# What was Einstein's principle of Equivalence?*

John Norton

## 1. Introduction

In October and November 1907, just over two years after the completion of his special theory of relativity, Einstein made the breakthrough that set him on the path to the general theory of relativity. While preparing a review article on his new special theory of relativity, he became convinced that the key to the extension of the principle of relativity to accelerated motion lay in the remarkable and unexplained empirical coincidence of the equality of inertial and gravitational masses. To interpret and exploit this coincidence, he introduced a new and powerful physical principle, soon to be called the "principle of equivalence" upon which his search for a general theory of relativity would be based. Moreover, with the completion of the theory and throughout the remainder of his life, Einstein insisted on the fundamental importance of the principle to his general theory of relativity.

Einstein's insistence on this point has created a puzzle for philosophers and historians of science. It has been argued vigorously that the principle in its traditional formulation does not hold in thc general theory of relativity, Consider, for example, a traditional formulation such as Pauli's in his 1921 *Encyklopädie* article. For Pauli the principle asserts that one can always transform away an arbitrary gravitational field in an infinitely small region of space-time, by transforming to an appropriate coordinate system (Pauli 1921, p. 145).

In response, such eminent relativists as Synge (1960, p. ix), and even Eddington before him (1924, pp. 39–41), have objected that a coordinate transformation or change of state of motion of the observer can have no effect on the presence or absence of a gravitational field. The presence of a "true" gravitational field is determined by an invariant criterion, the curvature of the metric. The gravitation-free case of special relativity is just the case in which this curvature vanishes, whereas the true gravitational fields of general relativity are distinguished by the nonvanishing of this curvature.

This objection has immediate ramifications for the "Einstein elevator" thought experiment, which is commonly used in the formulation of the principle of equivalence. In this thought experiment, a small chambers such as an elevator, is accelerated in order to transforms away a gravitational field present within it or, depending on the version at hand. to produce a gravitational field in an initially gravitation-free chamber. Now in general relativity, nonvanishing metrical curvature is responsible for tidal gravitational forces. Their effects can be used by an observer within the chamber to decide whether the gravitational field present is a true gravitational field or is due to the acceleration of the chamber in gravitation-free space. Alternatively, they can be used to determine whether an apparently gravitation-free chamber is in free fall in a gravitational field or moving uniformly in gravitation-free space. It is significant that the effects of these tidal forces do not vanish as the box becomes arbitrarily small. For example, the tidal bulges arising in a freely falling liquid droplet do not vanish as the droplet in made arbitrarily small, ignoring such effects as surface tension (Ohanian 1977).

Of course it has proved possible to retain a principle of equivalence in general relativity. But to do this, the principle might be given quite new formulations, which seem to carry us far from Einstein's original intentions. For example, in its "weak" form the principle merely asserts the equality of inertial and gravitational mass.[1] Or in another form, it asserts that all phenomena distinguish a unique affine structure for space-time (Anderson 1967, pp. 334–338). Alternatively, we can retain a traditional formulation of the principle, such as Pauli's, by reading the restriction to infinitely small regions of space-time as denying access to certain quantities such as curvature, which are constructed from the higher derivatives of the metric tensor. But then the principle is reduced to a simple and, as far as questions of foundations are concerned, not especially interesting theorem in general relativity. Certainly Einstein could not represent such a result as a fundamental principle of his theory.

My purpose in this paper is to determine precisely what Einstein took his principle of equivalence to be, to show how it figured historically in his discovery of the general theory of relativity, and to show the sense in which he took it to be fundamental to that theory. In particular I will seek to demonstrate that Einstein's version of the principle and the way he sought to use it are essentially different from the many later versions and applications of the principle. As a result, we shall set that the objections rehearsed earlier from the later debate over the principle of equivalence are peripheral to the concerns of Einstein's version of the principle and that this version does find completely satisfactory and uncontroversial expression in the general theory of relativity.

In the following section, as a focus for the remainder of the paper, l will present one of the clearest and most cautious of Einstein's formulations of the principle or equivalence and in Section 3, l will develop sufficient formal apparatus to negotiate certain ambiguities in it. In particular. I will introduce the concept of a three-dimensional relative space of a frame of reference, which is essential to the understanding of Einstein's principle and much of his early work on his general theory of relativity.

In Sections 4 and 5, I will review the role the principle played in the 1907 to 1912 period of Einstein's search for his general theory of relativity. In Section 4, I will outline how the principle enabled Einstein to construct a novel relativistic theory of static gravitational fields and, in Section 5, I will outline the sense in which he believed the principle would enable an extension of the principle of relativity to accelerated motion.

In Sections 6, 7, and 8, I will examine the principle of equivalence within Einstein's general theory of relativity, whose basic formal structure was laid down by Einstein and Marcel Grossmann in 1912 and 1913 and which achieved its 'final form in November 1915. In Section 6, I will review aspects of Einstein's transition from a three- to a four-dimensional formalism, and, in Sections 7 and 8, l will review the status of the principle in the theory. In particular, we shall see its crucial heuristic role in the transition from the special to the general theory.

In Sections 9 and 10, I will relate Einstein's version of the principle and the results he drew from it to the "—infinitesimal" principle of equivalence, such as that formulated by Pauli, and which is now commonly but mistakenly regarded as Einstein's version of the principle. In particular, I will analyze in some detail a devastating objection Einstein had to this version of the principle. It follows from the objection that, insofar as it can be precisely formulated, the infinitesimal principle is trivial. In Section 1 1, I will review Einstein's attitude to Synge's now popular identification of "true" gravitational fields with metrical curvature. Finally, in Section 12, I will draw together the threads of my story and answer the question posed in the title of this paper.

## 2. Einstein's Formulation of the Principle of Equivalence

Einstein has given us many statements of the principle of equivalence in his treatments and discussions of the general theory of relativity. But none is clearer or more cautious than the formulation he gives in a 1916 reply to Kottler's claim that Einstein had given up the principle of equivalence in the general theory of relativity (Einstein 1916b). Einstein began by introducing the limiting case of special relativity in which he defined a "Galilean system". I quote this here for later reference:

> *1. The Limiting Case of the Special Theory of Relativity.* Let a finite space-time region be free from a gravitational field, i.e., it is possible to set up a reference system $K$ ("Galilean system"), relative to which the following holds for the region considered. Coordinates are measured directly in the well-known way with unit measuring rods, times with unit clocks, as is customarily assumed in the special theory of relativity. In relation to this system an isolated material point moves uniformly and in a straight line, as was assumed by Galileo.

He then proceeded to his statement of the principle:

> *2. Principle of Equivalence.* 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 "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 uniformly 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"

For Einstein, the basic assertion of the principle of equivalence is that "one may treat $K'$ as at rest.... " I will defer discussion of exactly what he intended with this assertion until Section 5. The assumption upon which this assertion is based—that acceleration can produce a gravitational field—is at present more commonly associated with the principle of equivalence. The way in which it is used, however, is distinct from its use in "traditional" formulations of the principle such as Pauli's. In the latter, by reversing Einstein's argument, one assumes that one can always transform away an arbitrary gravitational field in general relativity within an infinitesimal region of space-time, Einstein however considers only the homogeneous gravitational field produced by uniform, nonrotating acceleration in the Minkowski space-time of special relativity. In addition, there is clearly no restriction to infinitesimal regions.

These last features are typical characteristics of Einstein's preferred formulation of the principle and appear in many of the statements of the principle that Einstein gave throughout the half century of his working life. These include his first published formulation of the principle in 1907, some live years prior to the completion of the general theory of relativity (Einstein 1907. p. 454), his well-known 1911 communication on gravitation (Einstein 1911 pp. 898–899, and his 1916 review of the just-completed theory (Einstein 1916a, pp. 772–773).[2] The principle is defined in these terms in The Meaning of Relativity the work which came closest to his "textbook" on relativity (Einstein 1922, pp. 57–58). Finally, it appears again in this form in one of his last discussions of the question, the 1952 appendix to his popular book, Relativity (Einstein 1952, pp. 151–152).

Einstein's next step in his reply to Kottler was to insist pointedly that his principle did not allow one to transform away arbitrary gravitational fields. Rather it dealt only with those gravitational fields that could be transformed away and which we would now identify as associated with Minkowski space-time.

> *3. Gravitational Fields not only Kinematically Conditioned.* One can also invert the previous consideration. Let the system $K'$, formed with the gravitational field considered above, be the original one. Then one can introduce a new reference system $K$, accelerated with respect to $K'$, with respect to which (isolated) masses (in the region considered) move uniformly in a straight line. But one may not go on and say: if $K'$ is a reference system provided with an arbitrary gravitational field, then it is always possible to find a reference system $K$, in relation to which isolated bodies move uniformly in a straight line, i.e., in relation to which no gravitational field exists. The absurdity of such an assumption is quite obvious. If the gravitational field with respect to $K'$, for example, is that of a stationary mass point, then this field certainly cannot be transformed away for the entire neighborhood of the mass point, no matter how refined the transformation artifice. Therefore, one may in no way assert that gravitational fields should be explained so to speak purely kinematically; a "kinematic, not dynamic understanding of gravitation" is not possible. Merely by means of acceleration transformations from a Galilean system into another, we do not become acquainted with *arbitrary* gravitational fields, but those of a quite special kind. which, however, must still satisfy the same laws as all other gravitational fields. This is only again another formulation of the principle of equivalence (in particular in its application to gravitation).

In short, he rules out an extension of the principle to arbitrary gravitational fields on the grounds that an acceleration of the reference system can only produce gravitational fields of a quite special kind. Such comments appear quite frequently in Einstein's writings, throughout his life. They appear in his publications [3] and in his correspondence, right up to the last years of his life.[4] What might seem striking to the modern reader here is Einstein's failure to consider the possibility of transforming away arbitrary gravitational fields in infinitesimal regions of space-time. The omission was not a peculiarity of this particular discussion of the principle, for I have been unable to find any sustained treatment by Einstein of such an extension of the principle.[5] Nevertheless we can readily infer Einstein's attitude to this possibility. In Section 9, we shall see that he believed that one cannot distinguish the motion of a point-mass uninfluenced by a gravitational field from other motions if one considers only infinitesimal regions of the manifold. It follows immediately from Einstein's comments above that it is meaningless to talk in any thoroughgoing sense of transforming away a gravitational field in such infinitesimal regions.

The task of explicating Einstein's formulation of the principle of equivalence and even some of the preceding discussion is by no means straightforward. To begin, we must deal with Einstein's failure to maintain such distinctions as those between frames of reference and coordinate systems and between three-dimensional and four-dimensional concepts.[6] For example, we shall see that when Einstein speaks of a four-dimensional coordinate system, he may be referring to a four-dimensional coordinate system *simpliciter*, a frame of reference, or even a three-dimensional space associated with the frame. In the following section, I will introduce sufficient formal apparatus to deal with this problem, and then with it, we shall find that there is little difficulty in understanding Einstein's intentions. Then we can turn to ask precisely what Einstein means when he talks of a gravitational field produced by acceleration and in what sense the associated states of acceleration can be regarded as being "at rest".

## 3. On Reference Systems and Relative Spaces

In this section, I will deal with structures associated with the semi-Riemannian manifolds of special and general relativity.

In such manifolds, it is now customary to represent the intuitive notion of a physical frame of reference as a congruence of timelike curves. Each curve represents the world line of a reference point of the frame. The velocity of these points is given by the tangent vectors to the curves, where defined. We shall usually deal with frames of reference in rigid-body motion and we can readily nominate the state of motion of such frames because of the limited number of degrees of freedom associated with them.[7] In particular, an inertial frame of reference in a Minkowski space-time is a congruence of time-like geodesics in rigid-body motion. and therefore its reference points move with constant velocity.

A coordinate system $\{x^i\}(i = 1, 2, 3, 4)$ is said to be "adapted" to a given frame of reference just in case the curves of constant $x^1$, $x^2$, and $x^3$ are the curves of the frame. These three coordinates are "spatial" coordinates and the $x^4$ coordinate a " "time" coordinate.

With these definitions, Einstein's talk of "accelerated coordinate systems" can be made precise. A coordinate system is "accelerated" just in case it is adapted to an accelerating frame of reference. In this manner of speaking, a transformation from one frame of reference to another can be represented at least locally by a transformation between coordinate systems adapted to each frame.

Similarly we can represent the "Galilean" reference system mentioned in the last section as a coordinate system in Minkowski space-time, adapted to an inertial frame of reference and chosen so that the metric has components $\mathrm{diag}(-1, -1, -1, c^2)$, where $c$ is a positive constant—the coordinate speed of light. In such a coordinate system, differences of coordinates along curves, for which all but one coordinate is held fixed, are equal to the proper time or proper length of that segment of the curve, according to whether the curve is space-like or time-like. This implements Einstein's requirement that the coordinates be given directly by clock readings and measuring operations with rigid rods.

Presumably Einstein required thc coordinates of his accelerated coordinate systems to have as much of a similar direct metrical significance as was possible. Methods and scope for constructing analogous coordinate systems in the context of Newtonian theory and special and general relativity are well known (sec, for example, Friedman 1983, pp. 79—84, 129–135, 181–183).

However this discussion of Galilean and other systems in four-dimensional space-time does not entirely capture Einstein's intentions. He was also concerned with certain three-dimensional spaces, which are alluded to throughout his discussion of the principle of equivalence. It is appropriate to call these spaces "relative spaces", because of their similarity to the "relative space" Newton defined to contrast with his absolute space (Newton 1729, p. 6).[8] Einstein himself introduces the concept of this space in the introductions to his accounts of relativity theory, where it is presented as our most primitive notion of space (Einstein 1922, pp. 3–4,. 1954a, pp. 5–8). It arises through our experience that a given physical body can be extended by bringing other bodies into contact with it. The space of all such possible extension is the relative space of the body.

lf we think of the time-like curves of a frame of reference as the world lines of physical bodies, then these bodies define a single relative space, insofar as each of the bodies can be extended to contact any other body of the frame. The geometric properties of this space can be investigated in the familiar manner by laying out infinitesimal rigid rods, which are at rest in the frame. An example of this, which Einstein discussed frequently, is the relative space of a uniformly and rigidly rotating frame of reference in Minkowski space-time. In particular one finds there that the geometry of the relative space is non-Euclidean.[9]

The properties of the relative space defined by a given frame of reference can be precisely specified, although not in general by isomorphism with a three-dimensional hypersurface in the space-time manifold with the associated induced geometrical structure. The nature candidates for such hypersurfaces—the three-dimensional hypersurfaces orthogonal to the curves of the frame of reference—simply fail to exist if the frame of reference is rotating even in Minkowski space-time, for example.

Rather, we formally define the relative space $R_F$ of a frame of reference $F$ in a four-dimensional manifold $M$ as follows. $F$ defines an equivalence relation $f$ under which points $p$ and $p'$ of $M$ are equivalent if and only if they lie on the same curve $c$ of $F$. The relative space $R_F$ is the quotient manifold $M/f$ and has the curves of $F$ as elements. Coordinate charts of $R_F$ are inherited directly from the coordinate charts of $M$, which are adapted to the frame, ensuring that $R_F$ has a well-defined local topology. That is, if $\{x^i\}(i = 1, 2, 3, 4)$ is a chart in a neighborhood of $M$ adapted to $F$, then there will be a chart $\{y^i\}(i = 1, 2, 3)$ in the corresponding neighborhood of $R_F$ for which $y^i(c) = x^i(p)(i = 1, 2, 3)$ whenever $p$ lies on $c$.

A positive-definite metric $g_r$ is induced on $R_F$ as follows. At any point p on c we define the (unique) orthogonal metric $g_{\text{orth}}$ as the restriction of the space-time metric $g$ to any three-dimensional hypersurface $H_c(p)$ orthogonal to $c$ at $p$. A diffeomorphism $h$, which maps points of $H_c(p)$ in a neighborhood of $p$ to points in a neighborhood of $c$ in $R_F$, is such that, if $p'$ lies on the curve $c'$ of $F$, then $h(p') = c' \cdot g_r$ at $p$ is defined as the image of $g_{\text{orth}}$ at $p$ under $h$.[10] (Intuitively, we take $g_r$ to be the three-dimensional spatial metric revealed to an observer co-moving with the frame through the laying out of infinitesimal rods.)

Since point $p$ of $c$ here is chosen arbitrarily, it is clear that the resulting induced metric will only be uniquely defined in certain special cases. These special cases turn out to be just those in which the frame of reference is in rigid-body motion, for the requirement of rigid-body motion can be expressed as the requirement of constancy of the orthogonal metric along the world lines of the body. More specifically, what is required is the vanishing of the Lie derivative of $g_{\text{orth}}$, that is, $L_V g_{\text{orth}} = 0$, where $V$ is the tangent vector field of $F$.[11] General relativity deals with space-times that do not always admit rigid-body motions. Obviously, in these cases we will be unable to construct a relative space with a well-defined metric. To deal with the phenomena Einstein considers, we need to define a few more structures in these relative spaces. A gravitational field will be represented by a scalar field in nearly all the cases we need consider. A moving point-mass M will be represented by a scalar, its rest mass, and an appropriately parameterized curve $C$, its trajectory in the relative space $R_F$. $C$ can be inferred readily from the points of intersection of $M$'s world line with the time-like curves of the frame. That is, if $M$'s world line $c$ at parameter value $x$ intersects the curve $c'$ of frame $F$, then $C$ is the map that takes $x$ to $c'$. The velocity and acceleration vectors of $C$ can now be defined in the usual way. lf $c$ is parameterized by proper time, we would then arrive at the point-mass's proper velocity and proper acceleration.

In certain important special cases, it is possible to introduce a "frame time" into the relative space $R_F$ of a frame $F$. These cases are those in which the relevant neighborhood of the manifold can be foliated by a family of hypersurfaces, orthogonal to the curves of the frame $F$. Pick any curve $c$ of $F$, parameterized by proper time. Informally, we shall think of this curve as the frame clock of $F$ and its relative space $R_F$. Disseminate the time it marks by the following procedure. Define a scalar field $T$ on the space-time manifold whose constant-value hypersurfaces coincide with the hypersurfaces of the foliation and whose value agrees with the proper-time parameterization of $c$. Of course $T$ will only be be defined up to an additive constant.

This frame-time can now be transferred to the structures defined in $R_F$ by obvious means. For example the trajectory $C$ of a moving point-mass $M$ in $R_F$ can be parameterized by $T$, if $T$ is also used to parameterize $M$'s world line in the procedure for constructing $C$. From this parameterization, we would then arrive at $M$'s frame velocity and frame acceleration. Through a similar procedure, a time-varying field in $R_F$, induced by a field defined in the space-time manifold, can be represented by a family of fields indexed by $T$. The parameterization and indexing of structures in $R_F$ by $T$ gives a criterion of simultaneity.[12]

Clearly, in general we shall not be able to define a frame time. A rotating frame, for example, has no orthogonal hypersurfaces. Even if there are such hypersurfaces, the frame time may not be unique. A rigid, uniformly accelerating frame in Minkowski space-time admits orthogonal hypersurfaces; but the frame times defined by each of its curves differ by a multiplicative constant, although they yield the same simultaneity criterion. However, if the frame is an inertial frame in Minkowski space-time then the same frame time is defined by all curves of the frappe, up to an additive constant.

We can recover a "standard formulation" of special relativity—corresponding to the original three-dimensional formulation of the theory introduced by Einstein in 1905—by writing the laws that govern physical processes in Minkowski space-time in terms of structures defined within the relative space of an inertial frame, using the relative space's frame time. This formulation will hold just in any relative space of an inertial frame. Quantities describing the same process viewed from two different inertial relative spaces will be related by the Lorentz transformation in the familiar manner.

Generalizing, we construct a standard formulation of a four-dimensional space-time theory, in any given relative space that admits a frame time, by re-expressing its laws in terms of structures defined in the relative space, parameterized where necessary by the frame time. Thus we can construct a standard formulation of special relativity in the relative space of a rigid uniformly accelerating frame—and it will look quite different from the standard formulation associated with an inertial frame.

Einstein commenced his description of the principle of equivalence in his reply to Kottler by mention of space-time. It is now clear, however, that the phenomena he proceeded to describe are considered in relation to the relative spaces of the frames of reference. An isolated material point in a Galilean system can only be properly described as "mov[ing] uniformly and in a straight line" in the relative space. There it is represented by a geodesic of the relative space ("straight line"); its proper time and its frame time parameterization are directly proportional to the metrical distance along the curve ("move uniformly"). Use of either parameterization in this way also gives two general definitions of "uniform straight-line motion" in relative spaces, which agree in this case.

Similarly it is more natural to understand Einstein's requirement that the coordinates of the Galilean system be "measured directly in the well-known way" with rods and clocks as referring to operations described in the relative space and out of which the Galilean space-time coordinate system is constructed.

But most important of all, when Einstein speaks of "the presence of a gravitational field" in his reply to Kottler, clearly we should understand it to be present in the relative space of the frame of reference in question. In Minkowski space-time, there is a gravitational field in the relative space of the accelerated reference system but not in the relative space of the Galilean system. This is certainly more satisfactory than trying to speak of the presence of a gravitational field in space-time in this context. For then we would have to assume that a change of frame of reference can "produce" a gravitational field in space-time even though it does not change the world line of the point-mass on which the newly produced field is supposed to act.

This somewhat cumbersome mixture of three- and four-dimensional concepts in Einstein's formulation of the principle of equivalence derives directly from the fact that, for the first five years of its life, the principle and the gravitation theory associated with it were treated entirely within the same three-dimensional formalism Einstein had used in his 1905 special relativity paper. In particular, the spaces Einstein dealt with in this period were invariably the relative spaces of frames of reference. Nevertheless, Einstein's 1916 formulation and his original 1907 formulation of the principle read almost identically, even though the former was associated with a theory that could not readily be written in a three-dimensional formalism. In the following section I turn to examine this early period of Einsteinian work. I will be concerned with showing precisely which structures Einstein chose to represent the gravitational field in the relative spaces he dealt with.

# 4. A New Theory of Gravitation

## 4.1. A New Concept of Gravitational Field

Einstein made clear from the inception of the principle of equivalence in 1907 that its main purpose was to enable the extension of the principle of relativity to accelerated motion. [13] But for the five years following 1907, his actual use of the principle involved the development of a novel relativistic theory of static gravitational fields out of which his general theory of relativity would emerge in 1912 and 1913. The principle assured him that a certain structure ("inertial field") arising in the relative space of a uniformly accelerated frame of reference in Minkowski space-time was just one special type of gravitational field. The properties of this structure could be examined minutely using the known results of special relativity and the properties of other types of gravitational fields could then be inferred.

That this structure (whose properties will be developed and outlined in Section 4.2) could be regarded as a gravitational field requires a change in our understanding of what a gravitational field is. We must now accept that gravitational fields can have an existence dependent on the relative space considered and that the choice of relative space may decide whether or not a single given process is regarded as acted on by a gravitational field. The obvious objection, which was put by Laue to Einstein in 1911, is that this type of gravitational field cannot be "real" since it has no source masses.[14] Einstein's later response to this objection was that it is essential to field theory to be able to conceive of fields, such as gravitational fields, as existing independently of their sources.[15]

In effect, Einstein asks us to give up the familiar concept of gravitational field as that which mediates the gravitational interaction of bodies. In its place in the relative space of frames of reference, regardless of whether they are accelerated or not, we infer the existence of a structure that is responsible for the deviations from uniform straight-line motion of a free point-mass, without concerning ourselves with what generates that structure. Following Einstein's lead, we would take such a structure to be a gravitational field by definition, if the deviations associated with it are independent of the point's mass.

Using this definition, we could now describe as gravitational fields the inertial fields arising in relative spaces of rigid frames of reference in arbitrary states of acceleration in Minkowski space-time. It is difficult to imagine that Einstein would contradict this result. Nevertheless, as I have pointed out, he formulated his principle of equivalence only for the case of uniform acceleration.

There were most probably several reasons for this additional restriction. In the early years of the principle of equivalence, in order to convince skeptical contemporaries that inertial fields could be regarded as gravitational fields, he had to show that they behaved exactly like known gravitational fields—that is, like Newtonian gravitational fields—aside of course from the question of source masses. If the principle of equivalence is formulated in a Newtonian space-time, as Einstein did sometimes in these earlier years, [16] the requirement that the inertial field behave exactly like a Newtonian field places severe restrictions on the allowed states of motion of the frame of reference.

In Newtonian mechanics, the inertial field induced on the relative space of a rotating frame of reference contains a Coriolis field, which exerts a force on a body dependent on its velocity. A structure representing such a field will contain vector potentials, such as those arising in electromagnetic theory, rather than the familiar scalar potential of the Newtonian gravitational field.[17] The inertial field induced on the relative space of a frame of reference in rectilinear acceleration can be represented by a scalar potential satisfying Laplace's equation. But if the acceleration is not uniform the resulting field will be nonconservative due to the explicit time dependence of the potential.

In this case of a Newtonian space-time, we are led directly to Einstein's choice of a uniformly accelerated frame of reference for the formulation of the principle of equivalence. For only in this case will the structure concerned in the relative space behave exactly like a Newtonian gravitational field. It will be a scalar field, it will satisfy Laplace's equation, and its gradient will be equal to the acceleration of otherwise free point-masses in the space.

It would be natural for Einstein to continue to formulate the principle of equivalence in terms of the special case of uniform acceleration in Minkowski space-time as well, if only in the interests of continuity. In addition, we can identify at least three complexities arising with the use of rotating frames of reference or those in nonuniform acceleration in Minkowski space-time.

First, the associated relative spaces would have non-Euclidean geometries, if they were well defined. This was a problem Einstein was well aware of from a very early stage. But he treated it as a separate issue from his principle of equivalence, usually by consideration of a rotating frame of reference.

Second, he would be unable to introduce a frame time into the relative space, making very difficult the description of phenomena in the space by a standard formulation of a theory such as he used in 1907–1912.

Third, the trajectory of a light signal exchanged between two points in the relative space would differ on the forward and return journeys. In a letter of June 1912 to Ehrenfest, in which Einstein discussed the failure of his 1912 gravitation theory to deal with the fields associated with rotating frames of reference, he mentioned this failure or the "reversibility of light paths" in such fields and described how dealing with them would be the next step (EA 9–333).

In any case, after the completion of the general theory of relativity, when the difficulties of the earlier gravitation theory had been resolved, there is a suggestion in one or two places in Einstein's writings that he was prepared to extend the formulation of the principle to the case of frames of reference in rotation or nonuniform acceleration (for example, Einstein 1922. p. 59; l952, pp. 151–154).

## 4.2. The 1907–1912 Theory

Einstein's 1907–1912 theory of static gravitational fields achieved its most developed form in two consecutive papers in the latter year (Einstein 1912a; 1912b). The theory may be represented most precisely in four-dimensional terms, although Einstein had not yet begun to use them. It was based on exploiting certain especially simple properties of uniformly accelerating frames of reference in Minkowski space-time. These special properties can be derived from the result that one can always find a coordinate system $\{x^i\}(i = 1, 2, 3, 4)$ adapted to a uniformly accelerating frame in Minkowski space-time in which the metric has the form

$$\mathrm{diag}(-1, -1, -1, c^2)$$

where $c = 1 + bx^1$ and $b$ is a constant. It follows immediately that the geometry of the relative space is Euclidean, inheriting the coordinates $\{x^i\}(i = 1, 2, 3)$ as Cartesian coordinates. Further, the space-time can be foliated by a family of hypersurfaces orthogonal to the frame, the hypersurfaces of constant $x^4$. Therefore we can introduce a frame time.

For convenience, select the world line of the frame for which $x^1 = x^2 = x^3 = 0$ as the frame clock and call $t$ the frame time disseminated by it. The choice as frame clock of any of the other world lines of the frame would alter $t$ by a constant multiplicative factor and thus not materially affect the results.

Thus Einstein could introduce a standard formulation of special relativity in the relative space. In particular, it followed in this standard formulation that the motion of a free point-mass, whose world line was a geodesic in the space-time, was governed by the equation

$$d/dt(\beta v^i/c) = -\beta \partial c/\partial x^i,$$

where $\beta = 1/(1 - v^2/c^2)^{1/2}, v^i = d/dt(x^i)$ is the three-velocity of the point-mass, and $v$ is its magnitude.

This relation closely parallels the relation

$$\text{acceleration} = -\text{gradient of scalar field}$$

governing the motion of a freely falling point-mass in traditional Newtonian gravitation theory and in which the point's mass also does not appear. Thus in accord with the discussion of Section 4.1, Einstein could view the motion of the point-mass in the relative space as under the influence of a gravitational field whose scalar potential was $c$ and which was responsible for the deviations from uniform straight-line motion.

Note that while the scalar field $c$ was introduced earlier via the $g_{44}$ component of the Minkowski metric in a particular coordinate system, it can be described in coordinate-free terms: $c$ is just the Minkowski norm of the tangent four-vector of the curves of the frame, when parameterized by the frame time. It can be seen that c will have a constant value along each of these curves and therefore a unique, well-defined value at each point of the relative space.

Recalling that the coordinates $\{x^i\}(i = 1, 2, 3)$ are inherited as Cartesian coordinates by the Euclidean relative space, the relation $c = 1 + bx^1$ now can be seen to assert that the gravitational potential $c$ varies linearly with (Euclidean) distance in one direction in the relative space. This is exactly the way a traditional Newtonian potential behaves in the case of inhomogeneous gravitational field.

There were some complications however, in addition to the usual relativistic corrections; $c$ turned out to be the isotropic speed of light in the relative space, measured with frame time, which it now followed must also vary with position in the relative space. It could be shown that the rates or clocks at rest in the relative space would vary with $c$ and, therefore, with position.

Now that Einstein had a firm grasp on relativistic gravitational fields in the one special case of homogeneous fields, it was a simple matter to infer the properties of arbitrary static gravitational fields by a natural and hopefully unproblematic generalization. To do this, Einstein left the standard formulation of the theory unchanged, except for relaxing the condition that $c$ vary linearly with distance in the direction of acceleration. Following the model of Newtonian theory, he now required that $c$ satisfy a weaker condition, the field equation

$$\Delta c = \kappa c \sigma,$$

where $\sigma$ is the mass density and $k$ a constant.

This step amounted to the transition to the relative spaces of more general semi-Riemannian manifolds with static space-time metrics of Lorentz signature. The relative spaces are those of frames of reference whose velocity vectors are Killing vector fields. The metric must be static rather than just stationary, since the space-time must admit a foliation by a family of hyper-surfaces orthogonal to these frames, in order for a frame time to be defined for use in the standard formulation. The requirement that the relative spaces still be Euclidean further restricts the space-time metric to those whose orthogonal metrics are Euclidean.

It follows that there always exists a coordinate system $(x, y, z, t)$ adapted to the frame in which the space-time metric has the form $\mathrm{diag}(-1, -1, -1, c^2)$ and the relative space inherits the coordinates $\{x, y, z\})$ as Cartesian coordinates. As a result, Einstein's 1912 theory is sometimes described as a theory of space-times with the line element

$$ds^2 = -dx^2 - dy^2 - dz^2 + c^2dt^2,$$

where $c = c(x, y, z)$, although his theory actually deals with the relative spaces of such space-times.

It is interesting that the field equation chosen here for the relative space corresponds to the field equation for the space-time metric

$$R = k'T$$

where $R$ is the Riemann curvature scalar, $T$ is the trace of the stress-energy tensor of a dust cloud, and $k'$ is a constant, although when Einstein formulated his theory he could not have known this.

In the second of the 1912 papers cited, Einstein described the difficulties his bold new theory soon encountered. In order to retain the equality of action and reaction of forces, that is, to retain a law of momentum conservation, Einstein found himself forced to a modified field equation

$$\Delta\sqrt{c} = (k/2)\sqrt{c\sigma}.$$

This new field equation no longer admitted the homogeneous field associated with uniform acceleration in Minkowski space-time as a solution, unless one considered only infinitely small regions of the relative space. Einstein confessed that he had resisted this development, since it now meant that his principle of equivalence could only be Formulated in infinitely small regions of the relative space, even though it still dealt only with the simplest case of uniform acceleration in Minkowski space-time.[18]

## 4.3. The Temporary Limitation To Infinitesimal Regions

Because of the superficial similarity between this version of the principle and the infinitesimal principle of equivalence now common in the context of arbitrary gravitational fields in general relativity, some writers have regarded this development as, for example, "the dawn of the correct formulation of the principle of equivalence as a principle that holds only locally" (Pais 1982, p. 205). It certainly was not as far as Einstein was concerned. The limitation to infinitesimal regions of the relative space was not introduced to homogenize inhomogeneous fields, as it is in the modern infinitesimal principle. His principle still dealt only with homogeneous fields produced by uniform acceleration. (Note that the inhomogeneous fields of his 1912 theory were not produced by acceleration but by generalizing the properties of homogeneous fields.) Therefore, the need for such a limitation, in the case of fields that were *already* homogeneous, was a source of some puzzlement to him and he dispensed with it as soon as he could. But before he could, there were yet more problematic developments concerning the principle of equivalence. I relate them here in the hope of nipping in the bud the myth of Einstein's 1912 introduction of the modern infinitesimal principle of equivalence.

In late 1912 and early 1913. In this climate of uncertainty about the principle, Einstein made his major breakthrough to the *Entwurf* theory with the mathematical assistance of his friend Marcel Grossmann (Einstein and Grossmann 1913). The new theory contained virtually all the essential features of the final general theory of relativity. However, they were unable to incorporate generally covariant gravitational field equations in it. Einstein was able to remove this defect only after nearly three years of intense work and thereby arrived at his final general theory or relativity (see Norton 1984).

During this period, Einstein omitted to mention the catastrophe that had befallen the principle of equivalence. Because of their restricted covariance, it can be shown that the field equations of the *Entwurf* theory do not hold in coordinate systems adapted to uniformly accelerating frames of reference in Minkowski space-time, even allowing restrictions to infinitely small regions of space-time. In the language of Einstein's 1916 formulation of the principle in his reply to Kottler, this meant that he could not regard such coordinate systems as "at rest". That is, according to his new theory, the principle of equivalence was false if formulated for this standard and simple case.

Therefore, in the introduction to the *Entwurf* paper, Einstein had to present the principle of equivalence as a result drawn from his earlier theory of static fields; for he still based the principle on the assumption that a uniform acceleration of the reference system in Minkowski space-time produced a homogeneous gravitational field even if only in an infinitely small region of the relative space. Presumably because of this problem, Einstein avoided the detailed discussion of the equivalence of the inertial field of uniform acceleration and homogeneous gravitational fields in the three years in which he held to the *Entwurf* theory, for this theory entailed no such equivalence. But he retained the principle of equivalence, for it was essential to the conceptual development of his theory. In addition, the notion of the equivalence of inertial and gravitational fields was central to the theory. However, the extent to which his *Entwurf* theory admitted this equivalence was not entirely clear.

This difficulty was resolved dramatically and completely with Einstein's November 1915 adoption of the generally covariant field equations of his completed general theory of relativity. The restriction of the principle of equivalence to infinitely small regions of space disappeared from his writings.

## 5. Extending the Principle of Relativity

Einstein's early success in constructing a new gravitation theory from his principle of equivalence is partly responsible for the still prevalent misconception that this was its essential purpose. To combat this, he frequently stressed that the principle did not provide a recipe for producing arbitrary gravitational fields by acceleration. The real point of the principle, as he had made clear in 1907, was that it enabled an extension of the principle of relativity to accelerated motion. Thus in the 1916 formulation of the principle quoted in Section 2, the principle itself is "the assumption that one may treat [the uniformly accelerated reference system] $K'$ as at rest in all strictness without any laws or nature not being fulfilled with respect to $K'$."

Prior to 1913 and the development of the basic formal structure of the general theory of relativity, Einstein gave no sustained discussion of precisely what he required in an extension of the principle of relativity and how the principle of equivalence was to help bring it about. However, we can reconstruct Einstein's position on these questions in this early period by considering the discussion he gave in an introductory section of his 1916 review of the general theory of relativity, called "On the grounds which suggest an extension of the postulate of relativity" (Einstein 1916a, pp. 771–773). This section concluded with a formulation of the principle of equivalence. Further, it dealt only with concepts that would have arisen in the pre-1913 period, suggesting that he was rehearsing arguments essentially from this period of his work. In particular, the discussion focused exclusively on the relative spaces of frames of reference.

Einstein began by pointing out an "epistemological defect" of classical mechanics and special relativity, enabling us to locate his arguments in Newtonian and Minkowski space-times. In a celebrated thought experiment, he considered two fluid spheres in relative rotation and noted that only one of them can be free of centrifugal distortion. But there is no observable difference between the relative spaces or the rest frames of each sphere, other than the state of motion of the distant masses of the universe, in which, he concluded, the cause of the centrifugal distortion is to be sought. This led to the following requirement for relative spaces

> Of all imaginable spaces $R_1, R_2$, etc, in any kind of motion relatively to one another, there is none which we may look upon as privileged *a priori* without reviving the above-mentioned epistemological objection. *The laws of physics must be of such a nature that they apply to systems of reference in any kind of motion.* (Einstein 1916a, p. 772)

Einstein then prceeded to formulate the principle of equivalence that enables a *uniformly* accelerated observer to avoid inferring that he is "really" accelerated and enables us to regard the uniformly accelerated reference system $K'$ as just as "privileged" or "stationary" as the unaccelerated system $K$.

Since Einstein's discussion was in terms of relative spaces, it is clear that the "laws of physics" were being considered in their "standard formulations" described in Section 3. The standard formulations of classical mechanics and special relativity in question would be those then generally available, that is, those defined in the relative spaces of inertial frames (henceforth "inertial spaces"). These standard formulations would hold *only* in inertial spaces and therefore fail to satisfy Einstein's requirement that they "apply to [the relative spaces of] systems of reference in any kind of motion". Thus they would single out inertial spaces and their associated inertial frames as privileged.

In response, Einstein used the principle of equivalence to propose a more general theory, a theory of homogeneous gravitational fields, whose standard formulation will hold not only in inertial spaces but in uniformly accelerated spaces as well. The relativistic version of this theory is quite familiar to us now from Section 4 and presumably also to Einstein's readers of 1916. It is just his 1907–1912 gravitation theory, restricted to the case of a homogeneous gravitational field. In this way, Einstein broadened the set of privileged frames and relative spaces to include those in uniform acceleration.

Precisely what Einstein achieved with this result has not always been properly understood. His point can be made more clearly by avoiding reference to the standard formulation of theories, which has proven to be confusing to modern readers steeped in the four-dimensional formulation of these theories.

The focus of Einstein's concern is the necessity in special relativity and classical mechanics of presuming an immutable division of relative spaces and frames of reference into the privileged inertial and the noninertial. The principle of equivalence enabled him to eliminate the immutability of this division, by reinterpreting the nature of the inertial effects which distinguish the privileged inertial spaces and frames from all others. He explained this to a correspondent in a letter of July 12, 1953, reminding him that the principle could not be used to generate arbitrary gravitational fields by acceleration:

> The equivalence principle does not assert that every gravitational field (e.g., the one associated with the Earth) can be produced by acceleration of the coordinate system. It only asserts that the qualities of physical space, as they present themselves from an accelerated coordinate system, represent a special case of the gravitational field. It is the same in the case of the rotation of the coordinate system: there is *de facto* no reason to trace centrifugal effects back to a 'real' rotation.[19]

Through the principle of equivalence, Einstein proposed that we do not regard these distinguishing inertial effects as depending on an immutable property of the accelerating relative space, but as arising from the presence of a field in the relative space, which was to be seen as a special case of the gravitational field. This view could be extended beyond the case of uniform acceleration of the principle. Within this view, relative spaces would have no intrinsic states of motion—none would be "really" rotating for example—and in this sense they would all be indistinguishable. However, any relative space could become inertial according to the particular instances of the gravitational field defined on the relative spaces. Similarly, all frames of reference would be indistinguishable, until the introduction of any particular instance of the gravitational field made some inertial and others not.

This crucial aspect of Einstein's account has been commonly misunderstood. The fact that an accelerated frame remains distinguishable from an unaccelerated frame in both special and general relativity is irrelevant to the extension of the principle of relativity. Einstein's account *requires* that each instance of the gravitational field distinguish certain frames as inertial and others as accelerating. The decision as to which frames will be inertial and which accelerated, however, must depend only on the particular instance of the gravitational field at hand and not on any intrinsic property of the frames.[20]

At this stage of his development of general relativity, Einstein's important innovation did not yet lie in the introduction of an empirically new theory. According to the principle of equivalence, his theory of static gravitational fields was predictively identical to special relativity in the case of homogeneous gravitational fields. Rather, it lay in a new way of looking at the division of structures between space and the fields it contains in the context of special relativity. Specifically, he no longer regarded the structures accounting for inertial effects as a part of space. Rather he now looked upon them as associated with the fields defined in space and, in particular, intimately related to gravitation. This move stripped space of the privileged frames to which he objected.

Einstein's "*Gestalt* switch" can be described more precisely if we present it more explicitly in four-dimensional terms. Of course, Einstein himself did not begin to work explicitly in such terms until five years after his original 1907 formulation of the principle of equivalence.

In the old view of special relativity, the background arena of space and time, against which physical processes unfold, is a Minkowski space-time, that is, a pair: $\langle M, g\rangle$, where $M$ is a four-dimensional manifold and $g$ a Minkowski metric. This background arena admits certain privileged structures: inertial frames of reference and their associated inertial spaces.

In the new view of special relativity, we are informed by the principle of equivalence that the structure responsible for inertial effects, the Minkowski metric $g$, is not an intrinsic part of the background arena of space and time. Rather, it is a field defined against that background and actually a special case of the field structure that also accounts for gravitational effects. The background arena of space and time is now just the bare space-time manifold $M$. In $M$ in the absence of a metric, we can still introduce frames of reference as congruences of curves, although we cannot require them to be time-like, and we can still define their relative space, although they will have no induced metric. Clearly in terms of $M$ alone, all such frames and correspondingly all relative spaces will be indistinguishable and therefore none will be privileged.

Following the model of classical gravitation theory, special relativity in this new view circumscribes the metric fields allowed on the manifold by a differential field equation. It requires a metric of Lorentz signature and with a vanishing Riemann curvature tensor

$$R_{iklm} = 0$$

This requirement does not specify a unique Minkowski metric, but a large set of Minkowski metrics. Because of this, the *theory* does not single out any frame of reference as privileged in a particular "background space" (i.e., manifold), even though each metric allowed by the theory will single out certain frames as inertial and others as noninertial. For, speaking informally, it can be shown that there is always a Minkowski metric allowed by the theory in which any well-behaved noninertial frame would become inertial. This result, given more precisely later, rests entirely on an active interpretation of the general covariance of the preceding field equation.

In a space-time manifold $M$, let $g$ be a Minkowski metric and $F$ an inertial frame of reference, that is, one whose time-like curves are geodesics in rigid-body motion. Let $F'$ be any frame of reference in the neighborhood $U'$ of $M$ (or even any congruence of curves which need not be all time-like), for which there exists a coordinate system $x'^i$ with domain $U'$ adapted to $F'$. (Such a frame is "well behaved".) Now in some neighborhood $U$ of $M$ there exists a coordinate system $\{x^i\}$ adapted to $F$ whose range coincides with that of $x'^i$. $h$ is a diffeomorphism that maps $p$ to $hp$ such that $x^i(p) = x'^i(hp)$. Then it follows that $F'$ is an inertial frame of reference, with respect to the Minkowski metric $g'$, which is the image of $g$ under $h$.[21]

The essential features of the old and new way of viewing special relativity are summarized in Table 1.

TABLE 1. Comparison of old view of special relativity with new view informed by principle of equivalence.

| | Old view | New view |
|---|---|---|
| Background arena of space and time | Minkowski space-time $= \langle M, g \rangle$ where $M$ = four dimensional manifold $g$ = Minkowski metric | Four dimensional manifold $M$ only |
| Examples of contents/ processes in space and time | Electromagnetic fields, matter in dust clouds, etc. | Electromagnetic fields, matter in dust clouds, etc. Any Minkowski metric = special case of structure inducing gravitational fields |
| Priviledged frames of reference in background of space and time? | Yes, each $\langle M, g \rangle$ has a unique set of inertial frames. | No, bare manifold $M$ has no priviledged frames of reference. Any well-behaved frame can be made inertial by defining an appropriate Minkowski metric on $M$ |

The equivalence of all frames embodied in this new view goes well beyond the result that Einstein himself claimed in 1916 from the principle of equivalence. He claimed only an equivalence of inertial and uniformly accelerated relative spaces, that is, of inertial and uniformly accelerated frames. The establishment of a wider equivalence would have been straightforward, even if inessential in view of the fact that he had the general theory of relativity in hand by then. But he most likely chose to avoid this extension because it would have required him to find standard formulations of a gravitation theory, similar to his 1907–1912 theory, which would hold in relative spaces of frames in rotation or nonuniform acceleration. I listed some of the difficulties Einstein would face in this task in the last section.

In any case, Einstein could not simply take special relativity. viewed in the new way, as a theory extending the principle of relativity in the way required for two reasons. First, the principle of equivalence clearly indicated that the theory was not complete. The structure accounting for inertia must also account for all gravitational effects. The Minkowski metric of special relativity, however, could only account for effects due to gravitational fields which could be transformed away over some neighborhood of a relative space by transforming to a new relative space. So Einstein immediately continued from his statement of the principle of equivalence, quoted earlier from his 1916 review article, by observing that "in pursuing the general theory of relativity, we shall be led to a theory of gravitation.. . . " We shall see that it was the completion of this task that yielded the general theory of relativity.

The second reason was more subtle but far more important and can only be touched on informally here. The theory was also causally incomplete. As we have seen, Einstein required a complete theory of inertia to account for the disposition of inertial frames in space-time in terms of the only available *observable* cause, the distribution and motion of the masses of the universe.

Special relativity in any of the forms described cannot be that theory. The disposition of inertial frames and the Minkowski metric which determines them is completely unaffected by any change in these masses. In some large neighborhood of space-time, such changes might include the setting of all masses into rotation about a central axis or even the conversion of all their energy into radiation and its resulting dissipation.

However it was natural for Einstein to expect that the extended theory, which dealt with general gravitational effects, would explain the observed disposition of inertial frames of reference in terms of the matter distribution of the universe. For the structure that determined this disposition would behave in many aspects like a traditional gravitational field and therefore be strongly influenced by any motion of its sources, the masses of the universe.

Although Einstein's hopes were not borne out by later developments, he made clear in his earliest relevant publications that he expected his new general theory of relativity to implement a "hypothesis of the relativity of inertia", which required inertia to be nothing other than the resistance of a body to acceleration with respect to other bodies (Einstein 1913b, pp. 1260–1262). This, of course, would forbid universes, all of whose masses were rotating about a local inertial compass. He had already sought and found small effects he felt were consistent with this hypothesis. They included the dragging of the inertial frames of reference inside a rotating shell of matter and were similar to those discussed in his *Meaning of Relativity* (Einstein 1922, pp. 100–103). Clearly he also related this hypothesis to his 1907–1912 theory of static gravitational fields, for in 1912 he had published a paper which demonstrated the existence of similar such effects in that theory too (Einstein 1912d).

## 6. The Breakdown of Relative Spaces

It was inevitable that Einstein would give up the use of standard formulations of theories in his search for a general theory of relativity. For the relative spaces used by these formulations would only have well-defined geometries if the associated frame is in rigid motion, which is by no means generally the case. Even in Minkowski space-time, no *non*uniformly rotating frame can move rigidly. Worse, the relative space will only have the frame time required by standard formulations if the space-time admits a foliation by hypersurfaces orthogonal to the frame. Even uniformly rotating frames in Minkowski space-time do not admit such a foliation.

In his general theory of relativity, Einstein turned to the four-dimensional space-time formulation of theories. As indicated in the last section, he now also came to regard the four-dimensional space-time manifold without further structure as the background of space and time against which physical processes unfold.

One can define very few reference structures in such a manifold. Frames of reference as congruences of world lines can be defined. But without further structure, such as a metric, they cannot be described as time-like or have an overall state of motion assigned to them. The richest reference structure available is the arbitrary space-time coordinate system, whose coordinate values can have no metrical significance, such as Einstein had required in his Galilean reference systems.

So in the general theory of relativity, Einstein proceeded to use arbitrary space-time coordinate systems as the reference structures from which to view physical processes and formulate physical principles. In his expositions of general relativity, Einstein typically made this transition from frame of reference and relative space to arbitrary space-time coordinate system by considering the relative space of a frame of reference in uniform rigid rotation in Minkowski space-time (for example, Einstein 1916a, pp. 773–776,. 1922, pp. 59–62). He would show that the spatial geometry is non-Euclidean and conclude that the coordinate system used there could not have the same direct metrical significance of spatial coordinates in his Galilean reference systems. Similar results followed from attempts to retain a time coordinate, presumably for space-time, whose value would coincide with the readings of clocks at rest in the frame. Einstein then introduced the use of arbitrary space-time coordinate systems as a natural extension of the methods developed in the nineteenth century for dealing with non-Euclidean spatial geometries.

This argument gave psychologically natural grounds for introducing the methods of differential geometry into relativity theory. However, it failed to demonstrate the completeness of the demise of relative spaces in general relativity. The relative space of the argument's uniformly rotating frame of reference still has a well-defined geometry, unlike the relative spaces of other frames of reference in space-times with more general semi-Riemannian metrics. Einstein turned to this problem in his popularization *Relativity* (1954a), most of whose discussion is set in terms of the relative spaces of "reference bodies" ( = frames of reference). In chapter 28 he points out that rigid reference bodies will in general no longer be available in general relativity and that "the Gauss coordinate system has to take the place of the body of reference". He then proceeds to describe the difficulties and artificiality of retaining the use of nonrigid reference bodies (and by implication their associated relative spaces with ill-defined geometries) through the discussion of what he calls "reference molluscs".

In the same chapter, Einstein gave his well-known reformulation of the extended principle of relativity—"All Gaussian co-ordinate systems are essentially equivalent for the formulation of the general laws of nature"—and proceeded to explain that this requirement was satisfied by a theory if its laws were written in a generally covariant form. Naturally, this meant that his generally covariant general theory of relativity realized the extended principle of relativity.

Einstein has taken the principle of equivalence to assert the equivalence of inertial and uniformly accelerated relative spaces, an assertion that is subsumed by the extended principle of relativity. So it was easy for Einstein to conclude, in continuing his reply to Kottler, that the principle of equivalence was automatically satisfied by his general theory of relativity:

> A gravitation theory violates the principle of equivalence, in the sense which I understand it, only then, if the equations of gravitation are satisfied in no reference system $K'$, which is moving non-uniformly relative to a Galilean reference system. That this reproach cannot be raised against my theory with *generally* covariant equations is evident; for here the equations are satisfied with respect to each reference system. *The requirement of general covariance of equations embraces the principle of equivalence as a quite special case.* (Einstein 1916b, p. 641) [22]

Einstein's reformulation of the extended principle of relativity as the requirement of general covariance is unproblematic in so far as it is based on the fact that the space-time manifold without any additional structure has no privileged coordinate systems. This fact immediately entails that there are no privileged frames of reference and, therefore, no privileged relative spaces. For were any frames privileged, the coordinate systems adapted to them would also be privileged.

However, as has been frequently objected, it is hard to see how this requirement could capture all that Einstein required in an extension of the principle of relativity, when there are simple generally covariant formulations of many other theories apart from general relativity. These include special relativity, Nordström's theory of gravitation, and Newtonian gravitation theory. Of course Einstein was aware of this at least in the case of the first two theories.

A thorough analysis of Einstein's intentions here and their refinement in his later work is a complex task that goes well beyond this paper. Nevertheless, I will make a few tentative comments concerning Einstein's early view of the question to make his remarks more plausible.

For Einstein, violations of the extended principle of relativity need not be limited to the laws of a theory. They could also arise in its solutions, that is, in models or classes of models of the theory. For example he pointed out in a 1917 paper on the cosmological problem that it was "contrary to the spirit of the relativity principle" to introduce solutions of the field equations of general relativity by imposing a boundary condition of a Minkowski metric at matter-free spatial infinity (Einstein 1917, p. 147). This introduces privileged coordinate systems in which the metric approaches the form diag(-1, -1, -1, 1) as the limit to spatial infinity is taken. In addition, these privileged coordinate systems were objectionable since there was no observable cause for their special status, contradicting the hypothesis of the relativity of inertia.

Clearly, solutions of generally covariant formulations of special relativity and Newtonian theory would necessarily involve the introduction of similarly objectionable privileged coordinate systems in one form or other. Minkowski space-time, even regarded as a model of general relativity, would be objectionable for the same reason. However, Einstein believed that the introduction of these boundary conditions would not always be needed in the case of his general theory of relativity. In his 1917 paper he continued to demonstrate how the field equations of general relativity, augmented with the cosmological term, admitted solutions without the use of boundary conditions at spatial infinity. To arrive at these solutions, one needed only to specify the mass and world lines of the universe's smoothed-out dust cloud of matter on the manifold and invoke other natural requirements, such as the symmetry of the metric with respect to these world lines, and its isotropy about them.

In 1918, Einstein described a solution generated in this way as satisfying "Mach's Principle" (Einstein 1918a, p. 241). This principle required that the metric tensor be determined completely by the matter of the universe and was taken to be the natural generalization of the hypothesis of the relativity of inertia. In a footnote, he pointed out that he had not previously distinguished this principle from the (extended) principle of relativity and that this had caused confusion. So, at least at this time, the general theory of relativity seemed to be the only viable theory satisfying all his requirements concerning the relativity of motion. It was clearly impossible for special relativity or Nordström's theory to exhibit such Machian behavior, irrespective of the covariance of their formulations.

# 7. Generating General Relativity

Einstein had come to recognize that a general theory of relativity was to be found as a four-dimensional theory of gravitation. The principle of equivalence provided the crucial starting point: the identification or the Minkowski metric as an *instance* of the four-dimensional space-time structure representing gravitational fields. For Einstein had found that the Minkowski metric can induce gravitational fields on the relative spaces of a Minkowski space-time.

Einstein's discovery of the gravitational properties of the Minkowski metric was a remarkable feat. Unlike so many other discoveries in physics. it seems to have been almost totally unanticipated by his contemporaries.

The role of the principle of equivalence in Einstein's development of his new gravitation theory remained essentially the same as in his earlier 1912 theory of gravitation. The principle yields a special case of the gravitational field, whose properties are then generalized in a natural way to arrive at a general theory of gravitation.

However, from the perspective of the general theory of relativity. Einstein had no prospect of arriving at the correct laws of a general theory of the gravitational fields of relative spaces, as long as he worked within the framework of his 1912 theory. This follows immediately if we recall that Einstein sought to characterize arbitrary static gravitational fields as structures induced onto relative spaces by the special type of static space-times I described in Section 4.2.

In these space-times, in the source-free case, one can readily demonstrate that the field equations of general relativity, that is, the requirement of the vanishing of the Ricci tensor

$$R_{im} = 0,$$

entails the vanishing of the Riemann-Christoffel curvature tensor

$$R_{iklm} = 0.$$

This in turn entails that the only source-free gravitational fields in relative spaces which the theory can deal with correctly, from the perspective of the general theory of relativity, are those induced by acceleration in Minkowski space-time. In addition, it follows from an evaluation of the components of the curvature tensor in a coordinate system adapted to the accelerating frame that this acceleration must be a uniform rectilinear acceleration.[23]

Unfortunately, in the period 1912 to 1915, Einstein believed that the arbitrary static space-times associated with his 1912 theory ought also to be solutions of the field equations of his new general theory of relativity. I have argued elsewhere in detail that this played a major role in his failure to adopt the generally covariant field equations of his final theory in this period. (sec Norton 1984).

Nevertheless, Einstein commonly used the principle of equivalence to recover and motivate the basic formal structure of his general theory of relativity in an argument whose strategy was essentially the same as that used in 1912. Einstein presents the argument in a compact and well-developed form in a 1951 letter to Becquerel, in which the role of the principle of equivalence is made especially clear.[24] He begins by using the equality of inertial and gravitational mass to justify introduction of the principle, which is formulated in terms of relative spaces: "An inertial space without gravitational field is physically equivalent to a uniformly accelerated space, in which there is a (homogeneous) gravitational field. (Equivalence hypothesis.)" Then after introducing the requirement of general covariance, he proceeds with the steps he numbers as the third and fourth of his argument:

> (3) One kind of space is completely known to us, that is empty Minkowski-space, in which the interval $ds$, as given by
>
> $$ds^2 = -dx_1^2 - dx_2^2 - dx_3^2 + dx_4^2$$
>
> can be measured immediately by resting clocks and measuring rods. Through a nonlinear transformation, this becomes
>
> $$ds^2 = g_{ik} dx_i dx_k,$$
>
> where $ds$ has the same value as a Minkowski system. The $g_{ik}$ depend on the coordinates and, according to the equivalence hypothesis, describe a gravitational field (of a more special kind),
> (4) In general coordinates, a gravitational field of the more special kind satisfies the differential equations
>
> $$R^i{}_{klm} = 0$$
>
> from the loosening of which the field law of an arbitrary pure gravitational field must follow. For this, only
>
> $$R_{kl} = R^s{}_{kls}$$
>
> comes into consideration. It is natural to assume that ds expresses the naturally measured interval also in the case of a general pure gravitational field.

Because of its extreme brevity, Einstein's argument requires some explication. In his step 3, he appears to identify a coordinate effect, the nonconstancy of the components $g_{ik}$, with the presence of a gravitational field. His real intention emerges, however, if we recall his practice of tacitly associating changes of frame of reference with coordinate transformations. In particular, a nonlinear coordinate transformation can represent the change from an inertial frame of reference to a rigidly and uniformly accelerated frame of reference, which is precisely the case considered in the statement of the principle of equivalence just given. In this case, the nonconstancy of the

$g_{ik}$ is now associated with the presence of a homogeneous gravitational field in the relative space of the accelerated frame, for as we have seen in Section 4, the potential of such a field is given by $g_{44}$ in a coordinate system adapted to the frame.

Thus Einstein's step 3 is multifaceted. The introduction of an arbitrary coordinate system makes the presence of a metric tensor in Minkowski space-time formally explicit as a matrix of components $g_{ik}$. At the same time Einstein uses the principle of equivalence to point out that this metric induces a gravitational field of a special type in the relative space of an accelerated frame of reference. This justifies interpreting the Minkowski metric as a particular instance of the four-dimensional generalization of such gravitational fields.

Interpreting the Minkowski metric in this way indicates that Einstein can arrive at a four-dimensional theory of arbitrary gravitational fields, which will also be his general theory of relativity, by generalizing the properties of the Minkowski metric in a manner analogous to the way that uniform gravitational fields can be generalized to nonuniform fields in Newtonian theory. He finds that the way to proceed is straightforward. The general theory will deal not only with Minkowski metrics, but also others of Lorentz signature.

This argument appears throughout Einstein's earlier work, but in a slightly less-developed form.[25] For it was only in his later years that he explicitly renounced the use of a separate stress-energy tensor as the source term in the field equations and used these equations only in their source-free form.

This source-free form of the field equations can be arrived at readily in the argument, as Einstein shows earlier, by merely contracting the flat space-time condition of special relativity. The argument appears commonly in this more complete form in his later writings.[26]

The earlier examples of the argument also contained an important addition to the example quoted earlier. Einstein would note that in the Galilean reference system of special relativity, a free point mass moves uniformly in a straight line. Such motion is represented in Minkowski space-time by a time-like geodesic, which satisfies the condition that the interval be extremal along the curve:

$$\delta \int ds = 0$$

It was natural to *assume*, the argument continued, that this requirement would also be satisfied by the world line of a free point-mass in the more general case of the general theory of relativity. I will return to the importance of this point in Section 9.

In short, we have seen in this section that the principle of equivalence enabled Einstein to see that one structure was responsible for inducing both inertial and gravitational fields and that the Minkowski metric was a special case of it. Einstein summarized this insight in a compact 1918 statement of the principle:

> *Principle of Equivalence*: inertia and gravity are *wesensgleich* [identical in essence]. From this and from the results of the special theory of relativity it necessarily follows that the symmetrical "fundamental tensor" ($g_{\mu v}$) determines the metrical properties of space, the inertial behavior of bodies in it, as well as gravitational action. (Einstein 1918a, p. 241)[27]

## 8. A Manner of Speaking

It was not uncommon for Einstein to associate the nonconstancy of the components of the metric tensor, or, equivalently, the nonvanishing of the Christoffel symbols in a given coordinate system with the presence of a gravitational field. In particular, he would describe the Christoffel symbols as the "gravitational field strengths" or "components of the gravitational field", for in a coordinate system in which these symbols vanished, free point-masses move "uniformly in a straight line". Therefore, these components "condition the deviation of the motion from uniformity" (Einstein 1916a, p. 802).

As in the last section, this association of the Christoffel symbols with gravitational field strengths can be explicated by recalling that Einstein often tacitly referred to frames of reference and their relative spaces when he talked explicitly only of a coordinate system adapted to them. If a coordinate system adapted to a uniformly accelerating frame of reference in Minkowski space-time is chosen so that its spatial coordinates are Cartesian, then the Christoffel symbols will contain only the spatial derivatives of the $g_{44}$. However, these derivatives together form a field strength, the three-vector gradient of the potential of the homogeneous gravitational field in the associated relative space.

The connection made here between the Christoffel symbols and the field strengths of the gravitational fields in relative spaces depends on a careful choice of space-time and coordinate system. Einstein, however, did not make this clear in his work and rarely qualified the identification of nonvanishing Christoffel symbol and gravitational field strength.

This practice has undoubtedly caused confusion. In a letter of January 1951, Laue challenged Einstein on this point.[28] He gave the example in Minkowski space-time of the transformation to curvilinear *spatial* coordinates from a Galilean coordinate system with no alteration in the time coordinate. Since this transformation is not associated with a change of state of motion, the resulting nonvanishing of "field strengths" is physically counterintuitive.

Einstein began his response by stressing that the Newtonian concept of gravitational field ("a1l the expressions obtained from the potential") is different from the concept of the relativistic gravitational field ("everything formed out of the symmetrical $g_{ik}$").[29] This corresponds to the distinction made here between the gravitational fields of relative spaces, which are usually represented by scalar fields, and their four-dimensional generalization, the metric field. Nevertheless, as he continued to explain, it was possible to forge a heuristic link between these two concepts and this link was the principle of equivalence:

> Heuristically, the interpretation of the field existing relative to a system. parallelly accelerated [*parallel beschleunigten*] against an inertial system (equivalence principle) was naturally of decisive importance, since this field is equivalent to a Newtonian gravitational field with parallel lines of force. In this case, the Newtonian field strengths are equal to the spatial derivatives of the $g_{44}$. Correspondingly, if one wants to, one can designate the first derivatives of the $g_{ik}$ or the displacement quantities $\Gamma$[affine connection] as gravitational field strengths, which certainly have no tensor character. In this manner of speaking, the introduction of cylindrical coordinates leads to the appearance of field strengths in a Galilean space. With this it is only a question of a manner of speaking.

Here Einstein uses the special case described earlier to justify speaking of the first derivatives of the $g_{ik}$ (which determine the Christoffel symbols and the affine connection in these space-times) as gravitational field strengths. One can continue to use this manner of speaking in other cases, but as Einstein's response indicates, it should be used with some caution.

This attitude to the description of the Christoffel symbols as gravitational field strengths was not a later development in Einstein's thought. It is also clearly evident in his 1916 reply to Kottler. There he says of this nomenclature, referring also to the nongenerally covariant stress-energy pseudotensor of the gravitational field, that "it is meaningless in principle and only intended to make concessions to our physical thought habits" but that it "appears to me, at least provisionally, not without value to maintain the continuity of thought" (Einstein 1916b, p. 641).

Today, some fifty years later, we insist that coordinate effects be carefully distinguished from physical effects. Examples such as Laue's show the confusion that would otherwise arise. Therefore, the provisional value of Einstein's manner of speaking is no longer evident. Einstein continued his response to Laue by stressing the important point beneath his manner of speaking, which involved no equivocation about coordinate effects:

> It is essential however, that a gravitational field exists in the sense of general relativity also in the case of a Galilei or a Minkowski space, even if the field strengths in the sense defined above vanish. In the theory of relativity, just the dimensionality of the field is the only thing that remains of the earlier physically independent (absolute) space.

In a given space-time, the nature, and even existence, of a gravitational field in a relative space will depend on the choice of frame of reference defining the relative space. But this relative-space dependence of these gravitational fields does not extend to their four-dimensional generalization, the space-time metric. All space-times of general relativity contain such a metric field—a gravitational field "in the sense of general relativity"—regardless of the frame of reference of relative space under consideration. This holds equally for Minkowski space-times, even though we can always find relative spaces in them that are gravitation-free in the older sense. In short, in general relativity a Minkowski space-time is not the gravitation-free special case.

## 9. The Infinitesimal Principle of Equivalence

Einstein's contemporaries of the early 1920s regarded the relative-space dependence of the gravitational field as the basic assertion of the principle of equivalence, rather than the occasion for inference to a more fundamental structure. Naturally, they were dissatisfied that Einstein dealt only with this relative-space dependence in the very simple case of the homogeneous gravitational fields of uniformly accelerated reference systems in Minkowski space-time. They sought an extended statement of this dependence that would apply directly to arbitrary gravitational fields (Pauli 1921, pp. 145–147; Silberstein 1922. pp. 10–13). They believed that this could be achieved in general relativity on the basis of the notion that special relativity holds in infinitesimally small regions of the space-time manifold, tacitly assuming that special relativity is a gravitation-free special case. As a result, their construct of the principle was very different from Einstein's and lays stress on the notion that a gravitational field can always be transformed away.[30] Pauli's classic formulation of the resulting principle reads:

> For every infinitely small world region (i,e., a world region which is so small that the space- and time-variation of gravity can be neglected in it) there always exists a coordinate system $K_0(X_1, X_2, X_3, X_4,)$ in which gravitation has no influence either on the motion of particles or any other physical processes (Pauli 1921, p. 145).[31]

Pauli continued to explain a little later that

> The special theory of relativity should be valid in $K_0$. All its theorems have thus to be retained, except that we have put the system $K_0$, defined for an infinitely small region, in place of the Galilean coordinate system.

In particular, this meant that the metric adopted the form diag(l, 1, 1, - 1) in $K_0$.

This "infinitesimal principle of equivalence" can be connected to Einstein's version at least superficially by noting that classical gravitational fields become homogeneous in infinitesimal regions of the relative space. Inverting Einstein's usual argument, they can then be transformed away at least infinitesimally by an appropriate acceleration of the reference system. One then regards the Pauli version of the principle as a four-dimensional restatement of these two results.

Of course this infinitesimal principle and the discussion of its connection to Einstein's version is beset with a number of serious technical difficulties. The notion of both three- and four-dimensional "infinitesimal regions" and the sense in which special relativity holds in such regions are unclear. Further, the actual statement of the principle makes it look as though it deals solely with a coordinate effect. These problems will be addressed shortly.

The popularity of the infinitesimal principle derives at least in part from its leading to a particularly attractive result; that it is possible to reconstruct much of the space-time manifold of general relativity as a patchwork of infinitesimal pieces in which special relativity holds.

Moritz Schlick, in his influential two-part article on space and time in the March 1917 issues of Die *Naturwissenschaften*, attempted just such a reconstruction (Schlick 1917). "We stipulated," he wrote, "that in an infinitely small region and in a reference system in which the bodies considered have no acceleration the special theory of relativity holds." It followed that in a "local" coordinate system, such as Pauli's $K_0$, the interval between two infinitesimally separated events is given by

$$ds^2 = (dX_1)^2 + (dX_2)^2 + (dX_3)^2 - (dX_4)^2.$$

Transforming to an arbitrary space-time coordinate system $\{x_i\}(1 = 1, 2, 3, 4)$, the expression for the interval became

$$ds^2 = g_{11}(dx_1)^2 + 2g_{12}dx_1dx_2 + \ldots + g_{44}(dx_4)^2,$$

where the symmetric coefficients $g_{ik}$(i,k = 1, 2, 3, 4) represent the components of the metric tensor in the new coordinate system. Schlick was thus able to infer that the new theory would involve a metric tensor and to arrive at many of its properties by considering the properties of the interval as given in special relativity.

In addition, Schlick considered the motion of a free material point. By reviewing its motion in the relative spaces of both local and accelerated coordinate systems and invoking the principle of equivalence, he concluded that the components of the metric tensor in the new coordinate system determine the gravitational field in the latter space. It also followed from special relativity that the world line of such a particle in the local coordinate system $(X_i)$ would be a geodesic. Since this was an invariant property, it would also be true of the world line in all coordinate systems, such as $(x_i)$. He then invoked the "principle of continuity" to justify the important conclusion that the world line of a free material point would be a geodesic in finite regions of the manifold as well.

Einstein has used arguments very similar to those just described. In particular, he used the assumption that special relativity holds in infinitesimal regions of the space-time manifold of general relativity in a manner close to that of Schlick, to introduce the metric tensor and some of its properties, especially those relating to the behavior of infinitesimal rods and clocks (Einstein 1916a, pp. 777–778; 1922, pp. 62–64).[32] However, this assumption was never related to the principle of equivalence, which was always formulated in Minkowski space-times. In addition, he was cautious in his use of this assumption, since he held that it was only true to a limited extent. This emerged in the correspondence between Einstein and Schlick following Schlick's article.

We know from this correspondence that Einstein had seen Schlick's article prior to its publication and that he approved of it wholeheartedly.[33] Six weeks after their initial exchange, however, Einstein wrote to Schlick to point out an error in one of the arguments sketched out here:

> The derivation of the law of motion of a point mass given on page 184 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. The correct derivation runs as follows: in principle there can exist finite (matter-free) parts of the world for which
>
> $$ds^2 = dX_1^2 + \cdots - dX_4^2$$
>
> 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 rel. 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 geodesic, with an arbitrary choice of coordinates.
>
> That the world line of a point is a geodesic in other cases too (if none other than gravitational forces act) is an hypothesis, even if a very obvious one.[34]

Einstein's objection bears directly on the assumption that special relativity does hold in an infinitesimal region of the space-time manifold of general relativity. He claims that it can only hold in a limited sense, for in such regions we cannot formulate the requirement that the world line of a free point-mass be a geodesic. (Note that Einstein called such lines "straight" in a Galilean reference system, since their spatial coordinates are linear functions of the time coordinate.)

Rather, as Einstein indicates here and as was his own practice elsewhere, when one discusses the motion of free point-masses, one must consider finite regions of the manifold in both special and general relativity. From the assumption that special relativity holds infinitesimally in general relativity, it does not follow that the world line of a free point-mass will be a geodesic in general relativity. Einstein's approach here and throughout his early work was to take this result in general relativity as strongly suggested by the corresponding result in special relativity, but in the last analysis still an independent assumption. (Of course, later he sought to derive this result in general relativity from the gravitational field equations.)

Finally, Einstein's comments here provide one more reason for his failure to retain an infinitesimal principle of equivalence after he briefly entertained one in 1912. As he came to realize, such a principle could not deal with the motion of bodies, the consideration of which formed the core of his principle. In the next section, I turn to examine whether Einstein's objection to Schlick holds. If it does, then he has pointed out a rarely acknowledged, but nevertheless devastating, difficulty for the traditional infinitesimal principle of equivalence.[35] If he is correct, then the restriction to infinitesimal regions makes it impossible to distinguish the geodesic world lines of free point-masses from other world lines and thus it is impossible to judge whether—in the words of Pauli's formulation—"gravitation has no influence on ... the motion of particles".

## 10. The Problem of Infinitesimal Regions

When Pauli and Schlick wrote of special relativity holding in infinitely small regions of the space-time manifold of general relativity, they could not have meant that special relativity holds in its usual sense. For whatever an infinitesimal or infinitely small region is, it must contain at least one point. Special relativity requires the vanishing of the Riemann-Christoffel curvature tensor. This requirement is well defined at every point of the manifold and is typically not satisfied in general relativity.

Rather they referred to a coordinate-dependent result, as is suggested by their qualification that special relativity hold in the region of an appropriately defined coordinate system. In a neighborhood of any given point $p$ in the space-time manifold in general relativity, it is possible to introduce a "local" coordinate system $K_0$ so that at $p$: the components of the metric $g_{ik}$ have the values diag(l, 1, 1, -1); the first (coordinate) derivatives of the components of the metric tensor $g_{ik,m}$ and thus also the Christoffel symbols vanish; but, in general, the second derivatives $g_{ik,mn}$ will not vanish.

When special relativity is said to hold in $K_0$ in an infinitesimal region around $p$, what is meant is the following. In $K_o$ at $p$, structures defined on the manifold, which do not deal with second and higher (coordinate) derivatives of the metric tensor, behave identically to their special relativistic counterparts at any point of a Minkowski space-time in a Galilean coordinate system. The criterion of identical behavior is equality of components of the quantities concerned. For example, in both cases the metric has components diag(l, 1, 1, -1), which means that the coordinate velocity of light will be unity. Both cases are commonly regarded as gravitation free insofar as the Christopher symbols, the "gravitational field strengths," vanish. And the world line of a free point-mass is a "straight" line, in the sense that it satisfies the condition $d^2X^i/ds^2 = 0$ at $p$, where $s$ is the interval. The two cases differ, however, when quantities containing $g_{ik,mn}$ are considered. Most notably the curvature tensor vanishes only in the case of Minkowski space-time.

The ignoring of second and higher derivatives of the metric tensor is usually justified by the introduction of a hierarchy of nested orders of quantities. Examples of first-order quantities contain the $g_{ik}$ alone; of second-order quantities, the $g_{ik}$ and $g_{ik,m}$, of third-order quantities, the $g_{ik}$, $g_{ik,m}$, and $g_{ik,mn}$; and so on. One must now imagine that the $g_{ik}$ are given at $p$ alone; the $g_{ik,m}$ are given by comparing the $g_{ik}$ at $p$ and at an infinitesimally close point; and the $g_{ik,mn}$ by comparing the $g_{ik}$ at two points infinitesimally close to $p$, the second more removed than the first. Then, finally, we imagine that access to quantities higher than any designated order can be denied by restricting consideration to sufficiently small infinitesimal regions around $p$.

It is now clear that the notion of these infinitesimal regions is problematic in differential geometry, since such regions cannot be equated with neighborhoods in their usual sense or any other structure commonly employed.

If we are to make a consistent evaluation of Einstein's objection to Schlick, the foregoing discussion must be made more precise. First, ambiguous restrictions concerning infinitesimal regions will be replaced by restrictions concerning orders of quantities. The assertion that special relativity holds infinitesimally in general relativity will be taken to mean only that special relativity holds at a point in the space-time manifold when quantities up to second order only are considered.

Second, we can eliminate the dependence on the coordinate system $K_0$ and on Galilean coordinate systems in Minkowski space-time by replacing the quantities $g_{ik}$, $g_{ik,m}$, and $g_{ik,mn}$ the examples of first-. second-, and third-order quantities mentioned earlier, by the covariant quantities $g_{ik}$, $D_i$, and $D_iD_k$, respectively.$D_i$ is the unique covariant derivative operator compatible with the metric $g_{ik}$. The coordinate-dependent notion of identity of quantities in the space-time manifold of general relativity with corresponding quantities in a Minkowski space-time is also naturally replaced by a requirement of diffeomorphic equivalence at the two corresponding points of each manifold.

Finally, we can extend the hierarchical ordering of quantities to those not constructed solely out of the metric and its derivatives by a technique based on one outlined by Geroch.[36] We generate subsets of the set of all diffeomorphisms $\{h\}$ whose domain is some neighborhood of $p$ and which map $p$ back onto itself. Let $g'$ be the image or $g$ under such a diffeomorphism and $D'_i$ the derivative operator constructed from $g$'. $\{h_1\}$ are all those diffeomorphisms for which $g' = g$ at $p$. $\{h_2\}$ are all those diffeomorphisms for which $D'_i = D_i$ at $p$. $\{h_3\}$ are all those for which $D'_iD'_k = D_iD_k$ and so on. We find [37]

$$\{h_1\} \supset \{h_2\} \supset \{h_3\} \supset \cdots$$

We can think of the members of $\{h_n\}$ as disturbing the manifold about $p$ in a way that will not affect the particular $n$th order quantity used at $p$ to define them. More figuratively, they leave undisturbed the infinitesimal region about $p$ needed to determine that quantity. Hence it is natural to use these sets of diffeomorphisms to define the hierarchy of orders of other quantities defined on the manifold. If $Q$ is a quantity defined at $p$, then the order of any quantity $F(Q)$ derived from it in the hierarchy of orders engendered by $Q$ is the smallest value of $n$ for which we always have $F(Q')$ $= F(Q)$, where $Q'$ is the image of $Q$ under any member of $\{h_n\}$. Let $c$ be a curve through $p$ differentiable to all orders with a tangent vector $X$. We can also classify the hierarchy of quantities generated by $c$ at $p$ by considering the images of $c$ under members of $\{h\}$. If an image curve $c'$ has the tangent vector $X'$, then we find that $X'$ is first order since $X' = X$ only under any member of $\{h_1\}$. Writing $D_x = X^iD_i$, we find $D_{X'}X' = D_xX$ only under the members of $\{h_2\}$. Hence $D_XX$ is a second-order quantity. Similarly $(D_X)^nX$ is of order $n + 1$ for all positive integers $n$.[38]

Now let the curve $c$ passing through $p$ be a geodesic parameterized by the interval $s$ and have tangent vector $X = d/ds$. By definition. at every point of $c$ in some neighborhood of $p$, $X$ will satisfy the condition

$$D_X X = 0.$$

It necessarily follows that at p

$$D_X D_X X = 0 \qquad D_X D_X D_X X = 0 \ldots (D_X)^n X = 0 \ldots$$

for all positive integers $n$ [39]

Einstein's objection that "in the infinitely small every continuous line is a straight line" can now be made more precise. If we restrict ourselves to quantities of first order, then at $p$ we can only characterize curves through $p$ by their tangent vectors, if defined. But if $c^*$ is *any* curve through $p$ with tangent vector $X^*$, then there will always lie a geodesic $c$ through $p$ with tangent vector $X$ equal to $X^*$. That is, as far as first-order quantities are concerned one cannot distinguish smooth curves from geodesics. If we read Einstein's "continuous line" as smooth curved then this first-order indistinguishability seems to express his point more precisely.

In the context of the infinitesimal principle of equivalence however, access to first- and second-order quantities is allowed. It follows that a geodesic $c$ with tangent vector $X$ will lie indistinguishable from any sufficiently smooth curve $c^*$ with tangent vector $X^*$, provided $X* = X$ and $D_{X*}X* = D_X X = 0$. Of course, the higher derivatives of $X^*$ along $c^*$ will not vanish in general. So $c^*$ need not be a geodesic. Since Einstein's objection was concerned in effect with this second-order case, it would have been better stated as "the world lines of any particles *unaccelerated* at $p$ (i.e.. $D_X = 0$) are indistinguishable from geodesics."

It is now also clear that any restriction on the order of quantities accessible at $p$ will make it impossible to distinguish geodesics from other curves. If quantities to order $n$ are allowed, then we cannot distinguish a geodesic $c$ from any other sufficiently smooth curve $c^*$ if they agree on quantities up to order $n$. Nevertheless, $c^*$ need not be a geodesic since any of the $(D_{X^*})^m X^*$ may fail to vanish for $m > n - 1$.

Another way to arrive at similar results is to consider $c'$, the image of $c$ under any member of $\{h_n\}$. By definition, $c'$ will be indistinguishable from $c$ to order $n$ at $p$. That is, they will agree on *any* quantity up to order $n$ that characterizes them. For example, $X' = X$, $D_{X'}X' = D_X X, \ldots (D_{X'})^{(n-1)}X' = (D_X)^{(n-1)}X = 0$. But as before, $c'$ will not be a geodesic in general since its derivatives of order greater than $n - 1$ need not vanish.

The results of this section vindicate Einstein's objection to Schlick. If we understand the infinitesimal principle of equivalence to assert that special relativity holds at a point in the space-time manifold of general relativity up to second-order quantities only, then it follows that we cannot formulate special relativity's requirement that the world line of a free point-mass be a geodesic.

In the terminology used by Pauli, Schlick, and Einstein, we would say that in the infinitesimal region concerned in the "local" coordinate system $K_0$, the fact that a world line satisfies the condition $d^2X^i/ds^2 = 0$ does not mean that it is a geodesic. This much is obvious once we realize that the restriction to infinitesimal regions effectively involves a restriction to the consideration of quantities at a single point in the manifold. However, we now also say that, under a consistent treatment of this restriction, the higher derivative terms, which might enable us to distinguish other curves satisfying this condition from geodesics, are not accessible from within these infinitesimal regions.

## 11. Real and Fictitious Gravitational Fields

The infinitesimal principle of equivalence tells us that the space-time manifolds of special and general relativity share the same first- and second-order structure at a point. For example, it tells us that metric $g$ and compatible derivative operator $D_i$ at a single point in each manifold are diffeomorphically equivalent. This result is not deep—it really only depends on the fact that both metrics have the same signature.

Presumably, this result is what Synge had in mind when he lamented in the introduction to his well-known text on general relativity that he never understood what I assume to be the infinitesimal principle of equivalence.

> Does it mean that the signature of the space-time metric is +2 (or -2 if you prefer the other convention)? If so, it is important, but hardly a Principle. Does it mean that the effects of a gravitational field are indistinguishable from the effects of an observer's acceleration? If so, it is false. In Einstein's theory. either there is a gravitational field or there is none according as the Riemann tensor does not or docs vanish. (Synge 1960, p. ix)

Synge's response to this difficulty is to insist that the effects of a true gravitational field are distinguishable from those of a fictitious field produced by the acceleration of the observers through an invariant criterion based on the Riemann-Christoffel curvature tensor.

It should now be clear that Einstein would not endorse this response to the difficulties of the infinitesimal principle of equivalence. For here Synge is proposing to resurrect precisely the distinction whose breakdown was crucial to Einstein's discovery of the general theory of relativity. Einstein explained his attitude to this question in correspondence with late, after Laue had pointed out that the Riemann-Christoffel curvature tensor vanishes in the context of the rotating disk problem:

> It is true that in that case the $R_{iklm}$ vanish, so that one could say: "There is no gravitational field present." However, what characterizes the existence of a gravitational field from the empirical standpoint is the non-vanishing of the $\Gamma^l_{ik}$ [coefficients of the affine connection], not the non-vanishing of the $R_{iklm}$. If one does not think intuitively in such a way, one cannot grasp why something like a curvature should have anything at all to do with gravitation. In any case, no reasonable person would have hit upon such a thing. The key for the understanding of the equality of inertial and gravitational mass is missing.[40]

Here Einstein reminds Laue that he had been able to recognize that the relativistic theory of gravitational fields should be a theory dealing with metrics of *nonvanishing* curvature, precisely because he was able to recognize that special relativity, the theory which dealt with a metric of *vanishing* curvature, was really also the theory of a special type of gravitational field. He could see this because, in turn, the Minkowski metric induced a structure identical to a classical gravitational field on the relative spaces of accelerating frames of reference and, unlike Synge, he had resisted the temptation of regarding this structure as somehow fictitious or different from "real" gravitational fields. (We have seen earlier how the $\Gamma^l_{ik}$ can appear as the field strengths of this structure in the relative spaces concerned.)

In the last analysis, over a half century after Einstein found and used this key, it matters little to one's application of the theory if one follows Synge and says that "the Riemann tensor ... *is* the gravitational field" (Synge 1960, p. viii) or if one follows Einstein and calls the metric tensor the gravitational field. For the connection between these structures and the gravitational fields of relative spaces which they generalize is essentially only a heuristic one. Perhaps Synge's approach is more comfortable for those who wish to continue thinking of special relativity as a gravitation-free case. For them, the presence of a gravitational field is the intrusion of some kind of perturbation into the Minkowski metric, in the same way as classical gravitational fields arise as anisotropies in otherwise constant scalar fields. If the curvature of a metric field is nonvanishing, then even a freely falling observer can detect this perturbation through the presence of tidal gravitational forces and he may well also be able to identify some nearby massive body that is largely responsible for it. [41]

Personally however, I find Einstein's attitude more comfortable and the association of gravitational fields only with metrics of nonvanishing curvature an arbitrary and unnecessary distinction. For such a distinction masks one of the most beautiful of Einstein's insights, that there is no essential difference between inertia and gravity. According to general relativity, the same structure—the metric—governs the motion of a body in free-fall in the "gravitation-free" case of special relativity or in free-fall in a classically recognizable gravitational field. If we are to call any structure "gravitational field" in relativity theory, then it should be the metric.

## 12. What was Einstein's Principle of Equivalence?

Einstein's principle of equivalence asserted that the properties of space that manifest themselves in inertial effects are really the properties of a field structure *in* space: moreover this same structure also governs gravitational effects. As a result, the privileged inertial states of motion defined by inertial effects are not properties of space but of this structure and the various possible dispositions of inertial motions in space are determined completely by it. Space of itself is to be expected to designate no states of motion as privileged.

This principle guided Einstein to seek his general theory of relativity as a gravitation theory of which special relativity was a special case. There the principle found precise theoretical expression. The structure responsible for inertial and gravitational effects is the metric tensor. The space-time manifold itself has no properties that would enable us to designate the motion associated with any given world line as privileged, that is as "inertial" or "unaccelerated." This designation depends entirely on the metric and the affine structure for space-time that it determines.

The purpose of the "Einstein elevator" thought experiment was to show that the structures associated with supposedly gravitation-free special relativity were already intimately connected with gravitation. To demonstrate this, he transformed from an inertial frame of reference to a uniformly accelerated frame and showed that a structure indistinguishable from a classical homogeneous gravitational field was induced by the Minkowski metric on the g associated relative space.

This property of the Minkowski metric enabled Einstein to identify it as an *instance* of the four-dimensional generalization of classical gravitational fields. This identification set Einstein on a royal road to his general theory of relativity. For it effectively reduced his task to that of finding a theory that generalized the properties of the Minkowski metric in a way enabling treatment of arbitrary gravitational fields.

Unfortunately, Einstein's contemporaries seized upon one of Einstein's intermediate results, that in certain cases the gravitational fields of relative spaces have a relative existence, dependent on the choice of frame of reference. They sought to generalize this result from the simple cases in Minkowski space-time that Einstein considered to arbitrary gravitational fields. It has rarely been acknowledged that Einstein never endorsed the principle that results, here called the "infinitesimal principle of equivalence". Moreover, his early correspondence contains a devastating objection to this principle: in infinitesimal regions of the space-time manifold it is impossible to distinguish geodesics from many other curves and therefore impossible to decide whether a point-mass is in free fall.

Some readers may feel dissatisfied that Einstein's principle of equivalence finds the uncontroversial expression indicated above in the general theory of relativity. On the contrary, I find it a source of great satisfaction and a testament to the coherence and clarity of Einstein's vision. For it shows that Einstein has been completely successful in taking an idea, which was quite extraordinary when conceived in 1907, and incorporating it completely into the body of a now universally accepted physical theory. In recent decades there has been much criticism of "the" principle of equivalence. But the principle under cogent attack has rarely been Einstein's version. For, to paraphrase Einstein's 1916 reflection on the critics of Mach, "even those who regard themselves as Einstein's opponents barely know how much of his views they have imbibed, so to speak, with their mother's milk (Einstein 1916c, p. 102).[42]

*Acknowledgments.* I am most grateful for the generous hospitality of the Einstein Project, Princeton, and the Center for the Philosophy of Science, University of Pittsburgh, and for the sponsorship of the Fulbright Exchange Program during the researching and writing of this paper. I wish to thank Michael Friedman, Al Janis, David Malament. John Stachel, and Roberto Torretti for helpful discussion and comments on earlier drafts. I also wish to thank the Hebrew University of Jerusalem for its kind permission to quote the material in this paper from Einstein's unpublished writings, to which it holds copyright.

## Notes

* Reprinted from *Studies in History and Philosophy of Science*, vol. 16. ©1985 by Pergamon Press. Ltd. with permission.

[1] For a compact discussion of some principles of equivalence, see Thorne, Lee, and Lightman 1973, pp. 3570–3572.

[2] This hypothesis is not labeled as the "principle of equivalence" in this article—the term does not appear anywhere in the article.

[3] For example, Einstein 1911, p. 899; 1954a, pp. 77-78.

[4] For example, Einstein to T. Levi-Civita, March 20, 1915, EA 16–233; to E. Klug. February 13, 1929, EA 25–126; to L.R. and H.G. Lieber, November 20. 1940, EA 15–135; to J. Reyntjens, August 26, 1950, EA 27–144; to A. Rehtz, July 12, 1953, EA 27–134.

[5] In all the places cited in this section, the only weak exception to this is in the letter to the Liebers where he allows that the gravitational field at a point is "in a certain way fictitious" because it can be transformed away.

[6] Earman and Glymour have also remarked on this (1978, p. 254).

[7] Specifically, six degrees of freedom in Newtonian space-times, three in Minkowski space-time and three or less (if any) in an arbitrary semi-Riemannian manifold. See Pauli 1921, pp, 130–132. So a "(rigid) uniformly accelerated frame of reference" in Minkowski space-time is specified by requiring the reference points to be in rigid motion and one of them to be uniformly accelerated. I shall always read "uniform (rectilinear) acceleration" in Minkowski space-time as referring to hyperbolic motion (Pauli 1921 , pp. 74–76).

[8] Torretti 1983, pp. 14–15, 28, defines a similar "relative spaces"

[9] Stachel (1980) has discussed Einstein's use of this example in detail.

[10] If $F$ is rotating, $H_c(p)$ will be orthogonal to $c$ only. So in general this mapping procedure must be repeated with a new orthogonal hypersurface for each $c$ in $R_F$. Most of the discussion of this section can be transferred to Newtonian space-times with little modification. Similar induced metrics could be defined in the relative spaces of Newtonian space-times by deriving them from the three-dimensional metrics of hypersurfaces of simultaneity.

[11] Pauli 1921, p. 131 writes this as the requirement of the constancy along $c$ of the components of $g_{\mathrm{orth}}$ in an adapted coordinate system. This condition is equivalent to the vanishing of the frame's expansion tensor, as defined in

Hawking and Ellis 1973, p. 82. Informally, the condition ensures constancy of the orthogonal interval between $c$ and an infinitesimally close curve $c'$ of $F$ in the hypersurfaces $H_c$.

[12] In Newtonian space-times, the scalar field $T$ is already given for all frames by the absolute time field. Therefore every relative space will have a frame time.

[13] Einstein 1907, pp. 414, 454. Then he wrote (p. 454): "This assumption extends the principle of relativity to the case of uniformly accelerated translational motion of the reference system". Einstein did not begin to describe his hypothesis with the compact labels ("equivalence principle" and "equivalence hypothesis" until 1912 and 1913.

[14] Laue to Einstein, December 27, 1911, EA 16–008.

[15] see Einstein 19l8b p. 700; 1950, p. 347; 1955b, p. 140.

[16] In his early (1911) version, Einstein notes that he will "disregard the theory of relativity" and confine himself to "customary" kinematics and "ordinary" mechanics.

[17] Einstein briefly rehearses the problem of characterizing such fields as Newtonian gravitational fields in 1920a.

[18] Einstein relayed his puzzlement at this result to Ehrenfest in a letter of June 1912, EA 9-333. See also Einstein 1912c.

[19] Einstein to A. Rehtz, July 12, 1953, EA 27-134. In his 1920b, Einstein summarizes the principle in similar terms: of . . . the physical properties of space prevailing relative to $K'$ are completely equivalent to a gravitational fielders $K'$ is a reference system in uniform rectilinear acceleration with respect to a Galilean system.

[20] Friedman 1983, pp. 191–195, has given a lucid analysis of the limited prospects of using a principle of equivalence to yield a generalized principle of relativity if the latter is understood to require this type of indistinguishability.

[21] $g'$ must be a Minkowski metric, since if $g$ has the form diag(-1, -1, -1, 1) in a coordinate system $\{y^i\}$, then $g'$ will have the same form in $\{y'^i\}$, the image of $\{y^i\}$ under $h$. Similarly the components of $g$ in $\{x^i\}$ at $p$ will equal the components of $g'$ in $\{x'^i\}$ at $hp$; therefore: (a) since the curves of constant $x^i (i = 1, 2, 3)$ are geodesics of $g$, the curves of constant $x'^i (i = 1, 2, 3)$ will be geodesics of $g'$; and (b) since the orthogonal metric of $g$ in the frame $F$ satisfies the rigid-body motion condition, the same will be true of the orthogonal metric of $g'$ in $F'$. From (a) and (b) it follows that $F'$ will be an inertial frame of $g'$.

[22] In his correspondence about his early work on the general theory, Einstein commented briefly that he saw the principle of equivalence incorporated into the new theory through its covariance properties; Einstein to P. Ehrenfest. Winter 1913–1914?, EA 9-347; Einstein to M. Besso, March 1914 (Speziali 1972, p. 53).

[23] These results also make plausible the failure of Einstein's first 1912 field equation to yield a conservation law, in spite of its similarity to the field equations of general relativity. From the perspective of general relativity, we would only expect his first field equation to yield consistent results in the

trivial case of Minkowski space-time.

[24] Einstein to Becquerel, August 16, 1951, EA 6-074 and 6-075, Einstein's argument is especially interesting and important, since it is intended to take a skeptic who accepts special but not general relativity step by step from the former to the latter, carefully delineating the assumptions of each step.

[25] See Einstein 1913a; pp. 285–286; 1913b, pp. 1255–1256; 1914a, p. 177; 1914b, pp. 1032-1033. See also Einstein 1954a, pp. 100-101, for a very clear exposition without formalism.

[26] See Einstein 1936, pp. 308–309; 1949, pp. 70–73; 1950, pp. 350–351., 1952, pp. 153–154., 1955a, pp. 14–15.

[27] Einstein used this same notion of identity of essence elsewhere in Einstein l912c, p. 1063; Einstein and Grossmann 1913, p. 226; and Einstein 1922, p. 58.

[28] Laue to Einstein, January 8, 1951, EA 16–152.

[29] Einstein to Laue. January 16, 1951, EA 16–154.

[30] Compare with Einstein's: "There is no space without gravitational or inertial field. What one calls empty space in the sense of classical or Maxwell's theory, is a gravitational field of a special kind, that is one in which the gravitational potentials are constant with an appropriate choice of coordinates." Einstein to H. Titze, January 16, 1954, EA 23–026/027.

[31] See also Silberstein 1922, p. 12.

[32] In a letter to P. Painlevé, December 7, 1921, EA 19–003, Einstein stresses that the general theory rests completely on the assumption that space-time behaves as it does in special relativity in infinitely small elements of the space-time manifold.

[33] Schlick to Einstein, February 4, 1917, EA 21–568; Einstein to Schlick. February 6, 1917, EA 21–612.

[34] Einstein to Schlick, March 21, 1917, EA 21–614. Schlick corrected the argument in accord with Einstein's remarks in the republication of the article in monograph form. See Schlick 1920, pp. 60-62.

[35] Torretti 1983, pp. 150-151, 316. has made the same objection in this context using virtually the same words as Einstein, but independently of him. Torretti writes: "In a Riemannian manifold, *every* curve is 'straight in the infinitesimal'." He illustrates his point vividly by pointing out that the streets which run along both parallels of latitude and meridians on the earth's surface are straight in the infinitesimal of such cities as Chicago, but only the meridians are geodesics.

[36] I am grateful to David Malament for making available to me mimeographed lecture notes of Robert Geroch, in which the technique is outlined.

[37] If members of $\{h\}$ map a point with coordinates $x^k$ to one with $y^i$, then at $p$ members of $\{h_1\}$ satisfy ${y^i}_{,k} = \delta^i_k$; members of $\{h_2\}$ satisfy the additional condition ${y^i}_{,km} = 0$; members of $\{h_3\}$ satisfy the additional condition ${y^i}_{,kmn} = 0$ and so on. Commas denote differentiation with respect to $x^k$.

[38] It is important to note that one can only consistently compare orders of quantities if their orders are assigned within a hierarchy generated by

the same structure. Any tensor will generate a hierarchy of quantities in which that tensor is of first order, since all tensors are invariant under the members of $\{h_1\}$. For example, the curvature tensor will be of first order in a hierarchy it generates, whereas it is of third order in the hierarchy generated by the metric tensor. In the text I tacitly assume that one can compare the orders of quantities in the metric tensor hierarchy with the orders of quantities in the hierarchy engendered by a geodesic through $p$. This is justified by the fact that these two hierarchies can be combined as follows. Each member of the set of geodesics $\{c\}$ through $p$ has a parameterization by the interval $s$ induced upon it by the metric tensor $g$. Conversely, given this same parameterization we can recover the original $g$, through the condition $g(X, X) = 1$ for all tangent vectors $X = d/ds$. Therefore, for the present purpose, we can consider $g$ and associated quantities as well as the set of tangent vectors $\{X\}$ and associated quantities as dependent on $\{c\}$ and its parameterization. In particular, the image of $\{c\}$ and its parameterization under a member of $\{h\}$ will generate a new metric tensor $g'$ and a new set of tangent vectors $\{X'\}$. We can now determine the orders of these and related quantities in the manner outlined earlier. The expected results do obtain. For example, both $g$ and $X$ are first order in this hierarchy.

[39] This argument establishes the necessity of these additional conditions. Their necessity can be illustrated in the example of a two-dimensional Euclidean space. In the usual Cartesian coordinate system, geodesics passing through the origin are $y = mx$, for $m$ a constant. However, the curves $y = x^n$ for all $n > 2$ satisfy the condition $D_X X = 0$ at the origin. The conditions $(D_X)^n X = 0$ for all positive integers $n$ are *not* sufficient. In the Euclidean space they are satisfied at the origin by the smooth curve $y = 0$ when $x = 0; y = \exp(-1/x^2)$ for all other $x$, but this curve is not a geodesic. (I am grateful to AI Janis for this last point.)

[40] Einstein to Laue, September 12, 1950, EA 16–148.

[41] Einstein and Rosen 1935 have added a curious twist to the standard objection that the gravitational fields produced by acceleration cannot be "true" gravitational fields since they have no sources. Recalling the principle of equivalence by name. they consider a coordinate system $\{x^i\}$ adapted to a uniformly accelerated frame of reference in Minkowski space-time and, in the now familiar manner, associate a homogeneous gravitational field with it. This accelerated frame cannot fill all of Minkowski space-time. In the case they consider, their frame fills the submanifold given by $(y_1)^2 \geq (y_4)^2$, where $\{y^i\}$ is the Galilean coordinate system used to define the frame (see their footnote, p. 74). They note that the Minkowski metric is a solution of the usual gravitational field equations of general relativity in the coordinate system $\{x^i\}$, but that certain components ($T_{22}$ and $T_{22}$) of the otherwise everywhere vanishing source stress-energy tensor become singular along the hypersurface $x_l = 0$, which is a boundary of the submanifold containing the accelerated frame. This represents a kind of source mass or energy distribution. They introduce the example so they can proceed to illustrate how such singularities can be removed. For further details see Einstein and Rosen

1935, p. 74.

[42] Of course, the original quotation is recovered by replacing "Einstein" by "Mach." This image may complement Synge's memorable image of the principle of equivalence as a midwife at the birth of general relativity who is now to suffer burial, but at least with appropriate honors. (Synge 1960, pp. ix–x).

## REFERENCES

Anderson, James L. ( 1967). *Principles of Relativity Physics.* New York: Academic Press.

Earman, John, and Glymour, Clark. (1978). "Lost in the Tensors: Einstein's struggles with Covariance Principles 1912-1916. " *Studies in History and Philosophy of Science* 9: 251–278.

Eddington, Arthur Stanley (1924). *The Mathematical Theory of Relativity,* 2nd ed. Cambridge: Cambridge University Press.

Einstein. Albert (1907). "Über das Relativitätsprinzip und die aus demselben gezogenen Folgerungen." *Jahrbuch der Radioactivität und Electronik* 4: 411–462.

(1911). "Über den Einfluss der Schwerkraft auf die Ausbreitung des Lichtes." *Annalen der Physik* 35: 898–908. Translated as: "On the Influence of Gravitation on the Propagation of Light" In Lorentz et al. 1923, pp. 97–108.

(1912a). "Lichtgeschwindigkeit und Statik des Gravitationsfeldes." *Annalen der Physik* 38: 355–369.

(1912b). "Zur Theorie des statischen Gravitationsfeldes." *Annalen der Physik* 38: 443–458.

(1912c). "Relativität und Gravitation. Erwiderung auf eine Bemerkung von M. Abraham." *Annalen der Physik* 38: 1059–1164.

(1912d). "Gibt es eine Gravitationswirkung die der elektrodynamischen Induktionswirkung analog ist?" *Vierteljahrsschrift für gerichtliche Medizin und öffentliches Sanitätswesen* 44: 37–40.

(1913a). "Physikalische Grundlagen einer Gravitationstheorie." *Naturforschende Gesellschaft Zürich. Vierteljahrschrift* 58: 284–290.

(1913b). "Zum gegenwärtigen Stande des Gravitationsproblems." *Physikalische Zeitschrift* 14: 1249–1266.

(1914a). "Prinzipielles zur verallgemeinerten Relativitätstheorie und Gravitationstheorie." *Physikalische Zeitschrift* 15: 176–180.

(1914b). "Die formale Grundlage der allgemeinen Relativitätstheorie." *Königlich Preussische Akademie der Wissenschaften* (Berlin). Sitzungsberichte: 1030–1085.

(1916a). "Die Grundlage der allgemeinen Relativätstheorie." *Annalen der Physik* 49: 769–822. Translated as "The Foundation of the General Theory of relativity" In Lorentz et al. 1923, pp. 109–164.

(1916b). "über Friedrich Kottlers Abhandlung 'Über Einstein's Äquivalenzhypothese und die Gravitation.' " *Annalen der Physik* 51: 639–642.

(l9l6c). "Ernst Mach." *Physikalische Zeitschrift* 17: 101–104.

(1917). "Kosmologische Betrachtungen zur allgemeinen Relativitätstheorie." *Königlich Preussische Akademie der Wissenschaften* (Berlin). *Sitzungsberichte*: 142–152. Translated as "Cosmological Considerations on the General Theory of relativity" In Lorentz et al. 1923, pp. 175–188.

(1918a). "Prinzipielles zur allgemeinen Relativitätstheorie." *Annalen der Physik* 55: 241–244.

(1918b) "Dialog über Einwände gegen die Relativitätstheorie." *Die Naturwissenschaften* 6: 697–702.

(1920a). "Inwiefern lässt sich die moderne Gravitationstheorie ohne die Relativität begrunden?" *Die Naturwissenschaften* 8: 1010–1011.

(1920b). "Grundgedanken und Methoden der Relativitätstheorie in ihrer Entwicklung dargestellt." Unpublished manuscript EA 2–070. This document can be dated to 1920, in part by Einstein's mention of the "English (eclipse) expedition of the previous year" on p. 32.

(1922). *The Meaning of Relativity.* Princeton: Princeton Univesity Press. Page numbers are cited from the 5th ed. Princeton: Princeton University Press, 1955.

(1936). "Physics and Reality." Journal of the Franklin institute 221: 349–382. Page numbers are cited from the reprint in Einstein 1954b, pp. 290–323.

(1949). "Autobiographical Notes." In *Albert Einstein: Philosopher-scientist.* Paul Arthur Schilpp, ed. Evanston, Illinois: The Library of Living Philosophers, pp. 1–95.

(1950). "On the Generalized Theory of gravitation": *Scientific American* 182: 13–17. Page numbers are cited from the reprint in Einstein 1954b, pp. 341–356.

(1952). "Relativity and the Problem of space." Appendix 5 in Einstein 1954a, pp. 135–157.

(1954a). *Relativity, The Special and the General Theory: A Popular Exposition,* 15th ed. Robert W. Lawson, trans. London: Methuen, 1954.

(1954b). *Ideas and Opinions.* Carl Seelig, ed. Sonja Bargmann, trans. New York: Crown; reprint London: Souvenir Press, 1973.

(1955a). "Erinnerungen—Souvenirs." *Schweizerische Hochschulzeitung* 28 (*Sonderheft*): 145–148, 151–153. Reprinted as "Autobiographische Skizze." In *Helle Zeit-Dunkle Zeit.* Carl Seelig, ed. Zurich: Europa Verlag, 1956, pp. 9–l7.

(l955b). "Relativistic Theory of the Non-symmetric field." Appendix 2 in *The Meaning of Relativity*, 5th ed. Princeton: Princeton University Press, pp. 133–166.

Einstein, Albert, and Grossmann, Marcel (1913). *Entwurf einer verallgemeinerten Relativitätstheorie und einer Theorie der Gravitation. I. Physikalischer Teil von Albert Einstein. II. Mathematischer Teil von Marcel Grossmann.* Leipzig and Berlin: B.G, Teubner.

Einstein, Albert, and Rosen, Nathan (1935). "The Particlc Problem in the General Theory of relativity." *Physical Review* 48: 73–77.

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

Hawking, Stephen W., and Ellis, George F.R. (1973). *The Large-scale Structure of Space-Time.* Cambridge: Cambridge University Press.

Lorentz. Hendrik Antoon, et al. (1923). *The Principle of relativity; A Collection of Original Memoirs on the Special and General Theory of relativity.* W. Perrett and G.B. Jeffery, trans. London: Methuen; reprint New York: Dover, 1952.

Newton, Isaac (l729). *Sir Isaac Newton's Mathematical Principles of Natural Philosopby & His System of the World.* Andrew Motte, trans. London. Reprint New York: Greenwood, 1969.

Norton, John (1984). "How Einstein Found His Field Equations: 1912–1915." *Historical Studies in the Physical Sciences* 14: 253–316. See this volume, pp. 101–159.

Ohanian, Hans C. (1977). "What is the Principle of Equivalence?" *American Journal of Physics* 45: 903–909.

Pais, Abraham (1982). *"Subtle is the Lord ...": The Science and the life of Albert Einstein.* Oxford: Clarendon Press.

Pauli, Wolfgang (1921). "Relativitätstheorie." In *Encyclopädie der mathematischen Wissenschaften, mit Einschluss ihrer Anwendungen.* Vol. 5, *Physik,* part 2. Arnold Sommerfeld, ed. Leipzig: B.G. Teubner, 1904–1922, pp. 539–775. (Issued November 15, 1921.] Page numbers cited from the English translation *Theory of Relativity.* With supplementary notes by the author. G. Field, trans. London: Pergamon, 1958; reprint New York: Dover, 1981.

Schlick, Moritz (1917). "Raum und Zeit in der gegenw'"artigen Physik. Zur Einführung in das Verständnis der allgemeinen Relativitätstheorie." *Die Naturwissenschahen* 5: 161–167 (March 16), 177–186 (March 23).

(1920). *Raum und Zeit in der gegenwärtigen Physik. Zur Einfürung in das Verständnis der Relativitäts- und Gravitationstheorie,* 3rd ed. Berlin: Julius Springer, 1920. English translation *Space and Time in contemporary physics: an Introduction to the Theory of Relativity and Gravitation.* Henry Brose. trans. New York: Oxford University Press, 1920.

Silberstein, Ludwik (1922). *The Theory of General Relativity and Gravitation.* Toronto: University of Toronto Press.

Speziali, Pierre, ed. (1972). *Albert Einstein-Michele Besso. Correspondence, 1903–1955.* Paris: Hermann.

Stachel, John (1980). "Einstein and the Rigidly Rotating disk." In *General Relativity and gravitation: A Hundred years after the Birth of Einstein.* A. Held, ed. New York: Plenum, pp. 1–15. See this volume, pp, 48–62.

Synge, John Lighton (1960). *Relativity: The General Theory.* Amsterdam: North-Holland.

Thorne, Kip S., Lee, David L., and Lightman, Alan P. (1973). "Foundations for a Theory of Gravitation theories" *Physical Review D* 7: 3563–3578.

Torretti, Roberto (1983). *Relativity and Geometry.* Oxford: Pergamon

Press.