. SCUOLA INTERNAZIONALE SUPERIORE DI STUDI AVANZATI

SISSA Digital Library

Dynamical diffeomorphisms

*Original*
Dynamical diffeomorphisms / Ferrero, R., Percacci, R.. - In: CLASSICAL AND QUANTUM GRAVITY. - ISSN 0264-9381. - 38:11(2021), pp. 1-25. [10.1088/1361-6382/abf627]

*Availability:*
This version is available at: 20.500.11767/128772 since: 2022-05-14T15:13:35Z

*Publisher:*

*Published*
DOI:10.1088/1361-6382/abf627

*Terms of use:*

Testo definito dall'ateneo relativo alle clausole di concessione d'uso

*Publisher copyright*
IOP- Institute of Physics

This version is available for education and non-commercial purposes.

note finali coverpage

(Article begins on next page)

11 September 2026

# Dynamical diffeomorphisms

Renata Ferrero[1] and Roberto Percacci[2,3]

[1]*Institut für Physik (THEP), Johannes Gutenberg - Universität Mainz, Staudingerweg 7, 55128 Mainz, Germany*
[2]*International School for Advanced Studies, via Bonomea 265, 34136 Trieste, Italy*
[3]*INFN, Sezione di Trieste, Italy*

**Abstract**

We construct a general effective dynamics for diffeomorphisms of spacetime, in a fixed external metric. Though related to familiar models of scalar fields as coordinates, our models have subtly different properties, both at kinematical and dynamical level. The energy-momentum tensor consists of two independently conserved parts. The background solution is the identity diffeomorphism and the energy-momentum tensor of this solution gives rise to an effective cosmological constant.

## 1 Introduction

In General Relativity, spacetime diffeomorphisms play the role of active gauge transformations, while coordinate transformations are usually viewed as passive gauge transformations. In this paper we shall discuss instead a possible role of diffeomorphisms as dynamical variables, a point of view that is closer to their applications in hydrodynamics and elasticity. Originally, we had in mind two main potential applications: the problem of dark energy in cosmology and the problem of observables in General Relativity. In this introduction we shall review these motivations. We shall discuss in the conclusions the extent to which dynamical diffeomorphisms can be of use in these contexts.

Modelling dark energy is important both for early and late cosmology. A term in the action proportional to the volume of spacetime is the simplest explanation, but it has drawbacks. For this reason a more dynamical origin is often preferred. See [1] for a review of many possible alternatives. The most popular models are based on the potential of dynamical scalars. There are also many models of scalar fields with derivative interactions. The models we will discuss in this paper can be seen as a special subclass of the latter, where the fields are restricted kinematically to be diffeomorphisms of spacetime. The idea of using the dynamics of coordinates or diffeomorphisms to generate an effective cosmological constant goes back at least to [2–4], where it was used to induce spontaneous compactification of certain directions in higher-dimensional theories. More recently, it has been used extensively in the literature on massive gravity [5]. A mass term for the graviton breaks diffeomorphism symmetry, but the theory can still be written in a diffeomorphism-invariant way by introducing four "Stückelberg" fields, much in the same way as massive QED can be written in a gauge-invariant way by introducing one real scalar field [6]. In an influential paper, Arkani-Hamed, Georgi and Schwartz [7] have constructed the effective field theory of the four Goldstone bosons that are used to restore diffeomorphism invariance in massive gravity. The theory is formulated in terms of two separate "sites", that can be viewed as two copies of spacetime, each endowed with a separate diffeomorphism invariance,

[1]e-mail address: rferrero@uni-mainz.de
[2]e-mail address: percacci@sissa.it

and a field linking the two sites. As with all Goldstone bosons, the Lagrangian of the fluctuation of the linking field is shift-invariant and therefore contains only derivative couplings.

The major issue with this idea is that one of the four scalars (namely the one associated to the time coordinate) is a ghost. [1] This is because the target space of the scalar fields has a Minkowskian metric. In the original model, the ghost starts propagating at energy scales higher than $(1000\text{km})^{-1}$ [8, 9], leading to strong conflict with observations. This problem can be circumvented in two different ways. The first goes back to the fully non-linear version of the massive Fierz-Pauli theory, constructed as a bi-metric theory with a specific potential [10–15]. This theory can be made diffeomorphism-invariant by adding suitable scalar fields and is ghost-free. *In vacuo* it is reliable up to energies of order $\sim \text{mm}^{-1}$; above this scale, it becomes invalid due to strong coupling effects. In the presence of sources the situation is much worse.

The second is to construct a ghost-free theory reducing diffeomorphism invariance to foliation-preserving diffeomorphisms [16–19]. This is perhaps not too high a price to be paid, since in the cosmological context a preferred foliation is singled out anyway. [2] The resulting models can be interpreted as describing the dynamics of a medium filling the Universe. See [21–27] for later developments and applications to accelerating cosmological models.

Restricting ourselves to the covariant models, the main building block of the action is the matrix

$$B^{\mu}{}_{\nu} = g^{\mu\rho}(\varphi^* h)_{\rho\nu} = g^{\mu\rho}\partial_\rho\varphi^\alpha\partial_\nu\varphi^\beta h_{\alpha\beta}(\varphi) \ , \tag{1.1}$$

where $\varphi^\alpha$ are the scalar fields, $g_{\mu\nu}$ is the spacetime metric, $h_{\alpha\beta}(\varphi)$ is the target space metric (usually assumed flat) and $\varphi^* h$ is the pullback of $h$ by $\varphi$. A necessary condition for $\varphi$ to be a diffeomorphism is that $\partial_\mu\varphi^\alpha$ be nondegenerate everywhere. The transformations of $B^{\mu}{}_{\nu}$ are the same as one would have in an ordinary nonlinear sigma model in curved space, and diffeomorphism-invariant actions can be constructed by taking functions of traces of this matrix. [3] For example, the simplest action is the action for harmonic maps

$$S = -\frac{1}{2}f^2\int d^4x\sqrt{g}\,\text{tr}B \ , \tag{1.2}$$

where $f$ is a constant with dimension of mass. In the non-covariant models with only foliation-preserving diffeomorphism invariance, the indices $\alpha$, $\beta$ run only from 1 to $d-1$ and are associated with the space coordinates.

In these models it is generally assumed that the topology of spacetime is $\mathbb{R}^4$. This is in line with the standard mathematical definition of coordinates as maps from a manifold to $\mathbb{R}^4$. How could the models be extended to topologically nontrivial spacetimes, for example cosmology with compact spatial sections? This seems hard as long as one sticks to the notion of coordinates as dynamical fields, because coordinates are local constructions and trying to extend a coordinate system on, e.g., a sphere, will necessarily lead to singularities. It is then better to jump from the passive view of gauge transformations to the active one, namely diffeomorphisms. Unlike coordinate transformations, diffeomorphisms have a global definition and one may expect that it is possible to discuss the dynamics of diffeomorphisms globally. This is what we do in the present paper.

The other general motivation for our study was the absence of local observables in (quantum) gravity [28–30]. This is in apparent blatant contradiction with the fact that we routinely perform

[1] In the models of [2–4], the problem of ghosts did not arise because the compactified dimensions are spacelike.

[2] We recall that the same symmetry reduction is also present in Hořava-Lifshitz gravity [20], and is sufficient to reconcile ghost freedom with perturbative renormalizability.

[3] Obviously, diffeomorphism invariance requires that the metric be transformed.

measurements of local fields: for example, one can measure the components of the Riemann tensor near the surface of the earth. Upon reflection, one immediately understands that such measurements are made possible by the existence of matter. For example, the Ricci scalar $R$ at a point $x$ is not diffeomorphism invariant: under a diffeomorphism $\psi$, the scalar is transformed by pullback $R \mapsto \psi^* R$, so $R(x) \mapsto \psi^* R(x) = R(\psi(x))$. So, the Ricci scalar is not an observable. However, if $X$ denotes the (spacetime) position of a particle, a diffeomorphism will map $X \mapsto \psi^{-1}(X)$. Thus $R(X)$, the Ricci scalar at the position of the particle, is diffeomorphism invariant, and hence observable. With a bunch of infinitesimally close particles one can construct a local frame at $X$ and give physical meaning to the components of any tensor at $X$. A sufficiently dense dust of such particles, which in the limit becomes a continuous fluid, can be used to set up a physical coordinate system on some patch of spacetime, and hence give physical meaning to continuous tensor fields. This is standard procedure in cosmology (see for example [32–34] or [35] for a review).

All measurements of tensor or scalar fields, referring to such "physical" coordinate systems, can be thought of as measurements of relational observables. This is abundantly discussed in the literature [28, 29, 36–43]. The matter content in the Universe represents a reference system or reference medium with respect to which the points of the target space are defined. The question arises: what is the dynamics of these fields? In many applications the "coordinate infrastructure" is sufficiently rarefied that its effect on the geometry can be safely ignored. In general, however, this is not the case and the backreaction of the "coordinate infrastructure" on the geometry must be accounted for, as for example in cosmology. In particular, in the classical theory one can assume that the coordinate fluid is as "thin" as one wants, but not in the quantum theory. Giddings et al. discuss relational approaches to locality based on diffeomorphism-invariant nonlocal operators [44, 45]. In [46] they have discussed absolute limits. A simplified version of their argument would go as follows: in order to be able to distinguish points in spacetime with a resolution $\ell$, one needs to fill space with radiation of wavelength $\lambda < \ell$, and there must be at least one quantum per volume $\ell^3$. This absolutely minimal coordinate infrastructure would have an energy density $\ell^{-4}$.

One can draw from these remarks two general conclusions. The first is that in order to meaningfully talk about local observables in the theory of gravity one has to include a form of matter as a coordinate system, and that this will always back-react onto the geometry. Pure gravity is an unphysical abstraction. This is actually one way of formulating Mach's Principle; if one takes away all matter, there is no space [31]. The second is that at least some of the conceptual issues that arise in (quantum) gravity, are in fact the result of this unphysical abstraction, namely neglecting the existence of matter. This is another motivation for models of gravity endowed with physical coordinates. In the hamiltonian framework, this is called "deparametrization". The most popular models of this type contain some form of dust [32] or phantom energy [47]. There is a close analogue of this in gauge theories, where gauge-invariant observables can be constructed by suitably dressing local operators [48, 49]. In the case of gravity this has been discussed to some extent in [50]. Whether our models can be useful in the construction of relational observables of this type is a question that we shall briefly return to in the conclusion.

After these motivations, we remark that treating diffeomorphisms as dynamical variables gives rise to a theory with rather different features from any other. Indeed, ordinary matter propagating in spacetime can be represented either by maps into spacetime, such as the worldlines of point particles or the worldsheets of strings, or maps from spacetime into some other "internal" space, as is the case with all the usual matter fields. Here we are in a peculiar situation where the "matter" field is both a map *on* spacetime and *in* spacetime. This gives rise to significant differences compared to the dynamics of ordinary matter fields, and also to the dynamics

of coordinates, regarded as scalar fields on spacetime. At the level of kinematics, we replace the target space $\mathbb{R}^4$ by spacetime itself. [4] As we shall see, quite aside from the topological issues, this subtly changes the invariance group. In writing the action, the main novelty of our treatment is that in (1.1) we put $h(\varphi) = g(\varphi)$. This is a rather significant difference, because there is a new dependence on the metric that affects the definition of the energy-momentum tensor of the scalars, and has further consequences on the relation between the equations of motion and energy-momentum conservation. In particular, we shall see that the energy-momentum tensor consists of two pieces that are separately conserved.

This paper is organized as follows. In Section 2 we write a class of actions for diffeomorphisms of a manifold to itself. As discussed above, these differ from the models that have been considered previously because the metrics in the domain and in the target space are one and the same (i.e., $g = h$). We derive the equations of motion (EOMs) and show that the energy-momentum (EM) tensor contains new terms that had not been considered previously, to the best of our knowledge. Interestingly, these new terms are conserved independently of the rest. We show that the equation of motion of the scalar is generically equivalent to the conservation of the EM tensor, but not always. In Section 3 we discuss the identity solution as a model for dark energy and its stability. In Section 4 we return to the motivations given above and discuss the extent to which the models may provide satisfactory answers. Appendix A contains a detailed proof of the diffeomorphism invariance of the action. In Appendix B we show that models where the domain and target space are not identified (and in particular $h \neq g$) have different properties from those discussed in the main text. However, the latter can be obtained from the former under some additional conditions.

## 2 The models

The models we shall discuss are very similar to nonlinear sigma models, except that the domain and the target space are the same manifold $M$. In spite of this identification, we still need to distinguish two types of geometric objects: tensors evaluated at a point $x$ and tensors evaluated at $\varphi(x)$. In order to better keep track of the difference, in component formulas, we shall use letters from the middle greek alphabet for the former and letters from the beginning of the greek alphabet for the latter. The pullback transforms tensors at $\varphi(x)$ to tensors at $x$, for example given a covariant vector $\omega$ (a one-form), its pullback is

$$(\varphi^*\omega)_\mu = \partial_\mu \varphi^\alpha \omega_\alpha(\varphi) \tag{2.1}$$

and

$$(\varphi^* g)_{\mu\nu} = \partial_\mu \varphi^\alpha \partial_\nu \varphi^\beta g_{\alpha\beta}(\varphi) \tag{2.2}$$

is the pullback of the metric. We will always use this notation: a tensor like $\omega$ without argument is understood to be evaluated at some point $x$, whereas $\omega(\varphi)$ has to be understood as a tensor evaluated at $\varphi(x)$.

In this setting the use of covariant derivatives requires a little explanation. The covariant derivative of tensors at $x$ will be called $\nabla$. It involves only Christoffel symbols with indices $\mu$, $\nu$, $\rho$ etc. Tensors evaluated at $\varphi(x)$ have to be treated as scalars (they are inert under changes of frame at $x$). Thus for example $\nabla_\nu g_{\mu\nu} = 0$ as usual, but

$$\nabla_\nu g_{\beta\alpha}(\varphi) = \partial_\nu \varphi^\gamma \partial_\gamma g_{\beta\alpha}(\varphi) \ . \tag{2.3}$$

[4] There is an analog of this in the Stückelberg treatment of massive QED. If one takes topology properly into account, the Stückelberg scalar is not an ordinary real-valued scalar, but has to be identified modulo $2\pi$.

There is also a notion of covariant derivative on tensors at $\varphi$. We will write $D$ for this type of covariant derivative. For example

$$D_\mu \omega_\alpha(\varphi) = \partial_\mu \varphi^\gamma D_\gamma \omega_\alpha \ ; \quad D_\gamma \omega_\alpha = \partial_\gamma \omega_\alpha - \Gamma_\gamma{}^\beta{}_\alpha \omega_\beta(\varphi) \ . \tag{2.4}$$

Note that in spite of carrying an index, $\varphi^\alpha$ are not vectors. Their covariant derivative is the same as the ordinary partial derivative. Thus we will use interchangeably the notation $\partial_\mu \varphi^\alpha$ and $\nabla_\mu \varphi^\alpha$, but when the derivative index is raised we always use $\nabla$, so $g^{\rho\mu}\partial_\mu \varphi^\alpha = \nabla^\rho \varphi^\alpha$.

We shall also encounter objects that have indices of both types. For example, $\nabla_\mu \varphi^\alpha$ is a covariant vector at $x$ and a contravariant vector at $\varphi(x)$. [5] In this case we find it convenient to use a covariant derivative $\nabla$ that covariantizes only the index $\mu$, while $\mathcal{D}$ is the full covariant derivative that covariantizes both indices:

$$\nabla_\rho \nabla_\mu \varphi^\alpha = \partial_\rho \nabla_\mu \varphi^\alpha - \Gamma_\rho{}^\sigma{}_\mu \nabla_\sigma \varphi^\alpha \ . \tag{2.5}$$

$$\mathcal{D}_\rho \nabla_\mu \varphi^\alpha = \nabla_\rho \nabla_\mu \varphi^\alpha + \partial_\rho \varphi^\gamma \Gamma_\gamma{}^\alpha{}_\beta \nabla_\mu \varphi^\beta \ , \tag{2.6}$$

Let $g_{\mu\nu}$ be a metric on $M$. The basic dynamical variable $\varphi$ is a diffeomorphism of $M$ to itself. In analogy to (1.1), the basic building block of the action is the tensor

$$B^\mu{}_\nu = g^{\mu\rho}(\varphi^* g)_{\rho\nu} = g^{\mu\rho}\partial_\rho \varphi^\beta g_{\beta\alpha}(\varphi)\partial_\nu \varphi^\alpha \ . \tag{2.7}$$

In the r.h.s. we have written the factors in a particular order that calls for the use of matrix notation. If we denote $J_\mu{}^\alpha = \partial_\mu \varphi^\alpha$ the Jacobian of $\varphi$, we can write (2.7) in the form

$$B = g^{-1}\,(\varphi^* g) = g^{-1} J g(\varphi) J^T \ .$$

This notation is often convenient in calculations.

Let us now construct scalars out of this tensor. Since $B$ is assumed to transform as a mixed tensor, the invariants are traces of products of powers of $B$. Only four such terms are algebraically independent. An arbitrary scalar can be expressed, generally in a nonlinear way, via four chosen independent scalars. The natural choice of four independent scalars is

$$\tau_k = \mathrm{Tr}(B^k) \ , \qquad k = 1, 2, 3, 4 \ . \tag{2.8}$$

Our action is then

$$S = \int dx \, \sqrt{g}\, \mathcal{L}(\tau_1, \tau_2, \tau_3, \tau_4) \ , \tag{2.9}$$

where $\mathcal{L}$ is the Lagrangian, an arbitrary function of the traces. In the following we will sometimes consider the simplest case

$$\mathcal{L} = -\frac{1}{2} f^2 \tau_1 \tag{2.10}$$

where $f$ is a coupling. In mathematical literature (where $g$ is a Riemannian metric) the stationary points are the harmonic maps [51]. Another particularly interesting Lagrangian is

$$\mathcal{L} = c(\tau_1^2 - \tau_2) \ . \tag{2.11}$$

We shall discuss some of its properties in Sect. 3.2.

Let us now discuss the invariances of the theory. As long as $\varphi$ are the only dynamical fields, and $g$ is a fixed metric, the only symmetries of the action are the isometries of $g$. Thus,

[5] It is a section of $T^*M \otimes \varphi^* TM$. The connection in this bundle is the tensor product of the Levi-Civita connection in $T^*M$ and the pullback of the Levi-Civita connection in $TM$.

generically, there are no symmetries. Even though we will not discuss dynamical gravity in this paper, we shall consider the invariances of the action when the metric is allowed to transform. This becomes relevant when one couples the fields to a dynamical metric, and will also be important later in the discussion of the energy-momentum tensor and its conservation.

A diffeomorphism $\psi$ acts on tensors on $M$ in the standard way, in particular the action on covariant tensors is by pullback. For example, the metric $g$ is transformed to $g' = \psi^* g$. The action on $\varphi$ can be either by left or by right composition. Under right composition

$$\varphi \mapsto \varphi' = \varphi \circ \psi \tag{2.12}$$

we have

$$\varphi^* g \mapsto \varphi'^* g' = \psi^* \varphi^* \psi^* g \ , \tag{2.13}$$

which is not the correct transformation of a covariant tensor. Thus $B$ does not transform properly as a mixed tensor. Similarly under left composition

$$\varphi \mapsto \varphi' = \psi^{-1} \circ \varphi \ , \tag{2.14}$$

we find that the pullback of $g$ is invariant:

$$\varphi^* g \mapsto \varphi'^* g' = \varphi^* \psi^{-1*} \psi^* g = \varphi^* g \ . \tag{2.15}$$

Thus again $B$ does not transform properly. This was to be expected, because right composition (pullback) is the natural transformation for maps *on* spacetime (i.e. maps having spacetime as domain) and left composition is the natural transformation for maps *into* spacetime (i.e. having spacetime as target). Since a diffeomorphism $\varphi$ is simultaneously a map on and in spacetime, one should act both ways. Indeed, consider now the "diagonal subgroup" acting by conjugation

$$\varphi \mapsto \varphi' = \psi^{-1} \circ \varphi \circ \psi \ . \tag{2.16}$$

In this case we find that the pullback of the metric transforms as a covariant tensor:

$$\varphi^* g \mapsto \varphi'^* g' = \psi^* \varphi^* \psi^{-1*} \psi^* g = \psi^* \varphi^* g \ . \tag{2.17}$$

This leads to the correct transformation of $B$. We have established that the action $S(\varphi)$ is invariant under $\mathcal{D}iff(M)$, acting on the metric by pullback and on $\varphi$ by conjugation.

It is important to appreciate the following point: whereas the actions of $\mathcal{D}iff(M)$ on itself by left and right composition are transitive (every $\varphi \in \mathcal{D}iff(M)$ can be mapped to any other $\varphi'$ by a right- or left-composition) and free (there are no fixed points), the diagonal action is not. In fact, the diagonal action leaves the identity map invariant.

In Appendix B we discuss closely related models where the domain and target spaces are kept separate, and show that the action is separately invariant under left- and right-diffeomorphisms. In contrast to those models, neither of these group actions leaves the action (2.9) invariant.

Finally we observe that the identification of the spacetime and target space metrics leads to peculiar properties also from the point of view of dimensional analysis. Since the two metrics appearing in $B$ are functions of $x$ and $\varphi(x)$, respectively, it is most natural to assume that the fields $\varphi^\alpha$ have the same dimension as the coordinates $x^\mu$. [6] We will assume this throughout this paper. Then, the tensor $B^\mu{}_\nu$ is dimensionless, and so are the traces $\tau_n$. The couplings $f^2$ and $c$

[6] It is generally the case that when the Lagrangian contains non-polynomial interactions, the fields should be dimensionless. This, together with $[\varphi] = [x]$ leads to dimensionless coordinates, a choice that we find most natural also for other reasons. However, we do not need to commit to this choice here.

in the Lagrangians (2.10) and (2.11) have mass dimension equal to the spacetime dimension $d$. [7] In fact, the coefficient of any monomial in the $\tau_n$ must have dimension $d$. Since the Lagrangian $\mathcal{L}$ must have dimension $d$, we could extract an overall factor $f^2$ with dimension $d$ and write $\mathcal{L} = f^2\tilde{\mathcal{L}}$, where $\tilde{\mathcal{L}}$ is a purely numerical function. As usual, one expects that $\tilde{\mathcal{L}}$ does not contain exceedingly large or exceedingly small coefficients. Then, $f^2$ is the only characteristic scale of the theory and it will be related to the scale of the cosmological constant.

## 2.1 Equation of motion of $\varphi$

Let $\varphi(t)$ be a one-parameter family of maps $M \to M$. The derivative

$$\left.\frac{d\varphi^\alpha(t)}{dt}\right|_{t=0} = v^\alpha \tag{2.18}$$

is a section of the vector bundle $\varphi^*TM$. The equation of motion is obtained by setting to zero the directional derivative of the action (2.9) along an arbitrary vector $v$:

$$\begin{aligned} 0 = vS &= \left.\frac{dS(\varphi(t))}{dt}\right|_{t=0} \\ &= \int dx\sqrt{g}\sum_n n\frac{\partial\mathcal{L}}{\partial\tau_n}\mathrm{Tr}B^{n-1}\left.\frac{\partial B}{\partial t}\right|_{t=0} \\ &= \int dx\sqrt{g}\sum_n n\frac{\partial\mathcal{L}}{\partial\tau_n}(B^{n-1})^{\sigma\nu}\left(2\partial_\nu v^\beta g_{\beta\alpha}(\varphi) + \partial_\nu\varphi^\beta v^\gamma\partial_\gamma g_{\beta\alpha}(\varphi)\right)\partial_\sigma\varphi^\alpha \ . \end{aligned} \tag{2.19}$$

We used the fact that $(B^{n-1})^{\sigma\nu}$ is symmetric. In the first term we have to extract the variation $v$ from the derivative. We use that $\partial_\nu v^\beta = \nabla_\nu v^\beta$ and use the standard rules for integration by parts. In this way we find

$$\begin{aligned} 0 = vS &= \int dx\sqrt{g}\, v^\alpha\Bigg\{ -\nabla_\mu\left[2\sum_n n\frac{\partial\mathcal{L}}{\partial\tau_n}(B^{n-1})^\mu{}_\rho g^{\rho\tau}\partial_\tau\varphi^\beta g_{\beta\alpha}(\varphi)\right] \\ &\qquad + \sum_n n\frac{\partial\mathcal{L}}{\partial\tau_n}(B^{n-1})^\mu{}_\rho g^{\rho\tau}\partial_\tau\varphi^\beta\partial_\alpha g_{\beta\gamma}(\varphi)\partial_\mu\varphi^\gamma\Bigg\} \\ &= -2\int dx\sqrt{g}\, v^\alpha g_{\alpha\beta}(\varphi)\mathcal{D}_\mu\left[\sum_n n\frac{\partial\mathcal{L}}{\partial\tau_n}(B^{n-1})^{\mu\nu}\partial_\nu\varphi^\beta\right] \ . \end{aligned} \tag{2.20}$$

In the last step, we used (2.3) in the terms where $\nabla_\mu$ hits $g_{\beta\alpha}(\varphi)$. The resulting terms combine with the second line to produce a Christoffel symbol that enters in the covariant derivative $\mathcal{D}_\mu$. Thus the equation of motion reads

$$\mathcal{D}_\mu\left[\sum_n n\frac{\partial\mathcal{L}}{\partial\tau_n}(B^{n-1})^{\mu\nu}\partial_\nu\varphi^\beta\right] = 0 \ . \tag{2.21}$$

[7] Note that in ordinary nonlinear sigma models $f^2$ would have dimension $d-2$ and $c$ would have dimension $d-4$. The difference is due to the identification $h = g$. Also note that one could absorb $f^2$ in the metric, thereby making it dimensionless.

One can be more explicit in special cases. For example, in the case of the action (1.2), the only term in the sum has $\frac{\partial \mathcal{L}}{\partial \tau_1} = 1$ and $B^0 = \mathbf{1}$, or $(B^0)^{\mu\nu} = g^{\mu\nu}$. Using the aforementioned rules, and recalling that $\partial_\mu \varphi^\alpha = \nabla_\mu \varphi^\alpha$, this leads to the equation for harmonic maps:

$$\mathcal{D}^\mu \nabla_\mu \varphi^\beta = 0 \ . \tag{2.22}$$

Applying this to the other special case (2.11), the equation of motion becomes

$$\mathcal{D}_\mu [\tau_1 \nabla^\mu \varphi^\beta - B^{\mu\nu} \partial_\nu \varphi^\beta] = 0 \ . \tag{2.23}$$

It will be useful to rewrite (2.22) in another way. Expanding the covariant derivatives one has

$$0 = g^{\mu\nu} \left( \partial_\mu \partial_\nu \varphi^\alpha - \Gamma_\mu{}^\rho{}_\nu \partial_\rho \varphi^\alpha + \partial_\mu \varphi^\beta \Gamma_\beta{}^\alpha{}_\gamma \partial_\nu \varphi^\gamma \right) \ . \tag{2.24}$$

Since $\varphi$ is a diffeomorphism, from the transformation properties of the connection we can rewrite the first two terms as

$$\partial_\mu J_\nu{}^\alpha - \Gamma_\mu{}^\rho{}_\nu J_\rho{}^\alpha = -J_\mu{}^\beta J_\nu{}^\gamma \Gamma'_\beta{}^\alpha{}_\gamma \tag{2.25}$$

and therefore the equation for harmonic diffeomorphisms amounts to the statement that

$$0 = g'^{\beta\gamma} \left( \Gamma_\beta{}^\alpha{}_\gamma - \Gamma'_\beta{}^\alpha{}_\gamma \right) \ , \tag{2.26}$$

where $g'^{\beta\gamma} = g^{\mu\nu} \partial_\mu \varphi^\beta \partial_\nu \varphi^\gamma$ and $\Gamma'$ are the Christoffel symbols of $g'$. Since the difference of two connections is a tensor, this is a covariant statement.

## 2.2 Energy-momentum tensor

Next, we vary the action with respect of $g_{\mu\nu}(x)$:

$$\begin{aligned} \delta S &= \int dx \sqrt{g} \left\{ \frac{1}{2} g^{\mu\nu} \delta g_{\mu\nu} \mathcal{L} + \sum_n n \frac{\partial \mathcal{L}}{\partial \tau_n} \mathrm{Tr} B^{n-1} \delta B \right\} \\ &= \int dx \sqrt{g} \left\{ \frac{1}{2} g^{\mu\nu} \delta g_{\mu\nu} \mathcal{L} + \sum_n n \frac{\partial \mathcal{L}}{\partial \tau_n} (B^{n-1})^\sigma{}_\mu \Big( - g^{\mu\rho} \delta g_{\rho\tau} g^{\tau\lambda} \partial_\lambda \varphi^\gamma g_{\gamma\delta}(\varphi) \partial_\sigma \varphi^\delta \right. \\ &\qquad \left. + g^{\mu\lambda} \partial_\lambda \varphi^\alpha \delta g_{\alpha\beta}(\varphi) \partial_\sigma \varphi^\beta \Big) \right\} . \end{aligned}$$

We can obtain the energy-momentum tensor by straightforwardly using the rules of variational calculus and evaluating

$$T^{\mu\nu} = \frac{2}{\sqrt{g}} \frac{\delta S}{\delta g_{\mu\nu}} \ . \tag{2.27}$$

It is perhaps more instructive to observe that in the last term the variation appears in the combination $\partial_\lambda \varphi^\alpha \partial_\sigma \varphi^\beta \delta g_{\alpha\beta}(\varphi) = (\varphi^* \delta g)_{\mu\nu}(x)$. We exploit the fact that the integral does not change if we replace the integrand by its transform under a diffeomorphism. Let $x'^\alpha = \varphi^\alpha(x)$, and denote by a prime all transformed tensors. Covariant tensors are pulled back by $\varphi^{-1}$ and contravariant tensors are pushed forward with $\varphi$: $A' = \varphi_* A$, $g' = \varphi^{-1*} g$, and $(\varphi^* \delta g)' = \delta g$. Then the last term can be manipulated as follows:

$$\int dx \sqrt{g(x)} \, A^{\mu\nu}(x) (\varphi^* \delta g)_{\mu\nu}(x) = \int dx' \sqrt{g'(x')} \, A'^{\alpha\beta}(x') \delta g_{\alpha\beta}(x') \ ,$$

where $x'$ is to be regarded as independent integration variable. Then, from the definition given above, we obtain

$$T^{\mu\nu} = g^{\mu\nu}\mathcal{L} - 2\sum_n n\frac{\partial\mathcal{L}}{\partial\tau_n}(B^{n-1})^{\sigma\mu}\nabla^\nu\varphi^\gamma g_{\gamma\delta}(\varphi)\nabla_\sigma\varphi^\delta + 2\frac{\sqrt{g'}}{\sqrt{g}}\sum_n n\left(\frac{\partial\mathcal{L}}{\partial\tau_n}\right)'(B'^{n-1})^{\mu\nu} \ . \quad (2.28)$$

In particular for the Lagrangian (2.10)

$$T^{\mu\nu} = f^2\left[\nabla^\mu\varphi^\alpha\nabla^\nu\varphi^\beta g_{\alpha\beta}(\varphi) - \frac{1}{2}g^{\mu\nu}g^{\rho\sigma}\nabla_\rho\varphi^\alpha\nabla_\sigma\varphi^\beta g_{\alpha\beta}(\varphi) - \frac{\sqrt{g'}}{\sqrt{g}}g'^{\mu\nu}\right] \ , \quad (2.29)$$

whereas for (2.11)

$$T^{\mu\nu} = c\,g^{\mu\nu}(\tau_1^2 - \tau_2) - 4c(\tau_1 g^{\sigma\mu} - B^{\sigma\mu})\nabla^\nu\varphi^\gamma g_{\gamma\delta}(\varphi)\nabla_\sigma\varphi^\delta + 4\frac{\sqrt{g'}}{\sqrt{g}}c(\tau_1' g'^{\mu\nu} - B'^{\mu\nu}) \ . \quad (2.30)$$

### 2.3 Diffeomorphism invariance, EOM and EM conservation

Let us begin by recalling the general argument relating diffeomorphism invariance to EM conservation. Given an action $S(\varphi, g)$ for the "matter" fields $\varphi$ coupled to a metric $g$, its variation under an infinitesimal diffeomorphism $\xi$ is

$$\delta_\xi S = \int dx\left[\frac{\delta S}{\delta\varphi^\alpha}\delta_\xi\varphi^\alpha + \frac{\delta S}{\delta g_{\mu\nu}}\delta_\xi g_{\mu\nu}\right] \ . \quad (2.31)$$

We define the equation of motion

$$E^\alpha = \frac{1}{\sqrt{g}}\frac{\delta S}{\delta\varphi^\alpha} \ . \quad (2.32)$$

An infinitesimal diffeomorphism is defined by

$$\delta_\xi x^\mu = -\xi^\mu(x) \ , \quad (2.33)$$

where $\xi$ is a vector field. The infinitesimal variation of $\varphi$ is

$$\delta_\xi\varphi^\alpha(x) = \xi^\lambda\partial_\lambda\varphi^\alpha - \xi^\alpha(\varphi) \ , \quad (2.34)$$

where the first term comes from the right composition and the second from the left composition. The variation of any tensor $T$ is its Lie derivative $\delta_\xi T = \mathcal{L}_\xi T$. For the metric

$$\delta_\xi g_{\mu\nu} = \xi^\rho\partial_\rho g_{\mu\nu} + g_{\mu\rho}\partial_\nu\xi^\rho + g_{\rho\nu}\partial_\mu\xi^\rho = \nabla_\mu\xi_\nu + \nabla_\nu\xi_\mu \ .$$

Inserting these formulae and (2.27) in (2.31), invariance of the action implies

$$0 = \int d^4x\sqrt{g}\left[(\xi^\tau\nabla_\tau\varphi^\alpha - \xi^\alpha(\varphi))E_\alpha - \xi_\mu\nabla_\nu T^{\mu\nu}\right] \ . \quad (2.35)$$

We find, as expected, that $E_\alpha = 0$ implies EM conservation.

In the case of our theory of diffeomorphisms, there is more to be learned. We observe that for a generic $\varphi$, the coefficient of $E_\alpha$ is non-vanishing and therefore, conversely EM conservation also generically implies the EOM. This makes sense, because both statements amount to four second order differential equations for the fields. However, for the identity map $\varphi^\alpha = x^\alpha$, the

coefficient of $E_\alpha$ vanishes and therefore this implication does not hold. This is a consequence of the identity map being a fixed point of the action of the diffeomorphism group.

One could further explicitly compute the divergence of the EM tensor. This calculation is very complicated in general, but we can do it in the case of the Lagrangian (2.10). It turns out to be useful to split the EM tensor in two parts:

$$T^{\mu\nu} = T^{\mu\nu}_{(R)} + T^{\mu\nu}_{(L)} \ , \tag{2.36}$$

where the first part arises from the variations with respect to $g_{\mu\nu}$ and consists of the first two terms in (2.29), the second part comes from variation with respect to $g_{\alpha\beta}(\varphi)$ and consists of the third term in (2.29). Interestingly, these two parts are separately conserved. The conservation of $T^{\mu\nu}_{(R)}$ works exactly as for a nonlinear sigma model:

$$\nabla_\mu T^{\mu\nu}_{(R)} = \nabla^2\varphi^\alpha \nabla^\nu \varphi^\beta g_{\alpha\beta}(\varphi) + \left( \nabla^\mu \varphi^\alpha \nabla^\nu \varphi^\beta \nabla_\mu \varphi^\gamma - \frac{1}{2} \nabla^\mu \varphi^\alpha \nabla_\mu \varphi^\beta \nabla_\nu \varphi^\gamma \right) \partial_\gamma g_{\alpha\beta}(\varphi) \ .$$

The two terms in parentheses reconstruct a Christoffel symbol, and the whole expression is then seen to be proportional to the EOM, written in the form (2.24).

For the second part,

$$\begin{aligned} \nabla_\mu T^{\mu\nu}_{(L)} &= -f^2 \frac{1}{\sqrt{g}} \nabla_\mu (\sqrt{g'} g'^{\mu\nu}) \\ &= -f^2 \frac{1}{\sqrt{g}} \left( \partial_\mu (\sqrt{g'} g'^{\mu\nu}) + \sqrt{g'}\, \Gamma_\mu{}^\nu{}_\rho g'^{\mu\rho} \right) \\ &= -f^2 \frac{\sqrt{g'}}{\sqrt{g}}\, g^{\mu\rho} \left( \Gamma_\mu{}^\nu{}_\rho - \Gamma'_\mu{}^\nu{}_\rho \right) \end{aligned} \tag{2.37}$$

which vanishes due to the EOM written in the form (2.26).

While these statements are easy to check in the case of the Lagrangian (2.10), calculating the divergence of the EM tensor in the general case is very complicated. Still, the preceding statements remain true. One can see this by following in detail the proof of invariance of the action under infinitesimal diffeomorphisms, which is given in Appendix A. One can see there that the terms coming from variations of $g_{\mu\nu}$ (i.e. the divergence of $T^{\mu\nu}_{(R)}$) and those coming from variations of $g_{\alpha\beta}(\varphi)$ (i.e. the divergence of $T^{\mu\nu}_{(L)}$) cancel separately against terms coming from the variation of $\varphi$ (i.e. the EOM). More precisely, the differential identity (2.35) can be seen as the sum of two separate identities

$$\int d^4x \sqrt{g} \left[ \xi^\tau \nabla_\tau \varphi^\alpha E_\alpha - \xi_\mu \nabla_\nu T^{\mu\nu}_{(R)} \right] = 0 \ , \tag{2.38}$$

$$\int d^4x \sqrt{g} \left[ -\xi^\alpha(\varphi) E_\alpha - \xi_\mu \nabla_\nu T^{\mu\nu}_{(L)} \right] = 0 \ . \tag{2.39}$$

This is surprising, because the invariance group of the action has four parameters but it seems to imply eight differential identities. This can be explained by looking at the models as special cases of theories with different domain and target space. In these cases, as explained in Appendix B, the invariance group consists separately of left and right diffeomorphisms and therefore implies eight differential identities. The models with identical domain and target space are obtained by choosing a preferred diffeomorphism, and this does not invalidate the identities.

# 3 The identity solution

The identity $\varphi = Id_M$ is represented, in any local coordinate system, by

$$\varphi^\alpha = x^\alpha \ . \tag{3.1}$$

In this case the Jacobian reduces to $\partial_\mu \varphi^\alpha = \delta^\alpha_\mu$, $B^{\mu\nu} = g^{\mu\nu}$ and all the traces become constant: $\tau_n = 4$ for $n = 1, 2, 3, 4$. Thus $\mathcal{L}$ and its derivatives are just constants and the EOM (2.21) are satisfied.

The energy-momentum tensor of the solution becomes proportional to the metric. The last two terms in (2.27) cancel out and

$$T^{\mu\nu} = g^{\mu\nu}\mathcal{L} - 2\sum_n n\frac{\partial \mathcal{L}}{\partial \tau_n} g^{\mu\nu} + 2\sum_n n\Big(\frac{\partial \mathcal{L}}{\partial \tau_n}\Big) g^{\mu\nu} = g^{\mu\nu}\mathcal{L}(4,4,4,4) \ . \tag{3.2}$$

We can therefore interpret this energy-momentum tensor as an effective cosmological constant $\Lambda = 8\pi G\, \mathcal{L}(4,4,4,4)$. As already anticipated, if we assume that $\tilde{\mathcal{L}}(4,4,4,4)$ is a number of order one, the effective cosmological constant is $\Lambda \approx 8\pi G f^2$.

We will now discuss the stability of the identity solution. This can have two possible meanings: in Euclidean signature one asks whether the action has a minimum at the solution; in Lorentzian signature one asks whether the energy has a minimum. We begin by discussing the simpler problem of Euclidean stability. Since the Euclidean action is also identical to the energy of a static field configuration (in one dimension more) this analysis also says something about the stability of static configurations under static deformations. Full Lorentzian stability will be briefly discussed in Section 3.3, where we shall refer to existing results in the literature.

## 3.1 The second variation

Here we compute the Hessian of the action at the identity solution. This is needed to establish whether a Euclidean solution is stable, and is also needed in the study of linearized perturbations. Let $\varphi(t,s)$ be a two-parameter family of maps. We take the double derivatives at $t = s = 0$. We let

$$\left.\frac{d\varphi^\alpha(t,s)}{dt}\right|_{t=s=0} = v^\alpha \ , \qquad \left.\frac{d\varphi^\alpha(t,s)}{ds}\right|_{t=s=0} = w^\alpha \ . \tag{3.3}$$

The Hessian is defined by

$$\begin{aligned} H(v,w) &= \left.\frac{\partial^2 S(\varphi(t,s))}{\partial t \partial s}\right|_{t=s=0} \\ &= -2\frac{\partial}{\partial s}\int dx\sqrt{g}\,\frac{\partial \varphi^\alpha}{\partial t} g_{\alpha\beta}\left(\mathcal{D}_\mu \sum_n n\frac{\partial \mathcal{L}}{\partial \tau_n}(B^{n-1})^{\mu\nu}\partial_\nu \varphi^\beta\right)\Bigg|_{t=s=0} \ . \end{aligned} \tag{3.4}$$

The derivatives with respect to $s$ of the terms $\frac{\partial \varphi^\alpha}{\partial t}$ and $g_{\alpha\beta}$ are proportional to the EOM, and since we are interested in the variation around a solution, we can neglect them. Acting with the $s$-derivative on all the remaining occurrences of $\varphi$, and evaluating at $t = s = 0$, which correspond to $\varphi = Id_M$, the round bracket becomes

$$\begin{aligned} &\sum_n n\frac{\partial \mathcal{L}}{\partial \tau_n}\mathcal{D}^\mu \partial_\mu w^\beta + \sum_n n\frac{\partial \mathcal{L}}{\partial \tau_n}\mathcal{D}_\mu\left(\frac{\partial}{\partial s}(B^{n-1})^{\mu\beta}\Big|_{s=0}\right) + \sum_{n,m} n\frac{\partial^2 \mathcal{L}}{\partial \tau_n \partial \tau_m} D^\beta\left(\left.\frac{\partial \tau_m}{\partial s}\right|_{s=0}\right) \\ &+ \sum_n n\frac{\partial \mathcal{L}}{\partial \tau_n} g^{\mu\delta}\left(\partial_\mu w^\gamma \Gamma_\gamma{}^\beta{}_\delta + w^\epsilon \partial_\epsilon \Gamma_\mu{}^\beta{}_\delta\right) \ . \end{aligned} \tag{3.5}$$

Now consider the general formula

$$\begin{aligned}\mathcal{D}_\mu D_\nu w^\delta &= \nabla_\mu\nabla_\nu w^\delta + \nabla_\mu\nabla_\nu\varphi^\gamma\Gamma_\gamma{}^\delta{}_\beta w^\beta + \partial_\nu\varphi^\gamma\partial_\mu\varphi^\beta\partial_\beta\Gamma_\gamma{}^\delta{}_\epsilon w^\epsilon \\ &\quad +\partial_\nu\varphi^\gamma\Gamma_\gamma{}^\delta{}_\beta\partial_\mu w^\beta + \partial_\mu\varphi^\gamma\Gamma_\gamma{}^\delta{}_\epsilon\left(\partial_\nu w^\epsilon + \partial_\nu\varphi^\zeta\Gamma_\zeta{}^\epsilon{}_\beta w^\beta\right) \end{aligned} \tag{3.6}$$

and specialize to the identity map. Comparing this with the preceding formula, and using $\nabla_\mu\nabla_\nu\varphi^\beta = -\Gamma_\mu{}^\beta{}_\nu$, we see that the first and the last term in (3.5) combine to give

$$\sum_n n\frac{\partial\mathcal{L}}{\partial\tau_n}\left(\mathcal{D}^\mu D_\mu w^\beta + R^\beta{}_\gamma w^\gamma\right) \ .$$

At the identity, we can convert all indices to $\mu$, $\nu$ etc and $\mathcal{D}^\mu D_\mu = \nabla^2$. In the remaining two terms we have

$$\frac{\partial}{\partial s}(B^{n-1})^{\mu\nu}\bigg|_{s=0} = (n-1)\left(\nabla^\mu w^\nu + \nabla^\nu w^\mu\right)$$

and

$$\frac{\partial\tau_m}{\partial s}\bigg|_{s=0} = 2m\nabla_\mu w^\mu \ .$$

Putting everything together, the second variation of the action around the identity is

$$\begin{aligned}H(v,w) &= 2\int dx\sqrt{g}\Bigg[\sum_n n^2\frac{\partial\mathcal{L}}{\partial\tau_n}v^\mu\left(-\nabla^2 g_{\mu\nu} - R_{\mu\nu}\right)w^\nu \\ &\quad +\left(\sum_n n(n-1)\frac{\partial\mathcal{L}}{\partial\tau_n} + 2\sum_{m,n} mn\frac{\partial^2\mathcal{L}}{\partial\tau_n\partial\tau_m}\right)(\nabla_\mu v^\mu)(\nabla_\nu w^\nu)\Bigg] \ . \end{aligned} \tag{3.7}$$

We can write $H(v,w) = \langle v, Lw\rangle$, where $\langle\ ,\ \rangle$ is the natural inner product in the space of sections of $\varphi^*TN$ and $L$ represents the differential operator

$$\begin{aligned}L &= \sum_n n^2\frac{\partial\mathcal{L}}{\partial\tau_n}\left(-\nabla^2 g_{\mu\nu} - R_{\mu\nu}\right) \\ &\quad -\left(\sum_n n(n-1)\frac{\partial\mathcal{L}}{\partial\tau_n} + 2\sum_{m,n} mn\frac{\partial^2\mathcal{L}}{\partial\tau_n\partial\tau_m}\right)\nabla_\mu\nabla_\nu \ . \end{aligned} \tag{3.8}$$

Note that the factors involving $\mathcal{L}$ are constants. Using the rules for integrations by parts, one sees that the Hessian is symmetric:

$$H(v,w) = H(w,v) \tag{3.9}$$

which also means that the differential operator $L$ is self-adjoint.

Clearly, stability hinges on the form of the Lagrangian. As a first example consider the action for harmonic diffeomorphisms (1.2). Its Hessian is [52, 53]:

$$H(v,w) = 2\int dx\sqrt{g}\, v^\mu\left(-\nabla^2 g_{\mu\nu} - R_{\mu\nu}\right)w^\nu \ . \tag{3.10}$$

Hence we have to study te spectrum of the Laplace-type operator

$$L_1 = -g_{\mu\nu}D^2 - R_{\mu\nu} \ . \tag{3.11}$$

This spectrum is known for spheres. It consists of transverse and longitudinal fields. The lowest transverse eigenfunctions are the $d(d+1)/2$ Killing vectors, which are zero modes. This is related to the fact that every isometry is harmonic. The lowest longitudinal eigenfunctions are

$$w^i_\nu = \partial_\nu \phi^i \ , \tag{3.12}$$

where $\phi^i$ are cartesian coordinates of the flat Euclidean space in which the sphere is embedded. These are conformal Killing vectors and have eigenvalue $-\frac{d-2}{d(d-1)}R$ and multiplicity $d+1$. In $d=2$ they are zero modes; this is related to the fact that in $d=2$ conformal isometries are harmonic [54]. All other eigenvalues are positive. Thus, the identity is unstable as a harmonic map of spheres in $d>2$; in $d=2$ it is stable and belongs to a six-parameter family of degenerate solutions.

As another example we consider the model (2.11). In this case the Hessian is

$$H(v,w) \;\; = \;\; 4\int dx\sqrt{g}\left[(d-2)v^\mu\left(-\nabla^2 g_{\mu\nu} - R_{\mu\nu}\right)w^\nu + (\nabla_\mu v^\mu)(\nabla_\nu w^\nu)\right] \tag{3.13}$$

and, integrating the second term by parts, the associated operator reads

$$L_2 = (-\nabla^2 g_{\mu\nu} - R_{\mu\nu}) - \frac{1}{d-2}\nabla_\mu\nabla_\nu \ . \tag{3.14}$$

Clearly the spectrum on transverse vectors is the same as that of $L_1$, but it may differ on longitudinal vectors, and the new contribution is positive, so that the identity may become stable. In fact, the additional non-minimal term, acting on the eigenfunctions (3.12), gives $R/(d-1)(d-2)$, so that the eigenvalue becomes $-\frac{d-4}{d(d-2)}R$. It is negative in all dimensions except four, were it is zero. Also in this case, the zero modes are related to the infinitesimal isometries and conformal isometries.

### 3.2 Global Euclidean bounds

The second variation of the action gives information about the local stability of a solution, but there are some cases where absolute bounds on the action can be derived. In this section we assume that $M$ is compact without boundary. We use the totally antisymmetric tensor $\eta_{\mu_1\ldots\mu_d} = \sqrt{g}\,\varepsilon_{\mu_1\ldots\mu_d}$ where $\varepsilon$ is the tensor density with components $\pm 1$. It is the volume form on $M$, such that $V = \int_M \eta = \int d^dx\sqrt{g}$ is the volume. The winding number is

$$W = \frac{1}{d!V}\int d^dx\sqrt{g}\,\eta^{\mu_1\ldots\mu_d} J_{\mu_1}{}^{\alpha_1}\ldots J_{\mu_d}{}^{\alpha_d}\eta_{\alpha_1\ldots\alpha_d} \tag{3.15}$$

and is equal to one for orientation-preserving diffeomorphisms.

The action for harmonic maps in two Euclidean dimensions is $S = \frac{1}{2}f^2\int d^2x\sqrt{g}\,\tau_1$. We define the double dual ${}^*J^*$ by ${}^*J^{*\alpha}_\mu = \eta_\mu{}^\rho J_\rho{}^\gamma \eta_\gamma{}^\alpha$. Integrating the square of $J \pm {}^*J^*$ one obtains the well-known bound $S \geq Vf^2|W|$, [55] and since $|W|=1$ for diffeomorphisms,

$$S \geq Vf^2 \ . \tag{3.16}$$

The absolute minima are the maps for which $J = \pm {}^*J^*$. Indeed, the identity solves this equation.

There is a parallel example in four Euclidean dimensions. This time we consider the Lagrangian (2.11). Defining the antisymmetric tensor $K_{\mu\nu}{}^{\alpha\beta} = J_{[\mu}{}^{[\alpha}J_{\nu]}{}^{\beta]}$, and the inner product

$$(K,K') = \frac{1}{2}\int d^4x\sqrt{g}\,g^{\mu\rho}g^{\nu\sigma}K_{\mu\nu}{}^{\alpha\beta}K'_{\rho\sigma}{}^{\gamma\delta}g_{\alpha\gamma}g_{\beta\delta} \ ,$$

this action can be rewritten

$$S = 4c||K||^2 \equiv 2c \int d^4x \sqrt{g}\, g^{\mu\rho} g^{\nu\sigma} K_{\mu\nu}{}^{\alpha\beta} K_{\rho\sigma}{}^{\gamma\delta} g_{\alpha\gamma} g_{\beta\delta} \ . \tag{3.17}$$

We can define ${}^*K^*$, the double dual of $K$, by

$${}^*K^*_{\mu\nu}{}^{\alpha\beta} = \frac{1}{4} \eta_{\mu\nu}{}^{\rho\sigma} \eta^{\alpha\beta}{}_{\gamma\delta} K_{\rho\sigma}{}^{\gamma\delta} \ . \tag{3.18}$$

Since $||{}^*K^*||^2 = ||K||^2$, we have

$$0 \leq ||K \pm {}^*K^*||^2 = 2||K||^2 \pm 2(K, {}^*K^*) = \frac{1}{2c} S \pm 6VW[\varphi] \ , \tag{3.19}$$

where $W$ is the winding number of $\varphi$. Since diffeomorphisms have $|W| = 1$, we have the absolute bound

$$S \geq 12\, cV \ . \tag{3.20}$$

The bound is saturated by maps for which $H$ is double-self-dual, and the identity has this property.

For a generic metric, the identity will be an isolated solution. In the presence of isometries and/or conformal isometries, it will be an element of continuous degenerate families of solutions, as we have seen in the end of the preceding section.

### 3.3 Lorentzian stability

As long as we restrict ourselves to dynamical diffeomorphisms in a fixed external metric, there is no difference, regarding the issues of Lorentzian stability, between the models considered here, where the target space is spacetime itself, and those where the target space is another manifold, as in nonlinear sigma models, or a copy of the same manifold. In all cases, when the target space metric has Minkowski signature $-+++$, the scalar associated to the time coordinate has a kinetic term with opposite sign of those associated to the space coordinates. Assuming that the sign of the Lagrangian is such that the latter have the correct dynamics, the time-like scalar will be a ghost.

As already mentioned, these issues have been discussed extensively in the literature on massive gravity. where two different strategies have been developed in order to avoid ghost instabilities: the dRGT models, that preserve spacetime covariance, and Lorentz-breaking models. Both strategies can be used also for our models.

In the context of dRGT models of massive gravity, it has been observed that certain nonlinear sigma models with Minkowskian target space and special actions built out of $\mathrm{Tr}(\sqrt{B})^n$ will be ghost-free [56]. By expanding $B = 1 + X$ in terms of $X$, one can write $\mathrm{Tr}(\sqrt{B})^n$ as an infinite series of $\mathrm{Tr}B^n$. Hence, these actions can be viewed as special cases of our action and the same construction can be applied also to our models.

In Lorentz-breaking massive gravity, one gives up Lorentz invariance and preserves only Euclidean symmetry of 3-dimensional space. Dubovsky and Rubakov [16, 17] extended the non-linear Stückelberg trick to Lorentz-violating massive gravity using the pullback of a 3-dimensional target space metric $h_{ab}$. Similarly, we can construct ghost-free (but Lorentz-violating) dynamics of the diffeomorphisms of space, rather than spacetime. The target space metric $h(\varphi)$ would be only 3-dimensional in that case. The solution $\varphi = Id_\Sigma$ is again a solution of the EOM and it generates an effective cosmological term in the space directions.

# 4 Conclusions

We have constructed a dynamical theory for diffeomorphisms of the spacetime manifold, which is closely related to models of dynamical coordinates. From a geometrical point of view, coordinates are (locally defined) maps of spacetime into a fixed Euclidean space, whereas diffeomorphisms are maps of spacetime to itself. Thus, dynamical diffeomorphisms naturally differ from dynamical coordinates because of the identification of spacetime with the target space.

One motivation given in the introduction was that in general, coordinates are only defined locally and to cover a manifold with nontrivial topology, several coordinate patches are needed. One could overcome this problem by considering generalized models with a target space that is homeomorphic to spacetime itself, rather than flat Euclidean space. One would then have a theory of diffeomorphisms from one copy of spacetime into another. Such models would still differ from the one considered here as long as the target space metric is fixed. The true difference lies in the identification of the spacetime and target space, and the respective metrics.

This lies at the root of all the other differences that we have encountered. First we have differences in the kinematics. Suppose the dynamical coordinates are thought of as scalar fields on spacetime. Then diffeomorphisms act on them by right composition. If the coordinates are viewed as a map into spacetime (e.g. as fluid, analogous to comoving coordinates in cosmology), diffeomorphisms act on them by left composition. In both cases, the action of the group is free and transitive. For a map of spacetime into itself, diffeomorphisms must act at both ends, resulting in an action by conjugation, which is neither transitive nor free.

These kinematical differences inevitably give rise to differences in the dynamics: the energy-momentum tensor of diffeomorphisms has an additional term that is not present in the theories of dynamical coordinates, and this term is conserved separately from the rest. The identity is always a solution, but the effective cosmological constant it generates is different in the two cases. For dynamical diffeomorphisms, it is given just by the term that comes from the variation of $\sqrt{g}$ in the action. [8]

In this paper we have restricted our attention to the dynamics of the diffeomorphisms by themselves. We observe that if we allow the metric to become dynamical, then the theory is no longer local: the tensor $B$ of (2.7), and consequently also the action (2.9), depends on the metric evaluated at two different points. It would be interesting to investigate the possible couplings of matter fields to diffeomorphisms. Here many physical applications are possible. For example it will be interesting to couple diffeomorphisms to the Einstein-Hilbert action and see how our treatment would differ from existing models of massive gravity. Coupling diffeomorphisms to matter fields can introduce a mild form of non-locality in the theory.

Another motivation given in the introduction was the issue of observables in gravity. As in the case of Yang-Mills theories, local, gauge-invariant expressions can be constructed if a suitable "compensator" or "Stückelberg" field is present. The configuration space of this field must be a copy of the gauge group, and this is the case for our models, but more is required: the action of the gauge group on the compensator field must be free and transitive. To see this, it is enough to consider again the simple example of a reference fluid, already mentioned in the introduction. Let $F$ be a manifold labelling the fluid elements and $\Phi : F \to M$ describe a configuration of the fluid. [9] If $R : M \to \mathbb{R}$ is the Ricci scalar, for example, $R \circ \Phi$ is the Ricci scalar evaluated in the reference frame provided by the fluid. This will be gauge invariant provided diffeomorphisms of

[8] We observe here that in unimodular gravity this contribution would be absent, giving rise to a vanishing effective cosmological constant. The resulting EM tensor is generally not conserved, but as long as we are only interested in the identity solution, it actually is.

[9] $\Phi$ is invertible and we can identify $\Phi^{-1}$ with the fields $\varphi$ of Appendix B.

$M$ act on $\Phi$ by left composition: $R\circ\Phi\mapsto R\circ\psi\circ\psi^{-1}\circ\Phi$. However, if we identify $F=M$, the diffeomorphisms of $M$ act on $\Phi$ by conjugation, and $R\circ\Phi$ is no longer invariant, being mapped to $R\circ\psi\circ\psi^{-1}\circ\Phi\circ\psi$. The lesson to be drawn is that in order to construct relational gravitational observables, it is not useful to consider diffeomorphisms of the spacetime manifold to itself. One has to keep the domain and target space separate, as in the models discussed in Appendix B.

We have limited ourselves here to general considerations. Other aspects of the dynamics of diffeomorphisms, especially the small fluctuations around the identity, have been discussed in a framework that is very similar to ours in [57]. It was shown there that for Ricci-flat manifolds the linearised field equations are Maxwell's equations in the Lorentz gauge. Furthermore, for Minkowski space explicit massive and massless solutions of the resulting nonlinear field equations were constructed; the massless ones coming in two distinct types, right-handed and left-handed. For quantum aspects and the effective field theory description of similar models we refer for example to [58–61].

## Acknowledgments

We would like to thank M. Celoria, S. Gielen, P. A. Höhn, F. Nesti, M. Reuter for useful discussions and correspondence.

## A Variation under infinitesimal diffeomorphisms

In this appendix we prove the invariance of the action (2.9) under infinitesimal diffeomorphisms. There are variations of $g$ and variations of $\varphi$:

$$\begin{aligned}
\delta_\xi S &= \int dx\sqrt{g}\Big\{\frac{1}{2}g^{\mu\nu}\delta_\xi g_{\mu\nu}\mathcal{L}-\sum_n n\frac{\partial\mathcal{L}}{\partial\tau_n}(B^{n-1})^{\sigma\mu}\delta_\xi g_{\mu\nu}\nabla^\nu\varphi^\gamma g_{\gamma\delta}(\varphi)\nabla_\sigma\varphi^\delta\\
&\qquad+\sum_n n\frac{\partial\mathcal{L}}{\partial\tau_n}(B^{n-1})^\sigma{}_\mu g^{\mu\lambda}\nabla_\lambda\varphi^\alpha\delta_\xi g_{\alpha\beta}(\varphi)\nabla_\sigma\varphi^\beta\\
&\qquad+\sum_n n\frac{\partial\mathcal{L}}{\partial\tau_n}\Big(2\nabla_\nu\delta_\xi\varphi^\beta g_{\beta\alpha}(\varphi)\nabla_\sigma\varphi^\alpha(B^{n-1})^\sigma{}_\mu g^{\mu\nu}\\
&\qquad+(B^{n-1})^\sigma{}_\mu g^{\mu\nu}\nabla_\nu\varphi^\beta\delta_\xi\varphi^\alpha\partial_\alpha g_{\beta\gamma}(\varphi)\nabla_\sigma\varphi^\gamma\Big)\Big\}\\
&= \int dx\sqrt{g}\Big\{-2\xi_\nu\nabla_\mu\Big[\frac{1}{2}g^{\mu\nu}\mathcal{L}-\sum_n n\frac{\partial\mathcal{L}}{\partial\tau_n}(B^{n-1})^{\sigma\mu}\nabla^\nu\varphi^\gamma g_{\gamma\delta}(\varphi)\nabla_\sigma\varphi^\delta\Big]\\
&\qquad+2D_\alpha\xi_\beta(\varphi)\sum_n n\frac{\partial\mathcal{L}}{\partial\tau_n}\nabla_\sigma\varphi^\beta(B^{n-1})^\sigma{}_\mu g^{\mu\lambda}\nabla_\lambda\varphi^\alpha\\
&\quad+\xi^\tau\nabla_\tau\varphi^\alpha\Big[-\nabla_\mu\Big(2\sum_n n\frac{\partial\mathcal{L}}{\partial\tau_n}(B^{n-1})^\mu{}_\rho g^{\rho\tau}\nabla_\tau\varphi^\beta g_{\beta\alpha}(\varphi)\Big)\\
&\qquad+\sum_n n\frac{\partial\mathcal{L}}{\partial\tau_n}(B^{n-1})^\mu{}_\rho g^{\rho\tau}\partial_\tau\varphi^\beta\partial_\alpha g_{\beta\gamma}(\varphi)\partial_\mu\varphi^\gamma\Big]
\end{aligned}$$

$$-\xi^\alpha(\varphi)\Bigg[-\nabla_\mu\Bigg(2\sum_n n\frac{\partial\mathcal{L}}{\partial\tau_n}(B^{n-1})^\mu{}_\rho g^{\rho\tau}\nabla_\tau\varphi^\beta g_{\beta\alpha}(\varphi)\Bigg) + \sum_n n\frac{\partial\mathcal{L}}{\partial\tau_n}(B^{n-1})^\mu{}_\rho g^{\rho\tau}\partial_\tau\varphi^\beta\partial_\alpha g_{\beta\gamma}(\varphi)\partial_\mu\varphi^\gamma\Bigg]\Bigg\}\,. \tag{A.1}$$

The first, third and fourth lines are the same as one would have in an ordinary nonlinear sigma model in curved space, where the target space metric is not affected by spacetime diffeomorphisms. By explicitly taking the derivatives of the first line, one finds that they cancel the third and the fourth line.

The second, fifth and sixth lines are novel. They are characterized by the fact that the infinitesimal parameter $\xi$ is evaluated at $\varphi(x)$. The second line can be expanded as follows:

$$2(\partial_\alpha\xi_\beta(\varphi)-\Gamma^\lambda_{\alpha\beta}(\varphi)\xi_\lambda(\varphi))\sum_n n\frac{\partial\mathcal{L}}{\partial\tau_n}(B^{n-1})^{\sigma\lambda}\nabla_\lambda\varphi^\alpha\nabla_\sigma\varphi^\beta\,.$$

As recalled earlier, $\xi_\lambda(\varphi)$ is a scalar under diffeomorphisms, so by the chain rule its covariant derivative is $\partial_\lambda\varphi^\alpha\partial_\alpha\xi_\beta(\varphi)=\nabla_\lambda\xi_\beta(\varphi)$. Therefore the first term can be integrated by parts and we obtain

$$-2\xi_\beta(\varphi)\nabla_\lambda\left(\sum_n n\frac{\partial\mathcal{L}}{\partial\tau_n}(B^{n-1})^{\sigma\lambda}\nabla_\sigma\varphi^\beta\right)-2\Gamma^\lambda_{\alpha\beta}(\varphi)\xi_\lambda(\varphi)\sum_n n\frac{\partial\mathcal{L}}{\partial\tau_n}(B^{n-1})^{\sigma\lambda}\nabla_\lambda\varphi^\alpha\nabla_\sigma\varphi^\beta \tag{A.2}$$

Next, in the fifth line we separate the covariant derivative acting on $g_{\alpha\beta}(\varphi)$ from the rest, and we get

$$2\xi_\beta(\varphi)\nabla_\mu\left(\sum_n n\frac{\partial\mathcal{L}}{\partial\tau_n}(B^{n-1})^{\mu\tau}\nabla_\tau\varphi^\beta\right)+2\xi^\alpha(\varphi)\sum_n n\frac{\partial\mathcal{L}}{\partial\tau_n}(B^{n-1})^{\mu\tau}\nabla_\mu\varphi^\gamma\nabla_\tau\varphi^\beta\partial_\gamma g_{\alpha\beta}(\varphi) \tag{A.3}$$

The first term cancels the first term of (A.2). The second term combines with the sixth line to reconstruct a Christoffel symbol, and the result cancels with the second term in (A.2). In this way also these terms cancel out, and we have proven the invariance of the action.

We can now also see the differential identities that follows from $\mathcal{D}iff(M)$-invariance. To this end, we have to work a little more on the second line of (A.1). Namely, as in the derivation of the EM tensor, we want to have the infinitesimal parameter $\xi$ evaluated at $x$ rather than $\varphi(x)$. This is dealt with by the method already used in Sect. 2.2, namely changing coordinates from $x$ to $x'=\varphi(x)$. Then the second line becomes

$$\int dx\sqrt{g'}\sum_n n\left(\frac{\partial\mathcal{L}}{\partial\tau_n}\right)'(B'^{n-1})^{\alpha\beta}\nabla_\alpha\xi_\beta=-\int dx\sqrt{g}\,\xi_\beta\nabla_\alpha\left(\frac{\sqrt{g'}}{\sqrt{g}}\sum_n n\left(\frac{\partial\mathcal{L}}{\partial\tau_n}\right)'(B'^{n-1})^{\alpha\beta}\right)\,.$$

Note that $\nabla$ is the connection obtained from the metric $g$, so $\nabla_\rho g_{\mu\nu}=0$, but $\nabla_\rho g'_{\mu\nu}$ is not zero in general. We recognize that the content of the round bracket is the tensor $T^{\mu\nu}_{(L)}$ consisting of the last term in (2.28), whereas the round bracket in the first line of (A.1), is just the tensor $T^{\mu\nu}_{(R)}$ consisting of the first two terms of (2.28). On the other hand, the expression in square brackets in the third and fourth lines, and in the fifth and sixth lines is nothing but the EOM. Therefore (A.1) is just (2.35) written more explicitly, and where we recognize the separate conservation of $T^{\mu\nu}_{(R)}$ and $T^{\mu\nu}_{(L)}$.

# B Models with different domain and target

In this appendix we discuss models of dynamical diffeomorphisms where, in contrast to the models discussed in the main text, the domain $M_R$ and the target space $M_L$ are viewed as different manifolds, and consequently also their metrics are different. Of course since the manifolds are diffeomorphic one can also view them as "the same manifold", but this presupposes a preferred identification, whereas here we will not assume one, at least not to begin with.

The formalism lends itself to two rather different interpretations. In the "field theoretic" interpretation, $M_R$ is spacetime and $M_L$ is some internal space, endowed with a fixed metric. In the "brane" interpretation, $M_R$ is the brane worldsheet and $M_L$ is spacetime.

## B.1 Left and right diffeomorphisms

We denote $\mathcal{D}iff(M_R, M_L)$ the space of diffeomorphisms of $M_R$ to $M_L$. The diffeomorphisms of $M_R$ and $M_L$ into themselves will be denoted $\mathcal{D}iffM_R$ and $\mathcal{D}iffM_L$ respectively. They act in the usual way on tensors on $M_R$ and $M_L$ respectively. The groups $\mathcal{D}iffM_R$ and $\mathcal{D}iffM_L$ act on $\mathcal{D}iff(M_R, M_L)$ by right and left composition

$$\begin{aligned} \varphi \mapsto \varphi' &= \varphi \circ \psi_R \ , \\ \varphi \mapsto \varphi' &= \psi_L^{-1} \circ \varphi \ . \end{aligned} \tag{B.1}$$

Each of these actions is free and transitive. In particular, $\mathcal{D}iffM_L \times \mathcal{D}iffM_R$ acts transitively on $\mathcal{D}iff(M_R, M_L)$, so we can transform any $\varphi \in \mathcal{D}iff(M_R, M_L)$ to any other $\varphi'$. Now pick some fixed $\bar{\varphi} \in \mathcal{D}iff(M_R, M_L)$. It defines an isomorphism $\iota : \mathcal{D}iffM_R \to \mathcal{D}iffM_L$ by

$$\iota(\psi_R) = \bar{\varphi} \circ \psi_R \circ \bar{\varphi}^{-1} \ . \tag{B.2}$$

The stabilizer of $\bar{\varphi}$ is the "diagonal" subgroup $\Delta\mathcal{D}iff(M_R, M_L)$ consisting of transformations $(\psi_L, \psi_R) \in \mathcal{D}iffM_R \times \mathcal{D}iffM_L$ of the form

$$(\psi_L, \psi_R) = (\iota(\psi), \psi) \ .$$

The diagonal subgroup acts on $\mathcal{D}iff(M_R, M_L)$ as follows: for $f \in \mathcal{D}iffM_R$,

$$f : \varphi \mapsto \iota(f^{-1}) \circ \varphi \circ f = \bar{\varphi} \circ f^{-1} \circ \bar{\varphi}^{-1} \circ \varphi \circ f \tag{B.3}$$

and indeed under such action $\bar{\varphi}$ is invariant

$$f : \bar{\varphi} \mapsto \bar{\varphi} \circ f^{-1} \circ \bar{\varphi}^{-1} \circ \bar{\varphi} \circ f = \bar{\varphi} \ .$$

We have shown that the configuration space $\mathcal{D}iff(M_R, M_L)$ can be regarded as a homogeneous space

$$(\mathcal{D}iffM_R \times \mathcal{D}iffM_L)/\Delta\mathcal{D}iff(M_R, M_L) \ . \tag{B.4}$$

All this bears a striking similarity to chiral models of particle physics, but here the groups are infinite dimensional.

## B.2 Dynamics

Let us denote $g$ and $h$ the metrics in $M_R$ and $M_L$, respectively. The actions we are interested in have the form

$$S = \int_{M_R} dx \, \sqrt{g} \, \mathcal{L}(\sigma_1, \sigma_2, \sigma_3, \sigma_4) \ , \tag{B.5}$$

where $\sigma_n = \mathrm{tr}B^n$ and $B^\mu{}_\nu$ is given by (1.1). The usual action for nonlinear sigma models corresponds to $\mathcal{L} = -\frac{1}{2}\sigma_1$ and one may keep this example in mind in the following.

As long as the metrics $g$ and $h$ are kept fixed, the action has, generically, no symmetries. In the field theoretic interpretation the target space metric $h$ can be interpreted as an infinite set of coupling constants. Symmetries of a theory correspond to transformations that act on the dynamical variables leaving the couplings fixed, so left diffeomorphisms are not symmetries, unless $h$ has some isometries. The proper interpretation of left (target space) diffeomorphisms is as field redefinitions. In the brane interpretation, $h$ represents the spacetime metric and its dynamics, in the quantum theory, comes from the beta functions of the worldsheet quantum field theory. Once again, left (spacetime) diffeomorphisms cannot be interpreted as symmetries. We shall use the word "invariances" for transformations that leave the action invariant, from a mere mathematical viewpoint, and irrespective of their physical interpretation.

With these cautionary remarks in mind, the action is separately invariant under $\mathcal{D}iffM_R$ and $\mathcal{D}iffM_L$, as we shall now see. Under a right-diffeomorphism the metric $g$ is pulled back but the metric $h$ is invariant. The pullback $\varphi^*h$ transforms like an ordinary tensor on $M_R$:

$$\varphi'^*h' = (\varphi \circ \psi_R)^*h = \psi_R^*(\varphi^*h) \ .$$

Since the integrand of the action is a scalar density on $M_R$, the action is $\mathcal{D}iffM_R$-invariant.

On the other hand under a left-diffeomorphism $h$ gets pulled back but $g$ is invariant. The pullback $\varphi^*h$ is invariant, because:

$$\varphi'^*h' = (\psi_L^{-1} \circ \varphi)^*(\psi_L^*h) = \varphi^* \circ \psi_L^{-1*} \circ \psi_L^*h = \varphi^*h \ .$$

Since $g$ is also invariant, the action is trivially $\mathcal{D}iffM_L$-invariant.

The equation of motion of these models has the form (2.21), except that $B$ is given by (1.1) instead of (2.7), i.e. $g_{\alpha\beta}(\varphi)$ is replaced everywhere by $h_{\alpha\beta}(\varphi)$ and the Christoffel symbol appearing in (2.5) is the Christoffel symbol of $h_{\alpha\beta}$ (whereas $\nabla_\mu$ in (2.6) is still constructed with the Christoffel symbols of $g_{\mu\nu}$). The EM tensor defined by (2.27) is equal to (2.28), but the last term is absent. It thus agrees with what we called $T^{\mu\nu}_{(R)}$. Separately, we define the tensor

$$T^{\alpha\beta}_{(L)}(y) = \frac{2}{\sqrt{h(y)}} \frac{\delta S}{\delta h_{\alpha\beta}(y)} \ . \tag{B.6}$$

In the next subsections we consider the infinitesimal versions of these transformations and the differential identities that follow from the invariances of the action.

### B.3 Consequences of right diffeomorphism invariance

A right diffeomorphism acts on points of $M_R$ by $x \mapsto x' = \psi^{-1}(x)$. The infinitesimal version is

$$\delta x^\alpha = -\xi^\alpha(x) \ .$$

The variation of any tensor $T$ on $M_R$ is its Lie derivative $\delta_\xi T = \mathcal{L}_\xi T$. For the metric

$$\delta_\xi g_{\mu\nu} = \xi^\rho \partial_\rho g_{\mu\nu} + g_{\mu\rho}\partial_\nu \xi^\rho + g_{\rho\nu}\partial_\mu \xi^\rho = \nabla_\mu \xi_\nu + \nabla_\nu \xi_\mu \ .$$

The infinitesimal variation of $\varphi$ is

$$\delta_\xi \varphi^\rho(x) = \xi^\lambda \partial_\lambda \varphi^\rho \tag{B.7}$$

and $\delta_\xi h_{\alpha\beta} = 0$. Now varying the pullback we have

$$\begin{aligned}\delta_\xi(\varphi^* h_{\mu\nu}) &= \delta_\xi(\partial_\mu\varphi^\alpha\partial_\nu\varphi^\beta h_{\alpha\beta}(\varphi)) \\ &= \partial_\mu\delta\varphi^\alpha\partial_\nu\varphi^\beta h_{\alpha\beta}(\varphi) + \partial_\mu\varphi^\alpha\partial_\nu\delta\varphi^\beta h_{\alpha\beta}(\varphi) + \partial_\mu\varphi^\alpha\partial_\nu\varphi^\beta\partial_\gamma h_{\alpha\beta}(\varphi)\delta\varphi^\gamma \ .\end{aligned}$$

Inserting the above formulae for the variation and expanding, one arrives after a few steps at

$$\xi^\rho\partial_\rho(\varphi^* h)_{\mu\nu} + (\varphi^* h)_{\mu\rho}\partial_\nu\xi^\rho + (\varphi^* h)_{\rho\nu}\partial_\mu\xi^\rho = \mathcal{L}_\xi(\varphi^* h)_{\mu\nu} \ . \tag{B.8}$$

which just confirms that $\varphi^* h$ transforms as a tensor.

As usual, from the diffeomorphism invariance one can obtain a differential identity. The derivation follows the steps of Appendix A, but with some terms now absent. In the end one obtains

$$0 = \int d^4x\sqrt{g}\left[\xi^\tau\partial_\tau\varphi^\alpha E_\alpha - \xi_\mu\nabla_\nu T^{\mu\nu}_{(R)}\right] \ . \tag{B.9}$$

In contrast to (2.35), since $\partial_\tau\varphi^\alpha$ is nondegenerate, the coefficient of the EOM is always nonzero and therefore we find that the EOM and EM conservation are completely equivalent. This is a consequence of the action of the group being free and transitive.

### B.4 Consequences of left diffeomorphism invariance

A left diffeomorphism acts on points of $M_L$ by $x \mapsto x' = x$ and $\varphi(x) \mapsto \psi^{-1}(\varphi(x))$. The infinitesimal versions are

$$\delta_\xi x^\mu = 0 \ , \quad \delta_\xi\varphi^\alpha(x) = -\xi^\alpha(\varphi(x)) \ .$$

The variation of any tensor $T$ on $M_L$ is its Lie derivative $\delta_\xi T = \mathcal{L}_\xi T$, and tensors on $M_R$ are invariant. For the metric in $M_L$

$$\delta_\xi h_{\alpha\beta} = \xi^\rho\partial_\rho h_{\alpha\beta} + h_{\alpha\rho}\partial_\beta\xi^\rho + h_{\rho\beta}\partial_\alpha\xi^\rho = D_\alpha\xi_\beta + D_\beta\xi_\alpha \ , \tag{B.10}$$

where we used the notation (2.4). Now varying the pullback and using this formula we have

$$\begin{aligned}\delta_\xi(\varphi^* h_{\mu\nu}) &= \delta_\xi(\partial_\mu\varphi^\alpha\partial_\nu\varphi^\beta h_{\alpha\beta}(\varphi)) \\ &= \partial_\mu\delta_\xi\varphi^\alpha\partial_\nu\varphi^\beta h_{\alpha\beta}(\varphi) + \partial_\mu\varphi^\alpha\partial_\nu\delta_\xi\varphi^\beta h_{\alpha\beta}(\varphi) \\ &\quad +\partial_\mu\varphi^\alpha\partial_\nu\varphi^\beta\delta_\xi\varphi^\gamma\partial_\gamma h_{\alpha\beta}(\varphi) + \partial_\mu\varphi^\alpha\partial_\nu\varphi^\beta\delta_\xi h_{\alpha\beta}(\varphi) \\ &= 0 \ .\end{aligned} \tag{B.11}$$

As we have already seen at the level of finite transformations, both $g$ and $\varphi^* h$ are invariant, and therefore the invariance of the action is trivial.

Since the metric $g$ is unaffected by these transformations, no consequence can be derived from $\mathcal{D}iffM_L$-invariance concerning the energy-momentum tensor $T^{\mu\nu}_{(R)}$. Nevertheless, we can obtain another differential identity involving the tensor (B.6).

Since in the action the metric $h$ always appears evaluated at $\varphi(x)$, it is convenient to change integration variable from $x$ to $x' = \varphi(x)$, and write

$$S = \int dx'\sqrt{g'(x')}\mathcal{L}'(x') \ .$$

Since $\mathcal{L}$ is a given function of the trace invariants $\mathcal{L}(x) = F(\tau_1(x), \tau_2(x), \tau_3(x), \tau_4(x))$, the transformed Lagrangian $\mathcal{L}'$ will be the same function of the transformed invariants $\mathcal{L}'(x') = F(\tau'_1(x'), \tau'_2(x'), \tau'_3(x'), \tau'_4(x'))$ and since $\tau'_n(x') = \tau_n(x)$, also $\mathcal{L}'(x') = \mathcal{L}(x)$.

For example, for the Lagrangian (2.10), we can write

$$\mathcal{L}'(x') = -\frac{1}{2}(\varphi_* g^{-1})^{\alpha\beta}(x') h_{\alpha\beta}(x')$$

and therefore

$$T^{\alpha\beta}_{(L)}(y) = -f^2 \frac{\sqrt{g'}}{\sqrt{h}} g'^{\alpha\beta}(y) \ ,$$

where we write $g'^{\alpha\beta}$ for the push-forward of the inverse metric, $(\varphi_* g^{-1})^{\alpha\beta}(x')$. In general

$$T^{\alpha\beta}_{(L)}(y) = 2\frac{\sqrt{g'}}{\sqrt{h}} \sum_{n=1}^{4} n \left(\frac{\partial \mathcal{L}}{\partial \tau_n}\right)' (B'^{n-1})^{\alpha\beta}(y) \ , \tag{B.12}$$

Using the infinitesimal variation (B.10) and the invariance of the action under left diffeomorphisms, one finds that

$$0 = \delta_\xi S = \int dx' \sqrt{h(x')} \left[ -\xi^\alpha(x') E_\alpha(x') + \xi_\beta(x') D_\alpha T^{\alpha\beta}_{(L)}(x') \right] \tag{B.13}$$

and therefore, on shell

$$D_\alpha T^{\alpha\beta}_{(L)} = 0 \ . \tag{B.14}$$

### B.5 Relation to the models in the main text

Even though the manifolds $M_R$ and $M_L$ are, by assumption, diffeomorphic, there will in general be no relation between the respective metrics. Consider the special case in which there exists a diffeomorphism $\bar{\varphi}$ such that:

$$\bar{\varphi}^* h = g \ . \tag{B.15}$$

If $h$ and $g$ have isometries, $\bar{\varphi}$ will not be unique. We disregard this case here. We can use $\bar{\varphi}$ to define a preferred identification of $M_R$ and $M_L$ and, via equation (B.2), a preferred identification of the respective diffeomorphism groups.

Since $\bar{\varphi}$ is a diffeomorphism, without loss of generality we can choose atlases on $M_L$ and $M_R$ to be related by $\bar{\varphi}$. This means that on any chart, if $x^\alpha$ are the coordinates of a point $x$ and $y^\alpha$ are the coordinates of $\bar{\varphi}(x)$, then

$$y^\alpha = x^\alpha \ . \tag{B.16}$$

If we use $\bar{\varphi}$ to identify $M_R$ and $M_L$, we have only one manifold $M$, $\bar{\varphi}$ can be thought of as the identity mapping of $M$ to itself and the action (B.3) becomes conjugation. Since $\bar{\varphi}$ is now a fixed element of the theory, the original invariance under $\mathcal{D}iffM_R \times \mathcal{D}iffM_L$ is broken to $\Delta\mathcal{D}iff(M_R, M_L)$, acting by conjugation. In this way we recover the models of the main text.

## References


[1] L. Amendola and S. Tsujikawa, "Dark Energy," Cambridge University Press (2010).

[2] C. Omero and R. Percacci, "Generalized nonlinear $\sigma$-models in curved space and spontaneous compactification," Nucl. Phys. **B165** 351-364 (1980).

[3] M. Gell-Mann and B. Zwiebach, "Space-time compactification due to scalars," Phys. Lett. B **141** (1984), 333-336.

[4] M. Gell-Mann and B. Zwiebach, "Dimensional Reduction of Space-time Induced by Non-linear Scalar Dynamics and Noncompact Extra Dimensions," Nucl. Phys. B **260** (1985), 569-592.

[5] K. Hinterbichler, "Theoretical Aspects of Massive Gravity," Rev. Mod. Phys. **84** (2012), 671-710 [arXiv:1105.3735 [hep-th]].

[6] E. Stückelberg, "Theory of the radiation of photons of small arbitrary mass," Helv. Phys. Acta **30** (1957), 209-215.

[7] N. Arkani-Hamed, H. Georgi and M. D. Schwartz, "Effective field theory for massive gravitons and gravity in theory space," Annals Phys. **305** (2003), 96-118 [arXiv:hep-th/0210184 [hep-th]].

[8] D. G. Boulware and S. Deser, "Can gravitation have a finite range?," Phys. Rev. D **6** (1972), 3368-3382.

[9] P. Creminelli, A. Nicolis, M. Papucci and E. Trincherini, "Ghosts in massive gravity," JHEP **09** (2005), 003 [arXiv:hep-th/0505147 [hep-th]].

[10] C. de Rham, G. Gabadadze and A. J. Tolley, "Resummation of Massive Gravity," Phys. Rev. Lett. **106** (2011), 231101 [arXiv:1011.1232 [hep-th]].

[11] C. de Rham, G. Gabadadze and A. J. Tolley, "Ghost free Massive Gravity in the Stückelberg language," Phys. Lett. B **711** (2012), 190-195 [arXiv:1107.3820 [hep-th]].

[12] S. F. Hassan and R. A. Rosen, "Resolving the Ghost Problem in non-Linear Massive Gravity," Phys. Rev. Lett. **108** (2012), 041101 [arXiv:1106.3344 [hep-th]].

[13] S. F. Hassan and R. A. Rosen, "On Non-Linear Actions for Massive Gravity," JHEP **07** (2011), 009 [arXiv:1103.6055 [hep-th]].

[14] S. F. Hassan and R. A. Rosen, "Confirmation of the Secondary Constraint and Absence of Ghost in Massive Gravity and Bimetric Gravity," JHEP **04** (2012), 123 [arXiv:1111.2070 [hep-th]].

[15] C. de Rham, A. J. Tolley and S. Y. Zhou, "The $\Lambda_2$ limit of massive gravity," JHEP **04** (2016), 188 [arXiv:1602.03721 [hep-th]].

[16] V. A. Rubakov, "Lorentz-violating graviton masses: Getting around ghosts, low strong coupling scale and VDVZ discontinuity," [arXiv:hep-th/0407104 [hep-th]].

[17] S. Dubovsky, "Phases of massive gravity," JHEP **10** (2004), 076 [arXiv:hep-th/0409124 [hep-th]].

[18] D. Comelli, F. Nesti and L. Pilo, "Massive gravity: a General Analysis," JHEP **07** (2013), 161 [arXiv:1305.0236 [hep-th]].

[19] D. Comelli, F. Nesti and L. Pilo, "Weak Massive Gravity," Phys. Rev. D **87** (2013) no.12, 124021 [arXiv:1302.4447 [hep-th]].

[20] P. Hořava, "Quantum Gravity at a Lifshitz Point," Phys. Rev. D **79** (2009), 084008 [arXiv:0901.3775 [hep-th]].

[21] S. Endlich, A. Nicolis and J. Wang, "Solid Inflation," JCAP **10** (2013), 011 [arXiv:1210.0569 [hep-th]].

[22] D. Comelli, F. Nesti and L. Pilo, "Cosmology in General Massive Gravity Theories," JCAP **05** (2014), 036 [arXiv:1307.8329 [hep-th]].

[23] D. Comelli, F. Nesti and L. Pilo, "Nonderivative Modified Gravity: a Classification," JCAP **11** (2014), 018 [arXiv:1407.4991 [hep-th]].

[24] G. Ballesteros, D. Comelli and L. Pilo, "Massive and modified gravity as self-gravitating media," Phys. Rev. D **94** (2016) no.12, 124023 [arXiv:1603.02956 [hep-th]].

[25] M. Celoria, D. Comelli and L. Pilo, "Fluids, Superfluids and Supersolids: Dynamics and Cosmology of Self Gravitating Media," JCAP **09** (2017), 036 [arXiv:1704.00322 [gr-qc]].

[26] M. Celoria, D. Comelli and L. Pilo, "Self-gravitating Λ-media," JCAP **01** (2019), 057 [arXiv:1712.04827 [gr-qc]].

[27] M. Celoria, D. Comelli, L. Pilo and R. Rollo, "Adiabatic Media Inflation," JCAP **12** (2019), 018 [arXiv:1907.11784 [gr-qc]].

[28] C. Rovelli, "What Is Observable in Classical and Quantum Gravity?," Class. Quant. Grav. **8** (1991), 297-316.

[29] C. Rovelli, "Quantum gravity", Cambridge University Press (2004).

[30] C. G. Torre, "Gravitational observables and local symmetries," Phys. Rev. D **48** (1993), 2373-2376 [arXiv:gr-qc/9306030 [gr-qc]].

[31] H. Bondi and J. Samuel, "The Lense-Thirring effect and Mach's principle," Phys. Lett. A **228** (1997), 121 [arXiv:gr-qc/9607009 [gr-qc]].

[32] J. D. Brown and K. V. Kuchar, "Dust as a standard of space and time in canonical quantum gravity," Phys. Rev. D **51** (1995), 5600-5629 [arXiv:gr-qc/9409001 [gr-qc]].

[33] K. V. Kuchar and C. G. Torre, "Gaussian reference fluid and interpretation of quantum geometrodynamics," Phys. Rev. D **43** (1991), 419-441.

[34] K. Giesel, S. Hofmann, T. Thiemann and O. Winkler, "Manifestly Gauge-Invariant General Relativistic Perturbation Theory. I. Foundations," Class. Quant. Grav. **27** (2010), 055005 [arXiv:0711.0115 [gr-qc]].

[35] J. Tambornino, "Relational Observables in Gravity: a Review," SIGMA **8** (2012), 017 [arXiv:1109.0740 [gr-qc]].

[36] B. S. DeWitt, "The Quantization of geometry" in 'Gravitation: An introduction to current research', John Wiley and Sons (1962).

[37] C. Rovelli, "Quantum Reference Systems," Class. Quant. Grav. **8** (1991), 317-332.

[38] B. Dittrich, "Partial and complete observables for Hamiltonian constrained systems," Gen. Rel. Grav. **39** (2007), 1891-1927 [arXiv:gr-qc/0411013 [gr-qc]].

[39] B. Dittrich, "Partial and complete observables for canonical general relativity," Class. Quant. Grav. **23** (2006), 6155-6184 [arXiv:gr-qc/0507106 [gr-qc]].

[40] H. Westman and S. Sonego, "Coordinates, observables and symmetry in relativity," Annals Phys. **324** (2009), 1585-1611 [arXiv:0711.2651 [gr-qc]].

[41] S. Gielen, "Group field theory and its cosmology in a matter reference frame," Universe **4** (2018) no.10, 103 [arXiv:1808.10469 [gr-qc]].

[42] P. A. Höhn, "Switching Internal Times and a New Perspective on the 'Wave Function of the Universe'," Universe **5** (2019) no.5, 116 [arXiv:1811.00611 [gr-qc]].

[43] P. A. Höhn, A. R. H. Smith and M. P. E. Lock, "The Trinity of Relational Quantum Dynamics," [arXiv:1912.00033 [quant-ph]].

[44] S. B. Giddings, "Hilbert space structure in quantum gravity: an algebraic perspective," JHEP **12** (2015), 099 [arXiv:1503.08207 [hep-th]].

[45] W. Donnelly and S. B. Giddings, "Observables, gravitational dressing, and obstructions to locality and subsystems," Phys. Rev. D **94** (2016) no.10, 104038 [arXiv:1607.01025 [hep-th]].

[46] S. B. Giddings, D. Marolf and J. B. Hartle, "Observables in effective gravity," Phys. Rev. D **74** (2006), 064018 [arXiv:hep-th/0512200 [hep-th]].

[47] T. Thiemann, "Solving the Problem of Time in General Relativity and Cosmology with Phantoms and k-Essence," [arXiv:astro-ph/0607380 [astro-ph]].

[48] J. Frohlich, G. Morchio and F. Strocchi, "Higgs phenomenon without symmetry braking order parameter," Nucl. Phys. B **190** (1981), 553-582.

[49] A. Maas, R. Sondenheimer and P. Törek, "On the observable spectrum of theories with a Brout–Englert–Higgs effect," Annals Phys. **402** (2019), 18-44 [arXiv:1709.07477 [hep-ph]].

[50] A. Maas, "The Fröhlich-Morchio-Strocchi mechanism and quantum gravity," SciPost Phys. **8** (2020) no.4, 051 [arXiv:1908.02140 [hep-th]].

[51] J. Eells and L. Lemaire "A report on harmonic maps", Bull. London Math. Soc. **10** 1-68 (1978).

[52] E. Mazet, "La formule de la variation seconde de l'énergie au voisinage d'une application harmonique" J. Diff. Geom. **8**, 279 (1973).

[53] R.T. Smith, "The second variation formula for harmonic mappings", Proc. Am. Math. Soc. **47**, 229 (1975).

[54] B. Chow and P. Lu "Harmonic maps near the identiy of $S^n$" Diff. Geom. and its Applications **42** 31–36 (2015).

[55] A.M. Polyakov, A.A. Belavin, "Metastable States of Two-Dimensional Isotropic Ferromagnets", JETP Lett. 22 245-248 (1975) Pisma Zh. Eksp. Teor. Fiz. 22 503-506 (1975).

[56] C. de Rham, A. J. Tolley and S. Y. Zhou, "Non-compact nonlinear sigma models," Phys. Lett. B **760** (2016), 579-583 [arXiv:1512.06838 [hep-th]].

[57] M. Capoferri and D. Vassiliev, "Spacetime diffeomorphisms as matter fields," [arXiv:1805.01303 [math-ph]].

[58] H. Leutwyler, "Phonons as goldstone bosons," Helv. Phys. Acta **70** (1997), 275-286 [arXiv:hep-ph/9609466 [hep-ph]].

[59] S. Endlich, A. Nicolis, R. Rattazzi and J. Wang, "The Quantum mechanics of perfect fluids," JHEP **04** (2011), 102 [arXiv:1011.6396 [hep-th]].

[60] S. Dubovsky, L. Hui, A. Nicolis and D. T. Son, "Effective field theory for hydrodynamics: thermodynamics, and the derivative expansion," Phys. Rev. D **85** (2012), 085029 [arXiv:1107.0731 [hep-th]].

[61] G. Ballesteros and B. Bellazzini, "Effective perfect fluids in cosmology," JCAP **04** (2013), 001 [arXiv:1210.1561 [hep-th]].