General covariance from the perspective of Noether's theorems.

Harvey R. Brown[†] and Katherine A. Brading[•]

> "Most arguments about the formulation and content of general covariance are mere shadow boxing. Since all viable theories must be generally covariant, the only question of interest is their relative simplicity and the nuts and bolts of their construction." J. Barbour (2001)

I. Introduction

Questions concerning the meaning of the principle of general covariance and, perhaps to a lesser extent, its precise historical role in the development of Einstein's general theory of relativity (GR), never quite seem to go away[1]. General covariance is a bit like the principle of equivalence: much cited, often misunderstood, and a noble, if treacherous, source of occupation for the philosopher of physics. After all, in the process of developing GR, Einstein himself got seriously confused about it in a number of ways, and his mature writings never laid the matter to rest. Our own ideas on the topic were largely the by-product of immersion in the 1918 theorems of Emmy Noether, whose work was inspired by the attempt amongst the Göttingen mathematicians to understand the technical role of general covariance in the variational approach to GR. The results of Noether's work provide an illuminating method for testing the consequences of what we shall refer to as "coordinate generality" and for exploring what else must be added to this requirement in order to give the general covariance of GR its far-reaching physical significance. The discussion takes us through Noether's first and second theorems, and then a third related theorem due to F. Klein (which we call the Boundary theorem). Along the way contact will be made with the contributions of, principally, J.L. Anderson, A.

[†] Faculty of Philosophy, University of Oxford, 10 Merton Street, Oxford OX1 4JJ, U.K. E-mail: harvey.brown@philosophy.ox.ac.uk.

[•] Wolfson College, Linton Road, Oxford OX2 6UD. E-mail: katherine.brading@wolfson.ox.ac.uk.

[1] Some significant recent discussions are found in Norton (1993), Rynasiewicz (1999), Norton (2001) and Saunders (2001).

Trautman, P.A.M. Dirac and R. Torretti, not to mention the father of the whole business, A. Einstein.

II. Preliminary considerations.

Let us start with the familiar electromagnetic action in a spacetime with the (possibly curved) metric tensor field $g^{\mu\nu}$ which has Lorentzian signature and whose determinant is denoted by *g*:

$$S_{EM} = \int_{\Omega} L_{EM} \mathrm{d}^4 x = -\frac{1}{4}\int_{\Omega} g^{\mu\rho} g^{\nu\sigma} F_{\mu\nu} F_{\rho\sigma} \sqrt{-g}\mathrm{d}^4 x\ , \tag{1}$$

where $-F_{\mu\nu} = A_{\mu,\nu} - A_{\nu,\mu}$ and Ω is an arbitrary compact region of spacetime. (We are adopting the Einstein summation convention for Greek indices throughout.) If, as is well-known, we apply Hamilton's principle with respect to variations in the 4-potential A_{μ}, then we obtain the covariant form of Maxwell's equations in the source-free case:

$$F^{\mu\nu}{}_{;\nu} = 0 \tag{2}$$

Now suppose we choose similarly to apply Hamilton's principle with respect to variations in $g^{\mu\nu}$, treating (just for the sake of argument) S_{EM} as the total action. Then we immediately obtain

$$T_{\mu\nu} = 0 \tag{3}$$

where as usual in this context the stress-energy tensor $T_{\mu\nu}$ is defined in terms of the relevant variational derivative of the lagrangian density L_{EM} :

$$T_{\mu\nu} := \frac{-2}{\sqrt{-g}} \frac{\delta L_{EM}}{\delta g^{\mu\nu}}. \tag{4}$$

The result (3) is pretty disastrous: it means that $F_{\mu\nu} = 0$. If we want the metric tensor to be a *bona fide* dynamical player, we need to add another term to the action which is a functional of at least the $g^{\mu\nu}$ and their derivatives, but not of the A_μ and their derivatives—which is of course what is done in GR. But note that the electromagnetic field variables do not require an analogous second contribution to the action, independent of $g^{\mu\nu}$. Indeed, general covariance goes some way to ruling out such a possibility, as we see in section V.

Recall now that the Lagrangian in S_{EM} is a scalar density, so S_{EM} is strictly invariant under general (infinitesimal) coordinate transformations

$$x^\mu \to x'^\mu = x^\mu + \varepsilon \xi^\mu \tag{5}$$

where ξ^μ is an arbitrary vector field, and ε taken to be small. This condition is sufficient, but not necessary, for the general covariance of the Euler-Lagrange equations (2) and (3), as we see below. Now a subgroup of this group of arbitrary coordinate transformations is that associated with rigid spacetime translations for which the ξ^μ in (5) are now independent of the coordinates. It follows that elements of this 'rigid' subgroup are also 'Noether symmetries', and are specifically of the kind that figure in Noether's first (and more celebrated) theorem that leads to a connection between symmetries and conservation principles[2].

The theorem, in a generalised form, is this. Suppose that the first order variation in the action *S* vanishes (up to a surface term, of which more below) under a given group of infinitesimal transformations of the dependent or independent variables that depend on

[2] The English translation of Noether's celebrated paper (Noether 1918) is found in Tavel (1971). Good accounts of the first theorem are found in, for example, Hill (1951), Trautman (1962) and Doughty (1990).

a number of constant parameters. Then for each such parameter there is a *linear combination of the Euler expressions* associated with each of the dependent variables (fields) that *is equal to the divergence of the associated 'Noether current'*. (Recall that the "Euler expression" is the variational derivative of the lagrangian density with respect to the chosen field variable; when it vanishes, as a result of applying Hamilton's principle to that field, the Euler-Lagrange equations are said to hold for the field. The Noether current is a quantity which depends on, *inter alia*, the way the lagrangian density in turn depends on the derivatives of all the field variables.) If the mentioned linear combination of Euler expressions happens to vanish, a continuity equation is obtained of the form:

$$\partial_\mu j_k^\mu = 0 \tag{6}$$

where the index k in the current picks out one of the constant parameters involved in the symmetry transformation. From (6) a time-independent Noether 'charge' can finally be constructed by integration of j_k^0 over a 3-dimensional spatial region with suitable boundary conditions.

Let's apply Noether's first theorem to the electromagnetic action (1), using the invariance under rigid spacetime translations. If we assume that the Euler expressions associated with variations in the A_μ vanish, i.e. if we assume Maxwell's equations (2) hold, then the Noether condition reduces to:

$$T_{\mu\nu} g^{\mu\nu}{}_{,\sigma} \sqrt{-g} = -2\partial_\mu j_\sigma^\mu \ , \tag{7}$$

where $T_{\mu\nu}$ is defined as in (4) and j_ν^μ is the Noether current associated with S_{EM}. This current, it turns out, takes the form

$$\begin{aligned} j^{\mu}_{\sigma} &= \frac{\partial L_{EM}}{\partial A_{\alpha,\mu}} A_{\alpha,\sigma} - \delta^{\mu}_{\sigma} L_{EM} \\ &= \sqrt{-g}\left(F^{\mu\alpha} A_{\alpha,\sigma} - \frac{1}{4}\delta^{\mu}_{\sigma} F_{\lambda\rho} F^{\lambda\rho} \right) \end{aligned} \tag{8}$$

and the form of $T_{\mu\nu}$ can be obtained directly from (1) and (4):

$$T_{\mu\nu} = -F_{\mu\rho} F^{\rho}{}_{\nu} + \frac{1}{4} g_{\mu\nu} F_{\alpha\beta} F^{\alpha\beta}\,. \tag{9}$$

Now using again the field equations (2) it can be shown from (8) and (9) that

$$\sqrt{-g}\, T^{\mu}_{\sigma} = j^{\mu}_{\sigma} + \partial_{\alpha}\left(\sqrt{-g}\, F^{\mu\alpha} A_{\sigma}\right). \tag{10}$$

We thus see that the Noether current j^{μ}_{σ} equals $\sqrt{-g}\, T^{\mu}_{\sigma}$ up to a divergence term. *But we still have not obtained a conservation principle*, because the left hand side of Noether's equation (7) does not vanish. This is also seen by using (10) together with (7), obtaining

$$T^{\mu}_{\nu;\mu} = 0. \tag{11}$$

The failure of this formulation to correspond to a true conservation principle resides in the fact that it is the covariant, not the ordinary derivative that appears in the equation. One way to get rid of this impediment is to go to the special case of the flat Minkowski metric $g^{\mu\nu} = \eta^{\mu\nu}$, in which there are global (inertial) coordinates such that $\eta^{\mu\nu}{}_{,\lambda} = 0$, or equivalently, such that partial and covariant derivatives coincide. The Noether condition (7) now reduces in these coordinate systems to the desired form

$$j^{\mu}_{\nu,\mu} = T^{\mu}_{\nu,\mu} = 0. \tag{12}$$

What does this example illustrate? *First*, that it is not enough that certain 'global' transformations of the dependent and/or independent variables are Noether symmetries—i.e. ones under which the action is invariant up to a surface term—for there to be a conservation principle, or even a continuity equation. (An analogous situation holds for the 'local' symmetries like general covariance that feature in Noether's second theorem, as we shall see.) What more is needed? Trautman[3] long ago recognised the importance of this question and emphasised that unless *all* the Euler expressions vanish on the LHS of the Noether condition—a necessary but not sufficient condition for all the associated fields to be dynamical[4]— a conservation principle need not ensue. But he also emphasised the relevance of the possible existence of "motions" (isometries) in the spacetime in cases like ours where the metric field is non-dynamical. Recall that in the above example, the conservation principle holds perfectly well, *and is only interesting*, when the Euler expression associated with $g^{\mu\nu}$—which is essentially $T_{\mu\nu}$—does *not* vanish. What secures the conservation principle, besides the vanishing of the remaining Euler expression associated with A_μ, is the condition that $g^{\mu\nu}$ is an *absolute* background geometry of a *special* kind: *it is flat*. (Actually, conservation laws exist more generally whenever the spacetime has constant curvature.)

Secondly, it is the flatness of the background geometry that permits the Noether symmetries, which involve certain, special transformations between global coordinate systems, to have an active interpretation. It is not enough for the symmetries themselves to be 'global' in the *other* sense that they do not depend on the coordinates. We return to this point in section VIII.

Thirdly, the example exhibits a connection between the Noether current and the variationally-defined stress-energy tensor (4). This connection is a consequence of a generic structural feature of generally covariant theories of matter, which is captured in a Noether-type theorem that was first demonstrated by Felix Klein in 1918. Discussion of this issue will be found in section VI.

---

[3] See Trautman (1962), sections 5-2 and especially 5-3.

Fourthly and finally, let us not forget the 1917 lesson of E. Kretschmann[5], which in this context is that the original Maxwell theory in Minkowski spacetime is no different from *any* theory in being susceptible to a generally covariant formulation. It is a commonplace that at the level of the field equations, the price of the move to general covariance in Minkowski spacetime is the explicit appearance of previously implicit geometric structure (in particular the affine connection coefficients) in (2):

$$F^{\mu\nu}{}_{;\mu} = F^{\mu\nu}{}_{,\mu} + \Gamma^{\mu}_{\alpha\mu}F^{\alpha\nu} + \Gamma^{\nu}_{\alpha\mu}F^{\mu\alpha} = 0 \,. \tag{13}$$

But implementing a generally covariant formulation of the theory adds nothing new to its empirical content. As Ohanian and Ruffini write[6], "We will obey this commandment [general covariance] for the best of all reasons—it costs us nothing to do so." And yet the fact that there is no cost is itself non-trivial. For, as we shall see in section IV, the move to general covariance immediately raises the spectre of underdetermination. It might seem odd that a mere reformulation of a well-behaved dynamical theory, such as Maxwell theory in Minkowski spacetime, should complicate the issue as to whether it has a well-defined initial value problem. Indeed, realisation that the complication in this case must be a mere artifact of the new generalised presentation, i.e. that it *must* be innocuous, forces one to make a crucial decision. One must accept either that a privileged class of global coordinate systems—the inertial systems— is required in the process of prediction, or one must embrace the 'Leibniz equivalence' of diffeomorphically related versions of the world.

As for the first option, the inertial coordinate systems clearly *are* privileged in Minkowski spacetime, but the question is whether they are *essential for the purposes of prediction* in Maxwell theory. If they are, there is hardly any point to imposing general covariance. But it seems at any rate more natural to adopt the second option. One

---

[4] An apparent example of a non-dynamical field that is nonetheless subject to Hamilton's principle is found in Sorkin (2001).

[5] An excellent, detailed analysis of Kretschmann's famous paper (Kretschmann 1917) is found in Rynasiewicz (1999); see also Norton (1993) in this connection.

amongst several possible reasons for this is that an analogous threat of underdetermination in the formulation of electromagnetism involving the 4-potential A_μ—which arises because the equations are covariant under local gauge transformations of the A_μ—is removed once it is realised that the empirical content of the theory is gauge-independent. (In section VI we will see a connection of sorts between the existence of gauge invariance of this sort and the requirement of general covariance for equations for vector matter fields.) Einstein's struggle from 1913 to 1914 with the implications of the "hole argument"[7] in GR leads one to surmise that until that period he had not properly considered either the question as to whether special relativity has a generally covariant formulation, or the significance of the gauge structure of Maxwell theory.[8]

III. General covariance vs. coordinate generality.

We saw in the previous section that the invariant action (1) must be supplemented with a further term if $g^{\mu\nu}$ is to play a valid dynamical role in the theory. Suppose we require that the ensuing equations of motion, both for $g^{\mu\nu}$ and A_μ, are generally covariant. What kind of restriction on the action is this?

A brief look at the early history of the action principle in GR is enlightening in this respect. In 1915 D. Hilbert had proposed a pure gravitational action whose lagrangian density is the scalar curvature density $R\sqrt{-g}$. This action is clearly invariant under arbitrary diffeomorphisms. But in 1915 and 1916, Einstein proposed two versions of what

---

[6] See Ohanian and Ruffini (1994), p. 373.

[7] Recent discussions of the hole argument can be found in, e.g., Norton (1993), Rynasiewicz (1999), and Saunders (2001).

[8] It was Einstein's belated insight that different spacetime structures related by diffeomorphisms are nothing other than different representations of the same reality that solved the undertermination problem in GR. Now, diffeomorphisms end up having a much more interesting *creative* role within the "best-matching" (Machian) approach to gravitational dynamics defended by Barbour. Here, they are essential in the process of comparing, not two representations of a given geometry, but two distinct geometries in order to capture what their intrinsic difference is. Further details can be found in Barbour (2001) and the works cited therein.

is sometimes called the "Γ-Γ" action. The first contained a lagrangian density of the form $g^{\mu\nu}\Gamma^{\alpha}_{\mu\beta}\Gamma^{\beta}_{\nu\alpha}$, and was defined only in relation to special coordinate systems for which $\sqrt{-g} = 1$. (It is remarkable that despite his commitment to general covariance, for a period in 1915 and 1916 Einstein thought that restriction to such special coordinates would lead to a significant simplication of gravitational physics. We return to this issue in section VII.) The second version of the action contained the lagrangian density $g^{\mu\nu}\left(\Gamma^{\alpha}_{\mu\beta}\Gamma^{\beta}_{\nu\alpha} - \Gamma^{\beta}_{\mu\nu}\Gamma^{\beta}_{\alpha\beta}\right)\sqrt{-g}$, defined now for an arbitrary coordinate system (and reducing to the first version for the mentioned special coordinates, since when $\partial_{\sigma}\left(\sqrt{-g}\right) = 0$ then $\Gamma^{\beta}_{\alpha\beta} = 0$).[9] What is interesting for our purposes is that the latter version of the Γ-Γ action is clearly not invariant with respect to arbitrary coordinate transformations, although it leads to the same generally covariant field equations for $g^{\mu\nu}$ as Hilbert's.

The fact that it is sufficient *but not necessary* that the first-order variation in the total action *S* strictly vanish under arbitrary infinitesimal diffeomorphisms in order for the Euler-Lagrange equations to be generally covariant is today no secret, but its acknowledgement in modern texts on GR is not guaranteed[10]. To see what is going on technically, recall that Hilbert's invariant action has a curious property. Despite being of second order (i.e. dependent on first and second derivatives of $g^{\mu\nu}$) it somehow gives rise to only second-order, rather than fourth-order, Euler-Lagrange field equations. The Einstein Γ-Γ action, which is first-order, demonstrates how this magic occurs. Subtracting Einstein's lagrangian density from Hilbert's, one is left with a term that takes the form of a total divergence of a first-order functional of $g^{\mu\nu}$. This term contains *all* the second-order quantities in the Hilbert action. But it is known from the calculus of variations that to any Lagrangian density can be added a total divergence without affecting the Euler-Lagrange equations. So the Hilbert and Einstein actions are

---

[9] The abbreviated form of the action appeared in Einstein (1915, 1916a), and the full form in a footnote in Einstein (1916b). Note that Einstein's connection coefficients were the negative of the usual Christoffel symbol.

effectively the same—as Einstein fully appreciated. (Actions of the Γ-Γ kind, which are invariant under general coordinate transformations up to a surface term, are sometimes said to be "quasi-invariant", or, perhaps regrettably, "covariant". Interesting quasi-invariant actions arise not just in GR of course; the standard action of the free Newtonian particle is quasi-invariant relative to Galilean boosts. It is noteworthy that Noether herself did not take into account such cases in her general treatment of the variational problem.)

Now in Lovelock (1969) we find the following result. In a spacetime of four or less dimensions, any strictly invariant, second-order gravitational action that gives rise to second-order field equations must be associated with a lagrangian density which is a linear combination of the Hilbert lagrangian density and a cosmological term:

$$S_{grav} = \int_\Omega \left(aR\sqrt{|g|} + b\sqrt{|g|} \right) d^4x \tag{14}$$

where R is the curvature scalar. (Note that $|g|$ appears rather than the usual $-g$, because Lovelock made no assumptions about the signature of $g^{\mu\nu}$.) A strengthened version of this result was reported in Grigore (1992), concerning the class of *first-order, quasi-invariant* gravitational actions. Grigore's result, the proof of which is especially complicated, states that independently of the dimensionality of spacetime, the lagrangian density appearing in this action must take the form of a linear combination of the 1916 Einstein Γ-Γ lagrangian density[11] and a cosmological term:

$$S_{grav} = \int_\Omega \left\{ ag^{\mu\nu}\left(\Gamma^\alpha_{\mu\beta}\Gamma^\beta_{\nu\alpha} - \Gamma^\beta_{\mu\nu}\Gamma^\beta_{\alpha\beta} \right)\sqrt{|g|} + b\sqrt{|g|} \right\} d^4x\,. \tag{15}$$

The Lovelock-Grigore theorems are evidently highly non-trivial and they share the premiss that the Euler-Lagrange equations must be generally covariant. But it is worth

---

[10] In Misner, Thorne and Wheeler (1973), p. 503, for instance, one reads that: "… the action integral … is a scalar invariant, a number, the value of which depends on the physics but not at all on the system of coordinates in which that physics is expressed".

emphasising that the mere requirement that diffeomorphisms are Noether symmetries is far too weak to engender anything like these results. Indeed we have seen that both results explicitly require in addition that the Euler-Lagrange equations are no higher than second-order. We recall, however, that the generally covariant Brans-Dicke (1961) theory of gravitation also contains second order equations in $g^{\mu\nu}$, but its second-order gravitational action is not Hilbert's. Consider, in the same spirit, the following two *first-order* actions:

$$\bar{S}_{grav} = \int g^{\mu\rho} g^{\nu\sigma} B_{\mu;\nu} B_{\rho;\sigma} \sqrt{|g|}\, \mathrm{d}^4 x\,; \quad \tilde{S}_{grav} = \int g_{\mu\alpha}\left(\Gamma^{\mu}_{\nu\lambda} - \bar{\Gamma}^{\mu}_{\nu\lambda}\right) g^{\nu\beta} g^{\lambda\gamma}\left(\Gamma^{\alpha}_{\beta\gamma} - \bar{\Gamma}^{\alpha}_{\beta\gamma}\right)\sqrt{|g|}\, \mathrm{d}^4 x \qquad (16)$$

where B_{μ} is some vector field, $\Gamma^{\mu}_{\nu\lambda}$ is the usual metric-compatible connection (Christoffel symbol), and $\bar{\Gamma}^{\mu}_{\nu\lambda}$ is some distinct connection. These actions too are invariant under diffeomorphisms , so they might appear at first sight to satisfy the Grigore conditions, even though each is quite different from the Einstein 1916 action. Of course the reason that these three cases circumvent the Lovelock-Grigore results is that each introduces a geometric object field over and above $g^{\mu\nu}$ (in the Brans-Dicke case a scalar field). *Both Lovelock and Grigore are assuming that the gravitational lagrangian density is constructed out of the* $g^{\mu\nu}$ *field and its derivatives, alone*. This assumption is, like the previous one regarding the second-order nature of the field equations, quite independent of the requirement that the Euler-Lagrange field equations be generally covariant.

That the gravitational interaction in GR should be associated with the existence of a metric field with Lorentzian signature finds its motivation in all those empirical results that are related to the equivalence principle, as long as the trajectories of freely falling test particles are taken to correspond to time-like geodesics. The *strong* version of the principle presupposes that *no more* than the metric field is needed to account for the

[11] Grigore does not mention Einstein's 1916 lagrangian density, but his expression is equivalent to it.

gravitational potential.[12] In particular, there is no need to introduce any geometrical objects into the gravitational action that are *absolute*, in the sense of not being subject to Hamilton's variational principle (i.e. in the sense of acting but not being acted upon). The significance of this point will be discussed in section VII below. In the meantime, the point we wish to reiterate—the point that Anderson (1964, 1966, 1967) and Trautman (1966) went to such pains to emphasise— is that general covariance, puny though it is as a constraint on the equations of motion *per se*, leads to highly non-trivial conditions when combined with further demands. The demand we are interested in here is consistency with the strong equivalence principle—or at any rate with the principle that nothing other than the dynamical $g^{\mu\nu}$ field is necessary in order to account for gravity. (This last principle has itself occasionally been referred to in the mainstream literature, for better or for worse, as "the principle of general covariance".[13] In order to avoid any confusion, we shall sometimes use the term "coordinate generality" when we particularly wish to emphasise that we mean general covariance in the weak, Kretschmann sense.)[14]

.

IV. Noether's second theorem

In the book based on his 1970s Florida lectures on GR[15], P.A.M. Dirac emphasised a particular feature of the coupling of gravity with matter fields: that the Euler-Lagrange equations, obtained by varying the total action with respect to each of the distinct dynamical fields and using Hamilton's stationarity principle, are not all

---

[12] For a particularly good discussion of the strong equivalence principle, see Ehlers (1973). Note that the Brans-Dicke theory, which postulates a scalar field as well as $g^{\mu\nu}$, satisfies the "medium-strong" or "semistrong" version of the equivalence principle. For a discussion of this distinction, see Rindler (1977), section 1.20. It should be stressed that in all these versions of the principle, it is being assumed that the geodesic deviation associated with tidal gravitational effects are curvature-related. In the so-called teleparallel approach to GR, it is not affine curvature that gives rise to geodesic deviation but torsion; for a brief review see Blagojevic (2002), pp 68-72.

[13] See Wald (1984), p. 57.

[14] Coordinate generality as we have defined it corresponds to what Saunders (2001) refers to as "diffeomorphism covariance".

[15] Dirac (1996), section 29.

independent of one another. Dirac considered the specific case of a continuous distribution of charged matter, interacting with the electromagnetic field, with both 'fields' coupled to gravity. He showed that the equations of motion of the elements of matter (incorporating the Lorentz force law) are not only derivable directly by varying the given action with respect to the appropriate matter variables. They are also a consequence of the Einstein field equations (obtained of course by varying with respect to $g^{\mu\nu}$) together with the covariant form of Maxwell's equations (obtained by varying with respect to the electromagnetic four-potential A_μ). Dirac realised that this *interdependence* of the equations of motion is a consequence of the fact that arbitrary diffeomorphisms are dynamical symmetries. (By "interdependence" we mean that not all the equations of motion are independent.)

Dirac did not mention it, but the issue he was highlighting dates back to the investigations, between 1915 and 1918, involving David Hilbert, Felix Klein, Hermann Weyl and Emmy Noether, concerning the role of coordinate generality in the variational approach to geometric theories of gravitation of the sort suggested by Einstein.[16] It was Hilbert in 1915 who apparently first realised that the interdependence of the Euler-Lagrange equations was a consequence of general covariance—even in the absence of matter fields—, and Noether in 1918 who succeeded in treating this and related issues with rigour and (almost) complete generality. It is interesting that these considerations represent the flipside of the *underdetermination* problem that had caused Einstein such headaches before he arrived in late 1915 at the final triumphant form of his gravitational field equations—the "hole argument" mentioned earlier. In a sense they are one and the same problem. If the Euler-Lagrange equations associated with the Hilbert gravitational action, say, were all independent, then one could determine uniquely for a given coordinate system (x^σ) the value of $g^{\mu\nu}(x^\sigma)$ throughout spacetime, given the values of the $g^{\mu\nu}$ and their first derivatives on a given spacelike hypersurface. But consider a different coordinate system (x'^σ) that happens to coincide with the first one only in the

[16] A useful historical account is found in Rowe (1999), although we are not in agreement with some of the technical analysis therein.

vicinity of the 'initial value' hypersurface. Then because of the general covariance of the matter-free field equations, at an arbitrary point far from the hypersurface, $g'^{\mu\nu}(x'^{\sigma})$ must be the same function of the variables (x'^{σ}) as $g^{\mu\nu}(x^{\sigma})$ is of (x^{σ}). But this is inconsistent with the rules of tensor transformation. Thus, the equations are not all independent, and the predictions appear to be underdetermined.[17]

Noether's second theorem—again in a generalised version— involves the determination of a necessary condition on the form of the lagrangian density L in order that the first-order variation of the action vanish (up to a possible surface term) under infinitesimal transformations of the dependent and/or independent variables which depend on arbitrary functions of the coordinates. We are interested specifically in the arbitrary (generally non-rigid) coordinate transformations (5). The resulting "Noether identity" demonstrates, not surprisingly, a dependence between the Euler expressions generated by these transformations. Let us suppose that the lagrangian density is some functional of the $g^{\mu\nu}$ field and a collection of matter fields ϕ_i —not necessarily scalar fields but presumably components of tensor fields—and their derivatives. Then the Noether identity can be shown to take the following form:

$$-g^{\mu\nu}{}_{;\alpha}E_{\mu\nu} + \sum_i a_{i\alpha}E^i = 2\left(g^{\mu\sigma}E_{\mu\alpha}\right)_{;\sigma} + \sum_i \left(b^{\sigma}_{i\alpha}E^i\right)_{,\sigma} \tag{17}$$

where $E_{\mu\nu}$ and E^i are the Euler expressions associated with (induced) variations in $g^{\mu\nu}$ and ϕ_i respectively: $E_{\mu\nu} = \delta L / \delta g^{\mu\nu}$; $E^i = \delta L / \delta \phi_i$. The coefficients $a_{\alpha i}$ and $b^{\mu}_{\alpha i}$ are determined by the form of the Lie drag of the fields ϕ^i (in a coordinate-dependent way).

---

[17] The Euler-Lagrange equations will have unique solutions to the Cauchy initial value problem for an appropriate initial data hypersurface if they take the "normal form" defined by Cauchy and Kovalevskaya. However, it can easily be shown that the condition ("identity") associated with Noether's second theorem rules out the normal form holding for generally covariant equations. A useful discussion of the underdetermination issue and its connection with Noether's second theorem is found in Anderson (1967), sections 4-6, 4-7. See also Brading and Brown (2001).

(i) Before we return to Dirac's example, let us first briefly rehearse the more-or-less familiar application of Noether's second theorem to matter-free GR. Here the lagrangian density can be written either in the usual Hilbert form $R\sqrt{-g}$, or in the Einstein Γ-Γ form. In both cases we get the familiar Einstein tensor $G_{\mu\nu}$ appearing in the Euler expression:

$$E_{\mu\nu} = G_{\mu\nu}\sqrt{-g} = \left(R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu}\right)\sqrt{-g} \qquad (18)$$

where $R_{\mu\nu}$ is the Ricci tensor. Then, using the fact that the metric is compatible with the connection $\left(g^{\mu\nu}{}_{;\lambda} = 0\right)$, we get from (17) the (twice-) contracted Bianchi identity:

$$G^{\sigma}_{\alpha;\sigma} = 0. \qquad (19)$$

There are two aspects of this result we wish to comment on.

First, the result does not depend on the form of the gravitational action, in the sense that in the absence of matter, metric compatibility $\left(g^{\mu\nu}{}_{;\lambda} = 0\right)$ allows us to infer from (17) that $E^{\sigma}_{\alpha;\sigma} = 0$, as long as removing the matter part from the total action does not affect its invariance properties. The point of mentioning this is that it sheds light on a question posed in 1973 by Ehlers[18]. The question was whether a "reasonable" alternative to the Einstein field equations could exist which would still take the generic form $V_{\mu\nu}\left(g^{\rho\sigma}, g^{\rho\sigma}{}_{,\lambda}, g^{\rho\sigma}{}_{,\lambda\gamma}\right) = T_{\mu\nu}$ but in which $V^{\sigma}_{\alpha;\sigma} = 0$ need not hold *identically* in $g^{\mu\nu}$. (It is still being taken for granted that $T^{\sigma}_{\alpha;\sigma} = 0$ is a consequence of the matter equations of motion.) Noether's second theorem implies that this cannot be the case if the purely gravitational part of the action—for which $V_{\mu\nu}$ is the Euler expression—is itself quasi-invariant or invariant.

[18] Ehlers (1973), p. 42.

Second, (19) is a mathematical identity in its own right: it is a consequence (for the metric-compatible connection) of the ordinary Bianchi identity. Its validity does not depend on any properties of the gravitational action![19] It may be interesting here to consider the case of the Palatini procedure of treating the connection and metric as independent in the Hilbert action:

$$\int_\Omega R\sqrt{-g}\,\mathrm{d}^4x = \int_\Omega \left(\Gamma^{\sigma}_{\mu\nu,\sigma} - \Gamma^{\sigma}_{\mu\sigma,\nu} + \Gamma^{\sigma}_{\lambda\sigma}\Gamma^{\lambda}_{\mu\nu} - \Gamma^{\sigma}_{\lambda\nu}\Gamma^{\lambda}_{\mu\sigma}\right) g^{\mu\nu}\sqrt{-g}\,\mathrm{d}^4x\,. \tag{20}$$

It is well-known that metric compatibility $\left(g^{\mu\nu}{}_{;\lambda} = 0\right)$ is now a consequence of the field equations generated by varying with respect to the connection. In this case, Noether's second theorem does not lead to the contracted Bianchi identity (19). We note that G. Svetlichny has recently shown that the Noether identities for the Palatini procedure take the covariant form:

$$R^{\mu}_{\nu\lambda\rho}E^{\nu\lambda}_{\mu} - E^{\nu\lambda}_{\rho;\nu\lambda} - g^{\mu\nu}{}_{;\rho}E_{\mu\nu} - 2\left(E_{\mu\rho}g^{\mu\nu}\right)_{;\nu} = 0 \tag{21}$$

where $R^{\mu}_{\nu\lambda\rho}$ is the Riemann tensor, $E^{\nu\lambda}_{\mu}$ is the Euler expression defined with respect to variations of the connection $\Gamma^{\mu}_{\nu\lambda}$, and $E_{\mu\nu}$ is defined as above. We see that if the connection is not varied independently we return to the special case of (17) in which the matter fields ϕ_i vanish.[20]

---

[19] We make this point in case Wald's claim that "the contracted Bianchi identity may be viewed as a consequence of the invariance of the Hilbert action under diffeomorphisms" (Wald 1984, p. 456) is taken too literally.

[20] In the same work, Svetlichny has also shown that in deriving metric compatibility $\left(g^{\mu\nu}{}_{;\lambda} = 0\right)$ *via* the Palatini procedure, the usual assumption of symmetry for the connection (vanishing torsion) is not sufficient in the case of two-dimensional spacetime. Furthermore, if torsion does exist, for spacetimes with two or more dimensions, the field equations for the metric field are the usual Einstein ones, so that the spacetime objects characterising the torsion are uncoupled from the metric. See Svetlichny (2001).

(ii) Let us return to the condition (17). Suppose that amongst the matter fields ϕ_i there is at least one scalar field φ. Because of the nature of the Lie drag of scalar fields, the *b*-coefficient associated with it in (17) vanishes. It follows that if the Euler expressions associated with all the remaining matter fields as well as $g^{\mu\nu}$ vanish—i.e. if the equations of motion of all fields but φ are assumed to hold—then (17) reduces to the form $a_{\varphi\alpha}E^{\varphi} = \varphi_{,\alpha}E^{\varphi} = 0$. Since the gradient of φ is arbitrary at any spacetime point, the only solution is $E^{\varphi} = 0$, meaning that the equation of motion of the scalar field must also hold, if those for all the other fields do.

This simple result based on general covariance is reminiscent of the Dirac example mentioned at the beginning of this section. We should stress however that Dirac's continuous distribution of charged matter, whose equation of motion involves the Lorentz force law, is not represented by a scalar field. It is not clear to us whether Dirac's result can be given a similarly simple underpinning with the use of the general Noether condition (17).

V. Noether's second theorem and the 'response equation'.

But Dirac himself was keen to further clarify the connection between the interdependence of the equations of motion and general covariance. To do so he proceeded, in his Florida lectures, to consider *arbitrary* matter fields constrained only by the requirement that their contribution to the total action is, like the (Einstein-Hilbert) gravitational action on its own, invariant under diffeomorphisms. Dirac claimed to show that in this case the covariant divergence of the symmetric stress-energy tensor $T_{\mu\nu}$, defined as above in terms of the variational derivative of the matter action, vanishes. (Since this relation $T^{\mu}_{\nu\ ;\mu} = 0$ determines the re-action of the metric field on its sources—the possibility of which is secured by the non-linearity of the theory—it is sometimes called the 'response equation'.) It is on account of this relation, said Dirac, that the gravitational field equations are not all independent of the matter field equations.

We find Dirac's derivation of $T^{\mu}_{\nu\ ;\mu} = 0$ hard to follow[21], and so we will now sketch a reconstruction of it. The result is consistent with the standard account of "conservation" principles in GR, but it is stronger than the usual textbook derivation of $T^{\mu}_{\nu;\mu} = 0$, as we shall see.

We met in the previous section the general Noether condition (17) on the lagrangian density which depends on both $g^{\mu\nu}$ and arbitrary matter fields ϕ_i, in order for the action to be invariant (up to a possible surface term) under arbitrary coordinate transformations leaving the bounding surface of integration unchanged. Rearranging (17), we get

$$\sum_i \left\{ E^i a_{\alpha i} - \left(E^i b^{\sigma}_{\alpha i} \right)_{,\sigma} \right\} = 2\left(E_{\mu\alpha} g^{\sigma\mu} \right)_{;\sigma} + g^{\mu\nu}{}_{;\alpha} E_{\mu\nu}. \tag{22}$$

Following Dirac, we further consider the Noether identities following from the (quasi-) invariance of the *matter part of the action*. We obtain, again after rearranging terms:

$$\sum_i \left\{ E^i a_{\alpha i} - \left(E^i b^{\mu}_{\alpha i} \right)_{,\mu} \right\} = 2\left(N_{\mu\alpha} g^{\sigma\mu} \right)_{;\sigma} + g^{\mu\nu}{}_{;\alpha} N_{\mu\nu}, \tag{23}$$

where $N_{\mu\nu}$ is the Euler expression associated with the matter action in respect of variations in $g^{\mu\nu}$. That is, $N_{\mu\nu} = \delta L_M / \delta g^{\mu\nu} = -\left(\sqrt{-g}/2\right) T_{\mu\nu}$, where L_M is the matter lagrangian density, and $T_{\mu\nu}$ is the stress-energy tensor (see above).

Metric compatibility $\left(g^{\mu\nu}{}_{;\lambda} = 0\right)$ allows us now to infer from (22) and (23) that

[21] A reader of the relevant section (Dirac 1999, section 30, particularly p. 60) might be forgiven for thinking that the result is obtained without appeal to Hamilton's principle, i.e. to *any* equations of motion! This would make it, in the modern parlance, a "strong" principle. Less importantly, but still significantly, it is also obtained without specifying the form of the gravitational action, although it seems to require non-trivial constraints on the matter fields. We leave it up to the reader to decide whether the proof we give below is what Dirac really had in mind.

$$\left(\{E_{\mu\alpha} - N_{\mu\alpha}\} g^{\sigma\mu}\right)_{;\sigma} = 0 \,. \tag{24}$$

So we reach the desired result

$$N^{\sigma}_{\alpha;\sigma} = T^{\sigma}_{\alpha;\sigma} = 0 \tag{25}$$

only when we assume the validity of the gravitational field equations $E_{\mu\nu} = 0$. The salient difference between this argument and the usual derivation of the response equation $T^{\sigma}_{\alpha;\sigma} = 0$ in GR involving the Einstein field equations $G_{\mu\nu} = \kappa T_{\mu\nu}$ and the contracted Bianchi identities $G^{\mu}_{\nu;\mu} = 0$,[22] is that the derivation here *does not depend on the form of the action* apart from the invariance properties (which now encompass quasi- as well as strict invariance) of both the gravitational and matter-related parts.[23]

So we see that the combination of (a) what we might call the *double invariance* condition (namely, that both the gravitational and matter parts of the action are individually invariant or quasi-invariant) and (b) the validity of the gravitational field equations leads to the conclusion that the matter fields *cannot be other than dynamical*. (In some cases, the response equation implies, without further ado, the full equations of motion of the matter fields. This is so for perfect fluids and hence for dust. The somewhat more complicated situation for more general sources has been the subject of considerable study[24]; the details need not concern us here.) But this conclusion depends of course on the appearance of $g^{\mu\nu}$ (and possibly its derivatives, if we overlook the minimal coupling requirement) in the matter part of the action. And again it is the requirement of invariance

[22] It is remarkable that when he arrived at the field equations in 1915, Einstein was unaware of the contracted Bianchi identity and viewed $T^{\sigma}_{\alpha;\sigma} = 0$ as a constraint on the equations. See Pais (1982), p. 236. The suggestion made by Ehlers (see above and footnote 15) is then in the spirit of Einstein's 1915 interpretation of his field equations.
[23] The derivation above of the response equation is a special case of what we elsewhere call the Weyl strategy (Brading and Brown 2001). For a treatment of its origins in Weyl's 1918 unified field theory, see Brading (2002).

or quasi-invariance that makes it at best difficult to construct this part purely out of the matter field variables—as Anderson stressed, the fact that "there are no free particles in general relativity" is tied up with general covariance.[25]

VI. General Covariance and the Boundary theorem.

We turn our attention now to another theorem in lagrangian dynamics whose roots go back to Noether herself, and in particular to the contemporary work of F. Klein (1918), who was exploring the general covariance properties of GR with Noether's assistance. In its generalised form we call this little-known result the Boundary theorem.[26]

Both Noether's second theorem and the Boundary theorem apply to any action invariant up to a surface term under a local transformation of the dependent and/or independent variables, where, to repeat, by local transformations we mean transformations that depend on arbitrary functions of space and time. The general solution to this variational problem is the vanishing of a certain integral, where that integral consists of two terms, a 'bulk' or interior term (depending on values of the fields in the interior of the region of integration), and a boundary or surface term. Since these functions are arbitrary, we must allow for the possibility of their vanishing on the boundary. This means that the interior and surface contributions to the general solution must vanish independently, and the vanishing of the interior contribution leads to Noether's second theorem. The Boundary theorem follows from the vanishing of the boundary contribution, and it leads to *three identities*.

Now let us imagine again an action that contains a pure gravitational part (depending on $g^{\mu\nu}$ and its derivatives) and a matter part (depending on both $g^{\mu\nu}$ and the

[24] Excellent surveys of this issue are found in Ehlers (1973) and Torretti (1983), section 5.8.
[25] Anderson (1967), p. 438.
[26] This or related results have appeared in various places in the literature since 1918, apparently largely independently. We draw special attention to Utiyama (1956, 1959), containing (we believe) the first general treatment . A detailed discussion of the Boundary theorem is found in Brading and Brown (2001).

fields, and their derivatives), *the 'double invariance' condition of the previous section being assumed to hold*. The tensor $T_{\mu\nu}$ is as usual defined in terms of the variational derivative of the matter lagrangian density in respect of $g^{\mu\nu}$. The second identity of the Boundary theorem, arising out of the quasi-invariance of the matter action alone, takes the form

$$\sqrt{-g}T_\alpha^\sigma + b_{i\alpha}^\sigma E^i = j_\alpha^\sigma - \partial_\rho \left\{ b_\alpha^{\mu\nu\sigma} \frac{\partial L}{\partial g^{\mu\nu}{}_{,\rho}} + b_{i\alpha}^\sigma \frac{\partial L}{\partial \phi_{i,\rho}} \right\}, \tag{27}$$

where j_α^σ is the Noether current familiar from Noether's first theorem (associated with the "canonical" stress-energy tensor), and as above the two *b*-coefficients depend on the form of the Lie drags of the ϕ_i and $g^{\mu\nu}$ fields. We see immediately that when the equations of motion for *all* the matter fields are satisfied ($E^i = 0$), *whatever those fields may be*, $\sqrt{-g}T_\alpha^\sigma$ differs from j_α^σ by a divergence term. Now from the third identity of the Boundary theorem, which is an antisymmetrisation condition, we can derive:

$$\partial_\sigma \partial_\rho \left\{ b_\alpha^{\mu\nu\sigma} \frac{\partial L}{\partial g^{\mu\nu}{}_{,\rho}} + b_{i\alpha}^\sigma \frac{\partial L}{\partial \phi_{i,\rho}} \right\} = 0, \tag{28}$$

so the divergence of j_α^σ vanishes if and only if the divergence of $\sqrt{-g}T_\alpha^\sigma$ does. (Of course, in the case of flat spacetime and inertial coordinates, we may replace "$\sqrt{-g}T_\alpha^\sigma$" by "T_α^σ" in the last sentence.) Returning to the particular case of S_{EM} discussed in section II, we recover (10) from (27), using Maxwell equations (2) and the fact that the second *b*-coefficient takes the form $\delta_\mu^\sigma A_\alpha$.

There is another remarkable consequence of the third (antisymmetrisation) identity related to the Boundary theorem which was effectively pointed out by Hilbert in

1915[27] and noted more recently in Barbashov and Nesterenko (1983). Suppose the lagrangian density L_M for matter is a functional of a vector field B_μ, its first derivatives and $g^{\mu\nu}$, and suppose furthermore that the matter-related part of the action is *strictly* invariant under arbitrary diffeomorphisms. Then the third identity yields

$$\frac{\partial L_M}{\partial B_{\mu,\nu}} + \frac{\partial L_M}{\partial B_{\nu,\mu}} = 0\,. \qquad (29)$$

This implies that the derivatives of B_μ can only appear in L_M in the combination $B_{\mu,\nu} - B_{\nu,\mu}$. Barbashov and Nesterenko stressed that this requirement is natural given the requirement of general covariance and the fact that the tensor $B_{\mu,\nu} - B_{\nu,\mu}$ is unaffected when the derivatives therein are replaced by covariant ones. But suppose that we introduce the further requirement that L_M be a functional only of $g^{\mu\nu}$ and the first derivatives of B_μ, and not of B_μ. *Then the covariance of the matter equations of motion under the local gauge transformation* $B_\mu \to B_\mu{}' = B_\mu + \partial_\mu \theta$, for an arbitrary scalar field θ, *is assured*. Under certain conditions then, we see the emergence of a connection between general covariance and local gauge symmetry. Indeed, the simplest invariant action consistent with these conditions is arguably (1)!

VII. Einstein's struggle with general covariance

Einstein's own struggle with general covariance had two main components. The first was *overcoming the threat of underdetermination* which arose in the hole argument—an episode well-rehearsed in the historical and philosophical literature. We will say no more about it except to mention, a little later, that its solution seems to have provided for Einstein not just the removal of a conceptual obstacle related to general covariance but a

[27] Hilbert (1915). In this connection, see the useful discussion in Sauer (1999), section 3.3.

positive, extra motivation for the principle. The second component was *overcoming the challenge of Kretschmann*, and specifically the vacuity charge. It is well-known that Einstein (1918) addressed this issue, and argued that there is good reason why general covariance had proved to have "considerable heuristic force" in his own work on gravitation. This reason has to do, once again, with an interpretation of the principle that transcends Kretschmann's concerns. The principle for Einstein, which he reiterated decades later in his Autobiographical Notes[28], was not just that a theory should have a coordinate general formulation, but that it be such that *this* formulation is the "simplest and most transparent" one available to it.

There has been debate in the literature as to precisely what Einstein meant here.[29] When he proceeds to cite the case of Newtonian mechanics and gravitation as being ruled out "practically if not theoretically" by this principle, it seems that the damage is being caused in the theory by the absolute nature of the flat affine structure of spacetime as well as of the metric structure of both space and time, all of which allows for significant simplification of the dynamical description when restricted to global inertial coordinate systems. Was then Einstein essentially ruling out the existence of absolute objects—entities which act on other objects but which are not acted back on—of a kind that would open up the possibility of preferred coordinate systems relative to which some or all the laws of physics would take an especially simple form? Explicit rejection of absolute objects would certainly become a feature of his 1920s writings.[30] If this was the core of Einstein's response to Kretschmann in 1918, it was essentially an anticipation of the view of Anderson and Trautman that has already been referred to in this paper, and that was also defended in the texts by Misner, Thorne and Wheeler, by Wald and by Ohanian and

[28] Einstein (1970), p. 69.
[29] For a helpful discussion see Norton (1993), sections 5.2 and especially 5.5.
[30] See particularly Einstein (1924), pp. 15-16 of the English translation. For further discussion of Einstein's commitment to the action-reaction principle, see Anandan and Brown (1995). Ohanian and Ruffini (1994), p. 374, make a useful distinction between absolute entities which vary under coordinate transformations and those that don't—it being only the former that are ruled out.

Ruffini.[31] (J. Barbour, who extolls Einstein's reply to Kretschmann[32], has made a compelling case for the *prima facie* startling claim that Newtonian theory itself is, when formulated generally covariantly, several distinct theories, depending on the constant numerical values assigned to the energy and angular momentum of the entire universe. These different theories have differing degrees of simplicity, the greatest simplification by far being obtained when the mentioned constants are zero. It is however far from clear to us whether even this version of Newtonian theory —whose conceptual merits were illuminated so brilliantly by Barbour and Bertotti (1982)[33]—would have satisfied Einstein's 1918 principle above.)

But there may have been more to Einstein's reasoning than this. Is it conceivable that a violation of the simplicity and transparency criterion might also be caused by something other than the existence of absolute objects? Einstein appears to have thought so until late 1916.

We mentioned in section III that the first abbreviated version of the Γ-Γ action that Einstein proposed for gravity was defined relative to those special coordinate systems for which $\sqrt{-g} = 1$. In his important 1916 review paper, Einstein promised an "important simplification of the laws of nature" produced by the choice of these coordinates, and towards the end of the paper he considered that he had indeed achieved a "considerable simplification of the formulae and calculations", and all in a manner consistent with the principle of general covariance.[34] *Had Einstein stuck to this line, he could not have consistently answered Kretschmann in the way that he did in 1918.* But he did not stick to it. In November 1916, Einstein wrote to Weyl saying he had come to the view that no restrictions on the choice of coordinates should be used in the action principle approach to GR.[35] The reasons he gave were ease of calculation and the transparency of the connection between general covariance and the conservation laws in

[31] See Misner, Thorne and Wheeler (1973), section 12.9; Wald (1984) p. 57; Ohanian and Ruffini (1994), section 7.1.
[32] See Barbour (2001).
[33] For a recent philosophical analysis of this paper, see Pooley and Brown (2002).
[34] See Einstein (1916a), pp. 130, 156 of the English translation.
[35] See Einstein (1998).

GR. This switch of thinking on Einstein's part deserves more analysis[36], as it suggests that the criteria of simplicity and transparency that he used in his 1918 reply to Kretschmann may have been more subtle than is commonly thought.[37]

VIII. General covariance as a gauge-type symmetry

Our final remarks, however, are stimulated by other aspects of Einstein's treatment of the principle of general covariance in the mentioned 1916 paper. First, any reader must be struck by the multiplicity of reasons that Einstein adduces in favour of the principle. He cites Mach's principle and the weak equivalence principle in section 2, the non-operational significance of coordinate differences for rotating frames as well as the coordinate-independence of physical happenings (the 'point-coincidence' argument familiar from his treatment of the hole problem) in section 3—all in justification of general covariance. Einstein's instinctive feel for the importance of the principle was still clearly outstripping his ability to articulate its fundamental motivation. Secondly, despite the multiplicity of arguments, it appears that Einstein is consistently viewing the principle as an "extension" of the relativity principle shared (as he correctly says) between classical mechanics and the special theory of relativity (SR). This view in particular has attracted considerable criticism—even, later, from Einstein himself.[38]

One such critic was Roberto Torretti, who in his monumental *Relativity and Geometry* stressed that there are "considerable differences of meaning and motivation"

---

[36] Some useful remarks concerning Einstein's use of coordinates satisfying the above determinant condition are found in Janssen (1997), pp. li-lii.

[37] Einstein's initial flirtation with privileged, simplifying coordinate systems in GR is a precursor of V. Fock's defence, forty years later, of the special status of what he called "harmonic" coordinate systems, i.e. those satisfying the *de Donder* relations $\left(\sqrt{-g}g^{\mu\nu}\right)_{,\mu} = 0$ and relative to which the Einstein field equations take the "reduced" form (c.f. Wald 1984, p. 261). Details of Fock's argument and its critical reception are found in Norton 1993, section 9.1. This issue deserves more discussion in the philosophical literature.

[38] Torretti cites a draft letter of Einstein to Sommerfeld of 1926 to this effect (Torretti 1983, note 4 , p. 316).

separating general covariance from the relativity principle. Torretti argued that the former can not be a constraint on theories in the manner of Lorentz covariance, on a number of technical grounds. But his principle argument was that the inertial coordinate systems involved in SR have a “fixed metrical meaning” and the invariance group of transformations is “a representation of the group of motions of the underlying flat spacetime”. The structure of the set of spacetime coordinate transformations involved in general covariance, on the other hand,

> betokens only the differentiable structure of the spacetime manifold, but has nothing to say about the spacetime metric or its group of motions (which if the Riemann tensor varies freely from point to point, is probably none other than the trivial group, consisting of the identity alone).[39]

Now it is surely a sign of the confusing nature of this whole issue that Anderson used essentially the same point to argue in favour of the view that general covariance is a symmetry in the same mould as Lorentz covariance! In SR, Lorentz transformations are 'symmetries' for Anderson in the sense that they preserve the absolute spacetime structure—the Minkowski metric. As the absolute structure is removed in the transition to GR, the symmetry group defined in this way now coincides with the covariance group.[40] (This accounts for Anderson referring to the principle as "general invariance", rather than "general covariance"[41] ; it has of course gone beyond mere coordinate generality, as we have seen.) Again, what is going on here is essentially the same as what part of Einstein's reply to Kretschmann may have been: just as the relativity principle prohibits the existence of a proper subset of the inertial frames in SR relative to which the fundamental laws of the non-gravitational interactions take an especially simple form, so the absence of absolute objects in GR prohibits the existence of privileged simplifying coordinate systems (locally inertial or otherwise) for gravitational physics.

But our instincts are more with Torretti. Our reason this time is not connected with the Noether theorems. If anything, these treat Lorentz transformations in SR and

---

[39] Torretti (1983), p. 154.
[40] See Anderson (1967), p. 87.
[41] See Anderson (1967) section 10-3. Anderson’s terminology is also used by Trautman (1966) and Ohanian and Ruffini (1994), p. 374.

diffeomorphisms in GR essentially on a par: as Noether symmetries of a given action. But before we go on, a word of clarification is in order.

We should not be comparing general covariance in GR with the relativity principle, say, in SR with its flat spacetime. We should be comparing general covariance with the familiar symmetries that emerge in the *special relativistic limit* of GR, i.e. the local structure of GR that is perfectly consistent with the existence of curvature. In this case, we immediately see a difference between the 'symmetries' under discussion. General covariance is an exact symmetry of gravitational physics; the traditional spacetime symmetries like the relativity principle are not. They hold approximately: more specifically they are concerned with the form of the laws of the non-gravitational interactions appropriate to 'small' regions of spacetime in which curvature can be neglected. But *this* is not the distinction we are after.

What in our view is lacking in many of the discussions of general covariance is recognition of a fact that does not appear directly in the relevant mathematical analysis. The essential difference between gauge-type symmetries (of which general covariance is an instance) and the usual continuous symmetries associated with the tangent space structure of spacetime (such as the relativity principle and the homogeneity of space and time) is that only the latter *have an active interpretation in terms of isolated subsystems of the universe*.

We hasten to clarify that the literature on the hole argument is full of reference to "active" diffeomorphisms, where it is the arrangement of the fields on the spacetime manifold that is altered, rather than the assignment of coordinate labels to events. Indeed the hole argument can hardly be formulated without such a notion. But the fundamental *lesson* of the hole argument is that this notion of "active" is purely mathematical.[42] What we have in mind, in contrast, is in the spirit of the Galilean ship experiment, where a laboratory is physically boosted in relation to some fixed part of the environment. So one essential aspect of this experiment is that not everything in the universe is being 'dragged' by the transformation. *The result is a 'selective' transformation—but one that* *can be seen*, and which may in principle affect the form of the laws of non-gravitational physics pertaining to processes occurring in the laboratory. The other important aspect is that the relativity principle etc. involve *rigid* transformations, which when the spacetime region of interest is approximately flat are actively implementable in practice.

The case for viewing the traditional symmetries as tied up with the possibility of rigidly translating, rotating, boosting etc. *isolated* 'laboratories' —proper *sub*systems of the universe—containing further subsystems undergoing mutual interaction has been made in detail elsewhere[43] so we will not elaborate further here.

Acknowledgements

We wish first to thank Prof Álvaro López Fernández for the kind invitation to contribute to this festschrift for Roberto Torretti, from whom we have learnt so much about relativity theory, and whose erudite work in the philosophy of physics generally we much admire.

We thank also Julian Barbour for his generous help in translating Klein's 1918 paper, and Thomas Ryckman for translating and discussing sections of Hilbert's 1915 paper. We are grateful to Peter Holland for stimulating discussions, and indebted to Michel Janssen and David Wallace for useful critical remarks on an earlier draft of this paper. Finally, special thanks go to George Svetlichny who engaged in a lengthy and illuminating correspondence on matters Noetherian with one of us (HRB).

Bibliography

Anandan, J.S.(1996), *Reality and geometry in quantum theory*, D.Phil. thesis, University of Oxford.

[42] As Anandan aptly put it, the solution of the hole argument "abolishes the distinction between passive and active general covariance"; c.f. Anandan (1996), p. 14.
[43] See Brown and Sypel (1995) and Budden (1997).

Anandan, J.S. and Brown, H.R. (1995), "On the reality of spacetime structure and the wavefunction", *Found. Phys*. **25**, 349-360.

Anderson, J.L. (1964), "Relativity Principles and the Role of Coordinates in Physics", in *Gravitation and Relativity*, H.-Y. Chui and W.F. Hoffmann (eds.) (Benjamin, New York); pp. 175-194.

(1966), "Maximal Covariance Conditions and Kretschmann's Relativity Group", in *Perspectives in Geometry and Relativity*, B. Hoffman (ed.) (Indiana University Press, Bloomington and London), pp. 16-27.

(1967), *Principles of Relativity Physics* (Academic Press, New York and London).

Barbashov, B.M. and Nesterenko, V.V. (1983), "Continuous Symmetries in Field Theory", *Fortschr. Phys*. **31**, 535-567.

Barbour, J.B. (2001), "On general covariance and best matching", in *Physics meets philosophy at the Planck scale*, C. Callender and N. Huggett (eds.) (Cambridge University Press, New York); pp. 199-212.

Barbour, J.B. and Bertotti, B. (1982), "Mach's Principle and the Structure of Dynamical Theories*", Proc. Roy. Soc. Lond. Ser*. **A382**, 295-306.

Blagojevic, M. (2002), *Gravitation and Gauge Symmetries* (Institute of Physics Publications, Bristol, UK)

Brading, K.A. (2002), "Which symmetry? Noether, Weyl, and the conservation of electric charge", to appear in *Stud. Hist. Phil. Mod. Phys*.

Brading, K.A. and Brown, H.R. (2001), "Noether's theorems, gauge symmetries and general relativity", in preparation.

Brans, C.H. and Dicke, R.H. (1961), "Mach's principle and a relativistic theory of gravitation", *Phys Rev* **124**, 925-935.

Brown, H.R. and Sypel, R. (1995), "On the meaning of the relativity principle and other symmetries", *Int. Stud. Phil. Sci*. **9**, 235-253.

Budden, T. (1997), "Galileo's Ship and Spacetime Symmetry", *Brit. J. Phil. Sci*. **48**, 483-516.

Dirac, P.A.M. (1996), *General Theory of Relativity* (Princeton University Press, Princeton).

Doughty, N.A. (1990), *Lagrangian Interaction* (Addison-Wesley, Singapore).

Ehlers, J. (1973), "Survey of General Relativity Theory", in *Relativity, Astrophysics and Cosmology*, W. Israel (ed.)(Reidel, Dordrecht).

Einstein, A. (1915), "Zur allgemeinen Relativitätstheorie.", *König. Preuss. Ak. Wiss. (Berlin). Sitzungsber*. pt 2, pp. 778-786.

(1916a), "Die Grundlage der allgemeinen Relativitatstheorie", *Ann. Phys*. 49, 769-822. An English translation entitled "The Foundation of the General Theory of Relativity" is found in Lorentz *et al*. (1952), pp. 111-164.

(1916b), "Hamiltonsches Prinzip und allgemeine Relativitätstheorie", *König. Preuss. Ak. Wiss. (Berlin). Sitzungsber*. pt 2, pp. 1111-1116. An English translation entitled "Hamilton's Principle and the General Theory of Relativity" is found in Lorentz *et al*. (1952), pp. 167-173.

(1918), "Prinzipielles zur allgemeinen Relativitätstheorie", *Ann. Phys*. **55**, 241-244.

(1924), "Über den Äther*", Schweizerische naturforschende Gesellschaft, Verhanflungen* **105**, 85-93. An English translation is found in Saunders and Brown (1991), 13-20.

(1970), "Autobiographical Notes", in P.A. Schilpp (Ed.) *Albert Einstein: Philosopher-Scientist* (Open Court, La Salle, Illinois).

(1998) Document 278 in *The Collected Papers of Albert Einstein, volume 8, The Berlin Years: Correspondence 1914-1918*, R. Schulmann *et al.* (eds.) English translation, A.M. Hentschel (translator), K. Hentschel (consultant), (Princeton University Press, Princeton, New Jersey).

Grigore, D.R. (1992), "The derivation of Einstein equations from invariance principles", *Class. Quantum Grav*. **9**, 1555-1571.

Hilbert, D. (1915), "Die Grundlagen der Physik (Erste Mitteilung)", *Göttinger Nachrichten, Math.-phys. Kl*. 395-407.

Hill, E.L. (1951), "Hamilton's Principle and the Conservation Theorems of Mathematical Physics", *Rev. Mod. Phys*. **23**, 253-260.

Janssen, M. (1997), "Introduction to Volume 8", in *The Collected Papers of Albert Einstein, volume 8, The Berlin Years: Correspondence 1914-1918*, R. Schulmann *et al.* (eds.), (Princeton University Press, Princeton, New Jersey, 1997).

Klein, F. (1918), "Über die Differentialgesetze für die Erhaltung von Impuls und Energie in der Einsteinshcen Gravitationstheorie", *Göttinger Nachrichten, Math.-phys. Kl*., H. 2, S. 171-189.

Kretschmann, E. (1917), "Über den physikalischen Sinn der Relativitätspostulate, A. Einstein neue und seine ursprungliche Relativitätstheorie", *Ann. Phys*. **53**, 575-614.

Lorentz H.A.L. *et al*. (1952), *The Principle of Relativity* (Dover, New York).

Lovelock, D. (1969), "The uniqueness of Einstein field equations in a four-dimensional space", *Arch. Rat. Mech. Anal*. **33**, 54-70.

Misner, C.W., Thorne, K.S. and Wheeler, J.A. (1973) *Gravitation* (Freeman, San Francisco).

Noether, E. (1918), "Invariante Variationsprobleme*", Göttinger Nachrichten, Math.-phys. Kl*, 235-257.

Norton, J. (1993) "General covariance and the foundations of general relativity: eight decades of dispute", *Rep. Prog. Phys*. **56**, 791-858.

(2001), "General Covariance, Gauge Theories and the Kretschmann Objection"; e-print archive PITT-PHIL-SCI00000380.

Ohanian, H. and Ruffini, R. (1994), *Gravitation and Spacetime*, second edition (W.W. Norton, New York).

Pais, A. (1982), *'Subtle is the Lord …'. The Science and the Life of Albert Einstein* (Oxford University Press, New York).

Pooley, O. and Brown, H.R. (2002), "Relationism rehabilitated? I: classical mechanics", to appear in *Brit. J. Phil. Sci*. E-print: PITT-PHIL-SCI00000220.

Rowe, D. (1999), "The Göttingen Response to general relativity and Emmy Noether's theorems", in *The Symbolic Universe: Geometry and Physics 1890-1930*, J. Grey (ed.) (Oxford University Press, Oxford); pp. 189-233.

Rindler, W. (1977), *Essential Relativity*, 2nd edtion (Wiley, New York).

Rynasiewicz, R. (1999), "Kretschmann's Analysis of Covariance and Relativity Principles", in *The Expanding Worlds of General Relativity (Einstein Studies, volume 7)*, H. Goenner, J. Renn, J. Ritter, and T. Sauer (eds.) (The Centre for Einstein Studies); pp. 431-462.

Saunders, S. (2001), "Indiscernibles, General Covariance, and Other Symmetries", to appear in *Revisiting the Foundations of Relativistic Physics: Festschrift in Honour of John Stachel*, A. Ashtekar, D. Howard, J. Renn, S. Sarkar, and A. Shimony, (eds.), (Kluwer).

Saunders, S. and Brown, H.R. (1991), *The Philosophy of Vacuum* (Clarendon Press, Oxford).

Sauer, T. (1999), “The Relativity of Discovery: Hilbert’s First Note on the Foundations of Physics”, *Arch. Hist. Exact Sci.* **53**, 529-575.

Sorkin, R. (2001), "An example relevant to the Kretschmann-Einstein debate", e-print archive PITT-PHIL-SCI00000371.

Svetlichny, G. (2001), "Some Noether Matters", manuscript.

Tavel, M.A. (1971), "Noether's theorem", *Transport Theory and Statistical Physics* **1**, 183-207.

Torretti, R. (1983*), Relativity and Geometry*, Pergamon Press, Oxford. A Dover (New York) paperback edition was published in 1996.

Trautmann, A. (1962), "Conservation Laws in General Relativity", in *Gravitation: An Introduction to Current Research*, L. Witten (ed.) (John Wiley and Sons).

(1966), "The General Theory of Relativity", *Sov. Phys. Uspekhi* **89**, 319-336.

Utiyama, R. (1956) "Invariant Theoretical Interpretation of Interaction", *Phys. Rev.* **101**, 1597-1607.

(1959), "Theory of Invariant Variation and the Generalized Canonical Dynamics", *Prog. Theor. Phys. Suppl*. **9**, 19-44.

Wald, R. (1984), *General Relativity* (University of Chicago Press, Chicago).