The Localisation of Gravitational Energy, Momentum, and Spin

Luke Matthew Butcher

Astrophysics Group, Cavendish Laboratory
and St. John's College, University of Cambridge

A dissertation submitted for the degree of
Doctor of Philosophy
in the University of Cambridge

March 2012

Declaration

This dissertation is the result of work carried out in the Astrophysics Group of the Cavendish Laboratory, Cambridge, between October 2007 and March 2012; it is my own work and includes nothing which is the outcome of work done in collaboration, except where specifically indicated in the text. Throughout the thesis, the plural pronoun "we" is used in the editorial sense, and should be taken to refer to the singular author, with the possible inclusion of the reader. No part of this dissertation has been submitted for a degree, diploma or other qualification at this or any other university. The total length of this dissertation does not exceed sixty thousand words.

Luke Butcher

March 2012

Acknowledgements

First and foremost, I wish to thank my supervisors, Anthony Lasenby and Mike Hobson, for all their insight, encouragement and guidance over the past four years. I have thoroughly enjoyed the flow and exchange of ideas in our weekly meetings, and have learnt a great deal from you both.

In addition to my supervisors, my research has brought me into contact with a number of distinguished members of the gravitational physics community, each of whom has played a small role in shaping this thesis. I thank Leonid Grishchuk for further explaining his work to me, and for offering an amusing thought experiment concerning a car crash and a hospital bill. Thanu Padmanabhan has my gratitude for a stimulating discussion of gravitational bootstrapping. I am also indebted to Stanley Deser (with support from Richard Arnowitt) for his lively and provocative communications, without which I may never have investigated the localisation of gravitational spin.

I am very grateful to STFC (and by extension, the British taxpayer) for funding my PhD studentship, and for the additional grants I have received from St John's College, Cambridge. I would also like to thank the fellows of Jesus College, Cambridge for awarding me a Junior Research Fellowship, which has supported me through the final six months of thesis writing, and will allow me to continue my research over the next few years.

On a more personal level, I would like to thank my friends in the Astrophysics Group: Bec (for the baby/thesis race), Carmen (for the Canarian cultural exchange, especially the burial of the sardine), Chris (for innumerable dinners in the Buttery), Doug (for his love of bad puns and good beer), Matt (for political arguments and a transatlantic trip), Phil (for replacing Jon Pober), Sarah (for organising RADAs), Tash (for Ricochet Robots), Tim (for being Tim), and everyone else who has made lunchtimes and pub-trips so entertaining. Similarly, I thank my friends at St John's (especially Caroline, David, Henri, Hugo, and Martin) for making my time there so uniquely enjoyable. Additional thanks go to my friends outside the university (Nick, Max, Dan, and Michael; Mike, Matt, and Sahan; and Simon) for keeping me sane, and for putting my theoretical problems in perspective.

Finally, I wish to thank my family for everything they have done for me. I thank my parents for their unconditional love and support, for the happy childhood that shaped my mind, and for all the praise and pride which has fuelled my ambitions. My father deserves special thanks for sparking my interest in physics at an early age (with explanations of atoms, and tales of inter-dimensional travel) and I am very grateful that my mother taught me the "rules" of the English language. I also thank my brother Adam for his friendship, for endless discussions of film, television, and games, for listening to my play, and for letting me be a part of his various projects.

Luke Butcher
March 2012

Summary

This dissertation documents my efforts to solve a long-standing problem in general relativity: the task of finding a physically meaningful local description of the energy and momentum of the gravitational field. I develop a solution, valid within the linear approximation to general relativity, which eliminates its own gauge-freedom and displays numerous desirable properties.

In chapter 1, I argue that an apparent solution, recently proposed by Babak and Grishchuk, is physically ill-defined. In chapter 2, I develop a formalism for generating perturbative expansions of arbitrary metric-based theories of gravity, and employ these techniques to construct a quadratic action (a covariantised Fierz-Pauli action) from which general relativity can be derived, order by order, following a simple procedure which couples the gravitational field to its own energy-momentum.

The remainder of the thesis focuses on localising the energy, momentum, and spin of gravity within the linear regime. In chapter 3, I derive a new gravitational energy-momentum tensor by requiring that the tensor account for the local exchange of energy-momentum between matter and the gravitational field. The gauge-freedom of this description is removed in a natural fashion: the harmonic gauge condition arises as an automatic consequence of the tensor's derivation, and transverse-traceless gauge is then motivated by comparison with the gauge-invariant exchange of energy-momentum between the gravitational field and an infinitesimal detector. I show that, once this gauge-fixing programme is employed, my gravitational energy-momentum tensor always describes non-negative energy-density, and causal energy-flux.

Chapter 4 extends this framework by developing a local description of the angular momentum (and moment-of-energy) carried by the linearised gravitational field. Analysing the local exchange of angular momentum between matter and gravity, I derive a tensor which localises gravitational intrinsic spin; once the aforementioned gauge-fixing programme is employed, this spin tensor is traceless (which I argue to be necessary for ensuring the absence of infinite pressure-gradients) and describes purely spatial spin. Recapitulating the analysis of the previous chapter, I also give a treatment of the gauge-invariant exchange of angular momentum (and moment-of-energy) between the gravitational field and an infinitesimal detector.

Finally, in Chapter 5, I investigate the roles played by the previously derived energy-

momentum tensor and spin tensor in a broader theoretical context, demonstrating that (a) they are indeed Noether currents associated with the translational and rotational symmetries of the linearised gravitational field, and that (b) they generate gravity alongside the energy-momentum and spin of matter. I confirm (a) by constructing a Lagrangian for linear gravity (a covariantised Fierz-Pauli Lagrangian) which generates the tensors according to standard variational definitions. I then demonstrate (b) by identifying the tensors as the quadratic terms in a perturbative expansion of the Einstein and Einstein-Cartan field equations. The form of perturbation used in these expansions is also analysed, suggesting a route by which the formalism may be extended beyond the linear/quadratic regime, and linking the Lagrangian of this chapter with the action constructed in chapter 2.

Luke Butcher
March 2012

Contents

Chapter 0

Introduction[1]

0.1 Historical Background

Conservation laws, particularly those of energy, momentum, and angular momentum, are some of the most powerful and fundamental laws of nature presently known. In their most basic form, these rules demand nothing more than the constancy of a few important quantities (the total energy, momentum, and angular momentum of an isolated system) regardless of what interactions or physical processes may occur. Stated as such, these *global* laws do not insist upon a microscopic mechanism to enforce this constancy, and would allow energy (or momentum, or angular momentum) to disappear from one location and appear elsewhere, provided only that the two effects occur *simultaneously*.

During the 19$^{\text{th}}$ century, however, *force fields* came to replace *action-at-a-distance* in our understanding of electromagnetic phenomena [57], and local descriptions of electromagnetic energy began to be developed, most notably by Poynting in 1884 [67]. Following the arrival of special relativity [33] in 1905, it soon became clear that a local description of energy was in fact a necessity, and that any notion of a purely global conservation law would have to be abandoned. To explain: given that the definition of simultaneity had been shown to depend on an observer's motion, any instantaneous transmission of energy over a finite distance would (a) violate the conservation of energy in all but one reference frame, and (b) entail not only faster than light communication but also, in some frames of reference, communication backwards in time. It was therefore necessary to insist that energy (and momentum, and angular momentum) not only be conserved *globally*, but *locally* also: any energy lost by the particles and fields within a particular region would need to be accompanied by a *flow* of energy across the region's boundary. Accordingly, a description of energetics would only be considered complete when one could quantify the *energy-density* and *energy-flux* at every point in space.

However, while special relativity required energy and momentum to be localised on grounds of logical consistency, it did not provide any means to test this localisation ex-

[1] In order to maintain the numerical correspondence between chapters 1 to 5 and the articles [1–5] from which they derive, it has been necessary to begin with a *zeroth* chapter; this also has the advantage of providing a notional dividing-line between introductory material and the main body of the dissertation.

perimentally; it would not be until 1915, and the development of a relativistic theory of gravitation, Einstein's *general relativity* [34], that such a framework would exist. As special relativity had unified the concepts of mass, energy, and momentum, and because mass-density was the source of *Newtonian* gravity, general relativity cast the flux and density of energy and momentum (packaged into a single object, the *energy-momentum tensor* of matter) as the source of the relativistic gravitational field. Thus, with Einstein's field equations at hand, and sufficiently accurate measuring apparatus, one could infer the location of material energy-momentum, at least in principle, by observing the gravity it generated.

It is here that the ironic twist in the story unfolds, because although gravity had lent an empirical *operational definition* to the local energy-momentum of matter, there is one form of energy-momentum that escapes this definition entirely: the energy-momentum of the gravitational field itself! In the absence of an operational definition of this form (and for a host of technical and conceptual reasons, explored in later chapters) even to this day, nearly a century after general relativity was first discovered, we do not know *where* the energy and momentum of the gravitational field resides.

Numerous attempts have been made to define a gravitational energy-momentum tensor (beginning with Einstein [35] in 1916, and with notable contributions from Møller [59] and Landau and Lifshitz [52]) but the dependence of these tensors on the *gauge freedom* of the gravitational field (historically, the choice of coordinate system) allows them to assign a vast range of localisations to any one gravitational field. Thus, without a gauge invariant description, or a strong argument for the choice of one gauge over another, these so-called *pseudotensors* remain wildly ambiguous, unable to assign a meaningful local measure of energy-momentum to the gravitational field.

This difficulty was not a peripheral concern; indeed, it was instrumental in the confusion which reigned, even as late as 1956, as to whether gravitational radiation was a genuine prediction of general relativity. Although propagating solutions clearly existed within the theory, in the absence of a measure of the energy they carried, arguments were made [70, 72] which purported to demonstrate that such waves carried no energy at all; by implication, they were physically unobservable, and presumably an artifact of a poor choice of coordinates. The dispute was famously put to rest by Feynman at the 1957 Chapel Hill conference, where he argued, in his characteristically pragmatic style, that a very simple detector, consisting of a rod and two "sticky beads", would clearly absorb energy from a gravitational wave [68].

Having succeeded in persuading the theoretical community that gravitational waves were indeed physically observable, and that the gravitational field carried energy (regardless of the subtlety of its localisation) Feynman's elegant argument, popularised by Bondi [19], set the tone for many of the field's developments over the following years. In 1958, Bel first wrote down what we now refer to as the Bel-Robinson tensor [14, 69], an object structurally analogous to the electromagnetic energy-momentum tensor, wherein the Weyl curvature plays the role of the field strength. Although it was clear from the outset that the

Bel-Robinson tensor did not describe gravitational energy-momentum (having the wrong dimensions and too many indices) it at least offered a local, covariant, gauge invariant measure of the "intensity" of a gravitational wave. Soon after, following unrelated lines of enquiry, meaningful definitions of the *total* energy of gravitating systems began to appear: the Komar mass [49] (for stationary spacetimes) in 1959, followed by the ADM mass [10] and Bondi mass [20] (for asymptotically flat spacetimes) both in 1962.

These global definitions served to highlight the role of gravitational energy-momentum in contributing to the total gravitational weight of a system, a concept which had originally been used by Kraichnan, in a more formal setting, to argue that the full non-linear theory of general relativity arose by coupling the linear spin-2 field to its own energy-momentum [50]. It was in this conceptual environment, then, that Brill and Hartle developed their "self-consistent field" approach [22], a technique which *averaged* the energy-momentum of the high-frequency modes of the gravitational field in order to predict their effect on the low-frequency modes, and allowed them to construct their *gravitational geon* (a concentration of gravitational waves, held together by its own gravity) in 1964. Extending this idea to a more generally applicable framework, Isaacson presented an approach to the propagation of gravitational waves in the low-amplitude high-frequency limit [46], and defined a gauge-invariant "effective" gravitational energy-momentum tensor by the very same method, taking an average over many wavelengths of the gravitational wave [47].

Throughout the 1970s, this "averaged" energy-momentum tensor remained the only available gauge-invariant object that even resembled a truly local picture of gravitational energy-momentum, and many had arrived at the conclusion that this was all that could be achieved within the theory. In 1973, Milsner, Thorne and Wheeler famously declared the whole enterprise to be "looking for the right answer to the wrong question" [58] and succeeded in dissuading a generation of physicists from tackling the problem.

Bolstered by the successes of the global definitions of energy-momentum (in particular, proofs of *positivity* for the ADM mass [24, 73, 82]) interest in the 1980s turned to *quasi-local* definitions of gravitational energy-momentum [13, 42, 78]. These methods attempt to quantify the total energy within compact *finite* regions, usually by means of an integral over the region's boundary, but stop short of the infinitesimal regions required to define energy-densities. Despite considerable progress in this area, the quasi-local approach still struggles with many of the technical and conceptual problems encountered in the truly local problem, suffering from an overabundance of definitions and constructions, and a lack of consensus over what properties they should obey [77].

Over the last two decades, the task of finding a physically well-defined gravitational energy-momentum tensor has garnered little attention, with a small number of physicists offering tentative solutions [11, 56] which, on closer inspection, have been found to inherit many of the major flaws of earlier approaches. In contrast to this inactivity, advances in experimental and computational physics have introduced numerous potential applications for a local description of gravitational energy-momentum. Considering that the first direct detection of gravitational radiation is expected to occur in a matter of years [75], the era of gravitational wave astronomy may soon be upon us. In this context, a gravitational energy-momentum tensor could prove invaluable in the design of new gravitational detectors, much as Poynting's work influenced our understanding of electromagnetic antennae in the previous century. Complementing this, the art of *numerical relativity* has entered something of a "golden age" in recent years [54]. These methods have bestowed upon us a wealth of new solutions to the Einstein field equations, the often unintuitive behaviour of which may be better understood were we to possess a meaningful gravitational energy-momentum tensor.[2] These factors considered, now is perhaps the time to reevaluate the wisdom of Misner, Thorne and Wheeler's definitive declaration, and examine whether there is in fact a "right" question to ask.

0.2 Overview

This dissertation draws together the various avenues of my research into the localisation of gravitational energy, momentum, and intrinsic spin. Chapters 1 to 5 (being the main body of the thesis) were written as research papers [1–5] and have either appeared in, or recently been submitted to, Physical Review D. As each chapter is self-contained in terms of exposition, beginning with its own motivations and ending with its own conclusions, there is little need for a detailed introduction here. Instead, the purpose of this overview is to briefly explain the relationship between the chapters, so they need not be viewed in isolation.

Chapter 1 can be thought of as a preliminary study, in which the failings of Babak and Grishchuk's gravitational energy-momentum tensor [11] are discussed in order that such pitfalls might be avoided in the rest of the thesis. The lessons drawn from this work are threefold: (i) the coordinate dependence of gravitational pseudotensors can be nullified, and genuine tensors defined, if one presents these tensors in an auxiliary background spacetime; (ii) this does not rid the description of its gauge freedom, however, as the tensors will now depend on the arbitrary mapping between the physical and background spacetimes; (iii) unless the gauge freedom is heavily constrained, or a gauge-invariant replacement for the tensor found, this ambiguity is so great as to render the description largely meaningless.[3] As such, this chapter lays out the basic truths on which much of the thesis is built.

Chapter 2 examines the coupling between the gravitational field and its own energy-momentum (as first described by Kraichnan [50]) and presents a procedure by which this coupling can generate the full theory of general relativity, order by order, starting with

[2] For example, recent simulations of black hole mergers [23] have displayed peculiar bobs and kicks that one might hope to explain in terms of the motion of gravitational energy-momentum.

[3] Babak and Grishchuk adopt a slightly unusual philosophical position, in which the flat background is considered to be the "real" spacetime, and gravitational forces act on matter in such a way as to reproduce the same physical predictions of general relativity, without the need for a curved spacetime. As this stance appears to eschew the arbitrary mappings that would otherwise link the physical spacetime to the background, much of the chapter is devoted to identifying the gauge freedom that has been obscured.

only an action for linearised gravity. The gravitational energy-momentum tensor defined by this algorithm cannot escape the problem of gauge dependence, but the aim of this particular chapter is not to define an unambiguous measure of local energy-momentum. Rather, by illuminating the connection between background coupling in the action of the linear spin-2 field, and the non-linear theory that evolves from repeatedly coupling this field to its energy-momentum, we gain a number of valuable results that prove useful in later chapters, and observe the importance of *linearised* general relativity as the seed of structure within the full theory.

Following this, chapters 3, 4, and 5 constitute the main investigation of the thesis: focusing exclusively on the *linear* approximation to general relativity, I succeed in localising gravitational energy, momentum, and spin. The key to this approach is the appearance of a natural gauge-fixing procedure, motivated by properties of the gravitational energy-momentum tensor and spin tensor I derive, and by comparison with the gauge-invariant energy-momentum (and angular momentum) exchanged with an infinitesimal detector. Chapter 3 concerns energy and momentum, chapter 4 concerns angular momentum, intrinsic spin, and moment-of-energy, and chapter 5 combines all these results, embedding them within a theoretical framework that connects to previous treatments of gravitational energy-momentum. It is in this last chapter that the results of chapter 2 find their main application, linking the gravitational Lagrangian (from which the energy-momentum and spin tensors of chapters 3 and 4 can be derived) to a perturbative expansion of the Einstein-Hilbert action, and hinting at a method by which the framework may be extended beyond linear order.

A more thorough summary of the content of these chapters can be found in section 0.3, where the abstracts of the papers [1–5] are reproduced.

Notation and Conventions

Throughout this dissertation, I work in units where $c = 1$, write $\kappa \equiv 8\pi G$, and adopt the sign conventions of Wald [79] and Misner, Thorne, and Wheeler [58]: $\eta_{\mu\nu} \equiv \mathrm{diag}(-1,1,1,1)$, $[\nabla_c, \nabla_d]v^a \equiv 2\nabla_{[c}\nabla_{d]}v^a \equiv R^a{}_{bcd}v^b$, and $R_{ab} \equiv R^c{}_{acb}$. Roman letters are used as abstract tensor indices [79, §2.4] and Greek letters as numerical indices running from 0 to 3.

While notation is broadly consistent throughout the thesis, and entirely consistent within chapters 3–5, it has been necessary to adopt a different definition of the gravitational field h_{ab} within chapter 1 (following Babak and Grishchuk) and within chapter 2 (for technical reasons)[4]. As a result, the differential operator $\widehat{G}_{abcd}$ of chapter 2 has the opposite sign to the operator $\widehat{G}_{abcd}$ of chapters 3–5.

[4] See footnote 4 of chapter 2 for details.

0.3 Abstracts

Chapter 1 **[1]**
Physical Significance of the Babak-Grishchuk Energy-Momentum Tensor

We examine the claim of Babak and Grishchuk [11] to have solved the problem of localising the energy and momentum of the gravitational field. After summarising Grishchuk's flat-space formulation of gravity, we demonstrate its equivalence to General Relativity at the level of the action. Two important transformations are described (diffeomorphisms applied to all fields, and diffeomorphisms applied to the flat-space metric alone) and we argue that both should be considered gauge transformations: they alter the mathematical representation of a physical system, but not the system itself. By examining the transformation properties of the Babak-Grishchuk gravitational energy-momentum tensor under these gauge transformations (infinitesimal and finite) we conclude that this object has no physical significance.

Chapter 2 **[2]**
Bootstrapping Gravity: a Consistent Approach to Energy-Momentum Self-Coupling

It is generally believed that coupling the graviton (a classical Fierz-Pauli massless spin-2 field) to its own energy-momentum tensor successfully recreates the dynamics of the Einstein field equations order by order; however the validity of this idea has recently been brought into doubt [64]. Motivated by this, we present a graviton action for which energy-momentum self-coupling is indeed consistent with the Einstein field equations. The Hilbert energy-momentum tensor for this graviton is calculated explicitly and shown to supply the correct second-order term in the field equations; in contrast, the Fierz-Pauli action fails to supply the correct term. A formalism for perturbative expansions of metric-based gravitational theories is then developed, and these techniques employed to demonstrate that our graviton action is a starting point for a straightforward energy-momentum self-coupling procedure that, order by order, generates the Einstein-Hilbert action (up to a classically irrelevant surface term). The perturbative formalism is extended to include matter and a cosmological constant, and interactions between perturbations of a free matter field and the gravitational field are studied in a vacuum background. Finally, the effect of a non-vacuum background is examined, and the graviton is found to develop a non-vanishing "mass-term" in the action.

Chapter 3 **[3]**
Localising the Energy and Momentum of Linear Gravity

A framework is developed which quantifies the local exchange of energy and momentum between matter and the linearised gravitational field. We derive the unique gravitational energy-momentum tensor consistent with this description, and find that this tensor only exists in the harmonic gauge. Consequently, nearly all the gauge freedom of our framework is naturally and unavoidably removed. The gravitational energy-momentum tensor is then shown to have two exceptional properties: (a) it is gauge-invariant for gravitational plane-waves, (b) for arbitrary transverse-traceless fields, the energy-density is never negative, and the energy-flux is never spacelike. We analyse in detail the local gauge invariant energy-momentum transferred between the gravitational field and an infinitesimal point-source, and show that these invariants depend only on the transverse-traceless components of the field. As a result, we are led to a natural gauge-fixing program which at last renders the energy-momentum of the linear gravitational field completely unambiguous, and additionally ensures that gravitational energy is never negative nor flows faster than light. Finally, we calculate the energy-momentum content of gravitational plane-waves, the linearised Schwarzschild spacetime (extending to arbitrary static linear spacetimes) and the gravitational radiation outside two compact sources: a vibrating rod, and an equal-mass binary.

Chapter 4 [4]
Localising the Angular Momentum of Linear Gravity

In the previous chapter we derived an energy-momentum tensor for linear gravity that exhibited positive energy-density and causal energy-flux. Here we extend this framework by localising the angular momentum of the linearised gravitational field, deriving a gravitational spin tensor which possesses similarly desirable properties. By examining the local exchange of angular momentum (between matter and gravity) we find that gravitational intrinsic spin is localised, separately from "orbital" angular momentum, in terms of a gravitational spin tensor. This spin tensor is then uniquely determined by requiring that it obey two simple physically-motivated algebraic conditions. Firstly, the spin of an arbitrary (harmonic-gauge) gravitational plane-wave is required to flow in the direction of propagation of the wave. Secondly, the spin tensor of any transverse-traceless gravitational field is required to be traceless. (This condition is shown to rid the field of infinite pressure gradients.) Additionally, the following properties arise in the spin tensor spontaneously: all transverse-traceless fields have purely spatial spin, and any field generated by a static distribution of matter will carry no spin at all. Following the structure of the previous chapter, we then examine the (spatial) angular momentum exchanged between the gravitational field and an infinitesimal detector, and develop a microaveraging procedure that renders the process gauge invariant. The exchange of non-spatial angular momentum (i.e. moment-of-energy) is also analysed, leading us to conclude that a gravitational wave can displace the centre-of-mass of the detector; this conclusion is also confirmed by a "first principles" treatment of the system. Finally, we discuss the spin carried by a gravitational plane-wave.

Chapter 5 [5]
Localised Energetics of Linear Gravity: Theoretical Development

Thus far, we have developed a local description of the energy, momentum and angular momentum carried by the linearised gravitational field, wherein the gravitational energy-momentum tensor displays positive energy-density and causal energy-flux, and the gravitational spin-tensor is traceless and describes purely spatial spin. We now investigate the role these tensors play in a broader theoretical context, with the aim of demonstrating that (a) they do indeed constitute Noether currents associated with the symmetry of the linearised gravitational field under translation and rotation, and (b) they are themselves a source of gravity, analogous to the energy-momentum and spin of matter. To prove (a) we construct a Lagrangian for linearised gravity (a covariantised Fierz-Pauli Lagrangian for a massless spin-2 field) and show that our tensors can be obtained from this Lagrangian using a standard variational technique for calculating Noether currents. This approach generates formulae that *uniquely* generalise our gravitational energy-momentum tensor and spin tensor beyond harmonic gauge: we show that no other generalisation can be obtained from a covariantised Fierz-Pauli Lagrangian without introducing second derivatives in the energy-momentum tensor. We then construct the Belinfante energy-momentum tensor associated with our framework (combining spin and energy-momentum into a single object) and as our first demonstration of (b) we establish that this Belinfante tensor appears as the second-order contribution to a perturbative expansion of the Einstein field equations, generating the gravitational field in a manner equivalent to the (Belinfante) energy-momentum tensor of matter. By considering a perturbative expansion of the Einstein-Cartan field equations, we then demonstrate that (b) can be realised without forming the Belinfante tensor: our energy-momentum tensor and spin tensor appear as the quadratic terms in *separate* field equations, generating gravity as distinct entities. Finally, we examine the role of field redefinitions within these perturbative expansions; in contrast to our tensors, the Landau-Lifshitz tensor is found to require a *non-local* field redefinition in order to be cast as a source of the gravitational field. In an appendix, we also give a brief treatment of the global quantities that our framework defines, and verify their equivalence (within the quadratic approximation) to the ADM energy-momentum and angular momentum.

Chapter 1

Physical Significance of the Babak-Grishchuk Gravitational Energy-Momentum Tensor

1.1 Introduction

Despite the central role played by the energy-momentum tensor of matter in general relativity, there is no widely accepted way to localise the energy and momentum of the gravitational field itself. In the place of a genuine solution to this problem, we are forced to make do with an over-abundance of energy-momentum pseudotensors, objects designed to display some or other property befitting a measure of gravitational energy-momentum, but whose coordinate dependence renders them of little physical significance beyond giving the correct integrals at infinity in asymptotically flat spacetimes. Even for weak gravitational waves, the best measures at our disposal only become meaningful once we have averaged over many wavelengths.

The canonical response to the gravitational energy-momentum problem is to dismiss it as "looking for the right answer to the wrong question"[58, §20.4]; but while the well-known argument presented by Misner, Thorne and Wheeler is certainly compelling, it is far from watertight. They remind us that the equivalence principle ensures that all "gravitational fields" ${\Gamma^\alpha}_{\beta\gamma}$ can be made to vanish at a point by a suitable choice of coordinates, and conclude that because gravity is locally zero, there can be no energy density associated with it. However, this argument fails to consider tensors containing second derivatives of the metric, which unlike ${\Gamma^\alpha}_{\beta\gamma}$ cannot be made to vanish by choice of coordinates, and really do reflect the local curvature of spacetime: for example, the Riemann tensor can be used to construct objects such as the Bel-Robinson tensor [69]. Misner, Thorne and Wheeler also point out that, while the matter energy-momentum tensor derives its physical significance by curving space, a similar tensor for gravity would not be a source term for the field equations. However, this stance is based around a prejudice for writing the Einstein field equations as $G^{ab} = \kappa T^{ab}$ with gravity on the left and matter on the right; there is nothing to stop us splitting up G^{ab} in a covariant fashion, grouping one part with T^{ab}, and interpreting *this* as the total energy-momentum source, taking the remainder of G^{ab} to be the gravitational 'response'. Despite these reservations, the argument in [58] remains vindicated as yet by the failure of these escape-routes to yield anything which can be physically interpreted as an energy-momentum tensor.

It might appear that the only straightforward solution to the problem is to extend the definition of the matter energy-momentum tensor T^{ab} (a functional derivative of the matter Lagrangian with respect to the metric) to the gravitational field, and conclude that the gravitational energy-momentum tensor is $-G^{ab}/\kappa$, where $\kappa = 8\pi G/c^4$. The Einstein field equations could then be interpreted as a constraint that everywhere sets to zero the sum of gravitational and matter energy-momentum. While one might claim this simple idea conveys some important physical insight, it suffers from numerous problems. Firstly, $-G^{ab}/\kappa$ lacks the analytical power one expects from an energy-momentum tensor: the ability to split the set of all physical systems at a particular time into classes of *different* total energy and momenta, so that conservation laws alone can reveal that two particular spacelike hypersurfaces could never be part of the same spacetime. Secondly, it leads us to conclude that the gravitational field only has energy where matter is also present, precluding the use of this prescription to describe the energetics of gravitational waves, or define a gravitational tension in the vacuum between massive bodies. Thirdly, the energy-momentum tensors for gravity and matter are conserved *separately* ($\nabla_a G^{ab} = 0$ and $\nabla_a T^{ab} = 0$) so that although there is a delicate balance that keeps their sum zero, it is not the case that energy or momentum simply 'flows' between gravity and matter, as $\nabla_a(T^{ab} - G^{ab}/\kappa) = 0$ *alone* would imply. Lastly, we note that the conservation law $\nabla_a G^{ab} = 0$ actually tells us nothing at all about the gravitational field; it is satisfied identically, without any need for the equations of motion to hold. Because of these drawbacks, if we are to regard $-G^{ab}/\kappa$ as a solution to the gravitational energy-momentum problem, we consider it rather a trivial one. Clearly, the reason for this triviality is that we have overworked the metric: we cannot use the functional derivative with respect to a dynamical field as a way of defining the energy-momentum tensor for that same field, as we will only end up writing down the equations of motion twice. This line of reasoning leads us to consider that one method of attack for this problem may be to separate the two roles played by g^{ab} in general relativity, that of dynamic field and spacetime metric.

In [40], Grishchuk develops a "field-theoretical" approach to gravitation, which expresses the physical content of general relativity (GR) in terms of a dynamical symmetric tensor field in flat Minkowski spacetime. Although this formulation has been carefully designed to agree with the empirical predictions of GR, in [11] Babak and Grishchuk claim that the flat-space approach allows them to define a unique, symmetric, and non-trivial energy-momentum tensor for the gravitational field. The major purpose of this chapter is to examine the extent to which this tensor is physically meaningful.

1.2 Flat-Space Gravitation

Babak and Grishchuk represent gravitation as the theory of a dynamical symmetric tensor field h^{ab} defined over a four-dimensional manifold $\mathcal{M}$ with a (non-dynamical) flat Lorentzian metric γ^{ab}. Translation between this picture and the dynamical metric g^{ab} of GR can be achieved using the following relation:

$$\sqrt{-g}g^{ab} = \sqrt{-\gamma}(\gamma^{ab} + h^{ab}), \tag{1.1}$$

where $g = 1/\det(g^{\alpha\beta})$ and $\gamma = 1/\det(\gamma^{\alpha\beta})$. It should be emphasised that Babak and Grishchuk consider this relation to be the *definition* of g^{ab}, a tensor to which they assign no particular fundamental or geometric significance.[1] Accordingly, they use γ^{ab}, rather than g^{ab}, to raise and lower tensor indices[2], and define a (torsion-free) covariant derivative $\check{\nabla}_a$ (denoted by indices following " $\check{;}$ "[3] and with Christoffel symbols $C^a{}_{bc}$) by $\check{\nabla}_c\gamma^{ab} = 0$. As γ^{ab} is flat,

$$\check{R}^a{}_{bcd} \equiv C^a{}_{bd,c} - C^a{}_{bc,d} + C^e{}_{bd}C^a{}_{ec} - C^e{}_{bc}C^a{}_{ed} = 0, \tag{1.2}$$

and $\check{\nabla}_a$ derivatives commute. This contrasts to the usual (GR) covariant derivative ∇_a, denoted by indices following " ; ", defined by $\nabla_c g^{ab} = 0$, and with curvature tensor $R^a{}_{bcd} \neq 0$ in general.[4]

To ensure that h^{ab} obeys an equation of motion consistent with Einstein's field equations, its dynamics are determined by an action S that is equivalent to the Einstein-Hilbert action. Specifically, in [11] Babak and Grishchuk use the action

$$S = \frac{-1}{2\kappa}\int\sqrt{-\gamma}\Big[h^{ab}{}_{\check{;}c}P^c{}_{ab} - (\gamma^{ab}+h^{ab})(P^c{}_{ad}P^d{}_{bc} - \frac{1}{3}P_aP_b)\Big]\mathrm{d}^4x, \tag{1.3}$$

where $P^a{}_{bc}$ and P_a are functions of h^{ab}, γ^{ab} and $h^{ab}{}_{\check{;}c}$ given in their paper. If we add to the Lagrangian the following surface term:

$$\begin{aligned}\mathcal{L}_{\text{surface}} &= \frac{1}{2\kappa}\left[\sqrt{-\gamma}(\gamma^{ab}+h^{ab})P^c{}_{ab}\right]_{,c}\\ &= \frac{1}{2\kappa}\left[\sqrt{-\gamma}(\gamma^{ab}+h^{ab})P^c{}_{ab}\right]_{\check{;}c},\end{aligned} \tag{1.4}$$

[1]Of course, because Babak and Grishchuk insist that this viewpoint does not contradict the predictions of general relativity, effects that are traditionally deemed the result of spacetime geometry (proper lengths of coordinate displacements, rates of clocks, geodesic deviation, etc.) will be viewed as arising from a 4-force that matter feels in response to the presence of h^{ab}; see [40] for details. The correspondence with GR inevitably means that predictions of this nature can always be expressed in terms of g^{ab} alone.

[2]A singular exception is made for g^{ab}: it is assigned the 'lowered' form $g_{ab} = (g^{ab})^{-1}$ to coincide with the GR definition.

[3]This notation differs from [11]: where they write ∇ and " ; ", we write $\check{\nabla}$ and " $\check{;}$ ".

[4]There is no contradiction in being able to define two different covariant derivatives on a manifold. Because both have been defined by a tensor equation (without any reference to coordinate systems) they must both produce genuine (abstract) tensor indices $;a$ and $\check{;}a$. The significance of the standard covariant derivative (in GR) is not just that it is covariant, but that it expresses the Equivalence Principle: in a system of local inertial coordinates $\{x^\alpha\}$ such that $g^{\alpha\beta} = \eta^{\alpha\beta} + O(x^2)$ near some point p, the Christoffel symbols for the ∇_a derivative vanish and we find that (at p) $\nabla_a = \partial_a$, the ordinary derivative of these coordinates. Thus $\nabla_c g^{ab} = 0$ picks out the coordinate independent derivative operator which coincides with local inertial coordinate derivatives. In contrast, a coordinate system $\{y^\alpha\}$ for which $\check{\nabla}_a = \partial_a$ at p will not necessarily have $g^{\alpha\beta} = \eta^{\alpha\beta} + O(y^2)$ there; however, as the flat-space picture eschews the geometric interpretation of g^{ab}, we can avoid assigning much significance to this point.

then, applying the flatness condition (1.2) to equation (53) of [11], we see that

$$\begin{aligned} S + S_{\text{surface}} &= \frac{1}{2\kappa} \int \sqrt{-\gamma}(\gamma^{ab} + h^{ab}) \left(P^{c}{}_{ab\breve{;}c} + P^{c}{}_{ad} P^{d}{}_{bc} - \frac{1}{3} P_a P_b \right) \mathrm{d}^4 x \\ &= \frac{-1}{2\kappa} \int \sqrt{-g} g^{ab} R_{ab} \mathrm{d}^4 x \\ &= S_{\text{EH}}, \end{aligned} \tag{1.5}$$

the Einstein-Hilbert action. Extremising S with respect to variations in h^{ab}, we have

$$\frac{\delta S}{\delta h^{ab}} = 0 \quad \Rightarrow \quad \frac{\delta S_{\text{EH}}}{\delta g^{cd}} \left(\frac{\partial g^{cd}}{\partial h^{ab}} \right)_{\gamma} = 0, \tag{1.6}$$

where the subscript γ indicates that γ^{ab} has been held constant. As an inverse of $\left(\frac{\partial g^{cd}}{\partial h^{ab}}\right)_{\gamma}$ exists, namely

$$\begin{aligned} \left(\frac{\partial h^{ab}}{\partial g^{cd}} \right)_{\gamma} &= \frac{1}{\sqrt{-\gamma}} \frac{\partial \sqrt{-g} g^{ab}}{\partial g^{cd}} \\ &= \frac{\sqrt{-g}}{\sqrt{-\gamma}} \left(2\delta^{(a}{}_{c} \delta^{b)}{}_{d} - \frac{1}{2} g^{ab} g_{cd} \right), \end{aligned} \tag{1.7}$$

the equations of motion (1.6) are equivalent to the Einstein Field Equations:

$$\frac{\delta S_{\text{EH}}}{\delta g^{ab}} = 0. \tag{1.8}$$

As presented in [40], the original motivation for this flat-space picture is that it allows physicists to study and predict gravitational phenomena in a framework that is free of the conceptual baggage of differential geometry, and has more in common with the language of particle physics and classical electrodynamics. However, the work presented in [11] elevates this framework beyond the status of a ‘linguistic trick’, as the metric γ^{ab} allows one to define the “metrical energy-momentum tensor” according to

$$\begin{aligned} \overset{m}{t}{}^{ab} &\equiv \frac{-2}{\sqrt{-\gamma}} \frac{\delta \mathcal{L}}{\delta \gamma_{ab}} \\ &\equiv \frac{-2}{\sqrt{-\gamma}} \left(\frac{\partial \mathcal{L}}{\partial \gamma_{ab}} - \partial_c \left(\frac{\partial \mathcal{L}}{\partial \gamma_{ab,c}} \right) \right). \end{aligned} \tag{1.9}$$

From this, a unique gravitational energy-momentum tensor t^{ab} can be constructed that is symmetric, free of second derivatives, and conserved by the equations of motion: see equation (65) of [11]. Having made this identification, the field equations (1.6) take on the simple form[5]

$$\kappa t^{ab} = \left[\frac{g}{2\gamma} (g^{ab} g^{cd} - g^{ac} g^{bd}) \right]_{\breve{;}c\breve{;}d}. \tag{1.10}$$

Although this equation does not *define* the energy-momentum tensor, it provides us with a simple method for calculating t^{ab}, given the gravitational field.[6]

[5]This result is a corrected version of equation (78) of [11]; it is easy to see that the original equation lacks a factor of $-1/\gamma$ by comparing it to the preceding equation in that paper.

[6]This relation also reveals a connection between t^{ab} and t^{ab}_{LL}, the gravitational energy-momentum

1.3 Physical Content of t^{ab}

We cannot fault Grishchuk's formulation of gravitational dynamics within the realm of general relativity, as agreement over predictions of 'geometrical phenomena' (as they would be interpreted in GR) has been achieved by design.[7] However, in comparison with general relativity, the flat-space theory possesses additional mathematical structure: two tensors h^{ab} and γ^{ab} fulfil the role played by g^{ab} alone. This extra structure endows the flat-space theory with an increased range of expression, making possible the definition of tensors that cannot be constructed within the framework of GR. As we shall show, the gravitational energy-momentum tensor is one of these 'non-GR' quantities.[8] We investigate here whether t^{ab} (or any non-GR quantity) can be physically significant, or whether it can only ever be interpreted as an artefact of the mathematics.

1.3.1 Gauge Transformations

Besides allowing us to interpret gravity as a force-field on flat space, the presence of γ^{ab} has had the important side-effect of increasing the space of gauge transformations of the theory. The core reason for this is that the flatness constraint (1.2) is not enough to define a unique γ^{ab} for a given g^{ab}, a tensor which, through the correspondence with GR, can be used *alone* to construct the observable predictions of the theory. In this section we examine two transformations and justify their status as gauge transformations, i.e. that they alter the mathematical representation of a physical system, but not the system itself.

Diffeomorphism gauge transformations

Given a diffeomorphism ϕ: $\mathcal{M} \to \mathcal{M}$, we transform all tensor fields $X^{a\cdots}{}_{b\ldots}$ according to

$$X^{a\cdots}{}_{b\ldots} \to (\phi^* X)^{a\cdots}{}_{b\ldots}, \tag{1.11}$$

where the action of ϕ^* on X is defined in the standard way by the action of the pullback of ϕ (and the pushforward of ϕ^{-1}) on the dual-vector (and vector) arguments of X; see [79, §C.1] for details. Although, as written, this transformation cannot be the result of a

pseudotensor of Landau and Lifshitz [52]: in Lorentzian coordinates ($\gamma^{\alpha\beta} = \eta^{\alpha\beta}$) it follows from (1.10) that $t^{\alpha\beta} = (-g)t^{\alpha\beta}_{\mathrm{LL}}$. Roughly speaking, t^{ab} corresponds to a Landau-Lifshitz energy-momentum that has been rendered coordinate-system independent (tensorial, rather than pseudotensorial) through the replacement of ordinary coordinate derivatives ∂_a with flat-space covariant derivatives $\check{\nabla}_a$.

[7] Of course, one may still wish to attack the *aesthetics* of a framework which, from the GR viewpoint, appears to obscure the geometric nature of gravity, and replaces the Equivalence Principle with a seemingly arbitrary coupling between h^{ab}, γ^{ab} and matter. However, the potential for a greater understanding of the local energy-momentum content of the gravitational field should be enough to temporarily assuage these objections.

[8] This statement might appear obvious due to the use of γ^{ab} in (1.9), or the presence $h^{ab}{}_{\check{;}c}$ in the definition of t^{ab} (equation (65) of [11]). However, a tensor defined in terms of γ^{ab}, h^{ab}, and $\check{\nabla}_a$ may also be expressible in GR, e.g. $P^c{}_{ab\check{;}c} + P^c{}_{ad}P^d{}_{bc} - \frac{1}{3}P_a P_b = -R_{ab}[g]$.

change of coordinate system[9], it transforms the *components* $X^{\alpha\dots}{}_{\beta\dots}$ in a typographically identical manner to that of a coordinate change. More precisely, the components of $\phi^* X$ at $\phi(p)$ in a coordinate system $\{y^\alpha\}$ will be equal to the components of X at p in coordinates $\{x^\alpha\}$ where $y^\alpha(q) = x^\alpha(\phi^{-1}(q))$. As such, if we had chosen to represent all our tensor equations in terms of components in some coordinate system, it would be impossible to tell (from the transformation law alone) whether we had performed the diffeomorphism (1.11) or simply changed coordinates. Therefore, because the physical content of a tensor field's *components* cannot depend on which coordinate system it is expressed in, so the physical content of *tensor fields* cannot depend on the action of (1.11). Thus, just as in general relativity, we find that Grishchuk's formulation contains the group of diffeomorphisms $\phi\colon \mathcal{M} \to \mathcal{M}$ as a gauge freedom.

The γ-transformation

Besides the diffeomorphism gauge transformation (DGT), it is also possible to use a diffeomorphism $\phi\colon \mathcal{M} \to \mathcal{M}$ to define a transformation that reflects the range of flat-metrics γ^{ab}, and gravitational fields h^{ab}, consistent with a particular g^{ab}; we apply the diffeomorphism to γ^{ab} alone, and demand that h^{ab} compensate in such a way that g^{ab} remains unchanged:

$$\begin{aligned} & \gamma^{ab} \to (\phi^*\gamma)^{ab}, \\ & h^{ab} \to h'^{ab} = \frac{\sqrt{-\gamma}}{\sqrt{-\phi^*\gamma}}\left(\gamma^{ab} + h^{ab}\right) - (\phi^*\gamma)^{ab}, \\ \Rightarrow \qquad & g^{ab} \to g^{ab}. \end{aligned} \tag{1.12}$$

To be consistent with the field equations, if we are to include matter fields $M^{a\dots}{}_{b\dots}$ in the theory, we must make them similarly invariant:

$$M^{a\dots}{}_{b\dots} \to M^{a\dots}{}_{b\dots}. \tag{1.13}$$

It is easy to verify that the flatness of γ^{ab} is maintained by this map, as $\check{R}^a{}_{bcd} \to (\phi^*\check{R})^a{}_{bcd}$ and $\phi^* 0 = 0$.[10] It should be noted that the replacement $\gamma^{ab} \to (\phi^*\gamma)^{ab}$ does not represent a coordinate change, but is a map between two different metric tensors. Obviously, because both metrics are flat, we can always find coordinates for each such that their components are those of the Minkowski matrix $\eta^{\alpha\beta} = \mathrm{diag}(+1, -1, -1, -1)$, but while $\gamma^{\alpha\beta} = \eta^{\alpha\beta}$ in some coordinates $\{x^\alpha\}$, in general $(\phi^*\gamma)^{\alpha\beta} = \eta^{\alpha\beta}$ in a different set of coordinates $\{y^\alpha\}$.

A key feature of the γ-transformation (1.12) is that it allows us to distinguish between the two types of tensors in Grishchuk's formulation: those that can be constructed in

[9] In this chapter we use the *abstract index notation* developed by Penrose and Rindler [66], so that the Roman indices of γ^{ab} indicate the tensor 'slots' of the metric, and do not refer to components of the tensor in any coordinate system. Thus the 'effect' of a coordinate transformations is completely invisible to a tensor equation notated with abstract indices. To notate the matrix of components of a tensor such as γ^{ab} in coordinates $\{x^\alpha\}$ we use Greek indices: $\gamma^{\alpha\beta} \equiv \gamma^{ab}(\mathrm{d}x^\alpha)_a(\mathrm{d}x^\beta)_b$.

[10] These transformations form a subgroup of a larger group of transformations for which $\gamma^{ab} \to \gamma'^{ab}$ (still flat) and h^{ab} compensates such that g^{ab} is held fixed. Because this larger group does not relate so simply to the diffeomorphism gauge freedom, it is not discussed here.

standard GR, and 'non-GR' tensors, which cannot. Because g^{ab} is invariant under (1.12), all GR tensors (which must be expressible in terms of g^{ab}, ∇_a and $M^{a\dots}{}_{b\dots}$ only) will be likewise unchanged:

$$\text{GR:} \quad A^{a\dots}{}_{b\dots} \to A^{a\dots}{}_{b\dots}. \tag{1.14}$$

Thus, any tensor which is not invariant under *all* transformations of the form (1.12) must be non-GR:

$$\text{non-GR:} \quad B^{a\dots}{}_{b\dots} \to B'^{a\dots}{}_{b\dots} \neq B^{a\dots}{}_{b\dots}, \tag{1.15}$$

for some γ-transformation.

From this identification, and the formula (1.10), we can confirm our suspicions that t^{ab} is a non-GR quantity: under a γ-transformation (1.12), the g's in the square brackets are untouched, but the $\check{;}$ derivatives are transformed according to

$$\begin{aligned} & \check{\nabla}_a \to \check{\nabla}'_a, \\ \text{where} \qquad & \check{\nabla}_a \gamma^{bc} = 0, \\ \text{and} \qquad & \check{\nabla}'_a (\phi^* \gamma)^{bc} = 0. \end{aligned} \tag{1.16}$$

Although there may be some ϕ for which the transformation of $1/\gamma$ in (1.10) cancels the effects of transformation of $\check{\nabla}_a$, this will not happen for all ϕ.[11] Thus t^{ab} is not in general an invariant of the transformation, and must be impossible to construct in GR without introducing additional structure in the form of γ^{ab}.

Clearly, it is important to know whether the γ-transformation should be thought of as a gauge transformation, or as map between physically inequivalent systems. This is not a trivial problem, however, because we must be careful to avoid the tacit assumption that the GR metric g^{ab} describes everything about the gravitational field. Because g^{ab} is invariant under (1.12), the physics traditionally thought of as spacetime 'geometry' (and, in the flat-space view, are the observable effects of h^{ab} on particle worldlines, rods and clocks) must be left invariant also. Thus, comparing the γ-dependence of t^{ab} with the γ-independence of spacetime 'geometry'[12], we can immediately conclude that that t^{ab} cannot be determined by spacetime 'geometry' alone. However, it does not immediately follow that t^{ab} is an unphysical tensor, as we must seriously examine the possibility that gravity is more than just g^{ab}, and that in performing the γ-transformation we have altered something physical about the system that standard general relativity simply does not 'see'.

If we suppose that (1.12) does effect a physically meaningful change, we must conclude that every physical system is associated with a 'true' γ^{ab}, or at least with a class of

[11] To demonstrate this rigorously it is sufficient to show that t^{ab} is not invariant under infinitesimal γ-transformations; this calculation is performed in appendix 1.A.

[12] We insist on writing 'geometry' in inverted commas because although the phenomena to which we are referring are traditionally deemed to be the result of spacetime geometry, we must stress that this interpretation is not endorsed by Grishchuk's formulation. The term 'geometry' in this sense should simply be taken as a short-hand for the observable predictions shared by general relativity and the flat-space formalism.

physically equivalent flat-metrics $\{\gamma^{ab}\}$ that is smaller than the complete space spanned by all possible γ-transformations. The question is, given a physical system, how can we know when we have chosen the correct γ^{ab}? Clearly, no 'geometric' measurements can ever reveal which γ^{ab} is hidden beneath the g^{ab} metric, because 'geometric' phenomena are invariant under the γ-transformation. The only possibility of revealing γ^{ab} empirically would be if we could directly measure a non-GR tensor like t^{ab}. However, to assume that such a measurement could be carried out would make our logic circular, as for that to be possible the tensor would certainly need to be *physically meaningful*, and it is the truth of precisely this assertion that we have been trying to determine!

Even if we cannot rely on an empirical method to reveal the 'true' flat-metric γ^{ab} of a particular physical system, there may still be a systematic way to *define* one, given knowledge of quantities we can measure. Such a definition would pick out a 'canonical' γ^{ab} and we would be forbidden from performing γ-transformations because the new γ^{ab} would no longer be canonical.[13] The situation is analogous to the following question in electrostatics: what is the potential V at a particular point x? Even though we can never measure this quantity directly, we can still define a canonical potential $V(x)$ in a natural and systematic way by demanding that $V \to 0$ as the distance from the sources $r \to \infty$, or equivalently, as the electric charges $q_i \to 0$. In the same sense that we have $V = 0$ (everywhere) synonymous with the absence of electric charges, we would certainly hope that we could choose a canonical γ^{ab} such that $h^{ab} = 0$ (everywhere) is synonymous with the absence of matter fields. Indeed, given a GR metric g^{ab} that satisfies the Einstein field equations with a matter energy-momentum tensor T^{ab} as the source, we can write:

$$g^{ab} = g^{ab}(T^{cd}), \tag{1.17}$$

and define the canonical flat metric by

$$\gamma^{ab} = g^{ab}(T^{cd})|_{T^{cd}=0 \text{ (everywhere)}}. \tag{1.18}$$

For example, we could view the Schwarzschild spacetime with central mass M as a family of spacetimes $g^{ab}(M)$ and identify γ^{ab} with $g^{ab}(0)$. For any other prescription for the canonical γ^{ab} there will arise the following peculiar situation: in the absence of matter, despite spacetime 'geometry' being flat, g^{ab} will be not be equal to γ^{ab}, and we will still have to use a non-zero h^{ab} field to convert between these two different flat metrics. In this sense (1.18) is the only natural prescription for a canonical flat metric.

However, it turns out that even this effort cannot force us to abandon (1.12) as a genuine gauge transformation, as (1.18) does not behave correctly under some diffeomorphism gauge transformations (DGTs). To see this, start with a GR metric $g^{ab} = g^{ab}(T^{cd})$ and a canonical flat metric defined by (1.18). Now, consider a family of diffeomorphisms $\{\phi_f : \mathcal{M} \to \mathcal{M} \quad \forall f \in \mathbb{R}\}$ such that ϕ_0 is the identity diffeomorphism: $\phi_0(p) = p \;\; \forall p \in \mathcal{M}$.

[13] One might expect $\gamma^{\alpha\beta} = \eta^{\alpha\beta}$ to be a perfectly good definition for a canonical flat-metric; however this does not really fix γ^{ab} at all, it only begs the question: in which coordinate system do we insist that this equation holds?

We change nothing physical about this system by performing a DGT with ϕ_f for any value of the parameter f, and we are free to have the value of f determined by some functional of T^{cd} such that $T^{cd} = 0$ (everywhere) gives $f = 0$. Then, having performed this DGT, we can calculate the canonical flat metric again:

$$\gamma'^{ab} = g'^{ab}(T'^{cd})|_{T'^{cd}=0} = \left[(\phi_f^* g)^{ab}(T'^{cd})\right]_{T'^{cd}=0} = (\phi_0^*\gamma)^{ab} = \gamma^{ab}. \tag{1.19}$$

Thus, our DGT, coupled with our definition of the natural canonical flat metric, has had the following effect:

$$\begin{aligned} g^{ab} &\to (\phi_f^* g)^{ab}, \\ \gamma^{ab} &\to \gamma^{ab}, \\ h^{ab} &\to h'^{ab} = \frac{\sqrt{-\phi_f^*\gamma}}{\sqrt{-\gamma}}\left((\phi_f^*\gamma)^{ab} + (\phi_f^* h)^{ab}\right) - \gamma^{ab}, \\ M^{a\dots}{}_{b\dots} &\to (\phi_f^* M)^{a\dots}{}_{b\dots}. \end{aligned} \tag{1.20}$$

Whereas, under the DGT, we should have recovered $\gamma^{ab} \to (\phi_f^*\gamma)^{ab}$ and $h^{ab} \to (\phi_f^* h)^{ab}$. We are left with a choice: either we completely abandon the idea of a natural canonical γ^{ab} on the grounds that it is not covariant under all DGTs (and thus accept that the γ-transformation (1.12) *is* a gauge transformation), or we agree that this 'γ-fixed' transformation (1.20) is on equal footing with a DGT and is therefore another gauge transformation of the formalism. Of course, this is not really a choice at all, as the γ-fixed transformation has precisely the same effect as performing a diffeomorphism gauge transformation with ϕ_f and then a γ-transformation with $(\phi_f)^{-1}$; thus, by agreeing that (1.20) is a gauge transformation, one has agreed that the γ-transformation is one also.

The key to this argument is that because the prescription (1.18) does not pick γ^{ab} in a diffeomorphism covariant fashion[14], we retain the ability to perform γ-transformations through our choice of which diffeomorphism gauge we use to express the $T^{ab} = 0$ spacetime when we apply the definition for the canonical flat metric.

It is interesting to note that when Grishchuk refers to the gauge transformations of his formalism in [40], he appears to mean the γ-fixed variety: in appendix 1.A we calculate that the effect of an infinitesimal γ-fixed transformation on h^{ab} is

$$h^{ab} \to h'^{ab} = h^{ab} + \left(\xi^c(\gamma^{ab} + h^{ab})\right)_{;c} - 2\xi^{(a}{}_{;c}\left(\gamma^{b)c} + h^{b)c}\right) \tag{1.21}$$

and on setting $\gamma^{\alpha\beta} = \eta^{\alpha\beta}$ (which can either be viewed as a coordinate choice, given γ^{ab}, or a choice of γ^{ab} given some coordinate system) we recover

$$h'^{\alpha\beta} = h^{\alpha\beta} + \eta^{\alpha\beta}\xi^\lambda{}_{,\lambda} + (h^{\alpha\beta}\xi^\lambda)_{,\lambda} - 2\xi^{(\alpha,\beta)} - 2\xi^{(\alpha}{}_{,\lambda}h^{\beta)\lambda}, \tag{1.22}$$

which is equation (38) of [40].

Thus we must finally conclude that the γ-transformation (1.12) *is* a gauge transformation of Grishchuk's formalism, and that not only is the flat metric γ^{ab} unobservable, it

[14] We implicitly picked a gauge when we wrote g^{ab} as a *particular* solution of the field equations with source T^{cd} in (1.17).

is impossible to define a 'canonical' choice of γ^{ab} in a diffeomorphism gauge covariant, systematic, and natural fashion.

1.3.2 Transformation Properties of t^{ab}

We have demonstrated that the γ-transformation should be thought of as a map between different mathematical representations of the same physical system. As t^{ab} is not invariant under this gauge change (i.e. non-GR) we might be suspicious that this 'energy-momentum tensor' has no physical significance. However, before we dismiss t^{ab}, it is worth considering the following possibility: even though t^{ab} is not invariant under γ-transformations, could the transformed tensor t'^{ab}, somehow, have the same *physical content* as the untransformed tensor t^{ab}? After all, we see exactly this behaviour for a DGT: no tensor field is *invariant* under (1.11), however we can consider tensor fields to be *covariant* under this transformation (and their physical content unaltered) because they allow for the construction of gauge invariant quantities.[15] We must therefore consider the possibility that the γ-transformation law for t^{ab} constitutes some form of 'generalised covariance' that would allow gauge invariant quantities to be constructed.

Of course, the expected form of these invariants rather depends on what one supposes the physical content of t^{ab} to be. If it is, indeed, an energy momentum tensor, then an observer with 4-velocity u^a would expect to 'find' some energy density $\rho = t^{ab}u^c u^d g_{ac} g_{bd}$, or possibly $\rho = t^{ab}u^c u^d \gamma_{ac}\gamma_{bd}$. It is easy to check that neither of these quantities is invariant under a γ-transformation, despite the fact that we were forced to conclude that these transformations do not alter whatsoever the physical system we are examining. From this we deduce that, whatever physical meaning t^{ab} may have, since it cannot define a meaningful energy-density in the standard way, it is definitely not an energy-momentum tensor.

Infinitesimal transformations

It is instructive to examine the transformation properties of t^{ab} for an arbitrary infinitesimal gauge transformation. We proceed by constructing a diffeomorphism very close to the identity by Lie dragging tensor fields along an infinitesimal vector field ξ^a:

$$(\phi^* X)^{a\dots}{}_{b\dots} = X^{a\dots}{}_{b\dots} + (\mathcal{L}_\xi X)^{a\dots}{}_{b\dots}, \tag{1.23}$$

[15] All measurements necessarily correspond to scalars, thus the action of a DGT is simply to move these scalars to different points of $\mathcal{M}$. Because all the worldlines of observers and test particles are similarly displaced, the correlations between these scalars will be diffeomorphism gauge invariant.

where $\mathcal{L}_\xi$ is the Lie derivative along ξ^a. Under a γ-fixed gauge transformation for an infinitesimal diffeomorphism ϕ defined by (1.23), we find that $t^{ab} \to t'^{ab}$, where

$$\kappa t'^{ab} = \kappa\left(t^{ab} + (\mathcal{L}_\xi t)^{ab}\right) + \left[\xi^e{}_{\check{;}e}\left(\frac{g}{\gamma}\left(g^{ab}g^{cd} - g^{a(c}g^{d)b}\right)\right)_{\check{;}c}\right]_{\check{;}d} - \xi^e{}_{\check{;}c\check{;}d}\left(\frac{g}{2\gamma}\left(g^{ab}g^{cd} - g^{ac}g^{db}\right)\right)_{\check{;}e}. \tag{1.24}$$

This result is calculated in appendix 1.A. An important point of (1.24) is that, unlike the γ-fixed behaviour of a GR field ($A \to A + \mathcal{L}_\xi A$), the transformation law for t^{ab} includes second derivatives of ξ. Thus, in a qualitative sense, the new t'^{ab} (evaluated at some point $p \in \mathcal{M}$) seems to depends much more on the details of the transformation than a GR quantity would; certainly the complex formula (1.24) cannot be interpreted as some simple algebraic or geometric operation. If we imagine producing a finite transformation by 'exponentiating' (1.24) then the GR part of the transformation $t^{ab} + (\mathcal{L}_\xi t)^{ab}$ would correspond (loosely speaking) to a diffeomorphism '$\phi^* = \mathrm{e}^{\mathcal{L}_\xi}$' which would, to first order in ξ, only depend on ξ and its first derivatives. The extra terms in (1.24), once exponentiated, would vastly increase our freedom to determine t'^{ab} at any particular p, possibly enough to set $t'^{ab}(p) = 0$ for any t^{ab}. If this were indeed shown to be the case, then t^{ab} could hardly represent a meaningful *local* property of any field.

A particularly undesirable feature of (1.24) is that t'^{ab} is not determined by ξ^a and t^{ab} alone; we also need to know the tensor $[(g/2\gamma)(g^{ab}g^{cd} - g^{ac}g^{db})]_{\check{;}e}$ from which t^{ab} has been constructed. This detail seems to preclude the assembly of invariants from t^{ab} and observer worldlines alone.[16]

Finite transformations

To study the effect of finite gauge transformations on t^{ab}, we focus on the Schwarzschild spacetime with a central point-mass M. Working in natural units ($c = G = 1$) and suppressing the abstract indices on the coordinate differentials $(\mathrm{d}x^\alpha)_a$, we write the GR metric as

$$g_{\alpha\beta}\mathrm{d}x^\alpha\mathrm{d}x^\beta = \frac{1}{(f_1g_1 - f_2g_2)^2}\Big((g_1^2 - g_2^2)\mathrm{d}t^2 + 2(f_1g_2 - f_2g_1)\mathrm{d}t\mathrm{d}r - (f_1^2 - f_2^2)\mathrm{d}r^2\Big) - r^2(\mathrm{d}\theta^2 + \sin^2\theta\mathrm{d}\phi^2), \tag{1.25}$$

where $\{f_1, f_2, g_1, g_2\}$ are functions of r and t only. Birkhoff's theorem [16] shows the Schwarzschild spacetime to be the *only* spherically symmetric vacuum solution to the Einstein equations; thus for any choice of $\{f_i, g_i\}$ consistent with $R^{ab} = 0$, the metric given by (1.25) represents the Schwarzschild spacetime. This form of g^{ab} will be particularly

[16] Because non-GR tensors can be combined to form GR tensors, it will always be possible to 'add in' some combination of γ^{ab}, h^{ab}, and $\check{\nabla}_a$ to create a gauge invariant quantity from t^{ab}. However, in this case we should not associate the invariants with t^{ab} by itself, but instead with the larger GR object we have assembled.

useful for the present discussion, as it will allow us to choose explicitly the 'gauge' in which to express the gravitational field. To illustrate this point, we record below the recipes for the commonly used representations of the Schwarzschild spacetime.

	Standard Schwarzschild	Advanced Eddington-Finkelstein	Painlevé-Gullstrand
f_1	$1/\sqrt{1-2M/r}$	$1+M/r$	1
f_2	0	M/r	0
g_1	$\sqrt{1-2M/r}$	$1-M/r$	1
g_2	0	$-M/r$	$-\sqrt{2M/r}$

As we have emphasised, there is no unique γ^{ab} hidden beneath the metric defined in (1.25). However, for the sake of concreteness, we fix the flat-metric as

$$\gamma_{\alpha\beta}\mathrm{d}x^\alpha \mathrm{d}x^\beta = \mathrm{d}t^2 - \mathrm{d}r^2 - r^2(\mathrm{d}\theta^2 + \sin^2\theta \mathrm{d}\phi^2), \tag{1.26}$$

so that altering the functions $\{f_i, g_i\}$ will give rise to γ-fixed transformations.[17]

In order to proceed, we remove some of the gauge freedom by demanding

$$\begin{aligned} 1. \qquad f_2 &= 0, \\ 2. \qquad \partial_t g_{\alpha\beta} &= 0. \end{aligned} \tag{1.27}$$

Then we find that the vacuum field equations $R^{ab} = 0$ enforce

$$f_1 g_1 = C, \tag{1.28}$$

$$g_1^2 - g_2^2 = 1 - 2M/r, \tag{1.29}$$

where C and M are constants, and we have identified the latter as the central mass by comparison with the Standard Schwarzschild and Painlevé-Gullstrand gauges. Under these conditions we find that all the components of t^{ab} vanish apart from the energy density:

$$t^{\alpha i} = t^{i\alpha} = 0, \tag{1.30}$$

$$t^{00} = -\frac{g_1^3 - g_1 + 2r\partial_r g_1}{g_1^3 r^2}. \tag{1.31}$$

This last formula makes manifest the large space of gauge equivalent energy-momentum tensors associated with the Schwarzschild spacetime, even after we have removed a large portion of gauge freedom by demanding (1.27). Notice in particular that the Standard Schwarzschild gauge yields

$$t^{00} = -\left(\frac{2M}{r(r-2M)}\right)^2, \tag{1.32}$$

[17] Equally we could have arranged for this process to run in the opposite direction. Starting with the standard form of the Schwarzschild metric in (t, r, θ, ϕ) coordinates, we could have performed a coordinate transformation to a system (T, R, θ, ϕ) that preserved the spherical symmetry. Working in these coordinates, a seeming natural choice of the flat-metric would have been $\gamma_{\alpha\beta}\mathrm{d}x'^\alpha \mathrm{d}x'^\beta = \mathrm{d}T^2 - \mathrm{d}R^2 - R^2(\mathrm{d}\theta^2 + \sin^2\theta \mathrm{d}\phi^2)$, a different tensor from the one defined by (1.26). The choice of coordinates used to represent g^{ab} would therefore determine γ^{ab} but not alter g^{ab} itself. The effect would be that of a γ-transformation.

whereas, in the Painlevé-Gullstrand gauge

$$t^{00} = 0. \tag{1.33}$$

The gauge equivalence of these two results leaves little room for a physical interpretation of this energy-momentum tensor. Because t^{ab} can be made to vanish everywhere by a gauge transformation, it cannot possibly convey any more gauge-invariant information than to tell us that this spacetime is empty of whatever it is that t^{ab} represents. While this is not unreasonable per se (as t^{ab} might only be sensitive to gravitational radiation or some other phenomena absent from the Schwarzschild spacetime) it then becomes very difficult to justify why the energy-momentum tensor should be non-zero in any gauge at all. This uncomfortable situation would force us to identify a whole host of non-trivial energy-momentum tensors with emptiness, of which (1.31) are only a small fraction.

As the Advanced Eddington-Finkelstein gauge has $f_2 \neq 0$, we cannot use (1.31) to calculate the energy-momentum tensor. Instead, we take the general formula (1.10) as our starting point, and recover $t^{ab} = 0$, just as we found in the Painlevé-Gullstrand gauge. This agreement suggests that the non-zero energy-momentum tensor (1.32) might only be an artefact of the 'horizon' present in the Standard Schwarzschild gauge: in the (t, r, θ, ϕ) coordinate system picked out by γ_{ab}, the components of the GR metric $g_{\alpha\beta}$ are singular at $r = 2M$. In contrast, Painlevé-Gullstrand and Advanced Eddington-Finkelstein are *global* gauges: the components $g_{\alpha\beta}$ are regular everywhere but at the origin. While a coordinate singularity is admissible within differential geometry, in Grishchuk's flat-space picture this would correspond to an infinite 'gravitational field' h^{ab}, which could be deemed unphysical. This line of reasoning allows us to reject (1.32) because it was derived in a gauge which transforms the gravitational field to infinity at some points, and we would then hope to confirm that the physical result ($t^{ab} = 0$) applies in all global gauges. Unfortunately, this turns out to be impossible, as we show by means of a counter-example. Consider a family of gauges parametrised by λ:

$$\begin{aligned}
f_1 &= \sqrt{r/(r + \lambda M)}, \\
f_2 &= 0, \\
g_1 &= \sqrt{(r + \lambda M)/r}, \\
g_2 &= \sqrt{(2 + \lambda)M/r}.
\end{aligned} \tag{1.34}$$

It is easy to check that these obey the restrictions (1.27) and the vacuum field equations (1.28) and (1.29). Furthermore, for $\lambda \geq 0$, $g_{\alpha\beta}$ defined by (1.25) is regular everywhere but the origin. Using (1.31) we find that (apart from $\lambda = 0$ which is just Painlevé-Gullstrand again) the energy-momentum tensor is non-zero:

$$t^{00} = -\left(\frac{\lambda M}{r(r + \lambda M)}\right)^2. \tag{1.35}$$

Not only can we make $t^{00}(r)$ take on a wide range of values by adjusting λ, we also note that in the limit $\lambda \to \infty$, we have the disconcerting situation of a non-zero energy density that is independent of M.

In light of all these results, it appears highly unlikely that the behaviour of t^{ab} would permit the extraction of gauge invariant information and allow us to view this tensor as maintaining some physical content under gauge transformations.

1.4 Conclusion

The formulation of gravity presented in [40] succeeds in recasting general relativity as a flat-space theory of a symmetric tensor field. While we do not find fault with the formalism itself, we assert that care must by taken in its interpretation, as we believe we have demonstrated that only those quantities which can be defined solely in terms of GR tensors are of any physical importance. The physically insignificant content of the flat-space formalism is a consequence of an unmeasurable field γ^{ab} which is not uniquely determined by the requirement that it be a flat metric tensor.

There are in principle two ways to deal with the non-uniqueness of γ^{ab}: 1. Pick a particular flat metric and declare that this is the immutable 'correct' choice, to be used in all situations; 2. Allow γ^{ab} to depend somehow on the physical system we are describing, or how we have chosen to represent the system mathematically.

The problem with the first stance is that the theory still retains γ-fixed gauge transformations. To see this, note that equation (53) of [11] expresses the equivalence of Grishchuck's equations of motion ($r_{ab} \equiv -P^c{}_{ab\check{;}c} - P^c{}_{ad}P^d{}_{bc} + \frac{1}{3}P_a P_b = 0$) with the Einstein field equations:

$$R_{ab}[g] = \check{R}_{ab}[\gamma] + r_{ab}[h, \gamma]. \tag{1.36}$$

Babak and Grishchuk interpret this relation as follows: given a flat-metric γ^{ab}, an h^{ab} that satisfies $r_{ab} = 0$ will enforce $R_{ab} = 0$, establishing the agreement with GR. However, one can always use this equation to make the converse argument: given a flat-metric γ^{ab}, a g^{ab} which solves $R_{ab} = 0$ will enforce Grishchuk's equation $r_{ab} = 0$. As $R_{ab}[\phi^* g] = 0$ if $R_{ab}[g] = 0$, we can construct a whole range of solutions $\{h'^{ab} : r_{ab}[h'] = 0\}$ from h^{ab} simply by applying diffeomorphisms to g^{ab}. Because we declared γ^{ab} to be immutable, these new solutions will correspond to γ-fixed transformations of h^{ab}. Crucially, as $g'^{ab} = \phi^* g^{ab}$, no 'geometric' experiment can tell any h'^{ab} apart from from h^{ab}. Thus, without a method to measure a non-GR quantity directly, we have to conclude that these new solutions represent physically equivalent systems, and that the γ-fixed transformation is a gauge transformation of the theory.

The second stance appears to be able to dodge this argument, because one can claim that we should have applied the same diffeomorphism to γ^{ab} that we applied to g^{ab}, forcing us to perform a harmless DGT instead of a γ-fixed transformation. However, if we take this view, we will need a heuristic for deriving γ^{ab} from measurable quantities, otherwise we will never know where to start with the 'correct' pairing (g^{ab}, γ^{ab}). In order that this heuristic be consistent with arbitrary DGTs (which are gauge transformations of any tensorial theory) a prescription for which $T^{ab} = 0 \Rightarrow h^{ab} = 0$ inevitably leads us to identify

γ-transformations as gauge transformations anyway, because we are free to represent the $T^{ab} = 0$ limit in any diffeomorphism gauge we choose.

Accepting that γ-transformations and γ-fixed transformation are maps between different mathematical representations of the same physical system, we conclude that the exotic gauge transformation properties of t^{ab} cannot allow us to interpret this tensor as a local measure of the energy and momentum content of the gravitational field. Although t^{ab} is a perfectly legitimate mathematical construction, its dependence on the unmeasurable and non-unique tensor γ^{ab} renders it ill-defined, and devoid of physical meaning.

1.A Appendix: Infinitesimal Transformations

Here we calculate how h^{ab}, t^{ab}, and $\check{\nabla}_a$ change under transformations defined by diffeomorphisms infinitely close to the identity: $\phi^* = 1 + \mathcal{L}_\xi$. In this limit, the γ-fixed transformation (1.20) for h^{ab} is

$$\begin{aligned} h^{ab} &\to h'^{ab} \\ h'^{ab} &= (-\gamma)^{-1/2} (1 + \mathcal{L}_\xi) \left(\sqrt{-\gamma} (\gamma^{ab} + h^{ab}) \right) - \gamma^{ab} \\ &= (-\gamma)^{-1/2} \mathcal{L}_\xi \left(\sqrt{-\gamma} (\gamma^{ab} + h^{ab}) \right) + h^{ab}. \end{aligned} \tag{1.37}$$

Thus,

$$\begin{aligned} \delta h^{ab} &\equiv h'^{ab} - h^{ab} \\ &= (\gamma^{ab} + h^{ab})(-\gamma)^{-1/2} \mathcal{L}_\xi \sqrt{-\gamma} + \mathcal{L}_\xi (\gamma^{ab} + h^{ab}) \\ &= (\gamma^{ab} + h^{ab}) \xi^c{}_{\check{;}c} + \xi^c \left(\gamma^{ab} + h^{ab} \right)_{\check{;}c} - 2\xi^{(a}{}_{\check{;}c} \left(\gamma^{b)c} + h^{b)c} \right) \\ &= \left(\xi^c (\gamma^{ab} + h^{ab}) \right)_{\check{;}c} - 2\xi^{(a}{}_{\check{;}c} \left(\gamma^{b)c} + h^{b)c} \right), \end{aligned} \tag{1.38}$$

proving (1.21).

To calculate the behaviour of the energy-momentum tensor under a γ-fixed transformation, we define the tensor

$$Y^{abcd} \equiv \frac{g}{\gamma} g^{(a[b)} g^{(c]d)} = \frac{g}{2\gamma} \left(g^{ab} g^{cd} - g^{a(c} g^{d)b} \right), \tag{1.39}$$

so that

$$\kappa t^{ab} = Y^{abcd}{}_{\check{;}c\check{;}d}. \tag{1.40}$$

Under the γ-fixed transformation, $t^{ab} \to t'^{ab}$ where

$$\begin{aligned} \kappa t'^{ab} &= \left[\gamma^{-1} (1 + \mathcal{L}_\xi) \left(g g^{(a[b)} g^{(c]d)} \right) \right]_{\check{;}c\check{;}d} \\ &= \kappa t^{ab} + \left[\mathcal{L}_\xi \left(Y^{abcd} \right) - g g^{(a[b)} g^{(c]d)} \mathcal{L}_\xi (\gamma^{-1}) \right]_{\check{;}c\check{;}d} \\ &= \kappa t^{ab} + \left[Y^{abcd}{}_{\check{;}e} \xi^e - 2\xi^{(a}{}_{\check{;}e} Y^{b)ecd} - 2Y^{abe(c} \xi^{d)}{}_{\check{;}e} + 2Y^{abcd} \xi^e{}_{\check{;}e} \right]_{\check{;}c\check{;}d}. \end{aligned} \tag{1.41}$$

In contrast, were t^{ab} a GR tensor, under the γ-fixed transformation we would have $t^{ab} \to t^{ab} + \mathcal{L}_\xi t^{ab}$, with

$$\kappa \mathcal{L}_\xi t^{ab} = \xi^e Y^{abcd}{}_{\check{;}c\check{;}d\check{;}e} - 2\xi^{(a}{}_{\check{;}e} Y^{b)ecd}{}_{\check{;}c\check{;}d}. \tag{1.42}$$

Thus, the non-GR part of $\kappa t'^{ab}$ is

$$\begin{aligned}
\Delta(\kappa t^{ab}) &\equiv \kappa \left(t'^{ab} - t^{ab} - \mathcal{L}_\xi t^{ab} \right) \\
&= \xi^e{}_{\check{;}c\check{;}d} Y^{abcd}{}_{\check{;}e} + 2\xi^e{}_{\check{;}c} Y^{abcd}{}_{\check{;}e\check{;}d} - 2\xi^{(a}{}_{\check{;}e\check{;}c\check{;}d} Y^{b)ecd} - 4\xi^{(a}{}_{\check{;}e\check{;}c} Y^{b)ecd}{}_{\check{;}d} \\
&\quad - 2\left[Y^{abe(c}\xi^{d)}{}_{\check{;}e} - Y^{abcd}\xi^e{}_{\check{;}e} \right]_{\check{;}c\check{;}d}.
\end{aligned} \tag{1.43}$$

Note that the third and fourth terms vanish because $Y^{abcd} = -Y^{acbd}$ and $\check{\nabla}_a$ operators commute. Expanding out the derivatives acting on the square brackets, then cancelling and collecting like terms, we arrive at

$$\begin{aligned}
\Delta(\kappa t^{ab}) &= 2\xi^e{}_{\check{;}e} Y^{abcd}{}_{\check{;}c\check{;}d} - \xi^e{}_{\check{;}c\check{;}d} Y^{abcd}{}_{\check{;}e} + 2\xi^e{}_{\check{;}e\check{;}d} Y^{abcd}{}_{\check{;}c} \\
&= 2\left[\xi^e{}_{\check{;}e} Y^{abcd}{}_{\check{;}c} \right]_{\check{;}d} - \xi^e{}_{\check{;}c\check{;}d} Y^{abcd}{}_{\check{;}e}.
\end{aligned} \tag{1.44}$$

Replacing Y^{abcd} with its definition (1.39), the transformation law (1.24) immediately follows.

Because the γ-fixed transformation is simply a DGT with ϕ followed by a γ-transformation with ϕ^{-1}, it is easy to use this result to calculate the behaviour of t^{ab} under an infinitesimal γ-transformation:

$$\begin{aligned}
(\gamma\text{-fixed})_\phi t^{ab} &= (\gamma\text{-trans})_{\phi^{-1}} (\mathrm{DGT})_\phi \, t^{ab} \\
&= (\gamma\text{-trans})_{\phi^{-1}} \left(t^{ab} + \mathcal{L}_\xi t^{ab} \right) \\
&= (\gamma\text{-trans})_{\phi^{-1}} \, t^{ab} + \mathcal{L}_\xi t^{ab},
\end{aligned} \tag{1.45}$$

for infinitesimal ξ. Thus, under a γ-transformation, κt^{ab} becomes

$$\begin{aligned}
\kappa t'^{ab} = \kappa t^{ab} - &\left[\xi^e{}_{\check{;}e} \left(\frac{g}{\gamma} \left(g^{ab} g^{cd} - g^{a(c} g^{d)b} \right) \right)_{\check{;}c} \right]_{\check{;}d} \\
&+ \xi^e{}_{\check{;}c\check{;}d} \left(\frac{g}{2\gamma} \left(g^{ab} g^{cd} - g^{ac} g^{db} \right) \right)_{\check{;}e},
\end{aligned} \tag{1.46}$$

which clearly demonstrates that t^{ab} is non-GR.

For completeness, we calculate how the derivative operator $\check{\nabla}_a$ changes under an infinitesimal γ-transformation. We shall proceed without using the flatness of γ^{ab}, in order that the result be in its most general form; only at the end we will set $\check{R}^a{}_{bcd} = 0$ to recover the formula applicable here. According to (1.16), we have

$$\begin{aligned}
\check{\nabla}_a \gamma_{bc} &= 0, \\
\check{\nabla}'_a \left(\gamma_{bc} + \mathcal{L}_\xi \gamma_{bc} \right) &= 0.
\end{aligned} \tag{1.47}$$

Any two torsionless derivative operators can be related by a symmetric connection; thus, in the same way one might write the figurative relation " $\nabla = \partial + \Gamma$ " to define the GR Christoffel symbols, we write " $\check{\nabla}' = \check{\nabla} + E$ " to define a connection $E^a{}_{bc} = E^a{}_{cb}$ between $\check{\nabla}'_a$ and $\check{\nabla}_a$. By continuity $E^a{}_{bc}$ must be at least first order in ξ, so (1.47) becomes:

$$\begin{aligned} \check{\nabla}_a \left(\gamma_{bc} + 2\gamma_{d(b}\check{\nabla}_{c)}\xi^d\right) - 2\gamma_{d(b}E^d{}_{c)a} = 0, \\ \Rightarrow \quad E_{(bc)a} = \left(\check{\nabla}_a\check{\nabla}_{(c}\xi_{b)}\right). \end{aligned} \tag{1.48}$$

However, because E_{abc} is symmetric in its last two indices,

$$\begin{aligned} E_{(ab)c} + E_{(ac)b} - E_{(bc)a} = E_{abc}, \\ \Rightarrow \quad E^a{}_{bc} = \gamma^{ae}\left(E_{(eb)c} + E_{(ec)b} - E_{(bc)e}\right). \end{aligned} \tag{1.49}$$

Substituting (1.48) into the right-hand-side and reorganising the derivatives, we find

$$E^a{}_{bc} = \check{\nabla}_{(b}\check{\nabla}_{c)}\xi^a + \gamma^{ae}\left(\check{\nabla}_{[b}\check{\nabla}_{e]}\xi_c + \check{\nabla}_{[c}\check{\nabla}_{e]}\xi_b\right). \tag{1.50}$$

Finally, using the defining property of the Riemann tensor, $\check{\nabla}_{[a}\check{\nabla}_{b]}\xi_c = -\frac{1}{2}\check{R}^d{}_{cab}\xi_d$, we arrive at the following compact formula:

$$E^a{}_{bc} = \left(\delta^a_d\check{\nabla}_{(b}\check{\nabla}_{c)} - \check{R}^a{}_{(bc)d}\right)\xi^d. \tag{1.51}$$

In the case where γ^{ab} is flat, this becomes

$$E^a{}_{bc} = \xi^a{}_{\check{;}b\check{;}c}. \tag{1.52}$$

Chapter 2

Bootstrapping Gravity: A Consistent Approach to Energy-Momentum Self-Coupling

2.1 Introduction

It is a standard view in particle physics that the non-linearity of a field theory, such as those of Yang and Mills, can be equated with the notion that the field in question carries the charge of the very interaction it mediates. This idea has been brought to bear on gravity many times, and various arguments [21, 25, 26, 36, 41, 50, 63] aim to derive general relativity from a linear starting point by coupling gravity to the energy and momentum of all fields, including the gravitational field itself. Despite the conventional wisdom that this self-coupling process is already well understood, Padmanabhan has uncovered a number of serious problems with the standard arguments [64]. Although we postpone an examination of Padmanabhan's analysis to appendix 2.A, it suffices to express here what is, in our view, his most pertinent observation: one cannot start with *linear gravity*, the Fierz-Pauli massless spin-2 action [37, 64], and generate the higher-order corrections of general relativity by coupling the gravitational field to its own Hilbert energy-momentum tensor. More succinctly: one cannot derive the Einstein equations by bootstrapping gravitons[1] to their own energy and momentum.

To clarify the content of this observation, consider a perturbative expansion of the Einstein field equations $G_{\alpha\beta} = \kappa T^{\text{matter}}_{\alpha\beta}$ about a Minkowski background: $g_{\alpha\beta} = \eta_{\alpha\beta} + h_{\alpha\beta}$. Working to second-order in $h_{\alpha\beta}$, we obtain

$$G^{(1)}_{\alpha\beta} = -G^{(2)}_{\alpha\beta} + \kappa T^{\text{matter}}_{\alpha\beta}, \tag{2.1}$$

where the numbers in parenthesis denote the powers of $h_{\alpha\beta}$ the term contains. Because $G^{(1)}_{\alpha\beta} = 0$ is the equation of motion for a massless spin-2 field $h_{\alpha\beta}$, the right-hand side of (2.1) can be interpreted as this field's source. Thus a satisfying physical picture suggests

[1] In discussions of this nature, the word *graviton* is often used as a shorthand for the classical massless spin-2 field. We follow this convention to cohere with the literature, but stress that this graviton is in no way quantum mechanical. What is actually being referred to is a *gravitational wave*, a classical fluctuation in the geometry of spacetime.

itself: the gravitational field $h_{\alpha\beta}$ is induced by the energy-momentum tensor of *all* fields $T_{\alpha\beta} = T^{\text{matter}}_{\alpha\beta} + t_{\alpha\beta}$, where $t_{\alpha\beta}$ is gravity's own energy-momentum tensor, identified as $-G^{(2)}_{\alpha\beta}/\kappa$. In actuality, however, this description cannot be formulated in a straightforward manner. Although the Fierz-Pauli action S_{FP} is typically used to prescribe the dynamics of a massless spin-2 field, its Hilbert energy-momentum tensor[2]

$$t_{\alpha\beta} \equiv \frac{-1}{\sqrt{-\gamma}} \frac{\delta S_{\text{FP}}}{\delta \gamma^{\alpha\beta}}, \tag{2.2}$$

is *not* proportional to $G^{(2)}_{\alpha\beta}$, and thus cannot be used as the source-term for the second-order field equations. As an alternative approach, one could introduce energy-momentum self-coupling at the level of the action: because $t_{\alpha\beta}$ is a function of $h_{\alpha\beta}$, adding the self-coupling term $t_{\alpha\beta}h^{\alpha\beta}$ to the Lagrangian yields a different result from adding $t_{\alpha\beta}$ directly to the equations of motion. Unfortunately, this procedure also fails to generate $-G^{(2)}_{\alpha\beta}/\kappa$ in the field equations.

Padmanabhan claims that these realisations bring to light a previously neglected object $S^{\alpha\beta}$ (see appendix 2.A) which appears to codify the self-coupling of the gravitational field. Unfortunately, this object has many undesirable features: it is not a tensor under general coordinate transformations, has no clear physical interpretation, and fails to reveal any equivalence between the coupling of gravity to matter, and gravity to itself.

We propose an alternative solution to this apparent inconsistency: the action for the graviton is not the Fierz-Pauli action but is instead S_2 given by (2.4), possessing a non-minimally coupled term that vanishes when the (vacuum) background equations are enforced.[3] We shall demonstrate that the energy-momentum tensor of this action is the correct second-order contribution to the equation of motion, and furthermore, that this action provides the starting point for a straightforward energy-momentum self-coupling procedure that generates the Einstein-Hilbert action (modulo surface terms) to *arbitrary order*. We conclude the discussion by extending our formalism to non-vacuum spacetimes.

Throughout the chapter we employ the abstract index notation [79, §2.4], with lower-case Roman indices indicating a tensor's 'slots', and Greek indices serving to enumerate its components in a particular coordinate system. The metric has signature $(-,+,+,+)$, $\kappa \equiv 8\pi G/c^4$, and the Riemann and Ricci tensor are defined with the following conventions: $R^a{}_{bcd}v^b \equiv 2\nabla_{[c}\nabla_{d]}v^a$, $R_{ab} \equiv R^c{}_{acb}$.

[2]Although other definitions of the energy-momentum tensor exist (see §2.2.3) we must define $t_{\alpha\beta}$ according to the Hilbert's prescription (2.2) in order to maintain the analogy with $T^{\text{matter}}_{\alpha\beta}$. This definition requires that S_{FP} be "covariantised" (represented in arbitrary coordinates using a *flat* metric $\gamma_{\alpha\beta}$) and a functional derivative taken with respect to the metric. It is important to realise that even though $\gamma_{\alpha\beta}$ is flat, the arbitrary variations $\delta\gamma_{\alpha\beta}$ required to construct the functional derivative inevitably explore *curved* metrics in a neighbourhood of $\gamma_{\alpha\beta}$. Thus "covariantisation" is not really sufficient: the action must be generalised to a curved background spacetime. One of the key aims of this chapter is to generalise S_{FP} to curved spacetime in such a way that energy-momentum self-coupling is consistent with general relativity.

[3]More precisely, S_2 is the action for the graviton in a background spacetime with metric in some small neighbourhood of the solutions of the vacuum field equations. We use the term *vacuum* to signify a region without matter; this does not necessarily imply the absence of spacetime curvature.

2.2 Graviton Action

Contrary to the standard approach, we represent the gravitational field as a perturbation h^{ab} of the *inverse* physical metric g^{ab} from the background $\check{g}^{ab}$:

$$g^{ab} = \check{g}^{ab} + h^{ab}. \tag{2.3}$$

This expression is *exact* in that we have not neglected terms $O(h^2)$; in contrast, the physical metric $g_{ab} = \check{g}_{ab} - h^{cd}\check{g}_{ca}\check{g}_{db} + O(h^2)$. Following this convention, we use the contravariant field h^{ab}, rather than h_{ab}, as the fundamental dynamical variable of the action.[4] In general we will write "caron" marks over tensors derived solely from the background geometry, and adopt the usual notational convenience of raising and lowering indices with $\check{g}^{ab}$ and $\check{g}_{ab}$.[5]

We posit that the dynamics, energy and momentum of the gravitational field h^{ab}, propagating in a background spacetime with metric $\check{g}_{ab}$, are all determined (to lowest-order) by the following action:

$$S_2[\check{g}^{ab}, h^{ab}] \equiv \frac{1}{2\kappa}\int \mathrm{d}^4x\sqrt{-\check{g}}h^{ab}(\widehat{G}_{abcd} + \check{H}_{abcd})h^{cd}, \tag{2.4}$$

where

$$\widehat{G}_{abcd} \equiv \tfrac{1}{2}(\check{g}_{a(c}\check{g}_{d)b} - \check{g}_{ab}\check{g}_{cd})\check{\nabla}^2 - \check{\nabla}_{(c}\check{g}_{d)(a}\check{\nabla}_{b)} + \tfrac{1}{2}\check{g}_{ab}\check{\nabla}_{(c}\check{\nabla}_{d)} + \tfrac{1}{2}\check{g}_{cd}\check{\nabla}_{(a}\check{\nabla}_{b)} \tag{2.5}$$

is a differential operator representing the linearised Einstein tensor (see appendix 2.B) and

$$\check{H}_{abcd} \equiv \tfrac{1}{2}\check{R}(\check{g}_{ac}\check{g}_{db} + \tfrac{1}{2}\check{g}_{ab}\check{g}_{cd}) - \check{R}_{ab}\check{g}_{cd}. \tag{2.6}$$

[4]Any metric theory of gravity will have an ambiguity as to which variable $g \in \{g^{ab}, g_{ab}, \sqrt{-g}g^{ab}, \ldots\}$ should be identified as the true "gravitational field". Such a distinction is of no physical consequence and is largely unnecessary for a non-perturbative calculation; however for the present discussion we are forced to single out a particular field variable for the expansion $g = \check{g} + h$. Our aim is to connect gravity to the particle physics notion of a spin-2 field and elucidate a simple energy-momentum self-coupling scheme that generates general relativity; to this end we are required to pick $g \in \{g^{ab}, g_{ab}\}$ as it is only for these that h is a genuine spin-2 field, i.e. a symmetric tensor (not a tensor density) with (lowest-order) infinitesimal gauge transformation $\delta h^{ab} = 2\check{\nabla}^{(a}\epsilon^{b)}$. Fortunately, it is precisely for $g \in \{g^{ab}, g_{ab}\}$ that the necessary energy-momentum self-coupling is its most simple: $h^{ab}t_{ab}$ (see §2.3). These considerations provide no criteria for choosing the metric over its inverse as our expansion variable, and while this choice only trivially alters the perturbation theory at first-order ($h^{ab} \leftrightarrow -h_{ab}$) to second-order (the relevant order for S_2, t_{ab}, and $G^{(2)}_{ab}$) the two definitions of the h-field differ by a term of the form $h^{ac}h^{b}{}_{c}$. Our choice of $g = g^{ab}$ is preferable for this chapter because it simplifies the mathematics of the action and energy-momentum tensor. The reason for this is explored in §2.3.5, and stems from the fact that any Lagrangian for pure gravity must contain more factors of g^{ab} than g_{ab} in order that all the derivatives ∂_a be contracted; thus an expansion in $g = g^{ab}$ will be algebraically simpler. Indeed, this observation still holds when coupling gravity to a scalar field ϕ or a 1-form A_a, and thus taking $g = g^{ab}$ simplifies many of the calculations of the non-vacuum case also (see §2.4).

[5]The only exception to this rule is the physical metric and its inverse, for which $g^{ab} \neq g_{cd}\check{g}^{ac}\check{g}^{db}$, but rather $g^{ab}g_{bc} = \delta^a_c$.

While $\check{H}_{abcd}$ has no obvious geometric interpretation, we intend to show that its contribution to the action is necessary for the consistency of energy-momentum self-coupling with general relativity. Further motivation for this ansatz is given in section 2.3.

Naturally, if we are to obtain general relativity without at first assuming it, we must begin by considering the graviton in a *flat* background spacetime. Nevertheless, we will see from the formalism of section 2.3 that (provided we use S_2 to describe the graviton) energy-momentum self-coupling generates the Einstein-Hilbert action even when the background is not flat; $\check{g}^{ab}$ need only satisfy the weaker condition

$$\check{G}_{ab} \equiv \check{R}_{ab} - \tfrac{1}{2}\check{g}_{ab}\check{R} = 0. \tag{2.7}$$

While this equation expresses the generality of the analysis that is to follow, it should be stressed that no knowledge of (2.7) will be required to assemble the Einstein-Hilbert action order by order: a flat background will serve as a perfectly satisfactory starting point.[6] No matter which background we use, however, it is absolutely crucial that we refrain from inserting this particular metric (or even equation (2.7)) into the action, thereby reducing S_2 to $\frac{1}{2\kappa}\int \mathrm{d}^4x\sqrt{-\check{g}}h^{ab}\widehat{G}_{abcd}h^{cd}$. This is because we will need to be able to perform arbitrary variations of $\check{g}^{ab}$, not just those consistent with $\check{R}_{abcd} = 0$ or $\check{R}_{ab} = 0$, to construct the energy-momentum tensor for h^{ab}. That said, it will be instructive to temporarily ignore this advice so that we may relate S_2 to the Fierz-Pauli action.

2.2.1 The Fierz-Pauli Action

For a flat background, $\check{H}_{abcd}$ vanishes, and we can choose coordinates $\{x^\alpha\}$ such that $\check{g}^{\alpha\beta} = \eta^{\alpha\beta}$ and evaluate S_2 as a functional of the components $h^{\alpha\beta}$. Integrating by parts and discarding surface terms, we find that S_2 reduces to $\frac{-1}{2\kappa}\int \mathrm{d}^4x\mathcal{L}_{\mathrm{FP}}$, where

$$\mathcal{L}_{\mathrm{FP}} = \tfrac{1}{2}\partial_\lambda h_{\alpha\beta}\partial^\lambda h^{\alpha\beta} - \tfrac{1}{2}\partial_\lambda h\partial^\lambda h - \partial_\lambda h^{\alpha\beta}\partial_\alpha h_\beta{}^\lambda + \partial_\alpha h\partial_\beta h^{\alpha\beta} \tag{2.8}$$

is the Fierz-Pauli Lagrangian [64].[7] Modulo surface terms and an overall rescaling, $\mathcal{L}_{\mathrm{FP}}$ is the unique specially relativistic Lagrangian for a symmetric tensor field $h^{\alpha\beta}$ that is invariant under the infinitesimal gauge transformation $\delta h^{\alpha\beta} = 2\partial^{(\alpha}\epsilon^{\beta)}$ (see [64] for proof); hence it is the Lagrangian for the graviton (massless spin-2 field) in flat spacetime.

Starting from (2.8), we can "covariantise" $\mathcal{L}_{\mathrm{FP}}$ by making the replacements $\eta_{\alpha\beta} \to \check{g}_{\alpha\beta}$, $\partial_\alpha \to \check{\nabla}_\alpha$ and multiplying by $\sqrt{-\check{g}}$. This process obviously generates a unique manifestly covariant Lagrangian density if $\check{g}^{ab}$ is flat, as in this case the procedure is equivalent to representing the same Lagrangian in arbitrary coordinates. However, for the

[6] Of course, once the self-coupling procedure is complete, and the Einstein-Hilbert action has been assembled starting from the graviton on a flat background, we will be in a excellent position to justify (2.7), as this is precisely the field equation (applied to the background) that we will have derived. With hindsight, then, we can see there was nothing special about our flat-space starting point: we may begin with any *one* solution to (2.7) and use energy-momentum self-coupling to derive the action (and field equation) that defines *all* the others.

[7] Here and elsewhere we use the customary shorthand $h \equiv h^a{}_a \equiv h^{ab}\check{g}_{ab}$.

purposes of calculating the energy-momentum tensor (via arbitrary variations of $\check{g}^{ab}$) it will be necessary to generalise $\mathcal{L}_{\text{FP}}$ to arbitrary backgrounds, and for a curved metric the covariantisation procedure is ambiguous. To see this, observe that we can transmute the third term of (2.8) by twice integrating by parts:

$$\partial_\lambda h^{\alpha\beta}\partial_\alpha h_\beta{}^\lambda \leftrightarrow \partial_\alpha h^{\alpha\beta}\partial_\lambda h_\beta{}^\lambda. \tag{2.9}$$

However this equivalence relies on the commutativity of partial derivatives, and does not occur for the covariant derivatives of a curved background; instead, integration by parts yields

$$\check{\nabla}_c h^{ab}\check{\nabla}_a h_b{}^c \leftrightarrow \check{\nabla}_a h^{ab}\check{\nabla}_c h_b{}^c - h^{ca}h^b{}_c\check{R}_{ab} - h^{ab}h^{cd}\check{R}_{acdb}. \tag{2.10}$$

Thus we are forced to make a seemingly arbitrary choice: do we to covariantise (2.8) as written, or should we do so after performing (2.9)? These two possibilities determine Lagrangians which differ by $h^{ca}h^b{}_c\check{R}_{ab} + h^{ab}h^{cd}\check{R}_{acdb}$; they lead to different (first-order) equations of motion if the background is curved,[8] and determine different energy-momentum tensors even if the background is flat.[9] This last problem is discussed by Padmanabhan [64], and is one of his many non-trivial objections to the conventional wisdom that general relativity is the unique energy-momentum self-coupled limit of the flat-space massless spin-2 field.

A greater problem than this ambiguity, however, is that neither choice (nor an admixture) leads to general relativity after coupling it to its own energy-momentum. As we shall see in section 2.3, the contribution from $h^{ab}\check{H}_{abcd}h^{cd}$ is necessary to achieve this, and it is impossible to use the covariantising ambiguity to produce this tensor because it does not contain $h^{ab}h^{cd}\check{R}_{acdb}$. Instead, the presence of $\check{H}_{abcd}$ represents a rather different coupling ambiguity faced when moving from a flat background to a curved one. Typically we would invoke the Einstein equivalence principal to banish from the action terms coupling matter fields and Ricci tensors; we would argue that, working in locally inertial coordinates about a point p, the Lagrangian at p should have the same form as the Lagrangian in flat spacetime. This amounts to a minimal coupling procedure: once we have covariantised a specially relativistic Lagrangian, the job of coupling the field to the gravity is complete. However, while this rule may make sense to curve the background spacetime of a spin-2 field that is "just another matter-field" and has nothing to do with gravitation, it is far from clear that the principal should hold for the graviton, for which it was only ever a convenient fiction to think of as a tensor field propagating over a background geometry.

In summary, the Fierz-Pauli action is insufficient to determine S_2 for an arbitrary background geometry; the principal of equivalence fails to give a unique solution, and cannot justify all the contributions necessary for an energy-momentum self-coupling procedure

[8] The first-order field equation only describes the spacetime perturbations of general relativity if the ambiguous term is covariantised to become $\check{\nabla}_c h^{ab}\check{\nabla}_a h_b{}^c$; see §2.2.2 and appendix 2.B.

[9] Note that all other terms of $\mathcal{L}_{\text{FP}}$ are invariant under the operation that generated (2.9) so do not introduce further ambiguity.

consistent with general relativity. However, it was never our aim to construct general relativity from $\mathcal{L}_{\mathrm{FP}}$, and we do not pretend to be able to derive a curved spacetime theory of gravity from purely specially relativistic concepts. S_2 will serve as our starting point, and the only significance we shall ascribe $\mathcal{L}_{\mathrm{FP}}$ is that of a special case.

2.2.2 Field Equations

Leaving the Fierz-Pauli action behind, we refocus our attention on S_2 and begin the process of deriving its advertised connection to general relativity. First, we shall calculate the associated field equations. As usual, the equations of motion are derived from the condition that their solutions be stationary configurations of S_2 with respect to variations in the dynamical field h^{ab}. As we will have no cause to vary $\breve{g}^{ab}$ in the derivation, we can enforce the background equations (2.7) immediately and discard $\breve{H}_{abcd}$. Next, observe that $\widehat{G}_{abcd}$ is "self-conjugate": for any tensor fields A^{ab} and B^{ab}

$$\int \mathrm{d}^4x\sqrt{-\breve{g}}A^{ab}\widehat{G}_{abcd}B^{cd} = \int \mathrm{d}^4x\sqrt{-\breve{g}}B^{ab}\widehat{G}_{abcd}A^{cd}, \tag{2.11}$$

provided either A^{ab} or B^{ab} has compact support. Therefore, holding $\breve{g}^{ab}$ constant and performing a variation δh^{ab} (a symmetric tensor field with compact support) gives rise to a variation in the action

$$\delta S_2 = \frac{1}{\kappa}\int \mathrm{d}^4x\sqrt{-\breve{g}}\delta h^{ab}\widehat{G}_{abcd}h^{cd}. \tag{2.12}$$

As $\widehat{G}_{abcd}$ is already symmetric in its first two indices, we can conclude that the equation of motion is

$$\frac{1}{\sqrt{-\breve{g}}}\frac{\delta S_2}{\delta h^{ab}} = \kappa^{-1}\widehat{G}_{abcd}h^{cd} = 0. \tag{2.13}$$

The centrally important feature of this equation is that $\widehat{G}_{abcd}h^{cd} = G^{(1)}_{ab}$, the linear approximation to the Einstein tensor under the inverse metric expansion (2.3). This is particularly easy to verify for the special case of a flat background in Lorentzian coordinates, but is shown to hold more generally for vacuum backgrounds in appendix 2.B. Thus S_2 prescribes the correct first-order equation of motion for the graviton. In the next section we show that by adding the energy-momentum tensor t_{ab} of h^{ab} (determined by S_2) to the right hand side of (2.13) we successfully generate the Einstein field equations correct to *second-order*.[10]

[10]Of course, the resulting field equation will no longer be a stationary configuration of the action S_2. In order that this self-coupled equation of motion can be derived from the principle of stationary action it will be necessary to introduce a third-order correction to the action S_3. Naturally, S_3 will alter the energy-momentum tensor of h^{ab} by a term $O(h^3)$; however, seemingly by miracle, this will be precisely the *third-order* part of the Einstein field equations. This process continues indefinitely and is explained systematically in §2.3. For the moment we content ourselves with exploring the theory to second-order only.

2.2.3 Energy-momentum Tensor

We will now calculate the energy-momentum tensor of the graviton and relate it to the second-order contribution to the Einstein field equations. We follow Hilbert's prescription and define the energy-momentum tensor as a functional derivative of the action with respect to the (background) metric:

$$t_{ab} \equiv \frac{-1}{\sqrt{-\breve{g}}} \frac{\delta S_2}{\delta \breve{g}^{ab}}, \tag{2.14}$$

where h^{ab} (rather than h_{ab} or $h^a{}_b$) is to be held constant when taking this derivative, as this is the field we have taken to be the fundamental dynamical variable.[11]

As an aside, it is worth contrasting the variational definition (2.14) with Noether's (canonical) energy-momentum tensor:

$$t^{\mu\nu}_{\text{can}} \equiv \frac{\partial \mathcal{L}}{\partial(\partial_\mu h^{\alpha\beta})} \partial^\nu h^{\alpha\beta} - \eta^{\mu\nu}\mathcal{L}, \tag{2.15}$$

comprising the four conserved currents associated with the invariance of the Lagrangian $\mathcal{L}$ under rigid spacetime translations. The canonical tensor cannot be used in the present discussion for a number of reasons. Firstly, it is not uniquely determined by the action for h^{ab}: as it depends directly on the Lagrangian, we are free to alter $t^{\mu\nu}_{\text{can}}$ by adding a four-divergence to $\mathcal{L}$, without changing either the dynamics of h^{ab} or S_2. Secondly, we require a *symmetric* tensor to act as the source for the first-order field equation (2.13), but the canonical tensor need not have this property.[12] Lastly, Noether's definition does not naturally generalise to curved spacetime in such a way that $t^{\mu\nu}_{\text{can}}$ inherits a *covariant* conservation law [51]. None of these issues arise with t_{ab}, and in any case our aim has been to connect the coupling between matter and gravity found in general relativity with a perturbative coupling of gravity to itself; it is the Hilbert energy-momentum tensor of matter, not the canonical tensor, that appears in the full Einstein field equations as the gravitational source. For these reasons we discard the canonical tensor and henceforth refer to t_{ab}, following Hilbert's prescription (2.14), as the energy-momentum tensor of h^{ab}.

To begin the calculation of t_{ab}, we divide the action into two pieces $S_2 = S_{2G} + S_{2H}$:

$$S_{2G} \equiv \frac{1}{2\kappa} \int \mathrm{d}^4x \sqrt{-\breve{g}} h^{ab} \widehat{G}_{abcd} h^{cd}, \tag{2.16}$$

$$S_{2H} \equiv \frac{1}{2\kappa} \int \mathrm{d}^4x \sqrt{-\breve{g}} h^{ab} \check{H}_{abcd} h^{cd}. \tag{2.17}$$

It will be convenient to perform the functional derivative (2.14) on these two components separately. Focusing first on S_{2G}, we integrate by parts[13] so as to remove the second

[11] In later sections, the tensor written here as t_{ab} will be notated t^2_{ab} to indicate that it is the energy-momentum contribution from the second-order action S_2 only. Here we need not make this distinction.

[12] It is true that the canonical tensor can be *made* symmetric by adding to it an identically conserved "correction" $\partial_\alpha \phi^{\mu[\nu\alpha]}$, a function of h^{ab} that cancels the antisymmetric part of $t^{\mu\nu}_{\text{can}}$. However, if we allow this sort of ad hoc adjustment of the energy-momentum tensor, we only exacerbate the problem of non-uniqueness.

[13] More precisely, one adds to the integrand a divergence of the form $\partial_a(\sqrt{-\breve{g}}[h\check{\nabla}h]^a) = \sqrt{-\breve{g}}\check{\nabla}_a[h\check{\nabla}h]^a$ that alters S_2 only by a function of the fields on the boundary (or at infinity) and thus may be neglected for the purposes of functional variation.

derivatives from the integrand:

$$S_{2G} = \frac{-1}{2\kappa}\int \mathrm{d}^4x\sqrt{-\breve{g}}\breve{\nabla}_c h^{ab}\breve{\nabla}_d h^{ef} K_{ab\ \ ef}^{\ \ c\ \ \ d}, \tag{2.18}$$

for which we have introduced the abbreviation

$$\begin{aligned} K_{ab\ \ ef}^{\ \ c\ \ \ d} &\equiv \frac{1}{2}\Big(\breve{g}^{cd}\breve{g}_{a(e}\breve{g}_{f)b} - \breve{g}^{cd}\breve{g}_{ab}\breve{g}_{ef} - 2\delta^c_{(e}\breve{g}_{f)(a}\delta^d_{b)} + \delta^c_{(e}\delta^d_{f)}\breve{g}_{ab} + \delta^d_{(a}\delta^c_{b)}\breve{g}_{ef}\Big) \\ &= K_{ba\ \ ef}^{\ \ c\ \ \ d} = K_{ab\ \ fe}^{\ \ c\ \ \ d} = K_{ef\ \ ab}^{\ \ d\ \ \ c}. \end{aligned} \tag{2.19}$$

An infinitesimal variation in the inverse background metric $\delta\breve{g}^{ab}$, vanishing on the boundary of the integral, induces a variation in the action

$$\begin{aligned} \delta S_{2G} = \frac{-1}{2\kappa}\int \mathrm{d}^4x\sqrt{-\breve{g}}\Bigg[&\delta\breve{g}^{pq}\breve{\nabla}_c h^{ab}\breve{\nabla}_d h^{ef}\left(\frac{\partial K_{ab\ \ ef}^{\ \ c\ \ \ d}}{\partial \breve{g}^{pq}} - \frac{1}{2}\breve{g}_{pq}K_{ab\ \ ef}^{\ \ c\ \ \ d}\right) \\ &+ 4\breve{\nabla}_c h^{ab} C^{(e}_{\ \ sd} h^{f)s} K_{ab\ \ ef}^{\ \ c\ \ \ d}\Bigg], \end{aligned} \tag{2.20}$$

where

$$\begin{aligned} C^a_{\ bc} &\equiv \tfrac{1}{2}\breve{g}^{ad}\left(\breve{\nabla}_b\delta\breve{g}_{cd} + \breve{\nabla}_c\delta\breve{g}_{bd} - \breve{\nabla}_d\delta\breve{g}_{bc}\right) \\ &= -\tfrac{1}{2}\left(2\delta^a_p\delta^r_{(b}\breve{g}_{c)q} - \breve{g}^{ar}\breve{g}_{bp}\breve{g}_{qc}\right)\breve{\nabla}_r\delta\breve{g}^{pq} \end{aligned} \tag{2.21}$$

is the connection that arises from the variation of the covariant derivative: $\nabla_{\breve{g}+\delta\breve{g}} = \breve{\nabla}+C$. We can move the covariant derivatives off $\delta\breve{g}^{pq}$ in the connection term using integration by parts, and arrive at an equation of the form $\delta S_{2G} = \int \mathrm{d}^4x\ \delta\breve{g}^{pq}[\ldots]_{pq}$; the tensor density in square brackets is then the functional derivative we seek:

$$\begin{aligned} \frac{\kappa}{\sqrt{-\breve{g}}}\frac{\delta S_{2G}}{\delta\breve{g}^{pq}} = &\frac{-1}{2}\breve{\nabla}_c h^{ab}\breve{\nabla}_d h^{ef}\left(\frac{\partial K_{ab\ \ ef}^{\ \ c\ \ \ d}}{\partial\breve{g}^{pq}} - \frac{1}{2}\breve{g}_{pq}K_{ab\ \ ef}^{\ \ c\ \ \ d}\right) \\ &- \breve{\nabla}_r\Big(\breve{\nabla}_c h^{ab}\Big(K_{ab\ \ (p|f|q)}^{\ \ c}h^{rf} + K_{ab\ \ f(p}^{\ \ c\ \ \ \ r}h_{q)}^{\ \ f} - K_{ab\ \ \ f(p}^{\ \ cr}h_{q)}^{\ \ f}\Big)\Big). \end{aligned} \tag{2.22}$$

Meanwhile, S_{2H} varies by

$$\delta S_{2H} = \frac{1}{2\kappa}\int \mathrm{d}^4x\sqrt{-\breve{g}}\delta\breve{R}_{ab}\left(\tfrac{1}{2}\breve{g}^{ab}\left(\tfrac{1}{2}h^2 + h_{cd}h^{cd}\right) - h^{ab}h\right), \tag{2.23}$$

where we have used the background equation (2.7) (*after* the variation) to remove the terms proportional to $\breve{R}_{ab}$; these would only be significant if we intended to perform further variations in the metric. Now, because

$$\begin{aligned} \delta\breve{R}_{ab} &= 2\breve{\nabla}_{[c}C^c_{\ \ b]a} \\ &= \left(\tfrac{1}{2}\breve{g}^{rs}\breve{g}_{ap}\breve{g}_{qb} + \tfrac{1}{2}\delta^r_{(a}\delta^s_{b)}\breve{g}_{pq} - \delta^r_p\delta^s_b\breve{g}_{aq}\right)\breve{\nabla}_r\breve{\nabla}_s\delta\breve{g}^{pq}, \end{aligned} \tag{2.24}$$

when we (twice) integrate by parts to alleviate $\delta\breve{g}^{ab}$ of its covariant derivatives, we generate a second-order differential operator

$$\widehat{R}_{pqab} \equiv \tfrac{1}{2}\breve{g}_{a(p}\breve{g}_{q)b}\breve{\nabla}^2 + \tfrac{1}{2}\breve{g}_{pq}\breve{\nabla}_{(a}\breve{\nabla}_{b)} - \breve{\nabla}_{(a}\breve{g}_{b)(p}\breve{\nabla}_{q)}, \tag{2.25}$$

with the property

$$\int \mathrm{d}^4x\sqrt{-\breve{g}}\delta\check{R}_{ab}A^{ab} = \int \mathrm{d}^4x\sqrt{-\breve{g}}\delta\breve{g}^{pq}\widehat{R}_{pqab}A^{ab} \tag{2.26}$$

for all A^{ab}. Therefore, we can conclude from (2.23) that

$$\frac{\kappa}{\sqrt{-\breve{g}}}\frac{\delta S_{2H}}{\delta\breve{g}^{pq}} = \frac{1}{2}\widehat{R}_{pqab}\left(\tfrac{1}{2}\breve{g}^{ab}\left(\tfrac{1}{2}h^2 + h_{cd}h^{cd}\right) - h^{ab}h\right). \tag{2.27}$$

Finally, we have only to combine equations (2.22) and (2.27), expand out all the products and derivatives, and assemble the outcome into a formula for t_{ab} as a function of $\check{\nabla}_c h^{ab}$. This is a straightforward but arduous calculation, and as such we chose to complete it with a computer algebra package. The result is

$$\begin{aligned}\kappa t_{pq} = {}& \tfrac{1}{4}\breve{g}_{pq}\Big(h\check{\nabla}_a\check{\nabla}_b h^{ab} + 2h^{ab}\check{\nabla}_a\check{\nabla}_b h - 2h_{ab}\check{\nabla}^2 h^{ab} - h\check{\nabla}^2 h - \tfrac{1}{2}\check{\nabla}_a h\check{\nabla}^a h - \tfrac{5}{2}\check{\nabla}_c h_{ab}\check{\nabla}^c h^{ab}\\ &+ \check{\nabla}_c h_a{}^b\check{\nabla}_b h^{ac} + 2\check{\nabla}_a h\check{\nabla}_b h^{ab}\Big) + \tfrac{1}{4}h\check{\nabla}_{(p}\check{\nabla}_{q)}h - \tfrac{1}{2}h_{pq}\check{\nabla}^2 h + \tfrac{1}{4}h\check{\nabla}^2 h_{pq} + h_{a(p}\check{\nabla}^2 h_{q)}{}^a\\ &- \tfrac{1}{2}h^{ab}\check{\nabla}_a\check{\nabla}_b h_{pq} + \tfrac{1}{2}h_{pq}\check{\nabla}_a\check{\nabla}_b h^{ab} - h_{a(p}\check{\nabla}^b\check{\nabla}_{q)}h^a{}_b + \tfrac{1}{2}h_{ab}\check{\nabla}_{(p}\check{\nabla}_{q)}h^{ab} - \tfrac{1}{2}h\check{\nabla}_a\check{\nabla}_{(p}h_{q)}{}^a\\ &+ \tfrac{1}{4}\check{\nabla}_a h\check{\nabla}^a h_{pq} + \tfrac{1}{2}\check{\nabla}_b h_{ap}\check{\nabla}^b h^a{}_q - \tfrac{1}{2}\check{\nabla}_a h_{pq}\check{\nabla}_b h^{ab} + \tfrac{3}{4}\check{\nabla}_p h_{ab}\check{\nabla}_q h^{ab} - \check{\nabla}_b h^a{}_{(p}\check{\nabla}_{q)}h_a{}^b\\ &- \tfrac{1}{2}\check{\nabla}_b h\check{\nabla}_{(p}h_{q)}{}^b + \tfrac{1}{2}\check{\nabla}_b h^a{}_p\check{\nabla}_a h^b{}_q.\end{aligned} \tag{2.28}$$

It is possible to render this formula rather more manageable by working in a gauge with $\check{\nabla}_a h^{ab} = 0$, $h = 0$:

$$\begin{aligned}\kappa t_{pq} = {}& \breve{g}_{pq}\Big(\tfrac{1}{4}\check{\nabla}_c h_a{}^b\check{\nabla}_b h^{ac} - \tfrac{5}{8}\check{\nabla}_c h_{ab}\check{\nabla}^c h^{ab} - \tfrac{1}{2}h_{ab}\check{\nabla}^2 h^{ab}\Big) + h_{a(p}\check{\nabla}^2 h_{q)}{}^a - \tfrac{1}{2}h^{ab}\check{\nabla}_a\check{\nabla}_b h_{pq}\\ &- h^{bc}\check{R}_{abc(p}h_{q)}{}^a + \tfrac{1}{2}h_{ab}\check{\nabla}_{(p}\check{\nabla}_{q)}h^{ab} + \tfrac{1}{2}\check{\nabla}_b h_{ap}\check{\nabla}^b h^a{}_q + \tfrac{3}{4}\check{\nabla}_p h_{ab}\check{\nabla}_q h^{ab} - \check{\nabla}_b h^a{}_{(p}\check{\nabla}_{q)}h_a{}^b\\ &+ \tfrac{1}{2}\check{\nabla}_b h^a{}_p\check{\nabla}_a h^b{}_q,\end{aligned} \tag{2.29}$$

but we will not need this partially gauge-fixed result for this present chapter.[14]

Our task now is to compare t_{ab} with $G^{(2)}_{ab}$ and demonstrate that the energy-momentum self-coupling of h^{ab} (determined by S_2) is consistent with general relativity. Details of the calculation of $G^{(2)}_{ab}$ can be found in appendix 2.B; the conclusion is

$$G^{(2)}_{ab} = -\kappa t_{ab} + O(h^3), \tag{2.30}$$

and thus, to second-order, the vacuum Einstein field equations are

$$\widehat{G}_{abcd}h^{cd} = \kappa t_{ab} \tag{2.31}$$

as advertised.

[14]Gauge transformations are covered in §2.3.3; we note here only that because t_{ab} is not invariant under the infinitesimal gauge transformation $\delta h^{ab} = 2\check{\nabla}^{(a}\epsilon^{b)}$, only the first formula (2.28) can be used in all gauges. Although gauge invariance would be a highly desirable property if we intended to argue that t_{ab} was a physically meaningful tensor in full general relativity, it is an impossible request to make of the tensor we seek, which should be proportional to the gauge dependent tensor $G^{(2)}_{ab}$.

As a corollary of (2.31), we can confirm Padmanabhan's observation that general relativity cannot be derived from energy-momentum self-coupling the Fierz-Pauli Lagrangian. Only once the contribution from $\check{H}_{abcd}$ is included will Einstein's gravity result from an energy-momentum self-coupled graviton. This realisation casts doubt on Mannheim's recent treatment of gravitational energy-momentum [56], in which a tensor is constructed by applying (2.14) to a covariantised Fierz-Pauli Lagrangian, rather than S_2.

2.3 Perturbative Gravity

Here we develop the formalism to uncover the root cause of the second-order energy-momentum self-coupling (2.31), and reveal how the process continues to arbitrary order. The vast majority of this section applies to any metric theory of pure gravity[15] and can be generalised to include interactions with matter (see §2.4). Only in section 2.3.5 will we commit to general relativity, fix our action $S = S_{\mathrm{EH}}$, the Einstein Hilbert action, and derive the formula (2.4) for S_2.

We shall concern ourselves with an expansion of the inverse metric g^{ab} about a non-dynamical background $\check{g}^{ab}$, which is itself an exact solution of the vacuum field equations:

$$g^{ab} = \check{g}^{ab} + \lambda h^{ab}, \tag{2.32}$$

$$0 = \frac{\delta S[\check{g}]}{\delta \check{g}^{ab}}, \tag{2.33}$$

where λ, a dimensionless expansion parameter, is constant over spacetime.

Following (2.32), the action of the exact theory $S[g]$ becomes a λ-dependent functional of $\check{g}^{ab}$ and h^{ab}, which can be Taylor expanded thus:

$$S[g] = S[\check{g} + \lambda h] = \sum_{n=0}^{\infty} \lambda^n S_n[\check{g}, h], \tag{2.34}$$

where S_n is the "n^{th} partial action" given by

$$S_n[\check{g}, h] = \frac{1}{n!} \left(\partial_\lambda^n S[\check{g} + \lambda h]\right)_{\lambda=0}. \tag{2.35}$$

The derivative ∂_λ acts on each instance of λh^{ab} in the integrand of $S[\check{g} + \lambda h]$ by Leibniz's law, removing the factor of λ. The 'bare' h^{ab} left behind may still be covered by spacetime derivatives ∂_a, but these can be moved onto the remainder of the integrand by integration by parts. This operation generates the usual functional derivative:

$$\partial_\lambda S[\check{g} + \lambda h] = \int \mathrm{d}^4x h^{ab}(x) \frac{\delta}{\delta \check{g}^{ab}(x)} S[\check{g} + \lambda h]. \tag{2.36}$$

In truth, the left hand side of this equation differs from the right by the surface term $\int \mathrm{d}^4x \partial_a J^a$ created when integrating by parts. As this is only a functional of the fields on

[15] We require only that the dynamics are determined by an action that is a coordinate-independent integral of the metric and its derivatives.

the boundary (or as $x^\mu \to \infty$ if the integral of S runs over the entire manifold) it will not contribute to equations of motion or energy-momentum tensors, the calculation of which are dependent only on variations of the field that vanish on the boundary (or have compact support). Hence these surface terms may be neglected for our present purposes.

It follows from the repeated application of (2.36) that

$$\partial_\lambda^n S[\breve{g}+\lambda h] = \left[\int \mathrm{d}^4x h^{ab} \frac{\delta}{\delta \breve{g}^{ab}}\right]^n S[\breve{g}+\lambda h], \tag{2.37}$$

and thus the partial actions (2.35) are given by

$$S_n[\breve{g}, h] = \frac{1}{n!}\left[\int \mathrm{d}^4x h^{ab} \frac{\delta}{\delta \breve{g}^{ab}}\right]^n S[\breve{g}]. \tag{2.38}$$

An important consequence of this relation is that, using S_2 as our starting point, we can generate the entire set of partial actions $\{S_n : n \geq 3\}$ by calculating

$$S_n[\breve{g}, h] = \frac{2}{n!}\left[\int \mathrm{d}^4x h^{ab} \frac{\delta}{\delta \breve{g}^{ab}}\right]^{n-2} S_2[\breve{g}, h], \tag{2.39}$$

which is possible provided S_2 is known in a *neighbourhood* of whichever particular background (a solution of (2.33)) we are interested in. Note that the first two partial actions do not contribute to the dynamics of h^{ab}: $S_0 = S[\breve{g}]$ is manifestly independent of h^{ab}, and S_1 vanishes once the background equation (2.33) has been enforced. We conclude, therefore, that S_2 contains all the information necessary to reconstruct the "dynamical" part of the action

$$S_{\mathrm{dyn}}[\breve{g}, h] \equiv \sum_{n=2}^{\infty} \lambda^n S_n[\breve{g}, h], \tag{2.40}$$

which itself contains all the dynamical information of the full action S. This is absolutely key to the calculations of section 2.2, in which we saw the first consequence of this reconstruction process, the recovery of the second-order equation of motion from an action that one would expect to encode only first-order dynamics.

2.3.1 Field Equations

In general, we could let λ be a free parameter and, on demanding $\delta S[g]/\delta g^{ab} = 0$ for fixed $\breve{g}^{ab}$, derive a λ-dependent equation of motion $E_\lambda[\breve{g}, h] = 0$ for our dynamical field h^{ab}. Any h^{ab} that solved this equation would correspond to a metric $g^{ab} = \breve{g}^{ab} + \lambda h^{ab}$ that solved the field equations *exactly*.[16] However, if we are interested in approximating small variations of the metric (i.e. the limit $\lambda h^{ab} \to 0$) we can choose some order N to which we want the equation of motion to hold:

$$\frac{\delta S[g]}{\delta g^{ab}} = O(\lambda^{N+1}). \tag{2.41}$$

[16] It is advisable to set $\lambda = 1$ before attempting to solve $E_\lambda[\breve{g}, h] = 0$, as this constant can always be absorbed into the magnitude of h^{ab}. Although this refinement was convenient for §2.2, here we shall keep λ as it provides a simple method for tracking the powers of h^{ab} in expressions and is useful as a variable for differentiation.

This is equivalent to

$$\frac{1}{\lambda}\frac{\delta S^{N+1}_{\rm dyn}[\breve{g},h]}{\delta h^{ab}} = O(\lambda^{N+1}), \tag{2.42}$$

where $S^{N+1}_{\rm dyn}$ is defined by discarding from $S_{\rm dyn}$ those terms that can be neglected in (2.41):

$$S^{N+1}_{\rm dyn}[\breve{g},h] \equiv \sum_{n=2}^{N+1} \lambda^n S_n[\breve{g},h]. \tag{2.43}$$

We shall adopt this "$N^{\rm th}$-order approximation" picture for the development of our formalism, as we can always write $N=\infty$ if we wish to discuss the exact theory.

For the sake of continuity with the previous section, we introduce the notation

$$\left.\frac{\delta S_2[\breve{g},h]}{\delta h^{ab}}\right|_{\delta S[\breve{g}]/\delta \breve{g}^{ab}=0} \equiv \kappa^{-1}\sqrt{-\breve{g}}\widehat{G}_{abcd}h^{cd}, \tag{2.44}$$

where, because S_2 is second-order in h^{ab}, $\widehat{G}_{abcd}$ will be a linear differential operator dependent only on $\breve{g}^{ab}$.[17] The equation of motion (2.42) now takes the form

$$\lambda\widehat{G}_{abcd}h^{cd} = -\frac{\kappa}{\lambda\sqrt{-\breve{g}}}\frac{\delta}{\delta h^{ab}}\sum_{n=3}^{N+1}\lambda^n S_n[\breve{g},h], \tag{2.45}$$

where it should be taken as given that terms $O(\lambda^{N+1})$ have been neglected. This is the $N^{\rm th}$-order approximation to the equation of motion for h^{ab} that is consistent with the dynamics of g^{ab} prescribed by the action S. The first-order contribution has been separated from the sum so as to evoke the picture of a wave equation $\lambda\widehat{G}_{abcd}h^{cd}=0$ with a source. In the next section we will see that the source term on the right of (2.45) is indeed the energy-momentum tensor of the field h^{ab}, neglecting terms $O(\lambda^{N+1})$.

2.3.2 Energy-momentum Tensor

First we shall demonstrate that the dynamical part of the action (2.40) can be generated from S_2 by a simple energy-momentum self-coupling procedure. Observe that, as a consequence of (2.38), we have

$$S_n[\breve{g},h] = \frac{1}{n}\int \mathrm{d}^4x\, h^{ab}\frac{\delta S_{n-1}[\breve{g},h]}{\delta \breve{g}^{ab}}. \tag{2.46}$$

Defining the $n^{\rm th}$ partial energy-momentum tensor t^n_{ab} by applying Hilbert's prescription to the $n^{\rm th}$ partial action,

$$t^n_{ab} \equiv \frac{-1}{\sqrt{-\breve{g}}}\frac{\delta S_n[\breve{g},h]}{\delta \breve{g}^{ab}}, \tag{2.47}$$

[17]The operator $\widehat{G}_{abcd}$ defined here coincides with the definition in (2.5) once $S=S_{\rm EH}$ has been fixed. This is shown in §2.3.5 by deriving S_2.

we conclude that

$$S_n[\breve{g},h] = \frac{-1}{n}\int \mathrm{d}^4x\sqrt{-\breve{g}}h^{ab}t^{n-1}_{ab}. \tag{2.48}$$

This makes manifest the energy-momentum self-coupling procedure that allows us to generate the dynamical part of the action (2.40) to arbitrary order, given only S_2. The n^{th} partial action is nothing more than the integral of the contraction of h^{ab} with the energy-momentum tensor of the previous partial action (divided by $-n$). The dynamical part of the action is therefore given by

$$S^{N+1}_{\text{dyn}}[\breve{g},h] = \lambda^2 S_2[\breve{g},h] - \int \mathrm{d}^4x\sqrt{-\breve{g}}h^{ab}\sum_{n=2}^{N}\frac{\lambda^{n+1}t^n_{ab}}{n+1}. \tag{2.49}$$

Note that, for the particular case of general relativity ($S = S_{\text{EH}}$), the background equation (2.7) also sets $S_0 = 0$, thus $S_{\text{dyn}} = S_{\text{EH}}$ (modulo surface terms) and the energy-momentum self-coupling procedure recovers the *entire* action of the full theory, not just the dynamical part.

Because of factors of $n+1$ dividing each t^n_{ab} in (2.49), it is not the case that in the action h^{ab} couples directly to its (N^{th}-order) total energy-momentum tensor, given by

$$T^N_{ab} \equiv \frac{-1}{\sqrt{-\breve{g}}}\frac{\delta S^N_{\text{dyn}}}{\delta \breve{g}^{ab}} = \sum_{n=2}^{N}\lambda^n t^n_{ab}. \tag{2.50}$$

Instead, the numerical denominators account for the $n+1$ factors of h^{ab} in $h^{ab}t^n_{ab}$, and ensure that the equations of motion do indeed have T^N_{ab} as the source. To prove this, note that for any symmetric field l^{ab} (vanishing on the boundary, or with compact support) we have

$$\begin{aligned}
\int \mathrm{d}^4x l^{ab}\frac{\delta S_n[\breve{g},h]}{\delta h^{ab}} &= \int \mathrm{d}^4x\frac{l^{ab}}{n!}\frac{\delta}{\delta h^{ab}}\left(\partial^n_\lambda S[\breve{g}+\lambda h]\right)_{\lambda=0} \\
&= \tfrac{1}{n!}\left(\partial_\mu\left(\partial^n_\lambda S[\breve{g}+\lambda(h+\mu l)]\right)_{\lambda=0}\right)_{\mu=0} \\
&= \tfrac{1}{n!}\left(\partial^n_\lambda\partial_\mu S[\breve{g}+\lambda(h+\mu l)]\right)_{\lambda=\mu=0} \\
&= \tfrac{1}{n!}\left(\partial^n_\lambda\left(\lambda\partial_\alpha S[\breve{g}+\lambda h+\alpha l]\right)\right)_{\lambda=\alpha=0},
\end{aligned}$$

where $\alpha \equiv \lambda\mu \Rightarrow \partial_\mu = \lambda\partial_\alpha$. Thus,

$$\begin{aligned}
\int \mathrm{d}^4x l^{ab}\frac{\delta S_n[\breve{g},h]}{\delta h^{ab}} &= \frac{1}{n!}\Big(\lambda\partial^n_\lambda\partial_\alpha S[\breve{g}+\lambda h+\alpha l] + n\partial^{n-1}_\lambda\partial_\alpha S[\breve{g}+\lambda h+\alpha l]\Big)_{\lambda=\alpha=0} \\
&= \frac{1}{(n-1)!}\left(\partial_\alpha\partial^{n-1}_\lambda S[\breve{g}+\lambda h+\alpha l]\right)_{\lambda=\alpha=0} \\
&= \left(\partial_\alpha S_{n-1}[\breve{g}+\alpha l,h]\right)_{\alpha=0} \\
&= \int \mathrm{d}^4x l^{ab}\frac{\delta S_{n-1}[\breve{g},h]}{\delta\breve{g}^{ab}}.
\end{aligned} \tag{2.51}$$

Hence we have the following important result:

$$\frac{\delta S_n[\breve{g},h]}{\delta h^{ab}} = \frac{\delta S_{n-1}[\breve{g},h]}{\delta\breve{g}^{ab}}. \tag{2.52}$$

Or, using definition (2.47),

$$\frac{\delta S_n[\breve{g}, h]}{\delta h^{ab}} = -\sqrt{-\breve{g}}t^{n-1}_{ab}. \tag{2.53}$$

Therefore the equation of motion (2.45) takes on the form

$$\lambda \widehat{G}_{abcd}h^{cd} = \kappa\lambda^{-1}\sum_{n=3}^{N+1}\lambda^n t^{n-1}_{ab}, \tag{2.54}$$

or, recalling (2.50),

$$\lambda \widehat{G}_{abcd}h^{cd} = \kappa T^N_{ab}. \tag{2.55}$$

We have derived the relation we sought, demonstrating that any metric theory of pure gravity can be formulated as a first-order wave equation with its own energy-momentum tensor as a source. For every $N \geq 1$, we can derive the equation of motion (2.55) by applying the variational principle to the action S^{N+1}_{dyn}; the left hand side is the wave equation for the linearised theory, and the right hand side is the energy-momentum tensor prescribed by the action S^N_{dyn}. This energy-momentum tensor is, to some extent, incomplete: it does not include the $O(\lambda^{N+1})$ contribution from the highest-order partial action S_{N+1}. This contribution could be calculated, if so desired, and added by hand to the field equations (2.55) so that the right hand side read κT^{N+1}_{ab}, but this equation would no longer be a stationary configuration of the action S^{N+1}_{dyn}. To remedy this, we could introduce a correction to the action $\lambda^{N+2}S_{N+2}$ that would generate the extra term in the equation of motion; the appropriate functional is given by (2.48) and couples h^{ab} to the highest-order partial energy-momentum tensor t^{N+1}_{ab}. But now once again the energy-momentum tensor T^{N+1}_{ab} is incomplete, and we can apply this same line of reasoning anew. So long as there is no N for which t^N_{ab} vanishes identically, this process can continue indefinitely, and as $N \to \infty$ the exact field equations are recovered, along with the action $S_{\text{dyn}} = S - S_0 - \lambda S_1$.

All that remains is to connect our formalism to the specific results of the previous section. For the sake of completeness, however, we shall first discuss the gauge symmetries of the theory, and deduce the conservation law for T^{N+1}_{ab}.

2.3.3 Gauge Transformations

Because the action $S[g]$ is a coordinate-system independent integral, any diffeomorphism $\phi : \mathcal{M} \to \mathcal{M}$ gives rise to a gauge transformation of the theory through the action of ϕ^*, the map comprising the pullback of ϕ on covector indices and the pushforward of ϕ^{-1} on vector indices:

$$S[\phi^* g] = S[g]. \tag{2.56}$$

Taylor expanding both sides about $\breve{g}^{ab}$ and applying the background equation reveals the gauge invariance of the dynamical part of the action:

$$S^{N+1}_{\text{dyn}}[\breve{g}, h'] = S^{N+1}_{\text{dyn}}[\breve{g}, h], \tag{2.57}$$

where

$$\lambda h'^{ab} \equiv \phi^* g^{ab} - \breve{g}^{ab}. \tag{2.58}$$

In the context of an N^{th}-order approximation, we must insist that $\phi^* = 1 + O(\lambda)$, otherwise these transformations will map the small metric fluctuations λh^{ab} onto fluctuations comparable in magnitude to $\breve{g}^{ab}$. We can write a general diffeomorphism of this form as $\phi^* = e^{\lambda \mathcal{L}_\xi}$, where $\mathcal{L}_\xi$ is the Lie derivative along a vector field $\xi^a = O(1)$. The gauge transformations of the theory are hence given by

$$\begin{aligned} h^{ab} &\to h'^{ab} = h^{ab} + \delta h^{ab}, \\ \delta h^{ab} &\equiv \lambda^{-1} \sum_{n=1}^{N} \frac{(\lambda \mathcal{L}_\xi)^n}{n!} \breve{g}^{ab} + \sum_{n=1}^{N-1} \frac{(\lambda \mathcal{L}_\xi)^n}{n!} h^{ab}, \end{aligned} \tag{2.59}$$

where we have discarded all terms $O(\lambda^N)$, as these will only contribute terms $O(\lambda^{N+1})$ to the equation of motion, and terms $O(\lambda^{N+2})$ to S^{N+1}_{dyn}. If we wish we can let $\xi^a = \epsilon^a$, an infinitesimal vector field, and derive the infinitesimal gauge transformation

$$\delta h^{ab} = \begin{cases} \mathcal{L}_\epsilon \left(\breve{g}^{ab} + \lambda h^{ab}\right) & N \geq 2, \\ -2\breve{\nabla}^{(a}\epsilon^{b)} & N = 1. \end{cases} \tag{2.60}$$

Because these gauge transformations (infinitesimal or otherwise) are symmetries of S^{N+1}_{dyn}, they map solutions of the equation of motion (2.55) to other solutions. We can therefore use the equation of motion to deduce the transformation law for T^N_{ab}:

$$\delta T^N_{ab} \equiv T^N_{ab}[\breve{g}, h'] - T^N_{ab}[\breve{g}, h] = \frac{\lambda}{\kappa} \widehat{G}_{abcd} \delta h^{cd}. \tag{2.61}$$

This verifies the earlier remark that the energy-momentum tensor is gauge dependent, except in the trivial case $N = 1$, for which $T^N_{ab} = 0$ by definition. It may come as a surprise that the energy-momentum tensor does not inherit the gauge invariance of the action from which it was derived. It should be stressed, however, that S^{N+1}_{dyn} is not *identically* gauge invariant: the relation (2.57) is only true when the background equation is obeyed. For general $\breve{g}^{ab}$, the diffeomorphism invariance of $S[g]$ only furnishes the gauge transformation law $\delta S^{N+1}_{\text{dyn}} = -\lambda \delta S_1$, the right-hand side of which has a non-vanishing energy-momentum tensor responsible for the variation in T^N_{ab}. Equivalently, the gauge dependence of T^N_{ab} can be seen to result from the non-commutativity of gauge transformations and the functional derivative $\delta/\delta\breve{g}^{ab}$ used to define T^N_{ab} [55]; these operations would only commute if the gauge invariance of S^{N+1}_{dyn} extended to a neighbourhood of the solutions of the background equation, rather than being confined to the solutions themselves.

2.3.4 Conservation Law

It should be expected that $S^{N+1}_{\text{dyn}}[\breve{g}, h]$ inherits the diffeomorphism invariance of $S[g]$, and that this symmetry endows the energy-momentum tensor with a covariant conservation law with respect to the background metric. The derivation proceeds in close analogy to the proof of $\nabla^a T_{ab}^{\text{matter}} = 0$ from general relativity.

We again appeal to the diffeomorphism invariance of the action (2.56) but this time expand $S[g]$ about $\breve{g}^{ab}$ (a solution of the background equation) and $S[\phi^* g]$ about $\phi^* \breve{g}^{ab}$ (which will also be a solution). The result,

$$S_{\text{dyn}}^{N+1}[\phi^* \breve{g}, \phi^* h] = S_{\text{dyn}}^{N+1}[\breve{g}, h], \tag{2.62}$$

affirms that S_{dyn}^{N+1} is diffeomorphism invariant.[18] Now let ϕ be an infinitesimal diffeomorphism: $\phi^* = 1 + \mathcal{L}_\epsilon$ for an arbitrary infinitesimal vector field ϵ^a with compact support. Then (2.62) becomes

$$0 = \int \mathrm{d}^4 x \left[\frac{\delta S_{\text{dyn}}^{N+1}}{\delta \breve{g}^{ab}} \mathcal{L}_\epsilon \breve{g}^{ab} + \frac{\delta S_{\text{dyn}}^{N+1}}{\delta h^{ab}} \mathcal{L}_\epsilon h^{ab} \right]. \tag{2.63}$$

Clearly the second term vanishes (to $O(\lambda^{N+1})$) if h^{ab} solves the equation of motion (2.55), and thus

$$\begin{aligned} 0 &= \int \mathrm{d}^4 x \frac{\delta S_{\text{dyn}}^{N+1}}{\delta \breve{g}^{ab}} \breve{\nabla}^a \epsilon^b + O(\lambda^{N+2}) \\ &= \int \mathrm{d}^4 x \sqrt{-\breve{g}} \epsilon^b \breve{\nabla}^a T_{ab}^{N+1} + O(\lambda^{N+2}). \end{aligned} \tag{2.64}$$

As this equation holds for any ϵ^a it follows that

$$\breve{\nabla}^a T_{ab}^{N+1} = 0 \tag{2.65}$$

is valid up to and including $O(\lambda^{N+1})$. Because this relation holds whenever h^{ab} solves its equation of motion, and because gauge transformations map solutions to solutions, the conservation law is gauge invariant.

It is important to recognise that (2.65) applies to the $(N+1)^{\text{th}}$-order energy-momentum tensor: this is the highest-order approximation to the energy-momentum tensor that can be constructed from our truncated action S_{dyn}^{N+1}, and is a better approximation than the tensor T_{ab}^N which features in the equations of motion appropriate to this order. Of course, the conservation law for T_{ab}^N follows from (2.65) by discarding the highest-order term, and ensures the consistency of the equation of motion (2.55) with the identity $\breve{\nabla}^a \widehat{G}_{abcd} h^{cd} = 0$, which holds for all h^{ab} once the background equation has been enforced.

2.3.5 Constructing the Graviton Action

It is now time to close the circle of our discussion and connect the abstract formalism to our earlier calculation. We shall derive here the graviton action S_2, the ansatz of section 2.2, by applying the perturbative formalism to the particular case

$$S[g] = \frac{1}{\kappa} \int \mathrm{d}^4 x \sqrt{-g} R \equiv S_{\text{EH}}[g], \tag{2.66}$$

[18] Note that diffeomorphism invariance is equivalent to being independent of coordinate system, and is a distinct property from gauge invariance as defined in §2.3.3.

the Einstein-Hilbert action. To proceed, we will use equation (2.38) to derive S_1, and then S_2, by successive functional derivatives $\delta/\delta\breve{g}^{ab}$ acting on $S_{\mathrm{EH}}[\breve{g}]$. The first derivative generates

$$S_1[\breve{g}, h] = \frac{1}{\kappa}\int \mathrm{d}^4x\sqrt{-\breve{g}}\breve{G}_{ab}h^{ab}, \tag{2.67}$$

which of course vanishes for all h^{ab} when $\breve{g}^{ab}$ solves the background equation $\breve{G}_{ab} = 0$. A second variation in $\breve{g}^{ab}$ gives rise to

$$\delta S_1 = \frac{1}{\kappa}\int \mathrm{d}^4x\sqrt{-\breve{g}}\big[\delta\breve{R}_{ab}\big(h^{ab} - \tfrac{1}{2}h\breve{g}^{ab}\big) + \delta\breve{g}^{cd}\tfrac{1}{2}\big(h_{cd}\breve{R} - h\breve{R}_{cd} - \breve{g}_{cd}\breve{G}_{ab}h^{ab}\big)\big]. \tag{2.68}$$

Replacing $\delta\breve{R}_{ab} \to \delta\breve{g}^{cd}\widehat{R}_{cdab}$ in accordance with (2.26), we determine $\delta S_1/\delta\breve{g}^{ab}$ and assemble

$$\begin{aligned} S_2 &= \frac{1}{2}\int \mathrm{d}^4x h^{cd}\frac{\delta S_1}{\delta\breve{g}^{cd}} \\ &= \frac{1}{2\kappa}\int \mathrm{d}^4x\sqrt{-\breve{g}}\big[h^{cd}\widehat{R}_{cdab}\big(h^{ab} - \tfrac{1}{2}h\breve{g}^{ab}\big) + \tfrac{1}{2}h^{cd}\big(h_{cd}\breve{R} - h\breve{R}_{cd} - \breve{g}_{cd}\breve{G}_{ab}h^{ab}\big)\big] \\ &= \frac{1}{2\kappa}\int \mathrm{d}^4x\sqrt{-\breve{g}}h^{ab}(\widehat{G}_{abcd} + \breve{H}_{abcd})h^{cd}. \end{aligned} \tag{2.69}$$

In the last line we referred to the definitions (2.5) and (2.6), and made use of the identity

$$\widehat{R}_{abef}(\delta^e_c\delta^f_d - \tfrac{1}{2}\breve{g}^{ef}\breve{g}_{cd}) \equiv \widehat{G}_{abcd}. \tag{2.70}$$

This completes the derivation of the graviton action (2.4) and confirms that it can be used as the starting point of an energy-momentum self-coupling procedure (2.48) that generates the Einstein field equations and the Einstein-Hilbert action (modulo surface terms) to arbitrary order.

The preceding calculation helps to reveal the advantage of using h^{ab}, a perturbation in the *inverse* metric, as our fundamental degree of freedom. Had we instead taken the usual approach, expanding $g_{ab} = \breve{g}_{ab} + \lambda\mathfrak{h}_{ab}$ and taking $\mathfrak{h}_{ab}$ as fundamental, the perturbative formalism would have unfolded identically but for the placement of indices. However, the calculation of S_2 from S_{EH} would have differed dramatically. The Lagrangian of S_1 would instead be proportional to $\breve{G}^{ab}\mathfrak{h}_{ab}$, and because the Ricci tensor is naturally covariant, the variation of $\breve{G}^{ab} = \breve{R}_{cd}\breve{g}^{ca}\breve{g}^{db} - \frac{1}{2}\breve{R}_{cd}\breve{g}^{cd}\breve{g}^{ab}$ under $\delta\breve{g}^{ab}$ would have been complicated by the extra two factors of $\breve{g}^{ab}$ on the first term, compared to the relevant tensor in our approach: $\breve{G}_{ab} = \breve{R}_{ab} - \frac{1}{2}\breve{R}_{cd}\breve{g}^{cd}\breve{g}_{ab}$. This trend continues at every order; the $\mathfrak{h}_{ab}$ convention leads to a greater proliferation of terms in each partial energy-momentum tensor because the Lagrangian of S_n has the form $(\breve{\nabla}_a)^2(\mathfrak{h}_{ab})^n$ so must be contracted with a further $n+1$ factors of $\breve{g}^{ab}$ to render it a scalar.[19] Each of these metric factors generates a term in the partial energy-momentum tensor, and thus act as compound interest for the process of energy-momentum self-coupling. In comparison, our convention leads to

[19]There are of course the instances of $\breve{g}^{ab}\partial_c\breve{g}_{de}$ in each $\breve{\nabla}_a$, but these occur equally in either convention.

Lagrangians of the form $(\check{\nabla}_a)^2(h^{ab})^n$, which need only $n-1$ additional factors of $\check{g}_{ab}$.[20] Clearly the inefficiency of the $\mathfrak{h}_{ab}$ approach stems from the natural covariance of derivative operators (∂_a or $\check{\nabla}_a$) and curvature tensors; the advantages of the contravariant expansion $g^{ab} = \check{g}^{ab} + h^{ab}$ are therefore not peculiar to the Einstein Hilbert action, and are expected to be even more distinguished in higher derivative theories of gravity.

2.4 Matter

To avoid over-complicating our discussion, we have so far focused exclusively on *pure gravity*. Here we will go some way to remedy this simplification, and generalise the formalism of the previous section to include the perturbations of matter fields, and the effects of non-vacuum backgrounds.

In the most general case, let the action S be a functional of g^{ab} and a generic matter field Ψ^A, where A will serve as a placeholder for any number of internal or spacetime indices. We then expand S about a background $(\check{g}^{ab}, \check{\Psi}^A)$ as follows:

$$g^{ab} = \check{g}^{ab} + \lambda h^{ab}, \tag{2.71}$$

$$\Psi^A = \check{\Psi}^A + \lambda \psi^A, \tag{2.72}$$

$$\Rightarrow \quad S[g, \Psi] = \sum_{n=0}^{\infty} \lambda^n S_n[\check{g}, h, \check{\Psi}, \psi], \tag{2.73}$$

where $\check{g}^{ab}$ and $\check{\Psi}^A$ satisfy the background equations

$$\frac{\delta S[\check{g}, \check{\Psi}]}{\delta \check{g}^{ab}} = 0, \qquad \frac{\delta S[\check{g}, \check{\Psi}]}{\delta \check{\Psi}^A} = 0. \tag{2.74}$$

As before, each partial action can be calculated from the partial action at the previous order; with matter included, the appropriate recurrence relation is

$$S_n = \frac{-1}{n} \int \mathrm{d}^4x \sqrt{-\check{g}} \left(h^{ab} t_{ab}^{n-1} + \psi^A j_A^{n-1} \right), \tag{2.75}$$

where

$$t_{ab}^n \equiv \frac{-1}{\sqrt{-\check{g}}} \frac{\delta S_n}{\delta \check{g}^{ab}}, \qquad j_A^n \equiv \frac{-1}{\sqrt{-\check{g}}} \frac{\delta S_n}{\delta \check{\Psi}^A}. \tag{2.76}$$

There are two aspects of this coupling scheme that differ from pure gravity. The first is immediately apparent: the $h^{ab} t_{ab}$ term has been joined by an analogous coupling between matter fluctuations ψ^A and its "source current" j_A. The second difference is hidden within the definitions of t_{ab} and j_A; because the $\{S_n\}$ now represent the partial actions for gravity and matter together, $h^{ab} t_{ab}$ and $\psi^A j_A$ are no longer just self-couplings, and will in general contain terms coupling h^{ab} to ψ^A. In particular, t_{ab}^n should now be interpreted as the

[20] This does not mean that *all* terms in such a Lagrangian will contain only $n-1$ additional factors of $\check{g}_{ab}$; there will often be cases in which $\check{g}^{ab}$ is contracted with $(\check{\nabla}_a)^2$ and thus $n+1$ factors of the metric (and its inverse) will be present. These cases only represent a small proportion of all possible terms, particularly as n becomes large, and are no worse than the terms afforded by the $\mathfrak{h}_{ab}$ convention.

(n^{th}-order) energy-momentum tensor due to *all* the fields: h^{ab}, ψ^A, and the background matter $\check{\Psi}^A$.

Proceeding as before, we can now demand that the dynamical fields h^{ab} and ψ^A solve the field equations of the action $S^{N+1}_{\text{dyn}} = \sum_{n=2}^{N+1} \lambda^n S_n$, and generate approximate solutions of the exact field equations (prescribed by S) accurate to $O(\lambda^N)$. Instead of using the definition (2.44) for $\widehat{G}_{abcd}$, we write the general form of S_2, modulo surface terms, as

$$S_2 = \frac{1}{2}\int \mathrm{d}^4x\sqrt{-\check{g}}\Big(h^{ab}\widehat{G}_{abcd}h^{cd}/\kappa - 2h^{ab}\widehat{I}_{abA}\psi^A + \psi^A\widehat{W}_{AB}\psi^B\Big), \tag{2.77}$$

once the background equations (2.74) have been enforced. In the above equation, $\widehat{G}_{abcd}$, $\widehat{I}_{abA}$, and $\widehat{W}_{AB}$ are linear operators that depend only on background fields, $\widehat{G}_{abcd}$ and $\widehat{W}_{AB}$ are self-conjugate, in the sense given by (2.11), and $\widehat{I}_{abA}$ is conjugate to $\widehat{I}^{\dagger}_{Aab}$:

$$\int \mathrm{d}^4x\sqrt{-\check{g}}A^{ab}\widehat{I}_{abA}B^A = \int \mathrm{d}^4x\sqrt{-\check{g}}B^A\widehat{I}^{\dagger}_{Aab}A^{ab}, \tag{2.78}$$

for all A^{ab} or B^{ab}, provided one has compact support. These definitions lead to equations of motion, accurate to $O(\lambda^N)$, as follows:

$$\lambda\widehat{G}_{abcd}h^{cd} = \kappa T^N_{ab} + \lambda\kappa\widehat{I}_{abA}\psi^A, \tag{2.79}$$

$$\lambda\widehat{W}_{AB}\psi^B = J^N_A + \lambda\widehat{I}^{\dagger}_{Aab}h^{ab}, \tag{2.80}$$

where

$$T^N_{ab} \equiv \sum_{n=2}^{N}\lambda^n t^n_{ab}, \qquad\qquad J^N_A \equiv \sum_{n=2}^{N}\lambda^n j^n_A. \tag{2.81}$$

Although this formalism is quite general, it is probably too general to be usefully employed. Indeed, the complications involved in describing matter as a background field *and* a dynamical perturbation generally serve to obscure the physical interpretation of the mathematics. An interesting example of this occurs when one tries to rederive $\check{\nabla}^a T^{N+1}_{ab} = 0$ by applying the argument of section 2.3.4. The result that now follows is

$$\check{\nabla}^a T^{N+1}_{ab} = \frac{1}{2\sqrt{-\check{g}}}\frac{\delta}{\delta\epsilon^b}\int \mathrm{d}^4x\sqrt{-\check{g}}J^{N+1}_A\mathcal{L}_\epsilon\check{\Psi}^A, \tag{2.82}$$

the physical interpretation of which is far from clear. Rather than continue with this formulation in its full generality, it will therefore be more instructive to examine two special cases. First, we set $\check{\Psi}^A = 0$ and consider small matter fields $\lambda\psi^A$ interacting with λh^{ab}. Second, by setting $\psi^A = 0$ we can study the effect of a background matter field $\check{\Psi}^A$ on the propagation of the graviton. In principal, one could reach these special cases starting from the formalism we have just described, but it will be simpler and more illuminating to build them up from scratch.

2.4.1 Matter Perturbations

In a region where the matter fields are small enough that their effects on spacetime curvature can be described by small perturbations λh^{ab} in the inverse metric, we can model the dynamics by taking $\check{\Psi}^A = 0$, and describe the matter field using $\lambda\psi^A$ alone. As it is often the case for gravitational theories, let us suppose that the action S is the sum of a gravitational action S_g and a matter action S_Ψ:

$$S[g, \Psi] = S_\mathrm{g}[g] + S_\Psi[g, \Psi]. \tag{2.83}$$

Moreover, for the sake of simplicity, we take Ψ^A to be a *free* field:

$$S_\Psi[g, \lambda\Psi] = \lambda^2 S_\Psi[g, \Psi] \quad \forall\ g^{ab}, \Psi^A. \tag{2.84}$$

This assumption will mean that the perturbative expansion of S can be described by an energy-momentum coupling procedure only. To see this explicitly, we expand the action about a background $(\check{g}^{ab}, 0)$:

$$S[\check{g} + \lambda h, \lambda\psi] = \sum_{n=0}^{\infty} \lambda^n \left(S_{\mathrm{g}n}[\check{g}, h] + S_{\Psi n}[\check{g}, h, \psi]\right), \tag{2.85}$$

where each gravitational partial action

$$\begin{aligned} S_{\mathrm{g}n}[\check{g}, h] &= \frac{1}{n!}\left(\partial_\lambda^n S_\mathrm{g}[\check{g} + \lambda h]\right)_{\lambda=0} \\ &= \frac{1}{n!}\left[\int \mathrm{d}^4x h^{ab} \frac{\delta}{\delta \check{g}^{ab}}\right]^n S_\mathrm{g}[\check{g}], \end{aligned} \tag{2.86}$$

much as before, and the matter partial actions

$$\begin{aligned} S_{\Psi n}[\check{g}, h, \psi] &= \frac{1}{n!}\left(\partial_\lambda^n S_\Psi[\check{g} + \lambda h, \lambda\psi]\right)_{\lambda=0} \\ &= \frac{1}{n!}\left(\partial_\lambda^n \left(\lambda^2 S_\Psi[\check{g} + \lambda h, \psi]\right)\right)_{\lambda=0} \\ &= \frac{1}{(n-2)!}\left(\partial_\lambda^{n-2} S_\Psi[\check{g} + \lambda h, \psi]\right)_{\lambda=0} \\ &= \frac{1}{(n-2)!}\left[\int \mathrm{d}^4x h^{ab} \frac{\delta}{\delta \check{g}^{ab}}\right]^{n-2} S_\Psi[\check{g}, \psi]. \end{aligned} \tag{2.87}$$

Defining the partial energy momentum tensors for h^{ab} and ψ^A as

$$t_{ab}^{\mathrm{g}n} \equiv \frac{-1}{\sqrt{-\check{g}}} \frac{\delta S_{\mathrm{g}n}}{\delta \check{g}^{ab}}, \qquad t_{ab}^{\Psi n} \equiv \frac{-1}{\sqrt{-\check{g}}} \frac{\delta S_{\Psi n}}{\delta \check{g}^{ab}}, \tag{2.88}$$

respectively, we see that the partial actions are coupled as

$$S_n[\check{g}, h] = -\int \mathrm{d}^4x \sqrt{-\check{g}} h^{ab} \left(\frac{t_{ab}^{\mathrm{g}n-1}}{n} + \frac{t_{ab}^{\Psi n-1}}{n-2}\right). \tag{2.89}$$

These partial actions lead to the N^{th}-order equations of motion

$$\lambda \widehat{G}_{abcd} h^{cd} = \kappa T_{ab}^N = \sum_{n=2}^{N} \lambda^n \left(t_{ab}^{\mathrm{g}n} + t_{ab}^{\Psi n}\right), \tag{2.90}$$

$$\lambda \widehat{W}_{AB} \psi^B = \sum_{n=2}^{N} \left[\frac{-\lambda^n}{(n-1)\sqrt{-\check{g}}} \frac{\delta}{\delta \psi^A} \int \mathrm{d}^4x \sqrt{-\check{g}} h^{ab} t_{ab}^{\Psi n}\right]. \tag{2.91}$$

The first equation confirms that the energy-momentum tensors of ψ^A and h^{ab} combine as the source for the graviton. The second equation describes how the coupling between h^{ab} and t^{Ψ}_{ab} acts as a source for ψ^A. Note that even when the matter field is not free, because S_Ψ never contains terms linear in the matter fields, $\widehat{I}_{abA}$ must be at least linear in $\check{\Psi}^A$, so we will always have $\widehat{I}_{abA} = 0$ when $\check{\Psi}^A = 0$.

2.4.2 Non-Vacuum Background

For a non-vacuum spacetime, we expect to be able to approximate (at least to first-order) the behaviour of a gravitational perturbation by ignoring the perturbations in the matter field that it might induce. Alternatively, we may have in mind a particular non-vacuum solution of the field equations $(\check{g}^{ab}, \check{\Psi}^A)$ and wish to find nearby solutions (approximate or exact) with precisely the same matter content. For these two scenarios, we can set $\psi^A = 0$ and investigate the effect that the background $\check{\Psi}^A$ has on the dynamics of h^{ab}.

Considerations of this nature highlight an interesting feature of our prior discussion of the graviton action. In section 2.2 we saw the importance of a contribution to the action $h^{ab}H_{abcd}h^{ab}$ that vanished in the vacuum; the obvious question to ask is whether a similar term exists in the non-vacuum case, and whether or not it will vanish on the *non-vacuum* background equations. To answer these questions we will derive the graviton action for a non-vacuum background, which will also include the cosmological constant as a special case.

Let us restrict our attention to general relativity in the presence of a matter field:

$$S[g,\Psi] = S_{\mathrm{EH}}[g] + S_\Psi[g,\Psi], \tag{2.92}$$

$$S_\Psi[g,\Psi] \equiv 2\int \mathrm{d}^4x\sqrt{-g}\mathcal{L}_\Psi(g^{ab},\Psi^A,\partial_a\Psi^A). \tag{2.93}$$

The factor of two in the definition of the matter Lagrangian $\mathcal{L}_\Psi$ compensates for our slightly unusual normalisation of S_{EH}.[21] It should be noted that we have assumed that $\mathcal{L}_\Psi$ does not depend on derivatives of the metric. This is the case for the Lagrangians of all the fields of the standard model except the spin-$\frac{1}{2}$ fermion, which in any case should be coupled to gravity using the vierbein formalism, e.g. [53]; such an approach is beyond the scope of this chapter. The results of this section can be generalised to allow $\mathcal{L}_m$ to depend on $\partial_c g^{ab}$ without any great difficulty, but this is an added algebraic complication that seems to add little insight to our investigation.

We proceed by expanding the action about a background $(\check{g}^{ab}, \check{\Psi}^A)$ just as in (2.71) and (2.72), but now, as $\psi^A = 0$, the coupling scheme (2.75) reverts to the familiar energy-momentum coupling of section 2.3. Following precisely the same method as section 2.3.5, we can compute S_2 by two successive functional derivatives (with respect to $\check{g}^{ab}$) applied

[21] All our actions are twice as large as the usual definition. This normalisation has no effect on the classical equations of motion, but has allowed us to define the energy-momentum tensor without a factor of two, simplifying the algebra of §§2.2&2.3.

to $S[\breve{g}, \breve{\Psi}]$. The first derivative yields

$$S_1 = \frac{1}{\kappa} \int \mathrm{d}^4x \sqrt{-\breve{g}} \left(\breve{G}_{ab} - \kappa \breve{T}^{\Psi}_{ab}\right) h^{ab}, \tag{2.94}$$

where

$$\breve{T}^{\Psi}_{ab} = \frac{-1}{\sqrt{-\breve{g}}} \frac{\delta S_{\Psi}[\breve{g}, \breve{\Psi}]}{\delta \breve{g}^{ab}} = -2 \frac{\partial \breve{\mathcal{L}}_{\Psi}}{\partial \breve{g}^{ab}} + \breve{g}_{ab} \breve{\mathcal{L}}_{\Psi} \tag{2.95}$$

is the energy-momentum tensor of the background matter. The second derivative yields the graviton action:

$$\begin{aligned} S_2 &= \frac{1}{2} \int \mathrm{d}^4x h^{ab} \frac{\delta S_1}{\delta \breve{g}^{ab}} \\ &= \frac{1}{2\kappa} \int \mathrm{d}^4x \sqrt{-\breve{g}} \Big[h^{ab} \widehat{G}_{abcd} h^{cd} - \left(\breve{G}_{ab} - \kappa \breve{T}^{\Psi}_{ab}\right) h^{ab} h + 2\kappa h^{ab} h^{cd} \frac{\partial^2 \breve{\mathcal{L}}_{\Psi}}{\partial \breve{g}^{ab} \partial \breve{g}^{cd}} \\ &\qquad + \left(\breve{R} + 2\kappa \breve{\mathcal{L}}_{\Psi}\right) \left(\tfrac{1}{2} h_{ab} h^{ab} - \tfrac{1}{4} h^2\right) \Big]. \end{aligned} \tag{2.96}$$

This is the action we sought: the generalisation of equation (2.4) to a non-vacuum background.

If we are only interested in the linear theory, and have no wish to calculate the energy-momentum tensor, then we are free to enforce the background equation

$$\breve{G}_{ab} = \kappa \breve{T}^{\Psi}_{ab}, \tag{2.97}$$

in the graviton action. In sharp contrast to the vacuum case, however, the background equation does not reduce S_2 to $\frac{1}{2\kappa} \int \mathrm{d}^4x \sqrt{-\breve{g}} h^{ab} \widehat{G}_{abcd} h^{cd}$, or indeed any other covariantisation of the massless spin-2 Fierz-Pauli action. Instead, it appears as though the background matter has endowed the graviton with mass:

$$S_2 = \frac{1}{2\kappa} \int \mathrm{d}^4x \sqrt{-\breve{g}} \left(h^{ab} \widehat{G}_{abcd} h^{cd} + \alpha \right), \tag{2.98}$$

where the "mass-term" α is given by

$$\alpha \equiv -\tfrac{1}{2} M \left(h^{ab} h_{ab} - \tfrac{1}{2} h^2 \right) + N_{abcd} h^{ab} h^{cd}, \tag{2.99}$$

with

$$M \equiv 2\kappa \left(\breve{\mathcal{L}}_{\Psi} - \breve{g}^{ab} \frac{\partial \breve{\mathcal{L}}_{\Psi}}{\partial \breve{g}^{ab}} \right), \qquad N_{abcd} \equiv 2\kappa \frac{\partial^2 \breve{\mathcal{L}}_{\Psi}}{\partial \breve{g}^{ab} \partial \breve{g}^{cd}}. \tag{2.100}$$

We refer to α as a "mass-term" because it is quadratic in h^{ab}, free from derivatives, and has been added to the kinetic term $h^{ab} \widehat{G}_{abcd} h^{cd}$ in the Lagrangian. However, as we will see for the specific case of the cosmological constant, α does not by itself determine whether the graviton is actually *massive*, i.e. whether it propagates *subluminally*; the curvature of the background will play an equally important role in the field equations. In particular, while it is tempting to identify a mass m for the graviton according to $m^2 = M$ (at least when $N_{abcd} = 0$) we will soon see that the background matter often sets $M < 0$, so this idea is essentially untenable.

To explore these issues, it will be instructive to calculate α for a few simple examples. First, consider a scalar field background $\check{\Phi}$ with Lagrangian

$$\check{\mathcal{L}}_\Phi = -\tfrac{1}{2}\breve{g}^{ab}\partial_a\check{\Phi}\partial_b\check{\Phi} - V(\check{\Phi}); \tag{2.101}$$

the mass-term is

$$\alpha_\Phi = \kappa V(\check{\Phi})\left(h_{ab}h^{ab} - \tfrac{1}{2}h^2\right). \tag{2.102}$$

To ensure that the scalar field has positive energy density, we must insist that $V(\check{\Phi}) \geq 0$; hence $M \leq 0$ as previously warned. Equation (2.102) can also be used to find the corresponding mass-term for a cosmological constant. In this case the Lagrangian is $\mathcal{L}_\Lambda = -\Lambda/\kappa$, which we can reach from $\mathcal{L}_\Phi$ by setting $\partial_a\check{\Phi} = 0$ and $V = \Lambda/\kappa$. Clearly this gives

$$\alpha_\Lambda = \Lambda\left(h_{ab}h^{ab} - \tfrac{1}{2}h^2\right), \tag{2.103}$$

which similarly suffers from $M < 0$ if the cosmological constant is positive.

At this point, the reader may be suspicious that the formulae for α_Φ and α_Λ (with $M < 0$ and $N_{abcd} = 0$) signify that h^{ab} is a *tachyon* in the presence of a scalar field background or a cosmological constant. Indeed, if the background were flat and M constant over spacetime, we could derive the field equations from (2.98), observe that their divergence enforces the de Donder gauge condition

$$\partial^\alpha h_{\alpha\beta} - \tfrac{1}{2}\partial_\beta h = 0, \tag{2.104}$$

and, substituting this back into the equations of motion, conclude that the dynamics of the graviton were described by

$$\left(\partial^2 - M\right)h^{\alpha\beta} = 0. \tag{2.105}$$

This argument appears to justify the relation $m^2 = M$ for the graviton's mass, and motivate the conclusion that $M < 0$ betrays tachyonic behaviour. It is important to realise, however, that the field equation above is of little relevance to the actual physical system we were discussing. In reality, M will not be constant, and the presence of background matter will inevitably preclude background flatness. To understand how this last consideration alters the dynamics of the graviton, we shall briefly examine the field equation for h^{ab} in the presence of a cosmological constant. First, we substitute (2.103) into (2.98) and derive the field equation

$$\widehat{G}_{abcd}h^{cd} + \Lambda\left(h_{ab} - \tfrac{1}{2}\breve{g}_{ab}h\right) = 0. \tag{2.106}$$

In contrast to the naive approach, the covariant divergence of this equation vanishes identically, and so cannot be used to relate $\check{\nabla}_b h$ and $\check{\nabla}^a h_{ab}$. In place of this, the gauge invariance of the vacuum theory remains intact[22], and the field equation may be simplified by setting $h = 0$, $\check{\nabla}_a h^{ab} = 0$:

$$\check{\nabla}^2 h_{ab} - 2\check{R}_{dabc} h^{dc} = 0. \tag{2.107}$$

Surprisingly, the contribution from α_Λ has been cancelled by a term proportional to the background Ricci tensor, resulting in a field equation that is identical in form to the first-order *vacuum* field equation (2.13) in this gauge. Of course, this does not indicate that the cosmological constant has no effect on the propagation of h^{ab}, only that these effects are limited to the constraints imposed on the background geometry by the background equation $\check{R}_{ab} = \Lambda \check{g}_{ab}$. For this reason, it does not seem particularly natural to interpret $2\check{R}_{dabc} h^{dc}$ as endowing the graviton with a mass; equation (2.107) can instead be understood as a (partially gauge-fixed) massless spin-2 field equation that has been generalised to cosmological backgrounds. Quite aside from this, there is also the technical issue of interpreting the four-index tensor $\check{R}_{abcd}$ as a mass: only if this tensor can be defined in terms of a single scalar variable (and the background metric) could the argument be made that this single variable described the graviton's mass. For a non-zero cosmological constant, the only background with this property is de Sitter space: $\check{R}_{dabc} = \frac{\Lambda}{3}(\check{g}_{db}\check{g}_{ac} - \check{g}_{dc}\check{g}_{ab})$, thus the gauge-fixed field equation (2.107) becomes

$$\left(\check{\nabla}^2 - \frac{2\Lambda}{3}\right) h^{ab} = 0. \tag{2.108}$$

If we were so inclined, we might interpret this as a field equation for a graviton with $m^2 = 2\Lambda/3$, and note that this relation has the *correct* sign for positive Λ, unlike the formula $m^2 = -2\Lambda$ suggested by our preliminary inspection of α_Λ. In truth, however, further investigation is needed before we can either adopt or discard this interpretation. This is not only because (2.107) (of which (2.108) is a special case) can be understood as a generalisation of a massless field equation to cosmological backgrounds, but also because of the subtleties involved in interpreting the wave operator $\check{\nabla}^2$ in curved space, and issues of whether or not to use a conformal coupling. Clearly, more work must be done to ascertain the physical ramifications of α_Λ, and the "mass-term" α in general, before we can understand the degree to which its effects can be thought of as giving mass to the graviton.

Although massive gravitons and the cosmological constant were historically viewed as entirely separate concepts, recent work has brought to light a number of interesting connections between the two. Deser and Waldron [28] have demonstrated that, in (anti-)de Sitter background spacetimes, a massive spin-2 field is stable if and only if $m^2 \geq 2\Lambda/3$, or $m = 0$. While it is intriguing that our de Sitter background field equation (2.108)

[22] If we wish to extend our discussion of gauge invariance (§2.3.3) to include background matter *in general*, we would need to account for the gauge-fixing implicit in our starting assumption $\psi^A = 0$, which is obviously not preserved by a (first-order) infinitesimal diffeomorphism $\delta\psi^A = \mathcal{L}_\epsilon \check{\Psi}^A$. However, because Λ is constant over spacetime, no such difficulty arises here, and the transformations $\delta h^{ab} = -2\check{\nabla}^{(a}\epsilon^{b)}$ remain a symmetry of the equations of motion.

suggests precisely the same special value of $m^2 = 2\Lambda/3$, Deser and Waldron's analysis differs significantly from our own, so this superficial observation may be misleading. In particular, whereas our mass-term arises as a direct result of the perturbative expansion, Deser and Waldron add their mass-term to the action *by hand*. Thus it is far from clear that the massive gravitons of their paper correspond to the physical system considered above. In contrast, Novello and Neves [62] claim to prove that $m^2 = -2\Lambda/3$, with the implication that $\Lambda \leq 0$. This approach considers an unusual generalisation of the spin-2 field equation to curved backgrounds, making a non-standard choice for the covariantisation ambiguous term discussed in section 2.2.1. Thus, while their calculations arguably describe a spin-2 field, this does not appear to be a natural way to describe the spin-2 field that results from perturbations of the metric (or its inverse) in Einstein's theory. It is our intention to disentangle the connections between these two approaches, and our own, in a later publication.

For the sake of completeness, we conclude this section with an example of a mass-term that can have $M > 0$, and $N_{abcd} \neq 0$. Unlike α_Λ, however, we shall not attempt to derive any of the implications for the equations of motion. Consider an electromagnetic 1-form background $\check{A}_a$, with Lagrangian

$$\check{\mathcal{L}}_A = -\tfrac{1}{4}\check{F}^2 = -\tfrac{1}{4}\check{g}^{ab}\check{g}^{cd}\check{F}_{ac}\check{F}_{bd}, \tag{2.109}$$

and note that $\check{F}_{ab} \equiv 2\partial_{[a}\check{A}_{b]}$ is independent of the metric. The calculation yields

$$\alpha_A = -\tfrac{1}{4}\kappa\check{F}^2\left(h_{ab}h^{ab} - \tfrac{1}{2}h^2\right) - \kappa h^{ab}h^{cd}\check{F}_{ac}\check{F}_{bd}, \tag{2.110}$$

which has the aforementioned properties.

2.5 Conclusion

Contrary to the prevailing maxim, coupling the classical Fierz-Pauli graviton to its own energy and momentum *does not* recreate general relativity order by order. However, there is an alternative action for the graviton (2.4) for which energy-momentum self-coupling *is* consistent with Einstein's theory. Using this action, the energy-momentum tensor of the graviton (2.28), added as a source to the graviton's first-order equation of motion (2.13), builds a field equation consistent with the Einstein equations to *second-order*. Furthermore, the perturbative formalism developed in section 2.3 reveals that our action provides sufficient information to reconstruct general relativity to *arbitrary* accuracy: a simple recurrence relation (2.48) identifies the energy-momentum tensor at one order as the appropriate contribution to the action at the next. To any order N, this scheme assembles an action that dictates field equations (2.55) in which the graviton's N^{th}-order energy-momentum tensor is the source.

The formal machinery used to understand vacuum perturbations is easily extended to include matter, although the physical interpretation of the most general approach, in which matter comprises both a background field and a small perturbation, is less than transparent. Focusing on matter perturbations separately from non-vacuum backgrounds serves to clarify the formalism significantly. In a vacuum background, the interactions between the graviton and perturbations of a free matter field lead to a field equation (2.90) in which the source for the graviton is the sum of gravitational and matter energy-momentum. This interaction inevitably induces a source in the field equations for matter (2.91). Alternatively, one may neglect matter perturbations and examine the consequences of a non-vacuum background. In this case, the dynamics and energy-momentum of the graviton are prescribed by the action (2.96), generalising our previous ansatz. Surprisingly, the background matter appears to induce a "mass-term" in the graviton action, although it is currently unclear to what extent its interpretation as a mass is valid at the level of the field equations. The mass-terms induced by a scalar field (2.102), a cosmological constant (2.103) and electromagnetism (2.110) have been calculated.

2.A Appendix: Padmanabhan's Analysis

The recent article by Padmanabhan [64] unearths many significant shortcomings of the well known arguments [25, 36, 41, 50] that supposedly derive Einstein's equations by coupling the Fierz-Pauli graviton to its own energy-momentum tensor. Here we attempt to summarise his observations, and explain their relation to this present work.

In broad terms, Padmanabhan's criticisms fall into three areas:

1. The Einstein-Hilbert action consists of a bulk term (the Γ^2 action) and a surface term. The latter includes a piece *linear* in $h_{\alpha\beta}$, so there can be no way to construct it from a self-coupling procedure that starts with an action that is already *quadratic* in $h_{\alpha\beta}$.[23]

2. The starting point, the Fierz-Pauli Lagrangian (2.8), describes a *Lorentz invariant* field theory, and yet the end result, general relativity, is *generally covariant*. It is claimed that this metamorphosis only occurs because general covariance has been *assumed* in the various derivations, in which case it is "no big deal to obtain Einstein's theory". More generally, the classic bootstrapping arguments wield ideas developed in general relativity (such as Hilbert's definition of the energy-momentum tensor) or use knowledge of the end result to achieve their goal. Hence they cannot be regarded as a derivation of general relativity *from first principles*.

3. The first-order field equation can only take a *symmetric* tensor as its source; the canonical energy-momentum tensor (2.15) is not necessarily symmetric, and although it can be made to be so, this process is not unique. Therefore the energy-momentum self-coupling procedure is ill-defined. The Hilbert definition *is* uniquely determined by the action, but to use it would violate criticism 2. *Crucially*, even if we allow

[23] The argument given by Padmanabhan is phrased in terms of non-analyticity in a dimensionful coupling constant. This form of the argument depends on his particular choice of normalisation for $h_{\alpha\beta}$ and S_{EH}, but is essentially equivalent to the statement given here.

ourselves to use Hilbert's definition, we still fail to recover the correct source term for the second-order field equation.

It is to this very last crucial point that we have devoted the bulk of this chapter. We now wish to explain our position with regards to the first two criticisms, and also Padmanabhan's proposed solution to the third.

1. Our approach expressly avoids discussing surface terms. This has greatly streamlined our formalism, and because such terms are completely irrelevant for determining field equations or energy-momentum tensors, the only price to pay for this simplicity is that we can only claim to reconstruct the Einstein-Hilbert action *modulo surface terms*.[24] In this sense, Padmanabhan's first criticism still stands, although it is unclear whether it has any great importance. If the action is an integral over the whole manifold, and asymptotic conditions apply to h^{ab} such that the surface term at infinity vanishes, then of course there is no distinction between the Einstein-Hilbert action and the action we have constructed. Even if the action is an integral over a manifold with a boundary, so long as we consider the action to be a functional over all fields with a particular boundary configuration (just as we might think of the action of a particle as a functional over all paths with particular end-points) the two actions differ only by an irrelevant constant. Besides, in situations where contributions from the boundary really are important, one does not typically use the Einstein-Hilbert action anyway: the Gibbons-Hawking-York boundary term [39, 83] must be included to remove the dependence on second derivatives of the metric. This allows the field equations to be derived using a variational principle that only demands that the variation in the fields (and not also their derivatives) vanish on the boundary.

Padmanabhan's major concern is that the surface term of the Einstein-Hilbert action has some quantum mechanical significance. As the nature of quantum gravity has yet to be understood, it remains unclear whether or not this is the case. We stress once again that the analysis in this chapter is purely classical, and that we make no claims as to a quantum mechanical interpretation. Furthermore, it is not even known whether the graviton is a useful theoretical object for describing quantum gravity. We note again that the Gibbons-Hawking-York boundary term is usually included in quantum gravity investigations for which the boundary is not negligible.

2. It is our view that Padmanabhan's concerns about general covariance are unjustified: we take the position of Weinberg [80], that "general covariance by itself is empty of physical content." Any theory (Lorentz invariant or not) can be expressed in arbitrary curvilinear coordinates, so the requirement of general covariance cannot, in and of itself, constrain the sort of theory one might construct. Rather, the kinematical content of general relativity is encapsulated by *the equivalence principle*, that the effect of gravity vanishes locally in an inertial coordinate system; thus expressing physical equations in coordinate invariant notation is an invaluable tool for describing how their dynamics are modified by gravity. It is possible that when Padmanabhan refers to 'general covariance' he is referring to the

[24] Note that this does not nessesarily mean that we have constructed the Γ^2 action, only that the integrand of the action differs from $\sqrt{-g}R$ by some total divergence.

equivalence principle also. As the latter is tantamount to identifying the gravitational field with a dynamical metric, he would certainly be correct to criticise any "derivation" that contained such a step; needless to say, we do not appeal to the equivalence principle in our approach.

General covariance aside, though, Padmanabhan's objection to the use of curved-space ideas is a valid one, indicating that none of the classic arguments constitute a derivation from first principles. Our approach certainly makes use of curved-space concepts; however our goals are perhaps not quite so bold as the other derivations that Padmanabhan has scrutinised: we do not pretend to derive general relativity purely from the ideas of Lorentz-invariant field theory. It should be stressed, however, that even if some of the *kinematical* content of general relativity is in some way assumed (curved spacetime, functional derivatives with respect to the metric, etc.) it is still a "big deal" to derive the *dynamical* content of the theory, Einstein's equations.

3. We have already explained our position with regards to the definition of the energy-momentum tensor in section 2.2.3; the only reason that Hilbert's definition is unpalatable to Padmanabhan is that his aim is to start with as little curved-space mathematics as he can. However, the failure of the Hilbert energy-momentum tensor to give the correct second-order term for the Einstein field equations is a more significant stumbling-block. We have explained our remedy, the use of a different starting action, in the body of this chapter. Padmanabhan, on the other hand, eschews energy-momentum self-coupling and introduces a new object $S^{\alpha\beta}$ that he defines with the following algorithm. Start with a Lorentz invariant Lagrangian $\mathcal{L}(\eta_{\alpha\beta}, h_{\alpha\beta}, \partial_\gamma h_{\alpha\beta})$ expressed in Lorentzian coordinates $\{x^\alpha\}$. Replace every instance of $\eta_{\alpha\beta}$ with the metric $\breve{g}_{\alpha\beta}$ to produce a new Lagrangian $\widetilde{\mathcal{L}}(\breve{g}_{\alpha\beta}, h_{\alpha\beta}, \partial_\gamma h_{\alpha\beta})$; note that this is *not* the same as expressing $\mathcal{L}$ in an arbitrary coordinate system because the partial derivatives ∂_α have not been upgraded to covariant derivatives $\breve{\nabla}_\alpha$. We can now define

$$S^{\alpha\beta} \equiv 2 \left. \frac{\partial \sqrt{-\breve{g}}\widetilde{\mathcal{L}}}{\partial \breve{g}_{\alpha\beta}} \right|_{\breve{g}=\eta} . \tag{2.111}$$

The subscript reminds us that we must set $\breve{g}_{\alpha\beta} = \eta_{\alpha\beta}$ after taking the metric derivative, as we are supposedly working in Lorentzian coordinates. Padmanabhan claims to be able to reconstruct the Γ^2 action by coupling $h_{\alpha\beta}$ to this new object $S^{\alpha\beta}$. Unfortunately $S^{\alpha\beta}$ has a number of highly undesirable properties, suggesting that it is a rather unnatural object, ill-defined in its current form.[25]

Firstly, as it has been constructed from a Lagrangian rather than an action, $S^{\alpha\beta}$ depends directly on surface terms. This introduces a very large ambiguity, as $S^{\alpha\beta}$ will depend on whether we write the integrand of the action in the form $(\partial h)^2$, as Padmanabhan does, in the form $h\partial^2 h$, or as some arbitrary combination of both. Each possibility defines a different $S^{\alpha\beta}$ and (presumably) leads to a different self-coupled limit for the graviton.

[25] In private communication, Padmanabhan has indicated that he shares our concerns about $S^{\alpha\beta}$ and does not believe it to be of any fundamental importance; hence we present the case against $S^{\alpha\beta}$ for the sake of completeness rather than rebuttal.

It seems that the only remedy for this ambiguity is to artificially stipulate that $\mathcal{L}$ contain no second derivatives, although we note in passing that even this leaves us free to add surface terms of the form $\partial^\alpha(\phi A_\alpha)$ in theories for fields other than the graviton.

The second troubling aspect to $S^{\alpha\beta}$ is the "half-covariantising" algorithm used to construct $\widetilde{\mathcal{L}}$. It should be clear that this procedure has only been defined in Lorentzian coordinates, thus the matrix $S^{\alpha\beta}$ does not really constitute the components of a tensor, as we have not explained how their values change when expressed in another coordinate system.[26] There are essentially two ways to extend the definition (2.111) to include curvilinear coordinates. The trivial solution is to construct the tensor $S^{ab} \equiv S^{\alpha\beta}(\partial_\alpha)^a(\partial_\beta)^b$ using the vectors $\{(\partial_\alpha)^a\}$, partial derivatives with respect to the Lorentzian coordinates used to calculate $S^{\alpha\beta}$ in the first place. This obviously defines a genuine tensor, so the components $S^{\alpha'\beta'}$ of S^{ab} in some curvilinear coordinate system $\{x^{\alpha'}\}$ can be calculated, and they will be related to $S^{\alpha\beta}$ by the usual transformation rules. It should be clear, however, that this solution is rather unnatural: suppose we have a Lagrangian expressed in a curvilinear coordinate system, then the only way to calculate the components $S^{\alpha'\beta'}$ in that system is to first transform to Lorentzian coordinates, calculate $S^{\alpha\beta}$ according to (2.111), and then transform back to our original coordinate system. Also, because this process picks out a special set of coordinates, there is also no reason to expect that S^{ab} can be written as a tensorial function of h_{ab}, $\check{g}_{ab}$ and $\check{\nabla}_a$. The *natural* way to proceed would be to generalise the definition (2.111) in such a way that we could calculate $S^{\alpha'\beta'}$ working in any coordinate system. It might seem that a viable solution would be to define the tensor

$$S^{ab} \equiv \frac{2}{\sqrt{-\check{g}}} \left. \frac{\partial \sqrt{-\check{g}}\mathcal{L}}{\partial \check{g}_{ab}} \right|_{\check{\Gamma}}, \tag{2.112}$$

where $\mathcal{L} = \mathcal{L}(\check{g}_{ab}, h_{ab}, \check{\nabla}_c h_{ab})$ is the *fully* covariant Lagrangian, and the subscript indicates that the Christoffel symbols $\check{\Gamma}^a{}_{bc}$ are to be treated as independent of the metric and held constant in the derivative. This expression generalises (2.111) to define a tensor S^{ab} in a coordinate invariant fashion; because the Christoffel symbols are held constant, no term arises from a variation of the covariant derivatives, and S^{ab} will reduce to $S^{\alpha\beta}$ in Lorentzian coordinates. This expression gives us some insight into the geometrical meaning of Padmanabhan's half-covariantised algorithm; in particular it reveals that the derivative $\partial/\partial\check{g}_{\alpha\beta}$ used to define $S^{\alpha\beta}$ is in fact exploring geometries (infinitesimally close to Minkowski spacetime) with connections that are not metric compatible.[27] It is perhaps unsurprising that this $\check{\Gamma}$-constant derivative introduces a new layer of ambiguity to the procedure, as we can now alter S^{ab} by adding terms proportional to $0 = \check{\nabla}_c\check{g}_{ab}$ to the Lagrangian. Although this might seem a rather contrived objection, it is in fact a very

[26] The insistence that we be able to calculate the components of this object in arbitrary coordinates has nothing to do with curved spacetime or general relativity. Rather, this reflects the perfectly reasonable expectation that we should be able to express Padmanabhan's self-coupling procedure in *flat-space* spherical polar coordinates, for example, or any other coordinate system we choose.

[27] This is the same operation as the derivative used to acquire the Einstein equations from the Palatini action [45, §19.10], although here we will have no cause to perform the complementary derivative $\partial/\partial\Gamma|_{\check{g}}$.

common consideration. For example, suppose the Lagrangian includes a term of the form $\check{\nabla}_a h^a{}_b$; should we calculate S^{ab} by acting with $\partial/\partial\check{g}|_{\check{\Gamma}}$ on $\check{\nabla}_a(\check{g}^{ac}h_{cb})$, or should we first commute the metric past the covariant derivative, and act on $\check{g}^{ac}\check{\nabla}_a h_{cb}$ instead? Note that this issue would have been invisible in Lorentzian coordinates because

$$\left.\frac{\partial\check{\nabla}_c\check{g}_{ef}}{\partial\check{g}_{ab}}\right|_{\check{\Gamma}} = -2\check{\Gamma}^{(a}{}_{c(e}\delta^{b)}{}_{f)}, \tag{2.113}$$

which we would have automatically set to zero. It seems the only way to avoid this uncertainty in S^{ab} is to introduce another artificial constraint on the Lagrangian: we insist that it be written in such a way that no derivatives act on the metric. This should be achieved by commuting covariant derivatives through the metric, rather than integrating by parts, due to the aforementioned issues with surface terms.

We shall take our analysis of $S^{\alpha\beta}$ no further at this time. It is still uncertain whether this object can be generalised, naturally and uniquely, to form a genuine tensor; without such a generalisation it is difficult to ascertain what sort of mathematical object the matrix of functions $S^{\alpha\beta}$ is supposed to represent. Although we cannot claim to have exhausted all possibilities, the evidence before us suggests, at the very least, that this goal is not easily achieved.

Aside from these technical issues, we should also emphasise that, unlike the energy-momentum tensor, $S^{\alpha\beta}$ has no apparent physical interpretation beyond its supposed role in a graviton self-coupling scheme. Energy-momentum self-coupling was justified by analogy with matter-gravity coupling, and advanced by the notion that the energy-momentum of *all* fields should source gravitation. In contrast, the self-coupling scheme involving $S^{\alpha\beta}$ only serves to set gravity apart from the other fields. Furthermore, our solution displays an unusual symmetry between the coupling terms in the action and source terms generated in the field equations as a result (see §2.3.2); this symmetry is broken by Padmanabhan's self-coupling procedure.

2.B Appendix: Expansion of G_{ab}

Here we determine the first two terms of the expansion of the Einstein tensor

$$G_{ab} = G^{(1)}_{ab} + G^{(2)}_{ab} + O(h^3), \tag{2.114}$$

induced by a perturbation of the inverse metric about a vacuum background:

$$g^{ab} = \check{g}^{ab} + h^{ab}, \tag{2.115}$$

$$\check{G}_{ab} = 0. \tag{2.116}$$

The perturbation in the metric is of course fixed by the relationship $g^{ab}g_{bc} = \delta^a_c$,

$$\Rightarrow \quad g_{ab} = \check{g}_{ab} - h_{ab} + h_{ac}h^c{}_b + O(h^3). \tag{2.117}$$

To begin, introduce a connection $E^a{}_{bc}$ between the derivative operators ∇_a and $\check{\nabla}_a$:

$$E^a{}_{bc} = \tfrac{1}{2} g^{ab}(\check{\nabla}_b g_{cd} + \check{\nabla}_c g_{bd} - \check{\nabla}_d g_{bc}). \tag{2.118}$$

This allow the Ricci tensor to be expressed as

$$R_{ab} = 2\left(\check{\nabla}_{[c} E^c{}_{a]b} + E^c{}_{d[c} E^d{}_{a]b}\right). \tag{2.119}$$

From (2.118) it is clear that

$$E^{a(0)}_{\ bc} = 0, \tag{2.120}$$
$$E^{a(1)}_{\ bc} = -\tfrac{1}{2}\check{g}^{ad}(2\check{\nabla}_{(b} h_{c)d} - \check{\nabla}_d h_{bc}), \tag{2.121}$$
$$E^{a(2)}_{\ bc} = -\tfrac{1}{2} h^{ad}(2\check{\nabla}_{(b} h_{c)d} - \check{\nabla}_d h_{bc}) + \tfrac{1}{2}\check{g}^{ad}(2\check{\nabla}_{(b}(h_{c)e} h^e{}_d) - \check{\nabla}_d(h_{be} h^e{}_c)). \tag{2.122}$$

Hence the terms of the expansion $R_{ab} = R^{(1)}_{ab} + R^{(2)}_{ab} + O(h^3)$ can be computed as follows:

$$R^{(1)}_{ab} = 2\check{\nabla}_{[c} E^{c(1)}_{\ a]b} \tag{2.123}$$
$$R^{(2)}_{ab} = 2\left(\check{\nabla}_{[c} E^{c(2)}_{\ a]b} + E^{c(1)}_{\ d[c} E^{d(1)}_{\ a]b}\right). \tag{2.124}$$

Thus,

$$\begin{aligned} G^{(1)}_{ab} &= R^{(1)}_{ab} - \tfrac{1}{2}\check{g}_{ab} R^{(1)}_{cd}\check{g}^{cd} \\ &= -\check{\nabla}_c \check{\nabla}_{(a} h_{b)}{}^c + \tfrac{1}{2}\check{\nabla}^2 h_{ab} + \tfrac{1}{2}\check{\nabla}_a\check{\nabla}_b h - \tfrac{1}{2}\check{g}_{ab}\left(-\check{\nabla}_c\check{\nabla}_d h^{cd} + \check{\nabla}^2 h\right), \end{aligned} \tag{2.125}$$

which confirms that $\widehat{G}_{abcd}$, as defined in (2.5), represents the linearised Einstein tensor:

$$\widehat{G}_{abcd} h^{cd} = G^{(1)}_{ab}. \tag{2.126}$$

In particular, note that both sides of this equation agree on the order of the derivatives in $\check{\nabla}_c\check{\nabla}_{(a} h_{b)}{}^c$; this is the descendant of the covariantisation ambiguous term discussed in section 2.2.1.

To find $G^{(2)}_{ab}$, start with

$$G^{(2)}_{ab} = R^{(2)}_{ab} - \tfrac{1}{2}\check{g}_{ab}\left(R^{(2)}_{cd}\check{g}^{cd} + R^{(1)}_{cd} h^{cd}\right) + \tfrac{1}{2} h_{ab} R^{(1)}_{cd}\check{g}^{cd}, \tag{2.127}$$

and substitute equations (2.123) and (2.124), followed by (2.121) and (2.122). The book-keeping for this calculation is characteristically laborious, but is easily accomplished using a computer algebra package; the result is

$$G^{(2)}_{ab} = -\kappa t_{ab} + \tfrac{1}{2} h \widehat{G}_{abcd} h^{cd}, \tag{2.128}$$

where t_{ab} is given by (2.28). As expounded in section 2.2.2, and now confirmed by direct calculation (2.126), the first-order approximation to the Einstein field equation is $\widehat{G}_{abcd} h^{cd} = 0$, so $\widehat{G}_{abcd} h^{cd} = O(h^2)$ must hold true at second-order. Clearly it follows from this that $h\widehat{G}_{abcd} h^{cd} = O(h^3)$, and hence (2.30) is verified.

The third-order difference between $G^{(2)}_{ab}$ and $-\kappa t_{ab}$ exists because the field equation approximated to second-order in (2.31) is actually $\sqrt{-g} G^{ab}/\sqrt{-\check{g}} = 0$; this is of course entirely equivalent to the usual form of the Einstein field equation $G_{ab} = 0$.

Chapter 3

Localising the Energy and Momentum of Linear Gravity

3.1 Introduction

Half a century ago, a simple argument established that gravitational waves carry energy and can exchange this energy with matter. Often attributed to Feynman (certainly popularised by Bondi [19]) the argument asked us to imagine a gravitational detector comprising a rigid rod along which two "sticky beads" are threaded. A passing gravitational wave then acts to alter the proper distance between the beads, and this motion, opposed by friction, heats the detector and thus mediates a transfer of energy from gravity to matter. Despite the simplicity of this idea, even after fifty years, it has not been possible to explain *where* in spacetime this gravitational energy resides, and it is generally accepted that attempts to do so are "looking for the right answer to the wrong question"[58, §20.4].

The elusiveness of the "right answer", and the wrongness of the question, are very often identified as arising from gravity's gauge freedom, the consequence of which is a one-to-many mapping between physical spacetime and whatever localisation of gravitational energy-momentum might be proposed. Historically this issue was cast in terms of coordinate dependence, and the multitude of non-covariant objects that were constructed (first by Einstein [35], and most famously by Landau and Lifshitz [52]) were termed energy-momentum *pseudotensors*. However, a more recent formulation [11] has made it clear that the construction of a genuine tensor (defined on some background spacetime) is not the central problem; rather, it is the tensor's dependence on the arbitrary diffeomorphism that maps physical spacetime to the background (see chapter 1).

Nevertheless, there is no reason *a priori* that gauge dependence should preclude the construction of a physically unambiguous tensor, provided we are prepared to remove the gauge freedom in some well-defined way. In cosmology this is frequently done by constructing new variables which are gauge invariant but equal to the relevant gauge-dependent fields (such as gravity or density fluctuations) in a particular gauge [12, 76]; however, it is just as effective to provide a physically unambiguous method by which the gauge may be fixed, and to then insist that the gravitational field be evaluated in this gauge when locating its energy and momentum. Unfortunately, no previous approach has supplied instructions of this nature, and more importantly, neither the construction of these energy-momentum objects, nor their key properties, appear to favour one gauge (or one set of gauge-invariants) over another; thus it appears impossible to justify any of these seemingly arbitrary choices as *natural*.

Besides gauge dependence, there is also a great range of choice over which properties, physical or mathematical, should define the gravitational energy-momentum tensor: should we be guided by a putative conservation law, or have in mind a particular role in the field equations? For instance, it is always possible to locate the energy-momentum of *matter* by measuring the gravity it generates, so one might suggest that gravity's energy-momentum should be localised in a similar fashion, by examining the interaction it has with itself. Following this idea to its conclusion, it has been shown (in chapter 2 and elsewhere [25, 27, 36]) that general relativity may be constructed from an initially linear (spin-2) field theory that is then systematically coupled to its own (Hilbert) energy-momentum tensor. Sadly, this scheme leads us to identify the non-linear part of the Einstein tensor $G_{ab} - G^{(1)}_{ab}$ as the gravitational energy-momentum, so (a) the gauge problem remains, and (b) the result is additionally ambiguous, as different choices of "gravitational field" (g_{ab}, g^{ab}, $\sqrt{-g}g^{ab}$, etc.) mix the linear and non-linear terms in G_{ab}.

In spite of these various difficulties, one aspect of this enduring problem stands opposed to conventional wisdom and motivates our present discussion: when gravity and matter interact, the *exchange* of energy is local! To see this we need look no further than the sticky bead detector: here, the energy exchange is certainly localised in so far as it takes place only within the confines of the detector. Furthermore, we can imagine a very small detector, much smaller than a wavelength of the incident gravitational radiation, and observe that at each instant a well-defined power is developed in the detector as heat; thus, at least in this case, the rate of energy exchange is associated with a particular point in spacetime. One might hope, therefore, that consistency with this phenomenon would be enough to localise the energy and momentum of the gravitational field outside the detector, or even when no detector is present. Moreover, even if a gravitational energy-momentum tensor could not be found, there would still be great value in constructing a framework for the description and analysis of local gravitational energy-momentum exchange. The purpose of this chapter is to develop precisely this framework, and to examine the gravitational energy-momentum tensor it brings to light. In doing so we uncover a simple and unambiguous "right answer" through which the effects of gravitational energy-momentum may be usefully understood. Conceivably, this was the "right question" to ask.

For the sake of simplicity, we have restricted our present discussion to *linearised* general relativity on a flat Minkowski background. It is only in this linear regime that the convenient fiction of a "gravitational field" propagating on a background spacetime can be taken seriously, a construction which is essentially unavoidable when localising gravitational energy-momentum.[1] On a technical level, the restriction to the linear approximation

[1] As long as there is some spacetime with everywhere vanishing gravitational energy-momentum, then

limits the space of gauge transformations to a manageable size, facilitating the analysis and eventual removal of our description's gauge dependence. Furthermore, our gravitational energy-momentum tensor will not be derived from non-linear terms in the field equations, so we avoid any ambiguity arising from field redefinition. We shall not attempt to extend our results beyond the linear theory at this time.[2]

The structure of the chapter will be as follows. We begin by building the foundations of our framework, deriving a gravitational energy-momentum tensor (3.23) by demanding consistency with the energy and momentum exchanged with matter. As we will see, most of the tensor's gauge freedom is eliminated immediately as a natural consequence of this derivation. Following this, we demonstrate two important additional properties of our tensor, further solidifying its interpretation as gravity's energy-momentum tensor. We then develop our framework more concretely by analysing the transfer of gravitational energy-momentum onto an infinitesimal detector; in the process of making this analysis gauge invariant, we will purge the last trace of gauge ambiguity from our energy-momentum tensor. Finally, we examine the gravitational energy-momentum in some specific examples. Throughout, we work in units where $c = 1$, write $\kappa \equiv 8\pi G$, and use the sign conventions of Wald [79]: the metric signature is $(-,+,+,+)$, and the Riemann and Ricci tensors are defined by $[\nabla_c, \nabla_d]v^a \equiv R^a{}_{bcd}v^b$, and $R_{ab} \equiv R^c{}_{acb}$.

3.2 Motivation and Derivation

The purpose of this section is to explain how, by considering the energy-momentum transferred between matter and gravity, we are led to a formula for the gravitational energy-momentum tensor. We begin by laying down some mathematical groundwork.

3.2.1 Preliminaries

As previously explained, this chapter focuses exclusively on *linear* gravity: we only consider physical spacetimes $(\mathcal{M}, g_{ab})$ in which the curvature $R^a{}_{bcd}$ is everywhere small. As usual, this allows us to identify the physical spacetime with a *flat* background spacetime $(\check{\mathcal{M}}, \check{g}_{ab})$, where $\check{R}^a{}_{bcd} = 0$, using a diffeomorphism $\phi : \mathcal{M} \to \check{\mathcal{M}}$. The "gravitational field"

this will naturally play the role of the background, and fluctuations away from this configuration will constitute the gravitational field. Although the most natural choice for this "ground-state" is flat spacetime, this does not necessarily preclude the extension of our formalism to less trivial backgrounds; however, we suspect there may be technical or conceptual problems with "ignoring" the energy-momentum of a non-trivial background. In particular, we anticipate issues analogous to those of associating energy-momentum with a fluctuation in the electromagnetic field δF_{ab} when the background $\check{F}_{ab}$ is non-zero: the energy-momentum tensor $T \sim \check{F}^2 + \check{F}\delta F + (\delta F)^2$, so the dominant contribution from the fluctuation will be linear in the field, rather than quadratic.

[2]Of course, it may not be possible to extend the framework we develop here to the full non-linear theory, and we accept that localising gravitational energy-momentum in this regime (where the distinction between background and fluctuation is virtually meaningless) may be an inherently flawed idea. Of course, this does not alter the validity of our work in the linear case, where the "field theoretic" view is justified.

h_{ab} is then defined on $\check{\mathcal{M}}$ by

$$\phi^* g_{ab} = \check{g}_{ab} + h_{ab}, \tag{3.1}$$

and we insist that ϕ be chosen such that h_{ab} is small everywhere, in order that the *linearised* Einstein field equations are a good approximation:[3]

$$\widehat{G}_{ab}{}^{cd} h_{cd} = \kappa \check{T}_{ab} + O(h^2). \tag{3.2}$$

In the above relation, $\check{T}_{ab} \equiv \phi^* T_{ab} = O(h)$ is the matter energy-momentum tensor T_{ab} mapped onto the background, and

$$\widehat{G}_{ab}{}^{cd} h_{cd} \equiv \check{\nabla}_c \check{\nabla}_{(a} h_{b)}{}^c - \tfrac{1}{2}\check{\nabla}^2 h_{ab} - \tfrac{1}{2}\check{\nabla}_a \check{\nabla}_b h + \tfrac{1}{2}\check{g}_{ab}\left(\check{\nabla}^2 h - \check{\nabla}_c \check{\nabla}_d h^{cd}\right) \tag{3.3}$$

is the linearised Einstein tensor $G^{(1)}_{ab}$. Our freedom of choice over ϕ will of course give rise to the usual gauge transformation $\delta h_{ab} = \check{\nabla}_{(a}\xi_{b)}$.

On the background it will be useful to define four vectors[4] $\{\check{e}_\mu{}^a\}$ obeying

$$\check{\nabla}_a \check{e}^b_\mu = 0, \tag{3.4}$$

$$\check{e}_\mu{}^a \check{e}_{\nu a} = \eta_{\mu\nu}, \tag{3.5}$$

which form the basis of a Lorentz coordinate system $\{x^\mu\}$ on $\check{\mathcal{M}}$: $\check{e}_\mu{}^a \equiv (\partial_\mu)^a$. From this starting point, we shall define a corresponding set of vector fields $\{e_\mu{}^a\}$ in the physical spacetime,

$$e_\mu{}^a \equiv (\phi^{-1})^* \check{e}_\mu{}^a, \tag{3.6}$$

the behaviour of which will only be determined once we have fixed the gauge ϕ, an issue to which we will return later.

3.2.2 Energy-Momentum Currents

Superficially, general relativity is a theory in which the energy and momentum of matter is always conserved:

$$\nabla^a T_{ab} = 0. \tag{3.7}$$

However, the sticky bead argument has already demonstrated that this is not the case; in reality, matter may gain (or lose) energy through interaction with the gravitational field. The reason for this apparent contradiction is as follows. In order to determine the energy of each part of the detector, one must first specify a timelike vector field

[3] We use $O(h^n)$ as an abbreviation of $O((h_{ab})^n)$; this should not be confused with the trace of the gravitational field $h \equiv h_{ab}\check{g}^{ab}$.

[4] We use Roman letters as abstract tensor indices [79, §2.4] and Greek letters as numerical indices running from 0 to 3. Tensor indices of fields defined on the background are of course raised and lowered with $\check{g}_{ab}$.

${e_0}^a$ (the "time direction" conjugate to the energy) with which to form an energy current-density $J^a \equiv T^a{}_b {e_0}^b$. The incoming gravitational wave will then prevent ${e_0}^b$ from satisfying $\nabla_a {e_0}^b = 0$, and we will find that $\nabla_a J^a = T^a{}_b \nabla_a {e_0}^b \neq 0$. This inequality indicates a mismatch between the energy of the matter flowing into a given point, and the change in energy of the matter at that point; in other words, it represents the appearance of *additional energy* which was not already present in the matter – this is the energy absorbed from the gravitational wave! What is needed, therefore, is a framework which can account for this gained energy by identifying a corresponding loss in the energy of the gravitational field. We devote the rest of this section to the development of this idea, which will form the basis of our description of gravitational energy-momentum.

Following the previous discussion, it should now be clear that we must define one energy current-density, and three momentum current-densities, by

$$ {J_\mu}^a \equiv T^a{}_b {e_\mu}^b, \tag{3.8} $$

using the vectors $\{{e_\mu}^a\}$ that get mapped to the Lorentz basis of the background. This is a generalisation of the practice of defining conserved currents by contracting T_{ab} with a killing vector in a spacetime with a continuous symmetry. Here, however, the vector fields $\{{e_\mu}^a\}$ only correspond to *approximate* symmetries (present because spacetime is nearly flat) and thus the currents will not be conserved. The real difficulty is choosing sensible behaviour for $\{{e_\mu}^a\}$ that sufficiently captures the "parallelism" of killing vectors in the absence of any gravitational symmetry. Because ${e_\mu}^a \equiv (\phi^{-1})^* {\check{e}_\mu}{}^a$, this question has been recast as a choice of gauge, which we will address later.

Having defined our energy-momentum currents (apart from specifying ϕ) we are now in a position to express the key idea of our approach. We seek a symmetric tensor field τ_{ab}, defined on the background, that is a quadratic function of the gravitational field h_{ab}. We wish to be able to interpret τ_{ab} as the energy-momentum tensor of the gravitational field, and we shall achieve this by insisting that its non-conservation (in the background) exactly balances the non-conservation of the ${J_\mu}^a$ in the physical spacetime. Specifically, we wish to be able to define gravitational energy-momentum current-densities ${j_\mu}^a$ by

$$ {j_\mu}^a \equiv \tau^a{}_b {\check{e}_\mu}{}^b, \tag{3.9} $$

such that

$$ \check{\nabla}_a {j_\mu}^a + \phi^* (\nabla_a {J_\mu}^a) = 0. \tag{3.10} $$

This equation captures the idea that energy-momentum is transferred between matter and the gravitational field. In particular, equation (3.10) indicates that knowing the behaviour of h_{ab} at some point will be sufficient to determine the fields $\nabla_a {J_\mu}^a$ that express the local change in energy-momentum of the matter at the corresponding point in the physical spacetime.

We proceed by calculating the two elements of (3.10). Because we are using a Lorentz basis in the background, $\check{\nabla}_a \check{e}^b_\mu = 0$ trivially gives

$$ \check{\nabla}_a {j_\mu}^a = {\check{e}_\mu}{}^b \check{\nabla}_a \tau^a{}_b. \tag{3.11} $$

The second term is a little less trivial; using (3.7),

$$\begin{aligned}\phi^*(\nabla_a J_\mu{}^a) &= \phi^*(T^a_{\ b}\nabla_a e_\mu{}^b),\\ &= \phi^* T^a_{\ b}\phi^*(\nabla_a e_\mu{}^b),\\ &= (\check{T}^a_{\ b} + O(h^2))(\check{\nabla}_a \check{e}_\mu{}^b + \check{e}_\mu{}^c C^b{}_{ac}),\end{aligned} \tag{3.12}$$

where $C^a{}_{bc} = \frac{1}{2}(\check{\nabla}_b h_c{}^a + \check{\nabla}_c h_b{}^a - \check{\nabla}^a h_{bc}) + O(h^2)$ is the connection between the two derivative operators: $\phi^*(\nabla_a v^b) = \check{\nabla}_a \phi^* v^b + C^b{}_{ac}\phi^* v^c$. Now, because $\check{\nabla}_a \check{e}^b_\mu = 0$, and $\check{T}_{ab} = \check{T}_{ba}$, we have

$$\begin{aligned}\phi^*(\nabla_a J_\mu{}^a) &= \tfrac{1}{2}\check{T}^a_{\ b}\check{e}_\mu{}^c(\check{\nabla}_c h_a{}^b + \check{\nabla}_a h_c{}^b - \check{\nabla}^b h_{ac}) + O(h^3)\\ &= \tfrac{1}{2}\check{e}_\mu{}^c \check{T}^a_{\ b}\check{\nabla}_c h_a{}^b + O(h^3).\end{aligned} \tag{3.13}$$

Finally, we use the field equations (3.2) to write

$$\phi^*(\nabla_a J_\mu{}^a) = \frac{1}{2\kappa}\check{e}_\mu{}^q(\widehat{G}_{ab}{}^{cd}h_{cd})(\check{\nabla}_q h^{ab}) + O(h^3). \tag{3.14}$$

Inserting (3.11) and (3.14) into (3.10), and discarding the $O(h^3)$ terms, we arrive at the defining relation of the gravitational energy-momentum tensor:

$$\kappa\check{\nabla}^a \tau_{aq} = -\tfrac{1}{2}(\check{\nabla}_q h^{ab})\widehat{G}_{ab}{}^{cd}h_{cd}. \tag{3.15}$$

The next step will be to use this equation to derive a formula for τ_{ab} in terms of h_{ab}. In order to do so, however, we must make one additional demand: τ_{ab} will not depend on second derivatives of h_{ab}, but will be a function of $\check{\nabla}_c h_{ab}$ and $\check{g}_{ab}$ only. The reason we must impose this condition is that equation (3.15) can only define τ_{ab} up to the addition of "superpotential" terms, those fields whose divergence vanishes *identically*. Because these terms are of the form $\check{\nabla}^c\check{\nabla}^d H_{[ac][bd]}$ (where $H_{[ac][bd]} = H_{[bd][ac]}$ is some function of h_{ab}) they necessarily contain second derivatives; thus our restriction on τ_{ab} is sufficient to remove this ambiguity. At the moment, it might be tempting to view this condition as a convenient way to tame the derivation, and keep in mind that we can always add in super-potentials later if we wish. However, in section 3.3 it will become clear that many of the interesting properties displayed by τ_{ab} will be unavoidably spoilt by the addition of such terms. For this reason we will not consider superpotentials further here.

3.2.3 Determining the Energy-Momentum Tensor

In truth, it will not be possible to construct a symmetric tensor τ_{ab} that satisfies (3.15) for all h_{ab};[5] to make progress we will need to impose some condition on $\check{\nabla}_c h_{ab}$ and specialise to this restricted set of gravitational fields. Although this forced restriction might appear to be a flaw in our formalism, as we shall soon see, it is actually a valuable asset.

There are only three linear conditions we can place on $\check{\nabla}_c h_{ab}$ which neither introduce extra fields, nor break Lorentz invariance: (a) $\check{\nabla}_c h_{ab} = 0$, (b) $\check{\nabla}_a h = 0$, or (c) $\check{\nabla}^a h_{ab} =$

[5] We will shortly describe how to check this assertion, which is simply a property of (3.15) and independent of the requirement that τ_{ab} contain no second derivatives.

$\lambda\check{\nabla}_b h$, for some constant λ.[6] Condition (a) is obviously far too restrictive: it does not allow us any gravitational field whatsoever. In contrast, condition (b) is not restrictive enough: there is no τ_{ab} that solves (3.15) for all gravitational fields with constant trace.[7] We must therefore focus on condition (c), which we repeat for later reference:

$$\check{\nabla}^a h_{ab} = \lambda\check{\nabla}_b h. \tag{3.16}$$

Using this relation, it will be possible to replace any occurrence of $\check{\nabla}^a h_{ab}$ with $\lambda\check{\nabla}_b h$; hence the most general formula for a symmetric tensor τ_{ab}, a quadratic function of $\check{\nabla}_c h_{ab}$, is as follows:

$$\begin{aligned}\kappa\tau_{pq} = {}& \check{g}_{pq}(A_0\check{\nabla}_c h_{ab}\check{\nabla}^c h^{ab} + A_1\check{\nabla}_a h\check{\nabla}^a h + A_2\check{\nabla}_c h_{ab}\check{\nabla}^b h^{ac}) + A_3\check{\nabla}_p h_{ab}\check{\nabla}_q h^{ab}\\ &+ A_4\check{\nabla}_p h\check{\nabla}_q h + A_5\check{\nabla}_a h\check{\nabla}_{(p}h_{q)}{}^a + A_6\check{\nabla}^a h^b{}_{(p}\check{\nabla}_{q)}h_{ab} + A_7\check{\nabla}_a h_{bp}\check{\nabla}^a h_q{}^b\\ &+ A_8\check{\nabla}_b h_{ap}\check{\nabla}^a h_q{}^b + A_9\check{\nabla}_a h\check{\nabla}^a h_{pq},\end{aligned} \tag{3.17}$$

where $\{A_n\}$ are arbitrary constants. We proceed by substituting this ansatz into (3.15) and solving for $\{A_n\}$. First, let us calculate $\check{\nabla}^p\tau_{pq}$ by taking the divergence of (3.17); using (3.16) to convert every $\check{\nabla}^a h_{ab}$ to $\lambda\check{\nabla}_b h$, and collecting terms, we find that left-hand side of (3.15) amounts to

$$\begin{aligned}\kappa\check{\nabla}^p\tau_{pq} = {}& (2A_0 + A_3)\check{\nabla}_q\check{\nabla}_c h_{ab}\check{\nabla}^c h^{ab} + (2A_1 + A_4 + \tfrac{1}{2}\lambda A_5 + \lambda A_9)\check{\nabla}_a h\check{\nabla}^a\check{\nabla}_q h\\ &+ (2A_2 + \tfrac{1}{2}A_6)\check{\nabla}_c h_{ab}\check{\nabla}_q\check{\nabla}^b h^{ac} + A_3\check{\nabla}^2 h_{ab}\check{\nabla}_q h^{ab} + A_4\check{\nabla}^2 h\check{\nabla}_q h\\ &+ (\tfrac{1}{2}A_5 + \lambda A_7 + \lambda A_8 + A_9)\check{\nabla}_a\check{\nabla}^b h\check{\nabla}^a h_{bq} + (\tfrac{1}{2}A_5 + \tfrac{1}{2}\lambda A_6)\check{\nabla}_a\check{\nabla}_b h\check{\nabla}_q h^{ab}\\ &+ \tfrac{1}{2}A_5\check{\nabla}^a h\check{\nabla}^2 h_{aq} + (\tfrac{1}{2}A_6 + A_7)\check{\nabla}^a\check{\nabla}_b h_{cq}\check{\nabla}_a h^{bc} + \tfrac{1}{2}A_6\check{\nabla}_a h_{bq}\check{\nabla}^2 h^{ab}\\ &+ A_8\check{\nabla}^c h_{ab}\check{\nabla}^a\check{\nabla}^b h_{cq}.\end{aligned} \tag{3.18}$$

Meanwhile, (3.16) simplifies the right-hand side of (3.15):

$$-\tfrac{1}{2}(\check{\nabla}_q h^{ab})\widehat{G}_{ab}{}^{cd}h_{cd} = -\tfrac{1}{2}\check{\nabla}_q h^{ab}\left((\lambda - \tfrac{1}{2})\check{\nabla}_a\check{\nabla}_b h - \tfrac{1}{2}\check{\nabla}^2 h_{ab} + \tfrac{1}{2}\check{g}_{ab}(1-\lambda)\check{\nabla}^2 h\right). \tag{3.19}$$

Comparing (3.18) with (3.19), term by term, we conclude that the *unique* solution to (3.15) is

$$\begin{aligned}&A_0 = -\tfrac{1}{8}, \quad A_1 = \tfrac{1}{16}, \quad A_3 = \tfrac{1}{4}, \quad A_4 = -\tfrac{1}{8},\\ &A_2 = A_5 = A_6 = A_7 = A_8 = A_9 = 0, \quad \lambda = \tfrac{1}{2}.\end{aligned} \tag{3.20}$$

We have therefore determined the formula for our gravitational energy-momentum tensor,

$$\kappa\tau_{pq} = \tfrac{1}{4}\check{\nabla}_p h_{ab}\check{\nabla}_q h^{ab} - \tfrac{1}{8}\check{\nabla}_p h\check{\nabla}_q h - \tfrac{1}{8}\check{g}_{pq}(\check{\nabla}_c h_{ab}\check{\nabla}^c h^{ab} - \tfrac{1}{2}\check{\nabla}_a h\check{\nabla}^a h), \tag{3.21}$$

[6] The only other possibility, $\check{\nabla}^a h_{ab} = 0$, can be achieved by taking (c) with $\lambda = 0$.

[7] For the sake of brevity, we will not prove this assertion here. Instead we will attend to condition (c) and derive the formula for τ_{ab} that it admits. After we have done so, we invite the reader to perform a similar calculation under condition (b) and verify that no solution exists.

and the condition,

$$\check{\nabla}^a h_{ab} - \tfrac{1}{2}\check{\nabla}_b h = 0, \tag{3.22}$$

that we must place on the gravitational field. Finally, we introduce the abbreviation $\bar{h}_{ab} = h_{ab} - \frac{1}{2}\breve{g}_{ab}h$ for the trace-reversed gravitational field, and $\bar{\tau}_{ab} = \tau_{ab} - \frac{1}{2}\breve{g}_{ab}\tau$ for the trace-reversed gravitational energy-momentum tensor, allowing us to compactly re-express our results:

$$\kappa\bar{\tau}_{pq} = \tfrac{1}{4}\check{\nabla}_p h_{ab}\check{\nabla}_q \bar{h}^{ab}, \tag{3.23}$$

$$\check{\nabla}^a \bar{h}_{ab} = 0. \tag{3.24}$$

We are now in a position to justify our earlier claim, that the restriction on the gravitational field is not a hindrance but a major advantage. The condition we have derived (3.24) is simply the defining equation of the *harmonic gauge*.[8] As this equation can *always* be satisfied by making a gauge transformation, it does not in any way limit the physical applicability of our approach! Moreover, we have received a valuable gift: this condition has appeared as a natural consequence of the derivation, and forces upon us a very strong restriction for the diffeomorphism ϕ that maps the physical spacetime onto the background. Specifically, ϕ^{-1} is required to map the Lorentz coordinates of the background onto *harmonic* coordinates in the physical spacetime.[9] We can therefore think of (3.24) as the condition that specifies the correct behaviour to demand of the basis $\{e_\mu{}^a\}$ needed to define sensible energy-momentum currents $J_\mu{}^a$. We should stress, however, that while the harmonic gauge condition has removed the vast majority of the gauge freedom, a small amount remains in the form of transformations $\delta h_{ab} = \check{\nabla}_{(a}\xi_{b)}$ which satisfy $\check{\nabla}^2\xi_a = 0$; we will return to this issue in section 3.4.

This completes the derivation of τ_{ab}. We have found the unique symmetric tensor, a quadratic function of $\check{\nabla}_c h_{ab}$, that describes the transfer of energy and momentum between matter and the gravitational field according to (3.10). In doing so we have found that this solution only exists if the harmonic condition $\check{\nabla}_a \bar{h}^{ab} = 0$ is obeyed, and this in turn has solidified the definition of energy-momentum currents $J_\mu{}^a$ as the contraction of T_{ab} with the basis vectors associated with *harmonic* coordinate systems of physical spacetime. In the next section we shall prove that τ_{ab} displays many other interesting properties very much in keeping with its interpretation as an energy-momentum tensor. In section 3.4 we shall examine energy-momentum exchange in detail, and address the last piece of gauge freedom.

[8]This is also commonly referred to as de Donder gauge or Lorentz gauge.

[9]To see this, let $\{x^\mu\}$ be Lorentz coordinates on the background ($\check{\nabla}_a\check{\nabla}_b x^\mu = 0$), and let $\{y^\mu\}$ be coordinates in physical spacetime defined by $y^\mu(p) = x^\mu(\phi(p))$ for all $p \in \mathcal{M}$. Then $\phi^*(\nabla^2 y^\mu) = \check{\nabla}^2 x^\mu - h^{ab}\check{\nabla}_a\check{\nabla}_b x^\mu - (\check{\nabla}^a h_{ab} - \frac{1}{2}\check{\nabla}_b h)\check{\nabla}^b x^\mu = 0$.

3.3 Properties

Here we will demonstrate that, in two important special cases, τ_{ab} exhibits interesting mathematical properties (beyond accounting for $\nabla_a J_\mu{}^a$) that further promote its interpretation as the energy-momentum tensor of the linear gravitational field.

3.3.1 Gauge Invariance of Plane-Waves

Consider an arbitrary gravitational *plane-wave*:

$$h_{ab} = h_{ab}(k_\mu x^\mu). \tag{3.25}$$

Here, k_a is a constant vector, and $\{x^\mu\}$ are Lorentz coordinates on the background. The linear vacuum field equation $\check{\nabla}^2 \bar{h}_{ab} = 0$, and the harmonic condition $\check{\nabla}^a \bar{h}_{ab} = 0$, enforce

$$k^a k_a = 0, \quad k^a \bar{h}'_{ab} = 0, \tag{3.26}$$

respectively, where the prime indicates differentiation with respect to the variable $k_\mu x^\mu$. We wish to consider the most general gauge transformation $\delta h_{ab} = \check{\nabla}_{(a}\xi_{b)}$ that maintains the plane-wave form of h_{ab}. Clearly we require $\xi_a = \xi_a(k_\mu x^\mu)$, and thus

$$\delta h_{ab} = k_{(a}\xi'_{b)}. \tag{3.27}$$

Note that $k^a k_a = 0$ now guarantees $\check{\nabla}^2 \xi_a = 0$, ensuring that the harmonic condition (3.24) is not broken by the transformation. Let us now calculate the effect of this transformation on the gravitational energy-momentum tensor; working from (3.23),

$$\begin{aligned}
\kappa \delta \bar{\tau}_{pq} &= \tfrac{1}{2} \check{\nabla}_p \delta h_{ab} \check{\nabla}_q \bar{h}^{ab} + \tfrac{1}{4} \check{\nabla}_p \delta h_{ab} \check{\nabla}_q \delta \bar{h}^{ab} \\
&= \tfrac{1}{2} k_p k_q k_{(a}\xi''_{b)} \bar{h}'^{ab} + \tfrac{1}{4} k_p k_q k_{(a}\xi''_{b)} (k^{(a}\xi''^{b)} - \tfrac{1}{2} \check{g}^{ab} k^c \xi''_c) \\
&= \tfrac{1}{8} k_p k_q ((k^c \xi''_c)^2 - (k^c \xi''_c)^2) = 0.
\end{aligned} \tag{3.28}$$

Thus the energy-momentum of an arbitrary gravitational plane-wave is completely invariant under the gauge freedom consistent with the harmonic condition and its plane-wave form. This is significant for a number of reasons. Firstly it reveals that, for the special case of plane-waves, we need not concern ourselves with the gauge freedom that remains after enforcing the harmonic condition: the requirement that the gauge be chosen such that the plane-wave form of the field be manifest is sufficient to unambiguously define the energy-momentum tensor τ_{ab} from the physical spacetime $(\mathcal{M}, g)$. Thus, even if one does not accept our method for resolving the last of the gauge ambiguity (to be presented in section 3.4) it is still possible to stop at this point and agree that a well-defined energy-momentum tensor for linearised gravitational plane-waves has been found. Secondly, this particular gauge invariance will prove useful when we wish to produce a *global* picture of the motion of energy-momentum: if the *source region* of a gravitational wave is very far from the *detection region*, we may use a different gauge in each and yet still produce a consistent picture of energy-momentum transfer – the field in intermediate region will approximate a plane-wave, and thus τ_{ab} in this region will agree with both end-point gauges.[10] There is also a third significance to this result, but this will only become apparent once we have demonstrated the second important property of gravitational energy-momentum: positivity.

3.3.2 Positivity

This section concerns the energy-momentum of *transverse-traceless* (TT) gravitational fields, those for which $h = 0$, $\check{\nabla}^a h_{ab} = 0$, and $u^a h_{ab} = 0$, for some constant timelike vector field u^a defined on the background.[11] For now let us simply suppose that these conditions apply to h_{ab} and derive the consequences for τ_{ab}. We shall justify our interest in this specialisation, and offer an interpretation of u^a, in section 3.4. Presently, let it suffice to say that because these conditions may always be imposed (at least locally) by a gauge transformation in regions where $\check{T}_{ab} = 0$, the results of this section will be generally applicable to vacuum regions, but it will not be necessary to demand that $\check{T}_{ab} = 0$ globally. We present our result as the following theorem.

Theorem. *If, at some point $p \in \check{\mathcal{M}}$, the gravitational field h_{ab} obeys the transverse-traceless conditions*

$$\check{\nabla}^a h_{ab} = 0, \qquad h = 0, \qquad u^a h_{ab} = 0, \tag{3.29}$$

for some timelike vector u^a, then τ_{ab} satisfies the following inequalities

$$v^a \tau_{ab} v^b \geq 0, \tag{3.30}$$
$$v^a \tau_{ac} \tau^c{}_b v^b \leq 0, \tag{3.31}$$

at p, for any timelike vector v^a.

Proof. Without loss of generality we can set $u^a u_a = -1$ and $v^a v_a = -1$. Now, introduce two Lorentzian coordinate systems at p, the first $\{x^0, x^i\}$ $(i = 1,2,3)$ such that $u^0 = 1$, $u^i = 0$, and the second $\{x^{0'}, x^{i'}\}$ $(i' = 1,2,3)$ such that $v^{0'} = 1$, $v^{i'} = 0$. The transverse-traceless conditions (3.29) reduce (3.23) to

$$\kappa \tau_{pq} = -\tfrac{1}{8} \check{g}_{pq} \check{\nabla}_c h_{ab} \check{\nabla}^c h^{ab} + \tfrac{1}{4} \check{\nabla}_p h_{ab} \check{\nabla}_q h^{ab}, \tag{3.32}$$

and set $h_{0i} = h_{00} = 0$. Using the primed basis to express the tensor indices of $\check{g}_{ab}$ and $\check{\nabla}_a$, the unprimed basis for h_{ab}, and writing $\dot{h}_{ij} \equiv \partial_{0'} h_{ij}$, we find that

$$\begin{aligned} \kappa v^a \tau_{ab} v^b &= \kappa \tau_{0'0'} \\ &= \tfrac{1}{8}((\partial_{i'} h_{ij})^2 - (\dot{h}_{ij})^2) + \tfrac{1}{4}(\dot{h}_{ij})^2 \\ &= \tfrac{1}{8}((\partial_{i'} h_{ij})^2 + (\dot{h}_{ij})^2) \geq 0, \end{aligned} \tag{3.33}$$

[10] This idea is explained fully in section 3.4.4.

[11] Clearly, the harmonic condition is satisfied as a result of these requirements.

because $(\partial_{i'}h_{ij})^2 \equiv \sum_{i,j,i'=1}^{3} \partial_{i'}h_{ij}\partial_{i'}h_{ij}$ and $(\dot{h}_{ij})^2 \equiv \sum_{i,j=1}^{3} \partial_{0'}h_{ij}\partial_{0'}h_{ij}$ are sums of squares. Similarly,

$$\begin{aligned}(4\kappa v^a\tau_{ab})^2 &= 16\kappa^2(-(\tau_{0'0'})^2 + (\tau_{0'i'})^2)\\ &= -\tfrac{1}{4}((\partial_{i'}h_{ij})^2 + (\dot{h}_{ij})^2)^2 + (\dot{h}_{ij}\partial_{i'}h_{ij})(\dot{h}_{kl}\partial_{i'}h_{kl})\\ &= -\tfrac{1}{4}((\partial_{i'}h_{ij})^2 - (\dot{h}_{ij})^2)^2 - (\dot{h}_{ij})^2(\partial_{i'}h_{kl})^2 + (\dot{h}_{ij}\partial_{i'}h_{ij})(\dot{h}_{kl}\partial_{i'}h_{kl})\\ &= -\tfrac{1}{4}((\partial_{i'}h_{ij})^2 - (\dot{h}_{ij})^2)^2 - \tfrac{1}{2}(\dot{h}_{ij}\partial_{i'}h_{kl} - \dot{h}_{kl}\partial_{i'}h_{ij})^2 \leq 0. \end{aligned} \tag{3.34}$$

□

The inequalities we have just deduced are the gravitational version of the *Dominant Energy Condition*: the first indicates that τ_{ab} only ever defines positive energy-densities; the second indicates that the flux of this energy can never be spacelike. Succinctly, they tell us that gravitational energy is positive and never flows faster than light. As the Dominant Energy Condition has always referred to *matter*, we will avoid confusion if we resist subsuming (3.30) and (3.31) under this name; instead, when the gravitational energy-momentum tensor obeys these inequalities (for all timelike v^a) we shall simply say that it is *positive*, and write $\tau_{ab} \geq 0$ as a shorthand.

That τ_{ab} is positive for all transverse-traceless h_{ab} is one of the major advantages our approach has over previous descriptions of gravitational energy-momentum [11, 35, 52, 56, 59]. Provided we work with a transverse-traceless field (with respect to some timelike vector u^a), which is always possible locally in a vacuum, τ_{ab} will always make good physical sense in that it will obey its own version of the Dominant Energy Condition. To some extent, this result supplies its own justification for choosing the TT-gauge whenever possible; however, we will see in the next section that these conditions arise naturally by considering the gauge invariant transfer of energy-momentum onto point-sources. Furthermore, the significance of u^a (in terms of energy-momentum transfer) will also be explored through these arguments.

Before we move on, however, we take this opportunity to present an important corollary of the plane-wave gauge-invariance of section 3.3.1. It is well known that there always exists exactly one gauge transformation of the form (3.27) that takes an arbitrary plane-wave (3.25) to one obeying the TT-conditions [45, §18.1]. Hence we can transform any gravitational plane-wave into transverse-traceless gauge without altering the energy-momentum tensor, at which point the positivity theorem ensures that $\tau_{ab} \geq 0$. Thus, all gravitational plane-waves have positive energy-momentum tensors, even if they are not transverse-traceless.

3.4 Interactions

In this section we apply our formula for the gravitational energy-momentum tensor to the interaction between gravity and an idealised matter distribution that we shall refer to as a *point-source*. The reader will be familiar with the *compact source*, an isolated body confined to a compact spatial region of radius d much smaller than the wavelength λ of the gravitational radiation it emits; point-sources are the limit of such systems as $d \to 0$, entirely analogous to the infinitesimal dipoles of electromagnetism.[12] Not only will this provide a useful example of the practical application of our approach, these considerations will finally allow us to rid ourselves of the last trace of gauge-dependence in our description.

From now on we will work almost exclusively in the flat background spacetime; as such it will generally be convenient to represent all tensors in some Lorentzian coordinate system $\{x^\mu\}$, and to drop the "caron" mark from $\check{T}_{\mu\nu}$. Thus, our formula for the gravitational energy-momentum tensor is written as

$$\kappa\bar{\tau}_{\mu\nu} = \tfrac{1}{4}\partial_\mu h_{\alpha\beta}\partial_\nu \bar{h}^{\alpha\beta}, \tag{3.35}$$

the harmonic condition becomes

$$\partial^\mu \bar{h}_{\mu\nu} = 0, \tag{3.36}$$

and the linearised field equations (3.2) are

$$\partial^2 \bar{h}_{\mu\nu} = -2\kappa T_{\mu\nu}. \tag{3.37}$$

Also, it will be useful to separate $\{x^\mu\}$ into a time coordinate $t = x^0$, and spatial coordinates $\vec{x} = (x^1, x^2, x^3)$, define a radial coordinate $r \equiv |\vec{x}|$, and to use lower-case Roman indices $(i, j, k \ldots)$ to indicate spatial components. Typically, the coordinates will be implicitly chosen to coincide with the rest-frame of the system under consideration.

3.4.1 Pulses and Point-Sources

The core of our analysis will be to examine the most localised gravitational interaction possible: an infinitesimal point-source (at $\vec{x} = 0$) met by an instantaneous *pulse* plane-wave (propagating along the x^1 direction, arriving at $\vec{x} = 0$ at $t = t_0$).

As a result of calculations in appendix 3.A, we know that a point-source has the following energy-momentum tensor:

$$\begin{aligned} T_{00} &= M\delta(\vec{x}) + \tfrac{1}{2}I_{ij}\partial_i\partial_j\delta(\vec{x}), \\ T_{0i} &= \tfrac{1}{2}(\dot{I}_{ij} - L_{ij})\partial_j\delta(\vec{x}), \\ T_{ij} &= \tfrac{1}{2}\ddot{I}_{ij}\delta(\vec{x}). \end{aligned} \tag{3.38}$$

Here M and $L_{ij} = L_{[ij]}$ are constants representing, respectively, the mass and angular momentum of the source, and $I_{ij} = I_{(ij)}(t)$ its (time dependent) quadrupole moment.[13] We do not intend to use this point-source as an actual source of gravitational radiation, but rather as a probe of the energy-momentum of the incident pulse. To this end, we are interested in the limit $M, I_{ij}, L_{ij} \to 0$, allowing us to neglect the self-interaction of

[12] We derive the energy-momentum tensor and gravitational field of the point-source in appendix 3.A.

[13] The reader should refer to appendix 3.A for definitions of these quantities in terms of the infinitesimal limit of the compact source.

the source. As this procedure is entirely analogous to using a "test-particle" to probe the geometry of spacetime, we shall refer to the point-source as a *test-source* in this limit.

The gravitational field will consist of two parts: $h_{\mu\nu} = h^{\text{source}}_{\mu\nu} + h^{\text{wave}}_{\mu\nu}$. The first, due to the test-source, is given in appendix 3.A by (3.85) and satisfies the inhomogeneous field equations

$$\partial^2 \bar{h}^{\text{source}}_{\mu\nu} = -2\kappa T_{\mu\nu}. \tag{3.39}$$

The latter is the incident pulse plane-wave,

$$h^{\text{wave}}_{\mu\nu} = A_{\mu\nu} H(k_\alpha x^\alpha - t_0), \tag{3.40}$$

where H is the Heaviside step function, $k_\mu = (1, -1, 0, 0)$ is a null vector in the x^1 direction, and $A_{\mu\nu}$ is a constant tensor satisfying $k^\mu \bar{A}_{\mu\nu} = 0$, as demanded by the harmonic condition. Obviously, $h^{\text{wave}}_{\mu\nu}$ satisfies the homogeneous field equations: $\partial^2 h^{\text{wave}}_{\mu\nu} = 0$.

Let us now compute $\partial^\mu \tau_{\mu\nu}$, which, via (3.10), quantifies the exchange of energy-momentum between the test-source and the gravitational wave. Starting from (3.35), we have

$$\begin{aligned}\kappa \partial^\mu \tau_{\mu\nu} &= \tfrac{1}{4} \partial^2 \bar{h}_{\alpha\beta} \partial_\nu h^{\alpha\beta} \\ &= -\tfrac{1}{2} \kappa T_{\alpha\beta} \partial_\nu (A^{\alpha\beta} H(k_\sigma x^\sigma - t_0)) + O((h^{\text{source}}_{\mu\nu})^2),\end{aligned} \tag{3.41}$$

and we neglect terms of order $(h^{\text{source}}_{\mu\nu})^2$ compared to those of order $h^{\text{wave}}_{\mu\nu} h^{\text{source}}_{\mu\nu}$ in the limit $M, I_{ij}, L_{ij} \to 0$.[14] Using $H' = \delta$, the Dirac delta function, we arrive at

$$\begin{aligned}\partial^\mu \tau_{\mu\nu} &= -\tfrac{1}{2} k_\nu \delta(k_\sigma x^\sigma - t_0) T_{\alpha\beta} A^{\alpha\beta} \\ &= -\tfrac{1}{4} k_\nu \delta(k_\sigma x^\sigma - t_0) \Big(\ddot{I}_{ij} A_{ij} \delta(\vec{x}) - 2(\dot{I}_{ij} - L_{ij}) \partial_j \delta(\vec{x}) A_{i0} \\ &\qquad + (2M\delta(\vec{x}) + I_{ij} \partial_i \partial_j \delta(\vec{x})) A_{00} \Big).\end{aligned} \tag{3.42}$$

This is the equation we sought. It determines the energy and momentum collected by our probe due to the incident pulse, and locates this transfer in spacetime. The key problem is that above relation is not, as it stands, gauge invariant; we address this issue the next section.

3.4.2 Gauge Invariance and Microaveraging

The incident wave possesses gauge freedom that neither breaks the harmonic condition nor spoils its pulse plane-wave form:

$$\delta h^{\text{wave}}_{\mu\nu} = \partial_{(\mu} \xi_{\nu)}; \qquad \xi_\mu = E_\mu \Delta(k_\alpha x^\alpha - t_0), \tag{3.43}$$

[14] There is a slight technical issue here. From (3.85), we can see that, as $r \to 0$, $h^{\text{source}}_{\mu\nu} \to \infty$; thus $h^{\text{source}}_{\mu\nu}$ inevitably becomes larger than $h^{\text{wave}}_{\mu\nu}$ at small enough distances. Strictly speaking, then, one should use a finite-size source (of radius d, say) when one takes $M, I, J \to 0$ and neglects $O((h^{\text{source}}_{\mu\nu})^2)$. As we can choose d to be as small as we like, however, we can always replace the finite source with an equivalent point-source after this limit has been taken.

where E_μ is any constant vector, and $\Delta' = H$. The effect of this transformation is to alter $A_{\mu\nu}$ by $\delta A_{\mu\nu} = k_{(\mu}E_{\nu)}$, and although the transverse components do not change ($\delta A_{22} = \delta A_{23} = \delta A_{33} = 0$) the right-hand side of (3.42) is clearly not invariant. The beauty of working with an instantaneous interaction, however, is that we can average over the (infinitesimal) interaction region

$$\lim_{\epsilon\to 0}\mathcal{B}_\epsilon(t_0), \qquad \text{where} \quad \mathcal{B}_\epsilon(t_0) \equiv \{(t,\vec{x}) : |t-t_0| \le \epsilon, |\vec{x}| \le \epsilon\}, \tag{3.44}$$

without sacrificing the localised description of $\partial^\mu\tau_{\mu\nu}$. Let us call this operation a *microaverage* (at $\vec{x}=0$, $t=t_0$) and denote it by $\langle\ldots\rangle_{t_0}$:

$$\langle f\rangle_{t_0} \equiv \delta(\vec{x})\delta(t-t_0)\lim_{\epsilon\to 0}\int_{\mathcal{B}_\epsilon(t_0)} f\mathrm{d}^4x. \tag{3.45}$$

For the interaction we are analysing, the integral $\int_{\mathcal{B}_\epsilon(t_0)}\partial^\mu\tau_{\mu\nu}\mathrm{d}^4x$ captures the key physical content of $\partial^\mu\tau_{\mu\nu}$. To elaborate: the divergence theorem equates this integral with $\int_{\partial\mathcal{B}_\epsilon(t_0)}\tau_{\mu\nu}\mathrm{d}^3S^\mu$ which measures the mismatch between the net flux of gravitational energy-momentum entering through a spherical surface $\mathcal{S}_\epsilon \equiv \{\vec{x} : |\vec{x}| \le \epsilon\}$ barely larger the source, and the gravitational energy-momentum contained within $\mathcal{S}_\epsilon$ that is gained between the times $t = t_0-\epsilon$ and $t = t_0+\epsilon$. By the defining property (3.15) of $\tau_{\mu\nu}$, this mismatch in gravitational energy-momentum precisely accounts for the energy-momentum absorbed by the source, which is what we wanted to know. The only information we have lost in taking the microaverage is the knowledge of precisely where *within the test-source* the energy-momentum is being absorbed. As we have let the size of this probe shrink to zero, however, this is of little concern.

The computational advantage of the microaverage is that the integration in (3.45) allows us to transfer derivatives off the delta-functions in (3.42); for example,

$$\begin{aligned}\int_{\mathcal{B}_\epsilon(t_0)}\delta(k_\alpha x^\alpha - t_0)\dot{I}_{ij}\partial_j\delta(\vec{x})A_{i0}\mathrm{d}^4x &= -\int_{\mathcal{B}_\epsilon(t_0)}\partial_j\delta(t-x^1-t_0)\dot{I}_{ij}\delta(\vec{x})A_{i0}\mathrm{d}^4x\\ &= \int_{\mathcal{B}_\epsilon(t_0)}\dot{\delta}(t-x^1-t_0)\dot{I}_{i1}\delta(\vec{x})A_{i0}\mathrm{d}^4x\\ &= -\int_{\mathcal{B}_\epsilon(t_0)}\delta(t-x^1-t_0)\ddot{I}_{i1}\delta(\vec{x})A_{i0}\mathrm{d}^4x\\ &= -\ddot{I}_{i1}(t_0)A_{i0}.\end{aligned} \tag{3.46}$$

Applying this technique to the whole of (3.42) yields

$$\langle\partial^\mu\tau_{\mu\nu}\rangle_{t_0} = -\tfrac{1}{4}k_\nu\delta(\vec{x})\delta(t-t_0)\left(\ddot{I}_{ij}A_{ij} + 2\ddot{I}_{i1}A_{i0} + \ddot{I}_{11}A_{00} + 2MA_{00}\right). \tag{3.47}$$

Finally, we unpack $k^\mu\bar{A}_{\mu\nu} = 0$,

$$A_{00} + A_{11} + 2A_{01} = 0, \qquad A_{22} + A_{33} = 0, \qquad A_{02} + A_{12} = 0, \qquad A_{03} + A_{13} = 0, \tag{3.48}$$

and substitute these into (3.47). The result is

$$\langle\partial^\mu\tau_{\mu\nu}\rangle_{t_0} = -\tfrac{1}{2}k_\nu\delta(\vec{x})\delta(t-t_0)\left(\ddot{I}_\times A_\times + \ddot{I}_+A_+ + MA_{00}\right), \tag{3.49}$$

where we have written the transverse components of the wave as $A_\times = A_{23}$ and $A_+ = (A_{22} - A_{33})/2$, and extended this notation to I_{ij}. We are almost done: $\delta A_\times = \delta A_+ = 0$ under the gauge transformation (3.43), so the first two terms in (3.49) are manifestly gauge invariant; however, the term proportional to MA_{00} is not.

Various arguments can be made to show that this "monopole term" is physically irrelevant to the energy-momentum transfer we are considering. At the simplest level, the fact that we are free to set A_{00} to any value (including zero) through gauge transformation (leaving $A_\times$ and A_+ untouched) indicates that the monopole term can have no bearing on the energy-momentum of the physical system under scrutiny. Furthermore, if we consider a wave for which $A_{\mu\nu} = k_{(\mu}E_{\nu)}$, then it is clear that, while such a pulse is gauge-equivalent to flat spacetime ($A_{\mu\nu} = 0$) it would nonetheless register a transfer of energy-momentum if the monopole term were to be believed.

The physical irrelevance of the monopole term should come as no great surprise, as there can be no way to extract energy-momentum from a gravitational wave using a monopole alone (i.e. a test-source with $I_{ij} = L_{ij} = 0$): an observer sitting on an isolated point mass could perform no local test to distinguish whether a gravitational wave had even passed, and in particular, must be unable to extract any energy.

In fact, all that the monopole term is responding to is a change in normalisation of the time coordinate in the physical spacetime: $\phi^*(e_0^a e_{0a}) = -1 + h_{00}$. Naively, we might expect this factor to be significant as it represents the Newtonian potential at the test-source. However, this is not a local effect. The only way an observer on the test-source could be aware of such a shift is by comparison with some standard clocks at spatial infinity. The pulse plane-wave prevents this idea from being well-defined, however, as it divides spatial infinity into two regions: $x^1 < t - t_0$, where wave has already been received, and $x^1 > t - t_0$, where it has not. Fortunately, a gauge can always be chosen that does not suffer from this inconsistency; setting $A_{00} = 0$ is the only way to ensure that the standard clocks at infinity all run at the *same* rate (relative to our coordinate t) and this inevitably removes all trace of the monopole term from the interaction. Thus the insistence that the clocks at infinity agree with each other amounts to a prescription that removes the gauge-dependence of our microaveraged energy-momentum transfer.[15] We can implement this procedure mathematically (without fixing the gauge, or setting $M = 0$, which is physically untenable) by acting on $\langle \partial^\mu \tau_{\mu\nu} \rangle_{t_0}$ with the operator $(1 - M\partial_M)$:

$$\begin{aligned} \langle \partial^\mu \tau_{\mu\nu} \rangle_{t_0}^{\not{M}} &\equiv (1 - M\partial_M)\langle \partial^\mu \tau_{\mu\nu} \rangle_{t_0} \\ &= -\tfrac{1}{2} k_\nu \delta(\vec{x})\delta(t - t_0)\left(\ddot{I}_\times A_\times + \ddot{I}_+ A_+\right). \end{aligned} \tag{3.50}$$

We shall call this the *monopole-free* microaverage. This is a local, completely gauge

[15] The reader should not be under the impression that the monopole term is *universally* insignificant. Thus far we have argued its irrelevance only for pulse plane-wave, and as we shall see at the beginning of the next section, this idea follows by linearity to general gravitational waves. However, should the gravitational field have a *time-independent* part, then it is possible for this to couple to the monopole in a physically meaningful way. This is due to the particularly limited gauge freedom available to h_{00} when the field is time-invariant. We will return to this issue in section 3.4.5.

invariant description of the energy-momentum transferred onto test-sources by pulse plane-waves. Furthermore, the right-hand side of (3.50) has an obvious physical interpretation: the coupling between $\ddot{I}_{ij}$ and A_{ij} can be understood, roughly speaking, as the product of a force (responsible for accelerating the constituents of the quadrupole moment) and a distance (actually an expansion/contraction of spacetime) and thus represents the work done on the test-source. For example, consider a test-source composed of two bodies of mass m separated by a light elastic rod of length $2d$ aligned with the x^2-axis; provided the amplitude of the motion of the masses is much smaller than d, then $\ddot{I}_{22} \cong 4dma$, where a is the (outward) acceleration of each mass. Due to the gravitational wave, the proper distance of each mass from the centre of the rod increases by $A_{22}d/2$; thus, counting the motion of both ends of the rod, the total work done on the source, by the wave, is $-2(ma)(A_{22}d/2) = -\ddot{I}_{22}A_{22}/4 = -\ddot{I}_+A_+/2$, which agrees precisely with (3.50). Thus we see that the monopole-free microaverage corresponds to the familiar physical quantities that we would intuitively use to define the energy and momentum of the test-source.

In the next section we will generalise the monopole-free microaverage to arbitrary gravitational fields, and uncover a substantial mathematical shortcut that will greatly simplify this procedure.

3.4.3 Arbitrary Gravitational Fields

Clearly, pulse waves are a special case, and one might expect that for an arbitrary (harmonic gauge) plane-wave

$$h^{\text{wave}}_{\mu\nu} = B_{\mu\nu}(k_\alpha x^\alpha), \qquad k^\mu \bar{B}_{\mu\nu} = 0, \tag{3.51}$$

we would need to perform *finite* averages, rather than microaverages, to remove the gauge dependence of our description; thus, we would be forced to sacrifice our *localised* picture of energy-momentum transfer. However, provided $B_{\mu\nu}(t) \to 0$ as $t \to -\infty$, we can always write

$$\begin{aligned} h^{\text{wave}}_{\mu\nu} &= \int_{-\infty}^{\infty} B_{\mu\nu}(t_0)\delta(k_\alpha x^\alpha - t_0)\mathrm{d}t_0 \\ &= \int_{-\infty}^{\infty} \dot{B}_{\mu\nu}(t_0)H(k_\alpha x^\alpha - t_0)\mathrm{d}t_0 - [B_{\mu\nu}(t_0)H(k_\alpha x^\alpha - t_0)]^{+\infty}_{-\infty} \\ &= \int_{-\infty}^{\infty} \dot{B}_{\mu\nu}(t_0)H(k_\alpha x^\alpha - t_0)\mathrm{d}t_0, \end{aligned} \tag{3.52}$$

and perform the monopole-free microaverage on each component of this sum:

$$\left\langle \partial^\mu \tau_{\mu\nu}[h^{\text{source}}_{\alpha\beta} + h^{\text{wave}}_{\alpha\beta}] \right\rangle^{\not M}_{\int} \equiv \int_{-\infty}^{\infty} \langle \partial^\mu \tau_{\mu\nu}[h^{\text{source}}_{\alpha\beta} + \dot{B}_{\alpha\beta}(t_0)H(k_\sigma x^\sigma - t_0)] \rangle^{\not M}_{t_0} \mathrm{d}t_0. \tag{3.53}$$

The result of this process is

$$\langle \partial^\mu \tau_{\mu\nu} \rangle^{\not M}_{\int} = -\tfrac{1}{2}k_\nu \delta(\vec{x}) \left(\ddot{I}_\times \dot{B}_\times + \ddot{I}_+ \dot{B}_+ \right), \tag{3.54}$$

which renders the interaction completely gauge invariant, and does not sacrifice the local character of our description of energy-momentum transfer.[16]

Although the operation of splitting the wave into a series of pulses and performing a monopole-free microaverage on each pulse may seem too complicated to be useful, the same result can be achieved by a simple alternative method: transform $h_{\mu\nu}^{\text{wave}}$ to transverse-traceless gauge, where the vector u^μ referred to by the TT-conditions (3.29) corresponds to the rest-frame of the test-source. Then, when we calculate $\partial^\mu \tau_{\mu\nu}$, we will automatically recover the monopole-free microaveraged result. To demonstrate this, we recalculate $\partial^\mu \tau_{\mu\nu}$, generalising (3.42) for use with arbitrary plane-waves (3.51),

$$\partial^\mu \tau_{\mu\nu} = -\tfrac{1}{4}k_\nu \Big(\ddot{I}_{ij}\dot{B}_{ij}\delta(\vec{x}) - 2(\dot{I}_{ij} - L_{ij})\partial_j \delta(\vec{x})\dot{B}_{i0} \\ + (2M\delta(\vec{x}) + I_{ij}\partial_i\partial_j\delta(\vec{x}))\dot{B}_{00} \Big)\,, \tag{3.55}$$

and substitute the TT-conditions $B_{0\nu} = B = 0$ (which, along with $k^\mu \bar{B}_{\mu\nu} = 0$, set $B_{1\nu} = 0$ and $B_{22} = -B_{33}$):

$$\begin{aligned} \partial^\mu \tau_{\mu\nu}^{\text{TT}} &\equiv \partial^\mu \tau_{\mu\nu}[h^{\text{source}} + (h^{\text{wave}})^{\text{TT}}] \\ &= -\tfrac{1}{4}k_\nu \delta(\vec{x})\ddot{I}_{ij}\dot{B}_{ij} \\ &= -\tfrac{1}{2}k_\nu \delta(\vec{x}) \left(\ddot{I}_\times \dot{B}_\times + \ddot{I}_+ \dot{B}_+ \right). \end{aligned} \tag{3.56}$$

Hence,

$$\langle \partial^\mu \tau_{\mu\nu} \rangle_{\int}^{\not{M}} = \partial^\mu \tau_{\mu\nu}^{\text{TT}}. \tag{3.57}$$

Furthermore, this equation is not only applicable to incident plane-waves. Because both sides are linear in $h_{\mu\nu}^{\text{wave}}$, equations (3.57) must also hold when $h_{\mu\nu}^{\text{wave}}$ is any *sum* of plane-waves, propagating in arbitrary directions. Locally, we can always express $h_{\mu\nu}^{\text{wave}}$ as a sum of plane-waves (and some time-independent part, which we will ignore until section 3.4.5) so, quite generally, we have

$$\langle \partial^\mu \tau_{\mu\nu} \rangle_{\int}^{\not{M}} = -\tfrac{1}{4}\delta(\vec{x})\ddot{I}_{ij}\partial_\nu h_{ij}^{\text{TT}}, \tag{3.58}$$

where $h_{\mu\nu}^{\text{TT}}$ is the incident gravitational field in transverse-traceless gauge. This equation provides an easy method for calculating the energy-momentum transferred onto the microaveraged test-source due to the presence of arbitrary incident gravitational radiation. Moreover, we see that (ignoring the time-independent field) our gauge invariant probe only exchanges energy-momentum with the transverse-traceless field; the other components of the field do not play a role in this process. We explore the wider significance of this result in the next section.

3.4.4 Energy-Momentum and Transverse-Traceless Gauge

Let us now take a step back from the fine details of the test-source interaction and assess the general picture that is unfolding. As we first saw in section 3.2.3, $\tau_{\mu\nu}$ is not in

[16] Just as we write I_{ij} for $I_{ij}(t)$, we have, in (3.54) and elsewhere, left the argument of $B_{\mu\nu}(t)$ implicit.

general invariant under the gauge freedom that remains after the harmonic condition has been enforced. As a partial remedy of this, the monopole-free microaveraged test-source emerged as a local, gauge invariant probe of gravitational energy-momentum exchange. It has now come to light that only the transverse-traceless field takes part in this process. From this standpoint, a method suggests itself which will remove the remaining ambiguity of $\tau_{\mu\nu}$ in a natural fashion: simply transform the incident gravitational field to transverse-traceless gauge! Consequently, only the degrees of freedom relevant to gauge invariant energy-momentum exchange will contribute to the gravitational energy-momentum tensor. We shall codify this idea as a gauge-fixing programme defined in terms of two "frames", one associated with gravitational detectors, the other with astrophysical sources. In what follows, the gauge-fixing only refers to the *dynamical part* of the gravitational field; as we explain in section 3.4.5, the time-independent part of the gravitational field is essentially gauge invariant and so does not need to be fixed in any way.

Detector-frame. Consider a gravitational detector D in a region $\mathcal{V}_D$ which contains no matter besides the detector. We shall suppose that the *incident field* (due to sources outside $\mathcal{V}_D$) is much larger than the field due to the detector itself; in other words, we model D as a test-source. The detector-frame is then obtained by transforming the *incident field* to TT-gauge, taking u^μ to be the four-velocity of the detector.[17] As a result, the energy-momentum transferred onto D will be exactly equal to the gauge invariant quantities defined by the monopole-free microaverage. What is more, we can imagine adding hypothetical test-sources (co-moving with D) anywhere within $\mathcal{V}_D$ in order to "measure" the gravitational energy-momentum there; because the field has been prepared in this gauge, the result will agree with the gauge invariants we have already defined. In this way, the detector-frame defines $\tau_{\mu\nu}$ through the gauge invariant energy-momentum that would be absorbed by furnishing $\mathcal{V}_D$ with an array of infinitesimal probes moving at the same velocity as the actual detector.

Source-frame. Now consider a compact source S in a region $\mathcal{V}_S$ which, as above, contains no other matter. In contrast to the detector, we shall assume any incident field can be neglected in comparison to the *outgoing field* due to S. The source-frame is obtained by transforming the *outgoing field* to TT-gauge, taking u^μ to be the four-velocity of the source. This gauge transformation can only be achieved by breaking the harmonic condition at S (see Appendix 3.B for details) so it will not be possible to use $\tau_{\mu\nu}$ to describe the energy-momentum lost by the source; however, this self-interaction is ill-defined for a point-like system anyway, and we would have to resolve the source into component parts before such a question could be answered. Outside the source, $h_{\mu\nu}$ will remain harmonic, so $\tau_{\mu\nu}$ will still represent the energy-momentum that could be absorbed by a hypothetical test-source (with the velocity of S) were we to insert one. Much like the detector-frame, we can think of this prescription as measuring $\tau_{\mu\nu}$ by filling $\mathcal{V}_S$ with infinitesimal probes

[17] Note that it is not the total field, but just the incident field, which is made transverse-traceless. We cannot alter the gauge of the field generated by the detector because it is impossible to produce outgoing spherical waves in the gauge field ξ^μ without breaking the harmonic condition at D. This cannot be allowed to happen if $\tau_{\mu\nu}$ is to account for the energy-momentum exchanged with the detector.

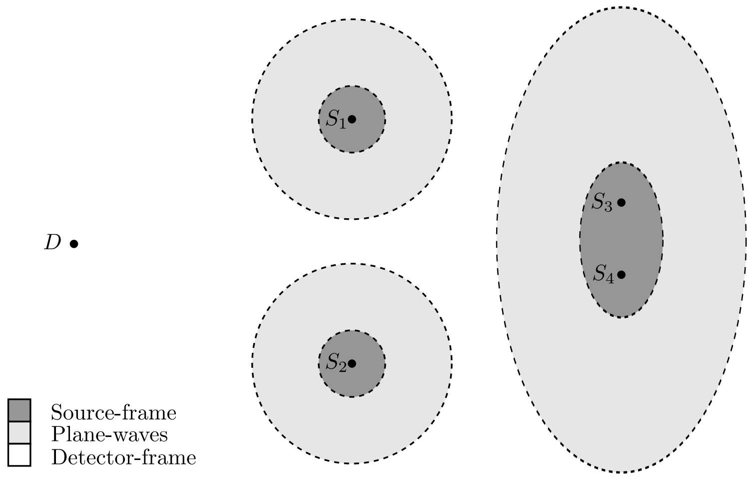

Figure 3.1: Schematic showing plane-wave regions sewing various source-frames to the detector-frame of D. S_1 and S_2 represent compact sources many wavelengths apart; if they have different velocities, then each will determine a separate source-frame. In contrast, S_3 and S_4 are only separated by a small number of wavelengths; as there is no plane-wave region dividing the sources, they must share their source-frame. It is natural to base this joint source-frame on the velocity of the centre of mass of the multi-component system.

that are co-moving with the source.

The gauge-fixing programme is simple: if one wishes to describe the energy-momentum of the gravitational field as it would be measured by some detector D, then adopt the source-frame near distant astrophysical sources, and the detector-frame of D everywhere else. This allows the energy-momentum in the vicinity of a source to be unambiguously determined by the source alone, whilst simultaneously adapting the gravitational field for a description of energy-momentum absorption by the detector. Remarkably, despite using *different* TT-gauges in the various regions, this programme still produces a self-consistent picture of the propagation of energy-momentum from the sources to the detector. This is because, many wavelengths from an isolated source, the gravitational field approximates a plane-wave, and so (as we saw in section 3.3.1) the gravitational energy-momentum is gauge invariant there. Thus, in this regime, the source-frame energy-momentum and the detector-frame energy-momentum *are equal*, regardless of the relative velocity between the source and the detector. The plane-wave regions will therefore "sew together" the different frames and produce a globally consistent description of energy-momentum flowing from source to detector. When the sources are not isolated, but are separated from each other by only a small number of wavelengths, no plane-wave region will exist between the sources. In this case, the sources must be treated as one extended source, with a joint source-frame that identifies u^μ with the four-velocity of the centre of mass of the many-body system. We illustrate the gauge-fixing programme schematically in figure 3.1.

Up to this point, we have justified our insistence on TT-gauge based purely on considerations of energy-momentum exchange with matter. Of course, there is another exceptional property of $\tau_{\mu\nu}$, derived in section 3.3.2, that also holds under these conditions: it is always *positive*. This is a peculiar and surprising result. It is something of a small miracle that transverse-tracelessness guarantees not only agreement with the monopole-free microaverage, but also ensures that $\tau_{\mu\nu}$ represents positive energy-density, and causal energy flux. Under these conditions it will always be possible to make physical sense of the gravitational energy-momentum tensor: we will never have to interpret (or explain) negative or superluminal energy.

We feel that the dual significance of transverse-traceless gauge leaves little doubt that this is the correct procedure by which to remove the final trace of ambiguity in the definition of $\tau_{\mu\nu}$. In section 3.5 we will apply this programme to a small number of examples, including two specific compact sources: a vibrating rod, and an equal-mass binary. First, however, we must address a technical issue regarding the time-independent part of the gravitational field.

3.4.5 Time-Independent Fields

Unlike the dynamical part of the gravitational field, the gauge of the time-independent mode (i.e. the time-averaged field) is completely fixed by insisting that (a) the harmonic condition $\partial^\mu \bar{h}_{\mu\nu} = 0$ holds everywhere, (b) $h_{\mu\nu} \to 0$ as $r \to \infty$, and (c) gauge transformations $\delta h_{\mu\nu} = \partial_{(\mu}\xi_{\nu)}$ are finite everywhere and bounded at infinity. To see this, suppose that our time-independent field $h_{\mu\nu}(\vec{x})$ obeys the harmonic condition and vanishes at spatial infinity. Then the transformed field $h'_{\mu\nu} = h_{\mu\nu} + \partial_{(\mu}\xi_{\nu)}$ will only satisfy the harmonic condition if $\partial^2\xi = 0$, and will only be time-independent if $\ddot{\xi}_\mu = 0$. Thus $\partial_i^2(\partial_{(\mu}\xi_{\nu)}) = 0$, the only bounded solutions of which are constants (by Liouville's theorem). Hence we are forced to take $h'_{\mu\nu} = h_{\mu\nu}$ if the new field is to also vanish at spatial infinity, and we thus conclude the gauge is unique.[18]

This result reveals that we are not required to perform any form of microaverage to remove the gauge dependence of the energy-momentum transfer associated with the time-independent mode of the gravitational field: this mode is already gauge invariant. In truth, this is a rather convenient situation. We could not microaverage a time-independent field even if we needed to, due to the caveat $B_{\mu\nu}(t) \to 0$ as $t \to -\infty$, encountered when deriving (3.53).

As there is no gauge freedom in the time-independent part of the gravitational field, we cannot expect this mode to have $h_{0\mu} = 0$ or $h = 0$ in general. This leaves open the possibility (at least in principal) that close to sources, where the time-independent mode can become comparable in amplitude to the dynamical field, the positivity of $\tau_{\mu\nu}$ may be compromised. However, in section 3.5.3 we will see that, even very close to (but not inside)

[18] This argument relies on our insistence that the harmonic condition be valid *everywhere*. In section 3.5.3 we will make use of the following mathematical trick: by relaxing the harmonic condition at the source itself, we will be able to combine many local TT-gauges to form a gauge in which the dynamical part of the gravitational field (outside the source) is transverse-traceless for all time. As we will see, however, it is impossible to apply this procedure to a time-independent field. Thus there is nothing to gain from weakening the harmonic condition on the time-independent mode, and it is therefore kept unbroken.

a compact source, the time-independent field obeys $\bar{h}_{00} \gg \bar{h}_{0i} \gg \bar{h}_{ij}$.[19] It is easy to show that such a field will not upset the positivity of $\tau_{\mu\nu}$. Neglecting the small quantities, the trace-reversed gravitational field will take the form

$$\bar{h}_{\mu\nu} = h^{\mathrm{TT}}_{\mu\nu} - 4\Phi u_\mu u_\nu, \tag{3.59}$$

where $h^{\mathrm{TT}}_{\mu\nu}$ is the dynamical field in transverse-traceless gauge, and $\Phi \equiv -\bar{h}_{00}/4$ is the the Newtonian potential, the only non-negligible contribution from the time-independent mode. The trace-reversed gravitational energy-momentum tensor therefore takes the form

$$4\kappa\bar{\tau}_{\mu\nu} = \partial_\mu h_{\alpha\beta}\partial_\nu \bar{h}^{\alpha\beta} = \partial_\mu h^{\mathrm{TT}}_{ij}\partial_\nu h^{\mathrm{TT}}_{ij} + 8\partial_\mu\Phi\partial_\nu\Phi, \tag{3.60}$$

which ensures that the positivity proof of section 3.3.2 can proceed almost exactly as before, with Φ effectively behaving as an additional component of h^{TT}_{ij}. Thus, even though it is not transverse-traceless, the time-independent mode does not give rise to any negative or superluminal energy.

3.5 Applications

This section is devoted to calculating the energy-momentum content of the gravitational field in a small number of examples, following the gauge-fixing programme of section 3.4.4.

3.5.1 Plane-Waves

Although we have already studied gravitational plane-waves in a variety of contexts, we have yet to evaluate the energy-momentum they carry. This calculation will serve as a simple first example, and will illustrate the use of the detector-frame.

We begin with an arbitrary (harmonic gauge) plane-wave,

$$h_{\mu\nu} = h_{\mu\nu}(k_\alpha x^\alpha), \quad k^\mu k_\mu = 0, \quad k^\mu \bar{h}'_{\mu\nu} = 0, \tag{3.61}$$

and substitute this field into equation (3.35):

$$\kappa\tau_{\mu\nu} = \tfrac{1}{4}k_\mu k_\nu h'_{\alpha\beta}\bar{h}'^{\alpha\beta}. \tag{3.62}$$

As we ascertained in section 3.3.1, the energy-momentum of plane-waves is gauge invariant. Consequently, we can simplify (3.62) by evaluating $h_{\mu\nu}$ in TT-gauge, thereby removing all components except for h_+ and $h_\times$:

$$\kappa\tau_{\mu\nu} = \tfrac{1}{2}k_\mu k_\nu((h'_+)^2 + (h'_\times)^2). \tag{3.63}$$

[19] These order-of-magnitude inequalities are not limited to the compact source. The time-independent mode of the gravitational field is always generated by the time-averaged energy-momentum tensor of matter $\langle T_{\mu\nu}\rangle$, and it is to be expected that this field will be dominated by the slow (i.e. non-relativistic) motion of matter, so that $\langle T_{00}\rangle \gg \langle T_{0i}\rangle \gg \langle T_{ij}\rangle$. Hence $\bar{h}_{00} \gg \bar{h}_{0i} \gg \bar{h}_{ij}$ will hold quite generally: whenever $\langle T_{\mu\nu}\rangle$ is dominated by non-relativistic motion.

In fact, because $\delta h_+ = \delta h_\times = 0$ under any gauge transformation that keeps $h_{\mu\nu}$ a plane-wave, the right-hand side of this equation is gauge invariant also. Hence, equation (3.63) must hold in any gauge, and all other terms on the right-hand side of (3.62) must cancel in general.[20] Using this formula for $\tau_{\mu\nu}$, every future-directed timelike unit-vector v^μ defines a gravitational energy current-density,

$$v^\mu \tau_{\mu\nu} = v^\mu k_\mu k_\nu ((h'_+)^2 + (h'_\times)^2)/2\kappa, \tag{3.64}$$

which is clearly future-directed and null; unsurprisingly, the energy of a gravitational plane-wave is positive and flows at the speed of light in the direction of propagation.

So far, the gauge invariance of $\tau_{\mu\nu}$ has made gauge-fixing unnecessary. The insistence that we evaluate the $h_{\mu\nu}$ in the detector-frame only becomes important when there are multiple plane-waves propagating in different directions. Suppose, for example, that there are two plane-waves:

$$h_{\mu\nu} = h^{\mathrm{I}}_{\mu\nu}(k^{\mathrm{I}}_\alpha x^\alpha) + h^{\mathrm{II}}_{\mu\nu}(k^{\mathrm{II}}_\alpha x^\alpha), \tag{3.65}$$

where k^{I}_μ and k^{II}_μ are non-parallel null vectors. As $\tau_{\mu\nu}$ is quadratic in $h_{\mu\nu}$, the energy-momentum of the total field takes the form

$$\kappa\tau_{\mu\nu} = \kappa\tau^{\mathrm{I}}_{\mu\nu} + \kappa\tau^{\mathrm{II}}_{\mu\nu} + \tfrac{1}{2} k^{\mathrm{I}}_{(\mu} k^{\mathrm{II}}_{\nu)} h^{\mathrm{I}\prime}_{\alpha\beta} \bar{h}^{\mathrm{II}\prime\alpha\beta} - \tfrac{1}{4}\eta_{\mu\nu} k^{\mathrm{I}}_\sigma k^{\mathrm{II}\sigma} h^{\mathrm{I}\prime}_{\alpha\beta} \bar{h}^{\mathrm{II}\prime\alpha\beta}, \tag{3.66}$$

where $\tau^{\mathrm{I}}_{\mu\nu}$ and $\tau^{\mathrm{II}}_{\mu\nu}$ are the individual energy-momentum tensors of $h^{\mathrm{I}}_{\mu\nu}$ and $h^{\mathrm{II}}_{\mu\nu}$ respectively. Now, any gauge transformation that preserves the form (3.65) of the gravitational field can be thought of as a pair of gauge-transformations that act on $h^{\mathrm{I}}_{\mu\nu}$ and $h^{\mathrm{II}}_{\mu\nu}$ separately, preserving their plane-wave forms; thus $\tau^{\mathrm{I}}_{\mu\nu}$ and $\tau^{\mathrm{II}}_{\mu\nu}$ must be invariant under gauge-transformations of this type. However, the "cross-terms" in (3.66) *are* gauge-dependent, as the (gauge-dependent) longitudinal components of $h^{\mathrm{I}}_{\mu\nu}$ will be transverse to $h^{\mathrm{II}}_{\mu\nu}$, and vice versa. This gauge ambiguity is removed, however, by the presence of a physical detector: once we demand that the energy-momentum exchanged with this detector is to equal the monopole-free microaverage, we fix the gauge completely. This is the detector-frame: $h_{\mu\nu}$ is transverse-traceless, with u^μ identified as the four-velocity of the detector. In this sense, the gauge-fixing programme is the procedure that enables us to "add together" the energy-momentum tensors of gravitational plane-waves (which, individually, are gauge invariant) to form the energy-momentum tensor of the total field.

For the sake of the concreteness, let us set $k^{\mathrm{I}}_\mu = (1,1,0,0)$, $k^{\mathrm{II}}_\mu = (1,0,1,0)$, and $u^\mu = (1,0,0,0)$. Then, once we have transformed $h_{\mu\nu}$ to TT-gauge, the energy-momentum tensor becomes

$$\kappa\tau_{\mu\nu} = \kappa\tau^{\mathrm{I}}_{\mu\nu} + \kappa\tau^{\mathrm{II}}_{\mu\nu} - \tfrac{1}{4}(2k^{\mathrm{I}}_{(\mu}k^{\mathrm{II}}_{\nu)} + \eta_{\mu\nu}) h^{\mathrm{I}\prime}_+ h^{\mathrm{II}\prime}_+,$$

where $h^{\mathrm{I}}_+ = h^{\mathrm{I}}_{22} = -h^{\mathrm{I}}_{33}$, and $h^{\mathrm{II}}_+ = h^{\mathrm{II}}_{33} = -h^{\mathrm{II}}_{11}$. Due to the positivity theorem of section 3.3.2, we already know this tensor describes a positive energy-density, and a causal energy

[20] This can be verified by taking $k_\mu = (1,-1,0,0)$ and using $k^\mu \bar{h}_{\mu\nu} = 0$ in the same form as (3.48).

flux. As a particular example of this, it is easy to calculate the energy-density associated with u^μ,

$$\kappa\tau_{00} = \tfrac{1}{2}\left((h_+^{\mathrm{I}\prime})^2 + (h_\times^{\mathrm{I}\prime})^2 + (h_+^{\mathrm{II}\prime})^2 + (h_\times^{\mathrm{II}\prime})^2\right) - \tfrac{1}{4}h_+^{\mathrm{I}\prime}h_+^{\mathrm{II}\prime},$$

and, as $(h_+^{\mathrm{I}\prime})^2 + (h_+^{\mathrm{II}\prime})^2 \geq h_+^{\mathrm{I}\prime}h_+^{\mathrm{II}\prime}/2$, we can confirm that this energy-density can never be negative.

3.5.2 Linearised Schwarzschild Spacetime

The Schwarzschild spacetime is the vacuum solution to the Einstein field equations outside any uncharged spherical non-rotating body of mass M. At distances much greater than κM, where the linear approximation is valid, the gravitational field must therefore correspond to that of the compact source with $I_{ij} = L_{ij} = 0$:

$$\bar{h}_{00} = \frac{\kappa M}{2\pi r}, \tag{3.67}$$

and $\bar{h}_{0i} = \bar{h}_{ij} = 0$. Obviously, this is an example of a gravitational field that is entirely time-independent; thus, as explained in section 3.4.5, there will be no possibility of transforming to transverse-traceless gauge, nor any need to do so.[21] The formula (3.35) for the gravitational energy-momentum tensor yields

$$\tau_{\mu\nu} = \kappa\left(\frac{M}{8\pi r^2}\right)^2(2\hat{x}_\mu\hat{x}_\nu - \eta_{\mu\nu}), \tag{3.68}$$

where $\hat{x}^\mu$ is the radial unit vector.[22] It is easy to confirm that this energy-momentum tensor is everywhere positive. Any timelike unit vector v^μ defines a positive gravitational energy-density,

$$\varrho \equiv v^\mu\tau_{\mu\nu}v^\nu = \kappa\left(\frac{M}{8\pi r^2}\right)^2\left(2(\hat{x}_i v_i)^2 + 1\right) \geq 0, \tag{3.69}$$

and an energy current-density $J^\nu \equiv v^\mu\tau_{\mu\nu}$ which is nowhere spacelike:

$$J^\nu J_\nu = -\kappa^2\left(\frac{M}{8\pi r^2}\right)^4 \leq 0. \tag{3.70}$$

It is also worth comparing equation (3.68) with the electromagnetic energy-momentum outside a point-charge: $T_{\mu\nu} \sim (g_{\mu\nu} + 2u_\mu u_\nu - 2\hat{x}_\mu\hat{x}_\nu)/r^4$. While both tensors diminish in proportion to $1/r^4$, they define very different stress profiles at each point. The gravitational field has $\tau_{rr} = -\tau_{\theta\theta} = -\tau_{ii} = \tau_{00} \geq 0$ and thus describes radial compression, tangential tension, and negative pressure; while the electromagnetic field has $-T_{rr} = T_{\theta\theta} = T_{ii} = T_{00} \geq 0$ and thus describes radial tension, tangential compression, and positive pressure. The physical significance of this difference is far from obvious, but may relate to the

[21] Of course, the linearised Schwarzschild spacetime can be represented in other gauges, but no others obey $\partial^\mu\bar{h}_{\mu\nu} = 0$ and $\dot{h}_{\mu\nu} = 0$ everywhere, and are well-behaved at infinity.

[22] This is a trivial extension of the notation $\hat{x}_i = x_i/r$ from appendix 3.A; we simply define $\hat{x}_0 = 0$.

like-attracts-like character of gravity: the negative gravitational pressure mediating the attraction of other masses, while positive electromagnetic pressure causes the repulsion of like-charges. In addition, it may be possible to understand the radial gravitational compression (and tangential tension) in terms of some "elastic" analogy for spacetime, as the Schwarzschild geometry "squeezes in" extra radial distance (between spheres of given area) in comparison to flat space. However, the theoretical value of such an analogy is unclear, and we do not intend to develop it any further here.

Although we have focused here on the linearised Schwarzschild spacetime as a particular example of a time-independent field, we note in passing that it is easy to evaluate the gravitational energy-momentum tensor associated with *any* static configuration of matter $T_{\mu\nu} = u_\mu u_\nu \rho(\vec{x})$: equation (3.35) simplifies to

$$\kappa\tau_{\mu\nu} = 2\partial_\mu\Phi\partial_\nu\Phi - \eta_{\mu\nu}\partial_\alpha\Phi\partial^\alpha\Phi, \tag{3.71}$$

where the Newtonian potential $\Phi \equiv -\bar{h}_{00}/4$ is determined by solving Poisson's equation $\partial_i^2\Phi = \kappa\rho/2$. Equation (3.71) reveals that the energy-momentum of the Newtonian potential is exactly that of a massless Klein-Gordon scalar field.

3.5.3 Gravitational Field of a Compact Source

We shall now calculate the energy-momentum content of the gravitational field (3.85) generated by a compact source.[23] The first step will be to enter the source-frame: we must transform the dynamical part of the outgoing field into TT-gauge, with u^μ identified as the four-velocity of the source. We can always make this transformation *locally* by choosing the gauge fields ξ_μ such that $h_{0\mu} = \dot{h}_{0\mu} = 0$ and $h = \dot{h} = 0$ at some time $t = t_0$; then $\partial^2\xi_\mu = 0$ (which preserves the harmonic condition) and the field equations $\partial^2 h_{\mu\nu} = 0$ (outside the source) ensure that $h_{0\mu} = 0$ and $h = 0$ continues to be true for $t \in (t_0 - r, t_0 + r)$.[24] This method is problematic in that it is based around an arbitrary special time t_0, and that transverse-tracelessness always breaks down within a time $\Delta t = 2r$; these issues prevent us from forming a *global* picture of the energy-momentum outside the source.

As we show in appendix 3.B, these problems can be completely avoided if we weaken the harmonic condition slightly, so that $\partial^\mu\bar{h}_{\mu\nu} = 0$ is only enforced *outside the source*. This trick allows us to find a gauge in which the dynamical field is transverse-traceless everywhere outside the source, for all t, and does not require us to choose a special time t_0. We can think of this gauge as a way of joining up the many possible local gauges (defined using the aforementioned method) in a mutually consistent fashion.[25] The process of

[23] This calculation should not be confused with the analysis performed in section 3.4, where a test-source (essentially a compact source in the limit $d, M, L_{ij}, I_{ij} \to 0$) interacted with an incident field, which presumably had been generated by another source, very far way. Here the compact source will represent an astrophysical source (with finite d, M, L_{ij} and I_{ij}) and by adopting the source-frame we will compute the energy-momentum of the outgoing field as it would be measured by microaveraged detectors co-moving with the source.

[24] See [79, §4.4] for details.

[25] Presumably, there is some topological obstruction which prevents us from joining these local gauges

transforming the gravitational field of the compact-source (3.85) can be found in appendix 3.B, here we simply display the result:

$$\begin{aligned}
\bar{h}_{00} &= \left(2M + \langle I_{ij}\rangle \partial_i \partial_j\right) \frac{\kappa}{4\pi r} \\
\bar{h}_{0i} &= -L_{ij}\partial_j \frac{\kappa}{4\pi r} \\
\bar{h}_{ij} &= \int_{-\infty}^{\infty} \frac{\mathrm{d}\omega}{2\pi} e^{i\omega t} \Bigg[\left(\tilde{I}_{kl}\delta_{ij} + \tilde{I}\delta_{ik}\delta_{jl} - 4\tilde{I}_{k(i}\delta_{j)l}\right)\partial_k\partial_l \\
&\qquad\qquad - \frac{1}{\omega^2}\tilde{I}_{kl}\partial_k\partial_l\partial_i\partial_j + \omega^2\left(\tilde{I}\delta_{ij} - 2\tilde{I}_{ij}\right)\Bigg] \frac{\kappa e^{-i\omega r}}{8\pi r},
\end{aligned} \tag{3.72}$$

where we have introduced the notation

$$\langle I_{ij}\rangle \equiv \lim_{\Delta\to\infty} \int_{-\infty}^{\infty} I_{ij}(t) \frac{e^{-t^2/\Delta^2}}{\sqrt{\pi\Delta^2}} \mathrm{d}t, \tag{3.73}$$

for the time-average of the quadrupole moment, and

$$\tilde{I}_{ij}(\omega) \equiv \int_{-\infty}^{\infty} e^{-i\omega t}\left(I_{ij}(t) - \langle I_{ij}\rangle\right) \mathrm{d}t, \tag{3.74}$$

for the Fourier transform of its dynamical part. Notice that the terms proportional to M, L_{ij}, and $\langle I_{ij}\rangle$ constitute the time-independent mode of the field, and have therefore not been transformed. At this point we can confirm the assertion of section 3.4.5, that the time independent field satisfies $\bar{h}_{00} \gg \bar{h}_{0i} \gg \bar{h}_{ij}$. To do so we note that, firstly, there is no time-independent term in $\bar{h}_{ij}$, and secondly, seeing as the radius of the source $d \gtrsim L/M$, and that we are outside the source (which is to say, $r \gg d$, the regime of validity of (3.72)) then we must have $L/r \ll M$.

Having rendered the dynamical field transverse-traceless outside the source, all that remains is to substitute (3.72) into (3.35) to calculate $\tau_{\mu\nu}$. As was shown in the process of deriving (3.60), the energy-momentum of the time-independent field adds *linearly* (i.e. without cross-terms) to that of the dynamical field. Given that we have already investigated the part due to the time-independent field in section 3.5.2, it is generally more interesting to discard this term, and focus on the additional energy-momentum due to the dynamical field. In figure 3.2 we show the results of a computation of this additional gravitational energy-density τ_{00} outside two monochromatic compact sources: a vibrating rod, and an equal-mass binary. It goes without saying that the energy-density is everywhere positive, and that the energy current-density is nowhere spacelike.

without violating the harmonic condition at the source. However, provided we do not intend to calculate the energy-momentum transferred between matter and gravity at the source, this is not an issue. Even if we were careful to keep the gravitational field harmonic at the source, more work would be needed to perform such a calculation, as this self-interaction only becomes well-defined by breaking down the compact source into component parts.

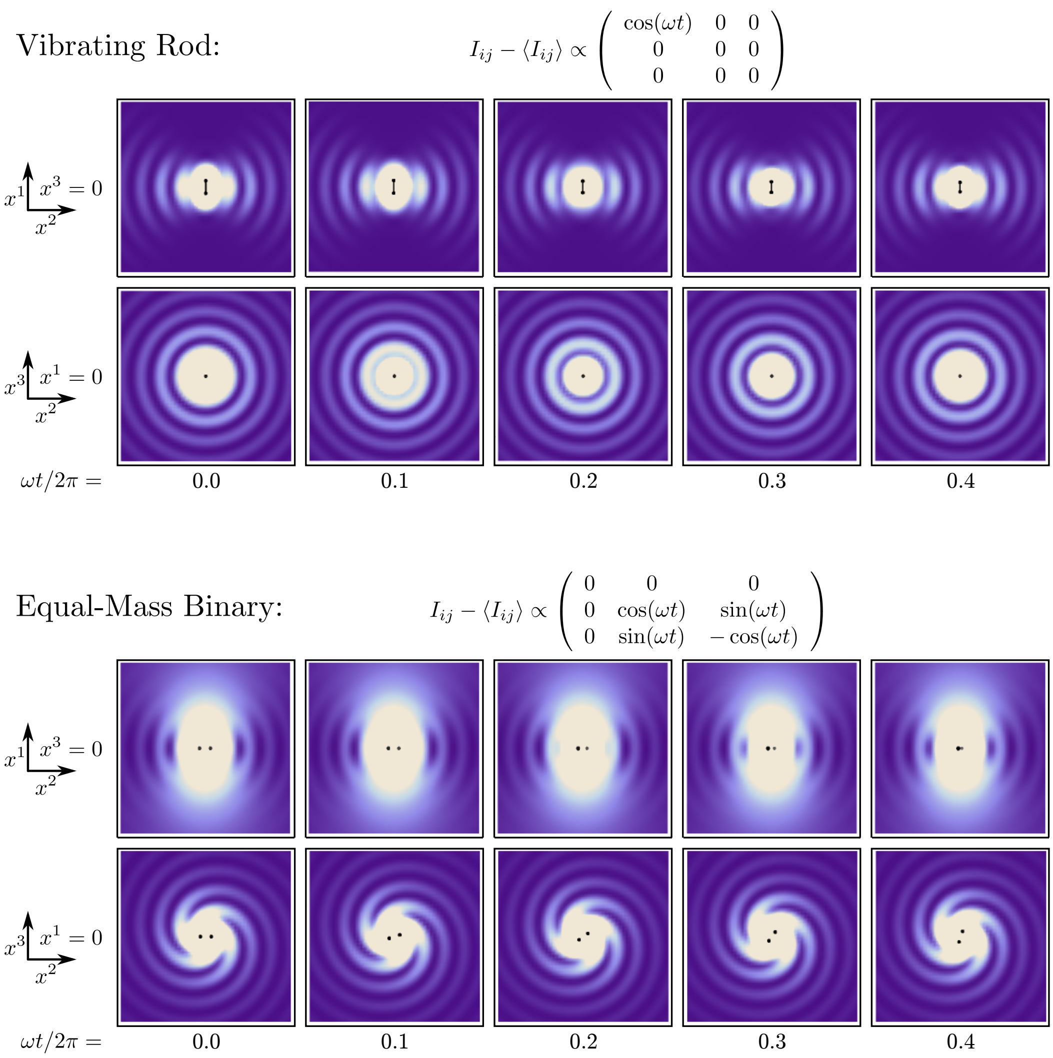

Figure 3.2: Plots of the energy-density of the dynamical gravitational field outside two monochromatic compact sources: a vibrating rod, and an equal-mass binary. Only half a period is shown, as $\tau_{\mu\nu}$ oscillates with twice the frequency of the source. Although the rod and the binary are much smaller than one wavelength, they have been magnified to illustrate the phase of their motion. The propagation of gravitational energy is more easily appreciated in the animated versions of these plots, available at `www.mrao.cam.ac.uk/~lmb62/animations`.

3.6 Conclusion

It is natural to suspect that wherever matter gains energy under the influence of gravity, a corresponding loss in the energy of the gravitational field must have occurred. By constructing a framework to quantify this idea, we have succeeded in localising the energy and momentum of the linear gravitational field, and have shown this energy to be positive and to not flow faster than light.

The core result of our investigation is the formula (3.23) for the gravitational energy-momentum tensor, the *unique* symmetric tensor, quadratic in $\check{\nabla}_c h_{ab}$, which accounts for the energy-momentum lost or gained by matter through its interaction with gravity (3.10). Crucially, a tensor satisfying these conditions only exists in the harmonic gauge (3.24) and thus, as a matter of necessity rather than choice, our framework discards nearly all its gauge freedom. A small set of viable gauge transformations still remain, however, and although these do not alter the energy-momentum of gravitational plane-waves (§3.3.1) this invariance does not extend to arbitrary gravitational fields.

In response to this ambiguity, the monopole-free microaverage was developed (§3.4.2); this constitutes a local and fully gauge-invariant description of energy-momentum transfer, and agrees with the intuitive notion that the "work done" on a gravitational detector is the product of the force (proper acceleration) and the proper distance through which the force is applied. Of the incident field, only the transverse-traceless part contributes to the microaveraged exchange (3.57), and thus a natural gauge-fixing programme is motivated, based around transverse-traceless gauge (§3.4.4). The effect of this programme is to prepare the field so that no microaverage is needed, and furthermore, to ensure that energy-momentum is only assigned to those components of the field whose energy-momentum can be measured by a microaveraged detector. Because the positivity property (§3.3.2) holds true wherever the field is transverse-traceless, the gauge-fixing procedure also ensures that (for the dynamical field at least) gravitational energy-density is positive, and gravitational energy flux is timelike or null. No longer burdened by gauge ambiguity, the gravitational energy-momentum tensor can be evaluated without difficulty: the energy-momentum of gravitational plane-waves (3.63), the linearised Schwarzschild spacetime (3.68), and the gravitational radiation outside compact sources (Fig. 3.2) have been provided as specific examples.

With regards to further investigation, there are two obvious directions in which our framework might be extended: beyond the linear approximation, and beyond the flat background.[26] However, it is currently unknown whether such extensions are possible, or even conceptually sound. On a more practical level, one could apply our formalism to

[26]To extend τ_{ab} beyond the linear regime, one would hope to construct a tensor t_{ab}, defined on the physical spacetime $\mathcal{M}$, such that $\phi^* t_{ab} = \tau_{ab} + O(h^3)$. Clearly, it will only be possible to make this identification if τ_{ab} is gauge invariant to second order, as $\delta(\phi^* t_{ab}) \sim O(h^2)\partial\xi$ under a change of gauge, whereas $\delta\tau_{ab} \sim O(h)\partial\xi$ unless it is invariant. Thus, only once τ_{ab} has been brought into the detector-frame, or the source-frame, can we proceed. Consequently, we should expect that t_{ab} will not only depend on the physical metric g_{ab}, but also on the four-velocity of the relevant detector or source.

the energetics of actual gravitational detectors, rather than the idealised test-sources that have so far dominated our discussion. In doing so, the framework developed here may benefit the design and analysis of future gravitational-wave experiments.

3.A Appendix: Sources

The aim of this Appendix is to derive the formula for $T_{\mu\nu}$ that defines a gravitational point-source (essentially an infinitesimal gravitational quadrupole) and the field $h_{\mu\nu}$ that it generates. The derivation comprises two parts: first, a calculation of the field due to a compact source; second, a calculation of the field due to a candidate $T_{\mu\nu}$ that vanishes everywhere but at $\vec{x} = 0$. As the field from the first calculation matches that of the second (within the region of validity of the compact source approximation) we will be able to conclude that our candidate $T_{\mu\nu}$ is indeed the energy-momentum tensor we sought, that of an infinitesimal compact source.

3.A.1 The Compact Source

A *compact source* is an isolated gravitational body confined to a compact spatial region $\mathcal{D}$ of radius d much smaller than the wavelength λ of the gravitational radiation it emits. Although calculations of the field $h_{\mu\nu}(\vec{x}, t)$ outside a compact source are available in many standard references, we present our own here for two reasons. Firstly, textbook treatments commonly conflate the slow-motion approximation ($d \ll \lambda$) with the far-field approximation ($|\vec{x}| \equiv r \gg \lambda$). Here we shall assume only that the source is very small ($d \ll r, \lambda$) but not anything about the ratio of λ to r.[27] Secondly, the standard approaches frequently omit a full calculation of $\bar{h}_{00}$ and $\bar{h}_{0i}$. Presumably, these components are ignored because they do not appear to contribute to the gravitational field in transverse-traceless gauge; however, they must be included if $h_{\mu\nu}$ is to satisfy the harmonic condition.

The retarded solution to the linearised field equations (3.37) is given by

$$\bar{h}_{\mu\nu}(\vec{x}, t) = \frac{\kappa}{2\pi} \int_{\mathcal{D}} \frac{T_{\mu\nu}(\vec{x}', t - |\vec{x} - \vec{x}'|)}{|\vec{x} - \vec{x}'|} \mathrm{d}^3 x'. \tag{3.75}$$

We shall proceed by expanding the right-hand side of this equation to second order in the small quantities d/λ and d/r, so that we have an integral of energy-momentum tensors $T_{\mu\nu} \equiv T_{\mu\nu}(\vec{x}', t - r)$ evaluated at the *same time* $t' = t - r$. Using

$$|\vec{x} - \vec{x}'| = r \left(1 - \frac{\vec{x} \cdot \vec{x}'}{r^2} + \frac{|\vec{x}'|^2}{2r^2} - \frac{(\vec{x} \cdot \vec{x}')^2}{2r^4} + O((d/r)^3) \right), \tag{3.76}$$

[27] As we are working with in the confines of linearised gravity, we should also insist that $d \gg 2\kappa M$, the Schwarzschild radius of the source. However, this will have little bearing on our calculation.

equation (3.75) expands to

$$\begin{aligned}\bar{h}_{\mu\nu} = \frac{\kappa}{2\pi r}\int_{\mathcal{D}} \mathrm{d}^3x' \Bigg[T_{\mu\nu}\left(1 + \frac{\vec{x}\cdot\vec{x}'}{r^2} - \frac{|\vec{x}'|^2}{2r^2} + \frac{3(\vec{x}\cdot\vec{x}')^2}{2r^4}\right) \\ + r\dot{T}_{\mu\nu}\left(\frac{\vec{x}\cdot\vec{x}'}{r^2} - \frac{|\vec{x}'|^2}{2r^2} + \frac{3(\vec{x}\cdot\vec{x}')^2}{2r^4}\right) \\ + r^2\ddot{T}_{\mu\nu}\frac{(\vec{x}\cdot\vec{x}')^2}{2r^4} + O((d/r)^3)\Bigg]. \end{aligned} \tag{3.77}$$

Although we have not written their arguments, it should be understood that the $T_{\mu\nu}$ terms in the integral are evaluated at $(\vec{x}', t-r)$, while $\bar{h}_{\mu\nu}$ is evaluated at $(\vec{x}, t)$.

In order to relate this integral to the basic physical properties of the source, we define its mass, momentum, and moment-of-energy by

$$M \equiv \int_{\mathcal{D}} T_{00}\mathrm{d}^3x', \qquad P_i \equiv -\int_{\mathcal{D}} T_{0i}\mathrm{d}^3x', \qquad X_i \equiv \int_{\mathcal{D}} T_{00}x'_i\mathrm{d}^3x', \tag{3.78}$$

respectively. Notice that, because the source is entirely contained within $\mathcal{D}$ (so $T_{\mu\nu} = 0$ on the boundary $\partial\mathcal{D}$) the conservation equation $\partial^\mu T_{\mu\nu} = 0$ (the linearised version of (3.7)) leads to the following relations:

$$\dot{X}_i = \int_{\mathcal{D}} \partial_0 T_{00}x'_i\mathrm{d}^3x' = \int_{\mathcal{D}} (\partial'_j T_{j0})x'_i\mathrm{d}^3x' = -\int_{\mathcal{D}} T_{j0}(\partial'_j x'_i)\mathrm{d}^3x' = P_i, \tag{3.79}$$

$$\dot{P}_i = -\int_{\mathcal{D}} \partial'_j T_{ji}\mathrm{d}^3x' = 0. \tag{3.80}$$

Thus $\ddot{X}_i = 0$, and we are free to fix $X_i = P_i = 0$ by our choice of coordinate system. Note also that $\dot{M} = 0$ follows by an identical argument. Next we define the quadrupole moment

$$I_{ij} \equiv \int_{\mathcal{D}} T_{00}x'_i x'_j\mathrm{d}^3x', \tag{3.81}$$

and then derive

$$\dot{I}_{ij} = -2\int_{\mathcal{D}} T_{0(j}x'_{i)}\mathrm{d}^3x', \tag{3.82}$$

$$\ddot{I}_{ij} = 2\int_{\mathcal{D}} T_{ij}\mathrm{d}^3x', \tag{3.83}$$

in a similar fashion. Finally we define the angular momentum of the source

$$L_{ij} \equiv -2\int_{\mathcal{D}} T_{0[j}x'_{i]}\mathrm{d}^3x', \tag{3.84}$$

and note that conservation sets $\dot{L}_{ij} = 0$.

Before substituting these definitions and results into (3.77), note that equations (3.82) and (3.83) indicate that $\int T_{0i}\mathrm{d}^3x' \sim \dot{I}/d \sim Md/\lambda$ and $\int T_{ij}\mathrm{d}^3x' \sim \ddot{I} \sim Md^2/\lambda^2$; hence the integrals of T_{0j} and T_{ij} already have (respectively) one and two extra factors of (d/λ) than the integrals of T_{00}. Thus, to second order, $\bar{h}_{ij}$ will include contributions from only the zeroth order quantities multiplying T_{ij} in (3.77), and $\bar{h}_{0i}$ will include only first and zeroth order quantities multiplying T_{0i}. The final result, accurate to second order in the small quantities (d/λ) and (d/r), is therefore

$$\begin{aligned}
\bar{h}_{00} &= \frac{\kappa}{4\pi}\left(\frac{2M + \ddot{I}_{ij}\hat{x}_i\hat{x}_j}{r} + \frac{3\dot{I}_{ij}\hat{x}_i\hat{x}_j - \dot{I}}{r^2} + \frac{3I_{ij}\hat{x}_i\hat{x}_j - I}{r^3}\right), \\
\bar{h}_{0i} &= -\frac{\kappa}{4\pi}\left(\frac{\ddot{I}_{ij}\hat{x}_j}{r} + \frac{\dot{I}_{ij}\hat{x}_j}{r^2} - \frac{L_{ij}\hat{x}_j}{r^2}\right), \\
\bar{h}_{ij} &= \frac{\kappa\ddot{I}_{ij}}{4\pi r},
\end{aligned} \tag{3.85}$$

where $\hat{x}_i = x_i/r$ is the radial unit vector, and all the I_{ij} terms are evaluated at the retarded time $t' = t - r$. Note that, while the fields $\bar{h}_{00}$ and $\bar{h}_{0i}$ are often omitted from standard calculations, even in the far-field limit $(r \to \infty)$, they still contain terms of equal size to $\bar{h}_{ij}$; these are necessary for consistency with the harmonic condition.

We have successfully derived the form of the gravitational field outside a compact source. However, because (3.85) was constructed under the approximation scheme $d \ll r$, we can only trust these equations at distances much larger than the size of the source. However, we can still ask the following question: what source would produce a field such that (3.85) was valid for all r, no matter how small? This is the *point-source* we have been interested in: the limit of the compact source as $d \to 0$. In the next section we present a candidate for the point-source, calculate its gravitational field, and show that this agrees with (3.85) for all r.

3.A.2 The Point-Source

Consider the following energy-momentum tensor for matter:

$$\begin{aligned}
T_{00} &= M\delta(\vec{x}) + \tfrac{1}{2}I_{ij}\partial_i\partial_j\delta(\vec{x}), \\
T_{0i} &= \tfrac{1}{2}(\dot{I}_{ij} - L_{ij})\partial_j\delta(\vec{x}), \\
T_{ij} &= \tfrac{1}{2}\ddot{I}_{ij}\delta(\vec{x}),
\end{aligned} \tag{3.86}$$

where M, $L_{ij} = L_{[ij]}$ are constants, $I_{ij} = I_{(ij)}(t)$ is independent of $\vec{x}$, and overdots indicate differentiation with respect to t. It is easy to check that this distribution obeys $\partial^\mu T_{\mu\nu} = 0$.

We wish to solve the linearised field equations

$$\partial^2\bar{h}_{\mu\nu} = -2\kappa T_{\mu\nu}, \tag{3.87}$$

looking for the retarded solution. Recalling that

$$\partial^2\left(f(t-r)/r\right) = -4\pi\delta(\vec{x})f(t), \tag{3.88}$$

for any twice differentiable function $f(t)$, we see that we can replace $f \to \kappa\ddot{I}_{ij}/4\pi$ to generate the result

$$\bar{h}_{ij} = \frac{\kappa\ddot{I}_{ij}(t-r)}{4\pi r}. \tag{3.89}$$

Also, from equation (3.88), we have

$$\partial^2(\partial_j(f(t-r)/r)) = -4\pi f(t)\partial_j\delta(\vec{x}). \tag{3.90}$$

Thus, setting $f \to \kappa(\dot{I}_{ij} - L_{ij})/4\pi$ gives

$$\bar{h}_{0i} = \frac{\kappa}{4\pi}\partial_j\left(\frac{\dot{I}_{ij}(t-r) - L_{ij}}{r}\right) = -\frac{\kappa}{4\pi}\left(\frac{\ddot{I}_{ij}\hat{x}_j}{r} + \frac{\dot{I}_{ij}\hat{x}_j}{r^2} - \frac{L_{ij}\hat{x}_j}{r^2}\right). \tag{3.91}$$

By the same method,

$$\begin{aligned}\bar{h}_{00} &= \frac{\kappa M}{2\pi r} + \frac{\kappa}{4\pi}\partial_i\partial_j\left(\frac{I_{ij}(t-r)}{r}\right)\\ &= \frac{\kappa}{4\pi}\left(\frac{2M + \ddot{I}_{ij}\hat{x}_i\hat{x}_j}{r} + \frac{3\dot{I}_{ij}\hat{x}_i\hat{x}_j - \dot{I}}{r^2} + \frac{3I_{ij}\hat{x}_i\hat{x}_j - I}{r^3}\right).\end{aligned} \tag{3.92}$$

Therefore the source (3.86) generates a gravitational field identical to that of the compact source (3.85), except that these equations are now valid for all $\vec{x}$ (except, possibly, $\vec{x} = 0$) not just $r \gg d$. The energy-momentum tensor (3.86) is the point-source we required and (3.85) the field it generates; the correspondence with the compact source allows us to validate the interpretation of M as the mass, I_{ij} the quadrupole moment, and L_{ij} the angular momentum of the source.

3.B Appendix: Persistent Transverse-Traceless Gauge

Here we describe a method by which the dynamical part of the gravitational field outside a compact source (centred at $\vec{x} = 0$) may be transformed to a gauge which remains transverse-traceless for all time, everywhere outside the source. This will be achieved by relaxing the harmonic condition slightly, so that $\partial^\mu \bar{h}_{\mu\nu} = 0$ only holds outside the source.

First, a point of notation. The gauge transformation described in this section is only applicable to the dynamical part of the gravitational field $h^{\text{dyn}}_{\mu\nu} \equiv h_{\mu\nu} - \langle h_{\mu\nu}\rangle$, where $\langle \ldots \rangle$ signifies a time average. Rather than crowd the notation, it will be convenient to assume that $\langle h_{\mu\nu}\rangle = 0$, and use $h_{\mu\nu}$ to stand for $h^{\text{dyn}}_{\mu\nu}$. For the compact source, this amounts to setting $M = L_{ij} = \langle I_{ij}\rangle = 0$ in (3.85). At the end of the calculation we will reinsert these time-independent terms to the transformed field without alteration.

The general procedure is as follows. To begin, take the Fourier transform of the dynamical part of the gravitational field:

$$\tilde{h}_{\mu\nu}(\omega, \vec{x}) \equiv \int_{-\infty}^{\infty} e^{-i\omega t} h_{\mu\nu}(t, \vec{x})\mathrm{d}t. \tag{3.93}$$

The Fourier transform renders the field equations as

$$(\omega^2 + \partial_i^2)\tilde{h}_{\mu\nu}(\omega, \vec{x}) = 0, \tag{3.94}$$

everywhere outside the source, i.e. for $\vec{x} \neq 0$. The harmonic condition becomes

$$-i\omega\tilde{\bar{h}}_{0\mu} + \partial_i\tilde{\bar{h}}_{i\mu} = 0, \tag{3.95}$$

and writing $\tilde{\xi}_\mu(\omega, \vec{x})$ for the Fourier transform of $\xi_\mu(t, \vec{x})$, the general gauge transformation $\delta h_{\mu\nu} = \partial_{(\mu}\xi_{\nu)}$ takes the form

$$\begin{aligned} \delta\tilde{h}_{00} &= i\omega\tilde{\xi}_0, \\ \delta\tilde{h}_{0i} &= \tfrac{1}{2}i\omega\tilde{\xi}_i + \tfrac{1}{2}\partial_i\tilde{\xi}_0, \\ \delta\tilde{h}_{ij} &= \partial_{(i}\tilde{\xi}_{j)}. \end{aligned} \tag{3.96}$$

To achieve transverse-tracelessness we set

$$\tilde{\xi}_0 = i\omega^{-1}\tilde{h}_{00}, \qquad \tilde{\xi}_i = 2i\omega^{-1}\tilde{h}_{0i} - \omega^{-2}\partial_i\tilde{h}_{00}. \tag{3.97}$$

From the field equations (3.94) it is clear that this gauge transformation obeys $(\omega^2 + \partial_i^2)\tilde{\xi}_\mu = 0$ for $\vec{x} \neq 0$, and thus the harmonic condition is preserved outside the source. It is also easy to check that (3.97) fixes $\delta\tilde{h}_{00} = -\tilde{h}_{00}$ and $\delta\tilde{h}_{0i} = -\tilde{h}_{0i}$, and hence ensures that the transformed field $h'_{\mu\nu} = h_{\mu\nu} + \delta h_{\mu\nu}$ has $h'_{0\mu} = 0$ everywhere. Furthermore,

$$\begin{aligned} \delta\tilde{h} &= -\delta\tilde{h}_{00} + \delta\tilde{h}_{ii} \\ &= \tilde{h}_{00} + \partial_i(2i\omega^{-1}\tilde{h}_{0i} - \omega^{-2}\partial_i\tilde{h}_{00}) \\ &= -\omega^{-2}\partial_i^2\tilde{h}_{00} - \tilde{h}_{ii}, \end{aligned} \tag{3.98}$$

where, in the last step, we have used the $\mu = 0$ component of (3.95). Thus, for $\vec{x} \neq 0$, where we may use (3.94), we have

$$\delta\tilde{h} = \tilde{h}_{00} - \tilde{h}_{ii} = -\tilde{h}, \tag{3.99}$$

so that $h' = 0$ outside of the source. In summary, the transformed field is

$$h'_{ij} = \int_{-\infty}^{\infty} \frac{\mathrm{d}\omega}{2\pi} e^{i\omega t} \left(\tilde{h}_{ij} + \frac{2i}{\omega}\partial_{(i}\tilde{h}_{j)0} - \frac{1}{\omega^2}\partial_i\partial_j\tilde{h}_{00} \right), \tag{3.100}$$

with all other components zero, and $h' = 0$, $\partial^\mu \bar{h}'_{\mu\nu} = 0$ everywhere outside the source.[28]

We are now in a position to apply this procedure to the gravitational field of the compact source (3.85). Before doing so, however, it is worth mentioning that the technique just described is not limited to compact sources. In generalising, the only adjustment needed is that (3.94) will only hold at $\vec{x}$ such that $T_{\mu\nu}(t, \vec{x}) = 0$ for all t. Figure 3.3 illustrates the difference between this technique and the standard method mentioned in section 3.5.3.

[28]It should now be clear why this method cannot be applied to the time-independent mode of the field: ill-defined contributions proportional to $\delta(\omega)/\omega$ or $\delta(\omega)/\omega^2$ would appear in the integral on the right-hand side of (3.100). Even without a delta-function at $\omega = 0$, this integral is not unambiguous until we explain how to deform the contour to avoid the poles there. We suggest the contour should dodge into the lower half of the complex plane, as this ensures that $h'_{ij}(t_1)$ is dependent only on $h_{\mu\nu}(t_2)$ for $t_2 \leq t_1$, which is to say, the transformed field does not depend on future values of the untransformed field. Using this "causal" contour, we can substitute (3.93) into (3.100) and perform the ω integral, arriving at $h'_{ij}(t, \vec{x}) = h_{ij}(t, \vec{x}) + \int_{-\infty}^{t} \mathrm{d}t'((t - t')\partial_i\partial_j h_{00}(t', \vec{x}) - 2\partial_{(i}h_{j)0}(t', \vec{x}))$. In general, this formula is less useful than (3.100), however it does reveal the asymptotic conditions that the dynamical field must obey for this gauge-transformation to be well-defined: as $t \to -\infty$, the non-oscillatory modes of $\partial_i\partial_j h_{00}$ and $\partial_{(i}h_{j)0}$ must vanish faster than t^{-2} and t^{-1} respectively.

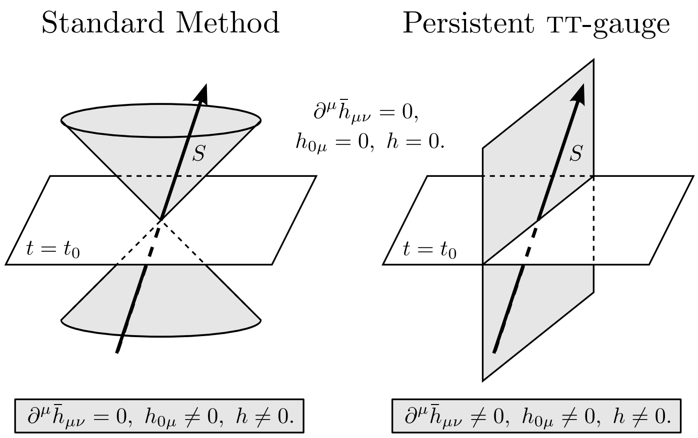

Figure 3.3: Comparison of the standard method for achieving transverse-traceless gauge in the vicinity of a source [79, §4.4], and the "persistent" method described here. In the two diagrams, S represents an arbitrary source (a region with $T_{\mu\nu} \neq 0$) moving relative to u^μ. The hypersurface $t = t_0$ used to define the gauge in the standard method is also shown, but plays no role in our method.

Continuing with the compact source, we write the dynamical part of (3.85) as

$$\begin{aligned} \bar{h}_{00} &= \partial_i \partial_j (\kappa I_{ij}(t-r)/4\pi r), \\ \bar{h}_{0i} &= \partial_j (\kappa \dot{I}_{ij}(t-r)/4\pi r), \\ \bar{h}_{ij} &= \kappa \ddot{I}_{ij}(t-r)/4\pi r, \end{aligned} \tag{3.101}$$

and take the Fourier transform:

$$\begin{aligned} \tilde{\bar{h}}_{00} &= \tilde{I}_{ij} \partial_i \partial_j (\kappa e^{-i\omega r}/4\pi r), \\ \tilde{\bar{h}}_{0i} &= i\omega \tilde{I}_{ij} \partial_j (\kappa e^{-i\omega r}/4\pi r), \\ \tilde{\bar{h}}_{ij} &= -\omega^2 \tilde{I}_{ij} \kappa e^{-i\omega r}/4\pi r, \end{aligned} \tag{3.102}$$

where $\tilde{I}_{ij}$ is the Fourier transform of the dynamical part of the quadrupole moment. Substituting this into (3.100) yields

$$\begin{aligned} h'_{ij} = \int_{-\infty}^{\infty} \frac{\mathrm{d}\omega}{2\pi} e^{i\omega t} \bigg[&\left(\tilde{I}_{kl} \delta_{ij} + \tilde{I} \delta_{ik} \delta_{jl} - 4 \tilde{I}_{k(i} \delta_{j)l} \right) \partial_k \partial_l \\ &- \frac{1}{\omega^2} \tilde{I}_{kl} \partial_k \partial_l \partial_i \partial_j + \omega^2 \left(\tilde{I} \delta_{ij} - 2 \tilde{I}_{ij} \right) \bigg] \frac{\kappa e^{-i\omega r}}{8\pi r}. \end{aligned} \tag{3.103}$$

Finally we recall that $h'_{\mu\nu} = \bar{h}'_{\mu\nu}$ (for $\vec{x} \neq 0$) and reinsert the time-independent mode

$$\begin{aligned} \langle \bar{h}_{00} \rangle &= (2M + \langle I_{ij} \rangle \partial_i \partial_j) \frac{\kappa}{4\pi r}, \\ \langle \bar{h}_{0i} \rangle &= -L_{ij} \partial_j \frac{\kappa}{4\pi r}, \\ \langle \bar{h}_{ij} \rangle &= 0, \end{aligned} \tag{3.104}$$

to confirm equation (3.72).

Chapter 4

Localising the Angular Momentum of Linear Gravity

4.1 Introduction

In the previous chapter, we developed a local description of energy and momentum in linear gravity, deriving a gravitational energy-momentum tensor τ_{ab} that describes positive energy-density and causal energy-flux. The purpose of this present chapter is to complete our picture of local linear gravitational energetics, extending our framework to quantify the *angular momentum* carried by the field. This approach will localise both the "orbital" angular momentum and the intrinsic spin of linear gravity, the former in terms of τ_{ab}, and the latter in terms of a gravitational spin tensor $s^a{}_{bc}$. Not only is this spin tensor vital if one is to account for the angular momentum possessed by gravity and exchanged locally with matter, the formula we derive for it will display a number physically desirable algebraic properties, closely analogous to those of τ_{ab}.

Armed with a local description of the energy, momentum, and angular momentum of linear gravity, we will be ready to tackle the task of chapter 5: to understand τ_{ab} and $s^a{}_{bc}$ in terms of the familiar theoretical apparatus that has been used to define gravitational energy-momentum in the past [7, 35, 52], and energy-momentum in general [15, 44, 61]. These developments will crystallise the tensors' physical interpretation, deepen our understanding of their theoretical underpinnings, and suggest a route by which our work might be generalised beyond the linear approximation.

Let us begin by summarising the key points of the programme developed in chapter 3.[1] We define the gravitational field h_{ab} on a *flat* background spacetime $(\check{\mathcal{M}}, \check{g}_{ab})$ by a diffeomorphism $\phi : \mathcal{M} \to \check{\mathcal{M}}$ that maps the physical spacetime $(\mathcal{M}, g_{ab})$ onto the

[1] As before, we work in units where $c = 1$, write $\kappa \equiv 8\pi G$, and use the sign conventions of Wald [79]: $\eta_{\mu\nu} \equiv \mathrm{diag}(-1, 1, 1, 1)$, $[\nabla_c, \nabla_d]v^a \equiv 2\nabla_{[c}\nabla_{d]}v^a \equiv R^a{}_{bcd}v^b$, and $R_{ab} \equiv R^c{}_{acb}$. We use Roman letters (except i, j, k, l) as abstract tensor indices [79, §2.4] and Greek letters as numerical indices running from 0 to 3. The indices i, j, k, l are reserved for spatial components, and run from 1 to 3.

background:[2]

$$\phi^* g_{ab} = \check{g}_{ab} + h_{ab}. \tag{4.1}$$

The physical spacetime is assumed to be "nearly flat", and ϕ chosen such that h_{ab} is small everywhere, so that the linear approximation to the Einstein field equations is valid:

$$\widehat{G}_{ab}{}^{cd} h_{cd} = \kappa \check{T}_{ab} + O(h^2), \tag{4.2}$$

where $\check{T}_{ab} \equiv \phi^* T_{ab} \sim O(h)$ is the matter energy-momentum tensor T_{ab} mapped onto the background, and

$$\widehat{G}_{ab}{}^{cd} h_{cd} \equiv \check{\nabla}_c \check{\nabla}_{(a} h_{b)}{}^c - \tfrac{1}{2}\check{\nabla}^2 h_{ab} - \tfrac{1}{2}\check{\nabla}_a \check{\nabla}_b h + \tfrac{1}{2}\check{g}_{ab}\left(\check{\nabla}^2 h - \check{\nabla}_c \check{\nabla}_d h^{cd}\right) \tag{4.3}$$

is the linearised Einstein tensor $G^{(1)}_{ab}$.

The gravitational energy-momentum tensor τ_{ab} is defined by seeking a symmetric tensor, quadratic in $\check{\nabla}_c h_{ab}$, which solves

$$\check{\nabla}_a j_\mu{}^a + \phi^*(\nabla_a J_\mu{}^a) = 0, \tag{4.4}$$

neglecting terms $O(h^3)$. In the above equation, $J_\mu{}^a \equiv T^a{}_b e_\mu{}^b$ are the (1 energy, 3 momentum) current-densities of matter, associated with the (1 timelike, 3 spacelike) vector fields $e_\mu{}^a \equiv (\phi^{-1})^* \check{e}_\mu{}^a$, the images of the Lorentzian coordinate basis $\check{e}_\mu{}^a \equiv (\partial/\partial x^\mu)^a$ that generate the translational symmetries of the background; the $j_\mu{}^a \equiv \tau^a{}_b \check{e}_\mu{}^b = \tau^a{}_\mu$ constitute the energy-momentum current-densities of the gravitational field. Consequently (4.4) indicates that the extent to which material energy-momentum fails to be conserved at a point in the physical spacetime is exactly equal and opposite to the extent to which gravitational energy-momentum fails to be conserved at the corresponding point in the background. Interactions between matter and gravity can then be understood in terms of a *local exchange* of energy and momentum between the two.

It is not possible to construct a τ_{ab} to solve (4.4) for all gravitational fields, so a condition must be placed on h_{ab} in order to proceed. Of all possible symmetric tensors τ_{ab}, quadratic in $\check{\nabla}_c h_{ab}$, and all (non-trivial, linear and Lorentz invariant) field conditions, only one combination solves (4.4):

$$\kappa \bar{\tau}_{ab} = \tfrac{1}{4}\check{\nabla}_a h_{cd} \check{\nabla}_b \bar{h}^{cd}, \tag{4.5}$$

$$\check{\nabla}^a \bar{h}_{ab} = 0, \tag{4.6}$$

where the overbars signify trace-reversal. Because (4.6) is simply the equation of *harmonic gauge*, which can always be satisfied through a choice of ϕ, the field condition does not restrict the physical applicability of our approach in any respect. In fact, the only

[2] As usual, fields defined on $\mathcal{M}$ have their indices raised and lowered with g_{ab}, and those on $\check{\mathcal{M}}$ with $\check{g}_{ab}$. Lorentzian coordinates $\{x^\mu\}$ are commonly deployed in $\check{\mathcal{M}}$, for which $\check{g}_{\mu\nu} = \eta_{\mu\nu}$.

effect of the field condition is to vastly reduce the gauge freedom in our description of gravitational energy-momentum (4.5). What at first appeared as a weakness, is in fact a great strength of our approach. Essentially, (4.6) indicates that ϕ is to be chosen such that it maps Lorentzian coordinates $\{x^\mu\}$ of the background onto harmonic coordinates $y^\mu(p) \equiv x^\mu(\phi(p))$ of the physical spacetime. This ensures that the energy-momentum currents ${J_\mu}^a$ are defined by the generators of a harmonic coordinate system; these represent the approximate translational symmetries of the physical spacetime (present due to its small curvature) and give a sensible replacement for killing vectors in the absence of an exact symmetry.

The gravitational energy-momentum tensor τ_{ab} has two notable mathematical properties, in addition to solving (4.4). Firstly, the energy-momentum tensor for any (harmonic gauge) gravitational plane-wave

$$h_{ab} = h_{ab}(x^\alpha k_\alpha), \qquad k^a \bar{h}_{ab} = 0, \qquad k^a k_a = 0, \tag{4.7}$$

is completely invariant under the remaining gauge freedom consistent with (4.6) and (4.7). Secondly, and most remarkably of all, τ_{ab} displays the following *positivity property*: all transverse-traceless (TT) gravitational fields have positive energy-density and causal energy-flux, for all observers. To state this rigorously: if, at some point $p \in \check{\mathcal{M}}$, the gravitational field h_{ab} obeys the transverse-traceless conditions

$$\check{\nabla}^a h_{ab} = 0, \quad h = 0, \quad u^a h_{ab} = 0, \tag{4.8}$$

for some timelike vector u^a, then τ_{ab} satisfies the following inequalities

$$v^a \tau_{ab} v^b \geq 0, \tag{4.9}$$

$$v^a \tau_{ac} \tau^c{}_b v^b \leq 0, \tag{4.10}$$

at p, for any timelike vector v^a.

In order to deal with the last trace of gauge freedom that remains after enforcing (4.6), we examined the energy-momentum transferred between the gravitational field and an infinitesimal probe, i.e. a matter “point-source” with energy-momentum tensor

$$\begin{aligned} \check{T}_{00} &= M\delta(\vec{x}) + \tfrac{1}{2} I_{ij} \partial_i \partial_j \delta(\vec{x}), \\ \check{T}_{0i} &= \tfrac{1}{2} (\dot{I}_{ij} - L_{ij}) \partial_j \delta(\vec{x}), \\ \check{T}_{ij} &= \tfrac{1}{2} \ddot{I}_{ij} \delta(\vec{x}), \end{aligned} \tag{4.11}$$

derived by shrinking a compact source down to a point.[3] The exchange is rendered completely gauge invariant by the *monopole-free microaverage*: the incoming wave is split into an sum of Heaviside step-functions, and the energy-momentum delivered by each is

[3] M, I_{ij} and L_{ij} are the mass, moment of inertia, and angular momentum of the source, respectively. Overdots indicate differentiation with respect to $t \equiv x^0$, and the three spatial coordinates are abbreviated $\vec{x} = (x^1, x^2, x^3)$.

integrated over a vanishingly small 4-volume centred on the probe.[4] The result

$$\langle \partial^\mu \tau_{\mu\nu} \rangle_{\int}^{M} = -\tfrac{1}{4}\delta(\vec{x})\ddot{I}_{ij}\partial_\nu h^{\text{TT}}_{ij} \tag{4.12}$$

is equal to the bare (i.e. not microaveraged) energy-momentum delivered by the incident field in TT-gauge. This motivated the programme of fixing the final piece of gauge freedom by insisting that the incident h_{ab} be transverse-traceless; consequently, τ_{ab} represents the gauge-invariant gravitational energy-momentum that is accessible to an infinitesimal probe at rest in the TT frame. Furthermore, due to the positivity property of τ_{ab}, this programme ensures that the gravitational field is always described with positive energy-density and causal energy flux.

The approach we will take for localising gravitational angular momentum will be very similar to the one we have just described. Section 4.2 of this chapter begins with the counterpart of (4.4) for the local exchange of angular momentum between matter and gravity. We will show that the local change in the angular momentum of matter is not entirely accounted for by the change in orbital angular momentum $2x_{[\mu}\tau_{\nu]}{}^a$ carried by the gravitational field: gravity's intrinsic spin $s^a{}_{\mu\nu}$ must be included to balance the exchange. This argument defines $s^a{}_{bc}$ up to the addition of total divergences, so further requirements must be placed on the tensor before we have a unique formula localising gravitational intrinsic spin. We achieve this in section 4.3 by demanding that $s^a{}_{bc}$ satisfy two simple, physically motivated, algebraic conditions, analogous to the algebraic properties of τ_{ab}. As a result, a formula (4.36) is derived for the spin tensor of the gravitational field. The gauge freedom of $s^a{}_{bc}$ is automatically nullified by the TT programme motivated in chapter 3; however, it is still enlightening to reprise our analysis of the infinitesimal probe and develop a microaverage procedure that renders the transfer of angular momentum gauge-invariant without the need of gauge-fixing. This is covered in section 4.4. In section 4.5 we examine the role of the non-spatial components of gravitational angular momentum, demonstrating that the exchange of non-spatial spin $s^a{}_{0i}$ can displace the centre-of-mass of a gravitational probe. We conclude our investigation with a calculation and analysis of the intrinsic spin carried by a gravitational plane-wave.

4.2 Local Angular Momentum Exchange

The purpose of this section, and the one that follows it, is to extend the basic framework of chapter 3 to include a local description of gravitational angular momentum. Unlike our work on τ_{ab}, eliminating gauge freedom will not be a major concern: we already know that harmonic gauge (4.6) is necessary, and that the last trace of freedom must be removed by insisting that the incident field be transverse-traceless. We begin by formulating the local exchange of angular momentum.

As noted in section 4.1, the material energy-momentum current-densities $J_\mu{}^a$ are formed by contracting $T^a_{\ b}$ with the vectors $e_\mu{}^b \equiv (\phi^{-1})^* \check{e}_\mu{}^b$, the push-forward of which

[4] Details are to be found in sections 3.4.2 and 3.4.3 of chapter 3.

under ϕ generate the translational symmetries of the background. Therefore, to define material angular momentum current-densities $J_{\mu\nu}{}^a$, we must contract $T^a_{\ b}$ with the vector fields $(\phi^{-1})^*(2x_{[\mu}\check{e}_{\nu]}{}^b)$, the push-forward of which generate the *rotational* symmetries of the background:[5]

$$J_{\mu\nu}{}^a \equiv T^a_{\ b}(\phi^{-1})^*(2x_{[\mu}\check{e}_{\nu]}{}^b) = 2T^a_{\ b}y_{[\mu}e_{\nu]}{}^b. \tag{4.13}$$

As usual, $\{x^\mu\}$ comprise a Lorentzian coordinate system on the background, and $y^\mu(p) \equiv x^\mu(\phi(p))$ are the image of these coordinates in the physical spacetime. The $\{y^\mu\}$ are harmonic (that is, $\nabla^2 y^\mu = 0$) as a result of the gauge condition (4.6).

We wish to explain the effect of the gravitational field on the angular momentum of matter in terms of a local exchange of angular momentum between the two. Just as (4.4) captured this idea for energy-momentum, we will require

$$\check{\nabla}_a j_{\mu\nu}{}^a + \phi^*(\nabla_a J_{\mu\nu}{}^a) = 0 \tag{4.14}$$

for angular momentum, where $j_{\mu\nu}{}^a$ is the angular momentum current-density of the gravitational field. Neglecting terms $O(h^3)$, equation (4.14) is equivalent to

$$\begin{aligned}
\check{\nabla}_a j_{\mu\nu}{}^a &= -\phi^*(\nabla_a(J_{\mu\nu}{}^a)) \\
&= -\phi^*(T^a_{\ b}\nabla_a(2y_{[\mu}e_{\nu]}{}^b)) \\
&= -\phi^*(T^a_{\ b})\left[(\check{\nabla}_c h_a{}^b + \check{\nabla}_a h_c{}^b - \check{\nabla}^b h_{ac})x_{[\mu}\check{e}_{\nu]}{}^c + 2\check{\nabla}_a(x_{[\mu}\check{e}_{\nu]}{}^b)\right] \\
&= -\check{T}^a_{\ b}(\check{\nabla}_c h_a{}^b)x_{[\mu}\check{e}_{\nu]}{}^c - 2(\check{T}^a_{\ b} - h^{ac}\check{T}_{cb})\check{\nabla}_a(x_{[\mu}\check{e}_{\nu]}{}^b),
\end{aligned} \tag{4.15}$$

where in the last line we used $\phi^*(T^a_{\ b}) = \phi^*(g^{ac}T_{cb}) = \check{T}^a_{\ b} - h^{ac}\check{T}_{cb} + O(h^3)$ and $\check{T}_{ab} = \check{T}_{ba}$. As we now have an equation relating tensors defined on the background, we can express these tensors in terms of their components in the Lorentzian coordinate system:

$$\begin{aligned}
\partial_\alpha j_{\mu\nu}{}^\alpha &= -\check{T}^\alpha_{\ \beta}(\partial_\gamma h_\alpha{}^\beta)x_{[\mu}\delta^\gamma_{\nu]} - 2(\check{T}^\alpha_{\ \beta} - h^{\alpha\gamma}\check{T}_{\gamma\beta})\partial_\alpha(x_{[\mu}\delta^\beta_{\nu]}) \\
&= -x_{[\mu}\check{T}^{\alpha\beta}\partial_{\nu]}h_{\alpha\beta} + 2h_{\beta[\mu}\check{T}_{\nu]}{}^\beta.
\end{aligned} \tag{4.16}$$

Finally, we recall the field equations (4.2) in harmonic gauge,

$$\partial^2\bar{h}_{ab} = -2\kappa\check{T}_{\alpha\beta}, \tag{4.17}$$

and eliminate $\check{T}_{\alpha\beta}$ from (4.16):

$$\begin{aligned}
\partial_\alpha j_{\mu\nu}{}^\alpha &= (\partial^2\bar{h}^{\alpha\beta}x_{[\mu}\partial_{\nu]}h_{\alpha\beta} - 2h_{\beta[\mu}\partial^2\bar{h}_{\nu]}{}^\beta)/2\kappa \\
&= \partial_\alpha\left[2x_{[\mu}\tau_{\nu]}{}^\alpha + h_{\beta[\nu}\partial^\alpha h_{\mu]}{}^\beta/\kappa\right].
\end{aligned} \tag{4.18}$$

[5] We use the term "rotational symmetry" here as a shorthand for both rotations and Lorentz boosts. The three independent vector fields $2x_{[i}\check{e}_{j]}{}^a$ generate rotations (so that $2x_{[1}\check{e}_{2]}{}^a$ rotates about the x^3-axis, for example) and hence define angular momentum current-densities. The $2x_{[0}\check{e}_{i]}{}^a$ generate boosts (in the x^i-direction) and define moment-of-energy current-densities, the interpretation of which we explore in section 4.5.

This is rather surprising result, and one that reveals the importance of the gravitational field's *intrinsic spin*. The first term in the square brackets clearly represents the *orbital angular momentum* of the field: it takes the familiar form $x \times p$ and is the result of the tangential linear momentum about the origin. The second term, in contrast, does not depend explicitly on x^μ; it measures the extent to which the field itself is spinning at a particular point, and contributes the same gravitational angular momentum without regard to *where* this spin is taking place. We are forced by (4.18) to accept that the angular momentum of the gravitational field is not simply orbital, but also has an *intrinsic* component:

$$j_{\mu\nu}{}^{\alpha} = 2x_{[\mu}\tau_{\nu]}{}^{\alpha} + s^{\alpha}{}_{\mu\nu}, \tag{4.19}$$

where $s^{\alpha}{}_{\mu\nu}$ is the gravitational *spin tensor* (composed of intrinsic spin current-densities) without which the local exchange of angular momentum would not balance. Of course, the division of angular momentum into orbital and intrinsic components is not a new idea, and the form of equation (4.19) originates from standard flat-space field theory [17, 18]. In general, the Noether current of a rotational symmetry cannot be constructed entirely from Noether currents of translational symmetries: the mismatch, born of the field's tensorial (or spinorial) structure, is called intrinsic spin.[6] More neatly, and of greater relevance to our later analysis, the energy-momentum tensor and the spin tensor can be derived separately from a Lagrangian by 'gauging' the translational and rotational symmetries of spacetime and taking the functional derivatives with respect to the two gauge fields. In chapter 5 we construct $\tau_{\mu\nu}$ and $s^{\alpha}{}_{\mu\nu}$ according to this method, confirming that our formulae for $\tau_{\mu\nu}$ and $s^{\alpha}{}_{\mu\nu}$ (soon to be derived) are in keeping with the established concepts of energy-momentum and spin.[7]

4.3 Gravitational Intrinsic Spin Tensor

Our immediate goal, of course, is to arrive at a formula for $s^{\alpha}{}_{\mu\nu}$ in terms of $h_{\alpha\beta}$. With this in mind, it is tempting to solve (4.18) simply by setting

$$\kappa s^{\alpha}{}_{\mu\nu} \stackrel{?}{=} h_{\beta[\nu}\partial^{\alpha}h_{\mu]}{}^{\beta}, \tag{4.20}$$

[6]Essentially, this is because a tensor field undergoes two types of transformation when it is rotated. A vector field $A^\mu(x)$, for example, becomes $\Lambda^{\mu}{}_{\nu}A^{\nu}(\Lambda^{-1}(x))$; in the parlance of quantum field theory, this can be understood as a displacement $x \to \Lambda(x)$ generated by the orbital angular momentum operator $x \times p$, and a pointwise Lorentz transformation $A^\mu \to \Lambda^{\mu}{}_{\nu}A^\nu$ generated by the spin operator.

[7]Because spin tensors are usually associated with *asymmetric* energy-momentum tensors, it is worth mentioning that the symmetry of $\tau_{\mu\nu}$ does not contradict the existence of $s^{\alpha}{}_{\mu\nu}$. Typically, one argues that $\tau_{[\mu\nu]} \neq 0$ describes finite torques acting on infinitessimal regions [58, §5.7], and then states that this is only acceptable if one can interpret these torques as generating intrinsic spin: $\partial_\alpha s^{\alpha}{}_{\mu\nu} = 2\tau_{[\mu\nu]}$. Clearly, this argument does not run in reverse: the presence of a spin tensor does not require that the energy-momentum tensor be asymmetric. A symmetric gravitational energy-momentum tensor simply indicates that there are no torques on infinitesimal regions due to gravity, and so (in the absence of matter) the spin-tensor is conserved: $\partial_\alpha s^{\alpha}{}_{\mu\nu} = 2\tau_{[\mu\nu]} = 0$.

and declare that we have found our local description of gravitational spin. However, this is not the only solution: the exchange equation (4.18) only defines ${s^\alpha}_{\mu\nu}$ up to terms with identically vanishing divergence, so further demands must be made of the spin tensor before it can be determined uniquely. Obviously, ${s^\alpha}_{\mu\nu}$ should have the same basic properties as $\tau_{\mu\nu}$: it should be a local, quadratic, Lorentz-covariant function of $h_{\alpha\beta}$, and contains no dimensionful constants other than κ.[8] The general solution to (4.18) is then

$$\kappa {s^\alpha}_{\mu\nu} = h_{\beta[\nu}\partial^\alpha {h_{\mu]}}^\beta + \partial_\beta {\Sigma^{\alpha\beta}}_{\mu\nu}, \tag{4.21}$$

where ${\Sigma^{\alpha\beta}}_{\mu\nu}$ is any local, quadratic, Lorentz-covariant function of $h_{\alpha\beta}$ (but not its derivatives) that obeys

$${\Sigma^{\alpha\beta}}_{\mu\nu} = -{\Sigma^{\beta\alpha}}_{\mu\nu} = -{\Sigma^{\alpha\beta}}_{\nu\mu}. \tag{4.22}$$

The most general tensor that can be formed from $h_{\alpha\beta}$ this way is

$${\Sigma^{\alpha\beta}}_{\mu\nu} \equiv A_1 h^\alpha_{[\mu} h^\beta_{\nu]} + A_2 h h^{[\alpha}_{[\mu}\delta^{\beta]}_{\nu]} + A_3 h^\gamma_{[\mu}\delta^{[\beta}_{\nu]} h^{\alpha]}_\gamma + \delta^\alpha_{[\mu}\delta^\beta_{\nu]}\left(A_4 h^2 + A_5 h_{\gamma\delta}h^{\gamma\delta}\right), \tag{4.23}$$

where the $\{A_n\}$ are arbitrary dimensionless constants. Equations (4.21) and (4.23) describe the range of possible gravitational spin tensors that account for the angular momentum exchanged with matter; the aim of this current section is to find a distinguished member of this set, deserving of its physical interpretation.

We encountered a similar "superpotential"[9] freedom when deriving $\tau_{\mu\nu}$ in chapter 3, and extinguished it immediately by insisting that the energy-momentum tensor should be free of second derivatives. Unfortunately, this tactic is of no use here: all the terms in ${s^\alpha}_{\mu\nu}$ have the same form $h\partial h$, and so cannot be distinguished from one another by their differential structure. Instead, we must place *algebraic* requirements on the spin tensor, and we shall do so by choosing two conditions that are physically well-motivated, and closely analogous to the algebraic properties of $\tau_{\mu\nu}$.

4.3.1 The Plane-wave Condition

Condition 1: The spin tensor of any (harmonic gauge) gravitational plane-wave (4.7), with wave-vector k^μ, must obey

$${s^\alpha}_{\mu\nu} \propto k^\alpha. \tag{4.24}$$

[8] This last stipulation (which forces the terms in ${s^\alpha}_{\mu\nu}$ to contain exactly one derivative, in order that they have the correct units) is essentially unavoidable within the context of classical general relativity: κ is the only dimensionful constant available. If we allow ourselves to use Planck's constant $\hbar$ (as we would for a quantum theory) or introduce a new dimensionful gravitational constant (as would arise in a higher-derivative theory of gravity) then higher derivative terms would be dimensionally permissible within the spin tensor; nonetheless, these higher-derivative terms would each be multiplied by small factors (such as the Planck length) that would ensure the terms were negligible within the low-curvature regime of the theory that corresponds to classical general relativity.

[9] These superpotentials are so called because they are total derivatives. They bear no relation to the homonymous concept from supersymmetric field theory.

Clearly, this ensures that spin flows in the direction of propagation of the wave, a physically reasonable request that reciprocates the property $\tau_{\mu\nu} \propto k_\mu k_\nu$ of plane-wave energy-momentum. Substituting (4.7) into equations (4.21) and (4.23), we find that the condition (4.24) holds for all harmonic gauge plane-waves if and only if

$$A_2 = -A_1, \qquad A_4 = A_1/4, \qquad A_3 = A_5 = 0. \tag{4.25}$$

This leaves us with a much smaller range of spin tensors

$$\kappa s^\alpha{}_{\mu\nu} = h_{\beta[\nu}\partial^\alpha h_{\mu]}{}^\beta + A_1\partial_\beta\left(\bar{h}^\alpha{}_{[\mu}\bar{h}^\beta{}_{\nu]}\right), \tag{4.26}$$

parametrised by A_1.

Of course, the influence of this first condition is not limited to gravitational plane-waves. In fact, the restriction (4.26) automatically endows the spin tensor with two highly desirable properties that apply to much more general gravitational fields. Furthermore, one can check that these two properties occur *only if* the spin tensor takes the form (4.26); hence the logic can be reversed, with both properties taken together as conditions on $s^\alpha{}_{\mu\nu}$, and (4.24) derived as a consequence.

Property 1a: The spin carried by a transverse-traceless gravitational field (4.8) is purely *spatial*:

$$h_{0\alpha} = 0,\ h = 0,\ \partial_i h_{ij} = 0 \quad \Rightarrow \quad s^\alpha{}_{0i} = 0. \tag{4.27}$$

Not only are transverse-traceless fields blessed with positive energy-density and causal energy-flux, now we see they carry only standard spatial spin! This result is akin to the Frenkel condition [38] that constrains the spin-tensor of a Weyssenhoff fluid [60, 81]: $S^\alpha{}_{0i} = 0$, in the rest-frame of the fluid.[10] The only difference here is that the gravitational field, being massless, has no rest-frame; in its place, the TT-frame defines the space/time split.

The reader should not be under the impression that the non-spatial spins $s^\alpha{}_{0i}$ are completely unphysical, however; as a matter of fact, they have a simple physical interpretation. In section 4.5, we explain that the non-spatial angular momentum current-densities $j_{0i}{}^\alpha$ localise gravity's *Moment-of-Energy*, the conserved quantity associated with the symmetry of the background under Lorentz boosts. Accordingly, the intrinsic current-densities $s^\alpha{}_{0i}$ signify an "internal displacement of energy" of the field. This alters the gravitational Moment-of-Energy just as the "internal spinning motion", signified by $s^\alpha{}_{ij}$, contributes to the total gravitational angular momentum. Due to (4.27) it is now clear that the transverse-traceless field does not carry these internal displacements, and hence, that the location of gravitational energy is determined by $\tau_{\mu\nu}$ alone. The $s^\alpha{}_{0i}$ still play an important role in the local exchange of Moment-of-Energy with matter (see §4.5.2) because TT-gauge cannot be adopted where $\check{T}_{\mu\nu} \neq 0$.

[10] The Weyssenhoff fluid is simply a perfect fluid with intrinsic spin. Note that the massive spin-1/2 field (described by the classical Dirac Lagrangian) also obeys the Frenkel condition, if one takes the charge current-density to define the field's 4-velocity [18].

Property 1b: All *static* distributions of matter give rise to *spinless* gravitational fields:[11]

$$\check{T}_{\mu\nu} = \rho(\vec{x})\delta^0_\mu\delta^0_\nu \quad \Rightarrow \quad s^\alpha{}_{\mu\nu} = 0. \tag{4.28}$$

The meaning of this statement is intuitively obvious: matter must be *in motion* if it is to generate gravitational intrinsic spin. It is worth remembering that the linearised gravitational field due to static matter is just the Newtonian potential Φ, so (4.28) is equivalent to the statement that Φ has no spin. This is in keeping with our observation in chapter 3 that the gravitational field corresponding to a static Newtonian potential has the energy-momentum tensor of a massless scalar field.

4.3.2 The Traceless Condition

So far we have placed one algebraic condition on the gravitational spin tensor and removed all but one of the superpotential degrees of freedom. Our second condition will fix A_1 and determine $s^\alpha{}_{\mu\nu}$ uniquely. Before we take this step, however, it will be valuable to examine the spin tensor of the transverse-traceless gravitational field in detail. The purpose of this analysis is to isolate an algebraic property of $s^\alpha{}_{\mu\nu}$ that signifies unphysical behaviour; we will then design our condition so that this possibility cannot arise.[12]

Because the spin of a transverse-traceless field is spatial (Property 1a) we can write

$$s^\alpha{}_{ij} \equiv s^\alpha{}_k \epsilon_{kij}, \tag{4.29}$$

where $s^\alpha{}_i$ is the *axial* spin tensor, the current-density of intrinsic spin about the x^i-axis. Each component s_{ij} represents the flux of x^j-axis spin in the x^i-direction; in other words, $s_{ij}\Sigma$ is the torque (along the x^j-axis) that acts on a small surface ($x^i = \text{const.}$) of area Σ.

Let us consider the $l \to 0$ limit of an $l \times l \times l$ cube of vacuum ($\check{T}_{\mu\nu} = 0$, $h_{\mu\nu} \neq 0$) as depicted in figure 4.1. The torque along the x^2-axis, acting on the $x^1 = 0$ face, is $G_1 = s_{12}l^2$. There will also be contributions from the $x_{[i}\tau_{j]k}$ part of the angular-momentum current density, but these terms will be of order l^3 and so can be safely neglected. It is convenient to think of G_1 as being generated by two equal and opposite forces $F_1 = 2s_{12}l$ acting on the points $(0, l/2, l/4)$ and $(0, l/2, 3l/4)$ as shown in the diagram. On the opposite face ($x^1 = l$) there will be a torque along the x^2-axis $G'_1 = -(s_{12} + l\partial_1 s_{12})l^2$, the minus sign arising as a result of the opposite direction of the outward normal, and the second term being negligible as long as s_{ij} is smooth in the cube. Again, this torque can be thought of as being generated by equal and opposite forces F_1 acting at the points $(l, l/2, l/4)$ and $(l, l/2, 3l/4)$. Following the same approach, we render the x^1-axis torques on the $x^2 = 0$ and $x^2 = l$ faces as forces $F_2 = 2s_{21}l$ acting on the appropriate points on the cube.

[11] In order that the distribution does not collapse under its own gravity, the matter will also have stresses $T_{ij} \sim O(\rho h) \sim O(h^2)$, but these can be neglected in the linear approximation.

[12] Note that we restrict our attention to the spin of the *transverse-traceless* gravitational field. The spin tensor can only be expected to have a sensible physical interpretation under the same conditions that $\tau_{\mu\nu}$ describes positive energy-density and causal energy-flux, i.e. for all TT-fields, arbitrary (harmonic-gauge) plane-waves, and static fields. It will be trivial to extend Condition 2 to include the last two cases, and since their inclusion does not constrain A_1, it is simpler to ignore them in what follows.

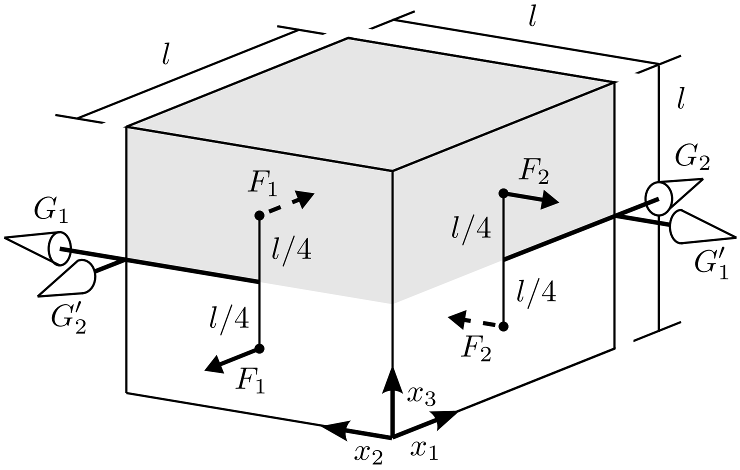

Figure 4.1: The torques on an infinitesimal cube of vacuum due to the flux of gravitational intrinsic spin.

We now split the cube along the plane $x^3 = l/2$, and consider the two "half-cubes" separately. The isotropic pressure acting on each half-cube can be evaluated using the formula

$$P = \frac{-1}{6} \sum_r \frac{\vec{f}_r \cdot \vec{n}_r}{A_r}, \tag{4.30}$$

where the index r enumerates the six faces of the half-cube (each with area A_r and outward unit normal $\vec{n}_r$) and $\vec{f}_r$ is the force acting on the r^{th} face.[13] For the upper half-cube ($x^3 \geq l/2$) both F_1 forces are directed inwards, while the two F_2 forces are outwardly directed; thus (4.30) gives

$$P_{\text{upper}} = \frac{-1}{6} \frac{2F_2 - 2F_1}{l^2/2} = \frac{4(s_{12} - s_{21})}{3l}, \tag{4.31}$$

where we have once again ignored the negligible forces, such as $\tau_{33} l^2$ on the $x^3 = l$ and $x^3 = l/2$ faces. The calculation for the lower half-cube ($x^3 \leq l/2$) is identical except that the forces F_1 point outward and F_2 point inward; as a result, $P_{\text{lower}} = -P_{\text{upper}}$. Therefore, within the cube we find a pressure gradient

$$\frac{\partial P}{\partial x^3} \approx \frac{P_{\text{upper}} - P_{\text{lower}}}{l/2} = \frac{16(s_{12} - s_{21})}{3l^2}, \tag{4.32}$$

which grows without bound as the limit $l \to 0$ is taken! The only way to avoid these infinite pressure gradients is to insist that $s_{[ij]} = 0$, or equivalently

$$s^{\alpha}{}_{\alpha\nu} = 0. \tag{4.33}$$

[13]To confirm the validity of this formula, describe the forces in terms of a stress tensor σ_{ij} by writing $f_{ri} = -\sigma_{ij} n_{rj} A_r$. The sum in (4.30) then becomes $-\sigma_{ij} \sum_r n_{ri} n_{rj} = -\sigma_{ij}(2\delta_{ij}) = -6P$.

This argument inspires the second algebraic condition that we place on the gravitational spin tensor.

Condition 2: The spin tensor of a transverse-traceless gravitational field (4.8) must be traceless:

$$h_{0\alpha} = 0,\ h = 0,\ \partial_i h_{ij} = 0 \quad \Rightarrow \quad s^{\alpha}{}_{\alpha\nu} = 0. \tag{4.34}$$

As we have just seen, this condition rids the transverse-traceless gravitational field of unphysical pressure gradients and is roughly analogous to the symmetry property of energy-momentum, $\tau_{[\mu\nu]} = 0$.[14] Furthermore, this condition strengthens the similarity between gravitation spin and standard examples of material spin: the spin tensors of the Weyssenhoff fluid [60, 81], and the spin-1/2 field [18], are also traceless.

The spin tensor (4.26) is consistent with Condition 2 if and only if

$$A_1 = -1; \tag{4.35}$$

as a result, we arrive at our final formula for the gravitational spin tensor:

$$\kappa s^{\alpha}{}_{\mu\nu} = 2\bar{h}_{\beta[\nu}\partial^{[\alpha}\bar{h}_{\mu]}{}^{\beta]}. \tag{4.36}$$

This is the *unique* local, quadratic, Lorentz-covariant function of $h_{\mu\nu}$ that accounts for the local exchange of angular momentum with matter (4.14), satisfies the two physically well-motivated algebraic conditions (4.24) and (4.34), and contains no dimensionful constants other than κ. This is an exceptionally compact formula, and one that embodies a remarkably parsimonious description of gravitational spin: for a transverse-traceless field, $s^{\alpha}{}_{\mu\nu}$ is specified by no more than 9 independent components (due to (4.27) and (4.34)) as opposed to the 24 that would be needed in the generic case.

Lastly, we should mention that equation (4.35) has a fundamental significance of its own: the superpotential term $-\partial_\beta \left(h^{\alpha}{}_{[\mu}h^{\beta}{}_{\nu]}\right)$ is essential for generalising the gauge invariance of $h_{\mu\nu}$ beyond the flat background, and ensures that a *local* field redefinition is sufficient to cast $\tau_{\mu\nu}$ and $s^{\alpha}{}_{\mu\nu}$ as the first self-interaction terms in the Einstein (and Einstein-Cartan) field equations. We give a full explanation of these statements in chapter 5.

This completes the foundational portion of the chapter. Following the structure of chapter 3, our next task is to apply our newly assembled framework to an investigation of the angular momentum absorbed by an infinitesimal gravitational detector. Section 4.4 will focus on the exchange of standard (i.e. spatial) angular momentum $j_{ij}{}^{\alpha}$, and the microaverage that renders this process gauge-invariant; section 4.5 concerns the interpretation of *non-spatial* angular momentum $j_{i0}{}^{\alpha}$, and the physical consequences of its exchange. A reader whose primary interests are the theoretical underpinnings of $\tau_{\mu\nu}$ and $s^{\alpha}{}_{\mu\nu}$ may wish to skip to chapter 5 at this point: knowledge of sections 4.4, 4.5, and 4.6 will not be necessary for the discussion therein.

[14] Of course, the argument [58, §5.7] that justifies $T_{[\mu\nu]} = 0$ is highly analogous to (and indeed, the main inspiration of) the argument given in this section.

4.4 Angular Momentum Microaverage

Having derived the formula (4.36) for gravitational spin, we now possess a complete description of the local energy, momentum, and angular momentum carried by the linear gravitational field. Our first application of this framework will be an analysis of the angular momentum exchanged with an infinitesimal probe. This will allow us to revisit the *monopole-free microaverage*, the procedure which defined the gauge-invariant energy-momentum transferred onto the probe, and motivated the (equivalent) programme of preparing the incident field in transverse-traceless gauge. Clearly, this gauge-fixing programme also provides us with an unambiguous definition of the angular momentum exchanged with the probe. What is not obvious, though, is whether a microaveraging procedure can also achieve this effect, allowing us to define a gauge-invariant exchange of angular momentum that does not rely on gauge-fixing. The aim of this section is to confirm the truth of this idea.

We shall consider a system that is almost identical to the one described in section 3.4.1 of chapter 3: a point-like detector in the path of a gravitational "pulse" plane-wave. The gravitational detector will once again consist of an infinitesimal point-source at $\vec{x} = 0$,[15] the energy-momentum tensor of which is given by (4.11) as $M, I_{ij}, L_{ij} \to 0$.[16] The gravitational field

$$h_{\mu\nu} = h^{\text{wave}}_{\mu\nu} + h^{\text{source}}_{\mu\nu}, \tag{4.37}$$

is the sum of the incoming gravitational wave,

$$h^{\text{wave}}_{\mu\nu} = A_{\mu\nu}\delta(k_\alpha x^\alpha - t_0), \quad A_{\mu\nu} = \text{const.}, \quad k_\mu = (1, -1, 0, 0), \quad k^\mu \bar{A}_{\mu\nu} = 0, \tag{4.38}$$

and the field $h^{\text{source}}_{\mu\nu}$ generated by the detector,

$$\partial^2 \bar{h}^{\text{source}}_{\mu\nu} = -2\kappa \check{T}_{\mu\nu}. \tag{4.39}$$

It is important to recognise that the plane-wave (4.38) is not quite the same as the one we used when defining the energy-momentum microaverage. There, the gravitational wave had the profile of a Heaviside step function H, and this brought about an exchange of energy-momentum $\partial^\mu \tau_{\mu\nu} \sim \partial h \partial^2 h \sim \delta(t - t_0)\delta(\vec{x})$ that was confined to an infinitesimal spacetime region over which we could average. The same is not true of angular momentum,

[15] It might appear that we risk a loss of generality in placing the probe at the origin, but this is not the case. To explain, let us consider a uniform translation of the coordinates $x^i \to x^i + a^i$; the probe then lies at $\vec{x} = \vec{a}$, and according to (4.19) the only effect on the gravitational angular-momentum current-density is $\Delta {j_{ij}}^\alpha = 2a_{[j}{\tau_{i]}}^\alpha$. Because a^i is constant, the exchange of angular-momentum associated with this term is simply $\partial_\alpha(\Delta {j_{ij}}^\alpha) = 2a_{[j}\partial^\alpha\tau_{i]\alpha}$, and we already know from (4.12) that $\partial^\alpha \tau_{i\alpha}$ (which quantifies the local exchange of linear momentum) is rendered gauge-invariant by the monopole-free microaverage: $\langle \partial_\alpha(\Delta {j_{ij}}^\alpha)\rangle^{\mathcal{M}}_{f} = 2a_{[j}\langle \partial^\alpha \tau_{i]\alpha}\rangle^{\mathcal{M}}_{f}$. Clearly, this term accounts for the angular momentum that results from the transfer of linear momentum onto the detector; by assuming that the probe is at $\vec{x} = 0$ in what follows, we are simply ignoring the trivial exchange of angular momentum associated with the detector's bulk motion.

[16] We take this limit as the size of the source shrinks to zero. The detector is then a form of generalised "test-particle" with negligible self-interaction in comparison to the effect of the external field.

however: a step function wave will give rise to a local exchange $\partial_\alpha j_{\mu\nu}{}^\alpha$ including a term $\partial_\alpha s^\alpha{}_{\mu\nu} \sim h\partial^2 h \sim H(t-t_0)\delta(\vec{x})$ that is not localised at $t = t_0$ and is therefore unsuitable for microaveraging. We have no choice but to use a delta-function wave to generate a point-like angular momentum exchange.[17] This will be the only modification needed to adapt the microaverage for angular momentum.[18]

Following the same reasoning that took us to equation (3.42) of chapter 3, we find that the exchange of spatial angular momentum for this system is given by

$$\begin{aligned}\partial_\alpha j_{ij}{}^\alpha &= 2x_{[i}\partial^\alpha \tau_{j]\alpha} + h^\beta{}_{[j}\partial^2 \bar{h}_{i]\beta}/\kappa \\ &= -\tfrac{1}{2}k_{[j}x_{i]}\dot{\delta}(k_\alpha x^\alpha - t_0)\Big[\ddot{I}_{kl}A_{kl}\delta(\vec{x}) - 2\big(\dot{I}_{kl} - L_{kl}\big)A_{k0}\partial_l\delta(\vec{x}) \\ &\qquad\qquad + \big(2M\delta(\vec{x}) + I_{kl}\partial_k\partial_l\delta(\vec{x})\big)A_{00}\Big] \\ &\quad - \delta(k_\alpha x^\alpha - t_0)\Big[A_{0[i}\big(\dot{I}_{j]k} - L_{j]k}\big)\partial_k\delta(\vec{x}) - A_{k[i}\ddot{I}_{j]k}\delta(\vec{x})\Big].\end{aligned} \tag{4.40}$$

As was the case with energy-momentum, the local exchange of angular momentum (4.40) is clearly not invariant under the gauge transformations

$$\delta A_{\mu\nu} = E_{(\mu}k_{\nu)}, \quad E_\mu = \text{const.}, \tag{4.41}$$

which neither break the harmonic condition (4.6) nor alter the form (4.38) of the wave. This gauge dependence can be dealt with in one of two ways. The simplest approach is to invoke the familiar TT programme, insisting that the incident field be transverse-traceless: $A_{\mu\nu} = A^{\text{TT}}_{\mu\nu}$. The alternative, which will now examine, is to integrate over the infinitesimal interaction region and render the exchange gauge invariant *without gauge-fixing*. The two methods give identical results, as we shall soon show.

The microaverage $\langle \ldots \rangle_{t_0}$ is defined, just as it was in chapter 3, by

$$\begin{aligned}\langle f\rangle_{t_0} &\equiv \delta(\vec{x})\delta(t-t_0)\lim_{\epsilon\to 0}\int_{\mathcal{B}_\epsilon(t_0)} f\,\mathrm{d}^4x, \\ \text{where}\quad \mathcal{B}_\epsilon(t_0) &\equiv \{(t,\vec{x}) : |t-t_0| \le \epsilon, |\vec{x}| \le \epsilon\}.\end{aligned} \tag{4.42}$$

Applying this definition to (4.40) and integrating by parts,[19] we arrive at

$$\langle \partial_\alpha j_{ij}{}^\alpha \rangle_{t_0} = \delta(\vec{x})\delta(t-t_0)\Big[k_{[j}\big(\ddot{I}_{i]k}A_{k0} + \ddot{I}_{i]1}A_{00}\big) + A_{0[i}\ddot{I}_{j]1} + A_{k[i}\ddot{I}_{j]k}\Big]. \tag{4.43}$$

[17] One might try to use a pulse based on derivatives of the delta-function, but the process of splitting a general wave into such pulses is non-local and introduces an arbitrary constant of integration.

[18] The lesson here is that the microaverage is not a process in which we split the incident wave into a particular sort of pulse: as we have seen, the profile of the pulse depends on what exchange we are microaveraging. Rather, it is a process in which we split the wave such that the local exchange (of energy-momentum $\partial_\alpha\tau^\alpha{}_\mu$, or angular momentum $\partial_\alpha j_{ij}{}^\alpha$) takes a particular form: a series of delta-function pulses (and possibly derivatives of delta-functions) each of which can then be averaged over a vanishingly small 4-volume.

[19] For each term, integrate by parts to move derivatives from $\delta(\vec{x})$ onto the $x_i\dot{\delta}(k_\alpha x^\alpha - t_0)$ or $\delta(k_\alpha x^\alpha - t_0)$ part of the term, convert $\partial_i\delta(k_\alpha x^\alpha - t_0) = -\delta_{1i}\dot{\delta}(k_\alpha x^\alpha - t_0)$, and integrate by parts once again to send the time-derivatives to the M, J_{ij}, I_{ij} part of the term, recalling that $\dot{M} = \dot{J}_{ij} = 0$. Note that at least one of the spatial derivative must act on the x_i in front of the orbital terms: those terms where x_i is left untouched will vanish because $\delta(\vec{x})$ will set $x_i = 0$ when the integral is finally evaluated.

Although it is far from obvious in its current form, this equation is in fact invariant under the gauge transformations given in equation (4.41). The easiest way to demonstrate this is to examine each component in turn and to use $k^\mu \bar{A}_{\mu\nu} = 0$ in the following form:

$$A_{00} + A_{11} + 2A_{01} = 0, \qquad A_{22} + A_{33} = 0, \qquad A_{02} + A_{12} = 0, \qquad A_{03} + A_{13} = 0. \tag{4.44}$$

After a great deal of cancelling, one finds that

$$\begin{aligned}
\langle \partial_\alpha j_{23}{}^\alpha \rangle_{t_0} &= \delta(\vec{x})\delta(t-t_0)\left(A_+ \ddot{I}_{23} - A_\times(\ddot{I}_{22} - \ddot{I}_{33})/2\right), \\
\langle \partial_\alpha j_{12}{}^\alpha \rangle_{t_0} &= -\delta(\vec{x})\delta(t-t_0)\left(A_+ \ddot{I}_{12} + A_\times \ddot{I}_{13}\right)/2, \\
\langle \partial_\alpha j_{13}{}^\alpha \rangle_{t_0} &= -\delta(\vec{x})\delta(t-t_0)\left(A_\times \ddot{I}_{12} - A_+ \ddot{I}_{13}\right)/2,
\end{aligned} \tag{4.45}$$

all of which depend only on the transverse components of the wave $A_\times = A_{23}$, $A_+ = (A_{22} - A_{33})/2$ which are invariant under (4.41). Considering that the microaverage was developed purely for the purposes of *energy-momentum* exchange, it is gratifying to discover that it renders the exchange of angular momentum gauge invariant as well.

It is possible to write the above relations (4.45) in a more compact form:

$$\langle \partial_\alpha j_{ij}{}^\alpha \rangle_{t_0} = \delta(\vec{x})\delta(t-t_0) A^{\mathrm{TT}}_{k[i} \ddot{I}_{j]k}, \tag{4.46}$$

where $A^{\mathrm{TT}}_{\mu\nu}$ is the transverse-traceless part of $A_{\mu\nu}$, the only non-zero components of which are $A^{\mathrm{TT}}_{22} = -A^{\mathrm{TT}}_{33} = A_+$ and $A^{\mathrm{TT}}_{23} = A^{\mathrm{TT}}_{32} = A_\times$. As previously advertised, this is exactly the same result as would be obtained from applying the TT programme to the bare angular momentum exchange (4.40):

$$\begin{aligned}
\partial_\alpha j^{\mathrm{TT}\,\alpha}_{ij} &\equiv \partial_\alpha j_{ij}{}^\alpha [h^{\mathrm{source}}_{\mu\nu} + A^{\mathrm{TT}}_{\mu\nu}\delta(k_\alpha x^\alpha - t_0)] \\
&= \delta(\vec{x})\delta(t-t_0) A^{\mathrm{TT}}_{k[i}\ddot{I}_{j]k}.
\end{aligned} \tag{4.47}$$

The only subtlety with this calculation is that one must set $x_i \dot{\delta}(k_\alpha x^\alpha - t_0)\delta(\vec{x}) = 0$, which is valid as an identity between distributions on test functions that are differentiable with respect to t at $(t_0, \vec{0})$.

The angular momentum microaverage need not be restricted to plane-wave pulses: we can generalise equation (4.46) following the same procedure as the energy-momentum case. First we note that an arbitrary (harmonic-gauge) plane-wave

$$h^{\mathrm{wave}}_{\mu\nu} = B_{\mu\nu}(k_\alpha x^\alpha), \qquad k^\mu \bar{B}_{\mu\nu} = 0, \tag{4.48}$$

can be split into a sum of individual pulses

$$h^{\mathrm{wave}}_{\mu\nu} = \int_{-\infty}^{\infty} B_{\mu\nu}(t_0)\delta(k_\alpha x^\alpha - t_0)\mathrm{d}t_0, \tag{4.49}$$

and the angular momentum exchange of each pulse microaveraged separately:[20]

$$\left\langle \partial_\alpha j_{ij}{}^\alpha [h^{\mathrm{source}}_{\mu\nu} + h^{\mathrm{wave}}_{\mu\nu}] \right\rangle_{\int\delta} \equiv \int_{-\infty}^{\infty} \langle \partial_\alpha j_{ij}{}^\alpha [h^{\mathrm{source}}_{\mu\nu} + B_{\mu\nu}(t_0)\delta(k_\alpha x^\alpha - t_0)] \rangle_{t_0} \mathrm{d}t_0. \tag{4.50}$$

[20] This microaverage carries the subscript $\int\delta$ to remind us that the wave has been split into δ-function pulses, rather than Heaviside steps.

Second we recall that any incident field $h^{\rm in}_{\mu\nu}$ can be expressed as a sum of plane-waves, at least locally. Because (4.46) is linear in the incident field, we can split any incident field into a sum of plane-waves, each of which can be split into a sum of pulses, then perform the microaverage on each element and reassemble the result. The general formula is therefore

$$\langle \partial_\alpha j_{ij}{}^{\alpha} \rangle_{\int\delta} = \delta(\vec{x}) h^{\rm TT}_{k[i} \ddot{I}_{j]k}, \tag{4.51}$$

where $h^{\rm TT}_{\mu\nu}$ is the transverse-traceless part of $h^{\rm in}_{\mu\nu}$.

This concludes our analysis of the spatial angular momentum transferred onto the probe. The non-spatial currents $j_{0i}{}^{\alpha}$ can also be absorbed by the detector; the exchange equation (4.14) then ensures that the shift in gravity's moment-of-energy is accompanied by a displacement in the detector's centre-of-mass. This is a rather surprising phenomenon, and one that, to our knowledge, has not been discussed in the literature. Under resonant conditions, this effect can cause the detector to "walk" in a direction transverse to the gravitational wave.[21] The next section is devoted to a detailed examination of this phenomenon.

4.5 Moment of Energy

Through its unification of space and time, and energy and momentum, special relativity fused together the once disparate notions of angular momentum and centre-of-mass. In this section we review this idea in terms of local currents, and offer an interpretation for the non-spatial intrinsic spin currents $s^{\alpha}{}_{0i}$. We also examine the local exchange of moment-of-energy between the gravitational field and an infinitesimal detector. In appendix 4.A we confirm that this phenomenon is also predicted by a "first principles" description of the system.

4.5.1 Definitions and Interpretation

It goes without saying that the non-spatial components $j_{0i}{}^{\alpha}$ and $J_{0i}{}^{\alpha}$ are needed to form the Lorentz-covariant currents $j_{\mu\nu}{}^{\alpha}$ and $J_{\mu\nu}{}^{\alpha}$; thus, at the most basic level, these non-spatial components carry the interpretation of standard "spatial" angular-momentum as seen by a moving observer. Beyond this, the non-spatial components carry an additional interpretation that is quite distinct from spatial angular momentum. They are the current densities of a conserved 3-vector quantity: *the Moment-of-Energy at* $t = 0$.

To explain, let us first define the moment-of-energy X_i, total linear momentum P_i, and total mass/energy M for matter:

$$X_i \equiv -\int \sqrt{-g} T^0{}_0 y_i \mathrm{d}^3 y, \qquad P_i \equiv \int \sqrt{-g} T^0{}_i \mathrm{d}^3 y, \qquad M \equiv -\int \sqrt{-g} T^0{}_0 \mathrm{d}^3 y, \tag{4.52}$$

[21] This should not be confused with the motion associated with the linear momentum that the probe gains according to (4.12). There, a resonance between the detector and the incident wave gives rise to a longitudinal acceleration, and the velocity gained in this process remains after the wave passed.

noting that the centre-of-mass $x_i^{(0)}$ is simply the moment-of-energy normalised by the total mass/energy:

$$x_i^{(0)} \equiv X_i/M. \tag{4.53}$$

The total non-spatial angular momentum of matter is then

$$\int \sqrt{-g} J_{0i}{}^{0} \mathrm{d}^3 y \equiv \int \sqrt{-g}(T^0{}_i y_0 - T^0{}_0 y_i)\mathrm{d}^3 y \equiv -tP_i + X_i, \tag{4.54}$$

where we have written $y^0 = t$.[22] In the absence of the gravitational field ($h_{\mu\nu} = 0$) the angular momentum currents are conserved,

$$\partial_\alpha(\sqrt{-g} J_{\mu\nu}{}^{\alpha}) = \sqrt{-g}\nabla_a J_{\mu\nu}{}^{a} = 0, \tag{4.55}$$

and as a result,

$$\partial_t \left(X_i - tP_i\right) = 0. \tag{4.56}$$

Furthermore, the conservation of energy-momentum ($\partial_\alpha(\sqrt{-g}T^\alpha{}_\mu) = \sqrt{-g}\nabla_a J_\mu{}^a = 0$) ensures that $\dot{P}_i = 0$, and leads to the following global conservation law:

$$\dot{X}_i - P_i = 0. \tag{4.57}$$

This equation integrates to $X_i = tP_i + X_i|_{t=0}$, which on substitution into (4.54) gives

$$\int \sqrt{-g} J_{0i}{}^{0} \mathrm{d}^3 y = X_i|_{t=0}, \tag{4.58}$$

which is constant by definition. In other words, the total non-spatial angular momentum is equal to the *moment-of-energy at* $t = 0$, a conserved quantity which we will refer to by the acronym MoE, where the stipulation "at $t = 0$" should be taken as given.

The same analysis can be performed for the gravitational field in the absence of matter. Working in the background, we define

$$\mathcal{X}_i^\tau \equiv \int \tau_{00} x_i \mathrm{d}^3 x, \qquad \mathcal{P}_i \equiv \int \tau^0{}_i \mathrm{d}^3 x, \qquad \mathcal{X}_i^s \equiv \int s^0{}_{0i} \mathrm{d}^3 x, \qquad \mathcal{X}_i \equiv \mathcal{X}_i^\tau + \mathcal{X}_i^s. \tag{4.59}$$

Then the total non-spatial gravitational angular momentum is given by

$$\int j_{0i}{}^{0} \mathrm{d}^3 x \equiv \mathcal{X}_i - t\mathcal{P}_i, \tag{4.60}$$

which, due to $\partial_\alpha j_{0i}{}^\alpha = 0$ and $\partial_\alpha \tau^\alpha{}_i = 0$, is conserved:

$$\dot{\mathcal{X}}_i - \mathcal{P}_i = 0, \qquad \int j_{0i}{}^{0} \mathrm{d}^3 x = \mathcal{X}_i|_{t=0}. \tag{4.61}$$

[22] Note that we use the same symbol t to represent the value of the time coordinate y^0 in physical spacetime and the time coordinate x^0 of the background. This has the advantage of allowing us to drop the distinction between the physical quantities X_i, P_i, M, $x_i^{(0)}$, and their background representations $\phi^*(X_i)$, $\phi^*(P_i)$, $\phi^*(M)$, $\phi^*(x_i^{(0)})$: the first set are functions of y^0 only, the second set of x^0 only, and the two sets are numerically equal when $x^0 = y^0$.

We conclude from this that the ${j_{0i}}^{\alpha}$ are the current-densities of the conserved quantities $\mathcal{X}_i|_{t=0}$ that constitute the gravitational MoE.

As (4.59) makes clear, the non-spatial spin densities ${s^0}_{0i}$ shift the gravitational MoE by $\mathcal{X}_i^s$, displacing it from the value $\mathcal{X}_i^\tau$ that would have been expected from τ_{00} alone. This suggest that the ${s^0}_{0i}$ represent an "internal displacement of energy" at a point (analogous to the notion of ${s^0}_{ij}$ as "internal spinning motion" at a point) so that the field's energy lies locally off-centre. The value of $\tau_{00}(p)$ still represents the density of gravitational energy at the point p, but an asymmetry in the distribution of the energy "within the point", quantified by ${s^0}_{0i}$, shifts the MoE by a small amount.[23] Because ${s^0}_{0i} = 0$ for any transverse-traceless gravitational field, these internal displacements rarely arise when describing the energetics of the gravitational field in vacuum. However, as TT-gauge cannot be adopted where $\check{T}_{\mu\nu} \neq 0$, the ${s^0}_{0i}$ inevitably play an active role in the exchange of MoE between matter and gravity.

4.5.2 Moment of Energy Exchange

When matter and gravity interact, neither ${j_{0i}}^{\alpha}$ nor ${J_{0i}}^{\alpha}$ are independently conserved, and MoE is exchanged between them according to (4.14). Consequently, the conservation laws (4.57) and (4.61) are broken,

$$\dot{X}_i - P_i \equiv \Delta\dot{X}_i \neq 0, \tag{4.62}$$

$$\dot{\mathcal{X}}_i - \mathcal{P}_i \equiv \Delta\dot{\mathcal{X}}_i \neq 0, \tag{4.63}$$

but the extent to which they are broken is exactly equal and opposite:[24]

$$\Delta\dot{\mathcal{X}}_i + \Delta\dot{X}_i = 0. \tag{4.64}$$

To understand this process in general, we turn once again to our preferred testing ground: an infinitesimal detector in the path of a gravitational plane-wave. Unlike our analysis of angular momentum for this system (§4.4) we will not employ the microaverage here. The reason for this is simple: the microaverage does not produce a gauge-invariant description of the exchange of MoE. In contrast to angular momentum and energy-momentum, the gauge invariant modes of the gravitational field do not deliver MoE evenly across the whole detector, they are biased by a dipole term proportional to $\partial_i\delta(\vec{x})$.[25] The microaverage is therefore unable to capture the exchange properly, as it can only produce quantities proportional to $\delta(\vec{x})$. This is an notable qualitative difference between the exchange of angular momentum and MoE, but in reality it poses no practical

[23] This pointwise internal structure (spinning motion and displacements) presumably takes place in the tangent space of the manifold, where the gravitational field is defined.

[24] This global exchange equation follows directly from the local exchange equations: multiply equation (4.14) by $\phi^*(\sqrt{-g}) = \sqrt{-\bar{g}} + O(h)$, discard terms $O(h^3)$, and integrate over the spatial coordinates. This gives $\Delta\dot{\mathcal{X}}_i + \Delta\dot{X}_i - t(\dot{\mathcal{P}}_i + \dot{P}_i) = 0$, and $\dot{\mathcal{P}}_i + \dot{P}_i = 0$ follows from the local exchange of linear-momentum (4.4) by exactly the same method.

[25] This can be seen in equation (4.67) below.

difficulty: we can still remove the gauge dependence by insisting that the incident field is transverse traceless.

With this in mind, we consider the same system as described in section 4.4 with one exception: the incident field is an arbitrary transverse-traceless plane-wave,

$$h^{\text{wave}}_{\mu\nu} = B^{\text{TT}}_{\mu\nu}(t - x_1), \qquad B^{\text{TT}}_{0\nu} = B^{\text{TT}}_{1\nu} = B^{\text{TT}} = 0, \tag{4.65}$$

rather than a pulse. Taking the same steps that were used to derive (3.42) of chapter 3, and deploying the distributional identity $x_i\delta(\vec{x}) = 0$, we find that the local exchange of non-spatial angular momentum is

$$\partial_\alpha j_{10}{}^\alpha = t\partial^\alpha \tau_{\alpha 1}, \tag{4.66}$$

$$i = 2, 3: \quad \partial_\alpha j_{i0}{}^\alpha = B^{\text{TT}}_{ik}\left(\dot{I}_{kj} - L_{kj}\right)\partial_j\delta(\vec{x})/2. \tag{4.67}$$

As the longitudinal (4.66) and transverse (4.67) equations represent two very different phenomena, we shall examine them separately.

Equation (4.66) is essentially trivial: it accounts for the extra MoE that arises from the exchange of linear momentum in the x^1 direction. To demonstrate this, let us take the time derivative of (4.60):

$$\dot{\mathcal{X}}_i - \mathcal{P}_i - t\dot{\mathcal{P}}_i = \int \partial_0 j_{0i}{}^0 \mathrm{d}^3x = \int \partial_\alpha j_{0i}{}^\alpha \mathrm{d}^3x. \tag{4.68}$$

Unlike the non-interacting case, we now have

$$\dot{\mathcal{P}}_i = \int \partial_0 \tau^0{}_i \mathrm{d}^3x = \int \partial_\alpha \tau^\alpha{}_i \mathrm{d}^3x, \tag{4.69}$$

which is nonzero in general. Consequently,

$$\Delta\dot{\mathcal{X}}_i \equiv \dot{\mathcal{X}}_i - \mathcal{P}_i = \int t\partial_\alpha\tau^\alpha{}_i + \partial_\alpha j_{0i}{}^\alpha \mathrm{d}^3x. \tag{4.70}$$

Thus, the quantity that describes the local exchange of MoE is in fact the sum

$$t\partial_\alpha\tau^\alpha{}_i + \partial_\alpha j_{0i}{}^\alpha, \tag{4.71}$$

as it is this combination which contributes the extra increase in $\mathcal{X}_i$ beyond what would be expected from simply integrating $\mathcal{P}_i(t)$ with respect to time. Because the gravitational wave only deposits momentum in the longitudinal direction (see equation (4.12)) this argument has no effect on the interpretation of (4.67); however, equation (4.66) reveals that

$$t\partial_\alpha\tau^\alpha{}_1 + \partial_\alpha j_{01}{}^\alpha = t\partial_\alpha\tau^\alpha{}_1 + (-t\partial_\alpha\tau^\alpha{}_1) = 0, \tag{4.72}$$

confirming that there is no exchange of MoE in the x^1-direction, only the exchange of linear momentum. The centre-of-mass of the detector will accelerate in the x^1-direction, but this acceleration will be exactly what one would expect from the linear momentum transfer discussed in chapter 3.

In comparison, the exchange of transverse MoE (4.67) is considerably less trivial. The first complication is that $\partial_\alpha j_{i0}{}^\alpha \propto \partial_j \delta(\vec{x})$, indicating that the transfer of MoE occurs within a dipole-like distribution, taking opposite signs at opposite ends of the detector. In general, these effects will partially cancel each other, so a more pertinent quantity to calculate (rather than the local exchange) is the *total* MoE exchange over the whole detector:

$$\begin{aligned}\Delta\dot{X}_i = -\Delta\dot{\mathcal{X}}_i &= -\int t\partial_\alpha \tau^\alpha{}_i + \partial_\alpha j_{0i}{}^\alpha \mathrm{d}^3x \\ &= \dot{B}^{\mathrm{TT}}_{ik}(t)\left(\dot{I}_{k1} - L_{k1}\right)/2,\end{aligned} \tag{4.73}$$

for $i = 2, 3$. This equation describes the transverse drift in gravitational MoE, and via (4.64), the opposite drift in the matter MoE.

In general, the centre-of-mass of the detector (4.53) will move according to

$$\dot{x}^{(0)}_i = (\Delta\dot{X}_i + P_i)/M - X_i\dot{M}/M^2, \tag{4.74}$$

under the influence of the gravitational wave. Focusing our interest on the transverse directions (for which $P_i = 0$ for all time) we note that the last term in (4.74) is the product of two small quantities (X_i and $\dot{M}$) and can therefore be neglected in comparison to the first term, which only contains one small quantity ($\Delta\dot{X}$).[26] Making these simplifications, and substituting (4.73) into (4.74), we finally arrive at a formula for the transverse motion of the detector's centre-of-mass:

$$i = 2, 3: \quad \dot{x}^{(0)}_i = \dot{B}^{\mathrm{TT}}_{ik}\left(\dot{I}_{k1} - L_{k1}\right)/2M. \tag{4.75}$$

It is important to realise that this motion is not simply a "coordinate effect". If we were to place a free particle at rest at the origin, then because the plane-wave is TT, this reference point will remain at $\vec{x} = 0$ indefinitely. Equation (4.75) therefore predicts the displacement of the centre-of-mass relative to this reference point, and the proper distance between the two points will be, to lowest order, equal to the Euclidean distance in the background.

In passing we also note that, when $I_{ij} = 0$, the acceleration of the centre-of-mass is exactly that of a spinning test-particle (of mass M and spin L_{ij}) as predicted by the linearised Papapetrou-Dixon equations [30, 65] in transverse-traceless gauge:

$$\begin{aligned}M\ddot{x}^{(0)}_i &= -L_{jk}R_{i0jk}/2 + O(h^2) \\ &= -L_{k1}\ddot{B}^{\mathrm{TT}}_{ik}/2 + O(h^2),\end{aligned} \tag{4.76}$$

[26] To argue this more rigorously, suppose that the incident wave has amplitude B and frequency Ω, and that the internal motions of the probe have frequency ω and amplitude l. We require $B \ll 1$ in the linear approximation, $\Omega l \ll 1$ to ensure that the probe is much smaller than the gravitational wavelength, and $\omega l < 1$ so that the internal motions are not superluminal. It follows from (4.73) that the first term $\Delta\dot{X}/M \sim Bl^2\Omega\omega$, and from (4.12) that the second term $X\dot{M}/M^2 = (X/M)(\int \partial^\alpha \tau_{\alpha 0}\mathrm{d}^3x/M) \sim (X/M)Bl^2\Omega\omega^2$. The factor $X/M = \int \Delta\dot{X}\mathrm{d}t/M$ can be no larger than $(\Delta\dot{X})_{\max}\Delta t/M \sim Bl^2\Omega\omega\Delta t$, where Δt is the duration of the interaction. From this we conclude that the second term $X\dot{M}/M^2 \lesssim B^2l^4\Omega^2\omega^3\Delta t$ is negligible in comparison with the first unless the wave and the probe interact for a very long time $\Delta t \sim (Bl^2\Omega\omega^2)^{-1}$. This becomes completely impossible as the length-scale of the probe $l \to 0$.

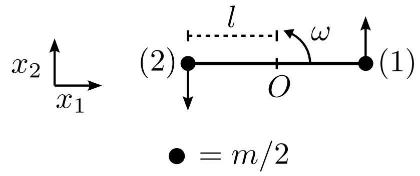

Figure 4.2: A toy model detector: two masses, connected by a light rod, rotate in the $x_3 = 0$ plane; a gravitational plane-wave, propagating in the x_1-direction, disturbs its centre-of-mass.

where, because the probe begins at rest, we have taken $\dot{x}_i^{(0)} = O(h)$. Thus (4.75) generalises this equation to include the effect of the quadrupole moment I_{ij} of the particle. Because this quantity is time dependent, this allows for the possibility of *resonance* between the probe and the wave, the consequence of which we shall explore in the following example.

Example: Rotating Rod

Let us consider the probe depicted in figure 4.2, a light rod (length $2l$) with bobs of mass $m/2$ at each end, spinning with angular frequency ω about the x_3-axis. A valuable feature of equation (4.75) is that one only needs the *unperturbed* motion of the detector (as captured by I_{ij} and L_{ij}) to calculate the motion of the centre-of-mass to lowest order in $h_{\mu\nu}$; this is not true of a "first principles" approach to the problem (see appendix 4.A) which complicates that calculation considerably. The unperturbed locations of the two masses are, in the background,

$$\vec{x}^{(1)} = l(\cos\omega t, \sin\omega t, 0) = -\vec{x}^{(2)}, \tag{4.77}$$

and assuming that the speeds are not relativistic (for the sake of simplicity) it is easy to confirm that

$$\dot{I}_{ij} - L_{ij} = m\omega l^2 \begin{pmatrix} -\sin(2\omega t) & \cos(2\omega t) - 1 & 0 \\ \cos(2\omega t) + 1 & \sin(2\omega t) & 0 \\ 0 & 0 & 0 \end{pmatrix}_{ij} .$$

Inserting this into (4.75) and setting the total mass/energy $M = m$ under the nonrelativistic assumption, we conclude that centre-of-mass of the spinning rod moves according to

$$\dot{x}_i^{(0)} = \frac{\omega l^2}{2} \dot{B}_{i2}^{\mathrm{TT}}(t)(\cos(2\omega t) + 1), \tag{4.78}$$

in the transverse directions $i = 2, 3$. For a generic gravitational wave, this equation predicts an oscillation in the centre-of-mass that averages to zero over many wavelengths. If the wave is of frequency 2ω, however, a resonance occurs in which the detector can steadily "walk" in the transverse-direction. A gravitational wave of the form

$$\begin{pmatrix} B_{22} \\ B_{23} \end{pmatrix} = \begin{pmatrix} \beta_+ \\ \beta_\times \end{pmatrix} \sin(2\omega(t - x_1)), \tag{4.79}$$

gives rise to an average transverse velocity

$$\begin{pmatrix} \langle \dot{x}_2^{(0)} \rangle \\ \langle \dot{x}_3^{(0)} \rangle \end{pmatrix} = \frac{\omega^2 l^2}{2} \begin{pmatrix} \beta_+ \\ \beta_\times \end{pmatrix}. \tag{4.80}$$

One of the most surprising aspects of this phenomenon is that the walking motion (4.80) is not associated with any transverse momentum: $P_2 = P_3 = 0$. The detector moves without being pushed, as it were: due to a careful conspiracy between the probe's internal motion, and the stretching and squeezing of space, the centre of the probe is displaced with each period.

To understand this on an intuitive level, let us imagine for a moment that the rod joining the masses does not exist, but that at $t = 0$ the masses have the same positions and velocities as before. Because the gravitational wave is invariant under translations in the transverse directions, the transverse momentum (i.e. the transverse components of the momentum covectors) of the two particles will be conserved, and hence the total transverse momentum remains zero. However, the velocity vectors of the masses are related to their conserved momentum covectors by the physical metric g_{ab}, which varies in the x^1-direction. Thus, because the physical metric differs between the positions of the two masses, while their momenta are equal and opposite, their velocities will not be. In this fashion, a gradient in the gravitational field across the detector can cause a drift in the centre-of-mass of the system. The role of the rod in our detector is simply to apply equal and opposite forces to the masses (again, having no effect on the total transverse momentum) so that once $t = \pi/2\omega$ and the gradient of the gravitational wave across the detector has reversed, the masses are now at the same value of x^1, and the drift that has occurred in the first quarter-wavelength will not be undone.

In appendix 4.A we substantiate this intuitive picture with a detailed rederivation of equation (4.73) from first principles. Not only does this further aid our understanding of the phenomenon, it should assuage any concerns that this unfamiliar effect might simply be an unphysical artifact of our formalism. In fact, the subtlety and complexity of this calculation emphasises the computational advantage of our approach, not only for MoE, but for angular momentum and energy-momentum also.

4.6 Gravitational Plane-Waves

As a final exploration of our formula (4.36) for gravitational intrinsic spin, we shall evaluate ${s^\alpha}_{\mu\nu}$ for a plane-wave. The motivation for this endeavour is to point out a number of interesting features, and to allow for a comparison with other descriptions of gravitational angular momentum.

A transverse-traceless gravitational plane-wave

$$h_{\mu\nu} = h_{\mu\nu}(k_\alpha x^\alpha), \qquad k_\mu = (1, -1, 0, 0), \qquad h_{\mu 0} = h_{\mu 1} = h = 0, \tag{4.81}$$

has an extremely simple spin tensor:

$$s^\alpha{}_{\mu 0} = s^\alpha{}_{\mu 1} = 0, \qquad \kappa s^\alpha{}_{23} = k^\alpha (h_\times \dot{h}_+ - h_+ \dot{h}_\times), \tag{4.82}$$

where $h_\times = h_{23}$ and $h_+ = h_{22} = -h_{33}$ are the transverse components of the wave. As one would expect, $s^\alpha{}_{\mu\nu}$ describes transverse spatial spin flowing in the direction of propagation of the wave. Furthermore, the amplitude of $s^\alpha{}_{23}$ quantifies the internal spinning motion of the field, as can be seen when we consider a monochromatic wave where the "plus" and "cross" polarisations differ by a phase θ:

$$h_+ = A_+ \cos\left(\omega(t - x^1)\right), \qquad h_\times = A_\times \cos\left(\omega(t - x^1) - \theta\right). \tag{4.83}$$

In this case, the spin-density is constant over spacetime,

$$\kappa s^0{}_{23} = \omega A_\times A_+ \sin\theta, \tag{4.84}$$

and is greatest in magnitude when the wave is *circularly polarised*, that is, when $\theta = \pm\pi/2$. Note that a wave with a purely linear polarisation will carry no spin at all.

In chapter 3 we saw that the energy-momentum tensor of a TT gravitational plane-wave was independent of the timelike vector u^μ that defines the wave's TT-frame (4.8). A similar property holds for the spin tensor, but it is complicated by the fact that spin is constrained to be spatial with respect to the TT-frame, that is, $u^\nu s^\alpha{}_{\mu\nu} = 0$. As we shall see, the longitudinal and non-spatial spins do transform as the TT-frame is changed, and in doing so they adapt the spin tensor to obey the spatial constraint for the new u^μ; however, the transverse spatial spin current $s^\alpha{}_{23}$ is left invariant. To demonstrate this invariance, we perform a gauge transformation on the field (4.81) that maintains its plane-wave form,

$$\delta h_{\mu\nu} = \partial_{(\mu}\left(\xi_{\nu)}(k_\alpha x^\alpha)\right) = 2k_{(\mu}\dot{\xi}_{\nu)}, \tag{4.85}$$

and note that the spin tensor changes by

$$\kappa\delta s^\alpha{}_{\mu\nu} = k^\alpha k_{[\mu}\left(h_{\nu]\beta}\ddot{\xi}^\beta - \dot{h}_{\nu]\beta}\dot{\xi}^\beta + k^\beta\left(\dot{\xi}_{\nu]}\ddot{\xi}_\beta - \ddot{\xi}_{\nu]}\dot{\xi}_\beta\right)\right), \tag{4.86}$$

confirming that $\delta s^\alpha{}_{23} = 0$.

Now suppose that the gravitational field (4.81) has been transformed to a new TT-frame, so that in some other Lorentz coordinate system $\{x^{\mu'}\}$ we have $h_{\mu' 0'} = h_{\mu' 1'} = 0$. Then by the same calculation that led us to (4.82) the transformed spin tensor $s'^a{}_{bc}$ will obey $s'^{\alpha'}{}_{\mu' 0'} = s'^{\alpha'}{}_{\mu' 1'} = 0$ exactly as the original tensor did in the original coordinate system. The only non-zero component of the transformed tensor (in the primed basis) will be $s'^{\alpha'}{}_{2'3'}$, and this quantity will also be gauge invariant by the same argument we used for $s^\alpha{}_{23}$. These two gauge-invariant currents are related by the constant factor $2\Lambda^{[2'}{}_2\Lambda^{3']}{}_3$, where $\Lambda^{\mu'}{}_\nu$ is the Lorentz transformation between the two coordinate bases:

$$s^\alpha{}_{23} = s'^\alpha{}_{23} = \Lambda^{\mu'}{}_2\Lambda^{\nu'}{}_3 s'^\alpha{}_{\mu'\nu'} = \left(2\Lambda^{[2'}{}_2\Lambda^{3']}{}_3\right) s'^\alpha{}_{2'3'}.$$

This constant of proportionality ensures that $s^a{}_{bc}$ and $s'^a{}_{bc}$ describe exactly the same spatial transverse spin current *in either basis*: $s^\alpha{}_{23} = s'^\alpha{}_{23}$ and $s^{\alpha'}{}_{2'3'} = s'^{\alpha'}{}_{2'3'}$. Thus, the only effect of a change in TT-frame is to re-express the same physical information (the transverse spin current of the wave) in terms of spin that is spatial with respect to a new rest-frame. In the absence of some material body (a detector or a source, for example) the massless gravitational plane-wave cannot define a preferred rest-frame, and so the spatial nature of its intrinsic spin will always have this ambiguity.

As a consequence of this, while a plane-wave region can "sew together" two different TT-frames to form a seamless picture of the propagation of gravitational *energy-momentum* (as described in section 3.4.4 of chapter 3) the same cannot be done for angular momentum: there will always be a discontinuity where the spatial spin of one frame is converted into the spatial spin of the other. Even so, one can construct a gravitational spin *pseudovector*

$$s^\alpha \equiv \epsilon^{\alpha\lambda\mu\nu} s_{\lambda\mu\nu}/2, \tag{4.87}$$

which is truly independent of TT-frame, and will therefore give a continuous description of gravitational spin within the sewing region. The invariance of s^α follows directly from the totally antisymmetric part of (4.86): $\delta s_{[\alpha\mu\nu]} = 0$. The physical interpretation of this pseudovector is not immediately clear, but suffice it to say that for a plane-wave, s^α captures only the spin that is linearly independent of the wave-vector k^μ.[27] For the plane-wave (4.81) we have been studying, the spin pseudovector is

$$\kappa s^\alpha = k^\alpha (h_\times \dot{h}_+ - h_+ \dot{h}_\times), \tag{4.88}$$

capturing all the physically pertinent information of (4.82) in a completely frame-independent fashion.

Finally, we should highlight the major difference that exists between the gravitational spin currents in (4.82) and the corresponding quantities given by the traditional approaches, including the Landau-Lifshitz tensor [52] and the integrand of the ADM energy-momentum [7]. In these descriptions, the local energy-momentum and spin of the gravitational field are packaged together in a single object, a *Belinfante* energy-momentum tensor $t_{\mu\nu} \sim \partial h \partial h + h \partial^2 h + O(h^3)$.[28] The local angular-momentum currents are then $x_{[\mu} t_{\nu]}{}^\alpha$ alone, with no extra "intrinsic" component. According to this viewpoint, there is no transverse angular momentum within a harmonic-gauge plane-wave: $x_{[2} t_{3]}{}^\alpha = 0$.[29] This differs dramatically from our description (4.82) and stands opposed to the intuitive

[27]We also note that s^α bears a resemblance to the Pauli-Lubanski pseudovector $S^\alpha \equiv \epsilon^{\alpha\lambda\mu\nu} P_\lambda L_{\mu\nu}/2$, which characterises the total spin of a particle or matter field, and reduces in the particle's rest-frame to (mass times) the familiar axial angular-momentum vector of non-relativistic mechanics [71].

[28]A Belinfante energy-momentum tensor can be constructed from any energy-momentum tensor and spin tensor, including our own: $t_{\mu\nu}[\tau, s] \equiv \tau_{\mu\nu} + \partial_\alpha (s_{\mu\nu}{}^\alpha + s_{\nu\mu}{}^\alpha - s^\alpha{}_{\mu\nu})/2$. We perform this calculation in chapter 5 and compare the result with the Landau-Lifshitz and ADM Belinfante tensors discussed here.

[29]This follows from simple index combinatorics. Within the plan- wave $t_{\mu\nu} \sim kk\dot{h}\dot{h} + kkh\ddot{h} + O(h^3)$, and because $k_\mu k^\mu = 0$ and $k^\mu \bar{h}_{\mu\nu} = 0$, both the free indices must occur on the wave-vectors, i.e. $t_{\mu\nu} \propto k_\mu k_\nu$. This continues to be true at higher order, where the terms in $t_{\mu\nu}$ are of the form $kk\dot{h}\dot{h}h^{n-2} + kk\ddot{h}h^{n-1}$. Consequently, the transverse angular momentum vanishes exactly: $x_{[2} t_{3]}{}^\alpha \propto x_{[2} k_{3]} k^\alpha = 0$.

notion of intrinsic spin as quantifying the internal spinning motion of the field. Without separating gravitational energy-momentum and spin into two separate tensors, $\tau_{\mu\nu}$ and $s^{\alpha}{}_{\mu\nu}$, the intrinsic spin carried by a (harmonic-gauge) plane-wave can never be manifestly present within the wave.

To be clear: the Belinfante-style descriptions still correctly quantify the *total* angular momentum of the wave, but they assign this angular momentum to the wave's boundary, not its interior.[30] Considering that the angular momentum currents along this boundary are given by $x_{[2}t_{3]}{}^{\alpha}$ as always, and are thus explicitly dependent on x^{μ}, even these currents cannot be thought of as a local and intrinsic property of the field. This perverse picture, in which all the spin of a gravitational wave resides on the edge of the wave, and this supposedly intrinsic quantity depends on the coordinate distance from the origin, only emphasises what was already well-known: the Landau-Lifshitz tensor and the integrand of the AMD energy-momentum should not be taken seriously as *local* descriptions of gravitational energy-momentum or spin. While they certainly define meaningful *global* quantities [10], the gauge-freedom of these Belinfante tensors cannot be fixed in a natural manner, and they commonly display negative energy-density and spacelike energy-flux.

4.7 Conclusion

Together, the energy-momentum tensor $\tau_{\mu\nu}$ and the spin tensor $s^{\alpha}{}_{\mu\nu}$ completely characterise the energy, momentum, and angular momentum carried locally by the linearised gravitational field:

$$\kappa\bar{\tau}_{\mu\nu} = \tfrac{1}{4}\partial_{\mu}h_{\alpha\beta}\partial_{\nu}\bar{h}^{\alpha\beta}, \tag{4.89}$$

$$\kappa s^{\alpha}{}_{\mu\nu} = 2\bar{h}_{\beta[\nu}\partial^{[\alpha}\bar{h}_{\mu]}{}^{\beta]}. \tag{4.90}$$

The gauge freedom of this description is highly constrained by the harmonic gauge condition,

$$\partial^{\mu}\bar{h}_{\mu\nu} = 0, \tag{4.91}$$

which arose as a consequence of the derivation of $\tau_{\mu\nu}$; the last remnant of this freedom is then eliminated by insisting that the incident gravitational field be transverse-traceless, a programme motivated in part by appealing to the gauge-invariant exchange of energy-momentum between gravity and an infinitesimal probe, and also distinguished by the numerous desirable properties that the tensors display in transverse-traceless gauge: positive energy-density, causal energy-flux, and traceless spatial spin.

We developed this framework around a simple principle: wherever the energy, momentum, or angular momentum of matter is changed under the influence of gravity, there

[30] To avoid a discussion of the boundary at infinity, suppose the plane-wave is in fact restricted to a spatially compact region; in this case, one will find that $x_{[2}t_{3]}{}^{\alpha} \neq 0$ at the boundary of the region, and the spatial integral of $x_{[2}t_{3]}{}^{0}$ will amount to the same total angular momentum described by $s^{0}{}_{23}$. In fact, it is generally true that (under suitable boundary conditions) $t_{\mu\nu}[\tau, s]$ gives the same global measure of energy-momentum and angular momentum as $\tau_{\mu\nu}$ and $s^{\alpha}{}_{\mu\nu}$; see chapter 5 for details.

must be an equal and opposite change in the energy, momentum, or angular momentum of the gravitational field. This idea, and the requirement that $\tau_{\mu\nu}$ be symmetric and free of second derivatives, was enough to determine the energy-momentum tensor (4.89) and the field condition (4.91). To determine the spin tensor uniquely, we demanded that it satisfy two physically-motivated conditions: first, the spin of a gravitational plane-wave must flow in the direction of propagation of the wave (4.24); second, a transverse-traceless field must possess a traceless spin tensor (4.34) and hence be free of infinite pressure gradients. In addition, the resulting spin tensor (4.91) displays a number of notable properties that were not required of it: the Newtonian potential has vanishing spin-tensor (4.28) and any transverse-traceless field carries purely spatial spin (4.27).

The microaverage, which defines the gauge invariant exchange of energy-momentum between gravity and an infinitesimal probe, also renders the exchange of spatial angular momentum gauge-invariant (4.51) without the need for gauge-fixing. In the same system, the exchange of non-spatial angular momentum can displace the centre-of-mass of the detector, beyond that which would be expected due to the exchange of linear momentum alone (4.75). Indeed, if the internal motions of the probe resonate with the incident wave, the detector may "walk" in a transverse direction, and acquire a net displacement over many wavelengths. We have explored this phenomenon for the specific example of a rotating rod (4.78) and rederived our predictions from first principles (see appendix 4.A).

Unlike $\tau_{\mu\nu}$, the spin tensor of a gravitational plane-wave is not completely independent of the TT-frame in which the wave is prepared. While the current-density of transverse spatial spin (in any frame) is invariant, the full tensor adapts so as to remain spatial with respect to whichever TT-frame is used. Thus, if a plane-wave region is used to sew together two TT-frames and produce a seamless picture of energy-momentum propagation, there will inevitably be a discontinuity in $s^{\alpha}{}_{\mu\nu}$ where the spin is projected from one spatial hypersurface to another; however, a spin psuedovector can be defined (4.87) that is conserved across this interface.

The spin carried by a plane-wave (4.82) is also an excellent example with which to compare our framework to the familiar "Belinfante" energy-momentum tensors of Landau and Lifshitz, and Arnowitt, Deser and Misner. Whereas $s^{\alpha}{}_{\mu\nu}$ describes spin that is manifestly present within the wave, the density of which depends on the rotational motion of $h_{\mu\nu}$ at each point, the Belinfante tensors assign all angular-momentum to the boundary of the wave, and its density there is not simply a function of $h_{\mu\nu}$ (as a truly local intrinsic property of the field would be) but is also dependent on the distance of the point from the origin.

Returning to chapter 3, it becomes clear that many of the remarkable properties of our gravitational energy-momentum tensor (including its positive energy-density and causal energy-flux) owe their existence to the careful separation of gravitational energy-momentum and gravitational spin. Now that we have made this separation explicit, and derived a formula for $s^{\alpha}{}_{\mu\nu}$, we have all the ingredients necessary to understand the broader theoretical picture in which our description resides. This is the task of chapter 5, the results of which, in many respects, are the main reward for our work here.

4.A Appendix: Moment of Energy Exchange from First Principles

In order to rederive equation (4.73) from first principles, we shall consider a detector, centred at the origin, composed of a set of N test-particles connected by some form of "light" mechanical apparatus.[31] The n^{th} particle has mass m_n, proper time τ_n, and follows a worldline $y_n^\mu(\tau_n)$ in the physical spacetime; its 4-velocity $u_n^\mu \equiv \mathrm{d}y_n^\mu/\mathrm{d}\tau_n$ has unit norm: $u_n^a u_n^b g_{ab} = -1$. In this approach, the detector is not truly infinitesimal, but we stipulate that the length-scale of the detector $l \sim y_n^i$ be sufficiently small that we may ignore terms $O(l^3)$ in our calculation, leaving us with a quadrupole approximation of the probe. As usual, a weak gravitational plane-wave is incident upon the detector, represented in transverse-traceless gauge in the background: $h_{\mu\nu} = h_{\mu\nu}^{\text{wave}}$ as given in (4.65). As $\partial_2 h_{\mu\nu} = \partial_3 h_{\mu\nu} = 0$, the physical spacetime is isometric under translations in the y^2 and y^3 directions of the $\{y^\mu\}$ coordinate system; for this section, we will reserve the index $k = 2, 3$ for these transverse directions.

In the physical spacetime, the energy-momentum tensor of the particles is

$$(T_{\text{particles}})^a{}_b = \sum_{n=1}^{N} \frac{1}{\sqrt{-g}} \int \mathrm{d}\tau_n \delta(y^\mu - y_n^\mu(\tau_n)) u_n^a p_{nb}, \tag{4.92}$$

where $p_{na} \equiv m_n g_{ab} u_n^b$ is the 4-momentum of the n^{th} particle. For the purposes of defining the moment-of-energy of the detector, we assume that the energy-momentum of the light apparatus is negligible:

$$X_i = -\int \sqrt{-g} (T_{\text{particles}})^0{}_0 y_i \mathrm{d}^3 y = -\sum_{n=1}^{N} p_{n0} y_{ni}.$$

In terms of background quantities, this is

$$X_i = -\sum_{n=1}^{N} m_n (\eta_{0\alpha} + h_{n0\alpha}) x_n'^\alpha x_{ni} = \sum_{n=1}^{N} m_n x_n'^0 x_{ni}, \tag{4.93}$$

where $x_n^\mu(\tau_n)$ are the coordinates of the particles in the background spacetime, primes indicate differentiation with respect to τ_n, and $h_{n\mu\nu} \equiv h_{\mu\nu}(x_n^\alpha(\tau_n))$ is the gravitational field evaluated at the n^{th} particle. The rate of change of the moment-of-energy is therefore

$$\dot{X}_i = \sum_{n=1}^{N} m_n \left(\frac{x_n''^0 x_{ni}}{x_n'^0} + x_{ni}' \right). \tag{4.94}$$

The normalisation of the 4-velocity,

$$-1 = -(x_n'^0)^2 + (h_{nij} + \delta_{ij}) x_n'^i x_n'^j, \tag{4.95}$$

[31] The adjective "light" is used to indicate that the total energy-momentum of the apparatus is negligible compared to that of the particles.

ensures that $x_n'^0 \sim O(1)$ and $x_n''^0 \sim O(l^2)$, so the first term (4.94) is $O(l^3)$ and can therefore be neglected. Consequently,

$$\ddot{X}_i = \sum_{n=1}^{N} \frac{m_n x''_{ni}}{x_n'^0} + O(l^3). \tag{4.96}$$

The accelerations x''_{ni} in (4.96) are be caused by both the gravitational field and the mechanical forces exerted on the particles by the apparatus. Our aim is to infer $\ddot{X}_i$ without assuming any detailed model of the latter. This might seem an impossible task, as it appears that we will need to know the motions of the particles (or the forces from the apparatus) to *first order* in $h_{\mu\nu}$ if we wish to calculate the first order contribution to $\ddot{X}_k$. Fortunately, because the apparatus is light, and the transverse momentum is conserved, only the *unperturbed* motions of the particles will be required. To see this, we start by calculating the linear momentum of the probe:

$$P_i = \int \sqrt{-g}(T_{\text{particles}})^0{}_i \mathrm{d}^3 y = \sum_{n=1}^{N} p_{ni}, \tag{4.97}$$

where once again we assume that the momentum of the apparatus can be neglected. Because the physical spacetime is isometric under translations in the y^2 and y^3 directions, the transverse momentum P_k will be conserved:[32]

$$0 = \dot{P}_k = \partial_t \left(\sum_{n=1}^{N} m_n (\delta_{ki} + h_{nki}) x_n'^i \right) = \sum_{n=1}^{N} \frac{m_n}{x_n'^0} \left(x''_{nk} + \partial_{\tau_n}(h_{nki} x_n'^i) \right), \tag{4.98}$$

which is equivalent to the statement that the mechanical forces on the particles (due to the apparatus) balance each another.[33] Substituting this constraint into equation (4.96) we find that

$$\begin{aligned} \ddot{X}_k &= -\sum_{n=1}^{N} \frac{m_n}{x_n'^0} \partial_{\tau_n}(h_{nki} x_n'^i) + O(l^3) \\ &= \partial_t \left(-\sum_{n=1}^{N} m_n h_{nki} x_n'^i \right) + O(l^3), \end{aligned} \tag{4.99}$$

which is easy to integrate:[34]

$$\dot{X}_k = -\sum_{n=1}^{N} m_n h_{nki} x_n'^i + O(l^3). \tag{4.100}$$

[32] This follows from the standard argument: $\partial_k g_{\alpha\beta} = 0$ guarantees that $(\partial_k)^a$ is a Killing vector, $\nabla^{(a}(\partial_k)^{b)} = 0$, and thus $0 = \sqrt{-g}\nabla_a (T^a{}_b (\partial_k)^b) = \partial_\alpha(\sqrt{-g} T^\alpha{}_k)$, the spatial integral of which is $\dot{P}_k = 0$.

[33] Although the total momentum of the apparatus is assumed to be negligible, we have not made any assumptions about the local flux of momentum between the apparatus and the particles, and so the individual mechanical forces on each particle cannot be neglected. The constraint (4.98) arises because the apparatus has much less mass than the particles, and so any momentum it were to gain would send it off with a very large velocity that would be impossible to maintain while in contact with the particles; in order to stay connected to the particles, the momentum of the apparatus must remain very small, and the forces acting on the apparatus must (approximately) sum to zero.

[34] The constant of integration is set to zero by the initial conditions: the probe is at rest ($\dot{X}_i = 0$) before the wave arrives ($h_{\mu\nu} = 0$).

This is the equation we sought: every instance of $x_n'^i$ is multiplied by $h_{\mu\nu}$, so only the unperturbed motions are needed to determine $\dot{X}_k$ to linear order in the gravitational field.

The last step is to relate the h_{nki} to the gravitational field at the origin:

$$\begin{aligned} h_{n\mu\nu} &= h_{\mu\nu}(t,\vec{0}) + x_n^i \partial_i h_{\mu\nu}(t,\vec{0}) + O(l^2) \\ &= B_{\mu\nu}^{\mathrm{TT}}(t) - x_n^1 \dot{B}_{\mu\nu}^{\mathrm{TT}}(t) + O(l^2); \end{aligned} \tag{4.101}$$

as a result, equation (4.100) becomes

$$\begin{aligned} \dot{X}_k &= -B_{ki}^{\mathrm{TT}} \left(\sum_{n=1}^{N} m_n x_n'^i \right) + \dot{B}_{ki}^{\mathrm{TT}} \left(\sum_{n=1}^{N} m_n x_n^1 x_n'^i \right) + O(l^3) \\ &= -B_{ki}^{\mathrm{TT}} \left(\dot{X}_k \right) + \dot{B}_{ki}^{\mathrm{TT}} \left(\sum_{n=1}^{N} m_n x_n^1 x_n'^i \right) + O(l^3) \\ &= \dot{B}_{ki}^{\mathrm{TT}} \left(\sum_{n=1}^{N} m_n x_n^1 x_n'^i \right) + O(h^2) + O(l^3). \end{aligned} \tag{4.102}$$

This simplifies even further when we notice that

$$\begin{aligned} \dot{I}_{ij} - L_{ij} &= \partial_t \left(-\int \sqrt{-g} (T_{\text{particles}})^0{}_0 y_i y_j \mathrm{d}^3 y \right) - 2 \int \sqrt{-g} (T_{\text{particles}})^0{}_{[j} y_{i]} \mathrm{d}^3 y \\ &= \partial_t \left(\sum_{n=1}^{N} m_n x_n'^0 x_n^i x_n^j \right) - 2 \sum_{n=1}^{N} m_n x_n'^{[j} x_n^{i]} + O(h) \\ &= 2 \sum_{n=1}^{N} m_n x_n'^{(i} x_n^{j)} - 2 \sum_{n=1}^{N} m_n x_n'^{[j} x_n^{i]} + O(h) + O(l^4) \\ &= 2 \sum_{n=1}^{N} m_n x_n^j x_n'^i + O(h) + O(l^4), \end{aligned} \tag{4.103}$$

which gives us our final result:

$$\dot{X}_k = \dot{B}_{ki}^{\mathrm{TT}} \left(\dot{I}_{i1} - L_{i1} \right) /2 + O(h^2) + O(l^3), \tag{4.104}$$

exactly as predicted by equation (4.73).

It should be clear that our formalism provides a much more direct route to this result: one needs only to produce (4.67) and integrate, a straight-forward operation that lacks the “insightful” steps of the first principles calculation, such as invoking conservation of transverse momentum (4.98) to remove degrees of freedom from (4.96). However, the moral of this appendix is not simply that our method is more computationally efficient; equally important is the *intuitive* power that our framework confers. Working from first principles, it is hard to imagine that one would have thought to derive (4.104) in the first place, as there is no obvious reason to expect that a gravitational wave would produce a transverse motion in the detector’s centre-of-mass. In comparison, our unified picture of local gravitational energetics brought this phenomenon to mind as readily as the exchange of energy, momentum, or angular momentum.

Chapter 5

Localised Energetics of Linear Gravity: Theoretical Development

5.1 Introduction

There are at least three ways to quantify the energy of a physical system. One approach is to consider interactions with a second system (the energy of which is known) and to seek a function-of-state of the first system which, by undergoing equal and opposite changes, accounts for the energy lost or gained by the second system. Alternatively, a Lagrangian for the physical system might be constructed, and the energy identified as the Noether charge associated with the Lagrangian's symmetry under translation in time. Thirdly, and most simply of all, one can "weigh" the system; the energy is then determined by the gravity it generates.

In chapters 3 and 4, we arrived at a local description of the energy, momentum, and angular momentum of the linearised gravitational field. The resulting gravitational energy-momentum tensor τ_{ab} and spin tensor ${s^a}_{bc}$ are particularly notable in that, whenever the field is transverse-traceless, they describe non-negative energy-density, causal energy flow, and traceless spatial spin; moreover, these properties, and the gauge invariant energetics of an infinitesimal probe, motivate a natural gauge-fixing procedure. These two tensors were derived by what is essentially the first method described above: we sought functions of the gravitational field which could account for the energy, momentum, and angular momentum exchanged locally with matter. The purpose of this present investigation is to explore the other two roles played by τ_{ab} and ${s^a}_{bc}$: as the Noether currents of local translations and rotations, and as self-interaction terms in the field equations that generate gravity alongside the energy-momentum of matter.

The chapter is organised as follows. In section 5.2, we construct a Lagrangian for linear gravity (a covariantisation of the Fierz-Pauli Lagrangian for a massless spin-2 field) and show that it generates τ_{ab} and ${s^a}_{bc}$ according to standard variational definitions of energy-momentum and spin. This process of "gauging" the translational and rotational symmetries of the background, and deriving τ_{ab} and ${s^a}_{bc}$ as functional derivatives of the Lagrangian with respect to the gauge fields, confirms their status as Noether currents of translational and rotational symmetry.[1] In section 5.3 we then demonstrate that, under a local redefinition of the gravitational field, τ_{ab} and $s^a{}_{bc}$ appear (combined into a single Belinfante energy-momentum tensor) as the *quadratic* part of the vacuum Einstein field equations. The same techniques are then applied to Einstein-Cartan gravity, with τ_{ab} and $s^a{}_{bc}$ appearing as the quadratic parts of two *separate* field equations. In both cases, τ_{ab} and $s^a{}_{bc}$ quantify the self-interaction of the gravitational field, and generate gravity in an identical fashion to material energy-momentum and spin. Finally, in section 5.4 we examine the significance of the field redefinitions used in section 5.3; in contrast to our tensors, we show that a *non-local* redefinition is required in order to cast the Landau-Lifshitz tensor [52] as a source of the gravitational field.

From all of this we learn that τ_{ab} and $s^a{}_{bc}$ stand on an equal footing with the energy-momentum and intrinsic spin of matter: they can be derived from the symmetries of a suitable Lagrangian, and behave as sources for the gravitational field. These developments solidify the tensors' physical interpretation, and embed them within the same theoretical apparatus that has been used to define gravitational energy-momentum in the past [7, 29, 35, 52]. Furthermore, through their role in the non-linear field equations, we gain insight into how τ_{ab} and $s^a{}_{bc}$ may be extended beyond the linear regime, and also uncover a new and possibly valuable set of gravitational field-variables.

Our notation and conventions are the same as those of chapters 3 and 4.[2] We will, however, introduce a new variety of index: Greek letters with overbars,[3] which we will use to enumerate the components of tensors in the non-holonomic basis defined by the tetrad $e^a_{\bar{\mu}}$; see appendix 5.A for details. As before, when working in the flat background spacetime $(\breve{\mathcal{M}}, \breve{g}_{ab})$ it will often be convenient to express our tensors in a coordinate system $\{x^\mu\}$ that is Lorentzian with respect to the background metric: $\breve{g}_{\mu\nu} = \eta_{\mu\nu}$.[4]

5.2 Lagrangian Formulation

The primary aim of this section is to reproduce the formulae obtained in chapters 3 and 4,

$$\kappa\bar{\tau}_{\mu\nu} = \tfrac{1}{4}\partial_\mu h_{\alpha\beta}\partial_\nu \bar{h}^{\alpha\beta}, \tag{5.1a}$$

$$\kappa s^\alpha{}_{\mu\nu} = 2\bar{h}_{\beta[\nu}\partial^{[\alpha}\bar{h}_{\mu]}{}^{\beta]}, \tag{5.1b}$$

[1] More precisely, the vectors $\tau^a{}_\mu$ are the Noether currents of translations in the x^μ direction, and $j_{\mu\nu}{}^a \equiv 2x_{[\mu}\tau_{\nu]}{}^a + s^a{}_{\mu\nu}$ are the Noether currents of rotations in the $x^\mu x^\nu$-plane.

[2] We work in units where $c = 1$, write $\kappa \equiv 8\pi G$, and use the sign conventions of Wald [79]: $\eta_{\mu\nu} \equiv \mathrm{diag}(-1,1,1,1)$, $[\nabla_c, \nabla_d]v^a \equiv 2\nabla_{[c}\nabla_{d]}v^a \equiv R^a{}_{bcd}v^b$, and $R_{ab} \equiv R^c{}_{acb}$. We use Roman letters as abstract tensor indices [79, §2.4] and Greek letters as numerical indices running from 0 to 3.

[3] These bars should not to be confused with those placed over 2-index tensors, which indicate trace-reversal: $\bar{h}_{ab} \equiv h_{ab} - \breve{g}_{ab}h/2$ and $\bar{\tau}_{ab} \equiv \tau_{ab} - \breve{g}_{ab}\tau/2$.

[4] In the next section we will need to implicitly vary $\breve{g}_{ab}$ in order to take functional derivatives with respect to the background tetrad $\breve{e}^a_{\bar{\mu}}$; when the background metric is curved, the coordinate system $\{x^\mu\}$ has no special properties.

from a Lagrangian $\mathcal{L}$ that also generates the field equations of linear gravity. Before we do this, however, we will have to make some accommodation for the harmonic gauge condition,

$$\partial^\mu \bar{h}_{\mu\nu} = 0, \tag{5.2}$$

which we have been diligently enforcing since its appearance as an unexpected consequence of the derivation of $\tau_{\mu\nu}$ in chapter 3. While this condition has been immensely valuable, naturally reducing the gauge ambiguity of our framework, it now has the potential to interfere with the arbitrary field variations that occur when taking functional derivatives of a Lagrangian. To avoid this problem, we shall temporarily relax the gauge condition (5.2) and aim to derive from $\mathcal{L}$ an energy-momentum tensor and spin tensor that *generalise* $\tau_{\mu\nu}$ and $s^\alpha{}_{\mu\nu}$ beyond the harmonic gauge, reducing to the familiar formulae (5.1) only once the harmonic condition is reintroduced. It is worth noting, however, that although the generalised forms of the tensors will be useful for technical reasons in later sections, they will not give us any further physical information than their restriction to harmonic gauge (5.1). This is because in order to interpret $\tau_{\mu\nu}$ and $s^\alpha{}_{\mu\nu}$ physically, we must first extinguish their gauge freedom; the only way to do this that produces sensible local properties (positive energy-density, causal energy-flow, and traceless spatial spin) is by insisting on transverse-traceless gauge, which obviously ensures that the harmonic condition is satisfied.

In addition to relaxing the harmonic condition, it will also be convenient to ignore matter ($T_{\mu\nu} = 0$) and work with gravity *in vacuo* for the entirety of this section. Even though the framework of our previous chapters was developed around the exchange of energy-momentum and angular momentum between matter and gravity, here we will be able to construct $\tau_{\mu\nu}$ and $s^\alpha{}_{\mu\nu}$ from the dynamics of the gravitational field alone.

5.2.1 The Fierz-Pauli Lagrangian

We begin in a flat background spacetime $(\check{\mathcal{M}}, \check{g}_{ab})$ with the Fierz-Pauli Lagrangian for a massless spin-2 field [37]:

$$\mathcal{L}_{\text{FP}} \equiv \frac{1}{4\kappa}\left(\partial_\mu h_{\alpha\beta}\partial^\mu \bar{h}^{\alpha\beta} - 2\partial_\mu \bar{h}^{\mu\alpha}\partial_\nu \bar{h}^\nu{}_\alpha\right). \tag{5.3}$$

From a non-gravitational standpoint, this Lagrangian can be derived by demanding invariance under the massless spin-2 gauge transformation:[5]

$$\delta h_{\mu\nu} = \partial_{(\mu}\xi_{\nu)} \quad \Rightarrow \quad \delta\mathcal{L}_{\text{FP}} = \text{surface terms.} \tag{5.4}$$

For our purposes, however, it suffices to observe that $\mathcal{L}_{\text{FP}}$ correctly reproduces the linearised vacuum Einstein field equations:

$$0 = \frac{\delta\mathcal{L}_{\text{FP}}}{\delta h_{\mu\nu}} = \frac{1}{\kappa}\widehat{G}^{\mu\nu\alpha\beta}h_{\alpha\beta}, \tag{5.5}$$

[5]Ignoring surface terms and an overall rescaling, $\mathcal{L}_{\text{FP}}$ is the unique scalar, quadratic in $\partial_\mu h_{\alpha\beta}$, which is invariant (up to surface terms) under $\delta h_{\mu\nu} = \partial_{(\mu}\xi_{\nu)}$; see [64] for a proof.

where

$$\widehat{G}_{\mu\nu}{}^{\alpha\beta}h_{\alpha\beta} \equiv G^{(1)}_{\mu\nu} \equiv \partial_\alpha\partial_{(\mu}h_{\nu)}{}^\alpha - \partial^2 h_{\mu\nu}/2 - \partial_\mu\partial_\nu h/2 + \eta_{\mu\nu}(\partial^2 h - \partial_\alpha\partial_\beta h^{\alpha\beta})/2 \tag{5.6}$$

is the linear part of the Einstein tensor when the physical metric g_{ab} is perturbed according to

$$\phi^* g_{ab} = \check{g}_{ab} + h_{ab}. \tag{5.7}$$

As before, $\phi : \mathcal{M} \to \check{\mathcal{M}}$ maps the physical spacetime $(\mathcal{M}, g_{ab})$ to the background.

To obtain an energy-momentum tensor and a spin tensor from $\mathcal{L}_{\mathrm{FP}}$, we will follow the standard definitions (5.121) from Einstein-Cartan gravity, a brief review of which can found in appendix 5.A.[6] We first need to covariantise $\mathcal{L}_{\mathrm{FP}}$. Invoking a tetrad $\check{e}^a_{\bar{\mu}}$ and spin connection $\check{\omega}_a{}^{\bar{\mu}\bar{\nu}}$ to "gauge" the translational and rotational symmetries of the background, we write the Fierz-Pauli Lagrangian in terms of quantities which are covariant under local translations and rotations:

$$\mathcal{L}'_{\mathrm{FP}} \equiv \frac{\check{e}}{4\kappa}\left(\check{D}_{\bar{\mu}}h_{\bar{\alpha}\bar{\beta}}\check{D}^{\bar{\mu}}\bar{h}^{\bar{\alpha}\bar{\beta}} - 2\check{D}_{\bar{\mu}}\bar{h}^{\bar{\mu}\bar{\alpha}}\check{D}_{\bar{\nu}}\bar{h}^{\bar{\nu}}{}_{\bar{\alpha}}\right), \tag{5.8}$$

where $\check{D}_a$ is a covariant derivatives with connection $\check{\omega}_a{}^{\bar{\mu}\bar{\nu}}$, the volume element $\check{e} \equiv 1/\det(\check{e}^a_{\bar{\mu}})$, and Greek indices with overbars enumerate the components of tensors in the non-holonomic basis $\{\check{e}^a_{\bar{\mu}}\}$.

So far, the background $(\check{\mathcal{M}}, \check{g}_{ab})$ is still flat and torsion-free: we have only rewritten the Fierz-Pauli Lagrangian in a more general language. We will soon need to perform arbitrary infinitesimal variations in $\check{e}^a_{\bar{\mu}}$ and $\check{\omega}_a{}^{\bar{\mu}\bar{\nu}}$, however, and in doing so we will inevitably explore backgrounds with curvature $\check{R}_{ab\bar{\mu}\bar{\nu}}$ and torsion $\check{\mathcal{T}}^a{}_{\bar{\mu}\bar{\nu}}$. For this reason, we must also decide how our Lagrangian should change when the background is no longer flat and torsion-free. The obvious response to this uncertainty is to follow the "minimal coupling" maxim, and insist that the Lagrangian remain as it is in equation (5.8) even when the background

[6] In standard general relativity, the rotational and translational symmetries of Minkowski spacetime are "gauged" into a single local symmetry: *diffeomorphism gauge invariance*. As a result, the energy-momentum and angular momentum of matter are all contained within a single *Belinfante* energy-momentum tensor T^{Bel}_{ab} (written simply as T_{ab} in chapters 3 and 4) which we derive according to Hilbert's definition: $T^{\mathrm{Bel}}_{ab} \equiv (1/\sqrt{-g})(\delta\mathcal{L}_{\mathrm{matter}}/\delta g^{ab})$. The Belinfante tensor includes intrinsic spin in so far as it integrates to give the correct *global* measure of angular momentum (see section 5.2.4) but as we saw in section 4.6 of the previous chapter, it cannot provide a physically sensible *local* description of the spin carried by a field. In order to derive two separate tensors, $\tau_{\mu\nu}$ and $s^\alpha{}_{\mu\nu}$, following a Hilbert-like approach, one must gauge the rotational and translational symmetries *separately*, and take derivatives of a Lagrangian with respect to the two gauge fields (5.121). These are the techniques of the Kibble/Sciama formulation of Einstein-Cartan gravity [17, 32, 43, 48, 53, 74], of which we give a brief summary in appendix 5.A. Einstein-Cartan gravity is a reformulation, and slight extension, of general relativity, in which spacetime torsion is generated by material intrinsic spin. As we will not require the spacetime to actually possess torsion, or matter to carry intrinsic spin, the results of this chapter do not depend on extending general relativity in this way. As we will see, however, torsion or no torsion, the *tetrad formalism* (that is, general relativity expressed in terms of a tetrad and a connection, rather than a metric) is often a more natural environment in which to understand $\tau_{\mu\nu}$ and $s^\alpha{}_{\mu\nu}$.

is curved and contorted. Despite the simplicity of this approach, the Lagrangian $\mathcal{L}'_{\mathrm{FP}}$ is actually a highly unnatural choice, as it deprives the field theory of its spin-2 gauge-invariance when the background is no longer flat. To see this, consider a curved vacuum background ($\check{R}_{a\bar{\mu}} = 0$, $\check{\mathcal{T}}^{a}{}_{\bar{\mu}\bar{\nu}} = 0$, $\check{R}_{ab\bar{\mu}\bar{\nu}} \neq 0$) and perform a covariantised spin-2 gauge transformation:

$$\delta h_{\bar{\mu}\bar{\nu}} = \check{D}_{(\bar{\mu}}\xi_{\bar{\nu})} \quad \Rightarrow \quad \delta\mathcal{L}'_{\mathrm{FP}} = -\frac{\check{e}}{\kappa}\check{D}^{\bar{\mu}}\xi^{\bar{\nu}}\check{R}_{\bar{\alpha}\bar{\mu}\bar{\nu}\bar{\beta}}\bar{h}^{\bar{\alpha}\bar{\beta}} + \text{surface terms.} \tag{5.9}$$

Thus, $\mathcal{L}'_{\mathrm{FP}}$ loses its spin-2 gauge invariance when one tries to extend the theory "minimally" beyond the flat background.

The gauge invariance of the field theory can be preserved, for vacuum backgrounds at least, if we allow $h_{\mu\nu}$ to couple directly the curvature of the background.[7] The Lagrangian (5.10) we will use to generate $\tau_{\mu\nu}$ and $s^{\alpha}{}_{\mu\nu}$ will do exactly this, although we should explain that it is not unique in this regard. If we had wanted to present the subsequent calculation as a genuine *ab initio* derivation of $\tau_{\mu\nu}$ and $s^{\alpha}{}_{\mu\nu}$, then we would need to justify our specialisation to (5.10) over the other possibilities. However, we have already derived $\tau_{\mu\nu}$ and $s^{\alpha}{}_{\mu\nu}$ from more concrete considerations (in chapters 3 and 4), and our aim here is only to show that a Lagrangian *exists* from which the tensors can be obtained. We will explore this curvature-coupling freedom in section 5.2.3, and by the end of section 5.4.2 we will be in a position to look back at $\mathcal{L}$ and better understand the significance of our "choice". For now, we shall simply write down our Lagrangian as an ansatz, justified by its being a covariantisation of $\mathcal{L}_{\mathrm{FP}}$ which preserves the field theory's gauge-invariance beyond the flat background, and proceed to calculate its energy-momentum tensor and spin tensor.

5.2.2 Energy-Momentum Tensor and Spin Tensor

Let us consider the following the Lagrangian for the linearised gravitational field:

$$\mathcal{L} \equiv \frac{\check{e}}{4\kappa}\left(\check{D}_{\bar{\mu}}h_{\bar{\alpha}\bar{\beta}}\check{D}^{\bar{\mu}}\bar{h}^{\bar{\alpha}\bar{\beta}} - 2\check{D}_{\bar{\mu}}\bar{h}^{\bar{\mu}\bar{\alpha}}\check{D}_{\bar{\nu}}\bar{h}^{\bar{\nu}}{}_{\bar{\alpha}} + 2\bar{h}^{\bar{\mu}\bar{\nu}}\check{R}_{\bar{\alpha}\bar{\mu}\bar{\nu}\bar{\beta}}\bar{h}^{\bar{\alpha}\bar{\beta}}\right). \tag{5.10}$$

This clearly reduces to the Fierz-Pauli Lagrangian (5.3) when the background is flat and torsion-free, and furthermore, successfully extends the spin-2 gauge-invariance of the theory to curved (vacuum) backgrounds:

$$\delta h_{\bar{\mu}\bar{\nu}} = \check{D}_{(\bar{\mu}}\xi_{\bar{\nu})} \quad \Rightarrow \quad \delta\mathcal{L} = \text{surface terms.} \tag{5.11}$$

Treating the fields $\{h^{\bar{\mu}\bar{\nu}}, \check{e}^{a}_{\bar{\mu}}, \check{\omega}_{a}{}^{\bar{\mu}\bar{\nu}}\}$ as independent variables,[8] we shall evaluate the

[7] No set of curvature-coupling terms (nor torsion-coupling terms) can extend the theory's gauge invariance to include *non-vacuum* backgrounds. This comes as no surprise, considering that we are studying a Lagrangian $\mathcal{L}_{\mathrm{FP}}$ that does not include matter. Evidently, the linearised vacuum field equations (5.5) can only be expected to be consistent when they describe perturbations from a vacuum background.

[8] It is slightly unusual to consider $h^{\bar{\mu}\bar{\nu}}$ as the independent field variable, rather than h_{ab}; however, the only difference between the two approaches are terms proportional to $\delta\mathcal{L}/\delta h_{ab}$ that appear in the energy-momentum tensor, and these vanish on the field equations (5.5) anyway.

energy-momentum tensor and spin tensor of $\mathcal{L}$ according to their definitions from Einstein-Cartan gravity:

$$\tau_a{}^\mu \equiv \left(\frac{1}{2\check{e}}\frac{\delta\mathcal{L}}{\delta\check{e}^a_{\bar{\mu}}}\right)_{\substack{\check{e}=\delta\\ \check{\omega}=0}}, \qquad s^a{}_{\mu\nu} \equiv \left(\frac{1}{\check{e}}\frac{\delta\mathcal{L}}{\delta\check{\omega}_a{}^{\bar{\mu}\bar{\nu}}}\right)_{\substack{\check{e}=\delta\\ \check{\omega}=0}}, \tag{5.12}$$

where the subscripts $\check{e}=\delta$ and $\check{\omega}=0$ signify that, once the functional derivatives have been taken, the background returns to its former state (flat and torsion-free) and the tetrad and spin connection become trivial (5.123) to reflect this. Substituting (5.10) into (5.12) we arrive at the following formulae for the energy-momentum tensor and spin tensor of the linearised gravitational field:

$$\kappa\bar{\tau}_{\mu\nu} = \tfrac{1}{4}\partial_\mu h_{\alpha\beta}\partial_\nu \bar{h}^{\alpha\beta} - \tfrac{1}{2}\partial_\mu \bar{h}_{\nu\alpha}\partial_\beta \bar{h}^{\alpha\beta}, \tag{5.13a}$$

$$\kappa s^\alpha{}_{\mu\nu} = 2\bar{h}_{\beta[\nu}\partial^{[\alpha}\bar{h}_{\mu]}{}^{\beta]} + \delta^\alpha_{[\nu}\bar{h}_{\mu]}{}^\beta\partial_\gamma\bar{h}^\gamma{}_\beta. \tag{5.13b}$$

This is precisely the result we needed: $\mathcal{L}$ has generated an energy-momentum tensor $\tau_{\mu\nu}$ and spin tensor $s^\alpha{}_{\mu\nu}$ which reduce to the familiar formulae (5.1) when the harmonic condition (5.2) is reintroduced.

We have achieved the main aim of this section, demonstrating that our energy-momentum tensor and spin tensor can be identified as translational and rotational Noether currents of a Lagrangian for linear gravity. In addition, the equations (5.13) reveal how our tensors (5.1) generalise beyond harmonic gauge. Before we study these generalised tensors in detail, we shall first examine the freedom that was available in our choice of covariantisation of $\mathcal{L}_{\text{FP}}$, and demonstrate that the formulae (5.13) constitute a suitably unique extension of (5.1).

5.2.3 Background Coupling and Superpotentials

We begin by considering the most general Lagrangian, quadratic in $h_{\mu\nu}$ and second-order in derivatives, which differs from the minimally coupled Lagrangian (5.8) only by terms which couple $h_{\mu\nu}$ to background curvature; ignoring surface terms, this is

$$\mathcal{L}_{\check{R}} \equiv \mathcal{L}'_{\text{FP}} + \frac{\check{e}}{2\kappa}\check{R}_{\bar{\alpha}\bar{\beta}}{}^{\bar{\mu}\bar{\nu}}\Sigma^{\bar{\alpha}\bar{\beta}}{}_{\bar{\mu}\bar{\nu}}, \tag{5.14}$$

where $\Sigma^{\alpha\beta}{}_{\mu\nu} = -\Sigma^{\beta\alpha}{}_{\mu\nu} = -\Sigma^{\alpha\beta}{}_{\nu\mu} = \Sigma_{\mu\nu}{}^{\beta\alpha}$ is a local quadratic Lorentz-covariant function of $h_{\mu\nu}$, the general form of which can be parametrised by five dimensionless constants $\{A_n\}$:

$$\Sigma^{\alpha\beta}{}_{\mu\nu} \equiv A_1 h^\alpha_{[\mu}h^\beta_{\nu]} + A_2 h h^{[\alpha}_{[\mu}\delta^{\beta]}_{\nu]} + A_3 h^\gamma_{[\mu}\delta^{[\beta}_{\nu]}h^{\alpha]}_\gamma + \delta^\alpha_{[\mu}\delta^\beta_{\nu]}\left(A_4 h^2 + A_5 h_{\gamma\delta}h^{\gamma\delta}\right). \tag{5.15}$$

If we recall the behaviour of $\mathcal{L}'_{\text{FP}}$ under a spin-2 gauge transformation (5.9), it is immediately clear that the field theory will retain its gauge invariance for curved vacuum backgrounds (5.11) if and only if $A_1 = -1$.

Inserting $\mathcal{L}_{\check{R}}$ into (5.12), we find that the energy-momentum tensor of this Lagrangian is identical to the tensor (5.13a) derived from $\mathcal{L}$, but that the spin tensor is given by

$$\kappa s^\alpha{}_{\mu\nu} = h_{\beta[\nu}\partial^\alpha h_{\mu]}{}^\beta + \delta^\alpha_{[\nu}\bar{h}_{\mu]}{}^\beta\partial_\gamma\bar{h}^\gamma{}_\beta + \bar{h}^\alpha{}_{[\mu}\partial^\beta\bar{h}_{\nu]\beta} + \partial_\beta\Sigma^{\alpha\beta}{}_{\mu\nu}. \tag{5.16}$$

In harmonic gauge this becomes

$$\kappa s^{\alpha}{}_{\mu\nu} = h_{\beta[\nu}\partial^{\alpha}h_{\mu]}{}^{\beta} + \partial_{\beta}\Sigma^{\alpha\beta}{}_{\mu\nu}, \tag{5.17}$$

revealing that $\{A_n\}$ are the very same constants that parameterised the *superpotential* freedom of $s^{\alpha}{}_{\mu\nu}$ in section 4.3 of the previous chapter. There, the value $A_1 = -1$ was derived by demanding that $s^{\alpha}{}_{\alpha\nu} = 0$ for all transverse-traceless $h_{\mu\nu}$, thereby ridding the gravitational field of infinite pressure gradients. Now we see that this special value of A_1 has a second significance: it ensures that the spin-2 gauge-invariance of the linearised theory extends beyond the flat background.[9]

Equation (5.17) also demonstrates that the parameters $\{A_n\}$ must take the values

$$A_1 = -1, \qquad A_2 = 1, \qquad A_4 = -1/4, \qquad A_3 = A_5 = 0, \tag{5.18}$$

(as they do in chapter 4) if the spin tensor (5.16) is to reduce to its original form (5.1b) in harmonic gauge; thus the freedom to add curvature terms (5.14) cannot, by itself, produce any other generalisation of $\tau_{\mu\nu}$ and $s^{\alpha}{}_{\mu\nu}$ than (5.13).

Now that we understand the role played by curvature terms in the Lagrangian, we must also explore the possibility of coupling $h_{\mu\nu}$ to background torsion. If the Lagrangian is to remain quadratic in $h_{\mu\nu}$ and second-order in derivatives, the only contribution we need to consider is

$$\Delta\mathcal{L} \equiv -\frac{\breve{e}}{\kappa}\breve{\mathcal{T}}^{a}{}_{\bar{\mu}\bar{\nu}}\Sigma_{a}{}^{\bar{\mu}\bar{\nu}}, \tag{5.19}$$

where $\Sigma_a{}^{\mu\nu} = -\Sigma_a{}^{\nu\mu}$ is composed of terms of the form hDh.[10] The torsion terms generate the superpotential freedom of the energy-momentum tensor:

$$\kappa\Delta\tau_a{}^{\mu} = \left(\frac{\kappa}{2\breve{e}}\frac{\delta\Delta\mathcal{L}}{\delta\breve{e}^{a}_{\bar{\mu}}}\right)_{\substack{\breve{e}=\delta\\ \breve{\omega}=0}} = \partial_{\nu}\Sigma_a{}^{\mu\nu}, \tag{5.20}$$

which is also accompanied by a change in the spin tensor,

$$\kappa\Delta s^{a}{}_{\mu\nu} = \left(\frac{\kappa}{\breve{e}}\frac{\delta\Delta\mathcal{L}}{\delta\breve{\omega}_a{}^{\bar{\mu}\bar{\nu}}}\right)_{\substack{\breve{e}=\delta\\ \breve{\omega}=0}} = 2\Sigma_{[\mu\nu]}{}^{a}. \tag{5.21}$$

Because the energy-momentum superpotentials are of the form $\partial(h\partial h)$, containing tensors of the form $h\partial^2 h$, their addition has the potential to spoil the homogeneous differential structure of (5.13a): $\tau_{\mu\nu} \sim \partial h\partial h$. In fact, there is no superpotential $\partial_\nu\Sigma_a{}^{\mu\nu}$, entirely composed of terms $\partial h\partial h$, which vanishes in harmonic gauge.[11] Thus, the freedom to add

[9] In section 5.4 of this chapter, we will also see that $A_1 = -1$ allows $\tau_{\mu\nu}$ and $s^{\alpha}{}_{\mu\nu}$ to be cast as the source terms of the quadratic field equations using a *local* redefinition of the gravitational field.

[10] Terms of the form $e\breve{\mathcal{T}}^2h^2$ are also allowable, but any contribution they make to $\tau_{\mu\nu}$ and $s^{\alpha}{}_{\mu\nu}$ will be $O(\breve{\mathcal{T}})$, vanishing when the background returns to being Minkowski spacetime. Also, note that no term can couple $h_{\mu\nu}$ to both torsion and curvature, as this would be at least third-order in derivatives.

[11] To prove this, construct the most general 2-index Lorentz-covariant tensor, composed entirely of terms of the form $\partial h\partial h$, which is at least linear in $\partial_\mu\bar{h}^{\mu\nu}$, and suppose that it is also a superpotential: $\kappa\Delta\tau_{\mu\nu} = \partial_\alpha\bar{h}^{\alpha\beta}(C_1\partial_\mu\bar{h}_{\nu\beta} + C_2\partial_\nu\bar{h}_{\mu\beta} + C_3\partial_\beta\bar{h}_{\mu\nu} + \eta_{\mu\nu}(C_4\partial_\beta h + C_5\partial_\gamma\bar{h}^{\gamma}{}_{\beta})) + C_6\partial_\alpha\bar{h}^{\alpha}{}_{\mu}\partial_\nu h + C_7\partial_\alpha\bar{h}^{\alpha}{}_{\nu}\partial_\mu h + C_8\partial_\alpha\bar{h}^{\alpha}{}_{\mu}\partial_\beta\bar{h}^{\beta}{}_{\nu}$, where $\{C_n\}$ are arbitrary dimensionless constants. Equation (5.20) informs us that $\partial_\nu\Delta\tau_\mu{}^{\nu} = 0$ for all $h_{\mu\nu}$; the only values of $\{C_n\}$ consistent with this are $C_n = 0$.

torsion terms to the Lagrangian is nullified by our insistence that the generalised $\tau_{\mu\nu}$ be free of second derivatives, and reduce to our original formula when the harmonic condition is enforced.

We therefore conclude that our Lagrangian (5.10) is the *unique* covariantisation of $\mathcal{L}_{\mathrm{FP}}$, quadratic in $h_{\mu\nu}$ and second-order in derivatives, which according to the definitions (5.12) generates an energy-momentum tensor that is free from second derivatives, and an energy-momentum tensor and spin tensor which agree with our original formulae (5.1) in harmonic gauge. Consequently, the resulting energy-momentum tensor and spin tensor (5.13) are the unique extensions of $\tau_{\mu\nu}$ and $s^{\alpha}{}_{\mu\nu}$ beyond harmonic gauge, which can be derived from a covariantised Fierz-Pauli Lagrangian according (5.12), and which do not introduce terms of the form $h\partial^2 h$ into $\tau_{\mu\nu}$.

Having demonstrated that (5.13) are indeed the unique extension of (5.1) beyond the harmonic gauge, it will be useful to review the basic properties of these generalised tensors, and construct their Belinfante tensor.

5.2.4 Basic Properties and Belinfante Tensor

Although the generalised energy-momentum tensor and spin tensor (5.13) can only be interpreted physically once the harmonic condition (and then transverse-traceless gauge) has been enforced, it will still be valuable to study their mathematical properties in this broader context, and compute the Belinfante tensor they define.

As our first observation, it is interesting to note that the gravitational energy-momentum tensor is *asymmetric* outside harmonic gauge: $\tau_{[\mu\nu]} \neq 0$. This is the usual context in which one encounters a spin tensor: the asymmetry of an energy-momentum tensor necessitates the existence of a spin tensor, as otherwise finite torques would act on infinitesimal regions of space [58, §5.7]. The framework we have developed is slightly unconventional in this regard, as originally the existence of $s^{\alpha}{}_{\mu\nu}$ was inferred by the exchange of angular momentum with matter, rather than any asymmetry in $\tau_{\mu\nu}$. Even though a symmetric energy-momentum tensor can be paired with a non-zero spin tensor (5.1) without contadiction,[12] it is perhaps reassuring to know that the moment we venture beyond harmonic gauge, the conventional inference returns: $\tau_{[\mu\nu]} \neq 0 \;\Rightarrow\; s^{\alpha}{}_{\mu\nu} \not\equiv 0$.

We should also review a few basic properties that the generalised $\tau_{\mu\nu}$ and $s^{\alpha}{}_{\mu\nu}$ possess by virtue of their definitions (5.12). We begin by observing that the energy-momentum of the linearised gravitational field is conserved when the vacuum field equations (5.5) are obeyed:

$$\partial^{\nu}\tau_{\mu\nu} = -\frac{1}{2\kappa}(\partial_{\mu}h^{\alpha\beta})\widehat{G}_{\alpha\beta}{}^{\gamma\delta}h_{\gamma\delta} = 0. \tag{5.22}$$

Next we find that the divergence of the spin tensor is given by

$$\partial_{\alpha}s^{\alpha}{}_{\mu\nu} = 2\tau_{[\mu\nu]} + \frac{2}{\kappa}h_{\alpha[\mu}\widehat{G}_{\nu]}{}^{\alpha\beta\gamma}h_{\beta\gamma}, \tag{5.23}$$

[12] The contradiction lies in having an *asymmetric* energy-momentum tensor and *not having* a spin tensor.

and hence, once the field equations have been applied,

$$\partial_\alpha(2x_{[\mu}\tau_{\nu]}{}^\alpha + s^\alpha{}_{\mu\nu}) = 0. \tag{5.24}$$

This confirms that the angular momentum current densities, defined in section 4.2 of chapter 4,

$$j_{\mu\nu}{}^\alpha \equiv 2x_{[\mu}\tau_{\nu]}{}^\alpha + s^\alpha{}_{\mu\nu}, \tag{5.25}$$

are also conserved as a consequence the field equations:

$$\partial_\alpha j_{\mu\nu}{}^\alpha = 0. \tag{5.26}$$

Of course, $\tau_{\mu\nu}$ and $j_{\mu\nu}{}^\alpha$ are only conserved because we are working with the *vacuum* field equations (5.5): there is no matter with which to exchange energy-momentum and angular momentum.

Because $\tau_{\mu\nu}$ and $s^\alpha{}_{\mu\nu}$ obey the properties displayed above, it is possible to construct a *Belinfante* energy-momentum tensor [15],

$$t_{\mu\nu} \equiv \tau_{\mu\nu} + \partial_\alpha(s_{\mu\nu}{}^\alpha + s_{\nu\mu}{}^\alpha - s^\alpha{}_{\mu\nu})/2, \tag{5.27}$$

which is symmetric by virtue of the field equations,

$$t_{[\mu\nu]} = \tau_{[\mu\nu]} - \partial_\alpha s^\alpha{}_{\mu\nu}/2 = \frac{1}{\kappa} h_{\alpha[\nu}\widehat{G}_{\mu]}{}^{\alpha\beta\gamma}h_{\beta\gamma} = 0,$$

and also conserved:

$$\partial^\nu t_{\mu\nu} = \partial^\nu \tau_{\mu\nu} = -\frac{1}{2\kappa}(\partial_\mu h^{\alpha\beta})\widehat{G}_{\alpha\beta}{}^{\gamma\delta}h_{\gamma\delta} = 0. \tag{5.28}$$

Furthermore, provided surface terms are negligible, the Belinfante tensor defines precisely the same *global* measure of energy, momentum, and angular momentum as $\tau_{\mu\nu}$ and $s^\alpha{}_{\mu\nu}$:

$$\int t_\mu{}^0 \mathrm{d}^3x = \int \tau_\mu{}^0 \mathrm{d}^3x, \tag{5.29a}$$

$$\int 2x_{[\mu}t_{\nu]}{}^0 \mathrm{d}^3x = \int (2x_{[\mu}\tau_{\nu]}{}^0 + s^0{}_{\mu\nu})\mathrm{d}^3x. \tag{5.29b}$$

The advantage of this Belinfante description is obvious: it combines energy-momentum and spin into a single symmetric tensor.[13] This apparent simplicity comes at a high price, however, because although the global picture remains intact (5.29) the Belinfante tensor is unable to reproduce the physically sensible *local* description that $\tau_{\mu\nu}$ and $s^\alpha{}_{\mu\nu}$ provide.

In general, the intermixture of spin and energy-momentum in (5.27) prevents us from localising the two quantities separately, and we are left with angular momentum currents

[13] That said, as far as our framework is concerned, the symmetry of the Belinfante tensor is not particularly impressive: our energy-momentum tensor $\tau_{\mu\nu}$ is already symmetric, by virtue of the harmonic condition (5.2) rather than the field equations.

$x_{[\mu}t_{\nu]}{}^{\alpha}$ which "contain" spin but do not assign it a local current,[14] and energy-momentum currents $t_{\mu}{}^{\alpha}$ which display *negative* energy-densities and *non-causal* energy-flux. Furthermore, because the gravitational Belinfante tensor has no special geometric or algebraic properties in either harmonic or transverse traceless gauge, it becomes impossible to justify a natural gauge-fixing programme. The tensor $t_{\mu\nu}$ can then be evaluated over the entire gauge space of $h_{\mu\nu}$, and will depend on the arbitrary mapping $\phi : \mathcal{M} \to \check{\mathcal{M}}$ as much as it depends on the physical properties of the gravitational field.

For these reasons, we cannot advocate interpreting $t_{\mu\nu}$ as the "true" energy-momentum of the gravitational field. The tensors $\tau_{\mu\nu}$ and $s^{\alpha}{}_{\mu\nu}$ are the local measures of gravitational energy-momentum and spin, describing positive energy-density, causal energy-flux, and traceless spatial spin; $t_{\mu\nu}$ is a derived quantity which packages spin and energy-momentum into a single object, losing some local information in the process.[15] The main application of the gravitational Belinfante tensor will arise in the next section, where we will also gain some insight into its physical interpretation. In brief, we will see that $t_{\mu\nu}$ appears as the quadratic contribution to the Einstein field equations, generating perturbations in the metric alongside the (Belinfante) energy-momentum of matter. In other words, it is the particular combination of energy-momentum and spin, $\tau_{\mu\nu} + \partial_{\alpha}(s_{\mu\nu}{}^{\alpha} + s_{\nu\mu}{}^{\alpha} - s^{\alpha}{}_{\mu\nu})/2$, that curves physical spacetime in a quadratic approximation to general relativity. It would be implausible to expect $\tau_{\mu\nu}$ alone to fulfill this role, as there is no other field equation in which $s^{\alpha}{}_{\mu\nu}$ could act as the source; only by considering perturbations in the Einstein-Cartan equations, as we do in section 5.3.2, will we find a setting in which $\tau_{\mu\nu}$ and $s^{\alpha}{}_{\mu\nu}$ arise as the self-interaction source-terms in two separate field equations.

In addition to its place in the Einstein field equations, $t_{\mu\nu}$ will also prove useful when we wish to compare our approach with the traditional treatments of gravitational energy-momentum, such as the Landau-Lifshitz tensor (see section 5.4.3) or the ADM energy-

[14] Even when the orbital angular momentum currents vanish ($x_{[i}\tau_{j]}{}^{\alpha} = 0$) the Belinfante "spin" currents $x_{[i}t_{j]}{}^{\alpha} \sim x\partial s$ still depend explicitly on the coordinates $\{x^{\mu}\}$; hence they cannot be interpreted as an intrinsic property of the gravitational field. As we saw in section 4.6 of the previous chapter, this defect is exemplified by the (harmonic-gauge) gravitational plane-wave, the Belinfante spin of which lies entirely on the boundary, with magnitude that is proportional to the coordinate distance from the origin.

[15] Historically, the Belinfante tensor has been preferred over the separate localisation of energy-momentum and spin. This is probably due to the privileged status of the Belinfante tensor of the electromagnetic field A_{μ}, which is invariant under the spin-1 gauge transformation $\delta A_{\mu} = \partial_{\mu}\lambda$, while (generically) the separate energy-momentum tensor and spin tensor are not. Of course, this criterion cannot be extended to the gravitational field, as there is no gravitational energy-momentum tensor, Belinfante or otherwise, which is invariant under the gauge transformation $\delta h_{\mu\nu} = \partial_{(\mu}\xi_{\nu)}$. In many respects, the electromagnetic field is a special case. Following the same methodology we used in this section, one can isolate the unique covariantised electromagnetic Lagrangian which (in contrast to a generic covariantisation) preserves spin-1 gauge invariance for arbitrary backgrounds. On applying the Einstein-Cartan definitions (5.121) to this particular covariantisation, one finds that the energy-momentum tensor is identical to the (gauge-invariant) Belinfante tensor, and that the spin tensor vanishes. This peculiar situation arises because the electromagnetic Lagrangian can be written in terms of differential forms, and so can be fully covariantised without introducing the spin connection: $\mathcal{L}_{\mathrm{EM}} \propto (dA)_{ab}(dA)_{cd}e^{a}_{\bar{\mu}}e^{\bar{\mu}c}e^{b}_{\bar{\nu}}e^{\bar{\nu}d}$. In contrast to electromagnetism, the spin tensor of the spin-1/2 field *does not* vanish, and both the energy-momentum tensor and spin tensor are gauge invariant [31].

momentum, which is discussed in appendix 5.C. As these approaches do not contain a gravitational spin tensor, they must also constitute Belinfante-style descriptions of gravitational energetics.

Proceeding with the calculation, we substitute the generalised tensors (5.13) into the definition (5.27) in order to obtain the Belinfante tensor associated with our framework. The resulting formula can be expressed most compactly in terms of the trace-reverse of $t_{\mu\nu}$:

$$\begin{aligned}\kappa\bar{t}_{\mu\nu} = &\frac{1}{4}\partial_\mu h_{\alpha\beta}\partial_\nu\bar{h}^{\alpha\beta} + \frac{1}{2}\partial_\alpha\bar{h}^{\alpha\beta}\left(\partial_{(\mu}\bar{h}_{\nu)\beta} - \partial_\beta h_{\mu\nu}\right) + \frac{1}{2}\partial_\alpha\bar{h}_{\beta(\mu}\left(\partial^\beta\bar{h}_{\nu)}{}^\alpha - \partial_{\nu)}\bar{h}^{\alpha\beta}\right)\\ &+ \frac{1}{2}\bar{h}^{\alpha\beta}\left(\partial_\alpha\partial_{(\mu}\bar{h}_{\nu)\beta} - \partial_\alpha\partial_\beta h_{\mu\nu}\right) + \frac{1}{2}\bar{h}_{\beta(\mu}\partial^\beta\partial^\alpha\bar{h}_{\nu)\alpha} + h_{\alpha[\nu}\widehat{G}_{\mu]}{}^{\alpha\beta\gamma}h_{\beta\gamma}.\end{aligned} \tag{5.30}$$

Although the last term can be removed by applying the field equations, we will retain it for the sake of generality. Incidentally, this calculation reveals another practical disadvantage of the Belinfante tensor: algebraic complexity. Not only is $t_{\mu\nu}$ composed of many more terms than the individual tensors $\tau_{\mu\nu}$ and $s^\alpha{}_{\mu\nu}$ combined, it also includes instances of second derivatives. As a result, computations involving $t_{\mu\nu}$ will often be considerably more demanding that those involving $\tau_{\mu\nu}$, or $s^\alpha{}_{\mu\nu}$, or both.

This concludes our analysis of the Lagrangian formulation of $\tau_{\mu\nu}$ and $s^\alpha{}_{\mu\nu}$. Armed with the results of this section, we are now in a position to "weigh" the gravitational field, and investigate the role our tensors play in the non-linear field equations.

5.3 Self-interaction in the Gravitational Field Equations

Here we examine how $\tau_{\mu\nu}$ and $s^\alpha{}_{\mu\nu}$ occur as the quadratic terms in a perturbative expansion of the Einstein field equations and also the Einstein-Cartan field equations, generating gravity in exactly the same fashion as material energy-momentum and spin.

As we move from a linear theory of gravity to a quadratic one, it will become important to fix the definition of $h_{\mu\nu}$ more precisely. Until this point, $h_{\mu\nu}$ has been used to signify a perturbation in the physical metric:

$$\phi^* g_{ab} = \breve{g}_{ab} + h_{ab}. \tag{5.31}$$

However, because our framework is based exclusively on gravity in the linear approximation, we could have defined $h_{\mu\nu}$ such that

$$\phi^* g_{ab} = \breve{g}_{ab} + h_{ab} + O(h^2), \tag{5.32}$$

and arrived at the very same results. For instance, suppose we had decided to work with the field $h'_{\mu\nu}$ that defines a (negative) perturbation in the *inverse* metric:

$$\phi^* g^{ab} = \breve{g}^{ab} - h'^{ab}. \tag{5.33}$$

The equation for the metric would then have been

$$\phi^* g_{ab} = \breve{g}_{ab} + h'_{ab} + h'_{ac}h'^c{}_b + O(h^3), \tag{5.34}$$

instead of (5.31), but because $h'_{\mu\nu} = h_{\mu\nu} + O(h^2)$, the linearised theory of $h'_{\mu\nu}$ would be the same as $h_{\mu\nu}$, and our framework would assign the same tensors (5.13) to describe its energy-momentum and spin. Only once we came to study the quadratic approximation of the field equations, as we do now, would any mathematical difference between the fields $h_{\mu\nu}$ and $h'_{\mu\nu}$ have been observed.

For the sake of concreteness, we will start with the standard definition of the gravitational field (5.31) and consider this to be true to all orders of approximation. As we will soon see, however, it is precisely the freedom to make field redefinitions of the form (5.32) that will allow us to cast $\tau_{\mu\nu}$ and $s^\alpha{}_{\mu\nu}$ as the sources of the gravitational field; by the end of the next section, we will have uncovered a new definition of $h_{\mu\nu}$, valid to quadratic order, that is specially selected by our local description of gravitational energetics. We will examine the wider significance of this definition, and the effects of field redefinition in general, in section 5.4.

5.3.1 The Einstein Equations

Consider the vacuum Einstein field equations (in the physical spacetime) expressed in terms of the "mixed" Einstein tensor density:

$$\sqrt{-g}G_a{}^b = 0. \tag{5.35}$$

Mapping this equation to the background, we apply (5.31) to every instance of $\phi^* g_{ab}$, and expand the result in powers of $h_{\mu\nu}$:

$$\widehat{G}_{\mu\nu}{}^{\alpha\beta} h_{\alpha\beta} + \widetilde{G}^{(2)}_{\mu\nu} + O(h^3) = 0, \tag{5.36}$$

where

$$\widetilde{G}^{(2)}{}_a{}^b \equiv \left[\phi^*(\sqrt{-g}G_a{}^b)\right]^{(2)} \tag{5.37}$$

is the quadratic part of the mixed Einstein tensor density.

Let us now redefine the gravitational field $h_{\mu\nu}$ by making the replacement

$$h_{\mu\nu} \to h_{\mu\nu} + h_{\mu\alpha} h^\alpha{}_\nu/2; \tag{5.38}$$

this gives rise to a corresponding change in the definition of the metric,

$$\phi^* g_{ab} = \check{g}_{ab} + h_{ab} + h_{ac} h^c{}_b/2, \tag{5.39}$$

and causes the vacuum equations (5.36) to become

$$\widehat{G}_{\mu\nu}{}^{\alpha\beta} h_{\alpha\beta} + \widehat{G}_{\mu\nu}{}^{\alpha\beta}(h_{\alpha\gamma} h^\gamma{}_\beta)/2 + \widetilde{G}^{(2)}_{\mu\nu} = 0, \tag{5.40}$$

when working to second order.[16] Moving all quadratic terms to the right-hand side, and making use of the following identity,

$$-\widetilde{G}^{(2)}_{\mu\nu} - \widehat{G}_{\mu\nu}{}^{\alpha\beta}(h_{\alpha\gamma} h^\gamma{}_\beta)/2 = \kappa t_{\mu\nu}, \tag{5.41}$$

[16] To clarify: the tensor $\widetilde{G}^{(2)}_{\mu\nu}$ is still the quadratic part of $\phi^*(\sqrt{-g}G_a{}^b)$ when the metric is expanded according to (5.31); this tensor has exactly the same formula in terms of the new $h_{\mu\nu}$ as it did the old because the replacement (5.38) only alters $\widetilde{G}^{(2)}_{\mu\nu}$ by quantities $O(h^3)$, which we neglect.

which is derived in appendix 5.B, we find that the quadratic vacuum field equations (5.40) are equivalent to

$$\widehat{G}_{\mu\nu}{}^{\alpha\beta} h_{\alpha\beta} = \kappa t_{\mu\nu}, \tag{5.42}$$

where $t_{\mu\nu}$ is the Belinfante energy-momentum tensor of the gravitational field (5.30) constructed from $\tau_{\mu\nu}$ and $s^{\alpha}{}_{\mu\nu}$.

Equation (5.42) is exactly what we had hoped to find: gravitational energy-momentum generates gravity in exactly the same fashion as material energy-momentum. To make this comparison transparent, we remind ourselves of the *non-vacuum* field equations at linear order:

$$\widehat{G}_{\mu\nu}{}^{\alpha\beta} h_{\alpha\beta} = \kappa T^{\text{Bel}}_{\mu\nu}, \tag{5.43}$$

where $T^{\text{Bel}}_{\mu\nu}$ (written simply as $T_{\mu\nu}$ in chapters 3 and 4) is the Belinfante energy-momentum tensor of matter, mapped to the background. Because $T^{\text{Bel}}_{\mu\nu}$ is Belinfante, any intrinsic spin carried by matter must be packaged inside this tensor according to the same formula (5.27) that defines the Belinfante tensor of the gravitational field, making the analogy with $t_{\mu\nu}$ extremely close. Furthermore, if one assumes that $T^{\text{Bel}}_{\mu\nu}$ is of the same order of magnitude as $t_{\mu\nu} \sim O(h^2)$, then at quadratic order the non-vacuum version of (5.42) is in fact

$$\widehat{G}_{\mu\nu}{}^{\alpha\beta} h_{\alpha\beta} = \kappa (t_{\mu\nu} + T^{\text{Bel}}_{\mu\nu}), \tag{5.44}$$

wherein the source of the gravitational field is the sum of the material and gravitational Belinfante tensors.[17]

It goes without saying that equation (5.42) can also be written as

$$\widehat{G}_{\mu\nu}{}^{\alpha\beta} h_{\alpha\beta} = \kappa \left(\tau_{\mu\nu} + \partial_\alpha (s_{\mu\nu}{}^{\alpha} + s_{\nu\mu}{}^{\alpha} - s^{\alpha}{}_{\mu\nu})/2\right), \tag{5.45}$$

making the function of $\tau_{\mu\nu}$ and $s^{\alpha}{}_{\mu\nu}$ absolutely clear: these tensors do not simply constitute a passive "kinematical" description of gravitational energy-momentum and spin, they actively determine the field's *dynamics*.

Despite the satisfying simplicity of this result, equation (5.45) is clearly not the best point at which to end our investigation. Having extolled the virtues of a formalism which keeps spin separate from energy-momentum, our real goal must be to find a formulation of gravity in which $\tau_{\mu\nu}$ and $s^{\alpha}{}_{\mu\nu}$ appear as source-terms in *separate* gravitational field equations. It should come as no surprise that Einstein-Cartan theory will provide precisely the environment in which to achieve this objective.

[17] If $T^{\text{Bel}}_{\mu\nu} \sim O(h) \gg t_{\mu\nu}$, a little more care must be taken: the right-hand side of (5.44) is then $\kappa(t_{\mu\nu} + T^{\text{Bel}(1)}_{\mu\nu} + T^{\text{Bel}(2)}_{\mu\nu} + T^{\text{Bel}(1)}_{\mu\alpha}\bar{h}^{\alpha}{}_{\nu})$ the last term arising from the density and index placement of the equation (5.35) on which (5.42) was based. These terms do not contradict the statement that $t_{\mu\nu}$ and $T^{\text{Bel}}_{\mu\nu}$ generate gravity in the same fashion, because $t^{(1)}_{\mu\nu} = 0$; furthermore, we would need to work to third-order in $h_{\mu\nu}$, and have a third-order definition of $t_{\mu\nu}$, before terms such as $t^{(3)}_{\mu\nu}$ and $t^{(2)}_{\mu\alpha}\bar{h}^{\alpha}{}_{\nu}$ could be seen in the field equations.

5.3.2 The Einstein-Cartan Equations

We will now disentangle the spin and energy-momentum in equation (5.45), formulating a quadratic approximation to Einstein-Cartan gravity in which $\tau_{\mu\nu}$ and $s^{\alpha}{}_{\mu\nu}$ appear as *separate* source-terms in the field equations.[18] In close analogy with the previous section, we proceed by expanding the field equations (5.120) to second order in $f_{\mu\nu}$ and $w_{\alpha}{}^{\mu\nu}$, where

$$\phi^* e^a_{\bar{\mu}} = \delta^a_\mu - f^a{}_\mu/2, \tag{5.46a}$$
$$\phi^* \omega_a{}^{\bar{\mu}\bar{\nu}} = w_a{}^{\mu\nu}, \tag{5.46b}$$

are initially considered to be true to all orders; we then perform a non-linear field redefinition,

$$f_{\mu\nu} \to f_{\mu\nu} + O(f^2), \tag{5.47a}$$
$$w_{\alpha}{}^{\mu\nu} \to w_{\alpha}{}^{\mu\nu} + O(f^2), \tag{5.47b}$$

to generate the field equations we desire. In order to identify $\tau_{\mu\nu}$ and $s^{\alpha}{}_{\mu\nu}$ in these equations, it will also be necessary to express tensors of the form $w\partial f + w^2$ in terms of $h_{\mu\nu}$. To this end, we will evaluate these tensors on torsion-free perturbations (5.130),

$$w_{\alpha}{}^{\mu\nu} = (\partial^{[\nu} f^{\mu]}{}_{\alpha} + \partial^{[\nu} f_{\alpha}{}^{\mu]} + \partial_\alpha f^{[\nu\mu]})/2 + O(f^2), \tag{5.48}$$

in the "symmetric" rotation gauge (5.132):

$$f_{[\mu\nu]} = O(f^2). \tag{5.49}$$

These relations allow us to identify

$$f_{\mu\nu} \equiv h_{\mu\nu} + O(f^2), \tag{5.50a}$$
$$w_{\alpha}{}^{\mu\nu} \equiv \partial^{[\nu} h^{\mu]}{}_{\alpha} + O(f^2), \tag{5.50b}$$

and thus convert the quadratic parts of the field equations into the corresponding tensors of perturbative general relativity: $w\partial f + w^2 = \partial h \partial h + O(f^3)$.[19]

To begin, let us focus our attention on the first Einstein-Cartan field equation (5.120a). Following the approach of section 5.3.1, we express the vacuum field equation in terms of the mixed Einstein tensor density,

$$eG_a{}^b = 0, \tag{5.51}$$

[18] The essentials of perturbative Einstein-Cartan gravity are reviewed in section 4 of appendix 5.A.

[19] The validity of the formulae (5.13) can only be guaranteed in the generally-relativistic regime, i.e. torsion-free gravity described by a symmetric tensor field $h_{\mu\nu}$. In the linearised Einstein-Cartan theory, this corresponds to the restriction (5.48) for $w_{\alpha}{}^{\mu\nu}$ and (5.49) for $f_{\mu\nu}$. If we already knew how to generalise $\tau_{\mu\nu}$ and $s^{\alpha}{}_{\mu\nu}$ beyond the generally-relativistic regime (where the gravitational field is described by unconstrained $f_{\mu\nu}$ and $w_{\alpha}{}^{\mu\nu}$) then it would be possible to recognise these tensors in the field equations without making such restrictions. The conversion $\partial h \to \partial f + w$ is not unique, however, so it is not immediately clear how this generalisation should be achieved.

where $G_a{}^b \equiv (R_a{}^{\bar{\mu}} - e_a^{\bar{\mu}} R/2) e_{\bar{\mu}}^b$ is a function of the physical tetrad $e_{\bar{\mu}}^a$ and physical spin connection $\omega_a{}^{\bar{\mu}\bar{\nu}}$. Mapping this equation to the background, we expand to quadratic order in $f_{\mu\nu}$ and $w_\alpha{}^{\mu\nu}$, and simplify the resulting equation by taking the trace-reverse:

$$2\partial_{[\mu} w_{\alpha]\nu}{}^\alpha = -f\partial_{[\mu} w_{\alpha]\nu}{}^\alpha + f_{\nu\beta}\partial_{[\mu} w_{\alpha]}{}^{\beta\alpha} + f^\alpha{}_\beta \partial_{[\mu} w_{\alpha]\nu}{}^\beta - 2w_{[\mu|}{}^\alpha{}_\nu w_{|\beta]}{}^\beta{}_\alpha. \tag{5.52}$$

We now redefine $w_\alpha{}^{\mu\nu}$ according to the replacement

$$w_\alpha{}^{\mu\nu} \to w_\alpha{}^{\mu\nu} - f w_\alpha{}^{\mu\nu}/2 + f^{\beta[\mu} w_{\alpha\beta}{}^{\nu]} + f_\beta{}^{[\mu}\partial_\alpha f^{\nu]\beta}/4; \tag{5.53}$$

the physical spin connection is then given by

$$\phi^* \omega_a{}^{\bar{\mu}\bar{\nu}} = w_a{}^{\mu\nu} - f w_a{}^{\mu\nu}/2 + w_{a\beta}{}^{[\nu} f^{\mu]\beta} + f_\beta{}^{[\mu}\partial_a f^{\nu]\beta}/4,$$

and the quadratic field equation (5.52) becomes

$$\begin{aligned} 2\partial_{[\mu} w_{\alpha]\nu}{}^\alpha = {} & \partial_{[\mu} f w_{\alpha]\nu}{}^\alpha - \partial_{[\mu|} f_{\nu\beta} w_{|\alpha]}{}^{\beta\alpha} - \partial_{[\mu} f^{\alpha\beta} w_{\alpha]\nu\beta} - 2w_{[\mu|}{}^\alpha{}_\nu w_{|\beta]}{}^\beta{}_\alpha \\ & - \partial_{[\mu|} f^\beta{}_\nu \partial_{|\alpha]} f^\alpha{}_\beta/4 + \partial_{[\mu} f^{\beta\alpha}\partial_{\alpha]} f_{\nu\beta}/4. \end{aligned} \tag{5.54}$$

Applying (5.50) to the terms on the right-hand side, we obtain

$$2\partial_{[\mu} w_{\alpha]\nu}{}^\alpha = \kappa\bar{\tau}_{\mu\nu}, \tag{5.55}$$

which is simply the trace-reverse of

$$2\partial_{[\mu} w_{\alpha]\nu}{}^\alpha - \eta_{\mu\nu}\partial_\alpha w_\beta{}^{\alpha\beta} = \kappa\tau_{\mu\nu}. \tag{5.56}$$

This is the field equation we had hoped to construct, mirroring the structure of the linearised non-vacuum field equation (5.129a) with gravitational energy-momentum $\tau_{\mu\nu}$ taking the place of the material energy-momentum tensor $T_{\mu\nu}$.

We now turn to the second vacuum Einstein-Cartan field equation (5.120b). Writing this as

$$\mathcal{T}^a_{\bar{\mu}\bar{\nu}} = 0 \tag{5.57}$$

in the physical spacetime, we once again use the standard definitions (5.46) and expand to second order in the background:

$$\partial_{[\mu} f^\alpha{}_{\nu]} + 2w_{[\mu}{}^\alpha{}_{\nu]} = f_{\beta[\mu}\partial^\beta f^\alpha{}_{\nu]}/2 + w_\beta{}^\alpha{}_{[\nu} f^\beta{}_{\mu]} + w_{[\mu}{}^\beta{}_{\nu]} f^\alpha{}_\beta. \tag{5.58}$$

Consistency with the first field equation (5.56) requires us to redefine $w_\alpha{}^{\mu\nu}$ as before (5.53) but places no constraint on the definition of $f_{\mu\nu}$; we are therefore free to make the replacement

$$f_{\mu\nu} \to f_{\mu\nu} - f_{\mu\alpha} f^\alpha{}_\nu/4, \tag{5.59}$$

and fix the tetrad expansion at quadratic order:

$$\phi^* e_{\bar{\mu}}^a = \delta_\mu^a - f^a{}_\mu/2 + f^a{}_\nu f^\nu{}_\mu/8. \tag{5.60}$$

Appling these redefinitions to the second field equation (5.58) and converting the quadratic terms using (5.50), one finds that

$$\begin{aligned}\partial_{[\mu}f^{\alpha}{}_{\nu]} + 2w_{[\mu}{}^{\alpha}{}_{\nu]} &= 2\bar{h}_{\beta[\nu}\partial^{[\alpha}h_{\mu]}{}^{\beta]} \\ &= \kappa(s^{\alpha}{}_{\mu\nu} + \delta^{\alpha}_{[\mu}s^{\beta}{}_{\nu]\beta}),\end{aligned} \tag{5.61}$$

from which the desired equation follows:

$$\partial_{[\mu}f^{\alpha}{}_{\nu]} + 2w_{[\mu}{}^{\alpha}{}_{\nu]} + \delta^{\alpha}_{[\mu|}(\partial_{|\nu]}f - \partial_{\beta}f^{\beta}{}_{|\nu]} - 2w_{\beta}{}^{\beta}{}_{|\nu]}) = \kappa s^{\alpha}{}_{\mu\nu}.$$

We have found a suitable counterpart to equation (5.56), in which the gravitational spin tensor $s^{\alpha}{}_{\mu\nu}$ takes on the role played by material spin $S^{\alpha}{}_{\mu\nu}$ in the linearised field equation (5.129b).

Combining these results, we conclude that under the perturbative expansions

$$\phi^{*}e^{a}_{\bar{\mu}} = \delta^{a}_{\mu} - f^{a}{}_{\mu}/2 + f^{a}{}_{\nu}f^{\nu}{}_{\mu}/8, \tag{5.62a}$$
$$\phi^{*}\omega_{a}{}^{\bar{\mu}\bar{\nu}} = w_{a}{}^{\mu\nu} - fw_{a}{}^{\mu\nu}/2 + w_{a\beta}{}^{[\nu}f^{\mu]\beta} + f_{\beta}{}^{[\mu}\partial_{a}f^{\nu]\beta}/4, \tag{5.62b}$$

the vacuum Einstein-Cartan field equations are approximated, to quadratic order, by

$$2\partial_{[\mu}w_{\alpha]\nu}{}^{\alpha} - \eta_{\mu\nu}\partial_{\alpha}w_{\beta}{}^{\alpha\beta} = \kappa\tau_{\mu\nu}, \tag{5.63a}$$
$$\partial_{[\mu}f^{\alpha}{}_{\nu]} + 2w_{[\mu}{}^{\alpha}{}_{\nu]} + \delta^{\alpha}_{[\mu|}(\partial_{|\nu]}f - \partial_{\beta}f^{\beta}{}_{|\nu]} - 2w_{\beta}{}^{\beta}{}_{|\nu]}) = \kappa s^{\alpha}{}_{\mu\nu}, \tag{5.63b}$$

wherein $\tau_{\mu\nu}$ and $s^{\alpha}{}_{\mu\nu}$ generate the gravitational fields $f_{\mu\nu}$ and $w_{\alpha}{}^{\mu\nu}$ in an identical fashion to the energy-momentum and spin of *matter* (5.129). We have found the analogue of (5.42) in the Einstein-Cartan theory of gravity, in which gravitational energy-momentum and spin act as the source terms of *separate* field equations.

5.4 Field Redefinition

We have succeeded in demonstrating that $\tau_{\mu\nu}$ and $s^{\alpha}{}_{\mu\nu}$ do indeed express the dynamical "weight" of the gravitational field, and have uncovered the field definitions, (5.39) and (5.62), which make this relationship manifest at level of the field equations. We now turn our attention to the field definitions themselves, investigating the importance of (5.39) and (5.62) in a broader context, and exploring the effects of field redefinitions in general.

As section 5.3.2 was restricted to torsion-free fields (5.48) and symmetric gauge (5.49), we will be limited in what we can say about the field redefinitions of Einstein-Cartan gravity; the majority of our analysis will therefore focus on the definition of $h_{\mu\nu}$ in general relativity.

5.4.1 The "Central" Expansion

In many respects, the most striking aspect of the new definition of $h_{\mu\nu}$, as displayed in (5.39), is how closely it relates to a linear perturbation in the metric (5.31) and a linear perturbation in the inverse metric (5.34); this is in comparison with the full range of local Lorentz-covariant field definitions consistent with (5.31) at linear order:

$$\phi^* g_{ab} = \breve{g}_{ab} + h_{ab} + B_1 h_{ac} h^c{}_b + B_2 h_{ab} h + \breve{g}_{ab}(B_3 h^2 + B_4 h_{cd} h^{dc}) + O(h^3), \tag{5.64}$$

where $\{B_n\}$ are arbitrary constants. One can argue, in fact, that the definition (5.39) lies at a natural "centre" of the four-dimensional space parameterised by $\{B_n\}$. This argument begins by observing that *a priori* there is no special variable which represents the "true" dynamical field of general relativity: one can equally well define the gravitational field $h_{\mu\nu}$ as a linear perturbation in a metric density $(-g)^\lambda g_{ab}$, or an inverse metric density $(-g)^\lambda g^{ab}$, for any value of λ. Of all these choices, perturbations in the metric and its inverse (i.e. $\lambda = 0$) are distinguished by the fact that they possess linear-order gauge-transformation of the form $\partial_{(\mu}\xi_{\nu)}$, without a part proportional to $\eta_{\mu\nu}\partial_\alpha\xi^\alpha$, and can therefore be identified with the Fierz-Pauli massless spin-2 field. However, once we have restricted our interest to these particular definitions ((5.31) or (5.34)) the decision to focus on one, and discard the other, is completely arbitrary. Instead of making a forced choice between two essentially equivalent options, one might instead consider the definition that lies exactly *half-way* between them, where the values of $\{B_n\}$ are the mean of those in (5.31) and (5.34). It is easy to see that this "centre point" is precisely the field (5.39) that casts $t_{\mu\nu}$ as the self-interaction term of the quadratic field equations (5.42)! This is an extraordinary coincidence, as $\tau_{\mu\nu}$ and $s^\alpha{}_{\mu\nu}$ (and consequently $t_{\mu\nu}$) were selected for their capacity to display positive energy-density, causal energy-flux, and traceless spatial spin; none of these criteria would be expected to determine a definition of $h_{\mu\nu}$ that is geometrically distinguished in this way.

5.4.2 Expansion of the Einstein Hilbert Lagrangian

The new metric expansion (5.39) also offers some perspective on the particular form of the Lagrangian (5.10) that generated the main results of section 5.2, including the generalised formulae (5.13) for $\tau_{\mu\nu}$ and $s^\alpha{}_{\mu\nu}$.

In chapter 2 we studied the expansion of the Einstein-Hilbert Lagrangian,

$$\mathcal{L}_{\mathrm{EH}} = -\sqrt{-g}R/\kappa, \tag{5.65}$$

under a linear perturbation of the inverse metric: $h'_{\mu\nu}$ as defined in (5.33).[20] Taking care to retain all terms proportional to background curvature, but ignoring surface terms, the Einstein-Hilbert Lagrangian was found to expand as follows,

$$\mathcal{L}_{\mathrm{EH}}[\breve{g}^{ab} - h'^{ab}] = -\sqrt{-\breve{g}}\breve{R}/\kappa + \mathcal{L}'_1[h'^{ab}] + \mathcal{L}'_2[h'^{ab}] + O(h'^3), \tag{5.66}$$

[20] The field written as $h_{\mu\nu}$ in chapter 2 is in fact $-h'_{\mu\nu}$ in our present notation; the Lagrangians of that chapter also take the opposite sign to those here.

where the linear and quadratic parts of the Lagrangian,

$$\mathcal{L}'_1[h'^{ab}] = \frac{\sqrt{-\breve{g}}}{\kappa}\breve{G}_{ab}h'^{ab}, \tag{5.67}$$

$$\begin{aligned}\mathcal{L}'_2[h'^{ab}] = \frac{-\sqrt{-\breve{g}}}{4\kappa}\Big(& h'_{ab}\breve{\nabla}^2 h'^{ab} - h'\breve{\nabla}^2 h' - 2h'^a{}_b\breve{\nabla}_c\breve{\nabla}_a h'^{bc} + h'\breve{\nabla}_a\breve{\nabla}_b h'^{ab} \\ &+ h'^{ab}\breve{\nabla}_a\breve{\nabla}_b h' - 2\breve{R}_{ab}h'^{ab}h' + \breve{R}(h'^{ab}h'_{ab} + h'^2/2)\Big),\end{aligned} \tag{5.68}$$

are given by equations (2.67) and (2.69) of chapter 2. By commuting the derivatives of the third term, and integrating the first five terms by parts (discarding surface terms), we now note that $\mathcal{L}'_2$ can be rewritten as

$$\mathcal{L}'_2[h'^{ab}] = \frac{\sqrt{-\breve{g}}}{4\kappa}\left(\breve{\nabla}_c h'^{ab}\breve{\nabla}^c\bar{h}'_{ab} - 2\breve{\nabla}^a\bar{h}'_{ab}\breve{\nabla}^c\bar{h}'_c{}^b + 2\bar{h}'^{ab}\breve{R}_{cabd}\bar{h}'^{cd} + 2\breve{G}_{ab}h'^a{}_c h'^{cb}\right). \tag{5.69}$$

To generate the expansion of $\mathcal{L}_{\text{EH}}$ under our newly defined perturbation $h_{\mu\nu}$, we need only compare its metric expansion (5.39) to that of $h'_{\mu\nu}$ (5.34),

$$h'_{ab} = h_{ab} - h_{ac}h^c{}_b/2 + O(h^3), \tag{5.70}$$

and substitute this relation into (5.66):

$$\begin{aligned}\mathcal{L}_{\text{EH}}[\breve{g}^{ab} - (h^{ab} - h^{ac}h_c{}^b/2 + O(h^3))] = -\sqrt{-\breve{g}}\breve{R}/\kappa + \mathcal{L}'_1[h^{ab} - h^{ac}h_c{}^b/2] + \mathcal{L}'_2[h_{ab}] \\ + O(h^3).\end{aligned} \tag{5.71}$$

Clearly, the quadratic part of this expansion is

$$\begin{aligned}\mathcal{L}_2[h^{ab}] &\equiv \mathcal{L}'_2[h^{ab}] + \mathcal{L}'_1[-h^{ac}h_c{}^b/2] \\ &= \frac{\sqrt{-\breve{g}}}{4\kappa}\left(\breve{\nabla}_c h'^{ab}\breve{\nabla}^c\bar{h}'_{ab} - 2\breve{\nabla}^a\bar{h}'_{ab}\breve{\nabla}^c\bar{h}'_c{}^b + 2\bar{h}'^{ab}\breve{R}_{cabd}\bar{h}'^{cd}\right),\end{aligned} \tag{5.72}$$

which is precisely the form of the Lagrangian $\mathcal{L}$ that we used to reproduce the formulae for $\tau_{\mu\nu}$ and $s^\alpha{}_{\mu\nu}$! Thus, the curvature term in (5.10), which we introduced in section 5.2 as an ansatz, can be understood as a consequence of the special definition of $h_{\mu\nu}$ associated with our framework: these are simply the terms proportional to the background curvature that appear when the Einstein-Hilbert action is expanded to quadratic order.

5.4.3 Field Redefinitions and Superpotentials

To understand the new metric expansion (5.39) in a wider context, we should also explain the relationship between field redefinitions and the superpotentials we encountered in section 5.2.3.

Recall that the addition of superpotentials to $\tau_{\mu\nu}$ and $s^\alpha{}_{\mu\nu}$ corresponds to the addition of curvature terms (5.14) and torsion terms (5.19) to the Lagrangian. Notice, however, that $t_{\mu\nu}$ is unaffected by torsion terms: the energy-momentum superpotentials (5.20) cancel those of the spin tensor (5.21) when they enter the formula (5.27). As a result, the superpotentials of the Belinfante tensor are characterised by curvature terms alone, which were determined by the five parameters $\{A_n\}$ of equation (5.15). Although these parameters are fixed according to (5.18), for the sake of argument let us relax these equations and alter each A_n by an amount ΔA_n; the gravitational Belinfante tensor then gains the superpotential term

$$\begin{aligned}\kappa\Delta t_\mu{}^\nu &= \kappa\partial_\alpha(\Delta s_\mu{}^{\nu\alpha} + \Delta s^\nu{}_\mu{}^\alpha - \Delta s^\alpha{}_\mu{}^\nu)/2 \\ &= \partial_\alpha\partial^\beta\Delta\Sigma^{\alpha\nu}{}_{\beta\mu},\end{aligned} \tag{5.73}$$

where, according to (5.15),

$$\begin{aligned}\Delta\Sigma^{\alpha\beta}{}_{\mu\nu} = \Delta A_1 h^\alpha_{[\mu}h^\beta_{\nu]} + \Delta A_2 h h^{[\alpha}_{[\mu}\delta^{\beta]}_{\nu]} + \Delta A_3 h^\gamma_{[\mu}\delta^{[\beta}_{\nu]}h^{\alpha]}_\gamma \\ + \delta^\alpha_{[\mu}\delta^\beta_{\nu]}\left(\Delta A_4 h^2 + \Delta A_5 h_{\gamma\delta}h^{\gamma\delta}\right).\end{aligned} \tag{5.74}$$

It is also possible to generate superpotentials in the quadratic approximation to Einstein's field equations: an arbitrary field redefinition

$$h_{\mu\nu} \to h_{\mu\nu} + \Delta h_{\mu\nu}, \tag{5.75}$$

adds the divergence-free tensor

$$-\widehat{G}_{\mu\nu}{}^{\alpha\beta}\Delta h_{\alpha\beta}, \tag{5.76}$$

to the right-hand side of the field equations (5.42) and thus defines a new Belinfante tensor,

$$\kappa t'_{\mu\nu} \equiv \kappa t_{\mu\nu} - \widehat{G}_{\mu\nu}{}^{\alpha\beta}\Delta h_{\alpha\beta}, \tag{5.77}$$

that acts as the source of the new gravitational field. Therefore, as long as we can find a field redefinition $\Delta h_{\mu\nu}$ to solve

$$-\widehat{G}_\mu{}^{\nu\alpha\beta}\Delta h_{\alpha\beta} = \partial_\alpha\partial^\beta\Delta\Sigma^{\alpha\nu}{}_{\beta\mu}, \tag{5.78}$$

we can produce the same superpotential in the field equations as the ones we have generated by altering $\{A_n\}$ in the Lagrangian.

First we shall try to solve equation (5.78) using *local* field redefinitions. Noting that $\Delta h_{\mu\nu}$ will also need to be Lorentz-covariant and quadratic in $h_{\mu\nu}$ to solve this equation, the most general field redefinition we need to consider is

$$\Delta h_{\mu\nu} = \Delta B_1 h_{\mu\alpha}h^\alpha{}_\nu + \Delta B_2 h_{\mu\nu}h + \eta_{\mu\nu}(\Delta B_3 h^2 + \Delta B_4 h_{\alpha\beta}h^{\alpha\beta}), \tag{5.79}$$

where the $\{\Delta B_n\}$ correspond to changes in the parameters $\{B_n\}$ of equation (5.64). Comparing the number of free parameters here with those of (5.74) it is immediately clear that these local redefinitions will not span the entire space of superpotentials. Indeed, if we rewrite (5.78) as[21]

$$0 = \partial_\alpha\partial^\beta\left(\Delta\Sigma^{\alpha\nu}{}_{\beta\mu} - 2\delta^{[\nu}_{[\mu}\Delta\bar{h}^{\alpha]}_{\beta]}\right), \tag{5.80}$$

[21]The useful identity $\widehat{G}_\mu{}^{\nu\alpha\beta}X_{\alpha\beta} \equiv -2\partial_\alpha\partial^\beta(\delta^{[\nu}_{[\mu}\bar{X}^{\alpha]}_{\beta]})$ holds for any symmetric tensor $X_{\mu\nu}$.

and insert (5.79) and (5.74), we arrive at

$$0 = \partial_\alpha \partial^\beta \Big[\Delta A_1 h^\alpha_{[\beta} h^\nu_{\mu]} + (\Delta A_2 - 2\Delta B_2) h h^{[\alpha}_{[\beta} \delta^{\nu]}_{\mu]} + (\Delta A_3 - 2\Delta B_1) h^\gamma_{[\beta} \delta^{[\nu}_{\mu]} h^{\alpha]}_\gamma \\ + \delta^\alpha_{[\beta} \delta^\nu_{\mu]} (\Delta A_4 + 2\Delta B_3 + \Delta B_2) h^2 + \delta^\alpha_{[\beta} \delta^\nu_{\mu]} (\Delta A_5 + 2\Delta B_4 + \Delta B_1) h_{\gamma\delta} h^{\gamma\delta} \Big], \tag{5.81}$$

which makes the mismatch of parameters unequivocal. Clearly, this equation can only hold for all $h_{\mu\nu}$ if

$$\Delta A_1 = 0, \tag{5.82}$$

and if we assume that this is the case, the local field redefinition $\Delta h_{\mu\nu}$ is determined uniquely:

$$\begin{aligned} &\Delta B_1 = \Delta A_3/2, && \Delta B_2 = \Delta A_2/2, \\ &\Delta B_3 = -\Delta A_4/2 - \Delta A_2/4, && \Delta B_4 = -\Delta A_5/2 - \Delta A_3/4. \end{aligned} \tag{5.83}$$

As the ΔA_n were defined relative to the values (5.18) that correspond to our Belinfante tensor $t_{\mu\nu}$, the condition (5.82) implies that

$$A_1 = -1, \tag{5.84}$$

which also arose in section 5.2.1 as the requirement that ensured the curvature terms would extend the spin-2 gauge invariance of the Lagrangian beyond the flat background.

If $A_1 \neq -1$ then it is still possible to solve equation (5.78) by inverting the differential operator $\widehat{G}_\mu{}^{\nu\alpha\beta}$:

$$\Delta\bar{h}_\mu{}^\nu = \frac{2}{\partial^2} \partial_\alpha \partial^\beta \Delta\Sigma^{\alpha\nu}{}_{\beta\mu}, \tag{5.85}$$

where the precise form of the propagator $1/\partial^2$ will depend on boundary conditions.[22] As we have seen, however, this $\Delta h_{\mu\nu}$ cannot be a *local* function of $h_{\mu\nu}$; hence it will no longer be possible to express the physical metric $\phi^* g_{ab}$ as a local function of the gravitational field (5.64).

From this vantage point, we can now appreciate another important property of our framework: the Belinfante tensor $t_{\mu\nu}$ (constructed from $\tau_{\mu\nu}$ and $s^\alpha{}_{\mu\nu}$) is the source of a gravitational field $h_{\mu\nu}$ of which the metric is a *local* function (5.39). The new field definition (5.39) is not only notable for its special location in the space of local field definitions (5.64): it is notable in that it even exists within this space!

In contrast, the same cannot be said for the celebrated tensor $t^{\rm LL}_{\mu\nu}$ of Landau and Lifshitz [52]:

$$\kappa t^{ab}_{\rm LL} \equiv -\phi^* G^{ab} + \frac{1}{\phi^* g} \check{\nabla}_c \check{\nabla}_d \Big(\phi^* (g g^{a[b} g^{c]d}) \Big). \tag{5.86}$$

[22] Because $\partial^\mu \widehat{G}_{\mu\nu}{}^{\alpha\beta} = 0$, the inverse of $\widehat{G}_{\mu\nu}{}^{\alpha\beta}$ is only defined on fields with vanishing divergence, such as $\partial_\alpha \partial^\beta \Delta\Sigma^{\alpha\nu}{}_{\beta\mu}$. Also, because $\widehat{G}_{\mu\nu}{}^{\alpha\beta}(\partial_{(\alpha}\xi_{\beta)}) = 0$, the inverse is only defined up to the addition of gauge fields $\partial_{(\mu}\xi_{\nu)}$; to remove this degeneracy, we have set $\partial^\mu \Delta\bar{h}_{\mu\nu} = 0$ and absorbed the fields $\partial_{(\mu}\xi_{\nu)}$ into the gauge freedom of $h_{\mu\nu}$.

The divergence on the right-hand side clearly contributes a term

$$\partial_\alpha\partial_\beta(h^{\mu[\nu}h^{\alpha]\beta}), \tag{5.87}$$

at second order, corresponding to a superpotential with $\Delta A_1 = 1 \Rightarrow A_1 = 0$. Thus, there can be no local field redefinition that will render the Landau-Lifshitz tensor as the source term of the vacuum Einstein field equations, and furthermore, the tensor cannot be derived from a covariantised Lagrangian which maintains its spin-2 gauge invariance beyond the flat background.

It is rather surprising that this deficiency is not more widely known. In their effort to construct a gravitational energy-momentum tensor that was symmetric and free of second derivatives *in all gauges*, it seems that Landau and Lifshitz were forced to include a superpotential (5.87) that would be impossible to generate in the field equations by a local redefinition of $h_{\mu\nu}$. The only way the Landau-Lifshitz tensor can be given an equal footing with the energy-momentum of matter, generating gravity alongside $T_{\mu\nu}$ in the Einstein field equations, would be to define the gravitational field $h_{\mu\nu}$ in terms of a *non-local* perturbation in the metric.

5.4.4 Beyond Second Order

It will be difficult to gain any further insight into the physical meaning of the new $h_{\mu\nu}$ without first deciding how the definition (5.39) should extend beyond quadratic order. As this question is intimately related to the issue of extending the formulae of $\tau_{\mu\nu}$ and $s^\alpha{}_{\mu\nu}$ to the full non-linear theory, we shall postpone a thorough investigation of this topic for another publication. For now, we only mention one particularly attractive possibility. As we have already explained, the new definition (5.39) lies on a point of symmetry between a linear perturbations in the metric (5.31) and a linear perturbation in the inverse metric (5.34); reflecting this, one finds that the expansion of the inverse metric, consistent with (5.39) is

$$\phi^* g^{ab} = \breve{g}^{ab} - h^{ab} + h^{ac}h_c{}^b/2 + O(h^3), \tag{5.88}$$

which, apart from a change in sign convention $h_{\mu\nu} \to -h_{\mu\nu}$, is identical to the metric expansion (5.39) to quadratic order. Thus, a particularly natural extension of (5.39) would be one which preserved this symmetry *exactly*, so that the metric and its inverse had an identical expansion to all orders, except for a sign change in $h_{\mu\nu}$. This idea can be realised by viewing the tensor $h^a{}_b/2$ as a linear map and forming its exponential $[e^{h/2}]^a{}_b$; a metric defined by

$$\phi^* g_{ab} \equiv [e^{h/2}]^c{}_a \breve{g}_{cd} [e^{h/2}]^d{}_b, \tag{5.89}$$

is then consistent with (5.39) to second order, and moreover, the associated expansion of the inverse metric,

$$\phi^* g^{ab} = [e^{-h/2}]^a{}_c \breve{g}^{cd} [e^{-h/2}]^b{}_d, \tag{5.90}$$

is clearly identical to (5.89) apart from a change in the sign of $h_{\mu\nu}$.[23] Roughly speaking, this $h_{\mu\nu}$ corresponds to a linear perturbation in the *logarithm* of the metric; indeed, this is literally the case for the metric determinant:

$$\phi^* \log(-g) = \log(-\breve{g}) + h. \tag{5.91}$$

5.4.5 New Fields for Einstein-Cartan

While it is certainly tempting to bring our analysis to bear on the new field definitions (5.62) that arose in Einstein-Cartan theory, unfortunately a full discussion of these variables will not be possible at this time. As we previously explained, the right-hand sides of (5.63) are only given by the formulae (5.13) when we restrict ourselves to the torsion-free perturbations (5.48) and symmetric gravitational field (5.49) of general relativity. We are therefore free to alter the quadratic parts of the new field definitions (5.62) by terms proportional to $\mathcal{T}_\alpha{}^{\mu\nu}$ and $f_{[\mu\nu]}$: such terms vanish under the aforementioned restrictions, and so do not interfere with our results. Until we fix the definitions of $\tau_{\mu\nu}$ and $s^\alpha{}_{\mu\nu}$ in terms of arbitrary perturbations $f_{\mu\nu}$ and $w^\alpha{}_{\mu\nu}$, this degeneracy will remain, and it will be difficult to offer a physical interpretation of the new field variables. We leave this generalisation of $\tau_{\mu\nu}$ and $s^\alpha{}_{\mu\nu}$, and the consequent analysis of the field variables they define, for another time. For now, we shall simply remark that the new tetrad expansion (5.62a) defines an inverse tetrad

$$\phi^* e_a^{\bar{\mu}} = \delta_a^\mu + f^\mu{}_a/2 + f^\mu{}_\nu f^\nu{}_a/8 + O(f^3), \tag{5.92}$$

which, to second order, differs from the tetrad expansion only by a change in sign convention $f_{\mu\nu} \to -f_{\mu\nu}$: mirroring the relationship between the metric expansion (5.39) and inverse metric expansion (5.88) for the new definition of $h_{\mu\nu}$ that arose in section 5.3.1. In fact, $f_{\mu\nu}$ defines a metric

$$\phi^* g_{ab} = \breve{g}_{ab} + f_{(ab)} + f_{(a|}{}^\mu f_{\mu|b)}/4 + f^\mu{}_a f_{\mu b}/4 + O(f^3),$$

which, if we set

$$f_{\mu\nu} = h_{\mu\nu} + O(f^3), \tag{5.93}$$

gives

$$\phi^* g_{ab} = \breve{g}_{ab} + h_{ab} + h_a{}^\mu h_{\mu b}/2 + O(h^3), \tag{5.94}$$

identical to the new metric expansion (5.39) of section 5.3.1. Note that the equivalence (5.93) between $f_{\mu\nu}$ and $h_{\mu\nu}$ now holds to *quadratic* order, whereas one would only expect a linear correspondence from the symmetric gauge (5.137) condition alone.

[23] The advantage of this "double sided" exponential, over the more obvious "one sided" definition $\phi^* g_{ab} \equiv [e^h]^c{}_a \breve{g}_{cb}$, is that the symmetry of the metric is automatic. For the one sided definition, we would need to impose the constraint $0 = h_{[ab]} = h^c{}_{[b}\breve{g}_{a]c}$, which prevents us from considering $h^a{}_b$ and $\breve{g}_{ab}$ as independent fields.

5.5 Conclusion

As a local description of the energy-momentum and spin of the linearised gravitational field, $\tau_{\mu\nu}$ and $s^{\alpha}{}_{\mu\nu}$ serve a number of purposes within the theory. In addition to accounting for the energy-momentum and angular momentum exchanged locally with matter (chapters 3 and 4), we can now confirm their status as Noether currents of translational and rotational symmetry, and as sources of gravity itself. Thus, our framework displays many of the fundamental properties possessed by previous treatments of local gravitational energetics [29, 35, 52], all the while endowing linear gravity with positive energy-density, causal energy-flux, and traceless spatial spin.

By "gauging" the translational and rotational symmetries of the Fierz-Pauli massless spin-2 field, we have constructed a Lagrangian

$$\mathcal{L} \equiv \frac{\check{e}}{4\kappa}\left(\check{D}_{\bar{\mu}}h_{\bar{\alpha}\bar{\beta}}\check{D}^{\bar{\mu}}\bar{h}^{\bar{\alpha}\bar{\beta}} - 2\check{D}_{\bar{\mu}}\bar{h}^{\bar{\mu}\bar{\alpha}}\check{D}_{\bar{\nu}}\bar{h}^{\bar{\nu}}{}_{\bar{\alpha}} + 2\bar{h}^{\bar{\mu}\bar{\nu}}\check{R}_{\bar{\alpha}\bar{\mu}\bar{\nu}\bar{\beta}}\bar{h}^{\bar{\alpha}\bar{\beta}}\right), \tag{5.95}$$

from which the Noether currents

$$\kappa\bar{\tau}_{\mu\nu} = \tfrac{1}{4}\partial_\mu h_{\alpha\beta}\partial_\nu\bar{h}^{\alpha\beta} - \tfrac{1}{2}\partial_\mu\bar{h}_{\nu\alpha}\partial_\beta\bar{h}^{\alpha\beta}, \tag{5.96a}$$
$$\kappa s^{\alpha}{}_{\mu\nu} = 2\bar{h}_{\beta[\nu}\partial^{[\alpha}\bar{h}_{\mu]}{}^{\beta]} + \delta^{\alpha}_{[\nu}\bar{h}_{\mu]}{}^{\beta}\partial_\gamma\bar{h}^{\gamma}{}_{\beta}, \tag{5.96b}$$

may be obtained. Thus a Lagrangian exists which encodes the dynamics of linear gravity and which also defines an energy-momentum tensor and spin tensor that reduce to our formulae (5.1) in harmonic gauge (5.2). Furthermore, by exposing the relationship between background-coupling (the freedom that arises when "covariantising" the Fierz-Pauli Lagrangian (5.3) beyond flat spacetime) and superpotentials, we have demonstrated that (5.95) is in fact the *only* covariantised Fierz-Pauli Lagrangian capable of reproducing our tensors in harmonic gauge, without introducing second derivatives into the generalised $\tau_{\mu\nu}$. Consequently, the formulae (5.96) comprise a suitably unique generalisation of $\tau_{\mu\nu}$ and $s^{\alpha}{}_{\mu\nu}$ beyond harmonic gauge.

In order to reveal the dynamical role of $\tau_{\mu\nu}$ and $s^{\alpha}{}_{\mu\nu}$ in general relativity, we then considered a non-linear pertubation of the metric,

$$\phi^* g_{ab} = \check{g}_{ab} + h_{ab} + h_{ac}h^{c}{}_{b}/2. \tag{5.97}$$

According to this expansion, the quadratic approximation of the vacuum Einstein field equations (5.35) takes the form

$$\begin{aligned}\widehat{G}_{\mu\nu}{}^{\alpha\beta}h_{\alpha\beta} &= \kappa\left(\tau_{\mu\nu} + \partial_\alpha(s_{\mu\nu}{}^{\alpha} + s_{\nu\mu}{}^{\alpha} - s^{\alpha}{}_{\mu\nu})/2\right)\\ &= \kappa t_{\mu\nu},\end{aligned} \tag{5.98}$$

wherein $\tau_{\mu\nu}$ and $s^{\alpha}{}_{\mu\nu}$ combine to form a Belinfante tensor (5.30) and this combination acts as the source of the gravitational wave-equation, generating gravity as though it were the energy-momentum and spin of matter (5.43). Considering the full range of local Lorentz-covariant metric expansions (5.64), the simplicity of (5.97) is rather remarkable.

Indeed, the gravitational field definition (5.97) lies on a special "central point" of the field definition parameter-space, exactly half-way between a linear perturbation in the metric (5.31) and a linear perturbation in the inverse metric (5.34). It is possible to preserve the special symmetry of this field definition to all orders by expanding the metric with an exponentiated gravitational field (5.89).

While (5.98) succeeds in casting $\tau_{\mu\nu}$ and $s^{\alpha}{}_{\mu\nu}$ as gravitational sources, it only achieves this objective by combining spin and energy-momentum into a single entity. Considering that many of the notable features of our framework depend on the separation of spin from energy-momentum, a more desirable result would be a formulation of general relativity in which $\tau_{\mu\nu}$ and $s^{\alpha}{}_{\mu\nu}$ appear as the sources of *separate* field equations. Einstein-Cartan gravity provided the framework needed to achieve this goal: by expanding the tetrad and spin connection according to

$$\phi^* e^a_{\bar{\mu}} = \delta^a_\mu - f^a{}_\mu/2 + f^a{}_\nu f^\nu{}_\mu/8, \tag{5.99a}$$
$$\phi^* \omega_a{}^{\bar{\mu}\bar{\nu}} = w_a{}^{\mu\nu} - f w_a{}^{\mu\nu}/2 + w_{a\beta}{}^{[\nu} f^{\mu]\beta} + f_\beta{}^{[\mu}\partial_a f^{\nu]\beta}/4, \tag{5.99b}$$

the quadratic approximation of the vacuum field equations (5.51) and (5.57) were shown to be

$$2\partial_{[\mu} w_{\alpha]\nu}{}^{\alpha} - \eta_{\mu\nu}\partial_\alpha w_\beta{}^{\alpha\beta} = \kappa\tau_{\mu\nu}, \tag{5.100a}$$
$$\partial_{[\mu} f^\alpha{}_{\nu]} + 2w_{[\mu}{}^{\alpha}{}_{\nu]} + \delta^\alpha_{[\mu|}(\partial_{|\nu]} f - \partial_\beta f^\beta{}_{|\nu]} - 2w_\beta{}^\beta{}_{|\nu]}) = \kappa s^\alpha{}_{\mu\nu}, \tag{5.100b}$$

where $\tau_{\mu\nu}$ appears as the quadratic term in the curvature equation, playing the role of material energy-momentum, and $s^{\alpha}{}_{\mu\nu}$ appears as the quadratic term in the torsion equation, playing the role of material spin. Thus, in a theory of gravity where the energy-momentum and spin of matter maintain their separate identities, $\tau_{\mu\nu}$ and $s^{\alpha}{}_{\mu\nu}$ occur as distinct objects in the dynamical equations, disentangling the Belinfante tensor of equation (5.98).

Around these main results, a secondary loop of logic has been threaded, connecting background couplings, superpotentials, and field definitions. The freedom to couple $h_{\mu\nu}$ to the curvature and torsion of the background (when covariantising the Fierz-Pauli Lagrangian (5.3)) generates superpotentials in the energy-momentum tensor and spin tensor (§5.2.3). By performing non-linear redefinitions of $h_{\mu\nu}$ (changing the expansion of the metric, for example) it is then possible to produce these superpotentials in the field equations (§5.4.3). Coming full-circle, the perturbative expansion of the metric (or the tetrad and spin connection) determines a perturbative expansion of the Einstein-Hilbert action about a curved vacuum background, fixing the curvature couplings that occur in the Lagrangian at quadratic order (§5.4.2). The interrelation of these ideas makes it possible to derive many of the key results of this chapter from various starting points: the formulae for $\tau_{\mu\nu}$ and $s^{\alpha}{}_{\mu\nu}$, the Lagrangian (5.10), or the field definitions (5.97) or (5.99).

In the process of navigating this mathematical circuit, the privileged status of a particular class of superpotentials ($A_1 = -1$ in (5.15)) has also become clear. At the level of the spin tensor, $A_1 = -1$ arose in chapter 4 when demanding that $s^{\alpha}{}_{\alpha\nu} = 0$ in transverse-traceless gauge, ensuring the absence of infinite pressure gradients. Within the Lagrangian, background couplings with $A_1 = -1$ were necessary to extend spin-2 gauge invariance to curved vacuum backgrounds (5.11). Thirdly, in the quadratic field equations, *local* field redefinitions (5.79) can only produce superpotentials with $A_1 = -1$; if $A_1 \neq -1$, a non-local redefinition (5.85) is required. Thus, the *traceless condition* we imposed on $s^{\alpha}{}_{\mu\nu}$ in chapter 4 has conveyed two unexpected but extremely important properties. First, the Lagrangian that generates our tensors (5.10) preserves the gauge invariance of $h_{\mu\nu}$ in curved vacuum backgrounds; second, it has been possible to derive field equations (5.98) and (5.100), in which $\tau_{\mu\nu}$ and $s^{\alpha}{}_{\mu\nu}$ generate the gravitational field, using *local* definitions of the gravitational field (5.97) and (5.99). In comparison, the Landau-Lifshitz tensor [52] has been constructed using superpotentials with $A_1 = 0$ and so can claim none of these advantages: in particular, no local field definition can cast the Landau-Lifshitz tensor as a source for the gravitational field.

Finally, in appendix 5.C, we briefly examine the *global* measures of energy, momentum and angular momentum that our local framework can define. Within the quadratic approximation to general relativity, and with appropriate boundary conditions, it is possible to confirm the equivalence of these global quantities to the energy, momentum and angular momentum of Arnowitt, Deser, and Misner [6, 9].

Having embedded $\tau_{\mu\nu}$ and $s^{\alpha}{}_{\mu\nu}$ within various aspects of the linear and quadratic approximations to gravity, the key goal that remains is to extend these ideas to the full non-linear theory. This is obviously an ambitious task, and at the present stage it is far from clear which of our framework's properties can survive in the exact theory. However, based on the results of this chapter, it seems more than likely that Einstein-Cartan gravity will provide a natural starting point from which to launch this undertaking, with the field definitions (5.39), (5.89) and (5.99) offering clues as to the new field variables into which this theory should be cast.

5.A Appendix: Einstein-Cartan Theory

The role of this appendix is to briefly introduce Einstein-Cartan gravity, establish notation, and serve as a reference for results needed in the body of the chapter. For a more complete treatment of the subject, see [17, 32, 43, 53].

5.A.1 Kinematics

Einstein-Cartan theory is a slight extension of general relativity, in which (as formulated by Kibble [48] and Sciama [74]) translational and rotational symmetries are gauged separately, rather than being subsumed into a single diffeomorphism gauge transformation. The gravitational field is represented by four vector fields $e^a_{\bar{\mu}}$ (the tetrad) and six covector fields $\omega_a{}^{\bar{\mu}\bar{\nu}} = -\omega_a{}^{\bar{\nu}\bar{\mu}}$ (the spin connection). Following the conventions of chapters 3 and 4 and Wald [79], Greek letters are used as numerical indices (running from 0 to 3, raised and lowered by $\eta_{\mu\nu} = \mathrm{diag}(-1, 1, 1, 1)$) while Roman indices represent the tensor "slots" of Penrose's abstract index notation [79, §2.4]. Note that the numerical indices now come in two varieties: the unadorned Greek letters are used to enumerate the components of tensors in a Lorentzian coordinate system $\{x^\mu\}$ of flat spacetime, whereas the Greek letters with overbars enumerate the components of tensors with respect to the non-holonomic basis $\{e^a_{\bar{\mu}}\}$ formed from the tetrad. Thus, a generic vector field v^a, or a generic covector field v_a, would define the following quantities:

$$v_{\bar{\mu}} \equiv e^a_{\bar{\mu}} v_a, \qquad v^{\bar{\mu}} \equiv e^{\bar{\mu}}_a v^a, \tag{5.101}$$

where $e^{\bar{\mu}}_a$ is the inverse tetrad, defined by $e^a_{\bar{\mu}} e^{\bar{\nu}}_a = \delta^{\bar{\mu}}_{\bar{\nu}}$, or equivalently $e^a_{\bar{\mu}} e^{\bar{\mu}}_b = \delta^a_b$. To make this convention consistent with general relativity, in which abstract indices are raised and lowered by the metric g_{ab}, we must make the following identification:

$$g_{ab} \equiv e^{\bar{\mu}}_a e_{\bar{\mu} b}. \tag{5.102}$$

From the point of view of Einstein-Cartan theory, this constitutes the *definition* of the metric. Alternatively, from a general relativistic standpoint, this can be thought of as an orthonormality condition on the tetrad: $e^a_{\bar{\mu}} e^b_{\bar{\nu}} g_{ab} = \eta_{\bar{\mu}\bar{\nu}}$.

The tetrad and the spin connection are gauge fields: they allow the global translational and rotational symmetries of flat space to be generalised to *local* symmetries. To explain, let us first consider an infinitesimal local translation generated by a vector field ξ^a, that is, a diffeomorphism $\varphi : \mathcal{M} \to \mathcal{M}$, the action of which on tensor fields is $\varphi^* = 1 + \mathcal{L}_\xi$, where $\mathcal{L}_\xi$ is the Lie derivative with respect to ξ^a. Under this local translation, a scalar field ψ transforms as follows:

$$\psi \to \varphi^* \psi = (1 + \xi^a \partial_a)\psi, \tag{5.103}$$

which has exactly the same form as a global translation ($\xi^a = \text{const.}$) because ξ^a is not differentiated. This is not the case for the gradient of ψ, however:

$$\partial_a \psi \to \varphi^*(\partial_a \psi) = (1 + \xi^b \partial_b)\partial_a \psi + \partial_a \xi^b \partial_b \psi, \tag{5.104}$$

which clearly depends on the derivatives of ξ^a. To remedy this, we can define a translationally covariant derivative $\partial_{\bar{\mu}} \equiv e^a_{\bar{\mu}} \partial_a$ which, because $\mathcal{L}_\xi e^a_{\bar{\mu}} = \xi^b \partial_b e^a_{\bar{\mu}} - e^b_{\bar{\mu}} \partial_b \xi^a$, defines a gradient $\partial_{\bar{\mu}} \psi$ that transforms exactly as ψ does:

$$\varphi^*(\partial_{\bar{\mu}} \psi) = (1 + \xi^a \partial_a)\partial_{\bar{\mu}} \psi. \tag{5.105}$$

In this fashion, we can convert all spacetime indices into basis indices, and render tensorial field equations in a form which is covariant under local translations.

Local rotations are embodied by position-dependent Lorentz transformations which act on the tetrad,

$$e^a_{\bar{\mu}} \to e^a_{\bar{\nu}} \Lambda^{\bar{\nu}}{}_{\bar{\mu}}(x), \qquad \Lambda^{\bar{\alpha}}{}_{\bar{\mu}} \Lambda^{\bar{\beta}}{}_{\bar{\nu}} \eta_{\bar{\alpha}\bar{\beta}} = \eta_{\bar{\mu}\bar{\nu}}. \tag{5.106}$$

These leave the metric (5.102) unchanged, but affect all quantities with basis indices:

$$v_{\bar{\mu}} \to \Lambda^{\bar{\nu}}{}_{\bar{\mu}} v_{\bar{\nu}}, \qquad v^{\bar{\mu}} \to (\Lambda^{-1})^{\bar{\mu}}{}_{\bar{\nu}} v^{\bar{\nu}}. \tag{5.107}$$

Neither ∂_a nor $\partial_{\bar{\mu}}$ transform covariantly under local rotations ($\partial_a v_{\bar{\mu}} \to \partial_a(v_{\bar{\nu}}\Lambda^{\bar{\nu}}{}_{\bar{\mu}}) = \Lambda^{\bar{\nu}}{}_{\bar{\mu}}\partial_a v_{\bar{\nu}} + v_{\bar{\nu}}\partial_a\Lambda^{\bar{\nu}}{}_{\bar{\mu}}$) so a rotation-covariant derivative D_a is constructed using the spin connection:

$$\begin{aligned} D_a v_{\bar{\mu}} &= \partial_a v_{\bar{\mu}} - \omega_a{}^{\bar{\alpha}}{}_{\bar{\mu}} v_{\bar{\alpha}}, \\ D_a v^{\bar{\mu}} &= \partial_a v^{\bar{\mu}} + \omega_a{}^{\bar{\mu}}{}_{\bar{\alpha}} v^{\bar{\alpha}}, \quad \text{etc.} \end{aligned} \tag{5.108}$$

If we declare that the spin connection should transform according to

$$\omega_a{}^{\bar{\mu}}{}_{\bar{\nu}} \to (\Lambda^{-1})^{\bar{\mu}}{}_{\bar{\alpha}} \left(\partial_a \Lambda^{\bar{\alpha}}{}_{\bar{\nu}} + \omega_a{}^{\bar{\alpha}}{}_{\bar{\beta}} \Lambda^{\bar{\beta}}{}_{\bar{\nu}} \right), \tag{5.109}$$

under local rotations, then it is easy to show that these derivatives are indeed covariant:

$$\begin{aligned} D_a v_{\bar{\mu}} &\to \Lambda^{\bar{\nu}}{}_{\bar{\mu}} D_a v_{\bar{\nu}}, \\ D_a v^{\bar{\mu}} &\to (\Lambda^{-1})^{\bar{\mu}}{}_{\bar{\nu}} D_a v^{\bar{\nu}}, \quad \text{etc.} \end{aligned} \tag{5.110}$$

Using the tetrad once again, we can now construct a derivative $D_{\bar{\mu}} \equiv e^a_{\bar{\mu}} D_a$ that is covariant under both local translations and rotations. Thus, by replacing all partial derivatives with covariant derivatives, and all spacetime indices with basis indices, we can "gauge" the global Poincaré-invariance of any flat-space field theory, and in doing so, extend the theory to a spacetime with curvature and torsion.

5.A.2 Curvature and Torsion

The rotation covariant derivatives define a curvature tensor $R_{ab}{}^{\bar{\nu}}{}_{\bar{\mu}}$:

$$[D_a, D_b] v_{\bar{\nu}} \equiv -R_{ab}{}^{\bar{\mu}}{}_{\bar{\nu}} v_{\bar{\mu}} \quad \forall v_{\bar{\mu}}, \tag{5.111}$$

from which it follows that

$$R_{ab}{}^{\bar{\mu}}{}_{\bar{\nu}} = 2\left(\partial_{[a}\omega_{b]}{}^{\bar{\mu}}{}_{\bar{\nu}} + \omega_{[a|}{}^{\bar{\mu}}{}_{\bar{\alpha}} \omega_{|b]}{}^{\bar{\alpha}}{}_{\bar{\nu}} \right). \tag{5.112}$$

An asymmetric Ricci tensor can then be formed by contraction with the tetrad,

$$R_{b\bar{\nu}} \equiv e^a_{\bar{\mu}} R_{ab}{}^{\bar{\mu}}{}_{\bar{\nu}}, \tag{5.113}$$

and the Ricci scalar by a further contraction:

$$R \equiv e^a_{\bar{\mu}} R_a{}^{\bar{\mu}} = 2 e^a_{\bar{\mu}} e^b_{\bar{\nu}} \left(\partial_a \omega_b{}^{\bar{\mu}\bar{\nu}} + \omega_a{}^{\bar{\alpha}[\bar{\mu}} \omega_b{}^{\bar{\nu}]}{}_{\bar{\alpha}} \right). \tag{5.114}$$

The major difference between Einstein-Cartan gravity and general relativity is that, in addition to curvature, spacetime may also possess torsion, which we quantify with the torsion tensor

$$\mathcal{T}^{\bar{\alpha}}_{ab} \equiv 2 D_{[a} e^{\bar{\alpha}}_{b]}, \tag{5.115}$$

or equivalently,

$$\mathcal{T}^{a}_{\bar{\mu}\bar{\nu}} \equiv -2 D_{[\bar{\mu}} e^{a}_{\bar{\nu}]}. \tag{5.116}$$

If this tensor vanishes everywhere, we can use $D_{[\bar{\mu}}e^a_{\bar{\nu}]} = 0$ to relate the spin connection to the tetrad:

$$\omega_a{}^{\bar{\mu}\bar{\nu}} = e^{[\bar{\mu}}_b \partial_a e^{\bar{\nu}]b} + e^{[\bar{\mu}|b}\partial^{|\bar{\nu}]} g_{ab} = e^{\bar{\mu}}_b \nabla_a e^{\bar{\nu}b}, \tag{5.117}$$

where ∇_a is the familiar (torsion-free) metric-compatible derivative of general relativity. Thus, when there is no torsion, the covariant derivatives $D_{\bar{\mu}}$ are equivalent to ∇_a,

$$D_{\bar{\mu}} v_{\bar{\nu}} = e^a_{\bar{\mu}} e^b_{\bar{\nu}} \nabla_a v_b \quad \text{etc.}, \tag{5.118}$$

and thus the curvature tensor $R_{ab}{}^{\bar{\nu}}{}_{\bar{\mu}}$ is equivalent to its counterpart from general relativity.

5.A.3 Dynamics

The dynamics of the tetrad, spin connection, and matter fields are determined by a Lagrangian $\mathcal{L}_{\mathrm{EC}}$ that closely resembles that of the Einstein-Hilbert action:

$$\mathcal{L}_{\mathrm{EC}} = -eR/\kappa + \mathcal{L}_{\mathrm{matter}}, \tag{5.119}$$

where $\mathcal{L}_{\mathrm{matter}}$ is the (covariantised) matter Lagrangian and $e \equiv \det(e^{\bar{\mu}}_a) = 1/\det(e^a_{\bar{\mu}}) = \sqrt{-g}$ is the volume element. Variation with respect to $e^a_{\bar{\mu}}$ and $\omega_a{}^{\bar{\mu}\bar{\nu}}$ generates the gravitational field equations,

$$G_a{}^{\bar{\mu}} = \kappa T_a{}^{\bar{\mu}}, \tag{5.120a}$$

$$F^a{}_{\bar{\mu}\bar{\nu}} = \kappa S^a{}_{\bar{\mu}\bar{\nu}}, \tag{5.120b}$$

where matter's energy-momentum tensor $T_a{}^{\bar{\mu}}$, and spin tensor $S^a{}_{\bar{\mu}\bar{\nu}}$, are the conjugate currents of the translational and rotational gauge fields,

$$T_a{}^{\bar{\mu}} \equiv \frac{1}{2e}\frac{\delta\mathcal{L}_{\mathrm{matter}}}{\delta e^a_{\bar{\mu}}}, \qquad\qquad S^a{}_{\bar{\mu}\bar{\nu}} \equiv \frac{1}{e}\frac{\delta\mathcal{L}_{\mathrm{matter}}}{\delta\omega_a{}^{\bar{\mu}\bar{\nu}}}, \tag{5.121}$$

and we have written

$$G_a{}^{\bar{\mu}} \equiv R_a{}^{\bar{\mu}} - e^{\bar{\mu}}_a R/2, \tag{5.122a}$$

$$F^a{}_{\bar{\mu}\bar{\nu}} \equiv 2e^{-1} D_b(e e^a_{[\bar{\mu}} e^b_{\bar{\nu}]}) = \mathcal{T}^a_{\bar{\mu}\bar{\nu}} + 2e^a_{[\bar{\mu}}\mathcal{T}^{\bar{\alpha}}_{\bar{\nu}]\bar{\alpha}}. \tag{5.122b}$$

Consequently, the energy-momentum of matter generates curvature, and the intrinsic spin of matter generates torsion. When $S^a{}_{\bar{\mu}\bar{\nu}} = 0$ everywhere, the second field equation (5.120b) ensures that torsion will vanish also; on substitution of (5.117) and (5.102), the first field equation (5.120a) then becomes the usual Einstein field equations.

5.A.4 Perturbations

When the curvature and torsion of the physical spacetime are small, it is often convenient to represent the gravitational fields as perturbations from a flat torsion-free background $(\check{\mathcal{M}}, \check{e}^a_{\bar{\mu}}, \check{\omega}_a{}^{\bar{\mu}\bar{\nu}})$. This spacetime is equipped with a constant tetrad

$$\check{e}^a_{\bar{\mu}} = \delta^a_\mu \equiv (\partial/\partial x^\mu)^a,$$

$$\check{e}^{\bar{\mu}}_a = \delta^\mu_a \equiv (\mathrm{d}x^\mu)_a, \tag{5.123a}$$

defined by a Lorentzian coordinate system $\{x^\mu\}$, and a vanishing spin connection,

$$\breve{\omega}_a{}^{\bar{\mu}\bar{\nu}} = 0. \tag{5.123b}$$

In the background spacetime it is customary to manipulate indices using the background tetrad; because $\breve{e}^\mu_{\bar{\nu}} = \delta^\mu_\nu$, the distinction between barred indices and unbarred indices can then be dropped.

Mapping the physical spacetime $(\mathcal{M}, e^a_{\bar{\mu}}, \omega_a{}^{\bar{\mu}\bar{\nu}})$ onto the background with a diffeomorphism $\phi : \mathcal{M} \to \breve{\mathcal{M}}$, we define an asymmetric tensor field $f^a{}_\mu$ as a perturbation in the physical tetrad,

$$\phi^* e^a_{\bar{\mu}} = \delta^a_\mu - f^a{}_\mu/2, \tag{5.124}$$

and the tensor field $w_a{}^{\mu\nu} = -w_a{}^{\nu\mu}$ as a perturbation in the physical spin connection,

$$\phi^* \omega_a{}^{\bar{\mu}\bar{\nu}} = w_a{}^{\mu\nu}. \tag{5.125}$$

The perturbation in the tetrad (5.124) will be accompanied by a perturbation in the inverse tetrad,

$$\phi^* e^{\bar{\mu}}_a = \delta^\mu_a + f^\mu{}_a/2 + f^\mu{}_\nu f^\nu{}_a/4 + O(f^3), \tag{5.126}$$

which in turn defines a perturbation in the physical metric:

$$\begin{aligned} \phi^* g_{ab} &\equiv (\phi^* e^{\bar{\mu}}_a)(\phi^* e_{\bar{\mu}b}) \\ &= \breve{g}_{ab} + f_{(ab)} + f_{(a|}{}^\mu f_{\mu|b)}/2 + f^\mu{}_a f_{\mu b}/4 + O(f^3). \end{aligned} \tag{5.127}$$

Thus, in the linear approximation, we can identify the symmetric part of $f_{\mu\nu}$ with the metric perturbation $h_{\mu\nu}$ of general relativity:

$$h_{\mu\nu} \equiv f_{(\mu\nu)} + O(f^2). \tag{5.128}$$

Working to first order in $f_{\mu\nu}$ and $w_a{}^{\mu\nu}$, the Einstein-Cartan field equations (5.120) take the following form:

$$2\partial_{[\mu} w_{\alpha]\nu}{}^\alpha - \eta_{\mu\nu}\partial_\alpha w_\beta{}^{\alpha\beta} = \kappa T_{\mu\nu}, \tag{5.129a}$$

$$\partial_{[\mu} f^\alpha{}_{\nu]} + 2w_{[\mu}{}^\alpha{}_{\nu]} + \delta^\alpha_{[\mu|}(\partial_{|\nu]} f - \partial_\beta f^\beta{}_{|\nu]} - 2w_\beta{}^\beta{}_{|\nu]}) = \kappa S^\alpha{}_{\mu\nu}, \tag{5.129b}$$

where, for the sake of notational brevity, we have dropped the ϕ^* from the tensors $\phi^* T_a{}^{\bar{\mu}}$ and $\phi^* S^a{}_{\bar{\mu}\bar{\nu}}$. To recover the linearised field equations of general relativity, we need only set $S^\alpha{}_{\mu\nu} = 0$: the solution to equation (5.129b) is then

$$w_\alpha{}^{\mu\nu} = (\partial^{[\nu} f^{\mu]}{}_\alpha + \partial^{[\nu} f_\alpha{}^{\mu]} + \partial_\alpha f^{[\nu\mu]})/2, \tag{5.130}$$

which can be substituted into (5.129a) to retrieve

$$\widehat{G}_{\mu\nu}{}^{\alpha\beta} f_{(\alpha\beta)} = \kappa T_{\mu\nu}. \tag{5.131}$$

The equivalence between $f_{\mu\nu}$ and $h_{\mu\nu}$ can be strengthened further by fixing the rotation gauge freedom (to linear order) with the condition

$$f_{[\mu\nu]} = O(f^2). \tag{5.132}$$

This can always be achieved by a local rotation (5.106) of the form

$$\Lambda^\mu{}_\nu = \delta^\mu_\nu + (f^\mu{}_\nu - f_\nu{}^\mu)/4 + O(f^2), \tag{5.133}$$

which is indeed a valid Lorentz transformation,

$$\Lambda^\alpha{}_\mu \Lambda^\beta{}_\nu \eta_{\alpha\beta} = \eta_{\mu\nu} + O(f^2), \tag{5.134}$$

and has the desired effect:

$$f_{\mu\nu} \to f_{(\mu\nu)} + O(f^2). \tag{5.135}$$

In this "symmetric" gauge, the torsion-free spin connection (5.130) is simply

$$w_\alpha{}^{\mu\nu} = \partial^{[\nu} f^{\mu]}{}_\alpha + O(f^2), \tag{5.136}$$

and the relationship

$$h_{\mu\nu} = f_{\mu\nu} + O(f^2) \tag{5.137}$$

ensures that $f_{\mu\nu}$ and $h_{\mu\nu}$ are equivalent to linear order.

5.B Appendix: An Identity

Here we derive an identity that relates the Belinfante tensor (5.30) to the Einstein tensor of physical spacetime. First, for notational purposes, let us define a tensor $\widetilde{G}^{(2)}{}_a{}^b$ to represent the quadratic part of the "mixed" Einstein tensor density:

$$\widetilde{G}^{(2)}{}_a{}^b \equiv \left[\phi^*(\sqrt{-g} G_a{}^b)\right]^{(2)}. \tag{5.138}$$

This definition expands to give

$$\widetilde{G}^{(2)}_{\mu\nu} = R^{(2)}_{\mu\nu} - R^{(1)}_{\mu\alpha}\bar{h}^\alpha{}_\nu - \eta_{\mu\nu}(R^{(2)} - R^{(1)}_{\alpha\beta}\bar{h}^{\alpha\beta})/2, \tag{5.139}$$

where

$$R^{(1)}_{ab} \equiv [\phi^* R_{ab}]^{(1)}, \qquad R^{(2)}_{ab} \equiv [\phi^* R_{ab}]^{(2)}, \tag{5.140}$$

are the linear and quadratic parts of the Ricci tensor, when expanded according to (5.31):[24]

$$R^{(1)}_{\mu\nu} = \partial_\alpha \partial_{(\mu} h_{\nu)}{}^\alpha - \partial^2 h_{\mu\nu}/2 - \partial_\mu \partial_\nu h/2, \tag{5.141a}$$

$$\begin{aligned} R^{(2)}_{\mu\nu} = &-h^{\alpha\beta}(2\partial_\alpha\partial_{(\mu}h_{\nu)\beta} - \partial_\mu\partial_\nu h_{\alpha\beta} - \partial_\alpha\partial_\beta h_{\mu\nu})/2 + \partial_\mu h^{\alpha\beta}\partial_\nu h_{\alpha\beta}/4 \\ &+ \partial^\alpha h^\beta{}_\mu(\partial_{[\alpha}h_{\beta]\nu}) - \partial_\alpha \bar{h}^{\alpha\beta}(\partial_{(\mu}h_{\nu)\beta} - \partial_\beta h_{\mu\nu}/2). \end{aligned} \tag{5.141b}$$

[24] The tensors $R^{(1)}_{ab}$, $R^{(2)}_{ab}$, $\widetilde{G}^{(2)}{}_a{}^b$ are defined on the background, and thus their indices are raised and lowered with $\breve{g}_{ab}$; the indices of physical tensors R_{ab} and G_{ab} are moved with the physical metric g_{ab}, however. Because of this, index manipulations do not commute with the $[\ldots]^{(2)}$ operation that isolates the quadratic terms: $R^{(2)} \equiv [\phi^* R_{ab}]^{(2)}\, \breve{g}^{ab} \neq [\phi^* R]^{(2)}$, for example.

Now consider the tensor

$$Q_{\mu\nu} \equiv R^{(2)}_{\mu\nu} - R^{(1)}_{\mu\alpha}\bar{h}^{\alpha}{}_{\nu} + \frac{1}{2}\partial_\alpha\partial_{(\mu}(h_{\nu)\beta}h^{\beta\alpha}) - \frac{1}{4}\partial^2(h_{\mu\alpha}h^{\alpha}{}_{\nu}) - \frac{1}{4}\partial_\mu\partial_\nu(h_{\alpha\beta}h^{\alpha\beta}), \tag{5.142}$$

the trace-reverse of which is

$$\begin{aligned}\bar{Q}_{\mu\nu} &= R^{(2)}_{\mu\nu} - R^{(1)}_{\mu\alpha}\bar{h}^{\alpha}{}_{\nu} - \frac{1}{2}\eta_{\mu\nu}(R^{(2)} - R^{(1)}_{\alpha\beta}\bar{h}^{\alpha\beta}) + \frac{1}{2}\partial_\alpha\partial_{(\mu}(h_{\nu)\beta}h^{\beta\alpha}) - \frac{1}{4}\partial^2(h_{\mu\alpha}h^{\alpha}{}_{\nu}) \\ &\quad - \frac{1}{4}\partial_\mu\partial_\nu(h_{\alpha\beta}h^{\alpha\beta}) - \frac{1}{4}\eta_{\mu\nu}(\partial_\alpha\partial_\gamma(h^{\gamma}{}_{\beta}h^{\beta\alpha}) - \partial^2(h_{\alpha\beta}h^{\alpha\beta})) \\ &= \widetilde{G}^{(2)}_{\mu\nu} + \frac{1}{2}\widehat{G}_{\mu\nu}{}^{\alpha\beta}(h_{\alpha\gamma}h^{\gamma}{}_{\beta}).\end{aligned} \tag{5.143}$$

Substituting equations (5.141) into (5.142), one finds

$$\begin{aligned}Q_{\mu\nu} &= -\frac{1}{2}h^{\alpha\beta}(\partial_\alpha\partial_{(\mu}h_{\nu)\beta} - \partial_\alpha\partial_\beta h_{\mu\nu}) - \frac{1}{2}h_{\alpha(\mu}(\partial^\alpha\partial^\beta h_{\nu)\beta} - \partial_{\nu)}\partial_\alpha h) \\ &\quad + \frac{1}{4}h(2\partial^\alpha\partial_{(\mu}h_{\nu)\alpha} - \partial^2 h_{\mu\nu} - \partial_\mu\partial_\nu h) - \frac{1}{4}\partial_\mu h^{\alpha\beta}\partial_\nu h_{\alpha\beta} - \frac{1}{2}\partial^\alpha h^{\beta}{}_{\mu}\partial_\beta h_{\nu\alpha} \\ &\quad - \frac{1}{2}\partial_\alpha h^{\alpha\beta}(\partial_{(\mu}h_{\nu)\beta} - \partial_\beta h_{\mu\nu}) + \frac{1}{4}\partial^\alpha h(2\partial_{(\mu}h_{\nu)\alpha} - \partial_\alpha h_{\mu\nu}) \\ &\quad + \frac{1}{2}\partial_\alpha h_{\beta(\mu}\partial_{\nu)}h^{\alpha\beta} - h^{\alpha}{}_{[\nu}R^{(1)}_{\mu]\alpha} \\ &= -\kappa\bar{t}_{\mu\nu},\end{aligned} \tag{5.144}$$

the last line of which can be confirmed by expanding out all the trace-reversed fields on the right-hand side of (5.30) and observing that $h^{\alpha}{}_{[\nu}R^{(1)}_{\mu]\alpha} = h^{\alpha}{}_{[\nu}G^{(1)}_{\mu]\alpha} = h^{\alpha}{}_{[\nu}\widehat{G}_{\mu]}{}^{\alpha\beta\gamma}h_{\beta\gamma}$. Comparing (5.143) with (5.144), we conclude that the following identity

$$\kappa t_{\mu\nu} = -\widetilde{G}^{(2)}_{\mu\nu} - \widehat{G}_{\mu\nu}{}^{\alpha\beta}(h_{\alpha\gamma}h^{\gamma}{}_{\beta})/2, \tag{5.145}$$

is valid for all $h_{\mu\nu}$.

5.C Appendix: ADM Energy-Momentum

In this chapter, and the those that have preceded it, we have focussed on the *local* aspects of gravitational energy-momentum and spin; although we will not attempt a thorough investigation here, it will be valuable to briefly examine the *global* energy and momentum that our framework defines, and compare these quantities to the well-known results of Arnowitt, Deser, and Misner (ADM) [6, 8, 9].

Recall that, as seen in (5.29), the Belinfante tensor $t_{\mu\nu}$ defines the same *total* energy, momentum, and angular momentum as $\tau_{\mu\nu}$ and $s^{\alpha}{}_{\mu\nu}$.[25] Thus, for the purposes of this appendix, we are free to use whichever set of tensors is convenient, and the results we derive will carry over to the other. With this in mind, let us define the total energy-momentum of gravity and matter by

$$\mathfrak{P}_\mu \equiv \int \mathrm{d}^3y\sqrt{-g}\left(T^{\mathrm{Bel}}{}_a{}^b + (\phi^{-1})^* t_a{}^b\right)(\mathrm{d}/\mathrm{d}y^\mu)^a(\mathrm{d}y^0)_b, \tag{5.146}$$

[25] To be precise: the integrals in (5.29) may in fact differ by surface terms quadratic in $h_{\mu\nu}$. However, as we will see, these can be neglected in comparison to the surface terms linear in $h_{\mu\nu}$ when the boundary of the integral is taken to spatial infinity.

where $y^\mu \equiv (\phi^{-1})^* x^\mu$ are the images of the Lorentzian coordinates $\{x^\mu\}$ in the physical spacetime.[26] To this definition we now apply the Einstein field equations and evaluate the integral in terms of background quantities:

$$\begin{aligned}\mathfrak{P}_\mu &= \frac{1}{\kappa}\int \mathrm{d}^3y\sqrt{-g}\left(G_a{}^b + \kappa(\phi^{-1})^* t_a{}^b\right)(\mathrm{d}/\mathrm{d}y^\mu)^a(\mathrm{d}y^0)_b \\ &= \frac{1}{\kappa}\int \mathrm{d}^3x\left(\phi^*(\sqrt{-g}G_a{}^b) + \kappa t_a{}^b\right)(\mathrm{d}/\mathrm{d}x^\mu)^a(\mathrm{d}x^0)_b \\ &= \frac{1}{\kappa}\int \mathrm{d}^3x\left(\widehat{G}_\mu{}^{0\alpha\beta}h_{\alpha\beta} + \widetilde{G}^{(2)}{}_\mu{}^0 + \kappa t_\mu{}^0\right),\end{aligned} \tag{5.147}$$

where, in the last line, we have expanded the metric according to (5.31) and neglected terms $O(h^3)$ under the assumption that the gravitational field is everywhere weak enough that the quadratic approximation to general relativity will suffice. We now use the identity (5.41) to write the energy-momentum as

$$\mathfrak{P}_\mu = \frac{1}{\kappa}\int \mathrm{d}^3x\widehat{G}_\mu{}^{0\alpha\beta}\left(h_{\alpha\beta} + O(h^2)\right). \tag{5.148}$$

Although terms $O(h^2)$ cannot be neglected in general (otherwise $t_{\mu\nu}$ should never have appeared in the integral (5.146) to begin with) we note that all the terms in (5.148) are total spatial derivatives, so $\mathfrak{P}_\mu$ will depend only on the behaviour of the gravitational field on the boundary of the integral. Thus, as the limit is taken in which this boundary moves to spatial infinity, we will require the linear surface terms $\partial h \sim 1/r^2$ in order that integral be finite, and as a consequence the quadratic surface terms $h\partial h \sim 1/r^3$ will be negligible in comparison.

Let us first consider the total energy of the system:

$$\begin{aligned}\mathfrak{P}^0 &= \frac{1}{\kappa}\int \mathrm{d}^3x\widehat{G}^{00\alpha\beta}h_{\alpha\beta} \\ &= \frac{1}{2\kappa}\int \mathrm{d}^3x(\partial_i\partial_j h_{ij} - \partial_i\partial_i h_{jj}) \\ &= \frac{1}{2\kappa}\int \mathrm{d}^2S_i(\partial_j h_{ij} - \partial_i h_{jj}),\end{aligned} \tag{5.149}$$

which the reader will recognise as the ADM mass [6, 9]. Furthermore, the total linear

[26] In equation (5.146), the contraction between the energy-momentum tensors and the covector $(\mathrm{d}y^0)_a$ defines, in the usual way, the energy-momentum *densities* on the surface of integration $y^0 = \text{const}$. The vectors $(\mathrm{d}/\mathrm{d}y^\mu)^a$ have assumed the role of killing vectors in the absence of an exact spacetime symmetry; these same vectors were denoted by e^a_μ in chapters 3 and 4, but this symbol is now being used for the Einstein-Cartan tetrad.

momentum

$$\begin{aligned}
\mathfrak{P}_i &= \frac{1}{\kappa}\int \mathrm{d}^3x \widehat{G}_i{}^{0\alpha\beta} h_{\alpha\beta} \\
&= \frac{-1}{2\kappa}\int \mathrm{d}^3x(\partial_j \dot{h}_{ij} - \partial_k\partial_k h_{0i} - \partial_i \dot{h}_{jj} + \partial_i\partial_j h_{0j}) \\
&= \frac{-1}{2\kappa}\int \mathrm{d}^2S_j(\dot{h}_{ij} - \partial_j h_{0i} - \delta_{ij}\dot{h}_{kk} + 2\delta_{ij}\partial_k h_{0k} - \partial_i h_{0j}) \\
&= \frac{-1}{\kappa}\int \mathrm{d}^2S_j(\Gamma^{(1)0}{}_{ij} - \delta_{ij}\Gamma^{(1)0}{}_{kk}) \\
&= \frac{-1}{\kappa}\int \mathrm{d}^2S_j \pi^{(1)}_{ij}, \qquad (5.150)
\end{aligned}$$

which is the familiar expression for the ADM momentum [6, 9] truncated at linear order. Thus, when the terms $O(h^3)$ can be neglected from the field equations, and the terms $O(h^2)$ can be neglected at spatial infinity, our gravitational Belinfante tensor $t_{\mu\nu}$ defines exactly the same total energy and momentum as ADM. Moreover, as we have already explained, these results would also arise if we had defined $\mathfrak{P}_\mu$ using $\tau_{\mu\nu}$, rather than the Belinfante tensor $t_{\mu\nu}$. Thus, $\tau_{\mu\nu}$ is able to cast the global information present in the ADM energy-momentum in terms of a local description with physically sensible properties, including gravitational energy-density that is nowhere negative and gravitational energy-flux that is nowhere spacelike. Although at present this idea is limited by the restriction of our results to the quadratic approximation to general relativity, if we can extend our framework to the full theory, while maintaining the positivity properties of $\tau_{\mu\nu}$ and consistency with ADM energy-momentum, then this observation has the potential to shed further light on the global positivity properties of ADM energy-momentum: $\mathfrak{P}^0 \geq 0$ and $\mathfrak{P}^\mu\mathfrak{P}_\mu \leq 0$ [24, 73, 82].[27]

Although the ADM energy-momentum is usually represented in terms of the asymptotic behaviour of the gravitational field, as above, the reader should also be aware that these global quantities can be cast as spatial integrals of a gravitational Belinfante tensor $t^{\mathrm{ADM}}_{\mu\nu}$ that emerges from the canonical formalism [6, 8, 9]. Although ADM did not propose that this tensor should be interpreted as a physically meaningful local measure of gravitational energy-momentum, it is nonetheless interesting to compare the quadratic part of $t^{\mathrm{ADM}}_{\mu\nu}$ with $\tau_{\mu\nu}$ and $t_{\mu\nu}$. The strongest resemblance occurs when we employ our gauge-fixing procedure (that is, we insist that $h_{\mu\nu}$ be transverse-traceless) and examine the $(0,0)$ components of the tensors:

$$\begin{aligned}
\tau_{00} = t_{00} &= \frac{1}{8\kappa}(\dot{h}_{ij}\dot{h}_{ij} + \partial_k h_{ij}\partial_k h_{ij}) \\
&= t^{\mathrm{ADM}}_{00} + O(h^3). \qquad (5.151)
\end{aligned}$$

Remarkably, we find that our gauge-fixed τ_{00} and t_{00} are in fact equal to the ADM

[27] The local positivity of $\tau_{\mu\nu}$ (and the Dominant Energy Condition for $T^{\mathrm{Bel}}_{\mu\nu}$) should be enough to guarantee these global inequalities, as a sum of future-directed non-spacelike vectors will itself be future-directed and non-spacelike.

"Hamiltonian density" t^{ADM}_{00} when working to quadratic order.[28] This is a rather surprising correspondence, particularly when one considers how little our framework has in common with the canonical 3+1 approach from which t^{ADM}_{00} arose. Note, however, that this equality does not extend to the other components of the tensors:

$$\kappa\tau_{i0} = \frac{1}{4}\dot{h}_{jk}\partial_i h_{jk} \tag{5.152a}$$

$$\kappa t_{i0} = \frac{1}{4}\left(\dot{h}_{jk}\partial_i h_{jk} - \dot{h}_{jk}\partial_j h_{ik} + h_{jk}\partial_j \dot{h}_{ki}\right) \tag{5.152b}$$

$$\kappa t^{\mathrm{ADM}}_{i0} = \frac{1}{4}\left(\dot{h}_{jk}\partial_i h_{jk} - 2\dot{h}_{jk}\partial_j h_{ik}\right) + O(h^3). \tag{5.152c}$$

Hence, $\tau_{\mu\nu}$, $t_{\mu\nu}$, and $t^{\mathrm{ADM}}_{\mu\nu}$ do not give rise to equivalent local descriptions of gravitational energy-momentum, even in transverse-traceless gauge; moreover, even though $t^{\mathrm{ADM}}_{\mu\nu}$ succeeds in defining a positive gravitational energy-density $t^{\mathrm{ADM}}_{00} \geq 0$ as seen by observers at rest with respect to the TT-frame, it does not display the full (Lorentz-invariant) positivity properties of $\tau_{\mu\nu}$: the energy-density $v^\mu v^\nu t^{\mathrm{ADM}}_{\mu\nu}$ (as seen by an observer moving with 4-velocity v^μ) may be negative, and the energy-flux $v^\mu t^{\mathrm{ADM}}_{\mu\nu}$ may be spacelike.

The formulae (5.152) also serve as another starting point from which to verify that all three tensors define the same total momentum $\mathfrak{P}_i$: those terms in t_{i0} and t^{ADM}_{i0} which do not appear in τ_{i0} can be integrated by parts (discarding a quadratic surface term, as usual) and then vanish due to the gauge condition $\partial_i h_{ij} = 0$. Of course, these terms *do* contribute to the total *angular momentum*: they correspond to the divergence $\partial_\alpha(s_{\mu\nu}{}^{\alpha} + s_{\nu\mu}{}^{\alpha} - s^{\alpha}{}_{\mu\nu})/2$ that packages intrinsic spin into the Belinfante tensor (5.27). Accordingly, when one neglects quadratic surface terms, one finds that

$$\begin{aligned}\int (2x_{[i}\tau_{j]}{}^0 + s^0{}_{ij})\mathrm{d}^3x &= \int 2x_{[i}t_{j]}{}^0\mathrm{d}^3x \\ &= \int 2x_{[i}t^{\mathrm{ADM}}{}_{j]}{}^0\mathrm{d}^3x + O(h^3),\end{aligned} \tag{5.153}$$

confirming that $\tau_{\mu\nu}$ and $s^\alpha{}_{\mu\nu}$ also give the same global description of angular momentum as ADM at second order.

In summary, whenever general relativity can be approximated to quadratic order, and quadratic surface terms can be neglected at spatial infinity, our gravitational energy-momentum tensor $\tau_{\mu\nu}$ and spin tensor $s^\alpha{}_{\mu\nu}$ provide the same *global* description of energy, momentum, and angular momentum as ADM, but localise these quantities in a physically sensible fashion, displaying positive gravitational energy-density, causal energy-flux, and traceless spatial spin.

[28] We thank Richard Arnowitt and Stanley Deser for bringing this to our attention.

Chapter 6

Closing Remarks

Within the linear approximation to general relativity, I have uncovered a compelling localisation of gravitational energy, momentum, and spin. The gravitational energy-momentum tensor $\tau_{\mu\nu}$ and spin tensor $s^{\alpha}{}_{\mu\nu}$ account for the exchange of energy, momentum and angular momentum between matter and linear gravity, constitute translational and rotational Noether currents of the linearised gravitational field, and generate gravity in quadratic approximations to the Einstein and Einstein-Cartan field equations. Moreover, the tensors motivate a natural gauge-fixing programme (removing all ambiguity from the description) and display numerous desirable properties befitting their physical interpretation, such as positive energy-density, causal energy-flux, and traceless spatial spin.

At the heart of the framework lie the formulae

$$\kappa\bar{\tau}_{\mu\nu} = \tfrac{1}{4}\partial_\mu h_{\alpha\beta}\partial_\nu \bar{h}^{\alpha\beta}, \tag{6.1a}$$

$$\kappa s^{\alpha}{}_{\mu\nu} = 2\bar{h}_{\beta[\nu}\partial^{[\alpha}\bar{h}_{\mu]}{}^{\beta]}, \tag{6.1b}$$

for the energy-momentum tensor and spin tensor of the linearised gravitational field. These were obtained, together with the harmonic gauge condition

$$\partial^\mu \bar{h}_{\mu\nu} = 0, \tag{6.2}$$

by considering the local exchange of energy, momentum, and angular momentum between matter and linear gravity. The gravitational energy-momentum tensor (6.1a) is the *unique* symmetric tensor, quadratic in $h_{\mu\nu}$ and free of second derivatives, which accounts for the energy-momentum exchanged locally with matter, as quantified by equation (3.10). This solution only exists when the field condition (6.2) is met, which has the highly beneficial effect of constraining the gauge of the gravitational field, without limiting the physical applicability of our results.

The spin tensor (6.1b) represents a collection of intrinsic spin current-densities; in combination with the "orbital" angular momentum defined by $\tau_{\mu\nu}$, this tensor determines the gravitational angular momentum current-densities

$$j_{\mu\nu}{}^{\alpha} = 2x_{[\mu}\tau_{\nu]}{}^{\alpha} + s^{\alpha}{}_{\mu\nu}, \tag{6.3}$$

which account for the local exchange of angular momentum with matter (4.14). In order to obtain a unique formula for the spin tensor, it was necessary to demand that ${s^\alpha}_{\mu\nu}$ obey two simple conditions: first, the spin of a gravitational plane-wave must flow in the direction of propagation of the wave (4.24), and second, the gravitational field must be free of infinite pressure gradients (4.34).

The tensors $\tau_{\mu\nu}$ and ${s^\alpha}_{\mu\nu}$ display an array of notable properties in various regimes, a summary of which can be found in table 6.1. The most remarkable of these is the *positivity property* of $\tau_{\mu\nu}$: for all observers, whenever $h_{\mu\nu}$ is transverse-traceless, the gravitational energy-density is non-negative, and gravitational energy-flux is timelike or null. This is an extremely valuable property that has not been seen in any previous localisation of gravitational energy-momentum. It extends the *dominant energy condition* to include the energy-momentum of the gravitational field, and ensures that $\tau_{\mu\nu}$ provides a description of gravitational energy-momentum that makes intuitive sense on local scales. It is also fortuitous that the positivity of $\tau_{\mu\nu}$ is displayed when the field is *transverse-traceless*, as transverse-tracelessness can always be achieved (for the dynamical part of the field at least) by a gauge transformation, extinguishing the small amount of gauge freedom that remains after enforcing the harmonic gauge condition (6.2).

This transverse-traceless gauge-fixing programme was further motivated by an analysis of the energy-momentum (and angular momentum) transferred onto an infinitesimal detector,

$$\begin{aligned} T_{00} &= M\delta(\vec{x}) + \tfrac{1}{2} I_{ij}\partial_i\partial_j\delta(\vec{x}), \\ T_{0i} &= \tfrac{1}{2}(\dot{I}_{ij} - L_{ij})\partial_j\delta(\vec{x}), \\ T_{ij} &= \tfrac{1}{2}\ddot{I}_{ij}\delta(\vec{x}). \end{aligned} \tag{6.4}$$

By splitting the incident gravitational field into a series of pulses, and averaging over the infinitesimal interaction region (4.42), one arrives at a local and completely gauge-invariant measure of the energy, momentum and angular momentum exchanged with the detector:

$$\langle \partial^\mu \tau_{\mu\nu} \rangle^{\not{M}}_{\int H} = -\tfrac{1}{4}\delta(\vec{x})\ddot{I}_{ij}\partial_\nu h^{\mathrm{TT}}_{ij}, \tag{6.5a}$$

$$\langle \partial_\alpha {j_{ij}}^\alpha \rangle_{\int \delta} = \delta(\vec{x}) h^{\mathrm{TT}}_{k[i}\ddot{I}_{j]k}. \tag{6.5b}$$

The same results can be achieved, without the microaveraging process, simply by insisting that the incident field be expressed in transverse-traceless gauge. Thus, the transverse-traceless programme removes the final gauge-ambiguity of $\tau_{\mu\nu}$ and ${s^\alpha}_{\mu\nu}$ in a manner that is consistent with the gauge-invariant energy-momentum absorbed by infinitesimal detectors, whilst simultaneously ensuring that the gravitational energy-density is positive and the gravitational energy-flux is causal.

In addition to accounting for the energy-momentum and angular momentum lost and gained by matter, $\tau_{\mu\nu}$ and ${s^\alpha}_{\mu\nu}$ play two other major roles within the theory. Firstly, they are Noether currents associated with the translational and rotational symmetry of linearised general relativity. Secondly, they occur as the quadratic terms in perturbative expansions of the Einstein and Einstein-Cartan field equations, generating gravity alongside material energy-momentum and spin.

I demonstrated that $\tau_{\mu\nu}$ and ${s^\alpha}_{\mu\nu}$ were indeed Noether currents by constructing a Lagrangian

$$\mathcal{L} \equiv \frac{e}{4\kappa}\left(\check{D}_{\bar{\mu}} h_{\bar{\alpha}\bar{\beta}} \check{D}^{\bar{\mu}} \bar{h}^{\bar{\alpha}\bar{\beta}} - 2\check{D}_{\bar{\mu}} \bar{h}^{\bar{\mu}\bar{\alpha}} \check{D}_{\bar{\nu}} {\bar{h}^{\bar{\nu}}}_{\bar{\alpha}} + 2\bar{h}^{\bar{\mu}\bar{\nu}} \check{R}_{\bar{\alpha}\bar{\mu}\bar{\nu}\bar{\beta}} \bar{h}^{\bar{\alpha}\bar{\beta}}\right), \tag{6.6}$$

which "gauged" the translational and rotational symmetries of the Fierz-Pauli massless spin-2 field. This Lagrangian prescribes the field equations of linear gravity (5.5) and, according to standard variational techniques (5.12), defines Noether currents

$$\kappa\bar{\tau}_{\mu\nu} = \tfrac{1}{4}\partial_\mu h_{\alpha\beta}\partial_\nu \bar{h}^{\alpha\beta} - \tfrac{1}{2}\partial_\mu \bar{h}_{\nu\alpha}\partial_\beta \bar{h}^{\alpha\beta}, \tag{6.7a}$$

$$\kappa {s^\alpha}_{\mu\nu} = 2\bar{h}_{\beta[\nu}\partial^{[\alpha}{\bar{h}_{\mu]}}^{\beta]} + \delta^\alpha_{[\nu}{\bar{h}_{\mu]}}^{\beta}\partial_\gamma {\bar{h}^\gamma}_\beta, \tag{6.7b}$$

which reduce to the familiar formulae (6.1) in harmonic gauge (6.2). In fact, these Noether currents are the *unique* generalisation of $\tau_{\mu\nu}$ and ${s^\alpha}_{\mu\nu}$ that can be derived from a covariantised Fierz-Pauli Lagrangian, in which $\tau_{\mu\nu}$ remains free of second derivatives.

To reveal the role played by $\tau_{\mu\nu}$ and ${s^\alpha}_{\mu\nu}$ in the *dynamics* of general relativity, I considered the following expansion of the metric:

$$\phi^* g_{ab} = \breve{g}_{ab} + h_{ab} + h_{ac}{h^c}_b/2. \tag{6.8}$$

Under this expansion, the quadratic approximation of the vacuum Einstein field equations (5.35) was seen to take the form

$$\begin{aligned}\widehat{G}_{\mu\nu}{}^{\alpha\beta} h_{\alpha\beta} &= \kappa\left(\tau_{\mu\nu} + \partial_\alpha({s_{\mu\nu}}^\alpha + {s_{\nu\mu}}^\alpha - {s^\alpha}_{\mu\nu})/2\right) \\ &= \kappa t_{\mu\nu},\end{aligned} \tag{6.9}$$

in which $\tau_{\mu\nu}$ and ${s^\alpha}_{\mu\nu}$ combine to form a Belinfante energy-momentum tensor (5.30) and this combination acts as a source term in the gravitational wave-equation. In order to disentangle the combination of spin and energy-momentum on the right-hand side of equation (6.9), I performed a similar calculation within the Einstein-Cartan formalism. Under the field expansion

$$\phi^* e^a_{\bar{\mu}} = \delta^a_\mu - {f^a}_\mu/2 + {f^a}_\nu {f^\nu}_\mu/8, \tag{6.10a}$$

$$\phi^* {\omega_a}^{\bar{\mu}\bar{\nu}} = {w_a}^{\mu\nu} - f{w_a}^{\mu\nu}/2 + w_{a\beta}{}^{[\nu} f^{\mu]\beta} + f_\beta{}^{[\mu}\partial_a f^{\nu]\beta}/4, \tag{6.10b}$$

the quadratic approximation to the vacuum Einstein-Cartan field equations ((5.51) and (5.57)) were found to be

$$2\partial_{[\mu} w_{\alpha]\nu}{}^\alpha - \eta_{\mu\nu}\partial_\alpha {w_\beta}^{\alpha\beta} = \kappa\tau_{\mu\nu}, \tag{6.11a}$$

$$\partial_{[\mu}{f^\alpha}_{\nu]} + 2w_{[\mu}{}^\alpha{}_{\nu]} + \delta^\alpha_{[\mu|}(\partial_{|\nu]} f - \partial_\beta {f^\beta}_{|\nu]} - 2{w_\beta}^\beta{}_{|\nu]}) = \kappa {s^\alpha}_{\mu\nu}. \tag{6.11b}$$

Thus, in a theory of gravity which maintains the separate identities of material energy-momentum and spin, $\tau_{\mu\nu}$ and ${s^\alpha}_{\mu\nu}$ likewise occur as distinct objects in the dynamical equations.

Many key features of the framework were illuminated by the metric expansion (6.8). In particular, if one applies this perturbation to the Einstein-Hilbert action, then the Lagrangian of equation (6.6) arises as the quadratic part of the integrand, modulo surface terms. The expansion (6.8) is also notable in its own right, as it occupies a "central" point between the linear metric perturbation (5.31) and the linear *inverse* metric perturbation (5.34). The symmetry of this expansion can be extended to all orders by defining the metric according to an *exponential* gravitational field,

$$\phi^* g_{ab} \equiv [e^{h/2}]^c{}_a \breve{g}_{cd} [e^{h/2}]^d{}_b; \tag{6.12}$$

without further investigation, however, it is unclear whether this is definition is particularly suited for the localisation gravitational energy-momentum.

The results of this thesis are limited in so far as they only apply to general relativity and Einstein-Cartan gravity in their linear and quadratic approximations. At present, it is far from clear what approach should be taken to extend the framework beyond this regime, nor which properties of $\tau_{\mu\nu}$ and $s^\alpha{}_{\mu\nu}$ can be preserved in the full theory. That said, the new field definitions (6.8), (6.10), and (6.12) provide valuable clues as to how progress might be made. For instance, one might attempt to construct a Belinfante tensor for the full theory by expanding the vacuum Einstein field equations (5.35) according to the exponential gravitational field (6.12), identifying all terms at quadratic order and higher with $\kappa t_{\mu\nu}$. Following this, one would need to disentangle the spin and energy-momentum from within this Belinfante tensor; hopefully, this separation can be rendered unique by the requirement that $\tau_{\mu\nu}$ be free of second derivatives. Alternatively, one could keep $\tau_{\mu\nu}$ and $s^\alpha{}_{\mu\nu}$ separate from the start, rewriting the Einstein-Cartan formalism in terms of fields which reduce to (6.10) at quadratic order. Whatever route is taken, the main aim of the endeavour should be to maintain the notable properties of $\tau_{\mu\nu}$ and $s^\alpha{}_{\mu\nu}$ at higher orders; hopefully, some combination will be sufficient to determine the tensors completely, in addition to the gauge-fixing programme that will render them physically well-defined in the full theory.

Table 6.1: Properties of $\tau_{\mu\nu}$ and $s^{\alpha}{}_{\mu\nu}$ under various conditions.

	Harmonic Gauge	Transverse-Traceless	Harmonic Plane-Wave	Static
$h_{\mu\nu}$	$\partial^{\mu}\bar{h}_{\mu\nu} = 0$	$\partial_i h_{ij} = 0$ $h_{\mu 0} = 0$ $h = 0$	$h_{\mu\nu} = h_{\mu\nu}(k_\alpha x^\alpha), \quad k^\mu \bar{h}_{\mu\nu} = 0$ $\left(\begin{array}{c} k_\mu = (1, -1, 0, 0) \\ h_+ \equiv (h_{22} - h_{33})/2, \quad h_\times \equiv h_{23} \end{array} \right)$	$\bar{h}_{00} = -4\Phi(\vec{x})$ $\bar{h}_{\mu i} = 0$
$\tau_{\mu\nu}$	$\kappa\bar{\tau}_{\mu\nu} = \frac{1}{4}\partial_\mu h_{\alpha\beta}\partial_\nu \bar{h}^{\alpha\beta}$	$v^\mu \tau_{\mu\nu} v^\nu \geq 0$ $v^\mu \tau_{\mu\alpha}\tau^{\alpha}{}_{\nu} v^\nu \leq 0$ $\forall \quad v_\alpha v^\alpha \leq 0$	$\kappa\tau_{\mu\nu} = \frac{1}{2} k_\mu k_\nu (\dot{h}_+^2 + \dot{h}_\times^2)$	$\kappa\bar{\tau}_{\mu\nu} = 2\partial_\mu \Phi \partial_\nu \Phi$
$s^{\alpha}{}_{\mu\nu}$	$\kappa s^{\alpha}{}_{\mu\nu} = 2\bar{h}_{\beta[\nu}\partial^{[\alpha}\bar{h}_{\mu]}{}^{\beta]}$	$s^{\alpha}{}_{0i} = 0$ $s^{\alpha}{}_{\alpha\mu} = 0$	$s^{\alpha}{}_{\mu\nu} \propto k^\alpha$ $\kappa s^{\alpha}{}_{23} = k^\alpha (h_\times \dot{h}_+ - h_+ \dot{h}_\times)$	$s^{\alpha}{}_{\mu\nu} = 0$

References

[1] L. M. Butcher, A. N. Lasenby, and M. P. Hobson. Physical significance of the Babak-Grishchuk gravitational energy-momentum tensor. *Phys. Rev. D*, 78(6):064034, 2008. *Chapter 1 of this thesis.*

[2] L. M. Butcher, M. P. Hobson, and A. N. Lasenby. Bootstrapping gravity: A consistent approach to energy-momentum self-coupling. *Phys. Rev. D*, 80(8):084014, 2009. *Chapter 2 of this thesis.*

[3] L. M. Butcher, M. P. Hobson, and A. N. Lasenby. Localizing the energy and momentum of linear gravity. *Phys. Rev. D*, 82(10):104040, 2010. *Chapter 3 of this thesis.*

[4] L. M. Butcher, A. N. Lasenby, and M. P. Hobson. Localizing the angular momentum of linear gravity. Submitted to *Phys. Rev. D,* 2012 *Chapter 4 of this thesis.*

[5] L. M. Butcher, M. P. Hobson, and A. N. Lasenby. Localized energetics of linear gravity: Theoretical development. Submitted to *Phys. Rev. D,* 2012 *Chapter 5 of this thesis.*

[6] R. Arnowitt, S. Deser, and C. Misner. Republication of: The dynamics of general relativity. *Gen. Rel. Grav.*, 40:1997–2027, 2008. 10.1007/s10714-008-0661-1.

[7] R. Arnowitt, S. Deser, and C. W. Misner. Dynamical structure and definition of energy in general relativity. *Phys. Rev.*, 116(5):1322–1330, 1959.

[8] R. Arnowitt, S. Deser, and C. W. Misner. Canonical variables for general relativity. *Phys. Rev.*, 117(6):1595–1602, 1960.

[9] R. Arnowitt, S. Deser, and C. W. Misner. Energy and the criteria for radiation in general relativity. *Phys. Rev.*, 118(4):1100–1104, 1960.

[10] R. Arnowitt, S. Deser, and C. W. Misner. Coordinate invariance and energy expressions in general relativity. *Phys. Rev.*, 122(3):997, 1961.

[11] S. V. Babak and L. P. Grishchuk. Energy-momentum tensor for the gravitational field. *Phys. Rev. D*, 61(2):024038, 1999.

[12] J. M. Bardeen. Gauge-invariant cosmological perturbations. *Phys. Rev. D*, 22(8):1882–1905, 1980.

[13] R. Bartnik. New definition of quasilocal mass. *Phys. Rev. Lett.*, 62:2346–2348, 1989.

[14] L. Bel. *CR Acad. Sci. Paris*, 247, 1958.

[15] F. Belinfante. On the spin angular momentum of mesons. *Physica*, 6:887–898, 1939.

[16] G. D. Birkhoff. *Relativity and Modern Physics*, page 253. Harvard University Press, 1927.

[17] M. Blagojević. *Gravitation and Gauge Symmetries*, chapter 3. IOP, 2002.

[18] N. N. Bogoliubov and D. V. Shirkov. *Introduction to the theory of quantized fields*, chapter 1. John Wiley, 3rd edition, 1980.

[19] H. Bondi. Plane gravitational waves in general relativity. *Nature*, 179(8):1072–1073, 1957.

[20] H. Bondi, M. G. J. van der Burg, and A. W. K. Metzner. Gravitational waves in general relativity. vii. waves from axi-symmetric isolated systems. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences*, 269(1336):21–52, 1962.

[21] D. G. Boulware and S. Deser. Classical general relativity derived from quantum gravity. *Annals of Physics*, 89(1):193–240, 1975.

[22] D. R. Brill and J. B. Hartle. Method of the self-consistent field in general relativity and its application to the gravitational geon. *Phys. Rev.*, 135:B271–B278, 1964.

[23] M. Campanelli, C. O. Lousto, Y. Zlochower, and D. Merritt. Maximum gravitational recoil. *Phys. Rev. Lett.*, 98:231102, 2007.

[24] S. Deser. Timelike character of gravitational field energy-momentum. *Il Nuovo Cimento B*, 55:593–594, 1968.

[25] S. Deser. Self-interaction and gauge invariance. *Gen. Rel. Grav.*, 1(1):9–18, 1970.

[26] S. Deser. Gravity from self-interaction in a curved background. *Class. Quant. Grav.*, 4(4):L99–L105, 1987.

[27] S. Deser. Gravity from self-interaction redux. *Gen. Rel. Grav.*, 42:641–646, 2010.

[28] S. Deser and A. Waldron. Stability of massive cosmological gravitons. *Physics Letters B*, 508(3-4):347–353, 2001.

[29] P. A. M. Dirac. *General Theory of Relativity*, chapter 31. Princeton University Press, 1975.

[30] W. G. Dixon. 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):59–119, 1974.

[31] C. Doran and A. N. Lasenby. *Geometric Algebra for Physicists*, chapter 13.5.4. Cambridge University Press, 2003.

[32] C. J. L. Doran, A. N. Lasenby, A. D. Challinor, and S. F. Gull. Effects of spin-torsion in gauge theory gravity. *J. Math. Phys.*, 39:3303, 1998.

[33] A. Einstein. Zur elektrodynamik bewegter körper. *Annalen der Physik*, 322(10):891–921, 1905.

[34] A. Einstein. Die feldgleichungen der gravitation. *Sitzungsberichte der Preussischen Akademie der Wissenschaften zu Berlin*, page 844847, 1915.

[35] A. Einstein. The Hamiltonian principle and general theory of relativity. *Sitzungsber. preuss. Akad. Wiss.*, 2:1111, 1916.

[36] R. P. Feynman, F. B. Morinigo, and W. G. Wagner. *Feynman Lectures on Gravitation*, pages 74–88. Addison-Wesley, 1995.

[37] M. Fierz and W. Pauli. On relativistic wave equations for particles of arbitrary spin in an electromagnetic field. *Proc. R. Soc. Lond. A*, 173(953):211–232, 1939.

[38] J. Frenkel. Die elektrodynamik des rotierenden elektrons. *Zeitschrift fr Physik A Hadrons and Nuclei*, 37:243–262, 1926.

[39] G. W. Gibbons and S. W. Hawking. Action integrals and partition functions in quantum gravity. *Phys. Rev. D*, 15(10):2752–2756, 1977.

[40] L. P. Grishchuk. Gravity-wave astronomy: some mathematical aspects. In N. Sanchez and A. Zichichi, editors, *Current topics in astrofundamental physics*, page 435. World Scientific, 1992.

[41] S. N. Gupta. Gravitation and electromagnetism. *Phys. Rev.*, 96(6):1683–1685, 1954.

[42] S. Hawking. Gravitational radiation in an expanding universe. *J. Math. Phys.*, 9:598–604, 1968.

[43] F. W. Hehl, P. von der Heyde, G. D. Kerlick, and J. M. Nester. General relativity with spin and torsion: Foundations and prospects. *Rev. Mod. Phys.*, 48(3):393–416, 1976.

[44] D. Hilbert. Die Grundlagen der Physik. *Nachr. Ges. Wiss. Gottingen*, 27:395–407, 1915.

[45] M. P. Hobson, G. P. Efstathiou, and A. N. Lasenby. *General Relativity: An Introduction for Physicists.* Cambridge University Press, 2006.

[46] R. A. Isaacson. Gravitational radiation in the limit of high frequency. I. the linear approximation and geometrical optics. *Phys. Rev.*, 166:1263–1271, 1968.

[47] R. A. Isaacson. Gravitational radiation in the limit of high frequency. II. nonlinear terms and the effective stress tensor. *Phys. Rev.*, 166:1272–1280, 1968.

[48] T. W. B. Kibble. Lorentz invariance and the gravitational field. *J. Math. Phys.*, 2:212, 1961.

[49] A. Komar. Covariant conservation laws in general relativity. *Phys. Rev.*, 113:934–936, 1959.

[50] R. H. Kraichnan. Special-relativistic derivation of generally covariant gravitation theory. *Phys. Rev.*, 98(4):1118–1122, 1955.

[51] K. Kuchar. Dynamics of tensor fields in hyperspace. III. *Journal of Mathematical Physics*, 17(5):801–820, 1976.

[52] L. D. Landau and E. M. Lifshitz. *The Classical Theory of Fields*, chapter 11.96. Pergamon, 1989.

[53] A. Lasenby, C. Doran, and S. Gull. Gravity, gauge theories and geometric algebra. *Phil. Trans. R. Soc. Lond. A*, 356(1737):487–582, 1998.

[54] L. Lehner. Numerical relativity: a review. *Classical and Quantum Gravity*, 18(17):R25, 2001.

[55] G. Magnano and L. M. Sokolowski. Symmetry properties under arbitrary field redefinitions of the metric energy-momentum tensor in classical field theories and gravity. *Class. Quant. Grav.*, 19(2):223–236, 2002.

[56] P. D. Mannheim. Gauge invariant treatment of the energy carried by a gravitational wave. *Phys. Rev. D*, 74(2):024019, 2006.

[57] J. C. Maxwell. A dynamical theory of the electromagnetic field. *Phil. Trans. R. Soc. Lond.*, 155:459–512, 1865.

[58] C. Misner, K. S. Thorne, and J. A. Wheeler. *Gravitation.* W. H. Freeman, 1973.

[59] C. Møller. *The Theory of Relativity*, chapter 11.9. Oxford University Press, 1972.

[60] Y. N.Obukhov and V. A. Korotky. The Weyssenhoff fluid in Einstein-Cartan theory. *Class. Quant. Grav.*, 4(6):1633, 1987.

[61] E. Noether. Invariant variation problems. *Transport Theory and Statistical Physics*, 1:186–207, 1971.

[62] M. Novello and R. P. Neves. The mass of the graviton and the cosmological constant. *Class. Quant. Grav.*, 20(6):L67–L73, 2003.

[63] T. Ortin. *Gravity and Strings*, chapter 3.2. Cambridge University Press, 2004.

[64] T. Padmanabhan. From gravitons to gravity: Myths and reality. *Int. J. Mod. Phys. D*, 17(2/3):367–398, 2008.

[65] A. Papapetrou. Spinning test-particles in general relativity. i. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences*, 209(1097):248–258, 1951.

[66] R. Penrose and W. Rindler. *Spinors and space-time, Vol. 1: two-spinor calculus and relativistic fields*, page 68. Cambridge University Press, 1984.

[67] J. H. Poynting. On the transfer of energy in the electromagnetic field. *Philosophical Transactions of the Royal Society of London*, 175:343–361, 1884.

[68] J. Preskill and K. S. Thorne. *Foreword to Feynman Lectures On Gravitation. Feynman et al.*, pages xxv–xxvi. Westview Press, 2002.

[69] I. Robinson. On the Bel-Robinson tensor. *Class. Quant. Grav.*, 14:A331–A333, 1997.

[70] N. Rosen. *Helv. Phys. Acta. Suppl.*, 4:171, 1956.

[71] L. H. Ryder. *Quantum Field Theory*, chapter 2.7. Cambridge University Press, 2nd edition, 1996.

[72] A. E. Scheidegger. Gravitational motion. *Rev. Mod. Phys.*, 25:451–468, 1953.

[73] R. Schoen and S. Yau. Positivity of the total mass of a general space-time. *Phys. Rev. Lett.*, 43(20):1457–1459, 1979.

[74] D. W. Sciama. The physical structure of general relativity. *Rev. Mod. Phys.*, 36:463–469, 1964.

[75] J. R. Smith and for the LIGO Scientific Collaboration. The path to the enhanced and advanced LIGO gravitational-wave detectors. *Classical and Quantum Gravity*, 26(11):114013, 2009.

[76] J. M. Stewart. Perturbations of Friedmann-Robertson-Walker cosmological models. *Class. Quant. Grav.*, 7(7):1169, 1990.

[77] L. B. Szabados. Quasi-local energy-momentum and angular momentum in general relativity. *Living Reviews in Relativity*, 12(4), 2009.

[78] K. Tod. Penrose's quasi-local mass. In T. Bailey and R. Baston, editors, *Twistors in Mathematics and Physics*, volume 156 of *London Mathematical Society Lecture Note Series*, pages 164–188. Cambridge University Press, Cambridge; New York, 1990.

[79] R. M. Wald. *General Relativity.* University of Chicago Press, 1984.

[80] S. Weinberg. *Gravitation and Cosmology*, pages 91–93. Wiley, 1972.

[81] J. Weyssenhoff and A. Raabe. Relativistic dynamics of spin-fluids and spin-particules. *Acta Phys. Polon.*, 9:7, 1947.

[82] E. Witten. A new proof of the positive energy theorem. *Comm. Math. Phys.*, 80(3):381–402, 1981.

[83] J. W. York. Role of conformal three-geometry in the dynamics of gravitation. *Phys. Rev. Lett.*, 28(16):1082–1085, 1972.