No-go theorem for bimetric gravity with positive and negative mass

Manuel Hohmann[*] and Mattias N. R. Wohlfarth[†]

Zentrum für Mathematische Physik und II. Institut für Theoretische Physik,
Universität Hamburg, Luruper Chaussee 149, 22761 Hamburg, Germany
(Received 24 September 2009; published 10 November 2009)

We argue that the most conservative geometric extension of Einstein gravity describing both positive and negative mass sources and observers is bimetric gravity and contains two copies of standard model matter which interact only gravitationally. Matter fields related to one of the metrics then appear dark from the point of view of an observer defined by the other metric, and so may provide a potential explanation for the dark universe. In this framework we consider the most general form of linearized field equations compatible with physically and mathematically well-motivated assumptions. Using gauge-invariant linear perturbation theory, we prove a no-go theorem ruling out all bimetric gravity theories that, in the Newtonian limit, lead to precisely opposite forces on positive and negative test masses.

I. MOTIVATION

In Newtonian mechanics and gravity, the notion of mass appears as a generic term in quite different physical contexts. Taking a closer look, one needs to distinguish three different types of mass [1]: active gravitational mass m_a is the source of gravitational fields; passive gravitational mass m_p determines the force acting on a test particle in a gravitational field; inertial mass m_i serves as the proportionality factor relating force and acceleration.

Experiment, however, shows that these at first unrelated types of mass are closely tied together. Both ratios m_a/m_p and m_i/m_p appear to be constant independent of material, see e.g. [2–8]. These observations are nicely explained by well-known theoretical principles. Newton's third law asserts that for every force there exists a reciprocal force of equal strength and opposite direction. Considering the gravitational force, this implies that the ratio m_a/m_p between active and passive gravitational mass must be equal for all bodies. The weak equivalence principle states that the acceleration of a physical body in a gravitational field is independent of its composition. This implies that the ratio m_i/m_p between inertial and passive gravitational mass is fixed. It is conventional to choose unit ratios so that all three masses become equal. Taking into account the observational evidence that gravity is always attractive, all mass can be chosen positive.

One may now argue that the experiments and observations mentioned above approve the proportionality and positivity of the different types of mass only for visible matter, i.e., for matter observed through nongravitational interactions, say through emitted light or other types of radiation. However, assuming Einstein gravity, modern astronomical observations [9] suggest that visible matter only constitutes a small fraction of about 5% of the total

matter content of the universe. The main constituents of the universe, known as dark matter and dark energy, have not been observed directly. So it remains unknown whether the same fixed ratio relations between the different types of mass are obeyed also in the dark universe. This lack of knowledge hence invites the interesting possibility to modify Newton's third law, the weak equivalence principle, or the positivity of mass.

In this article we will accept this invitation, and consider gravity theories with a modified weak equivalence principle. We will investigate a different ratio between inertial and passive gravitational mass in the dark sector, namely $m_i/m_p = -1$. Assuming that the inertial mass is still positive, we thus have $m_p = m_a$ and $|m_p| = m_i$ in both sectors. This modification introduces negative gravitational masses in such a way that like masses attract while unlike masses repel each other. Applied to cosmology, such a framework could be particularly rewarding. Effects usually attributed to dark matter and dark energy in principle could be explained by the presence of negative masses in the universe. Indeed, if the intergalactic space contained negative mass distributions, these could push positive matter back towards the respective centers of visible galaxies, mimicking the presence of dark matter sources within the galaxies themselves. Moreover, the repulsive gravitational force acting on the visible positive mass galaxies could also lead to accelerating expansion.

Neither the concept of negative mass nor its cosmological motivation are new in the literature, and have been discussed within several theoretical contexts and with different ratios and signs of m_a, m_p, and m_i. Already in 1897, Föppl [10] introduces negative mass as a natural extension of Newtonian gravity. Consequences within modified Newtonian dynamics are discussed e.g. in [11]. It has been observed that additional negative masses with $m_a = m_p = m_i < 0$ neither violate Newton's third law nor the weak equivalence principle. Forces on such bodies cause these to accelerate in the opposite direction. Probably the

_____
[*]manuel.hohmann@desy.de
[†]mattias.wohlfarth@desy.de

most striking example for this behavior is the gravitational dipole: a system of two bodies of positive and negative mass must forever accelerate in a common direction, the negative mass following the positive one. This effect also exists within Einstein gravity [1,12], where the ratio $m_i/m_p = 1$ is manifestly fixed by the weak equivalence principle. Considering only gravitational forces, such that the acceleration of a body is determined purely by the ratio m_i/m_p, the individual signs of m_p and m_i do not affect the trajectories. Thus, the only possibility to introduce negative mass into Einstein gravity is to choose negative sources for the gravitational field, i.e., $m_a < 0$. Various properties of such negative mass solutions have been discussed; for instance, gravitational lensing [13], gravitational collapse [14], and the stability [15,16] of negative mass black holes. More general repulsive gravitational effects are analyzed e.g. in [17–19].

A consequence of the weak equivalence principle is that all test masses, and all observers, follow the same set of preferred curves, namely, geodesics. In other words, there is only one type of observers. Since we wish to relax this condition by allowing the ratio $m_i/m_p = -1$ for a second type of matter, we have to extend the framework of Einstein gravity. Indeed, it seems natural to introduce a second metric to generate another set of geodesics describing the response of the negative mass observers to the gravitational field. Only then can the gravitational force of a given source be attractive for one class of test particles, while being repulsive for a second class. Such bimetric theories with an antisymmetry between the forces acting on positive and negative masses have become popular under the name "antigravity" [20], but their consistency, in particular their diffeomorphism invariance, has been doubted [21]. Moreover, the findings of this article place further strong constraints on the construction of such theories:

We will prove a theorem which rules out the possibility of gravitational forces of equal strength and opposite direction acting on the two classes of test particles in bimetric gravity theories.

The outline of this article is as follows. In Sec. II, we will state the assumptions that restrict the class of bimetric gravity theories under consideration. Each of these assumptions is motivated and its physical relevance and implications are discussed. At the end of this section we will formulate the no-go theorem, which will be proven in Sec. III. The proof is based on the most general form of linearized gravitational field equations consistent with our assumptions. Using gauge-invariant perturbation techniques, we will decompose the metric perturbations and identify the components relevant in the Newtonian limit. Then we will show by contradiction that it is not possible to obtain a theory with precisely opposite forces acting on positive and negative matter. We will conclude with a discussion in Sec. IV.

II. FORMULATION OF THE THEOREM AND ITS ASSUMPTIONS

We will begin by explaining the theoretical context in which our no-go theorem for antigravity is formulated. In particular, we will give strong motivation for why we consider bimetric theories in order to describe both attractive and repulsive gravitational forces. Then we will discuss in some detail the assumptions entering the no-go theorem which is formulated at the end of this section. The proof of the theorem will be contained in Sec. III.

We wish to consider gravity theories allowing for positive and negative gravitational masses, so that like or unlike masses attract or repel each other, respectively. Observations tell us that the standard model contains only one type of mass, say the positive type, and we will denote the fields of the standard model collectively by Ψ^+. We thus need to introduce a copy Ψ^- of the standard model fields but with negative mass.

Observers follow the curves of massive objects in the small mass limit where backreaction on the geometric background can be neglected. If fields $\Psi^\pm$ of positive and negative gravitational mass are available, this limit will produce two distinguished classes of curves $\gamma^\pm$ on the spacetime manifold. In extension of standard general relativity, it is reasonable to assume that these curves are described by the autoparallels of two linear connections $\nabla^\pm$. In order to make measurements, observers attach frames $\{e_\mu^\pm\}$ with $e_0^\pm = \dot{\gamma}^\pm$ to their spacetime paths $\gamma^\pm$; these frames must be orthonormalized with respect to some metrics $g^\pm$, i.e., $g^\pm(e_\mu^\pm, e_\nu^\pm) = \eta_{\mu\nu}$. In standard fashion, we define observers as inertial, when their frames are nonaccelerated and nonrotating; this is equivalent to Fermi-Walker transport according to $\nabla_{e_0^\pm}^\pm e_\mu^\pm = 0$. From the orthonormality relation, it now follows that the two metrics are covariantly constant, $\nabla^\pm g^\pm = 0$. Assuming vanishing torsion of both connections, we thus find that $\nabla^\pm$ are the Levi-Cività connections related to the metrics $g^\pm$.

Note that observers by this construction follow the timelike geodesics defined by just one of the metrics $g^\pm$. To achieve consistency of this notion of causality with that following from the matter field equations, we are led to assume that the fields Ψ^+ and Ψ^- couple exclusively to the respective metrics g^+ and g^-. We also assume that there is no nongravitational interaction between the two types of matter fields, which is consistent with the lack of direct nongravitational observational evidence for a second type of matter. In other words, matter of type Ψ^- should appear to be dark from the viewpoint of an observer measuring his world with the metric g^+, and vice versa.

This is arguably the most conservative framework for gravity and matter that we can choose to model attractive and repulsive gravitational forces. We simply double the standard ingredients Ψ^+, g^+ in the standard model by introducing additional negative mass fields Ψ^- and nega-

tive mass observers defined with respect to a second metric g^-. This yields a theory with two sectors in both of which gravity, in the absence of sources and observers of the second, nonstandard type, appears exclusively attractive. The interesting possible implications of the existence of both positive and negative masses are discussed in the Introduction, and would only arise from the gravitational interplay of the visible and the dark sector.

We will now formulate, and discuss the physical relevance of, a number of technical assumptions restricting the class of bimetric theories introduced above. These assumptions will be the basic ingredients for our no-go theorem below. For convenience we use in the following underlined quantities to denote two-component vectors (with + and − components), and doubly underlined quantities to denote two by two matrices. Our assumptions are:

(i) The gravitational field equations are a set of two symmetric two-tensor equations of the form

$$\underline{K}_{ab}[g^+, g^-] = \underline{M}_{ab}[g^+, g^-, \Psi^+, \Psi^-]. \quad (1)$$

This assumption is consistent with the naive counting argument by which the number of equations agrees with the number of algebraic components of the two metrics $g^\pm$. More importantly, one can think of these gravitational field equations as arising from a combined diffeomorphism-invariant matter and gravity action of the form $S_G[g^+, g^-] + S_M[g^+, g^-, \Psi^+, \Psi^-]$. Then variation with respect to the metric g^+ would provide the first vector component equation, variation with respect to the metric g^- the second.

(ii) The gravitational tensor $\underline{K}_{ab}$ in the field equations (1) contains at most second derivatives of the metrics g^+ and g^-.

This assumption is one of mathematical simplicity. It makes available a number of theorems on the solvability of partial differential equations, as are also relevant in Einstein gravity. Important for us in the proof of our no-go theorem will be the consequent restriction of the number of terms that can appear in the gravitational field equations.

(iii) The matter source in the field equations (1) is of the form $\underline{M}_{ab} = \underline{J} \cdot \underline{T}_{ab}$, where $\underline{J}$ is a constant invertible matrix, and the components of $\underline{T}_{ab}$ are the respective standard energy momentum tensors $T_{ab}^\pm[g^\pm, \Psi^\pm]$ of positive and negative mass fields.

This seemingly complicated assumption is easily motivated by recalling that matter fields Ψ^+ should only couple to the metric g^+, and fields Ψ^- only to g^-. If such field equations come from an action by variation, then this matter action would take the form $S_M[g^+, g^-, \Psi^+, \Psi^-] = j^+ S[g^+, \Psi^+] + j^- S[g^-, \Psi^-]$ for constant $j^\pm \neq 0$. Variation with respect to g^+ and g^- then produces precisely the assumed matter

source $\underline{M}_{ab}$, the matrix $\underline{J}$ having j^+, j^- on the diagonal.

(iv) The vacuum solution is given by two flat metrics $g_{ab}^\pm = \lambda^\pm \eta_{ab}$ with constants $\lambda^\pm > 0$.

This is another assumption of mathematical simplicity. This vacuum solution has the maximal number of Killing symmetries for both metrics $g^\pm$ simultaneously. The constants $\lambda^\pm$ correspond to the freedom of global rescalings of the Cartesian coordinates. Cosmological constants are excluded; after all, one of the motivations for this framework with both positive and negative mass is the potential explanation of cosmological constants.

(v) Stationary solutions with $\partial_0 g_{ab}^\pm = 0$ exist for arbitrary nonmoving dust matter $T_{00}^\pm = \lambda^\pm \rho^\pm$ with small energy densities $\rho^\pm \sim \mathcal{O}(h)$, so that the Newtonian potentials $\phi^\pm$ are small of the same order $\mathcal{O}(h)$ and the (gauge-fixed) linearly perturbed vacuum metrics are

$$
\begin{aligned}
g^\pm = \lambda^\pm[&-(1 + 2\phi^\pm)dt \otimes dt \\
&+ (1 - 2\gamma^\pm \phi^\pm)\delta_{\alpha\beta}dx^\alpha \otimes dx^\beta]. \quad (2)
\end{aligned}
$$

This simply states that the theory has a (post) Newtonian limit. Dust matter, nonmoving in a given coordinate system, has the energy momentum tensors $T^{\pm ab} = \rho^\pm u^{\pm a} u^{\pm b}$, where $u^\pm \sim \partial_0$. The normalizations $g^\pm(u^\pm, u^\pm) = -1$, corresponding to each type of observer, explain the occurrence of the factors $\lambda^\pm$ in $T_{00}^\pm$. That it should be possible to choose arbitrary small dust distributions $\rho^\pm$ reflects that metric solutions should exist for all appropriate choices of boundary conditions. We shall see below that it is not possible to gauge fix both metrics $g^\pm$ at the same time to have the form displayed above; the reason is that there are just the standard diffeomorphisms of the manifold available, but there is a second metric tensor. However, there exists a gauge-invariant vector of Newtonian potentials $\underline{I}_1$, and gauges can be chosen so that either $I_1^+ = \phi^+$ or $I_1^- = \phi^-$. Note that no specific values for the post-Newtonian parameters $\gamma^\pm$ are assumed, but experiment in our sector of the theory strongly supports the value $\gamma^+ = 1$ [22].

These assumptions suitably restrict the class of bimetric gravity theories in which we wish to analyze the behavior of attractive and repulsive gravitational forces. Using the normalization $8\pi G_N = 1$ for Newton's constant, we are now in the position to formulate our no-go theorem.

Theorem.—We assume a bimetric theory with positive and negative mass sources and observers satisfying points (i)–(v) detailed above. It is not possible to achieve a Newtonian limit with antisymmetric mass mixing in the Poisson equations for the vector $\underline{I}_1$ of gauge-invariant Newtonian potentials,

$$\triangle \underline{I}_1 = \frac{1}{2}\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} \cdot \underline{\rho}. \tag{3}$$

This is a very surprising no-go statement, and we will prove it in Sec. III. Antisymmetric mass mixing is precisely what one would want from a canonical extension of Newton and Einstein gravity. It implies that the Newtonian force on positive test masses m in any gravitational field is precisely opposite to the force felt by negative test masses $-m$ in the same place. Moreover, positive and negative mass sources generate precisely opposite forces on the same test mass. So the theory excluded by the no-go theorem is exactly that which would allow for a switch of sign in gravitational mass.

III. PROOF OF THE NO-GO RESULT

We will now prove by contradiction the no-go theorem for bimetric gravity theories with positive and negative mass as formulated at the end of Sec. II. Since this theorem takes recourse to the Newtonian limit, it is sufficient to use linearized field equations. After discussing the general form of the field equations we will apply the gauge-invariant linear perturbation formalism which is known from cosmological perturbation analysis [23–25]. This important technique enables us to avoid gauge ambiguities that otherwise might invalidate the proof. We will then show that the scalar, vector, and tensor modes of the metric perturbations decouple. Within the sector of scalar perturbations, which is relevant in the (post)Newtonian limit, we will finally construct the contradiction constituting the proof.

A. Field equations

The starting point of our proof is the most general gravitational field equations consistent with the assumptions of the no-go theorem. In agreement with assumptions (i) and (ii) these must be symmetric two-tensor equations containing at most second order derivatives of the metric tensors $g^{\pm}$. We may easily list all tensorial building blocks that may enter the equations according to their derivative order.

(1) no derivatives: the two metrics $g^{\pm}$, and the endomorphism $j = (g^{+})^{-1}g^{-}$;
(2) single derivative: the connection difference S which is a $(1, 2)$-tensor field defined by the decomposition $\nabla_X^- Y = \nabla_X^+ Y + S(Y, X)$, so that $S_{bc}^a = \Gamma_{bc}^{-a} - \Gamma_{bc}^{+a}$;
(3) double derivative: $\nabla^{\pm} S$, and the two Riemann curvature tensors $R^{\pm}$ of the two metrics.

Note that terms of the type $\nabla^{\pm} j$ can be combined from the connection difference S and the metrics; similarly, terms of type $\nabla^{\pm}\nabla^{\pm} j$, $\nabla^{\pm}\nabla^{\mp} j$ can be combined from derivatives of S.

We have already argued that it is sufficient to work with the linearized field equations. Because of assumption (iv), this is a weak field approximation around maximally sym-

metric Minkowski vacua. Then the metric tensors take the form

$$g_{ab}^{\pm} = \lambda^{\pm}(\eta_{ab} + h_{ab}^{\pm}) \tag{4}$$

for constants $\lambda^{\pm}$. In the course of the following calculation we will keep only terms linear in the perturbations $h^{\pm}$, which we assume are of the same order $\mathcal{O}(h)$.

The terms that may now occur in the linearized field equations are the symmetric two-tensors with at most second order derivatives formed from the linearization of the building blocks listed above. Looking at these in more detail, one realizes that terms without derivatives cannot appear; that terms of the type $\partial_a h_{bc}^{\pm}$ always appear quadratic, and so cancel; that all remaining terms are obtained by the various contractions of $\partial_a \partial_b h_{cd}^{\pm}$. With assumption (iii) this leads to the following most general form of the linearized field equations:

$$\begin{aligned}
\underline{K}_{ab} &= \underline{P} \cdot \partial^p \partial_{(a}\underline{h}_{b)p} + \underline{Q} \cdot \Box \underline{h}_{ab} + \underline{R} \cdot \partial_a \partial_b \underline{h} \\
&\quad + \underline{M} \cdot \partial_p \partial_q \underline{h}^{pq} \eta_{ab} + \underline{N} \cdot \Box \underline{h} \eta_{ab} \\
&= \underline{J} \cdot \underline{T}_{ab}. \tag{5}
\end{aligned}$$

Indices are raised with the metric η, and $\Box = \eta^{pq}\partial_p\partial_q$. The matrices $\underline{P}$, $\underline{Q}$, $\underline{R}$, $\underline{M}$, $\underline{N}$ on the geometry side $\underline{K}_{ab}$ of the equations are constant parameters. These are determined by the nonlinear field equations, and also absorb the factors $\lambda^{\pm}$ in the linearization ansatz (4). We neither need to know their precise form nor do we need to make any additional assumptions about these matrices to carry out our proof below.

B. Gauge-invariant formalism

Since our proof is based on linearized field equations we must take care to ensure that none of our conclusions finally depend on changes of gauge, i.e., on possible changes of coordinates that do not alter the structure of the linearization ansatz (4) as a small perturbation of maximally symmetric Minkowski vacua. Therefore we will now apply the formalism of gauge-invariant linear perturbation theory known from cosmology [23–25] to the ansatz (4) and equations (5).

First, we perform a purely algebraic $(1 + 3)$ split of the spacetime coordinates $x^a = (x^0, x^\alpha)$ into time and space, and decompose the corresponding components of the perturbations $\underline{h}_{ab}$ of the metric tensors as

$$\begin{aligned}
\underline{h}_{00} &= -2\underline{\phi}, & \underline{h}_{0\alpha} &= \underline{B}_\alpha, \\
\underline{h}_{\alpha\beta} &= -2\underline{\psi}\delta_{\alpha\beta} + 2\underline{E}_{\alpha\beta},
\end{aligned} \tag{6}$$

where $\underline{E}_{\alpha\beta}$ is trace free, i.e., $\delta^{\alpha\beta}\underline{E}_{\alpha\beta} = 0$. Under purely spatial coordinate transformations, the quantities $\underline{\phi}$ and $\underline{\psi}$ transform as scalars, $\underline{B}_\alpha$ as a vector, and $\underline{E}_{\alpha\beta}$ as a symmetric, trace-free two-tensor. The geometry side $\underline{K}_{ab}$ of the linearized equations (5) now decomposes as

$$\underline{K}_{00} = 2(\underline{P} + \underline{Q} + \underline{R} + \underline{M} + \underline{N}) \cdot \partial_0^2 \underline{\phi} - 2(\underline{Q} + \underline{N}) \cdot \triangle \underline{\phi}$$

$$- 6(\underline{R} + \underline{N}) \cdot \partial_0^2 \underline{\psi} + 2(\underline{M} + 3\underline{N}) \cdot \triangle \underline{\psi}$$

$$+ (\underline{P} + 2\underline{M}) \cdot \partial_0 \partial_\alpha \underline{B}^\alpha - 2\underline{M} \cdot \partial_\alpha \partial_\beta \underline{E}^{\alpha\beta}, \quad (7a)$$

$$\underline{K}_{0\alpha} = (\underline{P} + 2\underline{R}) \cdot \partial_0 \partial_\alpha \underline{\phi} - (\underline{P} + 6\underline{R}) \cdot \partial_0 \partial_\alpha \underline{\psi}$$

$$- (\tfrac{1}{2}\underline{P} + \underline{Q}) \cdot \partial_0^2 \underline{B}_\alpha + \tfrac{1}{2}\underline{P} \cdot \partial_\alpha \partial_\beta \underline{B}^\beta$$

$$+ \underline{Q} \cdot \triangle \underline{B}_\alpha + \underline{P} \cdot \partial_0 \partial^\beta \underline{E}_{\alpha\beta}, \quad (7b)$$

$$\underline{K}_{\alpha\beta} = -2(\underline{M} + \underline{N}) \cdot \partial_0^2 \underline{\phi} \delta_{\alpha\beta} + 2\underline{N} \cdot \triangle \underline{\phi} \delta_{\alpha\beta}$$

$$+ 2\underline{R} \cdot \partial_\alpha \partial_\beta \underline{\phi} + 2(\underline{Q} + 3\underline{N}) \cdot \partial_0^2 \underline{\psi} \delta_{\alpha\beta}$$

$$- 2(\underline{Q} + \underline{M} + 3\underline{N}) \cdot \triangle \underline{\psi} \delta_{\alpha\beta} - 2(\underline{P} + 3\underline{R}) \cdot \partial_\alpha \partial_\beta \underline{\psi}$$

$$- \underline{P} \cdot \partial_0 \partial_{(\alpha} \underline{B}_{\beta)} - 2\underline{M} \cdot \partial_0 \partial_\gamma \underline{B}^\gamma \delta_{\alpha\beta} + 2\underline{P} \cdot \partial^\gamma \partial_{(\alpha} \underline{E}_{\beta)\gamma}$$

$$+ 2\underline{Q} \cdot \Box \underline{E}_{\alpha\beta} + 2\underline{M} \cdot \partial_\gamma \partial_\delta \underline{E}^{\gamma\delta} \delta_{\alpha\beta}, \quad (7c)$$

where spatial indices are raised with the flat spatial metric δ.

In the second step we perform a differential decomposition of the spatial vectors $\underline{B}_\alpha$ and tensors $\underline{E}_{\alpha\beta}$ in (6)

according to

$$\underline{B}_\alpha = \partial_\alpha \tilde{\underline{B}} + \tilde{\underline{B}}_\alpha, \qquad \underline{E}_{\alpha\beta} = \triangle_{\alpha\beta} \tilde{\underline{E}} + 2\partial_{(\alpha} \tilde{\underline{E}}_{\beta)} + \tilde{\underline{E}}_{\alpha\beta}, \quad (8)$$

where $\triangle_{\alpha\beta} = \partial_\alpha \partial_\beta - \tfrac{1}{3} \delta_{\alpha\beta} \triangle$ denotes the trace-free second derivative and

$$\partial^\alpha \tilde{\underline{B}}_\alpha = \partial^\alpha \tilde{\underline{E}}_\alpha = 0, \qquad \partial^\alpha \tilde{\underline{E}}_{\alpha\beta} = 0, \qquad \delta^{\alpha\beta} \tilde{\underline{E}}_{\alpha\beta} = 0. \quad (9)$$

This differential decomposition is unique as has been shown in [25,26]. The essential fact entering the uniqueness argument is that the spatial sections in a Minkowski background are of constant curvature.

In consequence of the algebraic and differential decompositions, the perturbations now are summarized by the so-called scalar modes $\underline{\phi}$, $\underline{\psi}$, $\tilde{\underline{B}}$, $\tilde{\underline{E}}$, by the divergence-free (or transverse) vector modes $\tilde{\underline{B}}_\alpha$, $\tilde{\underline{E}}_\alpha$, and by the transverse trace-free tensor modes $\tilde{\underline{E}}_{\alpha\beta}$. These enter the geometry side of the linearized equations as follows:

$$\underline{K}_{00} = 2(\underline{P} + \underline{Q} + \underline{R} + \underline{M} + \underline{N}) \cdot \partial_0^2 \underline{\phi} - 2(\underline{Q} + \underline{N}) \cdot \triangle \underline{\phi} - 6(\underline{R} + \underline{N}) \cdot \partial_0^2 \underline{\psi} + 2(\underline{M} + 3\underline{N}) \cdot \triangle \underline{\psi} + (\underline{P} + 2\underline{M}) \cdot \partial_0 \triangle \tilde{\underline{B}}$$

$$- \frac{4}{3}\underline{M} \cdot \triangle \triangle \tilde{\underline{E}}, \quad (10a)$$

$$\underline{K}_{0\alpha} = (\underline{P} + 2\underline{R}) \cdot \partial_0 \partial_\alpha \underline{\phi} - (\underline{P} + 6\underline{R}) \cdot \partial_0 \partial_\alpha \underline{\psi} - (\tfrac{1}{2}\underline{P} + \underline{Q}) \cdot \partial_0^2 \partial_\alpha \tilde{\underline{B}} + (\tfrac{1}{2}\underline{P} + \underline{Q}) \cdot \partial_\alpha \triangle \tilde{\underline{B}} + \tfrac{2}{3}\underline{P} \cdot \partial_0 \partial_\alpha \triangle \tilde{\underline{E}}$$

$$- (\tfrac{1}{2}\underline{P} + \underline{Q}) \cdot \partial_0^2 \tilde{\underline{B}}_\alpha + \underline{Q} \cdot \triangle \tilde{\underline{B}}_\alpha + \underline{P} \cdot \partial_0 \triangle \tilde{\underline{E}}_\alpha, \quad (10b)$$

$$\underline{K}_{\alpha\beta} = -2(\underline{M} + \underline{N}) \cdot \partial_0^2 \underline{\phi} \delta_{\alpha\beta} + 2\underline{N} \cdot \triangle \underline{\phi} \delta_{\alpha\beta} + 2\underline{R} \cdot \partial_\alpha \partial_\beta \underline{\phi} + 2(\underline{Q} + 3\underline{N}) \cdot \partial_0^2 \underline{\psi} \delta_{\alpha\beta} - 2(\underline{Q} + \underline{M} + 3\underline{N}) \cdot \triangle \underline{\psi} \delta_{\alpha\beta}$$

$$- 2(\underline{P} + 3\underline{R}) \cdot \partial_\alpha \partial_\beta \underline{\psi} - \underline{P} \cdot \partial_0 \partial_\alpha \partial_\beta \tilde{\underline{B}} - 2\underline{M} \cdot \partial_0 \triangle \tilde{\underline{B}} \delta_{\alpha\beta} + \tfrac{4}{3}\underline{P} \cdot \partial_\alpha \partial_\beta \triangle \tilde{\underline{E}} + 2\underline{Q} \cdot \triangle_{\alpha\beta} \Box \tilde{\underline{E}} + \tfrac{4}{3}\underline{M} \cdot \triangle \triangle \tilde{\underline{E}} \delta_{\alpha\beta}$$

$$- \underline{P} \cdot \partial_0 \partial_{(\alpha} \tilde{\underline{B}}_{\beta)} + 2\underline{P} \cdot \triangle \partial_{(\alpha} \tilde{\underline{E}}_{\beta)} + 4\underline{Q} \cdot \Box \partial_{(\alpha} \tilde{\underline{E}}_{\beta)} + 2\underline{Q} \cdot \Box \tilde{\underline{E}}_{\alpha\beta}. \quad (10c)$$

In Sec. III C we will show that the scalar, vector, and tensor modes in this decomposition completely decouple, i.e., that they lead to equations that can be solved separately. This fact will be important for our proof because it will allow us to set vector and tensor modes to zero. As we will see in the final part of our proof in Sec. III E, the scalar equations then will provide the crucial information about the (post)Newtonian limit needed to prove the theorem.

C. Decoupling of modes

To demonstrate the decoupling of the scalar, vector, and tensor perturbations we consider in turn the 00, 0α and $\alpha\beta$ components of the linearized equations of motion $\underline{K}_{ab} = \underline{J} \cdot \underline{T}_{ab}$, see (5).

A quick inspection of the equation $\underline{K}_{00} = \underline{J} \cdot \underline{T}_{00}$ shows that only scalar modes occur; this becomes obvious from (10a) and by noting that $\underline{T}_{00}$ are scalar modes.

Next consider the equation $\underline{K}_{0\alpha} = \underline{J} \cdot \underline{T}_{0\alpha} = \underline{W}_\alpha$. Clearly the geometry side displayed in (10b) contains scalar and vector modes; schematically we have

$$\underline{K}_{0\alpha} = \partial_\alpha(\text{scalar containing only scalar modes})$$

$$+ (\text{divergence-free vector containing only}$$

$$\text{vector modes})_\alpha. \quad (11)$$

Also the matter side $\underline{W}_\alpha$ can be decomposed into appropriate scalar and vector modes, i.e., into a gradient and a transverse vector as $\underline{W}_\alpha = \partial_\alpha \tilde{\underline{W}} + \tilde{\underline{W}}_\alpha$ with $\partial^\alpha \tilde{\underline{W}}_\alpha = 0$. The uniqueness of these decompositions on both sides now implies that we obtain two separate equations,

$$\tilde{\underline{W}} = (\underline{P} + 2\underline{R}) \cdot \partial_0 \underline{\phi} - (\underline{P} + 6\underline{R}) \cdot \partial_0 \underline{\psi}$$

$$- (\tfrac{1}{2}\underline{P} + \underline{Q}) \cdot \partial_0^2 \tilde{\underline{B}} + (\tfrac{1}{2}\underline{P} + \underline{Q}) \cdot \triangle \tilde{\underline{B}} + \tfrac{2}{3}\underline{P} \cdot \partial_0 \triangle \tilde{\underline{E}}, \quad (12a)$$

$$\tilde{\underline{W}}_\alpha = -(\tfrac{1}{2}\underline{P} + \underline{Q}) \cdot \partial_0^2 \tilde{\underline{B}}_\alpha + \underline{Q} \cdot \triangle \tilde{\underline{B}}_\alpha + \underline{P} \cdot \partial_0 \triangle \tilde{\underline{E}}_\alpha, \quad (12b)$$

the first containing only scalar modes, the second only vector modes.

A very similar argument serves to show that the scalar, vector, and tensor modes in $\underline{K}_{\alpha\beta} = \underline{J} \cdot \underline{T}_{\alpha\beta} = \underline{Z}_{\alpha\beta}$ decouple. Both the geometry side explicitly displayed in (10c) and the matter contribution have to be decomposed as

$$\tfrac{1}{3}\underline{K}\delta_{\alpha\beta} + \triangle_{\alpha\beta}\tilde{\underline{K}} + 2\partial_{(\alpha}\tilde{\underline{K}}_{\beta)} + \tilde{\underline{K}}_{\alpha\beta}$$
$$= \tfrac{1}{3}\underline{Z}\delta_{\alpha\beta} + \triangle_{\alpha\beta}\tilde{\underline{Z}} + 2\partial_{(\alpha}\tilde{\underline{Z}}_{\beta)} + \tilde{\underline{Z}}_{\alpha\beta} \qquad (13)$$

into scalar modes $\underline{K}$, $\underline{Z}$ determining the traces, further

scalar modes $\tilde{\underline{K}}$, $\tilde{\underline{Z}}$, transverse vector modes $\tilde{\underline{K}}_{\alpha}$, $\tilde{\underline{Z}}_{\alpha}$, and transverse trace-free tensor modes $\tilde{\underline{K}}_{\alpha\beta}$, $\tilde{\underline{Z}}_{\alpha\beta}$. The important point to observe is that the respective modes on the curvature side only contain contributions from the same type of mode, e.g., the vector $\tilde{\underline{K}}_{\alpha}$ is fully determined by vector modes. The uniqueness of the decomposition on both sides finally yields four separate equations, each containing only a single type of perturbation modes, namely

$$\underline{Z} = 2(\underline{R} + 3\underline{N}) \cdot \triangle\underline{\phi} - 6(\underline{M} + \underline{N}) \cdot \partial_0^2\underline{\phi} - 2(\underline{P} + 3\underline{Q} + 3\underline{R} + 3\underline{M} + 9\underline{N}) \cdot \triangle\underline{\psi} + 6(\underline{Q} + 3\underline{N}) \cdot \partial_0^2\underline{\psi}$$
$$- (\underline{P} + 6\underline{M}) \cdot \partial_0 \triangle \tilde{\underline{B}} + (\tfrac{4}{3}\underline{P} + 4\underline{M}) \cdot \triangle \triangle \tilde{\underline{E}}, \qquad (14a)$$

$$\tilde{\underline{Z}} = 2\underline{R} \cdot \underline{\phi} - 2(\underline{P} + 3\underline{R}) \cdot \underline{\psi} - \underline{P} \cdot \partial_0\tilde{\underline{B}} + \tfrac{4}{3}\underline{P} \cdot \triangle\tilde{\underline{E}} + 2\underline{Q} \cdot \Box\tilde{\underline{E}}, \qquad (14b)$$

$$\tilde{\underline{Z}}_{\alpha} = -\tfrac{1}{2}\underline{P} \cdot \partial_0\tilde{\underline{B}}_{\alpha} + \underline{P} \cdot \triangle\tilde{\underline{E}}_{\alpha} + 2\underline{Q} \cdot \Box\tilde{\underline{E}}_{\alpha}, \qquad (14c)$$

$$\tilde{\underline{Z}}_{\alpha\beta} = 2\underline{Q} \cdot \Box\tilde{\underline{E}}_{\alpha\beta}. \qquad (14d)$$

The arguments above show that the decoupling of scalar, vector, and tensor perturbations is essentially a consequence of the uniqueness and cleverness of the algebraic and differential decomposition involved. The relevance of the decoupling will become obvious in Sec. III E, since it will allow us to limit our discussion to the scalar perturbations which determine the (post)Newtonian limit. Before we can approach this final part of our proof, however, we need to determine the gauge-invariant quantities containing the physical information contained in the metric perturbations.

D. Gauge invariance and consistency

In this section we will discuss gauge transformations. We will calculate how the scalar, vector, and tensor modes in the metric perturbations change under changes of gauge, and we will find the set of all gauge-invariant quantities. Since gauge transformations are special diffeomorphisms, we must require certain consistency conditions so that the gravitational field equations can be rewritten in terms of gauge-invariant quantities only. Otherwise the solutions of the field equations would not be diffeomorphism invariant.

Gauge transformations in linear perturbation theory are defined as diffeomorphisms that do not change the formal structure of the perturbation ansatz; here this means that the metrics should retain the form of Eq. (4),

$$\underline{g}_{ab} = \underline{\lambda}\eta_{ab} + \mathcal{O}(h). \qquad (15)$$

Every diffeomorphism is generated by a vector field ξ and changes a tensor field by the Lie derivative; hence $\delta_\xi \underline{g} = \mathcal{L}_\xi \underline{g}$. From the corresponding component expression for the Lie derivative, it is clear that the diffeomorphism generated by ξ is a gauge transformation only if $\xi^a \sim \mathcal{O}(h)$. Writing $\xi_a = \eta_{ap}\xi^p$, we then find $\delta_\xi \underline{g}_{ab} =$

$\underline{\lambda}(\partial_a\xi_b + \partial_b\xi_a)$, and so

$$\delta_\xi \underline{h}_{ab} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}(\partial_a\xi_b + \partial_b\xi_a). \qquad (16)$$

We will now compute how the different components of the metric perturbations $\underline{h}_{ab}$ transform under such a gauge transformation. As done above for the metric perturbations, we split the gauge transformation ξ into space and time components, and also employ the differential decomposition. We write

$$\xi_0 = \xi, \qquad \xi_\alpha = \partial_\alpha\tilde{\xi} + \tilde{\xi}_\alpha \qquad (17)$$

for a divergence-free spatial vector mode $\tilde{\xi}_\alpha$. According to (16), the metric perturbations h_{ab}^+ and h_{ab}^- transform in precisely the same way, namely, as would be the case for a single metric theory. Employing our previous definitions of scalar, vector, and tensor modes, we thus obtain for $+$ and $-$ components the same transformation behavior under gauge transformations as known from standard calculations [25],

$$\delta_\xi\underline{\phi} = -\partial_0\xi\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \qquad \delta_\xi\underline{\psi} = -\tfrac{1}{3}\triangle\tilde{\xi}\begin{pmatrix} 1 \\ 1 \end{pmatrix},$$

$$\delta_\xi\tilde{\underline{B}} = (\partial_0\tilde{\xi} + \xi)\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \qquad \delta_\xi\tilde{\underline{E}} = \tilde{\xi}\begin{pmatrix} 1 \\ 1 \end{pmatrix},$$

$$\delta_\xi\tilde{\underline{B}}_\alpha = \partial_0\tilde{\xi}_\alpha\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \qquad \delta_\xi\tilde{\underline{E}}_\alpha = \tfrac{1}{2}\tilde{\xi}_\alpha\begin{pmatrix} 1 \\ 1 \end{pmatrix},$$

$$\delta_\xi\tilde{\underline{E}}_{\alpha\beta} = \underline{0}.$$

$$(18)$$

We are now in the position to deduce gauge-invariant linear combinations of modes. A minimal set of such combinations in terms of which all gauge-invariant quantities can be expressed is

$$\underline{I}_1 = \underline{\phi} + \partial_0 \tilde{B} - \partial_0^2 \tilde{E}, \qquad \underline{I}_2 = \underline{\psi} + \tfrac{1}{3} \triangle \tilde{E},$$

$$I_3 = \tilde{B}^+ - \tilde{B}^-, \qquad I_4 = \tilde{E}^+ - \tilde{E}^-, \qquad (19)$$

$$\underline{L}_\alpha = \underline{\tilde{B}}_\alpha - 2\partial_0 \tilde{E}_\alpha, \qquad I'_\alpha = \tilde{E}^+_\alpha - \tilde{E}^-_\alpha, \qquad \tilde{E}_{\alpha\beta}.$$

Among the gauge-invariant quantities remain six scalars, three vectors, and two tensors. This matches expectations because the gauge transformation, via ξ, $\tilde{\xi}$, $\tilde{\xi}_\alpha$, contains two scalars and one vector which are eliminated from the originally eight scalars, four vectors, and two tensors in the metric perturbations.

The next step is to find the conditions under which the gravitational field equations can be rewritten in terms of the above gauge invariants. As discussed previously this is necessary to ensure diffeomorphism invariance of the solutions. First note that the tensor equations (14d) are already written in terms of gauge invariants. We will now illustrate how to proceed for the two vector equations (12b) and (14c). We replace all occurrences of $\tilde{B}_\alpha$ by $\underline{L}_\alpha + 2\partial_0 \tilde{E}_\alpha$, and then $\tilde{E}_\alpha$ by

$$\frac{1}{2}\begin{pmatrix} 1 \\ -1 \end{pmatrix} I'_\alpha + \frac{1}{2}\begin{pmatrix} 1 \\ 1 \end{pmatrix}(\tilde{E}^+_\alpha + \tilde{E}^-_\alpha). \qquad (20)$$

The two vector equations now read

$$\underline{\tilde{W}}_\alpha = -\left(\frac{1}{2}\underline{\underline{P}} + \underline{\underline{Q}}\right) \cdot \left(\partial_0^2 \underline{L}_\alpha + \begin{pmatrix} 1 \\ -1 \end{pmatrix}\partial_0 \Box I'_\alpha\right) + \underline{\underline{Q}} \cdot \triangle \underline{L}_\alpha$$
$$+ \left(\frac{1}{2}\underline{\underline{P}} + \underline{\underline{Q}}\right) \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix}\partial_0 \Box(\tilde{E}^+_\alpha + \tilde{E}^-_\alpha), \qquad (21a)$$

$$\tilde{\underline{Z}}_\alpha = -\frac{1}{2}\underline{\underline{P}} \cdot \partial_0 \underline{L}_\alpha + \left(\frac{1}{2}\underline{\underline{P}} + \underline{\underline{Q}}\right) \cdot \begin{pmatrix} 1 \\ -1 \end{pmatrix}\Box I'_\alpha$$
$$+ \left(\frac{1}{2}\underline{\underline{P}} + \underline{\underline{Q}}\right) \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix}\Box(\tilde{E}^+_\alpha + \tilde{E}^-_\alpha), \qquad (21b)$$

and are expressed in terms of gauge-invariant quantities provided that

$$(\underline{\underline{P}} + 2\underline{\underline{Q}}) \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \underline{0}. \qquad (22)$$

The procedure of rewriting the scalar equations (10a), (12a), (14a), and (14b) is very similar. One expresses $\underline{\phi}$ and $\underline{\psi}$ in terms of $\underline{I}_1$ and $\underline{I}_2$, and then uses I_3 and I_4 to substitute $\tilde{B}$ and $\tilde{E}$ as in (20). The additional conditions needed so that the scalar equations only contain gauge invariants are

$$(\underline{\underline{P}} + 2\underline{\underline{R}}) \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \underline{0}, \qquad (\underline{\underline{M}} + \underline{\underline{N}}) \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \underline{0}. \qquad (23)$$

Under these conditions the field equations for the scalar modes become

$$\underline{K}_{00} = 2(\underline{\underline{P}} + \underline{\underline{Q}} + \underline{\underline{R}} + \underline{\underline{M}} + \underline{\underline{N}}) \cdot \partial_0^2 \underline{I}_1 - 2(\underline{\underline{Q}} + \underline{\underline{N}}) \cdot \triangle \underline{I}_1 - 6(\underline{\underline{R}} + \underline{\underline{N}}) \cdot \partial_0^2 \underline{I}_2 + 2(\underline{\underline{M}} + 3\underline{\underline{N}}) \cdot \triangle \underline{I}_2$$
$$+ (\underline{\underline{P}} + \underline{\underline{Q}} + \underline{\underline{R}} + \underline{\underline{M}} + \underline{\underline{N}}) \cdot \begin{pmatrix} 1 \\ -1 \end{pmatrix}(-\partial_0^3 I_3 + \partial_0^4 I_4) + \left(\frac{1}{2}\underline{\underline{P}} + \underline{\underline{Q}} + \underline{\underline{M}} + \underline{\underline{N}}\right) \cdot \begin{pmatrix} 1 \\ -1 \end{pmatrix}\partial_0 \triangle I_3$$
$$- (\underline{\underline{Q}} - \underline{\underline{R}}) \cdot \begin{pmatrix} 1 \\ -1 \end{pmatrix}\partial_0^2 \triangle I_4 - (\underline{\underline{M}} + \underline{\underline{N}}) \cdot \begin{pmatrix} 1 \\ -1 \end{pmatrix}\triangle \triangle I_4, \qquad (24a)$$

$$\tilde{\underline{W}} = (\underline{\underline{P}} + 2\underline{\underline{R}}) \cdot \partial_0 \underline{I}_1 - (\underline{\underline{P}} + 6\underline{\underline{R}}) \cdot \partial_0 \underline{I}_2 - \left(\frac{3}{4}\underline{\underline{P}} + \frac{1}{2}\underline{\underline{Q}} + \underline{\underline{R}}\right) \cdot \begin{pmatrix} 1 \\ -1 \end{pmatrix}\partial_0^2 I_3 + \frac{1}{2}\left(\frac{1}{2}\underline{\underline{P}} + \underline{\underline{Q}}\right) \cdot \begin{pmatrix} 1 \\ -1 \end{pmatrix}\triangle I_3$$
$$+ \left(\frac{1}{2}\underline{\underline{P}} + \underline{\underline{R}}\right) \cdot \begin{pmatrix} 1 \\ -1 \end{pmatrix}\partial_0(\partial_0^2 + \triangle)I_4, \qquad (24b)$$

$$\underline{Z} = -6(\underline{\underline{M}} + \underline{\underline{N}}) \cdot \partial_0^2 \underline{I}_1 + 2(\underline{\underline{R}} + 3\underline{\underline{N}}) \cdot \triangle \underline{I}_1 + 6(\underline{\underline{Q}} + 3\underline{\underline{N}}) \cdot \partial_0^2 \underline{I}_2 - (\underline{\underline{Q}} - \underline{\underline{R}}) \cdot \begin{pmatrix} 1 \\ -1 \end{pmatrix}\partial_0^2 \triangle I_4 - 2(\underline{\underline{P}} + 3\underline{\underline{Q}} + 3\underline{\underline{R}}$$
$$+ 3\underline{\underline{M}} + 9\underline{\underline{N}}) \cdot \triangle \underline{I}_2 + 3(\underline{\underline{M}} + \underline{\underline{N}}) \cdot \begin{pmatrix} 1 \\ -1 \end{pmatrix}(\partial_0^3 I_3 - \partial_0^4 I_4) - \left(\frac{1}{2}\underline{\underline{P}} + \underline{\underline{R}} + 3\underline{\underline{M}} + 3\underline{\underline{N}}\right) \cdot \begin{pmatrix} 1 \\ -1 \end{pmatrix}\partial_0 \triangle I_3$$
$$+ (\underline{\underline{P}} + \underline{\underline{Q}} + \underline{\underline{R}} + 3\underline{\underline{M}} + 3\underline{\underline{N}}) \cdot \begin{pmatrix} 1 \\ -1 \end{pmatrix}\triangle \triangle I_4, \qquad (24c)$$

$$\tilde{\underline{Z}} = 2\underline{\underline{R}} \cdot \underline{I}_1 - 2(\underline{\underline{P}} + 3\underline{\underline{R}}) \cdot \underline{I}_2 - \left(\frac{1}{2}\underline{\underline{P}} + \underline{\underline{R}}\right) \cdot \begin{pmatrix} 1 \\ -1 \end{pmatrix}\partial_0 I_3 - (\underline{\underline{Q}} - \underline{\underline{R}}) \cdot \begin{pmatrix} 1 \\ -1 \end{pmatrix}\partial_0^2 I_4 + (\underline{\underline{P}} + \underline{\underline{Q}} + \underline{\underline{R}}) \cdot \begin{pmatrix} 1 \\ -1 \end{pmatrix}\triangle I_4. \quad (24d)$$

Now all equations are manifestly rewritten in terms of gauge invariants only.

We remark that there is a second argument that allows us to understand the matrix conditions (22) and (23) for gauge invariance: the vacuum equations $\underline{K}_{ab} = \underline{0}$ are tensor equations according to assumption (i), and so should not change

under diffeomorphisms, and, in particular, not under gauge transformations. Employing the transformation (16) in the field equations (5) we find the expression

$$\delta_\xi \underline{K}_{ab} = (\underline{P} + 2\underline{R}) \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix} \partial_a \partial_b \partial^p \xi_p + (\underline{P} + 2\underline{Q})$$

$$\cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix} \Box \partial_{(a}\xi_{b)} + 2(\underline{\underline{M}} + \underline{\underline{N}}) \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix} \Box \partial^p \xi_p \eta_{ab}$$

(25)

which should vanish; the necessary conditions for this precisely agree with (22) and (23).

E. Contradiction

We now come to the final part of our proof of the no-go theorem formulated at the end of Sec. II. To proceed, we will now employ the remaining assumption (v) to simplify the equations derived above. It will be sufficient to consider the scalar perturbations described by Eq. (24).

Recall that assumption (v) says that the theory should have a (post)Newtonian limit of stationary solutions with respect to the Killing vector field ∂_0, and so we may drop all terms containing time derivatives from the equations (24). Moreover, the Newtonian limit should hold for arbitrary nonmoving dust sources for which the spatial velocities and internal pressures in the energy momentum tensors vanish; the only nonvanishing components of the energy momentum tensors then are $\underline{T}_{00} = \underline{\lambda} \cdot \underline{\rho}$, for $\underline{\lambda} = \begin{pmatrix} \lambda^+ & 0 \\ 0 & \lambda^- \end{pmatrix}$ and energy densities $\underline{\rho}$. Also, the metric solutions in suitable gauges should be given by (2) for some post-Newtonian parameters $\gamma^\pm$.

The last point implies a very useful relation between gauge invariants, namely $\underline{I}_2 = \underline{\gamma} \cdot \underline{I}_1$ for $\underline{\gamma} = \begin{pmatrix} \gamma^+ & 0 \\ 0 & \gamma^- \end{pmatrix}$. To see why this is true, consider the form of the metric g^+ with scalar perturbations, which is

$$g^+ = \lambda^+[-(1 + 2\phi^+)dt \otimes dt + 2\partial_\alpha \tilde{B}^+ dt \otimes dx^\alpha$$
$$+ ((1 - 2\psi^+)\delta_{\alpha\beta} + 2 \triangle_{\alpha\beta} \tilde{E}^+)dx^\alpha \otimes dx^\beta]. \quad (26)$$

It is clear from the gauge transformations (18) that we can choose ξ and $\tilde{\xi}$ so that both $\tilde{B}^+ = 0$ and $\tilde{E}^+ = 0$; this gauge choice is called the longitudinal gauge. The metric g^+ then can be compared with the assumed post-Newtonian form

$$\lambda^+[-(1 + 2\phi^+)dt \otimes dt + (1 - 2\gamma^+\phi^+)dx_\alpha \otimes dx^\alpha]$$

(27)

relevant for nonmoving dust [22], and linearized in the Newtonian potential ϕ^+. Therefore, in longitudinal gauge, a given value of γ^+ for the observer related to g^+ implies $\psi^+ = \gamma^+ \phi^+$. The corresponding gauge-invariant statement is $I_2^+ = \gamma^+ I_1^+$. Repeating the same argument for the observer related to g^- then also shows that $I_2^- = \gamma^- I_1^-$.

The following argument is essentially unchanged by the value of the post-Newtonian parameters $\gamma^\pm$. In order to simplify the presentation maximally, we choose $\gamma^\pm = 1$ in which case $\underline{I}_2 = \underline{I}_1$. Then the equations (24) for the scalar perturbations simplify under the assumption (v) to

$$\underline{\underline{J}} \cdot \underline{\underline{\lambda}} \cdot \underline{\rho} = -2(\underline{\underline{Q}} - \underline{\underline{M}} - 2\underline{\underline{N}}) \cdot \triangle \underline{I}_1$$
$$- (\underline{\underline{M}} + \underline{\underline{N}}) \cdot \begin{pmatrix} 1 \\ -1 \end{pmatrix} \triangle \triangle I_4, \quad (28a)$$

$$\underline{0} = (\underline{P} + 2\underline{\underline{Q}}) \cdot \begin{pmatrix} 1 \\ -1 \end{pmatrix} \triangle I_3, \quad (28b)$$

$$\underline{0} = -2(\underline{P} + 3\underline{\underline{Q}} + 2\underline{R} + 3\underline{\underline{M}} + 6\underline{\underline{N}}) \cdot \triangle \underline{I}_1$$
$$+ (\underline{P} + \underline{\underline{Q}} + \underline{R} + 3\underline{\underline{M}} + 3\underline{\underline{N}}) \cdot \begin{pmatrix} 1 \\ -1 \end{pmatrix} \triangle \triangle I_4, \quad (28c)$$

$$\underline{0} = -2(\underline{P} + 2\underline{R}) \cdot \underline{I}_1 + (\underline{P} + \underline{\underline{Q}} + \underline{R}) \cdot \begin{pmatrix} 1 \\ -1 \end{pmatrix} \triangle I_4. \quad (28d)$$

We now eliminate the term containing $\triangle \triangle I_4$ in the third equation by substituting an appropriate combination of the first and last equations. This yields the simple result

$$-4\underline{\underline{Q}} \cdot \triangle \underline{I}_1 = \underline{\underline{J}} \cdot \underline{\underline{\lambda}} \cdot \underline{\rho}, \quad (29)$$

from which our contradiction will now follow. We must consider two possible cases:

(1) $\underline{\underline{Q}}$ is not invertible. In this case, the dimension of the image of $\underline{\underline{Q}}$, viewed as an endomorphism of $\mathbb{R}^2$, is less than two. From assumption (v) we know that $\underline{\rho}$ can be chosen arbitrarily; since $\underline{\underline{J}}$ and $\underline{\underline{\lambda}}$ by assumptions (iii) and (iv) are invertible, it follows that $\underline{\underline{J}} \cdot \underline{\underline{\lambda}} \cdot \underline{\rho}$ spans $\mathbb{R}^2$. This is a contradiction.

(2) $\underline{\underline{Q}}$ is invertible. In this case, we obtain the equation

$$\triangle \underline{I}_1 = -\frac{1}{4}\underline{\underline{Q}}^{-1} \cdot \underline{\underline{J}} \cdot \underline{\underline{\lambda}} \cdot \underline{\rho}, \quad (30)$$

which is the Poisson equation for the two Newtonian potentials $\underline{I}_1$ of the two different observers related to the metrics $\underline{g}$. Antisymmetric mass mixing, as defined in the no-go theorem, occurs if and only if

$$-\frac{1}{2}\underline{\underline{Q}}^{-1} \cdot \underline{\underline{J}} \cdot \underline{\underline{\lambda}} = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}. \quad (31)$$

Since the left-hand side of the equation is invertible while the right hand side is not, this immediately leads to the desired contradiction.

This concludes the proof of the Theorem of Sec. II that the construction of bimetric theories with antisymmetric mass mixing is not possible. ∎

IV. CONCLUSION

In this article, we have argued that the most conservative framework for a theory of antigravity containing both positive and negative gravitational mass sources and observers is a bimetric theory with two copies of standard model matter. The second metric generates another set of geodesics describing the response of negative mass observers to the gravitational field. The two matter sectors only interact gravitationally so that the negative mass sector appears dark from the point of view of a positive mass observer. Such theories could offer a potential explanation for the unaccounted dark part of the universe.

As our central result we have proven a no-go theorem ruling out all canonical bimetric extensions of Einstein gravity. In these excluded theories the gravitational forces acting on the two different types of matter are precisely opposite. Moreover, opposite matter sources yield opposite forces on a given test mass. Our proof proceeds from the most general form of the linearized field equations for a bimetric theory. Using gauge-invariant linear perturbation theory, we have computed the (post)Newtonian limit. We have obtained the Poisson equations for the Newtonian potentials, provided that the sources of the gravitational fields are constituted by nonmoving dust matter. It turns out that each copy of standard model matter contributes to both Newtonian potentials, but with an important restriction: the mixing matrix, which governs the contribution of each matter source to each gravitational field, is manifestly invertible. However, opposite forces on the different types of matter require a noninvertible mixing, and so are excluded.

Of course the strength of any no-go theorem highly depends on its assumptions. Our assumptions as stated in Sec. II are strongly motivated on physical and mathematical grounds, and apply to a very large class of bimetric theories. Hence, our no-go theorem is highly restrictive and puts very severe constraints on the construction of bimetric antigravity . Now that we have a clear picture of what is not possible, we might wonder whether antigravity theories exist at all. We will now discuss some possibilities using a smaller, or different, set of assumptions so that our no-go theorem does not apply.

One simple way of avoiding the conclusion of our no-go theorem is to allow for different strengths of the gravitational forces acting on positive and negative test masses in the same gravitational field. In the Newtonian limit, the Poisson equation would then read

$$\triangle \underline{L}_1 = \frac{1}{2} \begin{pmatrix} 1 & -\alpha \\ -\alpha & 1 \end{pmatrix} \cdot \underline{\rho} \qquad (32)$$

for $\alpha \neq \pm 1$. Now the mixing matrix that determines the contribution of the matter sources to the Newtonian potentials is invertible. So the only cases excluded by the proof of our no-go theorem are $\alpha = 1$ which corresponds to exactly opposite forces, and $\alpha = -1$ which means equal force on all observers (this is the situation modeled by Einstein gravity with a single metric). One may argue, however, that the introduction of an additional parameter α does not present a canonical extension of Einstein gravity.

A second possibility is to relax the assumption that the sources of the gravitational field originate from the standard action for matter fields. Instead one might use different actions containing both metric tensors, which would change the matter side of the equations. Of course, this would also change the equations of motion for matter fields so that all types of matter would be influenced by both metrics. This is problematic because it would change the causality of field propagation. However, one might argue that our observations of matter in gravitational fields are limited to particular settings, e.g. to the solar system. The theory might be constructed so that the changes in causality there might be weak or even cancel completely. In other words, our assumption (iii), which restricts the matter side of the field equations, could be valid within the bounds of current observations, but may not hold in general.

Third, we may consider a less conservative framework containing more than two metric tensors and a correspondingly higher number of standard model copies. The computation we have performed in our proof of the no-go theorem can be generalized to this case. It turns out that the Poisson equation in the Newtonian limit is formally the same as Eq. (30). As in the bimetric case, we could now demand that like masses attract while unlike masses repel each other with equal strength. This corresponds to the requirement that the Poisson equation should be

$$\triangle \underline{L}_1 = \frac{1}{2} \begin{pmatrix} 1 & -1 & \cdots & -1 \\ -1 & 1 & & -1 \\ \vdots & & \ddots & \\ -1 & -1 & & 1 \end{pmatrix} \cdot \underline{\rho}. \qquad (33)$$

If n metrics and n copies of the standard model are used, the mixing matrix has nonvanishing determinant $(2 - n) \cdot 2^{n-1}$ for $n \neq 2$. Hence it is invertible, and the conclusion of the no-go theorem only applies to the bimetric case $n = 2$.

Finally we wish to mention the recent comment [27] on this article, where it is argued that the antigravity theory of [28] also avoids our no-go theorem. This comes from the simple fact that this theory contains two additional (1, 1)-tensor fields besides our two metrics and two copies of the standard model.

The above discussion demonstrates that several ways of circumventing our no-go theorem might exist. Further research now must show whether one of these can be realized, i.e., whether it is actually possible to construct a concrete antigravity theory with attractive and repulsive gravitational forces. Once such a theory is available, its physical implications will have to be investigated. It would be particularly interesting to study the predictions of the

theory for cosmology, and whether the extra copies of the standard model may serve as an explanation for the dark universe.

ACKNOWLEDGMENTS

[1] H. Bondi, Rev. Mod. Phys. **29**, 423 (1957).
[2] L. B. Kreuzer, Phys. Rev. **169**, 1007 (1968).
[3] D. F. Bartlett and D. Van Buren, Phys. Rev. Lett. **57**, 21 (1986).
[4] C. M. Will, Astrophys. J. Lett. **393**, L59 (1992).
[5] R. V. Eötvös, D. Pekár, and E. Fekete, Ann. Phys. (Leipzig) **373**, 11 (1922).
[6] P. G. Roll, R. Krotkov, and R. H. Dicke, Ann. Phys. (N.Y.) **26**, 442 (1964).
[7] K. Nordtvedt, Phys. Rev. **169**, 1014 (1968).
[8] J. G. Williams, S. G. Turyshev, and D. H. Boggs, Int. J. Mod. Phys. D **18**, 1129 (2009).
[9] E. Komatsu *et al.* (WMAP Collaboration), Astrophys. J. Suppl. Ser. **180**, 330 (2009).
[10] A. Föppl, Sitzungsber. Kgl. Bayr. Akad. Wiss. **6**, 93 (1897).
[11] M. Milgrom, Astrophys. J. **306**, 9 (1986).
[12] G. W. Gibbons, Commun. Math. Phys. **35**, 13 (1974).
[13] D. F. Torres, G. E. Romero, and L. A. Anchordoqui, Mod. Phys. Lett. A **13**, 1575 (1998).
[14] R. B. Mann, Classical Quantum Gravity **14**, 2927 (1997).
[15] G. W. Gibbons, S. A. Hartnoll, and A. Ishibashi, Prog. Theor. Phys. **113**, 963 (2005).
[16] R. J. Gleiser and G. Dotti, Classical Quantum Gravity **23**, 5063 (2006).
[17] D. Harari and C. Lousto, Phys. Rev. D **42**, 2626 (1990).
[18] P. D. Mannheim, Found. Phys. **30**, 709 (2000).
[19] G. Preti and F. de Felice, Am. J. Phys. **76**, 671 (2008).
[20] S. Hossenfelder, Phys. Lett. B **636**, 119 (2006); S. Hossenfelder, arXiv:gr-qc/0605083.
[21] J. Noldus and P. Van Esch, Phys. Lett. B **639**, 667 (2006).
[22] C. M. Will, Living Rev. Relativity **9** (2006), http://www.livingreviews.org/lrr-2006-3.
[23] J. M. Bardeen, Phys. Rev. D **22**, 1882 (1980).
[24] K. A. Malik and D. Wands, Phys. Rep. **475**, 1 (2009).
[25] J. M. Stewart, Classical Quantum Gravity **7**, 1169 (1990).
[26] J. W. York, Ann. Inst. H. Poincaré (A) **21**, 319 (1974).
[27] S. Hossenfelder, arXiv:0909.2094.
[28] S. Hossenfelder, Phys. Rev. D **78**, 044015 (2008).