Pergamon

*Stud. Hist. Phil. Mod. Phys.*, Vol. 32, No. 1, pp. 33–51, 2001

Printed in Great Britain
1355-2198/01 $ - see front matter

# The Principle of Equivalence

***Michel Ghins*, Tim Budden*†***

We start from John Norton's analysis (1985) of the reach of Einstein's version of the principle of equivalence which is not a local principle but an extension of the relativity principle to reference frames in constant acceleration on the background of Minkowski spacetime. We examine how such a point of view implies a profound, and not generally recognised, reconsideration of the concepts of inertial system and field in physics. We then reevaluate the role that the infinitesimal principle, if adequately formulated, can legitimately be claimed to play in general relativity. We show that what we call the 'punctual equivalence principle' has significant physical content and that it permits the derivation of the geodesic law. © 2001 Elsevier Science Ltd. All rights reserved.

*Keywords*: Einstein; Equivalence Principle; General Relativity; Gravitation; Relativity Principle; Inertia; Spacetime.

## 1. Introduction

Einstein introduced his equivalence principle in 1907; see Stachel (1979). Ever since then, this principle has been the focus of intense debate and controversy among physicists and philosophers of physics. John Norton (1985) has shown that Einstein's principle of equivalence had nothing to do with a local or infinitesimal principle but is an extension of the principle of relativity to

(*Received 16 January 2000; revised 4 September 2000*)
* Centre de Philosophie des, Sciences, Institut Supérieur de Philosophie, Place Cardinal Mercier 14, B-1348 Louvain-la-Neuve, Belgium (*e-mail*: ghins@risp.ucl.ac.be).
† Centre de Philosophie des Sciences, Institut Supérieur de Philosophie, Université Catholique de Louvain, Place du Cardinal Mercier, 14, B-1348 Louvain-la-Neuve, Belgium.

PII: S1355-2198(00)00038-1

uniformly accelerated systems in Minkowski spacetime. Thus, Minkowski spacetime, as seen from a uniformly accelerated frame, is a particular case of gravitational field. We discuss how such a point of view implies a deep, and so far underestimated, reconsideration of the notion of inertial system and the concept of field in physics. We then ponder what could be an acceptable role for an infinitesimal equivalence principle in general relativity. Various formulations are examined and evaluated with respect to their soundness and physical content. We claim that what we call the 'punctual equivalence principle' has significant physical content and that it permits the derivation of the geodesic law and imposes strong constraints on a theory of gravitation.

## 2. Einstein's Version: An Extended Relativity Principle

### *2.1. Statement of the extended relativity principle and its validity*

Norton (1985) offers us a translation of a particularly clear statement of Einstein's equivalence principle and of how it can be satisfied, which we reproduce here. First Einstein asks us to consider what he calls 'the boundary case of the special theory of relativity': in a finite region of spacetime with no gravitational field present it is possible to set up an inertial ('Galilean') frame $K$ in which force-free bodies move uniformly along straight lines. Then he goes on to define what we shall call *Einstein's equivalence principle* (EEP) as an extension of the relativity principle and suggests a 'point of view' from which it can be satisfied in some theories of gravitation due to the special nature of the gravitational field.

> Starting from this limiting case of the special theory of relativity, one can ask oneself whether an observer, uniformly accelerated relative to $K$ in the region considered, must understand his condition as accelerated, or whether there remains a point of view for him, in accord with the (approximately) known laws of nature, by which he can interpret his condition as 'at rest'. Expressed more precisely: do the laws of nature, known to a certain approximation, allow us to consider a reference system $K'$ as at rest if it is accelerated with respect to $K$? Or, somewhat more generally: can the principle of relativity be extended also to reference systems which are (uniformly) accelerated relative to one another. The answer runs: as far as we really know the laws of nature, nothing stops us from considering the system $K'$ as at rest, if we assume the presence of a gravitational field (homogeneous in the first approximation) relative to $K'$; for all bodies fall with the same acceleration independent of their physical nature in a homogeneous gravitational field as well as with respect to our system $K'$. The assumption that one may treat $K'$ as at rest in all strictness without any laws of nature not being fulfilled with respect to $K'$, I call the 'principle of equivalence' (Einstein, 1916b; translated in Norton, 1985, p. 206).

For convenience, consider a laboratory, which, as tradition dictates, we call the *elevator*, at rest in $K$ for part of its history, then at rest in $K'$ for some later part.[1] For Einstein's equivalence principle to hold, $K$ and $K'$ must be physically equivalent in just the same way that inertial frames in relative motion are in special relativity, i.e. the same laws (in their standard vectorial form), in particular the Galilean law of inertia, should hold in the earlier and later stages in the elevator. In other words, $K$ and $K'$ are both inertial charts.

At first sight this presents us with a problem: clearly very different phenomena occur in the earlier and later stages in the elevator, for example isolated bodies move uniformly in a straight line in the earlier stage but accelerate in the later stage. This constitutes a striking dissimilarity with Galileo's ship thought experiment (Galileo, 1632, Salviati, 2nd day, pp. 186–187), in which we imagine experiments conducted below decks in a ship at rest, then set sail and conduct the same experiments once the ship is moving uniformly in a straight line. If, as we assume, the standard laws of special relativity hold, we witness exactly the same phenomena in the earlier and later stages, evidence *par excellence* that all inertial motions are equivalent. Does not the diversity of phenomena occurring in the earlier and later stages of the elevator thought experiment equally undermine Einstein's conception of $K$ and $K'$ as physically equivalent? Well, no. The relativity principle only requires that the same *laws* hold in the earlier and later stages of the elevator. And this can be the case despite the diversity of *phenomena* if we allow that the *physical environment* is correspondingly diverse in $K$ and $K'$: in $K'$ there is a gravitational-inertial field present; in $K$ there is not. On another day, or in a different world, the situation could be reversed: it could be the case that there was a gravitational field in $K$, in which case exactly the same phenomena that actually occur in $K'$ would occur in $K$. Despite contingent, factual differences, $K$ and $K'$ are nomologically equivalent and can therefore be both viewed as inertial (since $K$ is inertial *ex hypothesi*).[2]

To make this possibility more convincing, it is worth considering that similar factual distinctions can hold between inertial frames in special relativity while the physical equivalence of all inertial frames reigns at the level of laws. Take electrodynamics, for example—a paradigmatic example of a special relativistic theory. In order to obtain a theory more closely analogous to the gravito-particle mechanics Einstein had in mind, add to the laws of electrodynamics, gratuitously and for philosophical purposes only, a law which implies that the Faraday tensor $F_{ab}$ cannot vanish on any open region of spacetime. For example, we could simply add this statement itself, or the statement that only positive charges exist and that all worlds contain some of them. To our minds

[1] Recall that the elevator is presumed to occupy a region of space where special relativity holds and where, when it is at rest in $K$, no gravitational field is present. So do not think of an elevator suspended in a building on Earth! If you prefer, think of a spaceship in outer space, infinitely distant from any other matter.

[2] It is perhaps worth pointing out that the extension of the relativity principle at stake here does *not* concern an equivalence between pairs of accelerated frames which judge one another to be in constant motion. Such an equivalence between non-inertial frames does not generally hold in special relativity. (However, pairs of accelerating frames which are judged to be in constant relative motion *from the perspective of an inertial coordinate system* are equivalent in special relativity.)

there is no reason to think that such additions violate the relativity principle: we certainly do not need to explicitly refer to absolute space or, for that matter, *any* frozen geometric spacetime structure at all to formulate them and the Lorentz–Poincaré symmetry group of the theory remains unchanged. Nonetheless, in any world allowed by the strengthened theory there will be factual physical distinctions to be drawn between inertial frames or laboratories. Consider, for example, what we can think of as the *boundary case*: worlds with the simplest allowed configuration of electromagnetic fields, i.e. ones in which the Faraday tensor is constant but non-zero. There will always be a frame in such a world in which the Faraday tensor is purely electric, i.e. in which the $B$ field component vanishes, a fact which can be detected by experiments with test particles. All other frames, however, will host a $B$ field, another detectable state of affairs. Furthermore, there are no nearby allowed worlds in which our inertial laboratories are shielded from the influence of this constant cosmological electromagnetic field and thus oblivious to whether it has a $B$ component: for our additional laws effectively outlaw shielding.

But, despite these factual distinctions between inertial frames in any world of our electrodynamical theory, there are none at the level of laws. If we are willing to grant the physical equivalence of inertial frames in our strengthened electrodynamics, it seems reasonable to grant the physical equivalence of $K$ and $K'$ in gravito-particle mechanics as Einstein argues and thus grant his extension of the relativity principle in this context; just because the distinction between $K$ and $K'$ rests on the merely factlike, and not lawlike, presence of an inertial-gravitational field in $K'$.

One reaches the same conclusion by considering another analogous theory—a *weakened* version of Lorentzian electrodynamics. Recall that Lorentz postulated an ether as the seat of electromagnetic fields. Lorentz's ether was immutable or frozen, i.e. non-dynamic and partook of no internal motions, and thus singled out as special its rest frame. Imagine for vividness that this rest frame is experimentally detectable, i.e. that we are dealing with a Lorentz theory in which an ether wind is detectable, e.g. one in which the time dilation factor differed from that proposed by Larmor. Imagine also, again for philosophical reasons only, that we thaw the ether and allow it to flow just like a common-or-garden fluid. The frozen case is not ruled out by our theory: in some worlds—the boundary cases—the ether will behave inertially and hence pick out its rest frame. But this now has a different significance: the picking out of any frame, e.g. the rest frame of the Earth on 12 February 1997, would be merely accidental. In other possible worlds an immutable ether picks out other frames. This weakened theory, then, satisfies the relativity principle, even though, in special boundary cases, certain frames are singled out.

An interesting question is whether there is a point of view from which the Lorentzian theory *with* the frozen ether assumption (and still with a time dilation factor different from Larmor's) also satisfies the relativity principle. The fact that the ether is frozen does not prevent it from moving inertially! And the possibility of inertial motion of the ether through the frame it happens to be at rest in in the actual world, is all one apparently needs for the actual rest frame to have the same lawlike status as any other frame: possible observed differences would be contingent upon the state of inertial motion of the ether in the various frames and would not be grounded on a non-invariance of laws. Whatever one thinks about the existence of such a point of view from which relativity is satisfied in this case, it seems to us that the case for the existence of the point of view advocated by Einstein is stronger, since he explicitly considered a boundary case of a theory in which the gravitational field was independently known to have a very rich dynamics.

It is worth noticing at this point that although the equivalence of $K$ and $K'$ Einstein insightfully perceived at the level of particle mechanics might well *suggest* that $K$ and $K'$ enjoy full blown equivalence, i.e. also at the level of electrodynamics, quantum mechanics, etc., it certainly *does not imply* this. In other words, the satisfaction of the *weak* version of the equivalence principle[3] does not imply the satisfaction of the *strong* version. Although in some theory $T$ a uniform gravitational field might exert its influence on classical chargeless particles in a manner which can be mimicked simply by a change of coordinate perspective, this might not also be the case in $T$ for the influence of the gravitational field on the electromagnetic field or wave or quantum phenomena. Einstein seems erroneously to suggest otherwise in the quotation since he refers to 'the laws of nature' without restriction.

Notice also that the equivalence Einstein perceives between $K$ and $K'$ can only be the case if there is a gravitational field in $K'$ whilst there is none in $K$. In other words, the satisfaction of the equivalence principle in this case depends on the validity of a *frame-dependent* concept of a homogeneous gravitational field. There are similarities and dissimilarities to the electrodynamical case here. The presence of the $B$ field was frame-dependent in the electrodynamical worlds we discussed above. But the presence of $F$ was not. In fact $B$ is a component of the Faraday tensor. The gravitational field strength of a uniform gravitational field is frame-dependent and is not the component of a tensor. Einstein's equivalence principle requires that in some frame, all evidence of a uniform gravitational field vanishes. This means that a uniform gravitational field is not represented by a tensor, i.e. it does not add an extra field to the boundary case of special relativity.

### *2.2. The impact of Einstein's equivalence principle*

Even today Einstein's way of reasoning requires some mind-twisting: no doubt that the equivalence principle provides another striking illustration of

[3] The weak equivalence principle is not, as is often falsely claimed, a mere restatement of, but is stronger than, the observed equality of inertial and gravitational mass. Consider a theory of gravitation in which the latter held but in which the gravitational field is represented by an antisymmetric tensor just like the electromagnetic field, and particles feel its influence through a Lorentz force equation. The weak equivalence principle fails in such a theory. In general there exists no connection whose affine geodesics coincide with the paths of such particles; see Ehlers *et al.* (1972, p.76).

Einstein's stunning innovative abilities. EEP allows Einstein to see spacetime, or more accurately, the spacetime structure, in a new light. In an accelerated frame, the trajectory of a free particle is completely determined by the Christoffel symbols $\Gamma^{\sigma}_{\mu\nu}$ expressed in terms of the derivatives of the metric. The metric defines the spacetime structure: the motion of the free particle is then characterised in a purely geometrical manner. The motion of a particle in an accelerated frame unequivocally depends on the metric, in the same way that in a non-accelerated frame a free particle follows a straight line. If we can treat an accelerated frame as equivalent to a rest frame in which a constant gravitational field acts, it implies that, in the homogeneous case, the gravitational field is geometrised and is represented by the Christoffel symbols (or, in a coordinate-free language, by the affine connection).[4] As Norton (1985, p. 245) remarks, Einstein's EP states the complete physical indiscernibility, in a *particular* case, of an inertial and a gravitational field, that is the equivalence of physical structures. A constant gravitational field is exactly the inertial field which arises in a uniformly accelerating frame in Minkowski spacetime. As a consequence, the mathematical objects which represent the inertial field, i.e. the Christoffel symbols expressed in terms of derivatives of the metric, equally represent the gravitational field, and *vice versa*. Einstein's EP provides a clue, a heuristic device as was often pointed out, towards a geometrical representation of arbitrary gravitational fields by means of the affine connection of an appropriate spacetime structure. The spacetime structure, expressed by the metric, depends on the matter-energy distribution, becomes dynamical and uniquely determines, by the geodesic law, the paths of freely-falling bodies: this aim was attained in Einstein's general theory of relativity (see Stachel (1986) and Ghins (1990)).

The physical content of the equivalence principle crucially depends on what is meant by 'gravitational field'. Given the somewhat imprecise usage of the term 'gravitational field', this issue is not as trivial as it may seem. In fact, there are three candidates for the mathematical representation of the field: the metric tensor, the components of the connection (the Christoffel symbols) and the Riemann curvature tensor. If the metric, also sometimes called the 'gravitational potential', is constant in a frame, then no gravitational forces are present (since forces depend on the first-order derivatives of the metric). A constant gravitational field cannot be represented by a constant metric because this would simply imply that there are no forces present. However, if the Christoffel symbols are constant in a frame, we do have a constant gravitational force: this situation corresponds to the inertial-gravitational field generated by an acceleration in Minkowski spacetime and referred to in Einstein's principle. In this case the curvature is constant and nil. A constant connection can then be said to represent a constant gravitational field. The third possibility for representing the gravitational field, the curvature tensor, seems inappropriate since a vanishing curvature can correspond to either the absence of gravitational forces or the

[4] We will assume throughout this paper that affine and metrical structures are compatible in the sense that affine geodesics coincide with metrical geodesics.

presence of constant forces. Hereafter we will always assume that gravitational fields are represented by the affine connection (in a coordinate-free language) or the Christoffel symbols (in a component language). Einstein himself stressed that the Christoffel symbols were the correct mathematical representation of the gravitational fields.[5] Note that despite their similarities, the relativity principle and equivalence principle differ in the impact they have had on physics. Whereas the relativity principle urged us to delete absolute space from our physical spacetime ontology, the equivalence principle urged us to identify part of spacetime structure (the affine connection) with a dynamical field. There is no deletion of spacetime structure, merely an identification of it with the gravitational field. This leads Friedman (1983, pp. 191–193) to criticise the interpretation of the equivalence principle as a relativity principle, since he argues that it does not lead to a relativisation of acceleration. But in a sense it does: all acceleration is relative to the background of a *contingent* affine connection.

## 3. The Infinitesimal Principle of Equivalence

Although Einstein's equivalence principle gave him the clue how to develop his theory of gravitation—virtue enough for any principle!—many physicists have tried to liberate it from the restriction to homogeneous gravitational fields. After all, actual gravitational fields in nature are not constant and so Einstein's principle of equivalence does not apply to them. The idea behind *infinitesimal* principles of equivalence is that if we confine attention to infinitesimal regions in our theory of gravitation, then any gravitational inhomogeneity should become negligible and hence we can deploy Einstein's point of view in the infinitesimally small. The promise of retaining Einstein's insights in the small whatever the gravitational field is tantalising. This might provide us with a bridge from special relativity to general relativity, allowing us to read off the general relativistic laws for the behaviour of matter in the presence of a gravitational field (but not the behaviour of the gravitational field itself) from special relativity.

But there have been criticisms of the coherency of the infinitesimal equivalence principle and its ability to act as a bridge. Einstein himself was sceptical[6] and did not give a place to the infinitesimal principle of equivalence in his 1916 paper (Einstein, 1916a). Even so, we believe that there is a valid and powerful infinitesimal equivalence principle—which we shall call the punctual

[5] See Einstein's letter to Max von Laue (September 1950; translated in Stachel, 1986, p. 1858): 'what characterizes the existence of a gravitational field from the empirical standpoint is the non-vanishing of the $\Gamma^i_{jk}$, not the non-vanishing of the $R_{iklm}$. If one does not think in such intuitive ways, one cannot grasp why something like curvature should have anything to do with gravitation at all. In any case, no reasonable man would have hit upon anything in that way. The key to the understanding of the equality of inertial and gravitational mass would have been lacking'.

[6] See Norton (1985), though he did extend results he derived for the homogeneous gravitational field, e.g. the redshift formula, to inhomogeneous fields; see Einstein (1911, Section 3, p. 104f).

equivalence principle—and that a coordinate dependent approach helps to see this. Before we present our version we consider the difficulties with others.

### 3.1. The local equivalence principle

Consider the following as a first (perhaps obviously unsatisfactory) stab at formulating an infinitesimal equivalence principle:

> **Strict Local Equivalence Principle (SLEP)**: around any point $p$ in any model of a theory of physics which satisfactorily includes gravitation, there exists a neighbourhood $\mathcal{U}_p$ and a chart $x^\mu$ thereon such that special relativity holds in its standard vectorial form in $x^\mu$ at all points $q \in \mathcal{U}_p$.

Such a formulation might be suggested by a casual reading of some statements made by some physicists. Of course, we recommend neither such casual reading of the writings of physicists nor the formulation just given: for the latter is violated in general relativity. If special relativity holds on a region, no matter how small, Minkowski geometry holds in that region, as experiments with rods and clocks will reveal. Hence spacetime is flat and the curvature tensor vanishes; but this is commonly false in general relativity. This is clearly a serious defect in a formulation of a principle which is supposed to capture an essential aspect of general relativity!

A much better attempt, and one we think that many physicists would agree with, proceeds along the following lines:[7]

> $\forall\delta\exists\varepsilon$ **Equivalence Principle (DEEP)**: in any theory of physics which satisfactorily includes gravitation, for any order of approximation $\delta > 0$, we can always find a small neighbourhood of size $\varepsilon > 0$ around any point and a chart thereon in which the laws of motion of matter (but not gravitation) approximate their standard special relativistic form to order $\delta$.

For example, DEEP requires that if special relativity says a field $\phi$ obeys an equation of the form $f(\partial_\mu)\phi = 0$, then in a satisfactory theory of gravitation $\forall\delta$ there is a neighbourhood such that the field satisfies $f(\partial_\mu)\phi \leqslant \delta$ throughout. Weinberg (1972, p. 68) appears to have such a formulation in mind when he writes:

> Although inertial forces do not exactly cancel gravitational forces for freely falling systems in an inhomogeneous or time-dependent gravitational field, we can still expect an approximate cancellation if we restrict our attention to such a small region of space and time that the field changes very little over the region. Therefore we formulate the equivalence principle as the statement that *at every space-time*

[7] The degree of approximation $\delta$ is contextually defined by taking into account the nature of the phenomena under scrutiny and the degree of precision sought (see Brown, 1997, p. 74).

> *point in an arbitrary gravitational field it is possible to choose a 'locally inertial coordinate system' such that, within a sufficiently small region of the point in question, the laws of nature take the same form as in unaccelerated Cartesian coordinate systems in the absence of gravitation* (Weinberg's emphasis).

At first sight this stands a better chance of being satisfied in general relativity than SLEP. Let us assume that we may think of the gravitational field as represented in some chart by the Christoffel symbols associated with that chart. Now in a sufficiently small region around any point, there always is a system of coordinates in which all the Christoffel symbols are smaller than an arbitrarily small value, i.e. in which the gravitational field, in the sense under investigation, is arbitrarily close to being zero.

But it is of crucial importance to realise at this stage that the fact that the Christoffel symbols are approximately equal to zero in a small region does *not* imply that the curvature—the Riemann tensor—vanishes in that region: the nearly vanishing of the first-order derivatives of the metric does not entail that its second-order derivatives nearly vanish. Indeed arbitrary gravitational fields are represented in general relativity by curved metrics and this curvature shows up at *points* in all coordinate systems. Furthermore curvature has non-special relativistic effects on the behaviour of extended bodies; see Hawking and Ellis (1973, Sections 4.1–4.2, and especially the Landau–Raychaudhuri equation (eqn. 4.26, p. 84)).[8]

There are of course cases in which the curvature can be neglected in a finite region of spacetime. In such instances, the field can be practically treated as constant and Einstein's principle of equivalence can be applied. Readers interested in how the assumption of equivalence leads to interesting predictions in quantum mechanics should consult the paper by Greenberger and Overhauser (1979).[9]

Although it is perhaps natural to suppose that the effects of curvature on extended bodies—tidal effects—vanish as the size of the body shrinks, this is false, as has been shown by Ohanian (1977) (see also Ohanian and Ruffini (1994)). Consider a large fluid planet orbiting in close proximity to a massive star. The planet experiences tidal forces and hence takes up an ellipsoidal rather than a spherical form. As Ohanian shows, if we neglect surface tension effects, the ellipticity of the planet is *independent* of its radius and hence remains constant and finite regardless of how small we take its radius to be, so long as it has *some* radius. This is the case in general relativity, where the ellipticity of the planet is a measure of the curvature at its centre, and Newtonian gravitation alike and the ellipticity is, in principle at least, measurable.

[8] The curvature may also show itself in the behaviour of particles with spin (Ohanian, 1977, p. 908), but we shall leave aside the issue of spinning particles.

[9] It is our opinion that the argument in this paper is a great improvement on that of Greenberger (1983, Section VIII), which seems to be invalid.

Now it may be considered as a law of special relativity that

> **Special Relativistic Sphere Law**: all extended (and non-rotating) fluid bodies which are isolated and chargeless are spherical.

To the extent that DEEP, in the formulation given above, concerns all laws (except gravitation laws) and that the Special Relativistic Sphere Law is violated in GTR even for an arbitrarily small neighbourhood (leaving quantum limitations aside), DEEP is not satisfied by GTR. A way to get over this problem is to restrict its validity to fundamental laws. The Special Relativistic Sphere Law is not a fundamental law and can be derived from more basic laws only by means of additional assumptions on extended fluid bodies. Thus reformulated, DEEP requires that curvature does not occur in the basic laws of nature, in such a way that the coupling of spinless particles with the gravitational field is 'minimal' as Ohanian (1977, p. 905) puts it.

Granted, the notion of a 'fundamental law of nature' is problematic and requires a philosophical elucidation which we will not attempt here. At some stage, however, such an elucidation will have to rely on our intuitions: there is a clear intuitive sense in which the Special Relativistic Sphere Law is not fundamental but derivative.

It can be objected that there are other special relativistic laws which, unlike the Special Relativistic Sphere Law, may be considered fundamental and fail to be true, even at a point, if the underlying spacetime is curved. Think for example of the equation describing the temporal kinematics of a continuous matter distribution. If spacetime is curved, its temporal kinematics will be governed by the equation of geodesic deviation in which curvature explicitly appears. Or take the Landau–Raychaudhuyri equation mentioned above. These equations are expressed in terms of quantities defined at a point and seem to be *bona fide* fundamental laws of nature.

Yet, at a point and in an appropriate chart, the dynamical *field equations of matter* are curvature-free. Thus, there is a central part of special relativity which, being independent of curvature, is inherited in general relativity and which is expressed by Maxwell's laws and Lorentz's force law (see below). These certainly are fundamental laws and they govern the behaviour of pointlike classical particles in electromagnetic fields. Other special relativistic laws cease to be valid in curved spacetimes, even at a point, and are not retained. Nature tells us that the validity of the equivalence principle must be thus restricted. Nevertheless this restricted EP has significant physical content: Maxwell's laws could fail to be valid at a point (and in fact they do fail to account for quantum phenomena). Our position here parallels Eddington's views:

> It is likely that some phenomena will be determined by comparatively simple equations in which the components of curvature of the world do not appear [...]. It is to these that the Principle of Equivalence applies. But there are more complex phenomena governed by equations in which the curvatures of the world are

> involved [...]. For these the Principle of Equivalence breaks down. The Principle of Equivalence thus asserts that some of the chief differential equations of physics are the same for a curved region of the world as for an osculating flat region. There can be no infallible rule for generalising experimental laws; but the Principle of Equivalence offers a suggestion for trial, which may be expected to succeed sometimes and fail sometimes (Eddington, 1923, pp. 40–41).

Hereafter, the fundamental laws of special relativity will be taken in a restricted sense and will only include the simplest laws in which the curvature does not explicitly occur. Thus, the laws in question express relations between quantities defined at a point and which are curvature-independent. As a matter of fact, these equations are just the dynamical laws which govern the coupling of matter with the classical fields (except gravitation), that is the electromagnetic fields. This eliminates the geodesic deviation equation and the Landau–Raychaudhuri equation which involve kinematical quantities only, like the displacement vector field, and do not contain matter-field dynamical variables (gravitation aside). We now proceed to give a precise formulation of the punctual equivalence principle.

### *3.2. A punctual equivalence principle*

First we propose the following definition of what it is for special relativity to hold at a point $p$ in the restricted sense; see also Friedman (1983, pp. 199ff):

> **Special relativity holds at $p$:** there exists a local chart $x^\mu$ of a neighbourhood of $p$ such that the fundamental dynamical and curvature-free special relativistic laws hold in their standard vectorial form in $x^\mu$ at $p$.

Note that this requires only that the standard equations hold at the *single* point $p$ in $x^\mu$; they need not hold in that chart at any other point.

What we mean by 'special relativity holds in its standard vectorial form' is that the standard list of classical dynamical variables obey the standard list of the fundamental dynamical equations, e.g.:

**Maxwell's equations:**

$$\vec{\nabla}\cdot\vec{E} = 4\pi\sigma \tag{1}$$

$$\vec{\nabla}\times\vec{B} - \frac{\partial\vec{E}}{\partial t} = 4\pi\vec{j} \tag{2}$$

$$\vec{\nabla}\cdot\vec{B} = 0 \tag{3}$$

$$\vec{\nabla}\times\vec{E} + \frac{\partial\vec{B}}{\partial t} = \vec{0} \tag{4}$$

**Relativistic law of particle motion:**

$$\frac{d\vec{p}}{dt} = q(\vec{E} + \vec{v} \times \vec{B}). \tag{5}$$

There are some difficulties involved in defining special relativity with reference to these particular equations, since they do not account, for example, for the cohesion of solid bodies and thus render special relativity hopelessly empirically inadequate. Our aim in taking this list-based approach is to rule out the possibility that dynamical variables other than the electromagnetic field, etc., like spacetime curvature, be tolerated by the equivalence principle to play a significant rôle in some fundamental laws of physics at a point. In other words, there exist fundamental special relativistic laws—the chief dynamical laws—which hold at a point (in some charts), even in curved spacetimes, since they are independent of curvature.

This restricted way of understanding special relativity to hold at a point permits us to remove a serious obstacle to infinitesimal equivalence principles, namely that, as Norton (1985, p. 239) puts it: 'Special relativity requires the vanishing of the Riemann–Christoffel curvature tensor'. We agree that the holding of *all* special-relativistic laws at a point implies that the curvature vanishes at that point. Nevertheless some basic special relativistic laws, the dynamical curvature-free laws, do not carry that consequence, as we saw. While we also agree that the holding of special relativity in a *neighbourhood* is tantamount to the existence of a *curvature-free* Lorentzian spacetime metric on that neighbourhood[10] and, hence, incompatible with general relativity, we claim that good sense can be made of (restricted) special relativity holding *at a point*, independently of it holding on some neighbourhood containing that point, as we have done above, and that it just does not follow from special relativity holding in this way at a point that curvature vanishes there: special relativity holds perfectly well at a point in our restricted sense in general relativity. In the way we understand it, special relativity's holding at each point tells us nothing about curvature whatsoever.

This line of thought leads to the following definition of our *punctual equivalence principle*:

**Punctual Equivalence Principle (PEP):** for all $p \in M$ special relativity holds at $p$, in the restricted sense given above.

Perhaps this version is what physicists have intended all along. If so we will be happy to have made this clear. General relativity *does* satisfy our punctual version of the infinitesimal equivalence principle as special relativity holds at all points $p$ in general relativity. To see this, we must realise that the fact that special

[10] It does not follow from the latter that spacetime is globally Minkowskian: it could, for example, have a cylindrical global topology.

relativity holds at p in the restricted sense given above does not imply that the curvature tensor vanishes at p.

### *3.3. The physical content of PEP*

The punctual equivalence principle is rather strong and provides us with lots of physical information.

#### *3.3.1. The nature of gravitation*

The most profound insight is that gravitation is quite different from other physical effects: it is just not mediated by a force in the traditional sense, i.e. something which causes acceleration. That is, gravitational effects are not captured by a geometric equation of the form

$$\text{acceleration} = \text{Force} \div \text{mass}, \tag{6}$$

i.e. there is no tensorial gravitational force, since in the absence of non-gravitational forces, for any point $p$ we can always construct charts—the Lorentz charts—such that the acceleration of a particle under the influence of gravitation vanishes at $p$. True: in some other charts the acceleration will not vanish at $p$; but the crucial fact is its vanishing in the Lorentz charts, for this means that there is no non-zero *tensor* associated with it nor, thereby, with a gravitational force, quite unlike the case for the other interactions. In this sense, gravity is not a force. This non-force status of gravity follows from the fact that at any point $p$ in a region under the sway of an arbitrary gravitational field special relativity holds at $p$ as asserted by the punctual equivalence principle; is before clearly not a consequence of special relativity alone.

Einstein rejected some *prima facie* natural contenders for gravitational force expressions (4-vector and 6-vector theories) very early on in his investigation of gravitation, as they violated the principle that particles should accelerate vertically independently of their horizontal velocity. He seems to have rejected them by the time he wrote his 1907 paper (Einstein, 1907); see Norton (1992) for some wonderful exposition. PEP rules out all such expressions in one stroke, for it rules out the participation in Newton's second law of all forces above and beyond the special relativistic ones (i.e. gravitation does not count as a special relativistic force).

Einstein's general relativity offers an alternative approach to gravitation which avoids tensorial gravitational forces. Coordinate-independent gravitational effects are due to non-Minkowskian spacetime structure, more specifically, curvature. There are many ways a spacetime can differ from Minkowski spacetime. PEP constrains how we modify Minkowski structure in search for a theory of gravitation. For example, there are spacetimes with non-Riemannian, *Finslerian* metrics, unlike the Riemannian Minkowski metric. A Finslerian metric is one whose interval is not quadratic in the coordinate differentials, e.g.

$ds^2 = \sqrt{dx^4 + dy^4 + dz^4 - dt^4}$; see Synge (1956, p. 19).[11] And PEP forbids physical theories based on such spacetimes. Finsler geometry always introduces some anisotropy at each point of spacetime, anisotropy which is not present in special relativity, and so not allowed in theories of gravitation satisfying PEP, which must inherit special relativity at each point.

Thirdly it rules out many theories based on curved Riemannian spacetimes, namely those in which curvature enters the equations of motion for the matter fields. One such disallowed theory is the well-known conformally invariant scalar field theory whose field equation is:[12]

$$(\Box - \frac{1}{6}R)\phi = 0. \tag{7}$$

*3.3.2. Free fall motion*

Consider a theory of gravitation $T$ that satisfies the punctual equivalence principle and whose material dynamics unfolds against the backdrop of a Lorentzian spacetime $\langle M, \nabla, g \rangle$, where the connection $\nabla$ and metric $g$ are compatible, as in general relativity. We would like to know what the equivalence principle implies about the relationship between the material phenomena and the spacetime structure. In particular: does it imply that free-falling particles in the theory move along geodesics of the spacetime? Surprisingly, the answer to this question is 'yes'.[13]

Of course, in order to get off the ground with this project we have to assume *something* about how the metrical structure relates to the matter. (We want to avoid worlds, such as those envisioned by Poincaré, in which there is a scalar field which universally affects metrical devices and which would prevent them from correctly detecting the metric field. In other words, we want to only consider worlds in which the standard metrical devices are fair indicators of the metrical structure, and thus curvature.) The following seems reasonable, and is motivated by what is the case in general relativity.

> **Meshing Assumption:** for all points $p$ and local charts $x^\mu$ containing $p$, if the standard list of special relativistic fundamental equations of motion hold in their standard form in $x^\mu$ at $p$, then $g_{\mu\nu} = \text{diag}(-1,1,1,1)$ at $p$ and $\nabla_\mu = \partial_\mu$ at $p$, i.e. the chart should have zero Christoffel symbols at $p$.[14]

[11] See Bogoslovsky (1977) and Budden (1997a,b) for discussions of Finslerian generalisations of special relativity.

[12] See Wald (1984, p. 447).

[13] As is well known, the geodesic law of free fall can also be derived from Einstein's field equations, and more precisely from the vanishing of the covariant derivative of the stress-energy tensor. This is a remarkable property of general relativity which is not enjoyed by other field theories. See for example Stephani (1996, pp. 91–93).

[14] Together with metric compatibility, the meshing assumption implies that the first derivatives of the metric components also vanish at $p$. It is also to be noted that the vectorial, coordinate-independent, Maxwell equations hold not only in Lorentz charts but in all charts, e.g. polar coordinates, in which the flat geometry is preserved.

Given the meshing assumption, the geodesic law of free fall follows trivially from the punctual equivalence principle.

Since $T$ satisfies the punctual equivalence principle, special relativistic laws must hold at each point in $M$ (in the sense defined above). In other words, around each point $p$ in $M$ there exists a local chart $x^\mu$ such that the fundamental special relativistic laws hold in $x^\mu$ at $p$. Since the classical Galileo law is one such law (it follows from the Lorentz-force law when no electromagnetic field acts), we have

$$\frac{d^2x^i}{dt^2} = 0 \quad \text{at } p, \tag{8}$$

where $t = x^0$ and $i \in \{1, 2, 3\}$, for all worldlines of free-falling particles passing through $p$. This obviously implies

$$\frac{d^2x^\mu}{dt^2} = 0 \quad \text{at } p. \tag{9}$$

Given the meshing assumption ($\Gamma^\mu_{\nu\sigma} = 0$ at $p$ in $x^\mu$), this implies that the free-falling particles are non-accelerating at $p$ with respect to $\nabla$ (and that $t$ is an affine parameter at $p$). Since this is true for all $p$, it follows that all free-falling worldlines have zero acceleration along their entire lengths, i.e. that they are affine geodesics.

Our derivation of the geodesic law might seem *prima facie* objectionable, as Einstein himself rejected an attempt at a similar derivation by Moritz Schlick; see Norton (1985, pp. 235ff). In a letter to Schlick Einstein explained his objection as follows:

> The derivation of the law of motion of a point mass [...] proceeds from the motion of a point being a straight line, when considered in the local coordinate system. But from this nothing can be derived. In general, the local coordinate system has a meaning only in the infinitely small and in the infinitely small every continuous line is a straight line (Einstein to Schlick, 6 February 1917, EA 21 612; quoted by Norton (1985, p. 237)).

Our derivation certainly proceeds in just that way, but Einstein's objection that 'in the infinitely small every continuous line is a straight line' seems plainly false, at least in our context of full-blown Riemannian geometry. At every point on a continuous line in Riemannian geometry there is defined the acceleration vector, and straight lines have zero acceleration at every point whereas non-straight ones do not.

Norton (1985, Section 10) attempts a sophisticated, but, we believe, flawed, defense of Einstein's view. He is, of course, well aware that full-blown Riemannian geometry allows us to define an acceleration vector at each point on a continuous worldline, but he believes that the physicist who wishes to invoke an infinitesimal equivalence principle must work with less. As we mentioned above in Section 3.2, Norton thinks, rightly in our opinion, that the holding of special relativity always implies that curvature vanishes and, hence, that anyone wishing to assert that special relativity holds in general relativity must qualify this with something like 'ignoring third-order quantities', of which curvature is one, in Norton's classification scheme. Now, if we must ignore third-order quantities then we cannot draw all the distinctions we can draw if we allow ourselves the full resources of Riemannian geometry. Nonetheless, second-order assertions seem quite sufficient for the purposes of deriving the geodesic equation, as assertions of the form 'acceleration vanishes at point *p*' are amongst them. Any infinitesimal equivalence principle that asserts that special relativity holds everywhere up to second-order quantities, implies that freely-falling particles do not accelerate anywhere, i.e. that their worldlines should be geodesics. Norton apparently fails to see this and is seemingly misled by the fact that the vanishing of acceleration at some point on a particle's worldline does not imply its vanishing at distinct points. Infinitesimal equivalence principles like PEP do restrict what they say to assertions like $\forall p \in M$, $Sp$, where $M$ is the spacetime manifold and $S$ is some differential geometric property pertaining to single points, but they do not restrict themselves to assertions like $Sp$ where $p$ is a particular individual point!

## 4. The Principles Compared

We saw that Einstein's equivalence principle (EEP) is an extension of the principle of relativity to uniformly accelerated motion. It states that a spacetime with a homogeneous gravitational field is identical to a spacetime endowed with the inertial structure which appears in a frame uniformly accelerating in Minkowski spacetime. This allowed Einstein to identify a particular (homogeneous) case of gravitational field with an inertial field. As has been beautifully—and convincingly—shown by John Norton (1985), this insight led him to make the bold hypothesis that the spacetime metric is dynamical and contingent upon the matter-energy distribution and also that gravitational forces in general could be derived from a variable Riemannian metric and are thus geometrised. The next, and arduous, step was to formulate the mathematical connection between any matter-energy distribution and the metric tensor (the gravitational potentials), i.e. Einstein's field equations.

We do not intend to rediscuss here the well-documented heuristic rôle which the EP played in Einstein's hands.[15] Einstein's principle clearly had a different purpose than the infinitesimal principles. But, independently of the motivations behind the introduction of EEP and PEP, their physical contents (as well as the place of Einstein's principle in current gravitation theory) deserve a fresh look.

Einstein's principle of equivalence states that flat Minkowski spacetime is already a particular case of a spacetime with a gravitational field (when seen from a uniformly accelerating frame). This constant gravitational field, which is not produced by a source, must be allowed by the field equations. Thus, the

[15] Einstein (1911, pp. 100–101) himself insisted on this heuristic rôle.

general field equations must admit Minkowski spacetime among the matter-free solutions, i.e. when the stress-energy tensor is equal to zero. In this sense, Einstein's principle is satisfied in general relativity even though general relativistic physical spacetimes are not flat. This constraint also follows from the infinitesimal principles.

But do the infinitesimal principles physically impose more constraints on the field equations and the structure of the metric than EEP? It is commonly—and rightly—believed so, because the infinitesimal principles are claimed to restrict the admissible spacetime manifolds to (pseudo-)Riemannian manifolds endowed with a quadratic metric of negative signature which admit a tangent space at each point, whereas EEP does *not* impose that the spacetimes of the general theory of relativity are Lorentzian: EEP only requires that a source-free spacetime is a special case of Lorentzian spacetime, i.e. Minkowski spacetime. But field equations (different from Einstein's) which admit as non-vacuum solutions spacetime manifolds endowed with a Finsler metric are not ruled out by EEP, provided Minkowski spacetime is among the matter-free solutions. The infinitesimal principles of equivalence however require a Lorentzian metric for all solutions of the field equations.

Moreover, EEP by itself only strongly suggests—but does not require—that gravitation depends solely on derivatives of the metric, and is thus geometrised, also in the case of non-constant arbitrary fields. Such a geometrisation is a consequence of the infinitesimal principles.

Finally, Einstein knew that his equivalence principle did not permit the derivation of the geodesic law:

> [...] in principle, there can exist finite (matter-free) parts of the world for which
>
> $$ds^2 = dx_1^2 + \dots - dx_4^2 \tag{10}$$
>
> with an appropriate choice of the reference system. (If this were not the case, then the Galilean law of inertia and the special theory of relativity could not have held good.) In such a part of the world, the Galilean law of inertia holds with this choice of reference system; and the world line is a straight line and therefore a geodetic line, with an arbitrary choice of coordinates.
>
> That the world line of a point is a geodetic line in other cases too (if none other than gravitational forces act) is an hypothesis, even if a very obvious one (Einstein in a letter to Moritz Schlick; translated in Norton, 1985, pp. 237f).

We saw that the geodesic law follows from the punctual equivalence principle (PEP). Since PEP can be seen as the limiting, and thus a particular, case of DEEP when $\delta$ tends to zero, DEEP permits the derivation of the geodesic law as well. Both infinitesimal principles have the same consequences for the structure of the gravitational fields and its coupling to matter. The preference between the two principles, then, seems to be merely a matter of taste. Physicists may prefer DEEP since rather large regions of spacetime may in practice be considered physically equivalent to infinitesimal regions when the relevant approximations hold. Mathematicians (and perhaps philosophers too...) may feel more inclination towards the punctual principle since it does not involve the sometimes confusing notion of approximation. In any event, we can safely conclude that the infinitesimal principles, be it DEEP or PEP, have more physical content than EEP.

*Acknowledgements*—We are grateful to Jeeva Anandan, Erik Curiel, Jean-Marc Gérard, Jan Govaerts, René Lavendhomme and especially Harvey Brown and John Norton for their very helpful remarks and suggestions. One of us, Tim Budden, would like to thank the FSR (Fonds Spéciaux de Recherche) of the Université Catholique de Louvain for supporting this research.

## References

Bogoslovsky, G. Y. (1977) 'A Special-Relativistic Theory of the Locally Anisotropic Space-Time', *Il Nuovo Cimento* **40B**(1), 99–115, 116–133.

Brown, H. R. (1997) 'On the Role of Special Relativity in General Relativity', *International Studies in the Philosophy of Science* **11**, pp. 67–81.

Budden, T. (1997a) 'A Star in the Minkowskian Sky: Anisotropic Special Relativity', *Studies in History and Philosophy of Modern Physics* **28**(3), 325–361.

Budden, T. (1997b) 'Galileo's Ship and Spacetime Symmetry', *British Journal for the Philosophy of Science* **48**, 483–516.

Eddington, A. (1923) *The Mathematical Theory of Relativity* (Cambridge: Cambridge University Press).

Ehlers, J., Pirani, F. and Schild, A. S. (1972) 'The Geometry of Free Fall and Light Propagation', in L. O'Raifeartaigh (ed.), *General Relativity: Papers in Honour of J. L. Synge* (Oxford: Clarendon), pp. 63–84.

Einstein, A. (1907) 'Über das Relativitätsprinzip und die aus demselben gezogenen Folgerungen', *Jahrbuch der Radioaktivität und Elektronik* **4**, 411–462.

Einstein, A. (1911) 'Über den Einfluss der Schwerkraft auf die Ausbreitung des Lichtes', *Annalen der Physik* **35**, pp. 898–908. Translated in Perrett and Jeffery (1952), pp. 97–164.

Einstein, A. (1916a) 'Die Grundlagen der allgemeinen Relativitätstheorie', *Annalen der Physik* **49**, pp. 769–822. Translated in Perrett and Jeffery (1952), pp. 109–164.

Einstein, A. (1916b) 'Über Friedrich Kottlers Abhandlung "Über Einsteins Aequivalenzhypothese und die Gravitation" ', *Annalen der Physik* **51**, 639–642.

Friedman, M. (1983) *Foundations of Space-Time Theories* (Princeton: Princeton University Press).

Galileo, G. (1632) *Dialogue Concerning the Two Chief World Systems*, translated by S. Drake (Berkeley: University of California Press, 1953; 2nd edn 1967).

Ghins, M. (1990) *L'inertie et l'espace-temps absolu de Newton à Einstein. Une analyse philosophique* (Brussels: Académie Royale de Belgique).

Greenberger, D. M. (1983) 'The Neutron Interferometer as a Device for Illustrating the Strange Behavior of Quantum Systems', *Reviews of Modern Physics* **55**(4), 875–905.

Greenberger, D. M. and Overhauser, A.W. (1979) 'Coherence Effects in Neutron Diffraction and Gravity Experiments', *Reviews of Modern Physics* **51**(1), 43–78.

Hawking, S. W. and Ellis, G. F. R. (1973) *The Large Scale Structure of Space-Time* (Cambridge: Cambridge University Press).

Norton, J. (1985) 'What was Einstein's Principle of Equivalence?', *Studies in History and Philosophy of Science* **16**, 203–246.

Norton, J. (1992) 'Einstein, Nordström and the Early Demise of Scalar, Lorentz-Covariant Theories of Gravitation', *Archive for History of Exact Sciences* **45**, 17–94.

Ohanian, H. C. (1977) 'What is the Principle of Equivalence?', *American Journal of Physics* **45**(10), 903–909.

Ohanian, H. C. and Ruffini, R. (1994) *Gravitation and Spacetime*, 2nd edn (New York: Norton).

Perrett, W. and Jeffery, G. B. (1952) *The Principle of Relativity* (New York: Dover).

Stachel, J. (1979) 'The Genesis of General Relativity', in Nelkowski *et al.* (eds), *Einstein Symposion* (Berlin: Springer-Verlag), pp. 428–442.

Stachel, J. (1986) 'What a Physicist Can Learn from the Discovery of General Relativity', in R. Ruffini (ed.), *Proceedings of the Marcel Grossman Meeting on General Relativity* (Amsterdam: North-Holland), pp. 1857–1862.

Stephani, H. (1996) *General Relativity. An Introduction to the Theory of the Gravitational Field*, edited by J. Stewart; translated by M. Pollock and J. Stewart (Cambridge: Cambridge University Press).

Synge, J. (1956) *Relativity: The Special Theory* (Amsterdam: North-Holland).

Weinberg, S. (1972) *Gravitation and Cosmology* (New York: Wiley).

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