importantly the motion of Mercury) and the motion of a “gravitating body” in general. As already indicated by the above long quote, Einstein continued to view the geodesic principle as providing the “law of motion” not only for massless test particles but also for “arbitrarily large masses” well after the theory’s initial introduction.

During this period, the canonical view was likewise frequently articulated by Einstein’s colleagues. It takes only a brief survey of the literature from the first half of the 20th century to reveal the widespread general adoption of the canonical view, with most authors taking it for granted that the geodesic principle provided the dynamics of the theory regardless of its logical independence from

³By ‘logical dependence’ here I mean derivability, perhaps under certain conditions characterizing the body in question. Of course dissolving the conceptually suspect bifurcation of bodies into “background” charged sources, which determine the field, and “passive” charged bodies that then react to the field (without generating self-forces) in this caricature of electrodynamic evolution leads to well known significant complications that have (even after over a century of effort) yet to be fully resolved (for an historical presentation and philosophical discussion of this problem see e.g. (Frisch, 2005)). As we will see, similar complications involving self-force-like effects are relevant in determining the actual motion of free-fall bodies in general relativity. The independence resulting from such a bifurcation of bodies into background sources and passive test bodies is, nonetheless, a separate notion from the *logical* independence of the dynamical equations of motion from the field equations, which (at least according to the interpretation we are now considering) exists in the electrodynamics case but not in the relativistic case.

⁴This distinction has been highlighted by philosophers such as Brown (2005, pp140-1) as well as the physicists who worked on this problem in the early decades (e.g. (Einstein and Grommer, 1927, Infeld and Schild, 1949)). Unlike these physicists, however, Brown astutely notes what he describes as the “limited validity” of deductions establishing exact geodesic motion, a point that we will investigate in detail below.

the field equations.⁵ Expressions of this view were unmitigated (and sometimes even highlighted) by the apparent redundancy of postulating the geodesic principle as an independent assumption. This attitude is typified by the commentary of physicists such as Lanczos, for instance, who punctuates his demonstration by noting that his penultimate equation “is equivalent to the ‘law of the geodesic line’ which has always been considered *the natural dynamical law of general relativity.*” (Lanczos, 1941, p818 emphasis added) Moreover, this canonical view of the geodesic principle as providing the dynamics is frequently cited in text books on the subject both classical (e.g.(Bergmann, 1942, pp224-5)) and contemporary (e.g. (Hobson et al., 2006, pp188-90) and even (Misner et al., 1973, pp475-80)).⁶ Though our focus will be on Einstein’s interpretation of the geodesic principle and its role in providing the dynamics of the theory, he was not alone in this attitude well into the mid 20th century. It is thanks to this combination of endorsements that the account of the geodesic principle providing the dynamics of Einstein’s theory is plausibly characterized as “canonical.”

1.2 Whither Test Particles?

If, according to what I have been describing as the canonical account, the geodesic principle provides the dynamics of the general theory of relativity, we must figure out to what exactly such a dynamical principle is supposed to pertain. Who follows geodesics? A natural answer might be something like “test bodies,” the theoretical tool in the physicist tool box used to describe how certain “sources” react to the field without having to attend to the actual effects on the field values caused by the presence of the bodies in question. In the case of relativity theory, we might then answer that “it is test bodies who follow geodesics.”

While we will ultimately see that under a non-canonical interpretation something like this answer might be endorsed (section 5), in the following survey of geodesic demonstrations it will be of central importance to observe exactly why and in what manner ignoring the source effects of “test bodies” can be justified. That is to say, we will need to pay special attention (i) if a gravitating object is treated as a “test body” because its source effects are simply left unaccounted for, or (ii) if the object is treated as a “test body” because its source effects can be shown to be negligible (but non-vanishing) for the relevant purposes of the deduction. The hazards of leaving test body approximations unjustified

⁵E.g. references to the geodesic principle as providing the dynamics or law of motion in some form or another are evident in (Eisenhart, 1928, Eddington, 1923, Tolman, 1930, Dirac, 1938, Lanczos, 1941, Infeld and Schild, 1949). Some authors judiciously express the view in restricted form only as pertaining to “relatively small” masses or simply to “mass points” (e.g. (Weyl, 1922, p256) or (Bergmann, 1942, pp224-5)). As we will see, in the former case, there was still little real justification for such heuristic winks at sufficient smallness (see sections 1.2 and 4.2 below), whereas in the latter case significant difficulties abound when it comes to representing the massive point particle that supposedly follows a geodesic within the theory (see section 2).

⁶In the case of references found in contemporary texts there should be no doubt that the authors are well aware of gravitational multipole and “self force” effects resulting in non-geodesic motion (see section 3 below). (Misner et al. (1973, p479) are notably circumspect about some of these failures and later offer explicit instruction on calculating spin effects.) Hence, such references should be taken only as evidence of the pervasive popular endorsement of the canonical view and the fact that the effect of the view’s initial adoption still lingers in contemporary conceptions of Einstein’s theory.

(i.e. case (i)) become most vivid when we consider proofs of the geodesic principle. In cases where the field equations and dynamical equations are starkly separated, physicists have the luxury of an apparent distinction: bodies whose source behavior is “turned on” are governed by the field equations, whereas the behavior of test bodies can seemingly be restricted to the purview of the dynamical equations alone. However, if one attempts to *deduce* the dynamical laws from the field equations, this specious luxury evaporates. We are forced in the course of the proof to simultaneously discuss the matter-energy of the field equations as the matter-energy that we ultimately hope to show obeys the dynamical equations. Hence, it is not even an apparent option to treat test matter-energy as being *entirely free* of the field equations as might happen in case (i).

As we will see, under the canonical interpretation, ignoring source effects of a body (even when they are small) can often have significant impact on the general validity of the deductions. With his characteristically sardonic wit when discussing this subject, Jürgen Ehlers, in collaboration with Ekkart Rudolph, emphasizes this challenge as follows:

The *test body approximation* is usually defined by the requirement that the contribution of the body to the metric $g_{\alpha\beta}$ be negligible. The justification of this drastic simplification in any particular case is by no means trivial and is therefore rarely considered. Since, according to Einstein’s (and similar) field equations, the curvature within a body is of the order of the density, the “self curvature” usually dominates or is at least comparable to the “incident” or “external” curvature (even for a small iron ball near the Earth’s surface), and then it is wrong to take the metric within the body to be nearly equal to the “given,” external one in the local mechanical law $[T^{\alpha\beta}_{;\alpha} = 0]$. (Ehlers and Rudolph, 1977, p208)⁷

In case (ii) above the physicist will be able to explain why the effects of the test bodies are inconsequential in a relevant and rigorous sense and may hence be justifiably ignored. While the majority of attempts at geodesic demonstrations (certainly, at least, at the time of this quote) seemed to fall under case (i), Ehlers and Rudolph here explain that a supplementary justification of these test body approximations with heuristic winks at relative smallness will not typically suffice; far more work is left unfulfilled.

If the geodesic principle is to provide a dynamics that can be legitimately used to predict the paths of actual bodies, we must find a way to draw suitable inferences from how these test particles are supposed to behave to how actual bodies behave. Unfortunately, as is well known from the case of classical electrodynamics, paying attention to the actual field-creating abilities of our (in the electrodynamics case, charged) test bodies, things become increasingly messy. Shrinking the body down to “infinitesimal” volume results in a singular charge density, and extending the particle still results in having to grapple with non-analytically expressible expansions of the effects that the particle’s own field has on its motion.

⁷Ehlers and Rudolph go on to explain in a parenthetical that “For this reason the mathematically elegant argument given in (Geroch and Jang, 1975) is physically not very enlightening, in our opinion.” We will return to this point in section 4.1.

In the case of general relativity things are even more treacherous. As we shall see, not only are there self-force and spin effects to be grappled with, but also, in the case of general relativity, the presence of matter-energy, whose powers as a field source have not been artificially “turned off,” will effect the very metric that determines what counts as a geodesic. An infinite matter-energy density in general relativity is not just an aesthetically disheartening anomaly in our representation, it often results in our inability to coherently speak about the spacetime path where the singularity occurs. But if the metric becomes undefined wherever the source particle is located (if it can even be said to have a location), how are we supposed to say that it is following a “geodesic” of that metric? On the other hand, the representation of extended bodies in general relativity leaves a good deal more freedom available for how the body’s matter-energy is distributed, making it difficult to speak generally about such representations of the bodies (especially that they universally follow geodesics). As we shall see, some of these issues had already become manifest by the time of (Einstein and Grommer, 1927).

1.3 Einstein and Grommer’s “Three Ways” to get it Straight

In the introduction to their paper Einstein and Grommer lay out the same dichotomy in Newton’s theory of gravitation and classical electrodynamics between field equations and dynamical equations articulated decades later in the long quote discussed above from Infeld and Schild. Characterizing such matter-field “dualism” as “disturbing to any systematic spirit,” they proceed to identify three “ways” [Betrachtungsweisen] in the general theory of dealing with such duality (Einstein and Grommer, 1927, p3).⁸

Their “first way” is modeled after Newtonian gravitation, in which the field equations and the geodesic equation are posited independently.⁹ This approach is most similar to Einstein’s initial introduction of the theory in that the field equations and the geodesic equation are postulated independently. Unlike his initial introduction of the theory, however, in this method the field equations in question are not Einstein’s full field equations:

$$G_{ab} = T_{ab} \tag{1.1}$$

where the *Einstein curvature tensor* on the left hand side is defined by

⁸Unless noted otherwise, this and all below translations of (Einstein and Grommer, 1927) are thanks to the gracious assistance of Bihui Li.

⁹This method can be thought of as being modeled after the field theoretic accounts of Newtonian gravitation in the sense that there too background sources might be represented by singular points generating a gravitational potential field ϕ , where ϕ is a solution to Laplace’s equation (i.e. Poisson’s equation with $\rho = 0$) with suitable boundary conditions at the singular points. This potential field is then surfed by test bodies satisfying the equation of motion $-\nabla^i \phi = \frac{d^2 x^i}{dt^2}$, which of course comes from Newton’s laws of motion and gravitation. Einstein emphasizes this analogy explicitly in (Einstein and Infeld, 1949, p209-10), his final collaboration with Infeld on the subject.

$$G_{ab} := R_{ab} - \frac{1}{2}g_{ab}R \quad (1.2)$$

and the right hand side of the equation represents the flow of matter-energy from any perspective. Instead, for their “first way” they specify that the relevant field equations are the *vacuum* field equations where equation (1.1) reduces to the equation:

$$R_{ab} = \mathbf{0} \quad (1.3)$$

once the energy-momentum tensor field is made to vanish everywhere. At first blush, the appeal to the vacuum field equations in this account may strike the reader as somewhat backwards. According to the dichotomy discussed above, it would seem that in this method Einstein and Grommer are dealing with the uncomfortable distinction between “test energy-momentum” and the energy-momentum sources contributing to the gravitational field by eliminating the sources while keeping the test bodies. But in the discussion above, it was the energy-momentum *test bodies* facing conceptual complications, not the “background” sources. So in their “first way” it might appear that they are getting things the wrong way round, having eliminated the source energy-momentum while retaining only the conceptually suspect test bodies.

One way of seeing why they specify the vacuum equations in this case is to consider why (from the perspective of 1927) Einstein might have wished to employ the two sets of equations to generate predictions.¹⁰ In particular, we might consider how he would have calculated the perihelion of Mercury at that time. Once the Schwarzschild solutions had been discovered, it was possible to determine the stable geodesics of the metric and from there calculate the perihelion of the orbit. So modeling the sun (minus all the planets) with the Schwarzschild solution, we could apply this calculation schema to the case of a Mercurial test body in such a background metric. But observe, the Schwarzschild solution is a *vacuum* solution. According to our application of this model, all the matter and energy of the Sun is to be found not in the spacetime manifold but “in” the singularity of the Schwarzschild Solution “located” at the origin of the coordinate system. Einstein and Grommer could consider the first way to be a possibility even though there is literally no place for energy-momentum sources in the manifold, because hiding the sources in the singularity works so well in this kind of application. Einstein and Grommer’s selection of the vacuum equations in the “first way” is indicative of a significant shift in how Einstein in particular began to prefer to represent matter-energy in his (as he saw it, not yet complete) theory. This attitude becomes even more apparent in their response to the next method.

The problem with the “first way” of course is that rather than dissolve the aforementioned discomfort with matter-field dualities when it comes to the general theory, it exacerbates the dichotomy. In contrast, according to their “second way,” *all* matter-energy is represented via a continuous and

¹⁰See Einstein’s comments on this strategy in (Einstein, 1995, p310).

singularity free energy-momentum tensor field T_{ab} . Unlike their “first way” method, this time they seem to get things the right way round when it comes to the elimination of the potentially suspect test bodies. They keep only the source matter-energy of the tensor T_{ab} while eliminating all appeals to test matter-energy.

After noting that as a consequence of (1.1), the total divergence of the energy-momentum tensor vanishes, without any calculation or further explanation, they make the following claim:

If one assumes that matter is arranged along narrow “world-tubes” one obtains from this by an elementary consideration the theorem that the axes of those “world-tubes” are geodesic lines (in the absence of electromagnetic fields). This means: the law of motion is a consequence of the field law.¹¹

It is difficult to speculate which “elementary consideration” establishes their demonstration. Though by 1927 special-case derivations of the geodesic principle from Einstein’s non-vacuum field equations had gained substantial proliferation, in the intervening decade since the debut of his general theory, Einstein never published any discussion or recognition of such (apparent) redundancy. Given his well known reputation for neglecting the literature, it is possible (though remarkable in light of his familiarity with a number of the authors)¹² that Einstein was not even aware of the abundance of such results.¹³ Eddington’s plausibility result in particular would appear to be paradigmatic of Einstein and Grommer’s “second way” approach, but it requires significant symmetry assumptions about the world-tubes in question in order to establish geodesic motion. Whether they were aware of these earlier results or simply referring to their own margin calculations, for reasons that will

¹¹Translation of quote from (Havas, 1989, p240).

¹²See (Havas, 1989) for detailed discussions on this point.

¹³Einstein’s decade of silence (at least in publications) on the derivability should not be taken as evidence of his ignorance of special-case derivations. It has been recently argued by Kennefick (2005) that there is evidence that he was quite familiar with the possibility of special-case results. In particular the geodesic motion of pressureless dust matter, which transfers to the full theory, was derived within the *Entwurf* theory. The easily transferred *Entwurf* result in question can be understood by considering the following elementary derivation: Suppose matter takes the form of a “pressureless dust” such that the energy-momentum tensor field can be written

$$T^{ab} = \rho U^a U^b$$

where the U^a have been normalized to be unit timelike. Then, if the covariant derivative of the left hand side vanishes we have

$$0 = \nabla_a(\rho U^a)U^b + \rho U^a \nabla_a U^b$$

but contracting with U_b annihilates the second term leaving us with

$$0 = \nabla_a(\rho U^a).$$

So plugging this back into the second equation, at spacetime events where $\rho \neq 0$ we can divide through by ρ giving us that the “dust matter” there obeys the geodesic equation

$$0 = U^a \nabla_a U^b.$$

In a recently uncovered fragment of notes evidently intended for his (Einstein, 1922), Einstein claims that his field equations “already contains [sic.] the divergence equation and with it the laws of motion of material points,” suggesting that he remained aware of this kind of result during the intervening decade. It should go without saying that success in such a pressureless dust derivation does not generalize to arbitrary applications of the principle (nor is its application in certain cosmological models above reproach).

become evident in section 3, it is difficult to imagine that Einstein and Grommer's unexplicated "second way" derivations could have been terribly general, despite their tone to the contrary.

In any case, they immediately abandon this victory over the matter-field dualism in general relativity, rejecting such "second way" derivations on the grounds that the use of a continuous energy-momentum tensor field T_{ab} to represent the distribution of matter-energy throughout the manifold is suspect:

It looks as though the general theory of relativity has already overcome that annoying dualism. This would be the case if we had already arrived at a representation of matter through continuous fields, or if we were at least convinced that one day we will arrive at it. But there can be no question of that happening. All attempts in the last years to explain the elementary particles of matter through continuous fields are failures. The suspicion that this is ultimately not the correct route to understanding material particles has become very strong in us.

This suspicion of the energy-momentum tensor was by no means a sudden development in Einstein's attitude. Such comments echo cautions voiced by Einstein from the very beginning of his presentation of the general theory. He expresses wariness about such a representation of matter-energy, for instance, in his (Einstein, 1922) lectures as follows:

In reality, matter consists of electrically charged particles, and is to be regarded itself as a part, in fact, the principal part, of the electromagnetic field. It is only the circumstance that we have no sufficient knowledge of the electromagnetic field of concentrated charges that compels us, provisionally, to leave undetermined, in presenting the theory, the true form of this tensor. From this point of view it is at present appropriate to introduce a tensor, $T_{\mu\nu}$, of the second rank of as yet unknown structure, which provisionally combines the energy density of the electromagnetic field and that of ponderable matter; we shall denote this in the following as the 'energy tensor of matter'.(Einstein, 1922, p85)¹⁴

And in his perhaps most poetic (and well known) rejection of such a continuous energy-momentum tensor field representation, in 1936 Einstein offers the following illustration of this attitude:

[General Relativity] is sufficient - as far as we know - for the representation of the observed facts of celestial mechanics. But it is similar to a building, one wing of which is made of fine marble (left part of the equation [(1.1)]), but the other wing of which is built of low-grade wood (right side of equation [(1.1)]). The phenomenological representation of matter is, in fact, only a crude substitute for a representation which would do justice to all known properties of matter.(Einstein, 1995, p311)

Einstein's disparagement of the energy-momentum tensor as analogous to "low-grade wood" has to do with its representation of matter-energy by means of the continuous tensor field. He describes it as a "phenomenological representation" because such a continuum representation is so close to the representation of matter in continuum mechanics as taking the form of a continuous medium (as is phenomenologically apparent) rather than an atomistic or quantum form.¹⁵ Einstein's resistance to the "low-grade wood" representation of matter energy particularly in the context of determining the

¹⁴See also his reflection on these hesitations in (Einstein and Rosen, 1935, note 3)

¹⁵Cf. (Einstein, 1922, pp52-53)

motion of bodies was intimately tied to his hopes for a unified theory, and the final remark of (Einstein and Grommer, 1927) explicitly speculates about their preferred “third way” methods leaving room for integration with the “quantum theory of matter.”¹⁶ Though such hopes failed to come to fruition, save for a brief wavering in 1935,¹⁷ Einstein would continue to resist the representation of matter-energy by means of a continuous tensor field in favor of a singularity approach for the remainder of his life, frequently voicing his skepticism of “low-grade wood” approaches.¹⁸

Einstein and Grommer’s “third way” avoids both the “low-grade wood” representation of background sources with a continuous tensor field as well as suspect appeals to test bodies. Instead (in the absence of electromagnetism), it makes use of the vacuum field equations alone, attempting to hide all matter-energy along singular “world-lines” of the manifold. In the conclusion they characterize their result as a (special-case) demonstration that these “singular world-lines” obey the geodesic principle, stating that “[i]f one understands masses in the gravitational field as singularities, then the law of motion is fully determined by the field equations.”

Einstein’s ultimately preferred singularity approach to the representation of matter-energy significantly influenced his interpretation of the geodesic principle. His adoption (and somewhat mistaken interpretation) of the singularity method makes it clear why in his 1946 appendix to (Einstein, 1922) he thought he could characterize the geodesic “law” as applying not just to test matter-energy but to “arbitrarily large bodies.” By making use of the singularity results, he believed he was free to hide as much matter-energy as he likes in the singular “world-lines,” while still (ostensibly) being able to derive the geodicy of such curves. In section 2 we will critically review the incoherence of such “third way” strategies, particularly in attempting to show that such “world-lines” are geodesics.¹⁹ But for now it is worth noting that though their introduction of the singularity method is initially characterized as a representation of elementary particles, Einstein quickly shifts the auspice of his “derived” principle to include large composite bodies such as Mercury as well.²⁰ Einstein’s dynamical interpretation of the principle did not hinge on the ability to treat bodies obeying the principle as arbitrarily small, nor did he see the proper interpretation of the dynamical role of the principle as subject to the uncomfortable matter-field duality found in classical electrodynamics and Newtonian gravitation. In fact, his work with Grommer strongly indicates that by 1927 he viewed the evident

¹⁶His early hopes (later dashed) that attending to the motion of bodies may yield insight into such unification have been recorded by collaborators such as Infeld (1980). For discussion see (Pais, 2005, Howard, 1990, Earman and Eisenstaedt, 1999).

¹⁷This wavering was in response to a persistent challenge posed in his correspondence with Ludwik Silberstein (see (Havas, 1993) for a detailed review of this controversy) and only lasted for a period of months surrounding his publication of (Einstein and Rosen, 1935).

¹⁸See e.g. (Einstein et al., 1938, Einstein and Infeld, 1940, 1949)

¹⁹To avoid a tempting conflation, note that Einstein and Grommer’s “three ways” are distinct from what I will below (in sections 2,3, and 4) classify as the *three general families of deductions*. Though the method of “singularity proof” of section 2.1 uses Einstein and Grommer’s “third way” strategy, both the 0th-order proofs and limit operation proofs of sections 3 and 4 clearly count as “second way” strategies according to the Einstein and Grommer classification, breaking any compelling analogy.

²⁰Moreover, early post-Newtonian confirmations of the two-body motion of stellar objects is often credited back to the work in (Einstein et al., 1938) and its successors, which likewise adopts the singularity method.

derivability of the principle by means of the singularity method as a significant triumph in this respect. In his view, by making use of such singularity methods, Einstein could allow the geodesic principle to play a dynamical role *for actually massive bodies* without (any longer) having to succumb to such dualism. Unfortunately, singularity proofs, both those using Einstein’s methods as well as those using more sophisticated methods, ultimately fail to establish the geodesic principle in a way that is compatible with his field equations.

2 Singularity Proofs

The family of proofs which I will refer to as *singularity proofs* really consist of two distinct subclasses. The first subclass follows Einstein and Grommer’s original “third way” method in which they attempt to use true singularities in the manifold in order to represent matter-energy. These singularities in the manifold are then (somehow) supposed to be shown to be geodetic. With the mathematical advances in distribution theory, these true singularity proofs were succeeded by the second subclass, which attempts to leave the metric well defined at the location of the geodesic following particles by coupling it to energy-momentum tensor distributions. In the next two subsections, we will consider each of these in turn.

2.1 The Geodesic that Wasn’t There

As already hinted, the most perspicuous difficulty with Einstein’s method of deducing the geodesic principle for particles represented as singularities in vacuum solutions is that (strictly speaking) the supposed path of such geodesic following particles is not even in the spacetime manifold. In (1995, p12) Earman poetically summarizes this “perplexing” strategy with the explanation that “to speak of singularities in g_{ab} as geodesics of the spacetime is to speak in oxymorons.” The proponent of such a “vacuum-cum-singularity” technique is faced with the rather paradoxical challenge of explaining in what sense we can say that a singular curve (ostensibly constituted by the *missing* points in the manifold) is actually a geodesic of the spacetime from which it is absent. Not only is no metric defined at the singularity, but technically there are not even any spacetime events there: the “geodesic” doesn’t exist. By eliminating the “low-grade wood” representation of matter-energy sources, Einstein dodged the difficulties associated with using continuous representations of energy-momentum that might restrict the generality of the principle (see section 3 below) but only at the cost of having to justify the geodicty of a metric-less hole in spacetime.

Though Einstein and Grommer avoid elaborate consideration of this challenge, their strategy might (briefly) be characterized as follows: splitting the first order perturbations of the Minkowski metric in the neighborhood of the singularity into an “exterior” ($\gamma_{\alpha\beta}^{(ext)}$) part resulting from sources “far” from the singularity and an “interior” ($\gamma_{\alpha\beta}^{(int)}$) part resulting from the ostensible presence of the body “at” the singularity, they then argue that in their chosen coordinate system, $\gamma_{\alpha\beta}^{(ext)}$ obeys the

constraint that

$$\frac{\partial \gamma_{44}^{(ext)}}{\partial x_\mu} = 0$$

along coordinates of the x_4 -axis where they locate the singularity.²¹ The suggested implication then is that for a *second* singularity-free spacetime, whose metric is given by a $\gamma_{\alpha\beta}^{(ext)}$ correction to the Minkowski metric, the x_4 -axis (not in the domain of the first metric) is a geodesic. They seemingly take for granted, however, that we must associate the singular boundary in the former spacetime with the corresponding “filled in” points along the x_4 -axis in the latter spacetime, perhaps due to the appearance of an embedding relationship suggested by the similarity in the coordinatizations of the respective spacetimes. Unfortunately, the claim that Einstein and Grommer’s second singularity-free spacetime can tell us something about the nature of the singularity of the original spacetime is spurious. A similarity in the coordinates used to refer to the singularity in one spacetime and the coordinates of a second spacetime without a singularity at those coordinates is not enough to infer that the second spacetime is a “filling in of the singularity.”

In an attempt to vindicate the vacuum-cum-singularity strategy, Infeld and Schild concede that “[c]learly, the statement that a singular line is (or is not) a geodesic has no meaning” (Infeld and Schild, 1949, p410).²² They proceed to argue that the geodesic principle might nonetheless be proven by means of the vacuum-cum-singularity strategy, if it is once again asserted that the principle is (at least) germane for a certain kind of representation of test particles:

Physically, we can consider a sequence of particles, with masses tending to zero, and a corresponding sequence of gravitational fields. In the limit $m = 0$ we obtain a limiting world line along which the limiting gravitational field, the background field, is continuous. We must think of the background field as being assigned *a priori*; the geodesic “postulate” refers to the limiting world line in this continuous field and is thus meaningful.

Recall, Einstein claimed (as late as 1946) that the vacuum-cum-singularity method can be used to derive the postulate for “arbitrarily large masses.” The move of restricting their geodesic result only to this specific variety of test particles, which we will refer to as *Infeld-Schild* or *IS-particles*, constitutes a strategic retreat from Einstein’s position. Infeld and Schild’s derivation might hence be thought of as an attempt to embrace case (ii) considered in section 1.2 by trying to justify why the effects of the “test” body can be ignored. By restricting their results to these IS-particle sequences of spacetimes, Infeld and Schild were forced to limit their result to particles of *arbitrarily small*

²¹Of course, technically the metric is not well defined at those coordinates, but they claim to avoid this problem by stipulating that, since $\gamma_{\alpha\beta}^{(ext)}$ is generated by “external” sources, it should be regular in the neighborhood of the singularity. The (suspect) intimation being that for this reason it can be unproblematically extended across the coordinates of the singularity.

²²Infeld, Einstein’s long time collaborator on the motion of bodies, became one of the principle champions of singularity methods (both the vacuum-cum-singularity method and then later the distributional method) well after Einstein’s final contributions to the problem (Einstein et al., 1938, Einstein and Infeld, 1940, 1949, Infeld and Schild, 1949, Infeld, 1954, 1957, Infeld and Plebanski, 1960).

mass but now had the chance of explaining why we can associate the geodicity of a curve γ of the “background field” with the character of the singularities in spacetimes with $m \neq 0$: The background spacetime is a “limit” of the singular spacetimes, and the coordinates in each of these spacetimes demarcating the singularity are the same as those used to locate γ in the background spacetime.²³

The limiting procedure Infeld and Schild use is fatally flawed. Though there is a coordinate similarity of the “limit” spacetime and the singular spacetimes (speciously) suggesting an embedding relation, the singularity will exist for every one of the $m \neq 0$ spacetimes in the run up to the supposed “limit.” For every spacetime short of the background one, the “behavior at γ ” will remain undefined, obscuring the sense in which the singular spacetimes are “approaching” the background one. Again, there is no rigorous sense in which the singular behavior of the sequence of spacetimes converges to a non-singular background spacetime, making references to “*the* limiting behavior” literally nonsensical.²⁴

Infeld and Schild’s attempts to derive the geodicity of singularities in the manifold by considering perturbations in the boundary conditions that could be taken to indicate presence of arbitrarily small matter-energy located at the singularity ultimately failed. Though it is possible to use surface integral techniques, integrating around the singularity to suggest that there is (something like) matter-energy “hidden” so to speak at the undefined (singular) boundary region, there is no way of rigorously discussing what goes on “at the singularity” of such vacuum solutions, and in particular, no way of inferring geodicity. Einstein and Grommer’s “third way” vacuum method hence turns out to be unsalvageable even with a retreat to the arbitrarily small IS-particles.

By 1954, even Infeld had turned to a kind of compromise between the “second way” appeal to a non-vanishing energy-momentum tensor field and the “third way” attempts to concentrate matter energy onto a world-line where the metric diverges (Infeld, 1954, 1957). In this method Einstein’s “low grade wood” is replaced (metaphorically speaking) by a kind of sturdier “pressure treated wood” through the introduction of energy-momentum tensor distributions.

2.2 Pressure Treating the Wood: Distributional Energy-momentum

The idea behind distribution proofs of the geodesic principle is to concentrate all the matter-energy of a (would-be) geodesic following particle onto a one-dimensional (often timelike) curve γ . Once this is done, the task is to deduce from Einstein’s field equations (or a generalization of them) that γ must be a geodesic. In contrast to the singularity proofs of the last section, proofs using distributional energy-momentum do not use the vacuum field equations (1.3) ultimately preferred by Einstein. Instead a non-vanishing energy-momentum tensor distribution on the right side of the equation is used to represent the particle. In a sense then, distribution proofs are similar to Einstein

²³ Actually, unlike Einstein and Grommer, they attempt to use (proto-)geodesic completion methods to covariantly specify the singular “points.” Such completion methods unfortunately remain insufficient.

²⁴ There may not even be a pathology-free (e.g. Hausdorff) way to fill in the singularity (see (Geroch et al., 1982)). See (Torretti, 1996, pp178-9) for further discussion of this fallacy.

and Grommer’s so-called “second way” demonstrations in that, like those proofs, they appeal to field equations coupling the geometry of the manifold (via the Einstein curvature) to non-vanishing energy-momentum sources.

Unlike “second way” demonstrations, however, distribution proofs do not (strictly speaking) make use of Einstein’s original field equations (1.1). In “second way” demonstrations, the objects equated (G_{ab} and T_{ab}) are typically supposed to be smooth tensor fields defined on the manifold. In distributional (singularity) proofs, on the other hand, the energy-momentum field on the right hand side of the equation is a *distributional* object in some neighborhood of γ : assuming no electromagnetic contributions in that region, its support is restricted to a measure-zero region (i.e. the one-dimensional region γ). Yet in order to attribute non-vanishing matter-energy to the “particle,” we do not want integrals of the field in that region to vanish. Hence, such proofs make use of an energy-momentum *tensor distribution* that is defined by its action on a space of well behaved (mathematical) “test” objects.²⁵ The definition of the space of tensor distributions offers a natural extension of the space of locally integrable tensor fields in a way suggesting that we can “integrate” certain distributions concentrated on measure-zero regions without the integral necessarily vanishing (details are reviewed in appendix A).

In order to conduct a distributional derivation of the geodesic principle from “the” field equations of general relativity, Einstein’s original equations (1.1) must (at least implicitly) be modestly generalized by saying:

$$G_{ab} = T_{ab} \quad \text{as tensor distributions} \quad (2.1)$$

I.e. if \mathcal{M} is an orientable manifold, for all $\phi^{ab} \in \mathcal{T}_0^2(\mathcal{M})$ with compact support:

$$\int_{\mathcal{M}} G_{ab} \phi^{ab} \text{vol} = \int_{\mathcal{M}} T_{ab} \phi^{ab} \text{vol} \quad (2.1a)$$

where *vol* can be any volume element defined on \mathcal{M} .²⁶ Distributional proofs hence constitute a kind of compromise between Einstein and Grommer’s “second” and “third” way methods of derivation. Equation (2.1) is not a vacuum equation, but rather than taking a “low grade wood” approach to representing the matter-energy of the particle as a continuously distributed field, it (in a sense) allows us to specify the “particle’s” four-dimensional extension as being precisely restricted to the world-line (not tube) γ .

²⁵The use of the term ‘test’ in the context of distribution theory is only incidentally similar to the use of the term to refer to “test bodies.” The former are well behaved sets of mathematical objects on which distributions act, the latter (as already discussed) refers to a kind of theoretical representation signifying entities that react to physical fields but do not act as (significant) sources of those fields. Hence, there is no room for the two usages to be equivocated despite the unfortunately abundant opportunity for confusion.

²⁶Since smooth tensor fields $\mathcal{T}_2^0(\mathcal{M})$ are locally integrable and so have a natural embedding in the space of tensor distributions $\mathcal{D}'_2(\mathcal{M})$, using variational techniques, this relation trivially entails equation (1.1) in cases where the respective fields are smooth (see appendix A).

The earliest (implicit) use of distributional energy-momentum in the problem of motion in a relativistic context can be read into the derivations of the geodesic principle made by Myron Mathisson.²⁷ Infeld did not trade in Einstein’s vacuum-cum-singularity method for distributions until decades later in his (Infeld, 1954, 1957, Infeld and Plebanski, 1960).²⁸ Even still Infeld (1957, p399) characterized his reluctance about such a shift by noting that though it is technically “unfaithful” to “Einstein’s idea of not using the energy momentum tensor,” the introduction of a distributional energy-momentum tensor can be excuplated by the fact that it “tremendously simplifies the entire deduction of the equations of motion.”

The key to distribution proofs involves establishing a variational principle for integrals of the energy-momentum tensor for which, in the special case of energy-momentum tensor *distributions* concentrated on a world-line, the geodicity of the path is entailed. Specifically, assume one is able to establish that for all smooth co-vector fields ξ_b with compact support:

$$\int_{\mathcal{M}} T^{ab} \overset{g}{\nabla}_a \xi_b \text{vol}_g = 0 \quad (2.2)$$

where vol_g is the volume element for some metric g_{ab} and $\overset{g}{\nabla}_a$ is the derivative operator compatible with g_{ab} . It follows from (2.2) that if T^{ab} has (distributional) support restricted to a timelike curve γ in some neighborhood around it, then γ is a geodesic of g_{ab} .²⁹ So letting T^{ab} represent a point particle whose world-line is given by γ in that region of the manifold, we might interpret the result as saying that “point particles can only have a geodesic world-lines.”

Since condition (2.2) is sufficient for such a distributional representation of a point particle to follow only geodesic world-lines, it is worth considering how such a variational principle can be established in general relativity. Heuristically, we might first note that for *smooth* tensors, it is (in a sense) a *purely mathematical* consequence of the Bianchi identities that the total divergence of the Einstein tensor defined by (1.2) vanish (i.e. that $\nabla_a G^{ab} = \mathbf{0}$). Hence, Einstein’s original equations (1.1) immediately give us that

$$\overset{g}{\nabla}_a T^{ab} = \mathbf{0} \quad (2.3)$$

referred to as *the conservation condition*, which holds for any smooth solution $(\mathcal{M}, g_{ab}, T^{ab})$ to (1.1).

²⁷As we will discuss in section 3, Mathisson’s technique involved deriving a variational principle for the integral of an expansion (in gravitational multipole moments) of the energy-momentum tensor field, from which motion can then be deduced. In particular, in (1937, 1940) Mathisson casually shows that applying the principle only to the lowest order term in the expansion (because such a tensor field might be representative of a spinless “point particle”) entails the geodesic equation. Though Mathisson does not make explicit use of distributions in these 0^{th} -order derivations, such an appeal can naturally be read into this technique as was done later by Havas and Goldberg (1962). See also (Tulczyjew, 1959) for a distributional reconstruction of Mathisson’s work.

²⁸Infeld and Mathisson were colleagues in Poland when Mathisson had been developing the work from which his derivations follow. Infeld, who was familiar with the relevant papers, later conceded that “at the time” he had not understood Mathisson’s (significantly more advanced) methods (Infeld, 1968, p204).

²⁹The full proposition is given with a proof in appendix B .

Condition (2.2) follows for such smooth solutions from (2.3) by simply contracting with arbitrary test co-vector fields ξ_b and then integrating over the entire manifold.³⁰

The problem is that we want (2.2) to hold not just for smooth solutions to (1.1) but for *distributional* solutions to (2.1) as well. And in fact, the Bianchi identities do not automatically hold for distributional Einstein tensors for every solution of the generalized field equations (2.1).³¹ Hence, *the conservation equation (2.3) does not automatically follow in the generalized case of tensor distributions*. But both Mathisson’s implicit distribution result and Infeld’s explicit one overlook this nuance, conflating the distinction between solutions to (1.1) and (2.1) and then inferring that (2.3) (and so (2.2)) automatically follows even for distributional sources.³²

In 1974, Jean-Marie Souriau developed his own “proof” of the geodesic principle by making use of distribution techniques, which again (essentially) take advantage of condition (2.2). Unlike earlier attempts, however, Souriau justifies the condition not through the Bianchi identities, but by (rather ingeniously) formulating a “variational” method of expressing the condition that the field equations must be generally covariant (now referred to in the literature as *Souriau’s (local) covariance condition*)(Souriau, 1974).³³ In the case of the generalized field equations (2.1), such *Souriau covariance* easily reduces to condition (2.2).

Though Souriau’s method is able to avoid the particular invalidity of his predecessors’ arguments, his result is faced with an even more general threat to the use of distribution techniques in Einstein’s theory. In order to understand what goes wrong, observe that condition (2.2) does not just express a restriction on T^{ab} , but rather it expresses a restriction on the energy-momentum tensor field (or distribution) in relation to a metric g_{ab} . Hence, condition (2.2) (and Souriau’s subtly more general covariance condition) define a subset of *ordered pairs* of symmetric tensor fields or tensor distributions (g_{ab}, T^{ab}) defined on a manifold \mathcal{M} .³⁴ Let us refer to such pairs as *Souriau pairs*. Of course, solutions to the generalization of Einstein’s field equations (2.1) can also come in ordered pairs defined on \mathcal{M} . And Souriau’s covariance principle is supposed to establish that *for any solution (g_{ab}, T^{ab}) to (any) generally covariant field equations (such as (2.1)), (g_{ab}, T^{ab}) will constitute a Souriau pair*. So the logic works as follows:

³⁰In fact, as discussed in appendix A, if the connection is smooth, the (covariant) derivative of a distributional T^{ab} is calculated precisely by negating the left side of equation (2.2). Hence, condition (2.2) is the natural generalization of the classical conservation condition (2.3) and for this reason is sometimes referred to as the *generalized conservation condition*.

³¹Most importantly, they do not automatically hold for the important class of GT-regular solutions discussed in appendix C (see (Geroch and Traschen, 1987, p1020)).

³²See (Infeld, 1957, §4) and (Mathisson, 1937, §2). See also the explicitly distributional reconstruction of Mathisson’s demonstration in (Havas and Goldberg, 1962, §2). This equivocation can still occasionally be found in introductory texts offering what might be interpreted as “heuristic” derivations of the principle (see e.g. (Hobson et al., 2006, p188-9)).

³³The application of Souriau’s local covariance condition was further developed in both relativistic and non-relativistic contexts in (Guillemin and Sternberg, 1978, 1990, Sternberg, 1978, 1985b,a, 1999) and (Duval and Künzle, 1978, 1984) respectively.

³⁴Moreover, in the context of linear distributions, g_{ab} must be non-degenerate, and (at minimum) it must have a smooth connection wherever T^{ab} behaves singularly (as a distribution).

1. By Souriau’s covariance argument, if (g_{ab}, T^{ab}) is a solution to (any) generally covariant field equations (such as (2.1)), then the pair satisfies condition (2.2).
2. And by proposition 4,³⁵ if (g_{ab}, T^{ab}) satisfies condition (2.2), and T^{ab} is concentrated onto a timelike world-line, then it must be a geodesic of g_{ab} .

-
3. Hence, if (g_{ab}, T^{ab}) is a solution to (2.1), and T^{ab} is concentrated as a distribution onto a timelike world-line, then it must be a geodesic of g_{ab} .

Though this argument is valid, the antecedent of line 3 renders the conclusion (essentially) vacuous. The reason for this was ultimately demonstrated by Geroch and Traschen (1987, Thm. 1) in which they show that the only reasonable³⁶ solutions g_{ab} to the equations (2.2) cannot have sources (T^{ab}) with support concentrated on a one-dimensional curve.³⁷ *Hence, there is no (reasonable) distributional solution (g_{ab}, T^{ab}) to Einstein’s generalized field equations (2.1) such that T^{ab} can also be concentrated onto a timelike world-line.*

Geroch and Traschen’s theorem hence serves as an effective death nail for attempts to deduce the geodesic principle from Einstein’s field equations by using singular representations of one-dimensional “point” particles. As a result we seem to be left with the following option: We could conduct deductions from *inexact* (namely linearized) field equations.³⁸ It is possible to deduce exact geodesic motion in this case, *but not from Einstein’s actual field equations*. Alternatively, we might move to so called “second way” proofs attempting to deduce geodesic motion from Einstein’s (non-vacuum) field equations (1.1) using smooth energy-momentum tensors with four dimensional support to represent our geodesic following objects.³⁹ In the next two sections we will review the major strategies that have been used in such “second way” deductions. As we shall see, by moving to a context of extended representations of massive bodies, much more freedom in the behavior of the object is introduced, ultimately leading to the deduction of *non-geodesic* motion in generic cases. It will turn out that these additional modes of freedom can be later reduced by appealing to certain limiting cases, but this has the detrimental result of either bringing us back to the context of singularity proofs, or to

³⁵See appendix B

³⁶Recall, according to (2.1) the energy-momentum tensor is equated with the Einstein curvature tensor, and the curvature tensor in turn depends on the metric and its derivatives. Hence, the metric and its derivatives must meet certain minimal conditions on their integrability in order for “integration” (i.e. the action) of the Einstein tensor (and so the energy-momentum tensor) to make sense as a well defined tensor distribution. Geroch and Traschen define a class of metrics now called *GT-regular metrics* designed to meet such conditions so that the energy-momentum tensor distributions determined by these metrics can make sense. See appendix C for the precise definition of this class of metrics and a brief discussion of why this class in particular constitutes the appropriate class of “reasonable solutions” to (2.2).

³⁷They prove that distributional sources must have support of codimension no greater than 1 in \mathcal{M} .

³⁸The Geroch-Traschen proof crucially depends on the non-linear dependence of Einstein curvature on the metric. See appendix C for discussion.

³⁹A third possibility involving neither the original field equations (1.1) nor the (linear) distributional generalization (2.1) which might be developed is considered in Appendix C (see note 64).

limiting representations in which (contrary to the canonical account) the “gravitating” body simply vanishes.

3 0^{th} -Order Proofs

Einstein’s field equations have an initial value formulation. Under suitable conditions, this problem can be well posed so that we might use the field equations to deduce how a given tensor field defined on a particular hypersurface evolves over time (viz. the domain of dependence of that hypersurface). Though this works in principle, such a program is far more easily said than done in most cases and often numerically rather than analytically. As a consequence if physicists wished to predict, say, the motion of celestial bodies, it was more practical to figure out a way to approximate the structure of the bodies and the metric field affected by their presence by expanding both of these fields through various procedures and then dropping some of the “higher order” terms.⁴⁰ After the respective fields have been suitably simplified in this way, the physicist can take steps to determine the expected “approximate” paths of such bodies. For the most part *the resulting paths are not geodesic*. However, when *all* of the higher order terms of the tensor field representing the energy-momentum of a body (and the tensor field representing the body’s effect on the metric) are dropped, it is the case that one is able to “deduce” geodeticity from the reduced equations. Since these proofs share the feature that all higher order terms accounting for the energy-momentum of the body (and its effects on the metric) must be dropped in order to ensure geodeticity, this class of deductions will be referred to as 0^{th} -order proofs. The overwhelming majority of geodesic “derivations” in the literature can be classified as belonging to this family of proofs. The explanation for this pattern is entirely pragmatic: expansion methods are an efficient way of generating approximations of the motion of bodies under the influence of relativistic effects within particular margins of error.⁴¹ So, 0^{th} -order proofs are quite abundant in the literature, but ironically only thanks to the desire for approximations of motion accurate to degrees *higher* than 0^{th} -order, the lowest order geodesic deductions being offered as a kind of afterthought or perfunctory check.⁴²

By far, the earliest concerted attempts to approximate motion in relativity theory by expanding the energy-momentum of a body can be found in the works of Mathisson (1937, 1940). To understand the sense in which Mathisson “expanded” the energy-momentum tensor, consider a (timelike) world-

⁴⁰Aside from historical computational hurdles, there exists a further (epistemologically motivated) reason for why such approximation techniques might be advantageous. Namely, we might want to have a method of identifying the bulk behavior of general massive bodies, which may be obfuscated by attending every detail as in the initial value formulation. This further rationale will play an important role in the considerations of section 5. For now, however, we will delay discussion of this epistemological motivation.

⁴¹Bursts of progress in such techniques often appear to be motivated by concomitant instrumental advances demanding further accuracy. For example the advances in approximating “self-force” effects in the last dozen years seem to have been originally motivated by the promise of gravitational wave detectors (see (Quinn and Wald, 1999)).

⁴²Because of such abundance, in this section I will not attempt a comprehensive review of all major attempts. Instead, I will focus only on a few examples paradigmatic of the general methods of expansion techniques recovering geodesic motion (at the lowest order).

tube \mathcal{W} in a relativistic spacetime (\mathcal{M}, g_{ab}) .⁴³ We then consider a symmetric (locally integrable) tensor field T^{ab} with support contained in \mathcal{W} . This tensor field can be interpreted as representing the energy-momentum of a body moving along in the world-tube \mathcal{W} . At this point, Mathisson takes advantage of the fact that we can understand the properties of such a field by considering the following linear operation defined on arbitrary smooth tensor fields ϕ_{ab} also compactly supported in \mathcal{W} :

$$\langle T^{ab}, \phi_{ab} \rangle \mapsto \int_{\mathcal{M}} T^{ab} \phi_{ab} \text{vol}_g \quad (3.1)$$

Picking an arbitrary smooth timelike curve $I \ni s \mapsto \gamma(s) \in \mathcal{W}$ parametrized by proper time s , and letting $\Sigma(s)$ be a local foliation of \mathcal{W} parametrized by s and meeting certain orthogonality conditions with respect to γ , Mathisson proceeds to show that by integrating across the $\Sigma(s)$'s the action (3.1) can be equivalently approximated by a series of integrals along γ as follows:

$$\int_{\mathcal{M}} T^{ab} \phi_{ab} \text{vol}_g = \int_{\gamma[I]} \left(I_0^{ab} + I_1^{abm_1} \nabla_{m_1} + I_2^{abm_1m_2} \nabla_{m_2} \nabla_{m_1} + \dots \right) \phi_{ab} ds \quad \forall \phi_{ab} \quad (3.2)$$

Where the tensor fields $I^{abm_1 \dots m_n}$ satisfy certain symmetry conditions and orthogonality conditions with respect to γ . These tensor fields represent the 2^n -gravitational-multipoles of the body T^{ab} . The final move to arrive at Mathisson's variational equation of motion is to let $\phi_{ab} = \nabla_b \xi_a$ for arbitrary smooth, compactly supported co-vector fields ξ_a . But in these cases, the left side of (3.2) takes the form of the left side of the generalized conservation condition (2.2), which as we discussed above is equivalent to the traditional conservation condition for a smooth T^{ab} and metric. Based on this reason, Mathisson sets the left side of (3.2) to 0, giving us the final form of his *variational equation of motion*:

$$0 = \int_{\gamma[I]} \left(I_0^{ab} + I_1^{abm_1} \nabla_{m_1} + I_2^{abm_1m_2} \nabla_{m_2} \nabla_{m_1} + \dots \right) \nabla_b \xi_a ds \quad \forall \xi_a \quad (3.3)$$

Mathisson was able to use this variational equation, expanded to the first (dipole) and to the second (quadrupole) terms to generate explicit approximate equations of motion for a "small test body" with angular momentum and spherical asymmetries. These equations were (much later) derived in the better known (Papapetrou, 1951), using hyperbolic coordinates and somewhat similar expansion techniques, and are sometimes referred to as the *Mathisson-Papapetrou equations*. In the late 60's

⁴³Since our purpose in this section is primarily illustrative, in the following discussion I will assume that the spacetime is simply connected, orientable and that the metric is smooth. Some of these constraints might be relaxed, but as discussed above in section 2.2, doing so can lead to serious complications. Though Mathisson's work was quite sophisticated for his time, there are a number of mathematical ambiguities in his original formulation that I will not dwell on here, especially since such infelicities were eventually rectified by Dixon (see below). For example, the sense in which \mathcal{W} is timelike is made precise by Dixon through a construction that involves joining a particular set of local, convex, disjoint, hypersurfaces with compact closure that are normal to a timelike "baseline" curve.

W. G. Dixon took up the Mathisson project, eventually producing a fantastic series of papers Dixon (1964, 1967, 1970b,a, 1973, 1974, 1975), which rigorously put what he called the *multipole approximation* technique on fully matured mathematical footing.

The geodesic result (already mentioned in section 2) comes when we drop all higher order terms in the expansion of equation (3.3) to get the variational constraint:

$$0 = \int_{\gamma[I]} I_0^{ab} \nabla_b \xi_a ds \quad \forall \xi_a \quad (3.4)$$

From this constraint one can then deduce that if $I_0^{ab} \neq \mathbf{0}$ on $\gamma[I]$ then the curve is a geodesic of the metric compatible with the connection ∇_a .⁴⁴ So since (*ex hypothesi*) I_0^{ab} is supposed to be the only (significant) contribution to the energy-momentum of the body and our initial choice of γ was arbitrary, we can think of this result as telling us that certain “suitable”⁴⁵ timelike γ ’s contained in such a body must be geodesic. The victory is rather Pyrrhic, however, because approximations involving even one non-vanishing multipole term will no longer describe geodesic motion, and instead predict a kind of “wobbling” behavior inside \mathcal{W} .

In order to dodge this Pyrrhic foil, we might search for ways to justify the use of constraint (3.4) as an exact (or at least stable) description satisfying the canonical interpretation. One possibility is to justify the constraint by insisting that if the body is a point particle with only timelike extent, then the particle’s lack of spatial extent means that all higher order multipole terms must vanish, making the only curve left in the support of the point particle relevantly “suitable” and hence geodesic. This approach validates the inference to constraint (3.4) from Mathisson’s full variational equation (3.3), but it has the unfortunate side effect of turning the deduction into a variety of singularity proof which as we saw in section 2 is incoherent with Einstein’s field equations.

A second strategy is to accept that the body has spacelike extent, but then suggest that (3.4) holds of a certain “conceivable” type of material body, namely, one that is perfectly symmetrical about some timelike γ with respect to *every* higher order multipole moment. Though by construction such a representation would reduce the equation (3.3) to constraint (3.4), the proof is far from general. Such a “conceivable” body is not possible according to just about any serious theory of matter considered by physicists. For example, such extreme symmetry constraints would require that the body could not be composed of atomic or molecular constituents for such inhomogeneities would necessitate spherical asymmetries in the distribution of the body’s energy-momentum.⁴⁶ Moreover,

⁴⁴A deduction (somewhat different from that of Mathisson’s original proof) is contained in the proof of proposition 4 in appendix B (condition (3.4) is equivalent to condition (B.2), which in the course of the proof is shown to entail geodeticity, when we also assume that the body is non-vanishing on γ).

⁴⁵The *suitable* γ must be a curve with respect to which the higher order terms ${}_n I^{abm_1 \dots m_n}$ can be dropped. Intuitively this can be thought of as a timelike axis of symmetry with respect to all of the multipole moments of the body. As we shall see, when we expect perfect geodeticity (i.e. that all the higher order terms can be dropped because with respect to such a γ axis they actually vanish) this “suitability” condition becomes prohibitively strong.

⁴⁶The body would need to consist of something like the metaphysically curious homogeneous material famously

such restrictions necessary to generate geodesic motion are highly unstable under perturbations. Any change in the angular momentum (or any other multipole moment) would break the symmetry needed to exactly recover constraint (3.4). Though we might wish to identify these perfectly symmetrical constructs as a kind of “idealization” in Einstein’s theory, *the idealization is degenerate* with respect to recovering precise geodesic motion: the path becomes non-geodetic under *arbitrarily small* perturbations of the energy-momentum.⁴⁷

Before moving on to the final class of geodesic deductions, it is worth pointing out that by invoking a multipole approximation of the energy-momentum, the body considered is no longer coupled to the metric in accordance with Einstein’s equations (1.1). In other words, multipole approximations (even ones that do involve some higher order terms such as the Mathisson-Papapetrou equations) are guilty of adopting test body approximations that have not been explicitly justified. They ignore the so called “back reaction” or “self-force” effects that the body has on the metric.⁴⁸ Hence, when considering the path of extended bodies (i.e. “second way” deductions) we must not only worry about correcting for the possibility of unjustified test body approximations facing earlier proofs, but now must also manage the “spin effect” corrections to geodesic motion resulting from the internal degrees of freedom available to an energy-momentum tensor with spacelike extent. In the next section we will consider the final family of deductions, which attempts to manage such deviations from geodeticity by conducting certain limit operations. As we shall see, it is only by taking limits that appropriately manage both of these sources of non-geodeticity that success is achieved. Unfortunately for the canonical view, this will also mean that the body must vanish completely before we can recover such geodeticity.

4 Limit Operation Proofs

The strategy behind the final family of *limit operation proofs* is to avoid the complications arising from investigating the motion of “true” point particles with extent restricted precisely to one-dimensional timelike curves by instead considering sequences of energy-momentum representations of particles whose spacelike extent is confined to increasingly smaller neighborhoods of those curves. We can think of these infinite sequences of tensor fields as representing particles with *arbitrarily small (but non-zero) spacelike extent* in the sense that no matter how “narrow” we might want the

considered by Saul Kripke. In the context of Kripke’s spinning disk, it is ironic that one of the constraints that we are explicitly placing on such a material body is that it would have to be completely “spin free” in the specific sense that ${}_1J^{abm_1} = \mathbf{0}$.

⁴⁷Of course, as I shall argue below, this degeneracy only exists if we want to recover *exact geodesic motion* in accordance with the canonical interpretation. If alternatively we want to weaken our interpretation, as I will suggest, it may be possible to eliminate (or at least avoid) this kind of degeneracy.

⁴⁸Significant advances have been made in (Mino et al., 1997, Quinn and Wald, 1997, Gralla and Wald, 2008, Pound, 2010), which approximate the consequences of “self-force” effects as first order perturbations in the background metric. Though these methods are not without their own difficulties (particularly when it comes to justifying what is referred to as the “Lorentz gauge relaxation”) because these self-force effects lead to violations of geodesic motion, I shall not elaborate on these complications here. For a nice introduction to these issues see (Wald, 2011).

“particle” to be, eventually the sequence will list tensor fields with support entirely confined in such a narrow region. The strategy then is to show that if such a sequence of fields can be constructed for a given curve γ , then γ must be a (timelike) geodesic, allowing one (roughly) to claim that “arbitrarily small particles must follow (timelike) geodesics.”

4.1 Geroch-Jang Particles

In contrast to the overabundance of 0^{th} -order derivations, the class of limit proofs consists primarily of two elegant results.⁴⁹ The first result by Geroch and Jang (1975) considers sequences of ostensible energy-momentum tensor fields of ever narrowing support. More precisely, Geroch and Jang’s theorem can be formulated as follows:

Theorem 1. (Geroch-Jang, 1975) *Let $\gamma : I \rightarrow \mathcal{M}$ be a smooth curve in Lorentzian spacetime (\mathcal{M}, g_{ab}) . Suppose that given any open neighborhood \mathcal{O} of $\gamma[I]$, there exists a smooth symmetric tensor field T_{ab} defined on \mathcal{M} such that for all points $p \in \mathcal{M}$:*

1. T_{ab} has non-vanishing support contained in \mathcal{O} ,
2. For all timelike ξ^a : $T_{ab}\xi^a\xi^b \geq 0$ and if $T_{ab} \neq \mathbf{0}$, then $(T_{ac}\xi^a)(T^c_b\xi^b) > 0$,
3. $\overset{g}{\nabla}_a T^a_b = \mathbf{0}$,

then $\gamma[I]$ is the image of a timelike g_{ab} -geodesic.

In order to illustrate the significance of these conditions, consider any set of nested neighborhoods $(\mathcal{O}_i)_{i \in \mathbb{N}}$ that becomes arbitrarily narrow around the curve γ as $i \rightarrow \infty$. Next consider a sequence of tensor fields $(T_{ab})_{i \in \mathbb{N}}$ such that each T_{ab} satisfies conditions 2 and 3, and for each i , T_{ab} satisfies condition 1 for the neighborhood \mathcal{O}_i . Let us refer to such a sequence of tensor fields $(T_{ab})_{i \in \mathbb{N}}$ as a *Geroch-Jang* or *GJ-particle*. The existence of a GJ-particle sequence for an arbitrary sequence of nested neighborhoods tightening around γ is equivalent to the satisfaction of the conditions of theorem 1.

Let us consider what each of the conditions says about GJ-particles. First, condition 1 establishes the arbitrary smallness characteristic of GJ-particles. Since the nested neighborhoods $(\mathcal{O}_i)_{i \in \mathbb{N}}$

⁴⁹The self-similarity limit operations done by Gralla and Wald (2008) can (in part) also be classified as an enhancement of these limit proof strategies, though they then proceed to employ some of the expansion techniques discussed in section 3. Hence, (Gralla and Wald, 2008) appears to constitute a kind of borderline case between the two families. The work done far earlier by Robertson (1937) might also be considered a kind of proto-limit operation proof attempt, in that he follows the general strategy of considering the limiting behavior of an extended “corpuscle” as the spatial extent goes to 0. In contrast, the limits taken by Infeld and Schild (1949) in the “IS-particle” constructions discussed in section 2.1 would determinately not count as a member of the family of limit operation proofs we are considering in that (aside from the ill-defined convergence issues already discussed) for each of the sequence entries the test body is represented by a singularity rather than a smooth tensor field representation of an *extended* object. This distinction between considering the limiting behavior of sequences of *extended* bodies and merely attempting to appeal to limits in the course of the demonstration is significant, and the former more restrictive characteristic is required for our present classification (cf. Havas (1989, p254) who seems to overlook the significance this distinction).

become arbitrarily narrow, no matter how tight around γ we demand that the world-tube of the particle be confined, T_{ab}^i will eventually (for sufficiently large i) stay that close (or closer). If we interpret the symmetric field T_{ab}^i as an energy-momentum tensor, then condition 2 says that from the perspective of any observer the matter-energy in that part of the universe (a) is non-negative and (b) only flows in timelike directions (it doesn't go as fast or faster than the speed of light). So, roughly speaking, the theorem is only telling us about the behavior of a certain kind of body that is of positive mass (as opposed for example to photons, which we don't expect to follow timelike geodesics anyway or the kind of body we might wish to associate, say, with non-classical negative energy solutions). Hence, such conditions restricting the kind of matter-energy that can constitute a GJ-particle appear appropriate for the sort of material body about which we expect the principle to be relevant. However, it is worth observing that condition 2 *does not follow from Einstein's field equations*, and hence constitutes an additional assumption that must be obeyed in order to get Geroch and Jang's geodesic result.⁵⁰

Condition 3 is the familiar conservation condition (2.3), which (as discussed in section 2.2) follows directly from Einstein's original field equations for all smooth solutions. Condition 3 is the primary reason why one might say that theorem 1 constitutes a "deduction" of the geodesic principle from Einstein's field equations. *If (g_{ab}, T_{ab}^i) were a solution to Einstein's field equations, then condition 3 would be automatically satisfied for T_{ab}^i .* This might allow us (roughly) to characterize Geroch and Jang's result by claiming that "arbitrarily small bodies of positive mass that obey Einstein's field equations must follow geodesics."⁵¹

Unfortunately, this claim reveals a significant complication. Assume that the T_{ab}^i represent the entire contribution to the source side of Einstein's field equations (1.1) in their respective neighborhoods of γ . In this case, they each should be having a non-zero perturbative effect (ϵ_{ab}^i) on the metric of the spacetime manifold. The problem is that since the support of the GJ-particle is non-vanishing but continually shrinking down in spacelike extent, this non-zero perturbative effect ϵ_{ab}^i will not be stable in every region of spacetime even for sufficiently large i . That is to say, if solutions to the

⁵⁰A recent discussion of this logical independence can be found in (Malament, 2009), along with a demonstration that the existence of an "almost" GJ-particle (viz. ones that satisfy the first and third but not the second condition) fails to ensure that γ is a geodesic. See also (Weatherall, 2011) in which it is shown that condition 2 (as opposed to a slightly weaker energy condition originally used by Geroch and Jang) is necessary.

⁵¹A nuance worth noting that is imposed by condition 3 involves the question of electromagnetic (or other non-gravitational) field effects. One might think that this phrasing of the principle is too strong: though *neutral* massive bodies are supposed to follow geodesics, charged bodies under the influence of an electromagnetic field should not. Of course, if there is an electromagnetic field to influence our GJ-particle it would supply further energy-momentum ($T_{ab}^{(EM)}$) in the neighborhoods of γ (indeed, electromagnetic energy-momentum that should not stay confined to arbitrary neighborhoods of γ). So if, for example, we were talking about a charged body ${}_i T_{ab}$ interacting with an electromagnetic field $T_{ab}^{(EM)}$ near γ , then Einstein's field equations only ensure that the *total* energy-momentum is conserved (e.g. ${}^g \nabla_c g^{ca} ({}_i T_{ab} + T_{ab}^{(EM)}) = \mathbf{0}$). This means that when we interpret the GJ-particle entries ${}_i T_{ab}$ as representing the matter-energy flow of (small, massive) bodies, condition 3 can be thought of as requiring that the bodies are "free" in the sense that their energy-momentum does not change due to interactions with other non-gravitational energy carrying fields in the neighborhood of γ . Hence, we might paraphrase the result even more appropriately by saying that "arbitrarily small free bodies of positive mass must follow geodesics." (But see next.)

respective T_{ab} of a GJ-particles take the form $g_{ab} = g_{ab} + \epsilon_{ab}$, the perturbation field ϵ_{ab} must vary for some larger i value. This means in particular that ϵ_{ab} cannot vanish for arbitrarily large i , so no matter how far out in the sequence we look the remaining T_{ab} 's will never all couple to the metric g_{ab} in accordance with Einstein's field equations. Hence, the constraints placed on the existence of a GJ-particle do not prevent differences between (a) the geodesic structure of a spacetime in which a GJ-particle entry obeys Einstein's field equations and (b) the geodesic structure of the background metric g_{ab} according to which γ actually counts as a geodesic. In other words, geodesics of g_{ab} will not necessarily remain geodesics of the spacetimes with GJ-particles in them. We can think of imposing a GJ-particle in neighborhoods of γ as having a kind of "bending" effect on γ , ruining its geodocity.

So while it is nice to know that GJ-particles must follow geodesics of *some* spacetime, strictly speaking *the theorem does not ensure that GJ-particles must follow geodesics of the spacetime(s) in which they might actually exist (at least not without violating Einstein's field equations)*. Though Geroch and Jang's theorem is appealing with respect to its mathematical elegance and certain aspects of its representational fertility, in this form it is guilty of relying on a test body approximation that is ultimately left unjustified by the conditions. That being said, reflection on the theorem leaves one with the sense that "if only the perturbative effect ϵ_{ab} could be controlled somehow as the GJ-particle shrinks down in size, then we might at least be able to say that the spacetime metrics g_{ab} coupling to our GJ-particle entries should 'come close' to the background metric g_{ab} ."⁵² In the next section, we will consider a second limit operation proof achieving just this sort of result.

4.2 Ehlers-Geroch Particles

In 2004, Ehlers, who had evidently been concerned by the Geroch-Jang "test body approximation" for nearly three decades (see note 7), collaborated with Geroch to develop a second result that accommodates for the "geodesic bending" effects of GJ-like particles. If Geroch and Jang's theorem approaches things from the "source side" of Einstein's field equations (1.1), then the Ehlers & Geroch (2004) can be said to approach things from the geometry side of the equations. Specifically, instead of considering sequences of energy-momentum tensors, they consider sequences of *metrics* which converge in an appropriate way to a background metric and whose energy-momentum sources simultaneously satisfy (essentially) the same conditions as those placed on GJ-particles. Their result can be formulated as follows:⁵³

⁵²Note, for any GJ-particle sequence $({}_i T_{ab})_{i \in \mathbb{N}}$, there exists a second GJ-particle sequence $({}_i \hat{T}_{ab})_{i \in \mathbb{N}}$ whose matter-energy density vanishes arbitrarily quickly as $i \rightarrow \infty$. (Just define ${}_i \hat{T}_{ab} := (\alpha_i)({}_i T_{ab})$ for each i where $(\alpha_i)_{i \in \mathbb{N}}$ is sequence of scalars converging to 0 suitably quickly.)

⁵³Note, the theorem as stated originally in (Ehlers and Geroch, 2004) is slightly stronger than the following version: their result still goes through if the strict inequality in condition 2 is weakened to allow for equality as well. Of course the theorem as stated is an immediate consequence of the slightly stronger version. The difference is inconsequential to our current discussion.

Theorem 2. (Ehlers-Geroch 2004) Let $\gamma : I \rightarrow \mathcal{M}$ be a smooth timelike curve in Lorentzian spacetime (\mathcal{M}, g_{ab}) . Suppose that for any sufficiently small closed neighborhood $\mathcal{K} \subset \mathcal{M}$ of $\gamma[I]$ there exists a sequence of smooth Lorentzian metrics g_{ab} defined on \mathcal{K} such that for all points $p \in \mathcal{K}$:

1. For all j : G_{ab} has non-vanishing support contained in the interior of \mathcal{K} ,
2. For all j and all timelike ξ^a : $G_{ab}\xi^a\xi^b \geq 0$ and if $G_{ab} \neq \mathbf{0}$, then $g^{bd} \left(G_{ab}\xi^a \right) \left(G_{cd}\xi^c \right) > 0$,
3. The $g_{ab} \rightarrow g_{ab}$ as metrics in $\mathcal{C}^1(\mathcal{K})$ as $j \rightarrow \infty$,⁵⁴

where G_{ab} is the Einstein curvature tensor determined by g_{ab} , then $\gamma[I]$ is the image of a g_{ab} -geodesic.

As with GJ-particles, we might illustrate the content of the theorem by considering an arbitrary set of nested neighborhoods $(\mathcal{K}_i)_{i \in \mathbb{N}}$ converging down around $\gamma[I]$. For these neighborhoods, we can now consider the double indexed sequences of smooth metrics $(g_{ab})_{i,j \in \mathbb{N}}$ and corresponding curvature tensors $(G_{ab})_{i,j \in \mathbb{N}}$ defined for each i on \mathcal{K}_i . The latter sequence of curvature tensors $(G_{ab})_{i,j \in \mathbb{N}}$ can then be identified via Einstein's equations as a sequence of energy-momentum tensors $(T_{ab})_{i,j \in \mathbb{N}}$ which we will call *Ehlers-Geroch* or *EG-particles*. Observe, it follows from conditions 1 and 2 in theorem 2, that each T_{ab} satisfies conditions 1 and 2 of theorem 1 in (\mathcal{K}_i, g_{ab}) . Moreover, because each T_{ab} is equal to a smooth Einstein curvature tensor, they automatically satisfy condition 3 of theorem 1 in (\mathcal{K}_i, g_{ab}) as well. Hence, EG-particles are quite similar to GJ-particles, the difference with EG-particles is that the further condition 3 ensures that the perturbative effect of EG-particles on the “background metric” (g_{ab}) can be made *arbitrarily small* (in the relevant senses) for sufficiently large j . Most importantly, the convergence demanded in condition 3 ensures not only that the γ is a geodesic of the background metric g_{ab} , but also for sufficiently large j , γ will come arbitrarily close to being a timelike geodesic with respect to the “perturbed” metrics g_{ab} . So not only can EG-particles (like their GJ-particle cousins) be made arbitrarily small *in (spacelike) extent* around γ for sufficiently large i , unlike GJ-particles by picking sufficiently large j we can also control the “geodesic bending” effects resulting from their presence, ensuring that the EG-particle will be shrunken down around a curve that will come arbitrarily close to actually being a geodesic.

⁵⁴The $\mathcal{C}^1(\mathcal{K})$ topology $\tau_{\mathcal{C}^1(\mathcal{K})}$ is defined on the space of ordered pairs of symmetric tensor fields and covariant derivative operators defined for the closed region \mathcal{K} . $\tau_{\mathcal{C}^1(\mathcal{K})}$ consists of point-wise neighborhoods of the tensor fields and connections respectively, varying continuously with p but otherwise arbitrarily in the respective spaces (this can be done explicitly, for example, with the selection of arbitrary positive definite metrics defined on \mathcal{K}). We say that ${}_j g_{ab} \rightarrow g_{ab}$ as metrics in $\mathcal{C}^1(\mathcal{K})$ if for every neighborhood $\mathcal{N} \in \tau_{\mathcal{C}^1(\mathcal{K})}$ of (g_{ab}, ∇_a) , for sufficiently large j we have $({}_j g_{ab}, \nabla_a^{(j)}) \in \mathcal{N}$ where ∇_a and $\nabla_a^{(j)}$ are the unique derivative operators compatible with their respective metrics. Because these operators are uniquely determined by their metrics, explicit reference to them can be suppressed in the articulation of the theorem.

Hence, Ehlers and Geroch’s result can be characterized by the claim that “arbitrarily small bodies of positive mass come arbitrarily close to following geodesics.”

Unfortunately for the canonical view, the theorem cannot ensure the actual geodicity of γ in the presence of any massive body obeying the field equations. Though by “turning up” the i and j indices so to speak, we can make the EG-particle both as narrow and “close to straight” as we want, *we can never ensure actual geodicity for any finite j* . Actual geodicity is only achieved *at the limit* in the spacetime with the metric g_{ab} . Ehlers and Geroch’s theorem ensures near geodicity in the *approach to the limit*, but if in accordance with the canonical view we are looking to ensure massive bodies following actual geodesics, theorem 2 does not do the trick.

An ardent defender of the canonical view might attempt to get around the fact that geodicity is not acquired for $j < \infty$ by focusing on “the limit case” directly. One problem with this is that though condition 3 establishes convergence of the metrics and an approach to the geodicity of γ , it is insufficient for the convergence of EG-particles to a “limit” energy-momentum field.⁵⁵ In general, there is no energy-momentum limit of an EG-particle. In fact, the only way to get convergence of EG-particles in a way that does not allow for new energy-momentum to suddenly appear at the limit (but not before) and likewise does not violate Einstein’s field equations at the limit is by having the EG-particle converge to a tensor field that vanishes around γ .⁵⁶ In other words, EG-particles don’t generally converge, but even the ones that might converge (at least in any acceptable way) either violate Einstein’s field equations or vanish. Again the canonical view is left unable to

⁵⁵Ehlers and Geroch consider an explicit counterexample sequence of spacetimes whose metrics converge according to condition 3, but whose associated curvature tensors (and so energy-momentum tensors) become divergent. The reason for this possibility is that curvature tensors involve not just derivatives of the metric but also of the connection (see equations (C.1) and (C.2) below), but convergence in the \mathcal{C}^1 topology only ensures closeness of the metric and the connection, but not higher derivatives. They note that ruling out such examples would involve strengthening condition 3 to require convergence in a more restrictive “ $\mathcal{C}^2(\mathcal{K})$ topology.”

⁵⁶To see why this is the case, let \mathcal{K} be an arbitrary sufficiently small closed neighborhood of γ from theorem 2. If we want to preserve coherence with Einstein’s field equations, then any “converging” sequence of energy-momentum tensors $({}_jT_{ab})_{j \in \mathbb{N}}$ coupled to the metrics ${}_jg_{ab}$ defined for \mathcal{K} as in theorem 2 will have to converge to some energy-momentum tensor T_{ab} equal to the Einstein curvature tensor G_{ab} determined by the background metric g_{ab} . (Otherwise, it would be the case that the energy-momentum “in the limit” fails to couple to the limiting metric in accordance with equations (1.1).) Now let τ be any topology on the space of rank $(0, 2)$ -tensor fields with respect to which it might be claimed that ${}_jT_{ab} \rightarrow T_{ab}$ as $j \rightarrow \infty$. Of course there are numerous topologies with respect to which this might be claimed; some may be physically appropriate and others may not. Luckily, we need not determine here which particular topology (if any) is in fact most appropriate. Instead we will only require that any relevant convergence must at least ensure the following condition for all $p \in \mathcal{K}$:

Vacuum-point preservation: If there exists a $j_0 \in \mathbb{N}$ such that for all $j > j_0$, $T_{ab}|_p = \mathbf{0}$, then $T_{ab}|_p = \mathbf{0}$.

The vacuum-point preservation condition should strike us as a reasonable restriction on any τ -convergence in this context since it only precludes the sudden appearance of “new energy-momentum” at the limit that wasn’t already present in the approach as $j \rightarrow \infty$. (Vacuum-point preservation would be obeyed, for instance, if we wanted to focus our discussion specifically on convergence in $\mathcal{C}^2(\mathcal{K})$.) Now from condition 1 of theorem 2, we know that for every j , ${}_jT_{ab} = {}_jG_{ab} = \mathbf{0}$ on $\partial\mathcal{K}$. Hence, if the limit is vacuum-point preserving we have that $T_{ab}|_{\partial\mathcal{K}} = \mathbf{0}$. But \mathcal{K} was an arbitrary sufficiently small neighborhood of γ , which means that T_{ab} will vanish on the boundary of *every* sufficiently small neighborhood of γ . Moreover, since g_{ab} is smooth, we know that $T_{ab} = G_{ab}$ must be smooth. It follows from these two facts that T_{ab} must vanish in some neighborhood of γ . Hence, the only way to ensure that the energy-momentum tensors coupling to the sequences $({}_jg_{ab})_{j \in \mathbb{N}}$ defined in theorem 2 converge in a vacuum-point preserving way without violating Einstein’s field equations is by having them vanish around γ .

establish actual geodesic motion for massive bodies. Though EG-particles can be said to “come close” to exhibiting geodesic motion, the only way to establish actual geodcity is by violating the field equations or having energy-momentum of the particle completely disappear. As in the case of 0^{th} -order proofs, though the proofs can establish a kind of “approximation” to geodesic following in the case of extended bodies with matter-energy, when it comes to establishing genuine geodcity, both fall short, achieving such strict results (at best) only for inapposite or pathologically idealized special cases.

5 Conclusion: Towards a Geodesic Universality Thesis

In this paper I have argued that the canonical view that the geodesic principle provides the dynamics of general relativity theory fails. Under this interpretation, the commonly endorsed belief that the principle can be derived either from Einstein’s original field equations or a distributional generalization of them must be rejected (even if we allow for further background assumptions about the kind of matter-energy that is supposed to follow such geodesics). By reviewing the three major classes of proof, we have seen that would-be geodesic following bodies are forced either **(i)** to meet unrealistically restrictive special-case conditions, **(ii)** to have no matter-energy at all (i.e. vanish), **(iii)** to violate Einstein’s field equations, or **(iv)** to be located on “paths” that don’t just fail to be geodesic but fail to exist in the spacetime manifold at all.

Such arguments reveal that the claim that “massive bodies follow geodesics in Einstein’s theory” cannot be accepted as any sort of precise articulation of the actual motion of bodies in general relativity. However, this does not mean that there is no place for the principle. Instead, it is my suggestion that we weaken our interpretation of the geodesic principle, modestly “demoting” it from the status of dynamical principle governing the paths of massive bodies in Einstein’s theory to a less fundamental *geodesic universality thesis* to be interpreted in analogy with the kind of universality exhibited, for instance, in thermal systems during phase transitions. Though a comprehensive argument for this thesis will have to wait for a sequel to the current paper, let us briefly sketch what role such a less fundamental interpretation of the principle would serve, and how we might justify such an interpretation.

5.1 A New Place for the Principle?

The geodesic principle cannot accurately provide an account of the motion of massive bodies consistent with Einstein’s equations. Even in the case of “arbitrarily small” bodies, the principle fails to account precisely for the dynamics of general relativity theory. Moreover, if what we are interested in the precise evolution of bodies over time (viz. in the future domain of dependence of a neighborhood of a set of bodies on a particular folium), we have the initial value formulation. However, the availability of such a precise tool for determining the dynamical evolution of every bit of matter-energy

in a certain region may not always be optimal. As discussed in section 3 techniques for determining the evolution of bodies in their future domain of dependence can be less than pragmatic. Moreover, we generally do not have access to the appropriate initial data information that might be used in such formulations (e.g. it is difficult to ever attain precise information about the flow of matter-energy in the interior of Mercury). Recently Gralla and Wald (2008, p2) have highlighted some of the complications of relying exclusively on initial value formulations as follows:

The physical content of general relativity is contained in Einstein's equation, which has a well posed initial value formulation (see, e.g., [(Wald, 1984)]). 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.

Gralla and Wald's final point is particularly germane to answering the question of whether there is a proper place for the geodesic principle: Even if we do have access to sufficient initial data, and we were able to bound the pragmatic hurdles (numerically if not analytically), having the precise details about what every piece of matter-energy is doing at every single point in spacetime runs the risk of obscuring what we should expect of the general behavior of bodies under gravitational influences. Perhaps counterintuitively, by taking a step back from the deluge of details that we might receive from an initial value solution and instead searching for general approximations of the motion, we might actually capture further knowledge concerning the "basic general features" of these solutions. Such approximations of motion of course should lose something in the way of accounting for the precise evolution of a particular gravitating body. Nonetheless, if we wish to understand, not the precise evolution, but the broad characteristics of the paths of general gravitating bodies, approximate descriptions may in this respect have more potency than precise ones.

The geodesic principle understood not as a precise description of the evolution of massive bodies, but instead taken as a formula for characterizing the *approximate* behavior of (appropriately small) massive bodies in general may still offer important opportunities for insight into relativistic dynamics. Remarkably, if we wish to understand the principle as playing this role of characterizing this kind of general behavior, then limit operation proofs such as (Ehlers and Geroch, 2004) in particular help to provide a substantial justification for why we should expect such "approximate geodicity" of their paths. Recall, though theorem 2 was unable to establish perfect geodesic evolution of (actually massive) bodies, the interpretation of the theorem's significance offered above was able to establish that "*appropriately small* massive bodies will follow timelike paths that are *almost* geodesic." So

while such a result fails to tell us about the paths with absolute precision, it does enable us to draw the broad inference that large classes of bodies can be expected to stick “close to” a geodesic path for suitable time scales. If understood appropriately, being able to draw this kind of inference for such general classes of bodies without having to know about their exact constitution offers great opportunity for understanding about gravitation and gravitational dynamics despite the lack of attention to details at a certain level of precision. This insight can be achieved despite the fact that the geodesic principle fails to provide the kind of fundamental dynamics-defining role expected by the canonical interpretation.

Moreover, the kind of near-geodesic clustering of massive bodies that seems to be explained by results such as (Ehlers and Geroch, 2004) is well confirmed by our experimental observations.⁵⁷ Planetary bodies, whose relative size and gravitational effect compared to the sun is suitably small, exhibit near geodesic behavior. In particular, the Mercury confirmation shows that this clustering can be confirmed in relativistic regimes. Though a precise enough experiment should reveal divergence from perfect geodesic motion, once we have weakened the principle to a claim about the near-geodesic clustering of gravitating free bodies in general, the kind of subtle wobbling we might expect given the discussion in section 3 need not count as disconfirmation. Instead, a vast number of examples from Newton’s apple to gravitation on astronomical scales now constitute opportunities to confirm such general near-geodesic clustering.

5.2 Geodesic Universality

The suggestion considered above that the geodesic principle might be reinterpreted as a characterization of the general patterns of behavior of (small) gravitating bodies, despite significant possible variations in details of how the bodies are constituted or the type of external gravitational field they might be exposed to, is analogous to a prominent classification of certain phenomena studied in other fields of physics. Referred to as *universality phenomena*, such clustering patterns across multiple systems are studied most prominently in characterizing the similarities in behavior exhibited by thermal systems during phase transitions and near criticality. Kadanoff (2000, p225), often identified as one of the first to apply this concept in its contemporary sense in physics, has defined ‘universality’ as applying to patterns in which “[m]any physically different systems show the same behavior.” In the study of critical behavior, for instance, the phenomena is identified when numerous systems seem to cluster into what are called *universality classes*, despite possible vast disparities in the fundamental details characterizing members of a single class.

Robert Batterman who has done remarkable work in introducing universality phenomena to the philosophical literature argues that “[w]hile most discussions of universality and its explanation take

⁵⁷In fact, as long as we do not expect perfect geodicty, results from more sophisticated “0th-order type” proofs (e.g. those in (Gralla and Wald, 2008, Pound, 2010)) can be used to identify the appropriate regime scales for which we might expect such observed clustering.

place in the context of thermodynamics and statistical mechanics,... universal behavior is really ubiquitous in science” (Batterman, 2002). This claim has no shortage of vindicating examples: Universality identified in criticality models has been used to characterize the clustering behavior in astonishingly diverse contexts from avalanche and earthquake modeling (Kadanoff et al., 1989, Lise and Paczuski, 2001), to extinction modeling in population genetics (Sole and Manrubia, 1996), to modeling belief propagation in multi-agent networks (Glinton et al., 2007, 2010). Batterman has identified many examples of universality phenomena distinct from criticality models as well, including patterns in rainbow formation, semi-classical approximation, and drop breaking (Batterman, 2002, 2005, 2006, 2009).

It is my suggestion that the kind of near-geodesic clustering discussed above constitutes yet another example of (non-thermal) universality in physics. Though no free body of realistic matter-energy should be expected to exactly follow timelike geodesics, limit operation proofs establish that in general the world-tubes of bodies of sufficiently small relative size and gravitational effect can be expected to cluster around timelike geodesic curves despite otherwise potentially quite significant variations in their particular energy-momenta. Further argument is of course called for before the geodesic universality interpretation can be fully adopted.⁵⁸ However, such a thesis offers a promising avenue for redeploing the geodesic principle in a way that respects the positive results of a number of the geodesic “demonstrations,” which we have seen otherwise fall short of establishing the principle under its canonical construal. Most importantly, retaining the principle in the form of a universality thesis permits us to draw inferences and improve our understanding about the paths of general gravitating bodies despite their lack of perfect geodicity.

Appendix

A Tensor Distributions

Let \mathcal{M} be an orientable n -dimensional smooth manifold. The space $\mathcal{D}'(\mathcal{M})$ of *scalar distributions* on \mathcal{M} can be defined as the linear dual to the (LF-)space $\Omega_c^n(\mathcal{M})$ of smooth, compactly supported n -forms on \mathcal{M} . The space $\Omega_c^n(\mathcal{M})$ plays the role of our *test fields* in a differential geometry context in that (since \mathcal{M} is orientable) these test fields can essentially be thought of as products of smooth, compactly supported scalar fields (i.e. the test functions of typical distribution theory) and an arbitrary volume element $\epsilon := \epsilon_{abcd} \in \Omega^n(\mathcal{M})$.

This construction can be generalized to define linear spaces $\mathcal{D}'_r^s(\mathcal{M})$ of *tensor distributions of rank* (r, s) on \mathcal{M} as the dual of the space $\mathcal{T}_s^r(\mathcal{M}) \otimes \Omega_c^n(\mathcal{M})$ of *test tensor fields* consisting of exterior products of smooth tensors of rank (s, r) and compactly supported n -forms. Each element of $\mathcal{T}_s^r(\mathcal{M})$

⁵⁸It is worth emphasizing that in light of the arguments presented here, the geodesic universality thesis would need to *reject* the claim that actual gravitating bodies need ever follow geodesics precisely. Hence, a comprehensive defense of the thesis will involve explaining why this is compatible with the clustering behavior which it does affirm.

defines a mapping (via contraction) from the space $\mathcal{D}'_r(\mathcal{M})$ to the space $\mathcal{D}'(\mathcal{M})$. In fact, the space $\mathcal{D}'_r(\mathcal{M})$ is isomorphic (as a $\mathcal{C}^\infty(\mathcal{M})$ module) to the space of products of smooth tensors of rank (s,r) with elements of $\mathcal{D}'(\mathcal{M})$ (Grosser et al., 2001, Thm 3.1.15).

Hence, tensor distributions of rank (s,r) can be intuitively thought of as familiar smooth tensor fields of the same rank with scalar distributions as their coefficients. Analogous to the case in scalar distribution theory, a locally integrable tensor field $\alpha_{b_1 \dots b_r}^{a_1 \dots a_s}$ (not necessarily smooth) has a natural embedding in $\mathcal{D}'_r(\mathcal{M})$, where the action $\langle \alpha_{b_1 \dots b_r}^{a_1 \dots a_s}, \cdot \rangle$ on a test tensor field $\Phi_{a_1 \dots a_s}^{b_1 \dots b_r}$ is given by:

$$\langle \alpha_{b_1 \dots b_r}^{a_1 \dots a_s}, \Phi_{a_1 \dots a_s}^{b_1 \dots b_r} \rangle = \int_{\mathcal{M}} \alpha_{b_1 \dots b_r}^{a_1 \dots a_s} \Phi_{a_1 \dots a_s}^{b_1 \dots b_r}$$

It is this embedding that suggests that the action of distributions on test objects is like that of integrating the contraction of the tensor distribution with a test tensor field. The support of tensor distributions is likewise extended in the following way: a tensor distribution $\langle \alpha_{b_1 \dots b_r}^{a_1 \dots a_s}, \cdot \rangle$ is said to have *support on* \mathcal{K} if all test fields with support disjoint from \mathcal{K} are in the kernel of $\langle \alpha_{b_1 \dots b_r}^{a_1 \dots a_s}, \cdot \rangle$.

If ∇_b is any smooth derivative operator, the derivative of a tensor distribution $\alpha_{b_1 \dots b_r}^{a_1 \dots a_s} \in \mathcal{D}'_r(\mathcal{M})$ is a distribution $\nabla_a \alpha_{b_1 \dots b_r}^{a_1 \dots a_s} \in \mathcal{D}'_{r+1}(\mathcal{M})$ whose action is defined by:

$$\langle \nabla_b \alpha_{b_1 \dots b_r}^{a_1 \dots a_s}, \Phi_{a_1 \dots a_s}^{b_1 \dots b_r b} \rangle = - \langle \alpha_{b_1 \dots b_r}^{a_1 \dots a_s}, \nabla_b \Phi_{a_1 \dots a_s}^{b_1 \dots b_r b} \rangle \quad \forall \Phi_{a_1 \dots a_s}^{b_1 \dots b_r b} \in \mathcal{T}_s^{r+1}(\mathcal{M}) \otimes \Omega_c^n(\mathcal{M})$$

In the case that $\alpha_{b_1 \dots b_r}^{a_1 \dots a_s}$ is a locally integrable tensor field (not necessarily differentiable in the classical sense), and there exists a second locally integrable tensor field $\beta_{b_1 \dots b_r b}^{a_1 \dots a_s}$ such that

$$\langle \beta_{b_1 \dots b_r b}^{a_1 \dots a_s}, \Phi_{a_1 \dots a_s}^{b_1 \dots b_r b} \rangle = - \langle \alpha_{b_1 \dots b_r}^{a_1 \dots a_s}, \nabla_b \Phi_{a_1 \dots a_s}^{b_1 \dots b_r b} \rangle \quad \forall \Phi_{a_1 \dots a_s}^{b_1 \dots b_r b} \in \mathcal{T}_s^{r+1}(\mathcal{M}) \otimes \Omega_c^n(\mathcal{M})$$

then $\beta_{b_1 \dots b_r b}^{a_1 \dots a_s}$ is said to be the *weak derivative* of the tensor field $\alpha_{b_1 \dots b_r}^{a_1 \dots a_s}$.

Elements in the linear spaces $\mathcal{D}'_r(\mathcal{M})$ *do not* have a well defined product structure, and so unlike smooth tensors, they do not constitute an algebra. As a consequence, we can only consider exterior products and contractions of tensor distributions with non-distributional tensor fields.

B Proof from Generalized Conservation of T^{ab}

In order to represent a “point” particle by means of an energy-momentum tensor distribution T^{ab} , it will be useful to define the following scalar distribution in the space $\mathcal{D}'(\mathcal{M})$.

Definition 3. If $\gamma : I \rightarrow \mathcal{M}$ is a smooth curve in the spacetime (\mathcal{M}, g_{ab}) then we will refer to the linear mapping $\mathcal{D}_{(\gamma, g)} : \mathcal{C}_c^\infty(\mathcal{M}) \rightarrow \mathbb{R}$ given by following action on test functions:

$$\int_{\mathcal{M}} \mathcal{D}_{(\gamma,g)} \phi \text{vol}_g \mapsto \int_I \phi \circ \gamma \, ds \quad \forall \phi \in \mathcal{C}_0^\infty(\mathcal{M}) \quad (\text{B.1})$$

as the *concentrating distribution* for γ in spacetime (\mathcal{M}, g_{ab}) .⁵⁹

Proposition 4. *Let (\mathcal{M}, g_{ab}) be a Lorentzian spacetime, and let $\gamma : I \rightarrow \mathcal{M}$ be a smooth timelike curve in \mathcal{M} for some interval I . Then, if there exists a smooth symmetric tensor field T_γ^{ab} defined on \mathcal{M} and non-vanishing on $\gamma[I]$ such that,*

$$\int_{\mathcal{M}} \left(\mathcal{D}_{(\gamma,g)} T_\gamma^{ab} \right) \overset{g}{\nabla}_b \xi_a \text{vol}_g = 0 \quad \forall \xi_a \in (\mathcal{T}_1^0(\mathcal{M}))_c$$

where $(\mathcal{T}_1^0(\mathcal{M}))_c$ is the space of smooth co-vector fields on \mathcal{M} with compact support and $\mathcal{D}_{(\gamma,g)}$ is the concentrating distribution for γ in (\mathcal{M}, g_{ab}) , then $\gamma[I]$ is the image of a geodesic of g_{ab} .

Proof: Setting $\phi = T_\gamma^{ab} \overset{g}{\nabla}_a \xi_b$ for arbitrary $\xi_a \in (\mathcal{T}_1^0(\mathcal{M}))_c$, it follows from definition 3 that

$$\int_I T_\gamma^{ab} \overset{g}{\nabla}_b \xi_a \, ds = 0 \quad \forall \xi_a \in (\mathcal{T}_1^0(\mathcal{M}))_c \quad (\text{B.2})$$

Let $\mathcal{K}(\gamma)$ be the set of smooth functions on \mathcal{M} that vanish on $\gamma[I]$. Clearly for any $\xi_a \in (\mathcal{T}_1^0(\mathcal{M}))_c$ and $\alpha \in \mathcal{K}(\gamma)$ we have that $\alpha \xi_b \in (\mathcal{T}_1^0(\mathcal{M}))_c$ giving us the following:

$$\int_I T_\gamma^{ab} \overset{g}{\nabla}_b (\alpha \xi_a) \, ds = 0 \quad \forall \alpha \in \mathcal{K}(\gamma), \xi_a \in (\mathcal{T}_1^0(\mathcal{M}))_c \quad (\text{B.3})$$

And since α vanishes on $\gamma[I]$, (B.3) reduces to:

$$\int_I T_\gamma^{ab} \xi_a \overset{g}{\nabla}_b \alpha \, ds = 0 \quad \forall \alpha \in \mathcal{K}(\gamma), \forall \xi_a \in (\mathcal{T}_1^0(\mathcal{M}))_c \quad (\text{B.4})$$

We now observe that (B.4) holds (if and) only if for each $\xi_a \in (\mathcal{T}_1^0(\mathcal{M}))_c$ there exists a smooth scalar field ψ_ξ with compact support in $\gamma[I]$ such that:⁶⁰

$$T_\gamma^{ba} \xi_b = \psi_\xi U^a \quad \forall p \in \gamma[I] \quad (\text{B.5})$$

⁵⁹For any $\phi \in \mathcal{C}_c^\infty(\mathcal{M})$, the set of “test-function weighted” volume elements ϕvol_g is equivalent to the space of test 4-forms $\Omega_c^4(\mathcal{M})$. So though we have defined the action of $\mathcal{D}_{(\gamma,g)}$ relative to its action on test functions $\mathcal{C}_c^\infty(\mathcal{M})$ definition 3 clearly gives a well defined element of the space $\mathcal{D}'(\mathcal{M})$ constructed in appendix A. Of course despite this equivocation, it is worth observing that the action of $\mathcal{D}_{(\gamma,g)}$ does depend on the particular g_{ab} as well as γ .

⁶⁰The only if direction is satisfied by assuming for contradiction that for some ξ_a the vector $\gamma T^{ab} \xi_a$ is not proportional to the tangent vector to γ at some point $p_0 \in \gamma[I]$. Since $\gamma T^{ab} \xi_a$ is smooth this means that for some sub-interval $I_0 \subset I$ such that $p_0 \in \gamma[I_0]$, $\gamma T^{ab} \xi_a$ will not be proportional to the tangent vector. We now select an α which is positive at all points in a sufficiently small neighborhood of $\gamma[I_0]$ save those points on $\gamma[I_0]$ and vanishing everywhere else to give us a non-zero value for the integral $\int T^{ab} \xi_a \nabla_b \alpha \, ds$ in violation of (B.4).

where U^a is the unit tangent vector to the curve γ (recall γ is timelike). Hence, since this holds for arbitrary ξ_b , on $\gamma[I]$ we have that T_γ^{ab} must take the form:

$$T_\gamma^{ab} = U^a P^b \quad (\text{B.6})$$

for some smooth vector field P^a defined on $\gamma[I]$. Moreover, since $T_\gamma^{[ab]} = U^{[a} P^{b]} = \mathbf{0}$ on $\gamma[I]$, contracting with U_b entails that there exists a smooth scalar field $m = U_a P^a$ defined on $\gamma[I]$ such that:

$$P^a = m U^a \quad (\text{B.7})$$

So substituting into (B.2) and conducting an integration by parts we get:

$$\int_I U^b \overset{g}{\nabla}_b m U^a \xi_a ds - \int_I \xi_a U^b \overset{g}{\nabla}_b m U^a ds = 0 \quad \forall \xi_a \in (\mathcal{T}_1^0(\mathcal{M}))_c$$

In particular for all ξ_a compact on $\gamma[I]$, the first term vanishes and by arbitrarily varying these ξ_a with compact support on $\gamma[I]$, it follows from the second term that

$$U^b \overset{g}{\nabla}_b m U^a = 0 \quad (\text{B.8})$$

Last, the first integral of equation (B.8) gives us that the value m^2 (and so m) is constant along γ . So since $T_\gamma^{ab} \neq \mathbf{0}$ on the curve, we have that m is a non-zero constant on γ and the geodesic equation follows immediately from (B.8):

$$U^b \overset{g}{\nabla}_b U^a = 0 \quad (\text{B.9})$$

Hence, $\gamma[I]$ is the image of a g -geodesic.

□

C GT-regular and Semi-regular Metrics

In order for a tensor distribution source such as T_{ab} to be well defined as a distribution, it must be locally integrable.⁶¹ So since we want g_{ab} to be a solution to Einstein's generalized field equations (2.1), Geroch and Traschen tailor their class of GT-regular metrics by first looking at how the Einstein tensor, equated (as a distribution) to T_{ab} , depends on the metric and then considering what

⁶¹A tensor distribution $\alpha_{b_1 \dots b_s}^{a_1 \dots a_r} \in \mathcal{D}'_s(\mathcal{M})$ is said to *locally integrable* or in L^1_{loc} when scalar densities of the form $\alpha_{b_1 \dots b_s}^{a_1 \dots a_r} \Phi_{a_1 \dots a_r b c d}$ are Lebesgue measurable and integrable for arbitrary $\Phi \in \mathcal{T}_r^s(\mathcal{M}) \otimes \Omega_c^4(\mathcal{M})$. Similarly $\alpha_{b_1 \dots b_s}^{a_1 \dots a_r}$ will be said to be *locally square integrable* or in L^2_{loc} when $\alpha_{b_1 \dots b_s}^{a_1 \dots a_r} \alpha_{b'_1 \dots b'_s}^{a'_1 \dots a'_r}$ is locally integrable (and so on for elements in L^p_{loc}).

integrability properties the metric must satisfy in order to achieve integrability of the curvature. Specifically, let $\tilde{\nabla}_a$ be any smooth derivative operator with Riemann curvature $\tilde{R}_{abc}{}^d$. Now consider the dependence of another Riemann curvature tensor on an arbitrary metric g_{ab} (not necessarily smooth) in terms of $\tilde{\nabla}_a$:

$$R_{abc}{}^d = \tilde{R}_{abc}{}^d + 2C_{e[b}^d C_{a]c}^e + 2\tilde{\nabla}_{[b} C_{a]c}^d \quad (\text{C.1})$$

$$C_{bc}^a = g^{ae} \left(\tilde{\nabla}_{(b} g_{c)e} - \frac{1}{2} \tilde{\nabla}_e g_{bc} \right) \quad (\text{C.2})$$

Inspection of (C.1) reveals that $R_{abc}{}^d$ will be locally integrable if the tensor C_{bc}^a is locally square integrable.⁶² Moreover, (C.2) reveals that C_{bc}^a will be locally (square) integrable if g^{ab} is locally bounded and the weak derivative of g_{ab} exists and is locally (square) integrable. We hence have the following class of metrics:⁶³

Definition 5. (GT-regular metrics) A symmetric tensor field g_{ab} defined on \mathcal{M} is called a *GT-regular metric* if g_{ab} and g^{ab} are both in $L_{loc}^\infty \cap H_{loc}^1$.

In this definition L_{loc}^∞ is the space of locally bounded fields, and H_{loc}^1 is the Sobolev space of square integrable fields, whose weak first derivatives exist and are also square integrable. Hence, membership in the class of GT-regular metrics suffices for having a well defined Einstein tensor distribution.

The first nuance to note about this class is that though these metrics are *sufficient* for well defining curvature tensors as distributions, Geroch and Traschen's restrictions can be weakened a bit more. That is to say, we do not necessarily need the (weak) derivative of the metric to be square integrable, but only that the tensor fields C_{bc}^a and $C_{e[b}^d C_{a]c}^e$ (i.e. the *contraction* not the exterior product) exist and are locally integrable (though we still need g_{ab} and g^{ab} to be defined *almost* everywhere and be essentially bounded). Such a (strictly) wider class of metrics are referred to as *semi-regular* or *Garfinkle metrics* after his investigation in (Garfinkle, 1999). In contrast to GT-regular metrics, there do exist semi-regular metrics whose Einstein curvature tensor distribution can be concentrated on submanifolds of codimension 2.

Unfortunately, as previously observed in (Geroch and Traschen, 1987), even this meager weakening to semi-regularity faces representational complications. The reason for this has to do with why GT-regular metrics are so appropriately termed *regular*. A second important set of results proven

⁶²This condition directly suffices for the second term. Moreover, since $L_{loc}^2 \subset L_{loc}^1$ and the last term will be locally integrable if C_{bc}^a is locally integrable it also suffices for the final term. Since all smooth tensor fields are locally integrable, the first term is locally integrable without any further condition.

⁶³It should be noted that since GT-regular metrics are not in general Lipschitz in their first derivative, integral curves of "geodesic fields" U^a satisfying the condition $U^a \nabla_a U^b = 0$ will not always exist (or be uniquely determined for an initial value $U^a(p_0)$). In other words, geodesic curves will not always be well defined, particularly across regions of singular curvature.

by Geroch and Traschen (1987, Thm. 2-3), was that Cauchy sequences of regular metrics not only converge to a regular metric, but their respective curvature tensors converge to the curvature tensor of the limiting metric. In contrast, when we move to semi-regular metrics, this property is lost. This means that though we can define the action of curvature tensor distributions in semi-regular cases with support in less than three dimensions, such cases cease to have a natural interpretation as an extension of the classical framework of relativity theory.⁶⁴

As a final remark, it is worth noting that Geroch and Traschen’s proof strategy crucially depends on the required square integrability of the connection. It is because of this dependance (in part) that an analogous theorem preventing the existence of solutions with one-dimensional concentrations of energy-momentum (or mass-momentum) cannot be reconstructed for linearized approximations of Einstein’s equations (or for Newtonian gravitation). This means that (in a sense) it is thanks to the *non*-linearity of Einstein’s field equations that we are unable to coherently represent point particles in general relativity theory. In other words, *it is the non-linearity that precludes the possibility of using distribution proofs to deduce the geodesic hypothesis (in its most literal form) from the exact field equations*. Since historically the ability to deduce the geodesic principle from the field equations in the canonical account has typically been attributed to the fact that Einstein’s equations are non-linear, it is not without irony that this non-linearity is what stands in the way of the most literal variety of geodesic deduction.

⁶⁴ Recently in (Steinbauer and Vickers, 2006, Steinbauer, 2007, Steinbauer and Vickers, 2009) the authors have worked to generalize Einstein’s original equations even more than equation (2.1) in order to allow for solutions from non-linear tensor algebras (for clarity I will refer to elements of these algebras as *generalized tensors*) that can make sense of non-regular metric solutions such as GT-irregular Garfinkle metrics. Assuming this generalization project will come to fruition, such an end run around the Geroch-Traschen result would still fall short of vindicating the canonical view. The basic idea has to do with observing that actual material bodies have spacelike extent. That is to say massive bodies, even really small bodies or “atomic” constituents aren’t true points. This fact is germane to representations by means of tensor distributions and generalized tensors alike. Typically, physicists are able to avoid this problem when making use of distributional objects by arguing that objects of *very small* extent are “well approximated” by point particle representations through tensor distributions. But as we saw in section 3, any spacelike extent will generically molest the result of perfect geodesic motion expected of the canonical view (cf. Butterfield’s “atomism” thesis regarding limits of the arbitrary small in (Butterfield, 2011)).

The point is further punctuated in the case of generalized tensor algebras. Tensor distributions are embedded into these algebras by a process called *association* where it is shown that integrals of the infinite sequences constituting generalized tensors converge (in a specified way) to the action of a tensor distribution. Though a detailed investigation of this procedure will have to wait for further works, we may heuristically appreciate the complications of physically interpreting generalized tensors associated with one-dimensionally supported tensor distributions by observing that it is quite possible for such an association relation to exist even if every element of the sequence constituting the generalized tensor has support that extends outside the one-dimensional curve. So even if one-dimensionally supported sources could be associated with solutions to some such generalized field equations, a proponent of the canonical view would face interpretive challenges similar to the ones she faced in our discussion of EG-particles. That is to say, when it comes to interpreting the physical significance of such “associated solutions” she would still only be able to recover a point particle source that could count as being concentrated entirely on a geodesic curve *at the (associated distributional) limit*, but not before.

See also (Grosser et al., 2001, 2002, 2008) and references therein for further work developing generally covariant algebras of generalized tensors.

References

- Batterman, R., 2002. *The Devil in the Details: Asymptotic Reasoning in Explanation, Reduction, and Emergence*. Oxford University Press, USA.
- Batterman, R., 2005. Critical Phenomena and Breaking Drops: Infinite Idealizations in Physics. *Studies In History and Philosophy of Science Part B: Studies In History and Philosophy of Modern Physics* 36 (2), 225–244.
- Batterman, R., 2006. Hydrodynamics Versus Molecular Dynamics: Intertheory Relations in Condensed Matter Physics. *Philosophy of science* 73 (5), 888–904.
- Batterman, R., 2009. Idealization and Modeling. *Synthese* 169 (3), 427–446.
- Becquerel, J., 1922. *Le Principe de Relativité et la Théorie de la Gravitation*. Gauthier-Villars.
- Bergmann, P., 1942. *Introduction to the Theory of Relativity*. Prentice-Hall: Englewood N.J.
- Brown, H., 2005. *Physical Relativity: Space-time Structure from a Dynamical Perspective*. Oxford University Press, USA.
- Butterfield, J., 2011. Less is Different: Emergence and Reduction Reconciled. *Foundations of Physics* 41 (6), 1065–1135.
- Dirac, P., 1938. A New Basis for Cosmology. *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences* 165 (921), 199–208.
- Dixon, W., 1964. A Covariant Multipole Formalism for Extended Test Bodies in General Relativity. *Il Nuovo Cimento (1955-1965)* 34 (2), 317–339.
- Dixon, W., 1967. Description of Extended Bodies by Multipole Moments in Special Relativity. *Journal of Mathematical Physics* 8, 1591.
- Dixon, W., 1970a. Dynamics of Extended Bodies in General Relativity I. Momentum and Angular Momentum. *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences* 314 (1519), 499–527.
- Dixon, W., 1970b. Dynamics of Extended Bodies in General Relativity II. Moments of the Charge-current Vector. *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences* 319 (1539), 509–547.
- Dixon, W., 1973. The Definition of Multipole Moments for Extended Bodies. *General Relativity and Gravitation* 4 (3), 199–209.

- Dixon, W., 1974. Dynamics of Extended Bodies in General Relativity III. Equations of Motion. Philosophical Transactions of the Royal Society of London. Series A, Mathematical and Physical Sciences 277 (1264), 59.
- Dixon, W., 1975. On the Uniqueness of the Newtonian Theory as a Geometric Theory of Gravitation. Communications in Mathematical Physics 45 (2), 167–182.
- Duval, C., Künzle, H., 1978. Dynamics of Continua and Particles from General Covariance of Newtonian Gravitation Theory. Reports on Mathematical Physics 13 (3), 351–368.
- Duval, C., Künzle, H., 1984. Minimal Gravitational Coupling in the Newtonian Theory and the Covariant Schrödinger Equation. General Relativity and Gravitation 16 (4), 333–347.
- Earman, J., 1995. Bangs, Crunches, Whimpers, and Shrieks: Singularities and Acausalities in Relativistic Spacetimes. Oxford University Press, USA.
- Earman, J., Eisenstaedt, J., 1999. Einstein and Singularities. Studies In History and Philosophy of Science Part B: Studies In History and Philosophy of Modern Physics 30 (2), 185–235.
- Eddington, A. S., 1923. The Mathematical Theory of Relativity, 2nd Edition. Cambridge University Press: Cambridge.
- Ehlers, J., Geroch, R., 2004. Equation of Motion of Small Bodies in Relativity. Annals of Physics 309 (1), 232–236.
- Ehlers, J., Rudolph, E., 1977. Dynamics of Extended Bodies in General Relativity Center-of-mass Description and Quasirigidity. General Relativity and Gravitation 8 (3), 197–217.
- Einstein, A., 1913. Entwurf einer Verallgemeinerten Relativitätstheorie und eine Theorie der Gravitation, I. Physikalischer Teil von A. Einstein, II. Mathematischer Teil von M. Grossmann. Zeits. f. Math. u. Phys 62, 225–261.
- Einstein, A., 1916. The Foundation of the General Theory of Relativity. In: The principle of relativity. Dover Publications, Inc.: New York (1952), trans. W. Perrett and G.B. Jeffery.
- Einstein, A., 1922. The Meaning of Relativity, (1956) 6th Edition. Routledge Classics: New York, trans. Edwin P. Adams.
- Einstein, A., 1995. Ideas and Opinions. Three rivers press.
- Einstein, A., Grommer, J., 1927. Allgemeine Relativitätstheorie und Bewegungsgesetz. Preussische Akademie der Wissenschaften (Berlin), 2–13.
- Einstein, A., Infeld, L., 1940. The Gravitational Equations and the Problem of Motion II. Annals of mathematics 41 (2), 455–464.

- Einstein, A., Infeld, L., 1949. On the Motion of Particles in General Relativity Theory. *Can. J. Math* 1, 209.
- Einstein, A., Infeld, L., Hoffmann, B., 1938. The Gravitational Equations and the Problem of Motion I. *Annals of mathematics* 39 (1), 65–100.
- Einstein, A., Rosen, N., 1935. The Particle Problem in the General Theory of Relativity. *Physical Review* 48 (1), 73–77.
- Eisenhart, L., 1928. Dynamical Trajectories and Geodesics. *Annals of Mathematics* 30 (1), 591–606.
- Frisch, M., 2005. *Inconsistency, Asymmetry, and Non-locality: A Philosophical Investigation of Classical Electrodynamics*. Oxford University Press, USA.
- Garfinkle, D., 1999. Metrics with Distributional Curvature. *Classical and Quantum Gravity* 16, 4101.
- Geroch, R., Jang, P., 1975. Motion of a Body in General Relativity. *Journal of Mathematical Physics* 16, 65.
- Geroch, R., Liang, C., Wald, R., 1982. Singular Boundaries of Space-times. *Journal of Mathematical Physics* 23, 432.
- Geroch, R., Traschen, J., 1987. Strings and Other Distributional Sources in General Relativity. *Physical Review D* 36 (4), 1017–1031.
- Glinton, R., Paruchuri, P., Scerri, P., Sycara, K., 2010. Self-Organized Criticality of Belief Propagation in Large Heterogeneous Teams. *Dynamics of Information Systems* 40, 165–182.
- Glinton, R., Scerri, P., Scerri, D., Sycara, K., 2007. An Analysis and Design Methodology for Belief Sharing in Large Groups. In: *Information Fusion, 2007 10th International Conference on*. IEEE, pp. 1–7.
- Gralla, S., Wald, R., 2008. A Rigorous Derivation of Gravitational Self-force. *Classical and Quantum Gravity* 25, 205009.
- Grosser, M., Kunzinger, M., Obergugenberger, M., Steinbauer, R., 2001. *Geometric Theory of Generalized Functions with Applications to General Relativity*. Kluwer Academic Pub.
- Grosser, M., Kunzinger, M., Steinbauer, R., Vickers, J., 2002. A Global Theory of Algebras of Generalized Functions I. *Advances in Mathematics* 166 (1), 50–72.
- Grosser, M., Kunzinger, M., Steinbauer, R., Vickers, J., 2008. A Global Theory of Algebras of Generalized Functions II: Tensor Distributions. preprint.

- Guillemin, V., Sternberg, S., 1978. On the Equations of Motion of a Classical Particle in a Yang-Mills Field and the Principle of General Covariance. *Hadronic Journal* 1 (1).
- Guillemin, V., Sternberg, S., 1990. *Symplectic Techniques in Physics*. Cambridge Univ Press.
- Havas, P., 1989. The Early History of the "Problem of Motion" in General Relativity. In: Einstein and the history of general relativity: Based on The Proceedings of the 1986 Osgood Hill Conference, North Andover, Massachusetts, 8-11 May 1986. Birkhauser, p. 234.
- Havas, P., 1993. The General-Relativistic Two-Body Problem and the Einstein-Silberstein Controversy. *The Attraction of Gravitation: New Studies in the History of General Relativity* 1, 88.
- Havas, P., Goldberg, J., 1962. Lorentz-invariant Equations of Motion of Point Masses in the General Theory of Relativity. *Physical Review* 128 (1), 398–414.
- Hobson, M., Efstathiou, G., Lasenby, A., 2006. *General Relativity: An Introduction for Physicists*. Cambridge Univ Press.
- Howard, D., 1990. 'Nicht sein kann was nicht sein darf,' or the Prehistory of EPR, 1909–1935: Einstein's Early Worries about the Quantum Mechanics of Composite Systems. In: Miller, A. I. (Ed.), *Sixty-two years of uncertainty: Historical, Philosophical and Physical Inquiries into the Foundations of Quantum Mechanics*. Plenum: New York, pp. 61–111.
- Infeld, 1954. On the Motion of Bodies in General Relativity Theory. *Acta Physica Polonica* 13, 187–204.
- Infeld, L., 1957. Equations of Motion in General Relativity Theory and the Action Principle. *Reviews of modern physics* 29 (3), 398–411.
- Infeld, L., 1968. *Kordian, Fyzika I Ja*. Warsaw: Panstwowy Instytut Wydawniczy.
- Infeld, L., 1980. *Quest: An Autobiography*. Amer Mathematical Society.
- Infeld, L., Plebanski, J., 1960. *Motion and Relativity*. Pergamon Press, N.Y.
- Infeld, L., Schild, A., 1949. On the Motion of Test particles in General Relativity. *Reviews of Modern Physics* 21 (3), 408–413.
- Kadanoff, L., 2000. *Statistical Physics: Statics, Dynamics and Renormalization*. World Scientific Publishing Co.
- Kadanoff, L., Nagel, S., Wu, L., Zhou, S., 1989. Scaling and Universality in Avalanches. *Physical Review A* 39 (12), 6524–6537.

- Kennefick, D., 2005. Einstein and the Problem of Motion: A Small Clue. *The Universe of General Relativity*, 109–124.
- Kopff, A., 1923. *The Mathematical Theory of Relativity*. Methuen & Co.
- Lanczos, C., 1941. The Dynamics of a Particle in General Relativity. *Physical Review* 59 (10), 813–819.
- Lise, S., Paczuski, M., 2001. Self-organized Criticality and Universality in a Nonconservative Earthquake Model. *Physical Review E* 63 (3), 036111.
- Malament, D., 2009. A Remark About the "Geodesic Principle" in General Relativity. *Philsci-archive preprint*. Available at <http://philsci-archive.pitt.edu/5072/>.
- Mathisson, M., 1937. *Neue Mechanik Materieller Systeme*. *Acta Phys. Polon* 6 (3), 163–200.
- Mathisson, M., 1940. The Variational Equation of Relativistic Dynamics. In: *Mathematical Proceedings of the Cambridge Philosophical Society*. Vol. 36. Cambridge Univ Press, pp. 331–350.
- Mino, Y., Sasaki, M., Tanaka, T., 1997. Gravitational Radiation Reaction to a Particle Motion. *Physical Review D* 55 (6), 3457.
- Misner, C., Thorne, K., Wheeler, J., 1973. *Gravitation*. WH Freeman & Co.
- Pais, A., 2005. *Subtle is the Lord: The Science and the Life of Albert Einstein*. Oxford University Press, USA.
- Papapetrou, A., 1951. Equations of Motion in General Relativity. *Proceedings of the Physical Society*. Section A 64, 57.
- Pauli, W., 1921. *Theory of Relativity*, (1958) English Edition. Dover Publications.
- Pound, A., 2010. Self-consistent Gravitational Self-force. *Physical Review D* 81 (2), 24023.
- Quinn, T., Wald, R., 1997. Axiomatic Approach to Electromagnetic and Gravitational Radiation Reaction of Particles in Curved Spacetime. *Physical Review D* 56 (6), 3381.
- Quinn, T., Wald, R., 1999. Energy Conservation for Point Particles Undergoing Radiation Reaction. *Physical Review D* 60 (6), 64009.
- Robertson, H., 1937. Test Corpuscles in General Relativity. *Proceedings of the Edinburgh Mathematical Society* 5 (02), 63–81.
- Sole, R., Manrubia, S., 1996. Extinction and Self-organized Criticality in a Model of Large-scale Evolution. *Physical Review E* 54 (1), 42–45.

- Souriau, J., 1974. Modele de Particulea Spin dans le Champ Électromagnétique et Gravitationnel. *Ann. Inst. H. Poincaré Sect. A (NS)* 20, 315–364.
- Steinbauer, R., 2007. A Geometric Approach to Full Colombeau Algebras. In: *Proceedings of the International Conference on Generalized Functions*.
- Steinbauer, R., Vickers, J., 2006. The Use of Generalized Functions and Distributions in General Relativity. *Classical and Quantum Gravity* 23, R91.
- Steinbauer, R., Vickers, J., 2009. On the Geroch-Traschen Class of Metrics. *Classical and Quantum Gravity* 26, 065001.
- Sternberg, S., 1978. On the Role of Field Theories in our Physical Conception of Geometry. *Differential Geometrical Methods in Mathematical Physics II*, 1–80.
- Sternberg, S., 1985a. Magnetic Moments and General Covariance. *Letters in Mathematical Physics* 9 (1), 35–42.
- Sternberg, S., 1985b. The Interaction of Spin and Torsion II. The Principle of General Covariance. *Annals of Physics* 162 (1), 85–99.
- Sternberg, S., 1999. General Covariance and Harmonic Maps. *Proceedings of the National Academy of Sciences of the United States of America* 96 (16), 8845.
- Tolman, R., 1930. On the Use of the Energy-momentum Principle in General Relativity. *Physical Review* 35 (8), 875–895.
- Torretti, R., 1996. *Relativity and Geometry*. Dover Publications.
- Tulczyjew, W., 1959. Motion of Multipole Particles in General Relativity Theory. *Acta Phys. Pol* 18, 393.
- von Laue, M., 1921. *Die Relativitätstheorie*, 2nd Edition. Braunschweig: Friedrich Vieweg und Sohn.
- Wald, R., 1984. *General Relativity*. University of Chicago Press.
- Wald, R., 2011. Introduction to Gravitational Self-Force. In: Blanchet, L., Spallicci, A., Whiting, B. (Eds.), *Mass and Motion in General Relativity*. No. 162 in *Fundamental Theories of Physics*. Springer, pp. 253–262.
- Weatherall, J., 2011. A Brief Remark on Energy Conditions and the Geroch-Jang Theorem. *Foundations of Physics*, Forthcoming.
- Weyl, H., 1922. *Space, Time, Matter*. Dover Publications, translator H. L. Brose.