General Covariance, Diffeomorphism Invariance, and Background Independence in 5 Dimensions

Antonio Vassallo

University of Lausanne, Department of Philosophy, CH-1015 Lausanne
antonio.vassallo@unil.ch

Abstract

The paper considers the "GR-desideratum", that is, the way general relativity implements general covariance, diffeomorphism invariance, and background independence. Two cases are discussed where 5-dimensional generalizations of general relativity run into interpretational troubles when the GR-desideratum is forced upon them. It is shown how the conceptual problems dissolve when such a desideratum is relaxed. In the end, it is suggested that a similar strategy might mitigate some major issues such as the problem of time or the embedding of quantum non-locality into relativistic spacetimes.

Keywords: General relativity, general covariance, diffeomorphism invariance, background independence, Kaluza-Klein theory, induced matter theory, problem of matter.

Notation: In the following, Einstein's convention will apply. Lower-case indices will be taken to range from 0 to 3, while upper-case ones will range from 0 to 4. A semicolon before one or more indices will indicate covariant differentiation with respect to those indices, a comma will instead indicate ordinary differentiation. Moreover, all equations will be written in natural units such that c=G=1.

1

1 Introduction

General relativity (GR) is the most corroborated theory of gravitation we have so far. It is also considered a spacetime theory because it is taken to unify gravitational phenomena and physical geometry.[1] The dynamics of GR is encoded in the Einstein's field equations:[2]

$$\mathbf{G}[\mathbf{g}] = 8\pi\mathbf{T}[\mathbf{g}, \boldsymbol{\phi}]. \tag{1}$$

The left-hand side of the equation is the Einstein tensor defined over a 4-dimensional differentiable semi-Riemannian manifold M_4: since it depends on the 4-metric tensor $\mathbf{g}$, it is taken to encode information on the 4-geometry of physical spacetime. The right-hand side comprises the stress-energy tensor, which encodes information such as the energy-momentum density of some matter field $\boldsymbol{\phi}$ distributed over (a region of) spacetime. A solution (or model) of (1) is then a triple $< M_4, \mathbf{g}, \mathbf{T} >$ that represents a physical scenario where the geometry of spacetime (aka the gravitational field) interacts with matter fields distributed over it. In short: «Space acts on matter, telling it how to move. In turn, matter reacts back on space, telling it how to curve.» (Misner et al., 1973, p.5).

Equations (1) are formulated in terms of geometrical field-theoretic objects, more precisely tensor fields, defined over a 4-manifold M_4. In order to make them simpler to handle for sake of calculation, it is possible to introduce some coordinate system $\{x^i\}$ over a neighborhood U of M_4, and then rewrite them in terms of components of the geometrical objects in that coordinate system:

$$G_{ij}[g_{ij}] = 8\pi T_{ij}. \tag{2}$$

Unlike (1), the equations compactly written as (2) involve derivatives and symmetric matrices, which simplify (but not *over*-simplify) the work of the physicists dealing with calculations. The important fact is that, by construction, the physical information encoded in the relations between field-theoretic objects as described by (1) remain unchanged when switching to (2), independently of the particular coordinatization chosen. For example, for two coordinate systems $\{x^i\}$ and $\{\tilde{x}^i\}$ and two associated bases $\{e^i\}$ and $\{\tilde{e}^i\}$ on a neighborhood of a point $P \in M_4$, we have:

$$\mathbf{g} = g_{ij}|_P e^i \otimes e^j = \tilde{g}_{ij}|_P \tilde{e}^i \otimes \tilde{e}^j, \tag{3}$$

the (invertible) transformation law relating g_{ij} and $\tilde{g}_{ij}$, being:

$$g_{ij} \rightarrow \tilde{g}_{ij} = \frac{\partial x^h}{\partial \tilde{x}^i} \frac{\partial x^k}{\partial \tilde{x}^j} g_{hk}. \tag{4}$$

---

[1] I will not even try to address here questions regarding the meaning (if any) of the unification involved. The interested reader can take a look, for example, at Lehmkuhl (2008, 2014).

[2] I choose to disregard the cosmological constant, since it does not affect the analysis carried out in this paper.

Since we know from elementary tensor calculus that (3) holds for whatever tensorial object, it follows that it is possible to recover (1) from (2) for whatever coordinate system $\{x^i\}$ defined on a neighborhood of a point in the manifold.

Further, it can be shown that to any (sufficiently smooth) coordinate transformation $\{x^i\} \rightarrow \{\tilde{x}^i\}$ defined in (a neighborhood of) M_4 corresponds a self-diffeomorphism[3] f such that, for each point P in (a neighborhood of) M_4, it is the case that $x^i(f(P)) = \tilde{x}^i(P)$.[4]

In the following, we will take advantage of the duality between the geometrical or intrinsic formulation of GR (involving geometrical objects and diffeomorphisms) and its local formulation (involving components of geometrical objects and change of coordinate systems), and we will switch back and forth from the intrinsic to the coordinates-related language for the sake of simplicity, according to the specific circumstances considered.

Nowadays, GR is so widely and firmly accepted by the scientific community to be considered one of the pillars of modern physics. At the root of such a vast agreement lies not only the empirical success of the theory, but also its numerous applications to, just to mention two major fields, astrophysics and cosmology. In the words of Misner, Thorne, and Wheeler:

> Einstein's theory attracts the interest of many today because it is rich in applications. No longer is the attention confined to three famous but meager tests: the gravitational redshift, the bending of light by the sun, and the precession of the perihelion of Mercury around the sun. The combination of radar ranging and general relativity is, step by step, transforming the solar-system celestial mechanics of an older generation to a new subject, with a new level of precision, new kinds of effects, and a new outlook. Pulsars, discovered in 1968, find no acceptable explanation except as the neutron stars predicted in 1934, objects with a central density so high ($\sim 10^{14} g/cm^3$) that the Einstein predictions of mass differ from the Newtonian predictions by 10 to 100 per cent. About further density increase and a final continued gravitational collapse, Newtonian theory is silent. In contrast, [GR] predicted [...] the properties of a completely collapsed object, a

---

[3] A self-diffeomorphism (from now on called simply diffeomorphism), is a mapping $f : M_4 \rightarrow M_4$ that is bijective, continuous, and differentiable together with its inverse f^{-1}.

[4] Suggestively speaking, while in the "coordinate" case we "keep P fixed" and we evaluate it (in the sense of associating a 4-tuple of real numbers to it) under a coordinate system *different* from the starting one, in the "diffeomorphic" case we "move P around the manifold" and then we evaluate it under the *same* coordinate system. For the reader unsatisfied by this, indeed, very rough characterization, Norton (2011) has a section that nicely explains the difference between extrinsic (or passive) transformations expressed in terms of coordinate systems, and intrinsic (or active) transformations expressed in terms of diffeomorphisms.

"frozen star" or "black hole".
(Misner et al., 1973, pp. viii,ix)

However, the success of GR is not only due to its empirical adequacy and its richness in applications. Many physicists (and mathematicians) would in fact add that GR is an extremely elegant and mathematically beautiful theory. This further virtue of the theory should not be regarded as merely aesthetic, but it should be taken as a hint of the fact that GR points at some deep physical "truth" about the world. Dirac addresses this view as follows:

> Let us now face the question, suppose a discrepancy had appeared, well confirmed and substantiated, between the theory and observations. [...] Should one then consider the theory to be basically wrong? [...] I would say that the answer to the [q]uestion is emphatically NO. The Einstein theory of gravitation has a character of excellence of its own. Anyone who appreciates the fundamental harmony connecting the way Nature runs and general mathematical principles must feel that a theory with the beauty and elegance of Einstein's theory *has* to be substantially correct. If a discrepancy should appear in some applications of the theory, it must be caused by some secondary feature relating to this application which has not been adequately taken into account, and not by a failure of the general principles of the theory.
> (Dirac, 1978, pp. 6,7 (Dirac's emphasis))

The questions about how, in general, a physical theory becomes accepted by the community, and what features should be privileged over others when assessing the merits of such a theory (e.g. empirical over formal), are beyond the scopes of this essay (see, for example, Brush, 1999 for a historical reconstruction and a philosophical discussion of the reasons that led to the acceptance of Einstein's theory). Here we will just make the working hypothesis that there is "something" physical *and* formal about GR (what Dirac calls the «general principles of the theory») that, according to mainstream physicists, places a constraint on future generalizations of the theory (to include, for example, quantum effects). In the remainder of the section, we will briefly consider some candidates for this "something", and we will finally propose a tentative definition that captures the «character of excellence» of GR that should therefore be preserved in any decent generalization of the theory.

To start our inquiry, let us consider the most straightforward candidate to be such "something": general covariance. With the designation "general covariance" here we will intend the formal invariance of the dynamical equations (2) under an arbitrary coordinate transformation. Einstein originally thought that general covariance so intended was some sort of generalization of the principle of relativity

that holds in Newtonian mechanics and special relativity. Very simply speaking, while in these two latter theories the description of the dynamics of a physical system remained unchanged under coordinate transformations between inertial reference frames, so that these frames where considered physically equivalent,[5] in GR such a description is unchanged under whatever coordinate transformation, so one might claim that *all* reference frames are physically equivalent. Can the "virtue" of GR reside in its general covariance simpliciter?

The answer, as firstly pointed out by Kretschmann (1917), is no since it is possible to render generally covariant also a theory that patently accords a privileged status to inertial reference systems. Hence, general covariance cannot be taken as the implementation of the physical requirement that all reference frames are physically equivalent. To understand why it is so, consider a theory of a massless Klein-Gordon field ϕ over a manifold endowed with a Minkowski metric $\boldsymbol{\eta}$. We can write the field equations of the theory as:

$$\eta^{ij}\phi_{,ij} = 0. \tag{5}$$

Equation (5) is not generally covariant because it holds only in inertial coordinate systems. In order to make it generally covariant we just have to make the purely formal move of rewriting it in a way such that it preserves its form in whatever coordinate system *and* reduces to (5) in those coordinate system that are inertial. This means that we have to write $\boldsymbol{\eta}$ in a generalized form $\mathbf{g}$ and introduce the (unique) compatible covariant derivative operator $\boldsymbol{\nabla}$,[6] so that (5) finally reads:

$$g^{ij}\phi_{;ij} = 0. \tag{6}$$

It is easy to understand that (6) is invariant in form under whatever coordinate transformation but, still, the theory - being specially relativistic - accords to inertial frames of reference a privileged status. In the language of geometrical objects, (6) is written

$$\Box_{\mathbf{g}}\phi = 0, \tag{7}$$

and, as expected, it is invariant in form (that is, covariant) under diffeomorphic transformations. Since such a move can be performed for whatever spacetime theory - even Newtonian mechanics - then (i) general covariance per se has no physical import and (ii) the generally covariant nature of GR cannot be linked to a peculiar feature such as a generalized principle of relativity. If Kretschmann's point is correct, then what is it that renders GR such a peculiar and successful theory?

---

[5]The question of the equivalence between coordinate systems and frames of reference is another extremely delicate matter I will not touch upon here. See, e.g., Norton (1993, section 6.3).

[6]"Compatible" means that $\boldsymbol{\nabla}\mathbf{g} = 0$.

The most simple reaction would be to claim that the peculiarity of GR does not reside just in the mere formal invariance of its field equations, but also in the physical import of the solutions space of such equations. If, in fact, we consider two solutions of (6)/(7) related by a diffeomorphism f, they will represent one and the same physical situation if and only if f is an isometry[7] of the Minkowski metric. Not surprisingly, the group of isometries $iso(\boldsymbol{\eta})$ of the Minkowski metric is just the Poincaré group. In the case of equation (1), instead, since the metric $\boldsymbol{g}$ is not fixed a priori but it is subject to the dynamical evolution, the group $diff(M_4)$ is unrestricted and, as expected, whatever two solutions of (1) related by a diffeomorphism are physically indistinguishable.[8]

The above reply points at the fact that what makes the fortune of GR is not just its general covariance, but its diffeomorphism invariance: in virtue of having a dynamical geometry, GR has no fixed isometries different from the whole $diff(M_4)$ and, hence, diffeomorphisms are physical symmetries of the theory. Indeed, $diff(M_4)$ can be considered the *gauge group* of GR.

The question whether GR can be really considered a gauge theory in the sense commonly intended in particle physics is complex and highly debated (to have a substantive example of such debate, see Earman, 2006; Pooley, 2010); here we will just sketch a counter-reply to the above argument.

If the real difference between GR and previous spacetime theories is the dynamical status of spacetime's geometry, then we can just straightforwardly rewrite whatever spacetime theory by adding dynamical constraints on the spatiotemporal structures, so that such a difference can be obliterated. To see this, let us take the above theory of a Klein-Gordon field over a Minkowski spacetime and rewrite the dynamical equations as:

$$\mathbf{Riem[g]} = 0, \tag{8a}$$

$$\Box_g \phi = 0, \tag{8b}$$

$\mathbf{Riem[g]}$ being the Riemann curvature tensor. The Minkowski metric, now, is a dynamical object picked up as a solution of (8a). The theory whose dynamics

---

[7]Roughly speaking, $f \in diff(M_4)$ is an isometry for a geometrical object $\boldsymbol{\theta}$ if the application of such a transformation to the object does not change it: $f^*\boldsymbol{\theta} = \boldsymbol{\theta}$.

[8]Just to have a more concrete idea of what this means, consider a solution $\mathfrak{M}$ of (1) where we can construct a physical observable interpreted as a 4-distance function $d_{\mathbf{g}} : M_4 \times M_4 \rightarrow \mathbb{R}$ defined in terms of the metric $\mathbf{g}$. Let us now take a diffeomorphically transformed model $f^*\mathfrak{M}$, where $f \in diff(M_4)$; in such a model we can construct an observable diffeomorphically related to $d_{\mathbf{g}}$, namely a function $d_{f^*\mathbf{g}}$ defined by the diffeomorphic metric $f^*\mathbf{g}$. It is now easy to see that, taken two arbitrary points $P, Q \in M_4$, we have $d_{\mathbf{g}}(P, Q) = d_{f^*\mathbf{g}}(f(P), f(Q))$. It follows that no 4-distance measurement can discriminate $\mathfrak{M}$ from $f^*\mathfrak{M}$. By repeating the same reasoning for all the observables constructible in $\mathfrak{M}$ and $f^*\mathfrak{M}$, we can conclude either that the two models represent different physical possibilities which are empirically indistinguishable - and then we would immediately fall pray of arguments from indeterminism, such as the infamous "hole argument", cf. Earman and Norton (1987) - or that the two models represent one and the same physical scenario.

is encoded in (8) has the feature that two models related by whatever diffeomorphism represent in fact one and the same physical situation.

The immediate reaction to such an example is that it exploits a mere mathematical trick to render the transformations in $diff(M_4) \setminus iso(\boldsymbol{\eta})$ just *mathematical* symmetries of the theory and then mix them up to the *physical* symmetries in $iso(\boldsymbol{\eta})$ in order to patch them together into a fake gauge group. In this particular case, a possible implementation of this counter-argument is, following Anderson (1967); Friedman (1983), to point out that some geometrical objects have been artificially rendered dynamical and, thus, the theory can be deparametrized by quotienting out the redundant structures in the solutions space introduced by adding the fake dynamical sector.[9] Such a strategy makes sense in the present case: if we deparametrize (8) by quotienting out the solution space by $diff(M_4) \setminus iso(\boldsymbol{\eta})$, we end up again with (7). If this reasoning could be carried out for all spacetime theories prior to GR, we would then have a reasonable candidate to be the peculiar feature of this latter theory, namely, the absence of spacetime structures that are not genuinely dynamical. Some authors would dub such a feature "background independence".

However, the above reasoning cannot be the final word in the debate. The present case per se is, in fact, too artificial to represent a really difficult challenge. Here the Minkowski metric represents an elephant in the room in the sense that we are perfectly able to see that such a structure "persists" unaltered in all models of the theory modulo a diffeomorphism transformation.[10] In general, however, making the distinction between genuine dynamical and pseudo-dynamical structures varies from difficult to impossible so that defining background independence in an accurate manner is a hard task (Giulini, 2007; Belot, 2011, discuss at length the problem and try to propose some alternative solutions).

So, in the end, what is it that confers on GR a «character of excellence of its own»? Let us try to filter from the above discussion the salient uncontroversial claims regarding the features of GR, and come up with the following tentative characterization: Unless its predecessors, GR is a $diff(M_4)$-gauge theory whose only non-dynamical constraint is to be formulated over some 4-dimensional semi-Riemannian manifold, and that does not postulate more spatiotemporal structure than just that encoded in a Lorentzian 4-metric. If this is (or gets close to) what is at the root of the major success of GR, then we can consider it as a (minimal) constraint on any future (quantum) spacetime theory that seek to supersede GR:

---

[9]Just to have a rough idea of what adding redundant structure means when passing from (7) to (8), it is sufficient to point out that to each solution $< M_4, \mathbf{g}, \phi >$ of (7) corresponds a set of solutions $\{< M_4, \mathbf{g}, \phi >, < M_4, f^*\mathbf{g}, f^*\phi >, < M_4, h^*\mathbf{g}, h^*\phi >, \dots\}$ of (8) with $f, h, \dots \in diff(M_4) \setminus iso(\boldsymbol{\eta})$.

[10]Which means that, for whatever two models $< M_4, \mathbf{g}, \phi >$ and $< M_4, \mathbf{g}', \phi' >$, there is *always* some $f \in diff(M_4)$ such that $\mathbf{g}' = f^*\mathbf{g}$.

GR-desideratum. *Any theory which is a valid candidate for superseding GR must be [become in the classical limit] a $diff(M_d)$-gauge theory whose only non-dynamical constraint is to be formulated over some d-dimensional semi-Riemannian manifold and that does not postulate more spatiotemporal structure than just that encoded in a d-metric.*[11]

Note that, in general, besides arguments from physics, the GR-desideratum can be backed up by metaphysical arguments. One might for example argue that (1) is preferable to (8) because, in this latter theory, there is a "lack of reciprocity" between the Minkowski metric and the scalar field: in (8) spacetime tells matter how to move, but matter does not tell spacetime how to curve.

In the following, I will show how a too strict pursuit of the GR-desideratum might negatively affect the interpretation of prima facie genuine generalizations of GR.

2   5-Dimensional Extensions of General Relativity

The Kaluza-Klein approach, so called because, historically, the first attempts in this sense were made by Theodor Kaluza (1921) and, slightly later, by Oskar Klein (1926), started as a theoretical program that sought to unify gravitational and electromagnetic forces as curvature effects of a 5-dimensional semi-Riemannian manifold. This is achieved by considering the 5-dimensional vacuum Einstein's equations:[12]

$$R_{AB}[g_{AB}] = 0, \tag{9}$$

with the 5-Ricci tensor R_{AB} depending on a metric of the form:

$$g_{AB} = \begin{pmatrix} g_{ij} + \kappa^2\phi^2 A_i A_j & \kappa^2\phi^2 A_i \\ \kappa^2\phi^2 A_j & \phi^2 \end{pmatrix}, \tag{10}$$

where **A** is the 4-vector potential, ϕ is a scalar field, and κ is a free parameter. The signature of the 5-metric will be assumed to be Lorentzian, thus identifying the extra fifth dimension as space-like.

The original Kaluza-Klein theory placed two very strong constraints on the fifth dimension, namely, (i) that all partial derivatives with respect to the fifth coordinate are zero (cylinder condition), and (ii) that the fifth dimension has a closed short-scale topology (compactification condition).[13] The most important consequences of these conditions are that there is no change in 4-dimensional physical

---

[11]The g_{ij} part of such a metric having Lorentzian signature.

[12]In the 4-dimensional case, if we take (1) and we set the stress energy tensor equal to zero (that is, no matter is present), we obtain that the equations reduce to **Ric**[**g**] = 0, **Ric** being the so-called Ricci tensor. This is why they are in general said to be "vacuum" field equations.

[13]To be fair to historians of physics, (i) was proposed by Kaluza, while (ii) was added later by Klein, who further showed that, in fact, (ii) implies (i).

quantities that can be ascribed to the presence of an extra spatial dimension and that such a fifth dimension is unobservable at low energies. Condition (ii) was also a vital ingredient in the attempt to explain the quantization of electric charge.[14] En passant, it is interesting to note that condition (ii) obviously prevents whatever macroscopic object from spanning the fifth dimension.[15]

If we set $\phi = 1$ and $\kappa = 4\sqrt{\pi}$ in (10), and we substitute it in (9), the ij-components of the field equations become:

$$G_{ij} = 8\pi T_{ij}^{EM}, \tag{11}$$

with G_{ij} the Einstein 4-tensor and $T_{ij}^{EM} = \frac{1}{4}g_{ij}F_{kl}F^{kl} - F_i{}^k F_{jk}$ the electromagnetic stress-energy 4-tensor of standard GR. The i4-components, instead, become:

$$F_{ij}{}^{;i} = 0, \tag{12}$$

with $F_{ij} = A_{j,i} - A_{i,j}$ the Faraday 4-tensor. In short, (9) collapse into Einstein's field equations of GR and the (source free) Maxwell's equations of electromagnetism coupled together. This is the so called "Kaluza-Klein miracle" and represents a remarkable result in that it "geometrizes away" the electromagnetic field as GR does with gravity, thus suggesting that both gravity and electromagnetism are just a manifestation of the geometry of a 5-dimensional spacetime. Is this a right suggestion? Let us consider a reason why it does not seem so.

In order to see what is fishy about the interpretation of Kaluza-Klein theory as a genuine 5-dimensional theory, we just have to notice that conditions (i) and (ii) restrict the possible topologies of the 5-manifold to $M_4 \times S^1$, with M_4 whatever topology compatible with the ij-part of the metric tensor (10) and S^1 the topology of the circle. This hints at the fact that, as long as we restrict our attention to the ij-part of the theory, we are dealing with standard GR and, hence, whatever coordinate transformation $x^i \rightarrow \tilde{x}^i = f(x^i)$ we apply to the metric (10), we always end up with the same equations (11) *and* we do not change the physical information originally encoded in the 4-metric before the transformation. What about the components of the metric depending on the fifth coordinate? In this case, the freedom in choosing the coordinate transformations that leave (12) invariant in form and do not alter the physical information encoded in (10) is severely limited. In fact, the only possible choice that we can make is $x^4 \rightarrow \tilde{x}^4 = x^4 + f(x^i)$. To realize why it is so, we have

---

[14]Attempt which experiments proved unsuccessful: According to Kaluza-Klein theory, the electron had a mass of twenty-two orders of magnitude higher then the measured one.

[15]The reader is invited to think about the strange consequences for spatial 3-dimensional beings like us of living in an uncompactified 4-space. For example, if we were allowed to move in an extra spatial dimension, we could "switch" the left and right sides of our body without altering the up-down and front-back orientations. Moreover, from the perspective of the extra dimension, the interior of our bodies would be exposed (our skin represents just a 3-dimensional boundary).

just to plug this transformation in the 5-dimensional extension of the law (4) for the transformation of the metric tensor: it is easy to see that such a transformation induces just one change in the metric (10), namely, $A_i \rightarrow \tilde{A}_i = A_i + \frac{\partial f(x^i)}{\partial x^i}$, which obviously does not alter the electromagnetic part of the theory. In short, the gauge group of the theory is not $diff(M_5)$ but $diff(M_4) \times U(1)$.

If, now, we claim that a genuine 5-dimensional extension of GR should admit $diff(M_5)$ as gauge group, we cannot but reach the conclusion that Kaluza-Klein theory is a fake 5-dimensional theory in that it is a version of 4-dimensional GR that camouflages the gauge group of electromagnetism by "spatializing" the symmetries of the theory using an additional space-like dimension: there is no substantial unification here, just a mere algebraic play.[16] Note that the above reasoning is based on the simple fact that the conditions of cilindricity and compactification are fixed a priori, i.e., they are not part of the dynamics of the theory. Simply speaking, the theory is not really dealing with solutions of (9) but with a small subset of it, namely Ricci-flat spacetimes with topology $M_4 \times S^1$ where, additionally, S^1 has to "appear" at short scales. This situation is nothing but a more complex instance of that occurring with the theory (6)/(7). It is then clear that we can in principle modify the field equations (9) so that they can encode cilindricity and compactification conditions[17] in order to have a fully $diff(M_5)$-invariant theory, exactly as (8) is a $diff(M_4)$-invariant extension of (6)/(7).

Such a move is exactly as controversial as the one discussed in the previous section, but can it be charged of hiding an elephant in the room as it was the case with the Minkowski metric in (8)? Well, if we hold the firm commitment to the GR-desideratum, then we can clearly hear the $M_4 \times S^1$ structure trumpeting beneath the dynamics. Notice, however, that in this case the GR-desideratum cannot be backed up with compelling metaphysical arguments as, for example, the "reciprocity" one mentioned at the end of the previous section. Here we are dealing with an empty spacetime, so there is nothing acting on anything without being affected in return. Hence it would not be that weird to claim that the dynamically extended version of Kaluza-Klein theory (9) just depicts spacetimes that are, *as a bare matter of fact*, more structured then a generic (empty) general relativistic spacetime, in that it adds to such generic spacetime an extra-dimensional tiny ring for each point. The counter-reply would be that, in this case, such a theory would be unable to explain *why* spacetime has to have such a structure. This is fair enough, but any theory has to start from some assumptions in order to be developed: even standard GR does not explain why spacetime has to be 4-dimensional and Lorentzian; it just takes as a bare matter of fact that special relativity holds at

---

[16]This line of reasoning can be found, for example, in Maudlin (1989, p.87).

[17]For example, by adding some conditions that pick up at each point of 4-spacetime a fifth-dimensional-like vector field whose integral curves are circles with small radii.

10

quasi-point-sized regions. The whole discussion, then, seems to turn into meta-physical tastes about what has to be counted as a bare fact and what has not. We could, for example, endorse some variant of the regularity account of laws where the dynamical laws of Kaluza-Klein theory just supervene on a pure spatiotemporal Humean mosaic of local matters of particular fact. In such a context, asking why the Kaluza-Klein 5-spacetime has this structure is a misleading question: it just happen to show a pattern of regularities that is best described by the equations of the theory.

Let us recap: if we insists on the GR-desideratum, then we are driven to judge Kaluza-Klein theory as a mathematical trompe-l'œil that spatializes the symmetries of the electromagnetic theory, so it cannot count as a genuine unification of gravity and electromagnetism. However, such a judgment seems too harsh and does not involve in any way the empirical adequacy of the theory (whose failure is the only uncontroversial argument that historically led to the dismissal of Kaluza-Klein theory). It is in fact possible to make perfect physical and metaphysical sense of Kaluza-Klein theory as a unified theory by relaxing the GR-desideratum. The moral to draw from the discussion so far is that, all things considered, perhaps the GR-desideratum is not that strict a desideratum: we can relax such a requirement without necessarily falling into patently artificial theories such as (6)/(7). But let us go ahead in the analysis, and show a case where the GR-desideratum is not only too rigid, but has also metaphysically odd consequences.

Since the strongest source of resistance to the idea that Kaluza-Klein theory represents a genuine 5-dimensional theory lies in the conditions of cilindricity and compactification, let us now consider an implementation of the Kaluza-Klein program[18] that dispenses with such constraints. This is the case of induced matter theory, especially known in its "space-time-matter" version as firstly put forward by Wesson (1984). This theory (or, better, class of theories) rests on the mathematical result that any analytic N-dimensional [semi-] Riemannian manifold can be locally embedded in a $(N+1)$-dimensional Ricci-flat [semi-] Riemannian manifold (Campbell-Magaard theorem, see for example Romero et al., 1996); hence, the field equations of the theory are, again, (9).

If we now take a 5-metric in its diagonal form:

$$
g_{AB} = \begin{pmatrix} & & & & 0 \\ & & & & 0 \\ & g_{ij} & & & 0 \\ & & & & 0 \\ 0 & 0 & 0 & 0 & g_{44} \end{pmatrix}.
\tag{13}
$$

and we substitute it in (9), we obtain again the 4-dimensional Einstein's field equations plus an expression for the 4-stress-energy tensor of the form (see Wesson and

---

[18]See Overduin and Wesson, 1997 for an extensive survey of Kaluza-Klein gravity.

Ponce de Leon, 1992, for the detailed calculations):

$$T_{ij} = \frac{(\phi_{,i})_{;j}}{\phi} - \frac{1}{2\phi^2}\left\{\frac{\phi_{,4}g_{ij,4}}{\phi} - g_{ij,44} + g^{kl}g_{ik,4}g_{jl,4} - \frac{g^{kl}g_{kl,4}g_{ij,4}}{2} + \frac{g_{ij}}{4}\left[g^{kl}_{,4}g_{kl,4} + (g^{kl}g_{kl,4})^2\right]\right\},$$

(14)

where, for notational simplicity, it is assumed $g_{44} = \phi^2$.

From these constructions it follows that, whenever we consider a 4-hypersurface Σ_4 by fixing $x^4 = const.$, we obtain a 4-metric g_{ij} and a stress-energy tensor T_{ij} both well-defined on Σ_4.

Just to have a rough idea of how this works, let us consider an example taken from Ponce de Leon (1988) and, in order to simplify the notation, let us follow the author in naming the coordinates as follows: $x^0 = t, x^1 = x, x^2 = y, x^3 = z, x^4 = \psi$. In case of a homogeneous and isotropic 5-dimensional universe, a class of 5-line elements (parametrized by $\alpha \in \mathbb{R}_{\neq 0}$) correspondent to 5-metrics which are solutions to (9) is:

$$ds^2 = \psi^2 dt^2 - t^{\frac{2}{\alpha}}\psi^{\frac{2}{1-\alpha}}(dx^2 + dy^2 + dz^2) - \frac{\alpha^2}{(1-\alpha)^2}t^2 d\psi^2.$$

(15)

It is easy to see that, if we restrict to 4-hypersurfaces ($\psi = const., d\psi = 0$), (15) is reduced to:

$$ds^2 = dt^2 - R_\alpha^2(t)(dx^2 + dy^2 + dz^2),$$

(16)

which is just the family of Friedman-Lemaître-Robertson-Walker metrics in 4-dimensions corresponding to flat 3-geometries. Moreover, if we substitute (15) in (14), we recognize that the stress-energy tensor resembles that of a perfect fluid with density ρ and pressure p given by:

$$\rho = \frac{3}{8\pi(\alpha\psi t)^2},$$

(17a)

$$p = \left(\frac{2\alpha}{3} - 1\right)\rho.$$

(17b)

Hence, by fixing α and ψ we obtain different possible FLRW 4-models, with state equation for matter given by (17). For example, for $\alpha = \frac{3}{2}$, we would have a 4-dimensional universe filled with dust, while for $\alpha = 2$ we would have one filled with radiation.

The moral to be drawn from the above facts is that matter and energy are 4-dimensional physical properties of spacetime "induced" by the 5-metric tensor. Therefore, the cosmological solutions of (9) all agree in giving us a picture of a 5-dimensional empty universe whose restriction to a 4-surface is a spacetime curved by matter.

In conclusion, it seems that now we have a truly 5-dimensional generalization of GR: the theory is $diff(M_5)$-invariant and there are neither background geometrical objects camouflaged as dynamical, nor bizarre restrictions on the topology of 5-spacetime. However, such a complete fulfillment of the GR-desideratum does not come for free.

To see what is the price to be paid, consider, for example (Wesson, 1995), a 5-dimensional Minkowski spacetime. The correspondent line element in polar coordinates reads:

$$ds^2 = dt^2 - dr^2 - r^2 d\Omega^2 - d\psi^2, \tag{18}$$

whose sections $\psi = const.$ are 4-dimensional Minkowski spacetimes.

However, if we consider the transformation:

$$t' = t, \quad r' = \frac{r}{\psi}\Big(1 + \frac{r^2}{\psi^2}\Big)^{-\frac{1}{2}}, \quad \psi' = \psi\Big(1 + \frac{r^2}{\psi^2}\Big)^{\frac{1}{2}}, \tag{19}$$

then (18) becomes

$$ds^2 = dt'^2 - \psi'^2\Big(\frac{dr'^2}{1 - r'^2} + r'^2 d\Omega^2\Big) - d\psi'^2, \tag{20}$$

whose sections $\psi' = const.$ are FLRW spacetimes.

Using the intrinsic language, we can think that (18) belongs to a 5-model $\mathfrak{M} = < M_5, g_{AB} >$, and that the transformation (19) is dual to a 5-diffeomorphism f, such that the metric (20) belongs to the diffeomorphic model $f^*\mathfrak{M} = < M_5, f^*g_{AB} >$. We therefore have two 5-models which are diffeomorphically equivalent, but that induce 4-models, which are empty in the former case and non-empty in the latter. What is the physical meaning of this fact? None, and that is exactly the problem. If, in fact, we claim that f is a gauge transformation, then we cannot but accept that $\mathfrak{M}$ and $f^*\mathfrak{M}$ are physically indistinguishable, which means that (18) and (20) carry exactly the same physical information, namely, that related to the 5-geometry of spacetime. The fact that we are free to convey such an information by writing the line element either in the form (18) or (20) means that such freedom has no physical consequences: this is exactly what physicists refer to as "gauge freedom". The puzzling consequence is that it makes no physical difference whether we take the 5-spacetime under consideration as a "pile" of empty Minkowski 4-spacetimes or matter-filled FLRW 4-spacetimes, since, in the end, the only thing that counts for the theory is the 5-geometry of spacetime. Crudely speaking, the theory is "blind" to whatever change in 4-spacetimes (as long as their piling up leads to the same 5-spacetime), and if we insist that $\mathfrak{M}$ and $f^*\mathfrak{M}$ are ontologically distinct albeit physically indistinguishable, we commit the theory to indeterminism, as already mentioned in footnote 8. For let us assume for the sake of argument that the dynamics of the theory under scrutiny can be cast

13

in a (4+1) fashion, that is, that a 5-geometry can be recovered by considering a generic foliation of 4-hypersurfaces, specifying some initial data on one of such surfaces, and then evolving these data on the subsequent leaves of the foliation. In this case, the gauge freedom of the theory - as hinted above - would be translated in the freedom of choosing such a foliation without altering the dynamical evolution. Now let us consider a foliation $\mathcal{F}$ whose folios are Minkowski 4-spacetimes and another one $\mathcal{F}'$ which is identical to $\mathcal{F}$ except for the fact that some leaves in-between are FLRW 4-spacetimes. If we claim that $\mathcal{F}$ and $\mathcal{F}'$ are ontologically distinct, then so will be the corresponding dynamical evolutions and, hence, we are forced to accept that, even by specifying with arbitrary precision the data on an initial Minkowski 4-surface (which is common to $\mathcal{F}$ and $\mathcal{F}'$), the theory is unable to single out a unique dynamical evolution between $\mathcal{F}$ and $\mathcal{F}'$, which means that the theory is indeterministic. Such a situation is of course avoided if we take $\mathcal{F}$ and $\mathcal{F}'$ to represent one and the same physical situation.

In conclusion, since the theory is $diff(M_5)$-invariant, all the quantities that are not invariant under 5-diffeomorphisms are just "gauge fluff" that can be fixed at will without changing the physical information conveyed by the theory. This is the case for all the 4-dimensional quantities extracted by 5-dimensional ones by fixing the gauge (for example by setting $\psi = const.$): under a different gauge fixing, in fact, we would obtain different 4-quantities from the same 5-ones. Of course, the 4-dimensional stress-energy tensor (14) is among the non-gauge invariant quantities.

If we take seriously this picture, then we cannot but claim that, in this theory, matter is just an unphysical illusion which merely depends on the 4-dimensional "perspective" from which 5-dimensional spacetime is seen, just like a holographic sticker that shows a different image depending on how it is inclined with respect to the light source. One could say that the unreality of matter in this theory already stems from the fact that, at the fundamental level, there is nothing but 5-spacetime, but that would be a half-truth. While in fact we can agree that according to equations (9) *at the fundamental ontological level* there is nothing but 5-dimensional geometry, this does not rule out the possibility that matter configurations *supervene* on such a geometry, and hence are real albeit ontologically non-fundamental. Naively, we could solve such a "problem of matter" just by declaring our firm pre-theoretical commitment to the existence of matter as a substance and hence arguing for the dismissal of induced matter theory. This line of reasoning, however, would be weird to say the least, since it pushes us to reject a well-defined and consistent physical theory *based only* on metaphysical considerations. The other option would be to grant the theory physical dignity, in the sense that we should judge its validity based on its empirical adequacy. This second choice is potentially even worse, because it would drive us into something that smells like

14

a case of empirical incoherence:[19] the truth of this theory would undermine our empirical justification for believing it to be true. This is because, if induced matter theory fulfills the GR-desideratum, there is no way to account for whatever measurement in a 5-diffeomorphic invariant way.

Such a situation is all the more strange for the following reason. Since induced matter theory rests on the Campbell-Magaard theorem, GR is embedded in such a theory already at the mathematical level, which means that the former is a formal generalization of the latter and hence, if we forget for a moment the GR-desideratum, all the empirical tests that corroborate GR, would corroborate induced matter theory. Moreover, the theory would give novel testable predictions, mainly related to deviations in the geodesic motion of 4-dimensional objects due to the presence of the uncompactified fifth dimension (see, e.g., Wesson, 2006).

In order to find a solution to the problem of matter compatible with the GR-desideratum, some coherent account could in principle be put forward that saves the reality of 4-dimensional material entities (including measuring devices), perhaps by appealing to some distinction between "partial" and "complete" observables à la Rovelli (Rovelli, 2002), but that would look more like a patch rather than a real solution. Of course, we can prevent the problem from happening by putting in the theory 5-dimensional matter, thus retrieving a 5-dimensional analog of Einstein's field equations (1)/(2). However this would not count as a genuine solution because (i) it would just amount to dismissing induced matter theory in favor of another one, and (ii) it would not straightforwardly wash away any "perspectival" character from 4-dimensional matter: such a "full" 5-dimensional generalization of GR would for example fire up a metaphysical debate between 5- and 4-dimensionalists that would be very similar to that between 4- and 3-dimensionalists in standard relativistic physics (for an example of the literature, see Balashov, 2000; Gilmore, 2002).

Be it as it may, the moral to draw from the above analysis is that, in induced matter theory, the GR-desideratum has rather undesired consequences. Hence, if we are willing to solve the problem of matter without dismissing induced matter theory, we can do it in a straightforward way, i.e., by relaxing the GR-desideratum.

A possible strategy in this direction would be simply to add more structure to 5-spacetime, for example a 5-flow in spacetime. This would amount to selecting a privileged fifth-dimension-like vector field $\boldsymbol{\psi}$ with components $\psi^A \equiv (0, 0, 0, 0, \psi)$ by adding to the field equations (9), the following ones:

$$\psi_A \psi^A = -1, \tag{21a}$$

$$\psi^A{}_{;i} = 0. \tag{21b}$$

---

[19]The definition of empirical incoherence given here is taken from Huggett and Wüthrich (2013, section 1), who follow in turn Barrett (1999, section 4.5.2).

It is easy to see that ψ induces a privileged decomposition of the 5-manifold M_5 in 4-hypersurfaces that are normal to the vector field and, hence, spacetime-like: each 4-slice of the foliation will have equation $\psi = const.$. For this reason, equations (21) restrict the possible topologies compatible with (9) to $M_5 = M_4 \times \mathbb{R}$. In a sense, we are back to the case of the compactified Kaluza-Klein theory: we are using (21) as a dynamical constraint on the possible topologies. But there is more to that: (21) tells us that not only M_5 has to be foliable by spacetime-like hypersurfaces, but that there is, for each model, a *distinguished* way to foliate it. In short, according to (21), M_5 is a "pile" of 4-spacetimes.[20]

Under this new reading, a 5-model $\mathfrak{M}$ of the theory will not be just a couple $< M_5, g_{AB} >$ but a *triple* $< M_5, g_{AB}, \psi^A >$. Let us return to the two models whose line elements are (18) and (20), which are related by the intrinsic transformation $f \in diff(M_5)$ (dual to the extrinsic transformation (19)): we see that they correspond to the new situation where the first model is, say, $< M_5, g_{AB}, \psi^A >$ and the second model is $< M_5, f^* g_{AB}, \psi^A >$. We immediately notice that, due to the introduction of the structure ψ^A, the two models cannot be considered anymore as really related by a diffeomorphic transformation because the only diffeomorphically related objects in them are the 5-metrics. It is then clear that now the two models represent *distinct* physical situations unless $f^* \psi^A = \psi^A$.

Also in this case, we can be accused of cheating - not once but twice! Firstly because we are not considering the starting theory but a brand new one that admits a structure (namely, a privileged foliation) fixed a priori. Secondly because we are hiding such an a priori fixing behind (21), thus massaging a $diff(M_4) \times diff(\mathbb{R})$-invariant theory into a fake $diff(M_5)$-invariant one.

As regards the first charge, the plea is: guilty. It is undeniable that the starting theory dealt just with a generic 5-metric (13) solution to the field equations (9) and nothing more (e.g. no cilindricity or compactification conditions), but this is exactly at the root of the problem of matter: a 5-dimensional theory just resting on (13) and (9) has simply not enough structure to ground claims about the existence of 4-matter. Can a physical theory that places dynamical constraints so weak to risk to be empirically incoherent be considered really a theory in any useful physical sense? Under this reading, adding further structure to such a theory is more like *completing* it, rather than merely changing it.

As regards the second charge, the plea is: not guilty. ψ^A is not a "background" object as the Minkowski metric in (8) in that it does not persist unchanged (modulo diffeomorphisms) throughout all the models of the theory: (21) admits solutions that are in fact physically distinct 5-flows. To see this, it is sufficient to show that

---

[20]In this case, 4-matter would still supervene on 5-dimensional state of affairs, such as the rigging of the pile (e.g., a certain matter configuration would not have existed having the 5-flow been different).

there are two models $< M_5, g_{AB}, \psi^A >$ and $< M_5, g'_{AB}, \psi'^A >$ such that ψ^A and ψ'^A are not related by any diffeomorphism. Still, one may claim that the 5-flows affect the 4-matter distributions in each model without being affected in return. However, the "influence" we are talking about is not a physical interaction (5-flows do not really push 4-matter in any physical sense), but an ontological one: in each model 4-matter supervenes on the 5-dimensional structure comprising 5-metric *and* 5-flow.

To sum up, ψ^A is a genuine dynamical object, although it represents an additional structure besides the 5-metric: as in the Kaluza-Klein case, there is nothing particularly wrong in that as long as we do not mind relaxing the GR-desideratum. Indeed, the payoff for this relaxation is huge: the problem of matter just vanishes.

3   Conclusion

The starting point of this paper was considering the possibility that, although general covariance, diffeomorphism invariance, and background independence are not features uniquely ascribable to GR, there is something in the way GR implements such features that renders this theory better than its predecessors. We tentatively identified this "something" in the GR-desideratum, and we then discussed two cases in which forcing such a desideratum on theories that seek to generalize GR leads to conceptual difficulties. We finally suggested to overcome these difficulties just by relaxing the GR-desideratum, in particular by introducing further (dynamical) spatiotemporal structures besides the metric. From what has been discussed in section 1, it appears clear that questioning the GR-desideratum is not questioning general covariance, diffeomorphism invariance, or background independence by themselves, but the necessity of implementing these features in the way GR does.

The analysis developed in this paper is far from being just an otiose conceptual exercise, since fully understanding what the GR-desideratum really is might shed light on two huge open problems in the philosophy and physics of spacetime theories. Namely, (i) the problem of time in Hamiltonian GR and (ii) the problem of constructing a theory that combines the main tenets of relativistic physics with the empirically proven non-locality of quantum theories.

The first problem (debated for example in Earman, 2002; Maudlin, 2002), roughly speaking, arises when we decompose the 4-dimensional spacetime of GR by means of space-like 3-surfaces in the so-called ADM formulation of Hamiltonian GR (Arnowitt et al., 1962). The idea behind this (3+1) decomposition is that we can render the general relativistic dynamics simpler for calculational purposes by using the machinery of Hamiltonian dynamics. In such a setting, we specify a set of initial conditions on a starting 3-surface and we evolve it using Hamilton's

equations of motion. Of course, such a decomposition is merely formal: it is just a convenient way to recast the standard dynamics of GR. However, there is a conceptual problem lurking beneath the formalism.

To see this, let us consider some arbitrary model of GR; since the (3+1) decomposition is merely formal, we can imagine two different ways to foliate the model by means of the foliations $\mathcal{F}$ and $\mathcal{F}'$. Let us further assume that the foliations agree only on the initial surface Σ_0. We are now in the strange situation already experienced in the previous section: even by specifying with arbitrary precision the initial data on Σ_0, the dynamics is unable to single out one evolution among $\mathcal{F}$ and $\mathcal{F}'$. Does this imply indeterminism? Of course no, because both evolutions represent one and the same physical situation. However, if we take the freedom to foliate the model as a gauge freedom, then we are forced to say that $\mathcal{F}$ and $\mathcal{F}'$ are physically indistinguishable, which means that the only physically meaningful observables definable in this context are those that do not change whatever foliation we choose. This excludes that whatever physically significant quantity changes in time: taking this picture seriously seems to imply that the universe is a frozen block. Note how this problem stems from forcing a strict requirement of gauge invariance on the theory: from this point of view, the problem of time in Hamiltonian GR is strikingly similar to the problem of matter in induced matter theory. Also in this case, then, a way out of the conundrum would be to loosen or modify the gauge requirement so that other observables besides the unchanging ones could be defined.

The second problem (discussed at length in Maudlin, 1996, 2011), can be roughly summarized as follows. Let us imagine a quantum system made up of space-like separated parts A and B in an entangled state, and let us consider an event $E(A)$ such as a certain outcome of a measurement performed on A. It has been experimentally shown (Aspect et al., 1981) that $E(A)$ is not determined solely by the events in its past light-cone. This is what Bell called non-locality (Bell, 2004, ch. 2) and that might mean that $E(A)$ can be affected either by events in its future light-cone or by space-like separated events such as a measurement taking place on B. In any case, quantum non-locality seems to be at odds with the physical interpretation of the light-cone structure of relativistic spacetimes, so that some modification of such a structure might be called for.

If we consider that one of the most promising theoretical programs towards the construction of a quantum theory of gravitational phenomena (viz. canonical quantum gravity) seeks to quantize GR starting from its Hamiltonian formulation, we immediately realize that this program faces a conceptual tangle that is basically the combination of the above mentioned problems. This is a further reason for reflecting on the GR-desideratum and finding ways to relax it without giving up general covariance, diffeomorphism invariance, and background independence.

References

Anderson, J. (1967). *Principles of relativity in physics*. Academic Press.

Arnowitt, R., S. Deser, and C. Misner (1962). The dynamics of general relativity. In L. Witten (Ed.), *Gravitation: an introduction to current research*, Chapter 7, pp. 227–265. John Wiley & Sons. http://arxiv.org/abs/grqc/0405109.

Aspect, A., P. Grangier, and G. Roger (1981, August). Experimental tests of realistic local theories via Bell's theorem. *Physical Review Letters 47*(7), 460–463. http://qudev.phys.ethz.ch/content/courses/QSIT06/pdfs/Aspect81.pdf.

Balashov, Y. (2000). Enduring and perduring objects in Minkowski spacetime. *Philosophical Studies 99*, 129–166. http://www.yuribalashov.com/Papers/enduring_and_perduring_objects.pdf.

Barrett, J. (1999). *The quantum mechanics of minds and worlds*. Oxford University Press.

Bell, J. (2004). *Speakable and unspeakable in quantum mechanics* (Second ed.). Cambridge University Press. First edition 1987.

Belot, G. (2011). Background-independence. *General Relativity and Gravitation 43*, 2865–2884. http://arxiv.org/abs/1106.0920.

Brush, S. (1999). Why was relativity accepted? *Physics in Perspective 1*, 184–214. http://www.hep.yorku.ca/menary/courses/phys2040/misc/why_relativity_accepted.pdf.

Dirac, P. (1978). The excellence of Einstein's theory of gravitation. Paper given at the symposium on the impact of modern scientific ideas on society held in Munich, 18-20 September 1978. Available at http://unesdoc.unesco.org/images/0003/000311/031102eb.pdf.

Earman, J. (2002). Thoroughly modern McTaggart or what McTaggart would have said if he had read the general theory of relativity. *Philosopher's Imprint 2*(3). http://www.philosophersimprint.org/002003/.

Earman, J. (2006). Two challenges to the requirement of substantive general covariance. *Synthese 148*(2), 443–468.

Earman, J. and J. Norton (1987). What price spacetime substantivalism? The hole story. *British Journal for the Philosophy of Science. 38*, 515–525. http://bjps.oxfordjournals.org/content/38/4/515.full.pdf+html.

Friedman, M. (1983). *Foundations of Space-Time Theories. Relativistic Physics and Philosophy of Science*. Princeton University Press.

Gilmore, C. (2002). Balashov on special relativity, coexistence, and temporal parts. *Philosophical Studies 109*, 241–263. http://philsci-archive.pitt.edu/2983/.

Giulini, D. (2007). Remarks on the notions of general covariance and background independence. *Lecture notes in physics 721*, 105–120. http://arxiv.org/abs/gr-qc/0603087.

Huggett, N. and C. Wüthrich (2013). Emergent spacetime and empirical (in)coherence. *Studies in History and Philosophy of Modern Physics 44*, 276–285. http://arxiv.org/abs/1206.6290.

Kaluza, T. (1921). Zum unitätsproblem in der physik. *Sitzungsber. Preuss. Akad. Wiss. Berlin. Math. Phys. 33*, 966–972. http://homepage.uibk.ac.at/~c705204/pdf/kaluza-1921.pdf.

Klein, O. (1926). Quantentheorie und fünfdimensionale relativitätstheorie. *Zeitschrift für physik A 37*, 895–906.

Kretschmann, E. (1917). Über den physikalischen sinn der relativitätspostulate, A. Einsteins neue und seine ursprüngliche Relativitätstheorie. *Annalen der Physik 53*, 575–614. Italian translation by S. Antoci available at http://fisica.unipv.it/antoci/re/Kretschmann17.pdf.

Lehmkuhl, D. (2008). Is spacetime a gravitational field? In D. Dieks (Ed.), *The ontology of spacetime*, Volume 2 of *Philosophy and foundations of physics*, Chapter 5, pp. 83–110. Elsevier B.V.

Lehmkuhl, D. (2014). Why Einstein did not believe that general relativity geometrizes gravity. *Studies in History and Philosophy of Modern Physics 46*, 316–326. http://philsci-archive.pitt.edu/9825/.

Maudlin, T. (1989). The essence of space-time. *PSA 1988 2*, 82–91.

Maudlin, T. (1996). Space-time in the quantum world. In J. Cushing, A. Fine, and S. Goldstein (Eds.), *Bohmian mechanics and quantum theory: An appraisal*, Number 184 in Boston studies in the philosophy of science, pp. 285–307. Kluwer.

Maudlin, T. (2002). Thoroughly muddled McTaggart or how to abuse gauge freedom to generate metaphysical monstrosities. *Philosopher's Imprint 2*(4). http://www.philosophersimprint.org/002004/.

Maudlin, T. (2011). *Quantum non-locality and relativity. Metaphysical intimations of modern physics* (Third ed.). Wiley - Blackwell. First edition 1994.

Misner, C., K. Thorne, and J. Wheeler (1973). *Gravitation*. W.H. Freeman and Company.

Norton, J. (1993). General covariance and the foundations of general relativity: Eight decades of dispute. *Reports on Progress in Physics 56*, 791–858. http://www.pitt.edu/~jdnorton/papers/decades_re-set.pdf.

Norton, J. (2011). The hole argument. *The Stanford Encyclopedia of Philosophy*. http://plato.stanford.edu/archives/win2008/entries/spacetime-holearg/. See http://www.science.uva.nl/~seop/entries/spacetime-holearg/Active_Passive.html for a toy model explaining the difference between active and passive transformations.

Overduin, J. and P. Wesson (1997). Kaluza-Klein gravity. *Physics Reports 283*, 303–378. http://arxiv.org/abs/grqc/9805018.

Ponce de Leon, J. (1988). Cosmological models in a Kaluza-Klein theory with variable rest mass. *General Relativity and Gravitation 20*, 539–550.

Pooley, O. (2010). Substantive general covariance: Another decade of dispute. In M. Suàrez, M. Dorato, and M. Rèdei (Eds.), *EPSA Philosophical Issues in the Sciences: Launch of the European Philosophy of Science Association*, Volume 2, Chapter 19, pp. 197–209. Springer. http://philsci-archive.pitt.edu/9056/1/subgencov.pdf.

Romero, C., R. Tavakol, and R. Zalaletdinov (1996). The embedding of general relativity in five dimensions. *General Relativity and Gravitation 28*, 365–376.

Rovelli, C. (2002, June). Partial observables. *Physical Review D 65*(12), 124013. http://arxiv.org/abs/grqc/0110035.

Wesson, P. (1984). An embedding for general relativity with variable rest mass. *General Relativity and Gravitation 16*, 193–203.

Wesson, P. (1995). Consequences of covariance in Kaluza-Klein theory. *Modern Physics Letters A 10*(1), 15–24.

Wesson, P. (2006). The equivalence principle as a probe for higher dimensions. *International journal of modern physics D 14*, 2315–2318. `http://arxiv.org/abs/gr-qc/0601065`.

Wesson, P. and J. Ponce de Leon (1992). Kaluza-Klein equations, Einstein's equations, and an effective energy-momentum tensor. *Journal of Mathematical Physics 33*, 3883–3887.