Zeitschrift: Helvetica Physica Acta
Band: 48 (1975)
Heft: 2

Artikel: Remarks on the foundations of general relativity
Autor: Pietenpol, J.L. / Speiser, D.
DOI: https://doi.org/10.5169/seals-114667

Helvetica Physica Acta
Vol. 48, 1975.

Birkhäuser Verlag Basel

Remarks on the Foundations of General Relativity

by **J. L. Pietenpol** and **D. Speiser**

Université Catholique de Louvain, Institut de Physique Théorique, 2, Chemin du Cyclotron, 1348 Louvain-la-Neuve, Belgium

To E. C. G. Stueckelberg, physicist and teacher, for his 70th birthday

(20. XII. 74)

Abstract. A new way of introducing general relativity is presented, with emphasis on the physical meaning of the formalism and on its relation to classical mechanics. We also underscore the crucial role of the experimental tests by pointing out the existence of two inequivalent but equally simple theories, both of which satisfy the principles of general relativity.

1. Introduction

General relativity is still regarded by many physicists with reticence. This is probably due mainly to the extraordinary mathematical apparatus which seems to be needed, but also to the fact that such notions as affine connection and curvature seem remote from physics. Moreover, the impression is often given that, from a few vaguely worded principles, the whole theory is directly deduced and must then be accepted or rejected as a whole.

This paper has several purposes. First, we want to show that the formalism used in general relativity arises in fact in a very natural way in classical mechanics and that the Christoffel symbols and Riemann tensor have a direct physical meaning. It seems to us that this way to general relativity is the shortest one possible, and, in a way, also the one closest to experiment.

Second, we try to clarify the role played by the 'general principles' such as the principle of equivalence and general covariance on one hand, and the experiments on the other. Contrary to what is often stated, these principles *do not* determine general relativity uniquely even if we add simplicity postulates. There are, as we shall show, (at least) two equally simple theories which satisfy these principles, neither of which is a special case of the other and between which only experiment can decide.

The 'scalar theory' discussed below was proposed by Nordström [1] in the form of a special-relativistic potential theory. A first step toward its generally covariant formulation was made by Einstein and Fokker [2], but they lacked the notion of the Weyl tensor, which had not yet been developed by its author. Thus Pauli could call this theory 'artificial and complicated' [3] in comparison with Einstein's. His point would be well taken if one did not regard the metric as fundamental [4] but instead based the theory on the affine connection, for then the tensor $R_{\mu\nu}$ is irreducible, and the vacuum field equation of Einstein's theory (when expressed in the form $R_{\mu\nu} = 0$) is meaningful in the absence of a metric. But if the metric exists, then $R_{\mu\nu}$ reduces into a trace (the scalar curvature) and a traceless part, and there is no *a priori* reason why these two parts should be combined in any particular way in the equations of the theory. (As Pauli expressed it: 'Let no man join together what ...') The existence of reproducible standards of length (or time) in nature seems to point to the fundamental role of the metric. These standards owe their existence to phenomena outside the domain of classical physics, namely quantum mechanics and the atomistic structure of matter.

The discussion contains a criticism of the use often made of the notion 'equivalence principle'. We maintain that this principle has no precise meaning, as long as it remains a word or a sentence and is not supplemented by a formula. As an example we write two equations both containing an (active) equivalence principle but leading to different consequences. We emphasize that this discussion is not concerned with the distinction between active and passive equivalence principles.

2. Linear Relativity

Einstein's laws of mechanics (special relativity) are based on two principles:

a) Two observers describe a process by the same equations, if they are at rest or in uniform linear motion with respect to each other.
b) The velocity of light in vacuum is the same in all inertial systems.

The first of these has been aptly named the 'principle of linear relativity' by the mathematician O. Blumenthal [5]. Mathematically, it says that space-time has a linear affine structure. The second principle says that space-time possesses a conformal structure. As was recognized by Minkowski, both statements together imply that space-time is endowed with a pseudo-Euclidean metric, now usually written in the form:

$$l^2(x) = \eta_{\mu\nu} x^\mu x^\nu$$

where

$$\eta_{\mu\nu} \equiv \begin{pmatrix} 1 & & & \\ & -1 & & \\ & & -1 & \\ & & & -1 \end{pmatrix}$$

is the *constant* metric tensor of space-time. The isometry group of space-time, the Poincaré group, is then the group of transformations relating all inertial frames.

Both statements in this formulation contain an assumption, not always stated explicitly, that inertial frames exist, or at least that such systems can be approximated as closely as one wishes.

In analogy to Newton's first two axioms, Einstein's axioms of mechanics may then be formulated as follows:

E_{I}: In the absence of forces, the four-momentum of a mass-point remains constant:

$$p_\mu = c_\mu \quad \mu = 0, \dots 3 \tag{1}$$

where $p_\mu \equiv m_0 d_\tau x_\mu$ so that $p_\mu p^\mu = m_0^2 c^4$. I.e., in the absence of forces the point will move on a straight line.

E_{II}: Under the influence of forces, the rate of change of four-momentum with respect to proper time is equal to the four-force:

$$d_\tau p_\mu = F_\mu. \tag{2}$$

We emphasize that both equations, though written in coordinate form, are *vector equations*. They express relations between vectors in a linear space, namely Minkowski's space-time. Einstein's axioms E_I and E_{II} as we have written them are valid only in inertial frames. Only in inertial frames (and even then not always) can a sharp distinction be made between inertia (inertial force) and force. The Poincaré transformations which connect these frames are therefore distinguished from all other transformations by physical, i.e. *measurable*, facts.

3. Einstein's Axioms in Generalized Form

It is, however, important to ask: how can Einstein's axioms be formulated so as to be valid for all (differentiable) coordinate systems?

We shall answer this question first for the second axiom, where the answer is well known, in principle, but rarely fully pursued. For the force term it suffices to note that the differential $dA \equiv F_\mu dx^\mu$ is an invariant, and thus we may define 'generalized forces' $Q_\mu = Q_\mu(q^\nu)$ by writing

$$dA = Q_\mu(q^\nu)\, dq^\mu.$$

For the left-hand side the easiest way is to use a variational principle. One notices that the principle

$$\delta\left[m_0 \int \sqrt{\eta_{\mu\nu}\, dx^\mu dx^\nu}\right] = 0 \tag{3}$$

leads to the following Euler equation:

$$m_0 \ddot{q}^\mu = 0$$

where a dot denotes the derivative with respect to the *proper* time. If we define a metric tensor $g_{\mu\nu}(q)$ in such a way that the line element is invariant:

$$g_{\mu\nu}\, dq^\mu dq^\nu = \eta_{\mu\nu}\, dx^\mu dx^\nu,$$

then in an arbitrary coordinate system the variational principle takes the form

$$m_0 \delta \int \sqrt{g_{\mu\nu}(q)\, dq^\mu dq^\nu} = 0. \tag{4}$$

The fact that the metric components are now functions of the coordinates gives rise to additional terms in the Euler equation. One finds

$$m_0[g_{\nu\lambda}\ddot{q}^\lambda + \tfrac{1}{2}(g_{\rho\nu,\sigma} + g_{\sigma\nu,\rho} - g_{\rho\sigma,\nu})\dot{q}^\rho \dot{q}^\sigma] = 0$$

where a comma preceding an index denotes the partial derivative with respect to the corresponding coordinate.

In terms of the Christoffel symbol

$$\Gamma^\lambda_{\rho\sigma} \equiv \tfrac{1}{2} g^{\lambda\tau}(g_{\rho\tau,\sigma} + g_{\sigma\tau,\rho} - g_{\rho\sigma,\tau})$$

where

$$g^{\mu\nu} g_{\nu\lambda} = \delta^\mu_\lambda,$$

the Euler equation becomes

$$m_0(\ddot{q}^\mu + \Gamma^\mu_{\rho\sigma}\dot{q}^\rho\dot{q}^\sigma) = 0, \tag{5}$$

or, in the presence of forces:

$$m_0(\ddot{q}^\mu + \Gamma^\mu_{\rho\sigma}\dot{q}^\rho q^\sigma) = Q^\mu(E^g_{\text{II}}). \tag{6}$$

Thus the Γ's have made their appearance without any recourse to differential geometry.

Now, what is the significance of this second term on the left-hand side? This term represents what are called 'apparent forces', i.e., manifestations of inertia due to the use of non-Cartesian frames. The Γ's (or Christoffel symbols) are thus shown to have an important physical meaning. Written in this way, E^g_{II} is valid for all (differentiable) coordinate systems.

We now turn to the first axiom. E_{I} looks as if it were only a special case of E^g_{II}, i.e., the case where no force is present. But in fact it is much more. E_{I} is a physical statement about space-time. It says that, in the absence of any influence from some observable external system, the mass-point will move uniformly on a straight line. This implies that such straight lines exist and that the observer can recognize them. In other words, it says that space-time is Minkowskian and that it is possible to introduce a Cartesian coordinate system. This is indeed a *physical* statement. Its mathematical, invariant formulation, according to Riemann, is the vanishing of the curvature tensor at all points:

$$R^\lambda_{\gamma\rho\sigma} \equiv 0. \tag{7}$$

4. Gravitation

We now turn to gravitation. General relativity is a formulation of the gravitational interaction, compatible with Einstein's principles and such that the equivalence principle is automatically incorporated. This is admittedly only one aspect of general relativity but indeed the most tangible one. We therefore begin by considering this 'principle'.

In Newtonian mechanics it says that the 'heavy mass', i.e. the passive gravitational charge, is proportional to the inertial mass:

$$m_g = \text{constant} \cdot m_{\text{inertial}}$$

so that, for a mass-point subject only to the gravitational force, we have

$$m\ddot{\vec{x}} = \text{constant} \cdot m \cdot \vec{f}(\vec{x}),$$

and the inertial mass drops out of the equation. This is sometimes referred to as the *passive* aspect of the equivalence principle.

Newton's third axiom ('action = reaction') combined with this principle then demands that the force which a body exerts on another one also be proportional to its inertial mass. Thus,

$$m_1\ddot{\vec{x}}_1 = \text{constant} \cdot m_1 m_2 \vec{f}(|\vec{x}_1 - \vec{x}_2|).$$

This is sometimes referred to as the *active* aspect of the equivalence principle.

In the framework of special relativity it is indicated to conceive of the forces as transmitted by a field, as Einstein emphasized. The problem then splits up into the three questions:

1) What is this field and, in particular, what are its transformation properties?
2) How does a mass-point behave under the influence of the field?
3) How is the gravitational field excited, i.e., what are its sources and how is the field coupled to them?

To answer the first question, one observes that the equivalence principle says in effect that gravitation has locally all the characteristics of an apparent force: it can be transformed away by a suitable choice of coordinates. This is illustrated by Einstein's well-known elevator Gedanken experiment. This observation suggests that the gravitational field be described precisely by the $\Gamma^{\lambda}_{\mu\nu}$, which have the property that in a suitable coordinate system they do vanish at a point, as the Gedanken experiment requires. Of course, this last property would still hold if we added some tensor $t^{\lambda}_{\mu\nu} = t^{\lambda}_{\nu\mu}$ to the field: $\Gamma^{\lambda}_{\mu\nu} \rightarrow \Gamma^{\lambda}_{\mu\nu} + t^{\lambda}_{\mu\nu}$. But then the gravitational field would lose its direct connection with the metric. We note explicitly that the use of the Γs for describing gravitation is motivated by two different ideas:

1) Gravity has the characteristics of an apparent force. The field quantities which describe gravity must be removable in at least one coordinate system.
2) The field should be expressible in terms of the metric alone.

The equation of motion, i.e. the answer to question 2, is then simply

$$\ddot{q}^{\lambda} + \Gamma^{\lambda}_{\mu\nu}\dot{q}^{\mu}\dot{q}^{\nu} = 0 \tag{8}$$

where the $\Gamma^{\lambda}_{\mu\nu}$ now describe the inertial-gravitational field. Thus the passive equivalence principle is automatically contained in (8).

It remains to find the field equations and to describe the source of the field. If gravity and inertia are to be one, then obviously Newton's first axiom must be given up or at least be weakened. The inertial-gravitational field must no longer be kept frozen in but must be formulated as a dynamical entity which can be excited by external sources. How can this be done?

At this point we depart from what Einstein said and follow rather what he did. First we note that the connection between time and space as accepted by linear relativity for all space-time must be kept valid *locally*. For the existence of clocks and yardsticks seems to be guaranteed by facts outside the framework of gravity-inertia, such as the Bohr-radius of atoms, etc, and light fronts appear as spheres to all locally inertial observers, moving with uniform relative velocities. This means that we accept the existence of a metric tensor. (If the $g_{\mu\nu}$ did not exist, we would have to take the $\Gamma^{\lambda}_{\mu\nu}$, i.e. the forces themselves, as our starting point.) We note that if the $g_{\mu\nu}(q)$ exist, then equation (8) says that the path followed by a point is not only a straightest one but also (locally) a longest one. In fact, it can be derived from a variational principle:

$$\delta \int d\tau = 0. \tag{9}$$

The existence of $g_{\mu\nu}$ in turn implies that in *tangent space*, Lorentz transformations are distinguished from general linear transformations, and the special-relativistic equation (7) can now be written with the help of the $g_{\mu\nu}$ as

$$R_{\mu\nu\rho\sigma} = 0. \tag{10}$$

This system of $n^2(n^2-1)/12$ (=20 if $n=4$) equations must be relaxed. Since locally the Lorentz group is still operationally well defined, the reduction of $R_{\mu\nu\rho\sigma}$ with respect to this group is also well defined. One finds:

$$R_{\mu\nu\rho\sigma} = R^W_{\mu\nu\rho\sigma} - \frac{1}{n-2}(R^0_{\mu\rho}g_{\nu\sigma} - R^0_{\nu\rho}g_{\mu\sigma} + R^0_{\nu\sigma}g_{\mu\rho} - R^0_{\mu\sigma}g_{\nu\rho})$$

$$- \frac{1}{n(n-1)} R(g_{\mu\rho}g_{\nu\sigma} - g_{\nu\rho}g_{\mu\sigma})$$

where $R^W_{\mu\nu\rho\sigma}$ is the Weyl tensor and $R^0_{\mu\nu}$ is the Ricci tensor without trace:

$$R^0_{\mu\nu} \equiv R_{\mu\nu} - \frac{1}{n} R g_{\mu\nu}.$$

In formulating a field theory of gravitation, any of these tensors may be used. In particular, to describe the field in space-time free of matter and radiation, *any* of these tensors or any combination of them might be put equal to zero. Thus, according to Einstein, for empty space we must put:

$$\begin{cases} R^0_{\mu\nu} = 0 \\ R = 0 \end{cases}$$

or simply:

$$R_{\mu\nu} = 0. \tag{11}$$

But it must be noted that a priori we could equally well put

$$\begin{cases} R^W_{\mu\nu\rho\sigma} = 0 \\ R = 0. \end{cases} \tag{12}$$

For, such a theory would also be 'generally covariant', and we shall show presently that it too incorporates an 'equivalence principle'. To show this we must determine the source of the field and write down inhomogeneous field equations.

The source of a gravitational field theory must be a covariant (scalar, vector, tensor, etc.), such that in the non-relativistic approximation we recover Poisson's equation:

$$\Delta\psi = 4\pi G\rho$$

where ψ is Newton's potential and ρ is the mass density. The 'equivalence of mass and energy' then tells us that there are (at least) two candidates for the source term of the field equation:

1) The trace of the energy-momentum tensor:

$$T \equiv g^{\mu\nu}T_{\mu\nu}.$$

2) The traceless part of the energy-momentum tensor:

$$T^0_{\mu\nu} \equiv T_{\mu\nu} - \frac{1}{n} T g_{\mu\nu}.$$

There is no suitable vector, and such a theory must be ruled out at once any way since we know from Maxwell's theory that it would lead to repulsion between heavy bodies if the field energy is to be positive. However, possibilities 1 and 2 should both be considered.

The full scalar theory reads:

$$S: \begin{cases} R^{W}_{\mu\nu\rho\sigma} = 0 \\ R = -3\kappa T \end{cases} \tag{13}$$

where $\kappa = 8\pi G/c^2$ is Einstein's constant, and the factor 3 is included so that in the non-relativistic limit we recover Poisson's equation.

For the tensor theory one may try:

$$R^{0}_{\mu\nu} = -\kappa T^{0}_{\mu\nu}. \tag{14}$$

Now, in special relativity one has the continuity equation:

$$T^{\mu\nu}_{,\nu} = 0$$

and it is indicated to generalize this for non-linear manifolds by postulating

$$T^{\mu\nu}{}_{|\nu} = 0 \tag{15}$$

where the bar denotes the covariant derivative. On the other hand, from Bianchi's identities we have

$$R^{0\mu\nu}{}_{|\nu} = \frac{n-2}{2n}(Rg^{\mu\nu})_{|\nu}.$$

Thus (14) and (15) imply

$$\frac{n-2}{2n}(Rg^{\mu\nu})_{|\nu} = \left(\kappa \frac{1}{n} Tg^{\mu\nu}\right)_{|\nu}$$

or

$$\frac{n-2}{2} R_{,\lambda} = \kappa T_{,\lambda}$$

and (for $n = 4$)

$$R - \kappa T = \text{const} \equiv 4\Lambda. \tag{16}$$

Multiplying (16) by $\frac{1}{4}g_{\mu\nu}$ and subtracting it from (14), we obtain finally

$$T: R_{\mu\nu} - \tfrac{1}{2}Rg_{\mu\nu} + \Lambda g_{\mu\nu} = -\kappa T_{\mu\nu}. \tag{17}$$

The so-called cosmological constant appears here as an integration constant, a point emphasized first by Einstein himself [6] and later by Lemaître. Our derivation parallels the one by Anderson and Finkelstein [7].

The scalar and the tensor theory both fulfil the passive equivalence principle, which says that gravity is indistinguishable from inertia up to and inclusive of first derivatives of the $g_{\mu\nu}$. Measurement of the second derivatives – i.e., a measurement of the Riemann tensor – allows such a distinction.

But the *active* equivalence principles of the scalar and tensor theories are not the same. Corresponding to the two different possible source terms, there are (at least) two different theories, both 'generally covariant' (whatever that may mean in precise mathematical language), between which only experiment can decide (see the appendix). It must be emphasized that neither of these two theories is a special case or a limiting case of the other.

In this respect the situation is somewhat analogous to the theory of weak interactions, where only experiment could decide between the possibilities S, V, T, A, P, etc.

In spite of what has been said, one may be tempted to guess at a reason why nature prefers the tensor theory and why Einstein was guided toward it. It is the same as the reason why nature seems to prefer the V–A coupling for weak interactions and why Fermi hit on the vector theory. In both cases, as in Maxwell's theory, the source satisfies a continuity equation (or almost), and as a consequence the charges are not renormalized. The deeper significance of these charge conservation laws (or almost-conservation laws) is still hidden, but the fact that we find them in three or almost four different places seems to indicate that they must be crucial for a future unification of such theories.

Acknowledgments

It is a pleasure to acknowledge discussions with R. Penzlin, W. Theis, L. Radicati and L. Michel and to acknowledge the hospitality of the Freie Universitaet (Berlin), the Scuola Normale Superiore (Pisa) and the Institut des Hautes Etudes Scientifiques (Bures-sur-Yvette).

APPENDIX

We summarize the predictions of the scalar and the tensor (Einstein's) theory. In the Eddington–Robertson expansion, the static, spherically symmetric metric is parametrized as follows:

$$ds^2 = \left(1 - 2\alpha \frac{GM}{c^2 r} + 2\beta \frac{G^2 M^2}{c^4 r^2} + \dots\right) c^2\, dt^2$$

$$- \left(1 + 2\gamma \frac{GM}{c^2 r} + \dots\right) (dr^2 + r^2\, d\theta^2 + r^2 \sin^2\theta\, d\phi^2).$$

To obtain the Newtonian limit one must put $\alpha = 1$. The red shift ('third test') is then predicted automatically. Thus, as is well known, the red shift is not really a test of the tensor theory at all.

For the constants β and γ, the scalar and tensor theories yield the values given in Table I.

Table I
Values of the Eddington–Robertson parameters for the scalar and tensor theories

	S	T
β	1/2	1
γ	−1	1

We then have the predictions listed in Table II. Here, the second column contains the combination of constants relevant to the effect (in first order), the third and fourth columns contain the corresponding numerical values for the scalar and tensor theories, and the last column gives the experimental result.

Table II
Comparison of the scalar and tensor theories with experiment

Test	Relevant combination of parameters	S	T	Experimental
Perihelion precession	$(2+2\gamma-\beta)/3$	$-1/6$	1	1.005 ± 0.02†)
Light deflection	$(1+\gamma)/2$	0	1	1.0 ± 0.1‡)
Red shift	α	1	1	1.00 ± 0.01§)
Time delay	$(1+\gamma)/2$	0	1	1.015 ± 0.05‖)
Gyroscope precession	$(1+2\gamma)/3$	$-1/3$	1	—¶)

†) I. I. Shapiro et al., Phys. Rev. Letters *28*, 1954 (1972). Value obtained by neglecting the effect of a possible solar quadrupole moment. Cf. R. H. Dicke, *The Theoretical Significance of Experimental Relativity* (Gordon and Breach, New York and London 1964), p. 25.

‡) For a review, see C. W. Misner, K. S. Thorne and J. A. Wheeler, *Gravitation* (W. H. Freeman and Company, San Francisco 1973), pp. 1103–1105.

§) Terrestrial red shift, R. V. Pound and J. L. Snider, Phys. Rev. Letters *13*, 539 (1964). A value of 1.05 ± 0.05 for the solar red shift was measured by J. Brault, Bull. Am. Phys. Soc. *8*, 28 (1963); see also R. H. Dicke, op. cit., p. 25.

‖) Test proposed by I. I. Shapiro, Phys. Rev. Letters *13*, 789 (1964). Experimental value quoted from I. I. Shapiro et al., Phys. Rev. Letters *26*, 1132 (1971).

¶) Test proposed by L. I. Schiff, Proc. Nat. Acad. Sci. *46*, 871 (1960). The experiment has not yet been performed.

References

[1] G. NORDSTRÖM, Ann. Phys. Lpz. *42*, 533 (1913); *43*, 1101 (1914).

[2] A. EINSTEIN and A. D. FOKKER, Ann. Phys. Lpz. *44*, 321 (1914).

[3] In German: 'künstlich und verwickelt'. W. PAULI, in *Encyklopädie der Mathematischen Wissenschaften*, Vol. V19 (B. G. Teubner, Leipzig 1921), p. 722. For the English translation, see Ref. 4, p. 160.

[4] But Pauli himself did consider the metric, and not the affine connection, as having a 'direct physical meaning'. Cf. W. PAULI, *Theory of Relativity* (Pergamon Press, London 1958), p. 227.

[5] 'Einstein hat das lineare Relativitätsprinzip zum allgemeinen erweitert.' O. BLUMENTHAL, *Fortschritte der Mathematischen Wissenschaften in Monographien*, Heft 2 (B. G. Teubner, Leipzig and Berlin 1922).

[6] A. EINSTEIN, Sitzungsberichte der Preussischen Akad. Wissenschaften (1919), 433. English translation in H. A. LORENTZ, A. EINSTEIN et al., *The Principle of Relativity* (Methuen and Company 1923), p. 189. Reprinted by Dover Publications.

[7] J. L. ANDERSON and D. FINKELSTEIN, Amer. J. Phys. *39*, 901 (1971).