Rep. Prog. Phys. 56 (1993) 791–858. Printed in the UK

General covariance and the foundations of general relativity: eight decades of dispute

John D Norton

Department of History and Philosophy of Science, University of Pittsburgh, Pittsburgh, PA 15260, USA

Abstract

Einstein offered the principle of general covariance as the fundamental physical principle of his general theory of relativity and as responsible for extending the principle of relativity to accelerated motion. This view was disputed almost immediately with the counter-claim that the principle was no relativity principle and was physically vacuous. The disagreement persists today. This article reviews the development of Einstein's thought on general covariance, its relation to the foundations of general relativity and the evolution of the continuing debate over his viewpoint.

This review was received in March 1993.

0034-4885/93/070791+68$30.00

Contents

1. Introduction

In November 1915, Einstein completed his general theory of relativity. Almost eight decades later, we universally acclaim his discovery as one of the most sublime acts of human speculative thought. However, the question of precisely what Einstein discovered remains unanswered, for we have no consensus over the exact nature of the theory's foundations. Is this the theory that extends the relativity of motion from inertial motion to accelerated motion, as Einstein contended? Or is it just a theory that treats gravitation geometrically in the spacetime setting? When Einstein completed his theory, his own account of the foundations of the theory was adopted nearly universally. However, among the voices welcoming the new theory were small murmurs of dissent. Over the brief moments of history that followed, these murmurs grew until they are now some of the loudest voices of the continuing debate.

In any logical system, we have great freedom to exchange theorem and axiom without altering the system's content. Thus we need no longer formulate Euclidean geometry with exactly the definitions and postulates of Euclid or use precisely Newton's three laws of motion as the foundations of classical mechanics. However, some two millennia after Euclid and three centuries after Newton, we still find their postulates and laws within our systems, now possibly as theorems and sometimes even in a wording remarkably close to the original.

The continuing disagreement over the foundations of Einstein's theory extends well beyond such an orderly expansion of our understanding of a theory's foundations. It is far more than a squabble over the most perspicacious way to reorganize postulate and theorem or to clarify brief moments of vagueness. The voices of dissent proclaim that Einstein was mistaken over the fundamental ideas of his own theory and that the basic principles Einstein proposed are simply incompatible with his theory. Many newer texts make no mention of the principles Einstein listed as fundamental to his theory; they appear as neither axiom nor theorem. At best, they are recalled as ideas of purely historical importance in the theory's formation. The very name 'general relativity' is now routinely condemned as a misnomer and its use often zealously avoided in favour of, say, 'Einstein's theory of gravitation.'

What has complicated an easy resolution of the debate are the alterations of Einstein's own position on the foundations of his theory. At different times of his life, he sought these foundations in three principles and with varying emphasis. They were the principle of equivalence, Mach's principle and the principle of relativity. By his own admission (Einstein 1918), he did not always distinguish clearly between the last two. Again, he lost completely his enthusiasm for Mach's principle, abandoning it unequivocally in his later life.

The reception and development of Einstein's account in the literature has been anything but a graceful evolution. It has been more a process of uncontrolled mutation, fragmentation and even disintegration. The principle of equivalence took root in so many variant forms that Anderson and Gautreau (1969, p 1656) eventually lamented that there are 'almost as many formulations of the principle as there are authors writing about it.' This dissipation is at least partially fuelled by skeptical attacks on the principle such as Synge's (1960, p ix) famous complaint that he has never been able to find a version of the principle that is not false or trivial.

The locus of greatest controversy has been at the core of Einstein's interpretation, the principle of relativity. Does the general theory extend the principle of relativity to accelerated motion and is this extension captured by the general covariance of its laws? It is routinely allowed that the special theory of relativity satisfies the principle of relativity of inertial motion simply because it is Lorentz covariant: its laws remain unchanged in form under a Lorentz transformation of the space and time coordinates. Now Einstein's general theory is generally covariant: its laws remain unchanged under an arbitrary transformation of the spacetime coordinates. Does this formal property allow the theory to extend the relativity of motion to accelerated motion? Until recent decades, the majority of expositions of general relativity answered yes and some still do.

As early as 1917, Kretschmann (1917) argued that general covariance has no real physical content and no connection to an extension of the principle of relativity. Rather, the finding a generally covariant formulation of a theory amounts essentially to a challenge to the mathematical ingenuity of the theorist. Skeptical sentiments such as these drove a dissident tradition that has grown from a minority in Kretschmann's time to one of the dominant traditions at present. It has derived further support from the development of more sophisticated mathematical techniques that are now routinely used to give generally covariant formulations of essentially all commonly discussed spacetime theories, including special relativity and Newtonian spacetime theory.

Finally, to many, Einstein's statements of his views seemed too simple or abbreviated to stand without further elaboration or repair; whereas their flat rejection by the skeptics seemed too easy. Thus much energy has been devoted to finding ways in which the general covariance of Einstein's theory can be seen to be distinctive even in comparison with the generally covariant formulations of special relativity and Newtonian spacetime theory. The best developed of these attempts is due to Anderson (1967) and is based on the distinction of absolute from dynamical objects. General relativity satisfies Anderson's 'principle of general invariance' which entails that the theory can employ no non-trivial absolute objects. This principle is offered as a clearer statement of Einstein's real intentions and as giving a precise interpretation of Einstein's repeated disavowal of the absolutes of Newton's space and time.

The purpose of this article is to review the development of Einstein's views on general covariance, their relation to the foundations of general relativity and the evolution of the continuing debate that sprang up around these views. Sections 2 and 3 will review the development of Einstein's views. Section 4 will outline the ways in which attempts were made to receive and assimilate Einstein's views in a favourable manner. Section 5 will review Kretschmann's famous objection, Einstein's response and the diverse ways in which both were received in the literature. It includes discussion of modern geometrical methods that ensure automatic general covariance. Section 6 reviews the development of the characterization of a relativity principle as a symmetry principle rather than a covariance principle. Section 7 explores the tradition of exposition of general relativity that simply ignores the entire debate and makes no mention of principles of general relativity or of general covariance. Section 8 develops Anderson's theory of absolute and dynamical objects as it relates to Einstein's views. Section 9 examines Fock's and Arzeliès proposals for alterations to the covariance of general relativity and gives an historical explanation of why so many of Einstein's pronouncements on coordinates and covariance are puzzling to modern readers.

In the time period covered in this review article, the mathematical methods used in relativity theory evolved from a coordinate based calculus of tensors to a coordinate free, geometric approach. The mathematical language and sensibilities used in various stages of the article will match those of the particular subject under review. The alternative of translating everything into a single language would harmfully distort the subject (see section 9.2).

2. The background of special relativity

2.1. Lorentz covariance and the relativity of inertial motion

Einstein's (1905) celebrated paper on special relativity brought the notion of the covariance of a theory to prominence in physics and introduced a theme that would come to dominate Einstein's work in relativity theory. The project of the paper was to restore the principle of relativity of inertial motion to electrodynamics. In its then current state, the theory distinguished a preferred frame of rest, although that frame had eluded all experiment and even failed to appear in the observational consequences of electrodynamics itself. Einstein's renowned solution was not to modify electrodynamics, but the background space and time itself. He devised a theory in which inertial frames of reference were related by the Lorentz transformation. If an inertial frame has Cartesian spatial coordinates (x, y, z) and time t and a second frame moving at velocity v in the x direction has spatial coordinates (ξ, η, ζ) and time coordinate τ, then, under the Lorentz transformation,

$$\xi = \gamma(x - vt) \qquad \tau = \gamma(t - vx/c^2) \qquad \eta = y \qquad \zeta = z \tag{1}$$

where $\gamma = (1 - v^2/c^2)^{-1/2}$ and c is the speed of light. Hitherto classical theory had in effect employed what was shortly called (by, for example, Laue (1911, p3)) the Galilei-transformation

$$\xi = x - vt \qquad \tau = t \qquad \eta = y \qquad \zeta = z.$$

Selecting suitable transformation laws for the field and other quantities, Einstein was able to show that the laws of electrodynamics remained unchanged under the Lorentz transformation. That is, they were Lorentz covariant. Therefore, within the space and time of special relativity, electrodynamics could no longer pick out any inertial frame of reference as preferred. Each inertial frame was fully equivalent within the laws of the theory. Anything said about one by the laws of electrodynamics must also be said of all the rest. Electrodynamics was now compatible with the relativity of inertial motion.

With the example of electrodynamics as its paradigm, the task of constructing a special relativistic version of a physical theory reduced essentially to formulating its laws in such a way that they remained unchanged under Lorentz transformation. Thus Einstein's (1905, section 10) original paper proceeded to formulate a modified mechanics for slowly accelerated electrons with this property. Thermodynamics soon also received some of its earliest relativistic reformulations in the same manner (see Einstein 1907, part IV, for example).

The lesson of Einstein's 1905 paper was simple and clear. To construct a physical theory that satisfied the principle of relativity of inertial motion, it was sufficient to ensure that it had a particular formal property: its laws must be Lorentz covariant. Lorentz covariance became synonymous with satisfaction of the principle of relativity of inertial motion and the whole theory itself, as Einstein (1940, p329) later declared:

> The content of the restricted relativity theory can accordingly be summarized in one sentence: all natural laws must be so conditioned that they are covariant with respect to Lorentz transformations.

2.2. Minkowski's introduction of geometrical methods

In Einstein's hands, Lorentz covariance was a purely algebraic property. Space and time coordinates were, in effect, variables that transformed according to certain formulae.

Hermann Minkowski (1908, 1909) was responsible for introducing geometric methods and thinking into relativity theory. He explained the background to his approach in his more popular (1909) lecture. It amounted to an inspired but essentially straightforward application of then current ideas in geometry. Minkowski's colleague at Göttingen, Felix Klein, had brought a fertile order to the world of 19th century geometry. That world was beginning to fragment after the discovery that geometry did not have to be Euclidean. In his famous *Erlangen* program, Klein (1872) proposed categorizing the new geometries by their characteristic groups of transformations. Euclidean geometry, for example, was characterized by the group of rotations, translations and reflections. The entities of the geometry were the invariants of these transformations.

Minkowski pointed out that geometers had concentrated on the characteristic transformations of space. But they had ignored the groups of transformations associated with mechanics, those that connected various inertial states of motion. Minkowski proceeded to treat these groups in exactly the same way as the geometric groups. In particular he constructed the geometry associated with the Lorentz transformation. To begin, it was not the geometry of a *space*, but of a *spacetime*, and the notion of spacetime was introduced into physics almost as a perfunctory by-product of the *Erlangen* program. Moreover he found the spacetime had the hyperbolic structure now associated with a Minkowski spacetime.

From this geometric perspective, the formulation of a theory that satisfied the principle of relativity became trivial. One merely needed to formulate the theory in terms of the geometric entities of the spacetime—in effect the various types of spacetime vectors Minkowski had defined—and the theory would be automatically Lorentz covariant. Thus Minkowski (1908, appendix; 1909, section V) could write down a gravitation theory without even needing to consider whether it was compatible with the principle of relativity, for the theory was constructed purely geometrically. Thus, in his exposition of four-dimensional vector algebra and analysis, Sommerfeld (1910, p 749) could state:

> According to Minkowski, as is well known, one can formulate the content of the principle of relativity as: only *spacetime vectors* may appear in physical equations . . .

2.3. *Covariance versus invariance in special relativity*

The difference between Einstein and Minkowski's approach to the same theory and even the same formalism is a polarity that will persist in various manifestations throughout the hole development of relativity theory, both special and general. Einstein's emphasis is on the algebraic properties of the theory, the equations that express its laws and their behaviour under transformation, its *covariance*. Thus the satisfaction of the principle of relativity is established by often arduous algebraic manipulation. The equations of the theory are transformed under the Lorentz transformation and the resulting equations are shown to have preserved their form. Minkowski's emphasis is on the geometric properties of the theory, on those geometric entities which remain unchanged behind the transformations, its *invariance*. Thus Minkowski ensures satisfaction of the principle of relativity by quite different means. The only structures allowed in constructing a theory are spacetime invariants. This restriction ensures compatibility with the principle of relativity and that its satisfaction can be settled by inspection.

3. Einstein's development of general relativity

While it may have been some years in preparation, the special theory of relativity coalesced into its final form quite suddenly so that Einstein's first paper on the theory remains one of its classic expositions. The development of general relativity was far slower and more tangled. Eight years elapsed between the inception and completion of the theory, during which time Einstein published repeated reports on the intermediate phases, false turns and unproven expectations. Even after the completion of the theory Einstein's account of its foundations continued to evolve. The modern image of Einstein's view of the foundations of general relativity is drawn fairly haphazardly from pronouncements that were made at differing times in this evolution. As a result, they are not always compatible. Indeed the pronouncements were sometimes as much expressions of results anticipated as demonstrated. For this reason, it would be misleading to construct any single edifice and proclaim it. Einstein's account of the foundations of general relativity. Rather we shall have to trace the evolution of Einstein's views as they were elaborated and modified in pace with the development of his theory.

In developing general relativity, Einstein sought to satisfy many requirements. However we shall see that his efforts were dominated by a single theme, covariance, and they reduced essentially to an enduring task, expanding the covariance of relativity theory beyond Lorentz covariance.

3.1. The early years 1907–1912: principle of equivalence and the relativity of inertia

Two years after his completion of the special theory, Einstein began developing ideas that would ultimately lead him to the general theory of relativity. In a final speculative section of a 1907 review article on relativity theory, he raised the question of whether the principle of relativity could be extended to accelerated motion (Einstein 1907, part V). The question was immediately understood as asking whether he could expand the covariance group of relativity theory. Feeling unable to tackle the general question, Einstein considered the simple case of a transformation from an inertial reference frame of special relativity to a reference frame in uniform rectilinear acceleration. In the accelerated frame of reference a homogeneous inertial field arises. Because of the key empirical fact of the equality of inertial and gravitational mass, Einstein was able to identify this field as a gravitational field. He then made the postulate that would dominate the early years of his work on gravitation. In the wording of Einstein (1911, section 1)

> . . . we assume that the systems K [inertial system in a homogeneous gravitational field] and K' [uniformly accelerated system in gravitation free space] are physically exactly equivalent, that is, . . . we assume that we may just as well regard the system K as being in a space free from gravitational fields, if we then regard K as uniformly accelerated.

This assumption soon acquired the name 'hypothesis of equivalence' (Einstein 1912a, p 355) and then 'principle of equivalence' (Einstein 1912b, p 443). Through it, Einstein generated a novel theory of static gravitational fields (Einstein 1907, part V, 1911, 1912a, b). In it, the now variable speed of light played the role of the gravitational potential; light from a heavy body such as the sun would be red shifted; and light grazing a heavy body such as the sun would be deflected.

For our purposes, the important point is that Einstein saw in the principle an extension of the principle of relativity. Continuing the above passage, he observed

> This assumption of exact physical equivalence makes it impossible for us to speak of the absolute acceleration of the system of reference, just as the usual theory of relativity forbids us to talk of the absolute velocity of a system . . . .

The principle of equivalence formed just one part of Einstein's assault on the problem of extending the principle of relativity. He had also to answer the more general worry that acceleration seemed distinguishable from inertial motion by observable consequences, whereas no such consequences enable us to distinguish inertial motion from rest. Newton had driven home the point in the Scholium to the Definitions of Book 1 of his *Principia* (1687). He noted that the absolute of rotation of water in a bucket was revealed by the observable curvature of the water's surface. The inertia of the water was responsible for this effect, leading it to recede from the axis of rotation.

Einstein found his answer to Newton in his reading of Ernst Mach. Mach (1893, p 284) pointed out that all that was revealed in Newton's bucket thought experiment was a correlation between the curvature of the water and its rotation with respect to the earth and other celestial bodies. Thus Einstein (1912c) was delighted to report his 1912 theory entailed certain weak field effects that promised to convert this correlation into a physical interaction, with the rotation of the stars with respect to the water directly causing the curvature of its surface. He found that the inertia of a test mass is increased if it is surrounded by a shell of inertial masses and that, if these same masses are accelerated, they tend to drag the test mass with it. These results raised the possibility of an idea which he attributed (p 39) directly to Mach:

> . . . the entire inertia of a point mass is an interaction with the presence of all the remaining masses and based on a kind of interaction with them.

Einstein (1913, p 1261) soon called this idea the 'hypothesis of the relativity of inertia.'

Clearly if a theory could be found that implemented this hypothesis, Einstein would have succeeded in generalizing the principle of relativity to acceleration. For, in such a theory, the preferred set of inertial frames would cease to be an absolute feature of the background space and time; the disposition of inertial frames of reference would merely be an accident of the overall distribution of matter in the universe. However, by the middle of 1912, Einstein was still far from such a theory. In concluding his response to a polemical assault by Max Abraham, Einstein (1912d, pp 1063–4) described his project in terms of the expansion of the covariance of the current theory of relativity and his hope that 'the equations of theory of relativity that also embraced gravitation would be invariant with respect to acceleration (and rotation) transformations.' However he confessed that 'it still cannot be foreseen what form the general spacetime transformations equations could have.' The Einstein who wrote these words in July 1912 had not yet foreseen that his name would be irrevocably associated with a generally covariant theory.

3.2. The 'Entwurf' theory 1912–1915: general covariance gained and lost

All this changed with Einstein's move to Zurich in August 1912. There he began collaborating with the mathematician Marcel Grossmann, a good friend from his student days. Grossmann discovered for Einstein the existence of the 'absolute differential calculus'† of Ricci and Levi-Civita (1901) and pointed out that this calculus would enable Einstein to construct a generally covariant theory.

The focus of this calculus was the fundamental quadratic differential form

$$\varphi = \sum_{r,s=1}^{n} a_{rs}\, dx_r\, dx_s \tag{2}$$

which was assumed to remain invariant under arbitrary transformations of the variables $x_1, \ldots, x_n$. Of course the modern reader immediately associates this form with the invariant

† The Ricci–Levi-Civita calculus only later acquired its modern name of 'tensor calculus' after Einstein and Grossmann (1913) renamed all of Ricci and Levi-Civita's 'contravariant and covariant systems' as 'tensors' thereby extending the formerly rather restricted compass of the term 'tensor.' See Norton (1992, appendix).

line element of a non-Euclidean surface of variable curvature, such as was introduced by Gauss and developed by Riemann. However Ricci and Levi-Civita's $x_1, \ldots, x_n$ were *variables* and not necessarily geometric coordinates. They were at pains to emphasize that what was then called infinitesimal geometry was just one of many possible applications of their calculus.

As late as 1912, Einstein had not adopted the four-dimensional methods of Minkowski, even though these methods had already found their first text book exposition (Laue 1911). Einstein's 1912 static gravitational theory had been developed using essentially the same mathematical techniques as his 1905 special relativity paper. Thus it is an odd quirk of history that, when Einstein did finally immerse himself in the four-dimensional spacetime approach, he turned to exploit a calculus whose creators sought to skirt its geometric interpretation in favour of a broader interpretation.

Einstein and Grossmann published the results of their joint research early the following year with Einstein writing the 'Physical Part' and Grossmann the 'Mathematical Part.' The theory of the resulting paper (Einstein and Grossmann 1913) is commonly known as the '*Entwurf*' theory from the title of the paper, '*Entwurf einer verallgemeinerten Relativitätstheorie und einer Theorie der Gravitation*' ('outline of a generalized theory of relativity and a theory of gravitation'). Its central idea involved the introduction of Ricci and Levi-Civita's fundamental form (2). They started with the invariant interval of Minkowski in differential form

$$ds^2 = c^2\,dt^2 - dx^2 - dy^2 - dz^2 \tag{3}$$

where (x, y, z, t) are the space and time coordinates of an inertial frame of reference in a Minkowski spacetime. Transforming to arbitrary coordinates x_μ, for $\mu = 1, \ldots, 4$, (3) becomes†

$$ds^2 = g_{\mu\nu}\,dx_\mu\,dx_\nu. \tag{4}$$

Einstein employed his principle of equivalence to interpret the matrix of quantities $g_{\mu\nu}$ that had arisen with the introduction of arbitrary coordinates. In the special case of the principle, the transformation from (3) to (4) is from an inertial coordinate system to a uniformly accelerated coordinate system. In that case, the matrix of coefficients $g_{\mu\nu}$ reduces to that of (3), except that c now is a function of the coordinates (x', y', z'). That is, (4) becomes

$$ds^2 = c^2(x', y', z')dt'^2 - dx'^2 - dy'^2 - dz'^2. \tag{3'}$$

According to the principle of equivalence, the presence of a gravitational field was the only difference between the spacetime of (3′) and that of special relativity (3). Therefore Einstein interpreted the coordinate dependent c of (3′) as representing a gravitational field and, more generally, the $g_{\mu\nu}$ of (4) as representing a gravitational field.

Einstein and Grossmann proceeded to develop essentially all the major components of the final general theory of relativity. Just one eluded them. The spacetimes represented by (3), (3′) and (4) are all flat. To treat the general case of the gravitational field, non-flat metrics must also be admitted and, in the final theory, the decision of which are admitted is

† Henceforth summation over repeated indices is implied. Einstein himself did not introduce this summation convention until 1916.

made by the gravitational field equations. Einstein expected these equations to take the now familiar form

$$G_{\mu\nu} = \kappa T_{\mu\nu} \tag{5}$$

where $T_{\mu\nu}$ is the stress-energy tensor and $G_{\mu\nu}$ a gravitation tensor constructed solely from the metric tensor $g_{\mu\nu}$ and its derivatives. Einstein and Grossmann considered the Ricci tensor as their gravitation tensor—just a hair's breadth away from Einstein's final choice of the Einstein tensor. However they reported that the resulting field equations failed to give the Newtonian limit in the case of weak, static gravitational fields. In their place, to the astonishment of modern readers, they offered a set of gravitational field equations that was not generally covariant. Einstein then descended into a long struggle with his imperfect theory that lasted almost three intense years before he emerged victoriously with the final generally covariant theory in hand†.

3.3. The hole argument: general covariance condemned

During these three years, Einstein formulated an argument that would decisively redirect his understanding of general covariance. He and Grossmann had been unable to find acceptable generally covariant field equations. The so-called 'hole argument' purported to show that this circumstance need not worry them since all generally covariant field equations would be physically uninteresting. Einstein published the argument four times in 1914, appearing, for example, as a later appendix to the journal printing of Einstein and Grossmann (1913). Its clearest exposition was in a review article (Einstein 1914, pp 1066–7)††.

The argument was beguilingly simple. Einstein asked us to imagine a region of spacetime devoid of matter—the 'hole'—in which the stress energy tensor $T_{\mu\nu}$ vanished. He now assumed that we had generally covariant gravitational field equations and that $g_{\mu\nu}$ was a solution for this spacetime in a coordinate system x_τ. Einstein transfromed to a new coordinate system x'_τ which agreed with x_τ outside the hole but came smoothly to differ from it within the hole. In the new coordinate system the metric would be $g'_{\mu\nu}$ and constructed according to the usual tensor transformation law. That is, the same gravitational field would be represented by $g_{\mu\nu}$ in coordinate system x_τ and by $g'_{\mu\nu}$ in coordinate system x'_τ.

At this point Einstein effected a subtle manipulation that is the key to the hole argument. One could consider the symmetric matrix $g_{\mu\nu}(x_\tau)$ as a set of ten functions of the variable x_τ and $g'_{\mu\nu}(x'_\tau)$ as a set of ten functions of the variable x'_τ. One can now construct a new set of ten functions $g'_{\mu\nu}(x_\tau)$. That is, take the ten functions of the *new* matrix $g'_{\mu\nu}$ and consider them as functions of the *old* coordinates x_τ. The original $g_{\mu\nu}(x_\tau)$ and the construction $g'_{\mu\nu}(x_\tau)$ cannot represent the same gravitational field in different coordinate systems. They are both defined on the same coordinate system x_τ, yet they have different components, since $g_{\mu\nu}$ and $g'_{\mu\nu}$ are different functions. That is, $g_{\mu\nu}(x_\tau)$ and $g'_{\mu\nu}(x_\tau)$ represent *different* gravitational fields in the *same* coordinate system. Now, by their construction, the functions $g_{\mu\nu}(x_\tau)$ and $g'_{\mu\nu}(x_\tau)$ will be the same outside the hole, but they will come smoothly to differ within the hole. Thus the two sets of functions represent distinct gravitational fields. Let us call them g and g'. The fields g and g' are the same outside the hole but come smoothly to differ within the hole.

† This fascinating episode has been dissected in some detail with some help from his private calculation (see Stachel 1980 and Norton 1984).

†† For further discussion see Norton (1987).

Einstein has presumed the field equations general covariant. Therefore, if they are solved by the $g_{\mu\nu}(x_\tau)$, then they must be solved by $g'_{\mu\nu}(x'_\tau)$ and therefore also by the construction $g'_{\mu\nu}(x_\tau)$. That is, generally covariant gravitational field equations allow as solutions the two distinct gravitational fields g and g'. Einstein found this outcome unacceptable. For the one matter distribution outside the hole now clearly fails to determine what the gravitational field would be within the hole. That is, we could specify the matter distribution and gravitational field everywhere in spacetime excepting some matter-free hole that could be arbitrarily small in both spatial and temporal extent. Nonetheless generally covariant field equations would be unable to determine what the gravitational field would be within this hole. This was a dramatic failure of what he called the law of causality and we might now call determinism. Einstein deemed the failure sufficiently troublesome to warrant rejection of generally covariant gravitational field equations as physically interesting†.

3.4. Einstein's 1916 account of the foundations of general relativity: general covariance regained

In November 1915, Einstein's long struggle with his 'Entwurf' theory came to a close. His resistance to general covariance finally broke under the accumulating weight of serious problems in his 'Entwurf' theory. His return to general covariance and the final general theory of relativity were reported to the Prussian Academy in a series of hasty communications that chronicle the tense confusions of these last desperate days††. Early the following year, Einstein (1916) sent *Annalen der Physik* a review article on the final theory.

The article's account of the theory's foundations was written with a freedom unavailable to Einstein in the dark years of the 'Entwurf' theory. Throughout those years, Einstein had maintained his allegiance to the relativity of inertia. That allegiance had to rest principally on a sincere hope of what might be demonstrable. He had not demonstrated the unconditional relativity of inertia in his 'Entwurf' theory; he was still sure only of weak field effects compatible with the relativity of inertia (Einstein 1913, section 9) and similar to those he had found in his 1912 theory. More vexing, however, was the very public failure of general covariance, which compromised the claim that he was extending the principle of relativity. Einstein did not report on equally serious problems that had befallen the principle of equivalence. The simple 1907/1911 version of the principle required only equivalence of *uniform* acceleration and a *homogeneous* gravitational field. Yet in the final version of the 1912 theory, the principle had to be restricted to infinitesimally small regions of space. Einstein found the need for this restriction extremely puzzling since the restriction was not invoked to homogenize an inhomogeneous field. Worse, in the 'Entwurf' theory, even this restriction failed to save this form of the principle, which had to be reported as a result of his earlier 1912 theory (see Norton (1985, section 4.3) for a discussion).

By 1916, Einstein's problems with general covariance had evaporated and with them the problems with the principle of equivalence. Thus the 1916 review article could commence with a more confident account of the theory's foundations which remains today one of the most widely known of Einstein's accounts. The exposition began with a series of now

† It was pointed out much later by Stachel (1980), using mathematical notions not available to Einstein in 1913, that the new gravitational field g' was generated from g as the carry along $g' = h^*g$ under the diffeomorphism h induced by the coordinate transformation x_τ to x'_τ. The indeterminism that worried Einstein so profoundly is now routinely obliterated as a gauge freedom associated with arbitrary diffeomorphism so that, while g and g' may be mathematically distinct, they are not judged to represent physically distinct gravitational fields (see Wald 1984, p 438).

†† Einstein 1915. For dissection of this episode, see Norton (1984, sections 7, 8).

familiar considerations all of which drove towards general covariance.

Both special relativity and classical mechanics, Einstein reported, suffered an epistemological defect. It was illustrated with Einstein's variant of Newton's bucket. Two fluid bodies hover in space. They are in an observable state of constant relative rotation about a line that connects them. In spite of the obvious symmetry of this set up, Einstein supposed that one sphere S_1 proves to be spherical when surveyed and the other S_2 proves to be an ellipsoid of revolution. Classical mechanics and special relativity could explain the difference by supposing that the first sphere is at rest in an inertial frame of reference, introduced by Einstein into the argument as a 'privileged Galilean space,' and that the second is not. This explanation, Einstein objected, violates the 'demand of causality,' for these privileged frames are 'merely factitious causes' and not an observable thing. The true cause of the difference must lie outside the system, Einstein continued, immediately identifying the true cause in the disposition of distant masses. In effect Einstein used his example to conclude that the only theory that could satisfactorily account for this example was one that satisfied the requirement of the relativity of inertia. Any such theory, Einstein continued, cannot single out any inertial frame as preferred. Therefore:

> *The laws of physics must be of such a nature that they apply to systems of reference in any kind of motion.* Along this road we arrive at an extension of the postulate of relativity (Einstein's emphasis).

Einstein then introduced the principle of equivalence in the form given above in section 3.1 in which it asserts the equivalence of uniform acceleration and a homogeneous gravitational field. The principle is used to suggest that a theory which implements a generalized principle of relativity will also be a theory of gravitation. Einstein then turns to deal with a complication that arises from using accelerated frames of reference in special relativity. In accelerated frames, in particular in rotating frames, geometry ceases to be Euclidean and clocks are slowed in a position-dependent manner. As a result it turns out that one can no longer easily define space and time coordinate systems by the familiar operations of laying out rods and using standard clocks. This apparent complication—and not the need for a generalization of the principle of relativity—leads Einstein to propose general covariance†:

> The method hitherto employed for laying co-ordinates into the space-time continuum in a definite manner thus breaks down, and there seems to be no other way which would allow us to adapt systems of co-ordinates to the four-dimensional universe so that we might expect from their application a particularly simple formulation of the laws of nature. So there is nothing for it but to regard all imaginable systems of co-ordinates, on principle, as equally suitable for the description of nature. This comes to requiring that:–
>
> *The general laws of nature are to be expressed by equations which hold good for all systems of co-ordinates, that is, are co-variant with respect to any substitutions whatever (generally covariant).*
>
> It is clear that a physical theory which satisfies this postulate will also be suitable for the general postulate of relativity. For the sum of *all* substitutions in any case includes those which correspond to all relative motions of three-dimensional systems of co-ordinates (Einstein's emphasis).

† A footnote at the word 'imaginable' was omitted from the standard Perrett and Jeffrey English translation. It says: 'Here we do not want to discuss certain restrictions which correspond to the requirement of unique coordination and of continuity.' This now essentially unknown footnote shows that Einstein did at least once apologize for his failure to specify precisely which group of transformations was intended by 'any substitutions whatever.'

Why did Einstein not simply insist that the generalization of the principle of relativity to accelerated motion forces general covariance? Following the analogy with Lorentz covariance, the generalized principle of relativity would require an extension of the covariance of the theory to include transformations between frames in arbitrary states of motion. But general covariance extends it even further. It includes transformations that have nothing to do with changes of states of motion, such as the transformation between Cartesian and polar spatial coordinates. But, as Einstein indicates, he feels compelled to go to this larger group since he can see no natural way of restricting the spacetime coordinate system.

3.5. The point-coincidence argument

Immediately following the above statement of the requirement of general covariance, Einstein gave another argument for general covariance which John Stachel has conveniently labelled the 'point-coincidence argument.'

> That this requirement of general co-variance, which takes away from space and time the last remnant of physical objectivity, is a natural one, will be seen from the following reflexion. All our space-time verifications invariably amount to a determination of space-time coincidences. If, for example, events consisted merely in the motion of material points, then ultimately nothing would be observable but the meetings of two or more of these points. Moreover, the results of our measurings are nothing but verifications of such meetings of the material points of our measuring instruments with other material points, coincidences between the hands of a clock and points on the clock dial, and observed point-events happening at the same place and the same time.
>
> The introduction of a system of reference serves no other purpose than to facilitate the description of the totality of such coincidences. We allot to the universe four space-time variables x_1, x_2, x_3, x_4, in such a way that for every point-event there is a corresponding system of values of the variables $x_1 \ldots x_4$. To two coincident point-events there corresponds one system of values of the variables $x_1 \ldots x_4$, i.e. coincidence is characterized by the identity of the co-ordinates. If, in the place of the variables $x_1 \ldots x_4$, we introduce functions of them, x'_1, x'_2, x'_3, x'_4, as a new system of co-ordinates, so that the system of values are made to corespond to one another without ambiguity, the equality of all four co-ordinates in the new system will also serve as an expression for the space-time coincidence of the two point-events. As all our physical experience can be ultimately reduced to such coincidences, there is no immediate reason for preferring certain systems of co-ordinates to others, that is to say, we arrive at the requirement of general co-variance.

This point-coincidence argument is cited very frequently in the literature since 1916. However its real purpose was essentially completely forgotten until it was rediscovered and revealed by Stachel (1980). Einstein's 1916 exposition of general relativity contained a very puzzling omission. In the years immediately preceding, by means of the hole argument, Einstein had apparently proved that any generally covariant theory would be physically uninteresting. Yet here was Einstein extolling exactly such a theory without explaining where the hole argument went astray.

That melancholy task of correcting his past error was the real function of the point coincidence argument. This was precisely the use to which the argument was put in Einstein's correspondence of December 1915 and January 1916 (see Norton 1987, section 4). According to Einstein's assumption, the physical content of a theory is fully exhausted by a catalogue of the spacetime coincidences it sanctions. Therefore any transformation that preserves these coincidences preserves its physical content. Now the transformation used in the hole argument from the field g to the mathematically distinct field g' is more than a mere transformation of coordinates. For g and g' are mathematically distinct fields in the same coordinate system. However the transformation from g to g' is one that preserves all coincidences. Therefore g and g' represent the same physical field. Whatever indeterminism is revealed in the hole argument is a purely mathematical freedom akin to a gauge freedom and offers no obstacle to the physical interest of a generally covariant theory.

Einstein scarcely ever mentioned the debacle of the hole argument again in print. However it continued to inform his ideas about covariance, spacetime, fields and coordinate systems. For example, in executing the hole argument, in order to effect the transition from $g_{\mu\nu}(x_\tau)$ to $g'_{\mu\nu}(x_\tau)$, one has to assume, in effect, that the coordinate system x_τ, has some real existence, independent of the $g_{\mu\nu}$ or $g'_{\mu\nu}$. For, figuratively speaking, one has to remove the field $g_{\mu\nu}$, leaving the bare coordinate system x_τ, and then insert the new field $g'_{\mu\nu}$. In a letter of December 26, 1915, to Paul Ehrenfest, Einstein explained that one defeats the hole argument by assuming among other things that 'the reference system signifies nothing real'.† We hear these echoes of the hole argument when Einstein (1922, p 21) proclaims in a May 1920 address in Leiden:

> There can be no space nor any part of space without gravitational potentials; for these confer upon space its metrical qualities, without which it cannot be imagined at all.

These same echoes still reverberate in the 1952 appendix to Einstein's popular text *Relativity: The Special and the General Theory*, when Einstein (1952, p 155) insists

> . . . a pure gravitational field might have been described in terms of the g_{ik} (as functions of the co-ordinates), by solution of the gravitational equations. If we imagine the gravitational field, i.e. the functions g_{ik}, to be removed, there does not remain a space of the type (1) [Minkowski spacetime], but absolutely *nothing*, and also no 'topological space' (Einstein's emphasis).

Most recently, the hole argument has enjoyed a revival in the philosophy of space and time literature where, in variant form, it provides a strong argument against the doctrine of spacetime substantivalism (Earman and Norton 1987). For further discussion of the background and ramifications of the hole and point-coincidence arguments see Howard (1992) and Ryckman (1992).

3.6. The Göttingen defence of general covariance

The most prominent legacy of the hole argument in the literature on general relativity does not arise from Einstein's analysis, however. In 1915 and 1917, David Hilbert (1915, 1917) published a two-part paper on general relativity which proved to be enormously influential. Citing the hole argument, Hilbert (1917, pp 59–63) turned to the question of the 'principle of causality'. He observed that his formulation of general relativity employed fourteen independent variables, that is, ten metrical components for the gravitational field and four potentials for the electromagnetic field. However in the joint theory of gravitational and electromagnetic fields, four identities reduced the fourteen field equations to only ten independent equations. The indeterminism lay in the freedom to set the four remaining conditions. These four conditions could, however, be absorbed in four stipulations used to specify a coordinate system.

Hilbert insisted that this underdetermination of the field was not physical. Echoing the geometric themes of his Göttingen colleagues Klein and the late Minkowski, he recalled (p 61) '. . . an assertion that does not remain invariant under any arbitrary transformation of the

† As quoted in Norton (1987, p 169).

coordinate system is marked as *physically meaningless*' (Hilbert's emphasis). He then argued that the four degrees of freedom did not leave the invariant content of the theory underdetermined. His example was an electron at rest in some coordinate system. A coordinate transformation leaves the electron unchanged in the past of some instant specified by time coordinate $x_4 = 0$, but sets it in motion in the future. The two coordinate descriptions are the same in the past, the electron is at rest, but in the future only one describes the electron as moving. The one past can extend to different futures. The differences, however, have no physical significance, since the relevant assertions about the electron's motion are not invariant. One could make them invariant by introducing an invariant coordinate system adapted to the spacetime geometry, such as the Gaussian system Hilbert considered. Coordinate based assertions of the electron's motion would now be invariant, but they would no longer be underdetermined since the introduction of the Gaussian system used up the four remaining degrees of freedom.

Hilbert's depiction of the indeterminism of a generally covariant theory was in terms of a count of independent field variables and independent field equations. It is the version that rapidly came to appear most often in the literature (e.g. Pauli 1921, section 56). The four identities among the field equations that allowed the underdetermination were only later connected with the contracted Bianchi identities (see Mehra 1974, section 7.3). Again Hilbert's discussion and his example of the electron was the first treatment of the Cauchy problem in general relativity, so that the literature on the Cauchy problem can trace its descent back to Einstein's hole argument (see Stachel 1992)†.

3.7. Einstein's three principles of 1918

In March 1918, Einstein (1918) returned to the question of the fundamental principles of general relativity. As he made clear in his introductory remarks, the paper was provoked by Kretschmann's (1917) criticism (see section 5.2 below). However its purpose was to lay out his understanding of the foundations of his theory. This exposition differed from the 1916 account in at least one major area. In 1916, Einstein assumed that his generally covariant theory would satisfy the relativity of inertia, although no proof had been given. At best Einstein would have been able to point to weak field effects compatible with the relativity of inertia. (These weak field effects are of the same type as those he reported in the 'Entwurf' theory in Einstein (1913, section 9) and are described in his text (Einstein 1922a, p 100)).

By 1917, Einstein had found that a simple reading of the relativity of inertia was incompatible with his theory. He reported this failure in an introductory section (section 2) to his famous paper on relativistic cosmology (Einstein 1917). On the basis of the relativity of inertia, he expected that the inertia of a body would approach zero if it was moved sufficiently far from other masses in the universe. This expectation would be realized in the theory if the spacetime metric adopted certain degenerate values at a mass-free spatial infinity. However Einstein found that such degenerate behaviour was inadmissible in his theory. Instead he seemed compelled to postulate some non-degenerate boundary conditions for the metric at a mass-free spatial infinity, such as Minkowskian values.

This Minkowskian boundary condition became the embodiment for Einstein of the failure of the relativity of inertia. For this boundary condition made a definite contribution to the inertia of a test body that could not be traced to other masses. That is, with these boundary conditions, the inertia of a body was influenced by the presence of other masses, in

† Howard and Norton (forthcoming) conjecture that there was an encounter in 1915 between the Göttingen resolution of the hole argument and an unreceptive Einstein, still convinced of the correctness of the hole argument.

so far as they affected the metric field. However its inertia was not fully determined by the other masses. Therefore, if the relativity of inertia was to be satisfied, it was necessary to abolish these arbitrarily postulated boundary conditions. (The question of whether this was also sufficient remained unaddressed.) Einstein succeeded in abolishing these boundary conditions at spatial infinity by a most ingenious ploy: he abolished spatial infinity itself. He introduced the first of the modern relativistic cosmologies, the one we now call the 'Einstein universe', which is spatially closed and finite. The price Einstein had to pay turned out to be high. In order for his field equations to admit the Einstein universe as a solution, he needed to introduce the extra 'cosmological' term in his field equations. In his notation and formulation of 1917, with $G_{\mu\nu}$ representing the *Ricci* tensor and κ a constant, this meant that the old field equations

$$G_{\mu\nu} = -\kappa(T_{\mu\nu} - \tfrac{1}{2}g_{\mu\nu}T)$$

were replaced by

$$G_{\mu\nu} - \lambda g_{\mu\nu} = -\kappa(T_{\mu\nu} - \tfrac{1}{2}g_{\mu\nu}T)$$

The cosmological term is $\lambda g_{\mu\nu}$ and λ is the cosmological constant.

This development was essential background to understanding the three principles Einstein listed in (Einstein 1918, pp 241–2) as those on which his theory rested.

(*a*) *Principle of relativity.* The laws of nature are only assertions of timespace coincidences; therefore they find their unique, natural expression in generally covariant equations.

(*b*) *Principle of equivalence.* Inertia and weight are identical in essence. From this and from the results of the special theory of relativity, it follows necessarily that the symmetric 'fundamental tensor' ($g_{\mu\nu}$) determines the metric properties of space, the inertial relations of bodies in it, as well as gravitational effects. We will call the condition of space, described by the fundamental tensor, the 'G-field.'

(*c*) *Mach's principle.* The G-field is determined *without residue* by the masses of bodies. Since mass and energy are equivalent according to the results of the special theory of relativity and since energy is described formally by the symmetric energy tensor ($T_{\mu\nu}$), this means that the G-field is conditioned and determined by the energy tensor.

The separation of the principle of relativity and Mach's principle into two distinct principles was clearly the product of Einstein's experience with the cosmological problem. If the Einstein of 1916 had assumed that the relativity inertia would be satisfied automatically within a generally covariant theory, then the Einstein of 1918 no longer harboured such delusions. The 1918 version of the principle of relativity seems to assert something less than a fully generalized relativity of the motion of bodies. In effect it merely asserts the key thesis of the point-coincidence argument: the physical content of a theory is exhausted by its catalogue of allowed spacetime coincidences. General covariance follows from this thesis as a consequence. The principle of relativity (a) is now supplemented by the new Mach's principle (c) and it is only their conjunction that begins to resemble Einstein's original goal of a fully generalized relativity of motion. In effect Mach's principle (c) was intended to capture in a field theoretic setting the old, Mach-inspired requirement of the relativity of inertia. It was to rule out the arbitrary postulation of boundary conditions for the metric field at spatial infinity, which, Einstein reported in 1917, compromised the relativity of inertia. All this was alluded to by Einstein in a footnote to the title 'Mach's Principle,' which also announced that he was introducing the name for the first time:

> Up to now I have not distinguished principles (a) and (c) and that caused confusion. I have chosen the name 'Mach's principle' since this principle is a generalization of Mach's requirement that inertia be reducible to an interaction of bodies.

Einstein's wording of the principle of equivalence (b) was an interesting departure in so far as it now emphasized that the principle depended on the empirical equality of two quantities, inertial and gravitational mass, and that the effect of the principle had been to unify them completely. However there was little real change from Einstein's earlier use of the principle, as was shown by the remainder of the paragraph that described the principle. In effect it gave a synopsis of the transition from the line element (3) to (3′) and (4) and the resulting interpretation of the non-constant coefficients of (3′) and (4) as representing the gravitational field, as well as the inertial and geometric properties of spacetime.

3.8. Mach's principle forsaken

For all his efforts, Einstein's portrayal of the foundations of general relativity had still not reached its final form with the 1918 list. Over the years following, the principle of relativity and of equivalence retained their 1918 forms. However Einstein came to abandon Mach's principle.

The seeds of Einstein's disenchantment with Mach's principle were becoming apparent as early as 1919. Einstein (1919, section 1) described its offspring, the cosmological term added to his 1915 field equations, as 'gravely detrimental to the formal beauty of the theory'. With the discovery of the expansion of the universe, Einstein formally disowned the cosmological term (Einstein and de Sitter 1932). In any case, the augmentation of his field equations with the cosmological term had forced neither the relativity of inertia nor Mach's principle into his theory. For it had not eliminated the possibility of essentially matter-free solutions of the field equations. In such solutions, the inertia of a test body could not be attributed to other masses. These solutions were the subject of an extended exchange in publication and in private between Einstein and de Sitter towards the end of the 1910s (see Kerszberg 1989).

Einstein also began to distance himself from the relativity of inertia. Whereas the idea was urged without reservation up to 1916, he soon came to describe it as a very significant idea, but one of essentially historical interest only. For example, Einstein (1924, p 87) attributed to Mach the idea that inertia arose as an unmediated interaction between masses. But he dismissed it casually as 'logically possible, but cannot be considered seriously any more today by us since it is an action-at-a-distance theory'.† Einstein (1924, p 90) did still maintain that the metric is fully determined by ponderable masses in a spatially finite cosmology according to his theory, although the term 'Mach's principle' was not used. As time passed, Einstein had fewer and fewer kind words for this Machian approach to inertia. He explained in 1946 for example in his *Autobiographical Notes* (1949, p 27)

> Mach conjectures that, in a truly reasonable theory, inertia would have to depend upon the interaction of the masses, precisely as was true for Newton's other forces, a conception that for a long time I considered in principle the correct one. It presupposes implicitly, however, that the basic theory should be of the general type of Newton's mechanics: masses and their interaction as the original concepts. Such an attempt at a resolution does not fit into a consistent field theory, as will be immediately recognized.

His 1918 Mach's Principle had been an attempt to translate this requirement on masses and their interactions into field theoretic terms, but he soon seemed to lose enthusiasm even for

† The same point is made less forcefully in Einstein (1922, pp 17–18) and Einstein (1922a, p 56).

this enterprise. The difficulty was that the 1918 principle required that the metric field $g_{\mu\nu}$ be determined by the masses of bodies as represented by the stress-energy tensor $T_{\mu\nu}$. However this gave a primary determining function to a quantity, $T_{\mu\nu}$, which Einstein (1949, p 71) reported he had always felt was 'a formal condensation of all things whose comprehension in the sense of a field theory is still problematic' and one that was 'merely a makeshift'. Einstein gave a final synopsis of Mach's principle in a letter of February 2, 1954 to Felix Pirani in the year prior to his death. Citing the above difficulty with the stress-energy tensor and the fact that this tensor presumes the metric, he labeled his 1918 version of Mach's principle 'a ticklish affair' and concluded 'In my opinion we ought not to speak about Mach's principle any more.'†

3.9. Einstein's causal objection to absolutes

When Einstein disowned the relativity of inertia and Mach's principle, he actually disowned somewhat less than it first seemed. Both these principles were introduced to solve a problem in earlier theories of space and time: these theories were defective in the way they used inertial systems as causes. Einstein still clearly maintained that the problem was serious and that his general theory of relativity had solved it. However he had originally thought the solution was best expressed in terms inspired by his reading of Mach; that is, as a generalized relativity of the motion of bodies. As he put it in Einstein (1913, p 1260)

> To talk of the motion and therefore also acceleration of a body A in itself has no meaning. One can only speak of the motion or acceleration of a body A relative to other bodies B, C *etc*. What holds in kinematic relation for acceleration ought also to hold for the inertial resistance, with which bodies oppose acceleration . . .

He was led away from this Machian characterization of the solution by his work on Mach's principle and the cosmological problem, as well as his preference for field rather than body as a primitive notion. We shall see that his mature characterization of the solution was that general relativity allowed space and time to be mutable. They no longer just acted causally, they could also be acted upon and, in this sense, had lost their absolute character. In Einstein's mature view, it is this special causal property that distinguishes general relativity from earlier theories and possibly even justifies the name 'general relativity', in so far as it is the field theoretic translation of Einstein's original notion of the generalized relativity of the motion of bodies.

In the early years of Einstein's theory, the causal defect was located most prominently in the mere fact of the older theories' use of an inertial reference system as a cause. Thus in Einstein's 1916 review article, he sought to account for the centrifugal bulges in a rotating fluid body (see section 3.4 above). To say that the body bulges because it rotated with respect to an inertial frame of reference is to introduce a 'merely *factitious* cause, and not a thing that can be observed (1916, p 113)'. This same example is treated similarly in Einstein (1914a, pp 344–6). Einstein (1917a) makes clear the sort of cause that he would find acceptable in his popular exposition of relativity. In ch XXI he asks for the reason for the preferred status of inertial systems. He draws an analogy with two pans of water on a gas range. One is boiling; one is not. The difference, Einstein insists, only becomes satisfactorily explained when we notice the bluish flame under the boiling pan and none under the other.

Einstein soon came to stress a different aspect of these earlier theories as causally defective. He identified this aspect with their absolute character. In his *Meaning of Relativity* (1922a, p 55) he wrote in parody of Newton's Latin

† Translation from Torretti (1983, p 202) with '*dem Mach'schen Prinzip*' rendered as 'Mach's principle'.

> . . . from the standpoint of the special theory of relativity we must say, *continuum spatii et temporis est absolutum*. In this latter statement *absolutum* means not only 'physically real', but also 'independent in its physical properties, having a physical effect, but not itself influenced by physical conditions'.

and continued to explain that such absolutes are objectionable since (pp 55–6)

> . . . it is contrary to the mode of thinking in science to conceive of a thing (the space-time continuum) which acts itself, but which cannot be acted upon.

The text immediately turned to Mach's ideas and, later (pp 99–108) to the weak field effects compatible with the relativity of inertia and his 1917 field formulation of this idea in a spatially closed cosmology. Around this same time, Einstein's briefer summaries advertised general relativity as eliminating the absoluteness of space and time (Einstein 1972, p 260)†:

> Space and time were thereby divested not of their reality but of their causal absoluteness—i.e. affecting but not affected.

In these briefer summaries, Einstein was no longer insisting that the spacetime metric was to be *fully* determined by the distribution of masses. Space and time had lost their absoluteness simply because they were no longer immutable. By the 1950s, as Einstein explained to Pirani, he no longer endorsed his 1918 Mach's principle. However he did retain the idea that the earlier theories were causally defective in admitting such absolutes (e.g. Einstein 1950, p 348) and, as he explained in the 'completely revised' (p 0) 1954 appendix to his *Meaning of Relativity* (1922, pp 139–40), general relativity had resolved the problem as its essential achievement:

> It is the essential achievement of the general theory of relativity that it has freed physics from the necessity of introducing the 'inertial system' (or inertial systems). . . . Thereby [in earlier theories], space as such is assigned a role in the system of physics that distinguishes it from all other elements of physical description. It plays a determining role in all processes, without in its turn being influenced by them.

This view of the deficiency of earlier theories and general relativity's achievement is not one that grew in the wake of Einstein's disenchantment with Mach's principle. Rather, it was present even in his earliest writings beneath the concerns for the relative motion of bodies and the observability of causes. Einstein (1913, pp 1260–1) makes the essential point:

> . . . in [theories current today], the inertial system is introduced; its state of motion, on the one hand, is not conditioned by the states of observable objects (and therefore caused by nothing accessible to perception) but, on the other hand, it is supposed to determine the relations of material points.

A footnote earlier in the paragraph also tried to identify what was so unsatisfactory about inertial systems

> What is unsatisfactory about this is that it remains unexplained *how* the inertial system can be singled from other systems.

Thus we have here the enduring core of the cluster of ideas that led Einstein to the relativity of inertia and Mach's principle: his concern that, through their introduction of inertial systems, earlier theories allowed absolutes that acted but could not be acted upon.

Finally, we may ask whether the 'essential achievement' of general relativity, the elimination of the absolute inertial systems, follows automatically from general covariance in Einstein's view, so that general covariance would then truly amount to a generalized principle of relativity in a form adapted to a field theory. It is hard to find a clear answer in Einstein's writings. His 1918 catalogue of three principles suggested that the requirement of general covariance ('(a) principle of relativity') needed to be supplemented by something

†See also Einstein (1922, p 18, 1924, p88).

additional ('(c) Mach's principle') to realize fully the general relativity of motion. Einstein's text suggests this without clearly stating it, for Einstein (1918, p 241) introduces the three principles with the remark that they are 'in any case in no way independent of one another'. Thus it is not clear whether these particular two of the three principles really are independent or, if they are not, whether general covariance somehow leads to Mach's principle. Perhaps the best answer we will find is Einstein's repeated insistence that general covariance, in conjunction with a requirement of simplicity, leads us directly to general relativity (see, for example, Einstein (1952, pp 152–3, 1949, pp 71–3; 1933, p 274)) And it is this theory that eliminates the absoluteness of the inertial system.

4. The favourable text-book assimilation of Einstein's view: fragmentation and mutation

Although Einstein had to struggle to gain acceptance of this theory in its earliest years (especially prior to 1916), by 1920 Einstein's new theory was widely celebrated. The extravagant publicity surrounding the success of Eddington's 1919 eclipse expedition had even launched Einstein into the popular press and public eye. During this period, the vast majority of accounts of Einstein's theory merely sought to recapitulate Einstein's own account. Thus began the tradition of writing in what I call the favourable assimilation of Einstein's view and which is to be reviewed in this section. I shall consider an account of the foundations of general relativity favourably to Einstein's view if it names some or all of Einstein's three principles of 1918 as foundations of the theory: principle of relativity/covariance, principle of equivalence and Mach's principle; it must include at least the first principle.

Two things will become clear about the favourable reception of Einstein's account of the foundations of general relativity. First, it is very widespread and still a major tradition today. Second, what is often offered as a recapitulation of Einstein's account—even if only tacitly—can differ in very significant ways from what Einstein really said. Most prominently, the relativity of inertia and Mach's principle is only infrequently reported as part of the foundations of general relativity in more technical expositions. This disfavour is not a response to Einstein's own later disillusionment with Mach's principle. From the earliest moments, the principle failed to find a place in the majority of accounts within more technical expositions. Rather the favourable accounts rapidly stabilized, most commonly, into locating the foundations of general relativity in the principle of equivalence and the principle of general covariance, with the latter understood as a generalization of the principle of relativity. Even here, these accounts have failed to remain faithful to Einstein's viewpoint. They almost exclusively employ an infinitesimal principle of equivalence, a variant form that Einstein never endorsed and was quite different in outlook from Einstein's own form.

In order to gauge the magnitude and character of the favourable reception, this section will review the favourable accounts of the foundations of general relativity as they have appeared in the textbooks on general relativity. The review is also limited principally to expositions that either provide a self-contained exposition of tensor calculus or sufficient differential geometry for general relativity or presume such knowledge in the reader and that proceed at least as far as a formulation of the gravitational field equations. We should note also that the favourable reception extends beyond the realm of relativity theory. Aguirre and Krause (1991, p 508) are prepared to label a mechanics 'general relativistic' merely because it is generally covariant.

Jean Eisenstaedt (1986, 1989) has described the rising and falling fortunes of general relativity. After an initial period of great interest and activity in the late 1910s and early 1920s, the theory fell into decades of neglect because of many factors: a sense that the theory had only slender confirmation, that its practical utility to physicists was small and that the theory had been eclipsed by the developments in quantum theory. The 1960s saw a new vigour in work on the theory, in part due to a renewed interest in empirical test of the theory and to the exploitation of new, more sophisticated mathematical tools. In the following, the favourable reception is divided into periods reflecting these shifts in intensity of work. First, however, I will review the special problem of the principle of equivalence.

4.1. Einstein's principle of equivalence as a covariance principle and its later misrepresentation

There are many instances of later accounts misrepresenting Einstein's ideas. None is as universal and complete as the later treatments of Einstein's principle of equivalence. In his *Meaning of Relativity*, Einstein gives a statement of the principle typical of all his writing. K is an inertial system in special relativity and K' a system of coordinates uniformly accelerating with respect to K. Having noted that free masses in K' are accelerated 'just as if a gravitational field were present and K' unaccelerated', Einstein (1922a, p 57–8) then writes:

> . . . there is nothing to prevent our conceiving this gravitational field as real, that is, the conception that K' is 'at rest' and a gravitational field is present we can consider as equivalent to the conception that only K is an 'allowable' system of co-ordinates and no gravitational field is present. The assumption of the complete physical equivalence of the systems of coordinates, K and K', we call the 'principle of equivalence'; . . . [it] signifies an extension of the principle of relativity to co-ordinate systems which are in non-uniform motion relatively to each other. In fact, through this conception we arrive at the unity of nature of inertia and gravitation.

Einstein, however, is nearly universally understood as urging a rather different principle, which I shall call the 'infinitesimal principle of equivalence'. A canonical formulation is given in Pauli (1921, p 145):

> 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 process.

The key idea here is that in adopting a sufficiently small region of spacetime, an *arbitrary* gravitational field becomes homogenous and can be transformed away by a suitable choice of coordinate system. This principle exists in many variant forms. Sometimes it is strengthened to require that when the gravitational field is transformed away, we recover special relativity locally (for example, Misner *et al.*, 1973, p 386). With somewhat different qualifications, Pauli's infinitesimal principle corresponds to Dicke's 'strong equivalence principle (Roll *et al.*, 1964, p 444). Dicke's 'weak equivalence principle', however, requires only the uniqueness of gravitational acceleration, which amounts to requiring that the trajectories of free fall of suitably idealized bodies are independent of their constitutions.

Unlike most other writers, Pauli (1921, p 145) acknowledged that his infinitesimal version of the principle of equivalence differed from Einstein's, suggesting that, where Einstein's principle applied only to homogeneous gravitational fields, Pauli's version was for the 'general case'. However the differences ran far deeper than Pauli allowed and pertain to quite fundamental questions of the role of the principle of equivalence in general relativity. These differences can be summarized in three essential aspects of the principle which remained fixed throughout Einstein's writings on general relativity, from the earliest moments in 1907, to his final years in the 1950s†:

- Einstein's principle of equivalence was a covariance principle.

Special relativity required the complete physical equivalence of all inertial coordinate systems; for Einstein, general relativity required the complete equivalence of all coordinate systems. Einstein's principle of equivalence required the complete equivalence of a set of coordinate systems of intermediate size: inertial coordinate systems plus uniformly accelerated coordinate systems. That is, the principle sanctioned the extension of the covariance of special relativity beyond Lorentz covariance but not as far as general covariance. Thus, for Einstein, the principle of equivalence was a relativity principle intermediate in range between the principle of relativity of special relativity and of general relativity.

The point is so important for our concerns here that it is helpful to have it in Einstein's own words of (1950, p 347):

> This is the gist of the principle of equivalence: In order to account for the equality of inert and gravitational mass within the theory it is necessary to admit non-linear transformations of the four coordinates. That is, the group of Lorentz transformations and hence the set of 'permissible' coordinate systems has to be extended.

Or, more succinctly, in an article devoted to explicating precisely what he intended with his principle of equivalence, Einstein (1916a, p 641) wrote in emphasized text:

> The requirement of general covariance of equations embraces that of the principle of equivalence as a quite special case.

The function of the alternative, infinitesimal principle of equivalence is to stipulate that a spacetime of general relativity with an arbitrary gravitational field is in some sense locally—that is, in infinitesimal regions—like the spacetime of special relativity. (Einstein objected in correspondence with Schlick to the latter's use of this idea, pointing out to Schlick that the sense in which special relativity holds locally must be so weak that accelerated and unaccelerated particles cannot be distinguished. For details, see Norton (1985, section 9).) As a covariance principle, Einstein's version of the principle served no such function. Therefore it was invariably restricted in the following related ways:

- Einstein's principle of equivalence was applied only in special relativity to what we now would call Minkowski spacetimes.

That is, the inertial coordinate system K of Einstein's formulation of the principle is not some kind of free fall coordinate system of general relativity. It is simply an inertial coordinate system of special relativity. Thus the coordinate systems K and K' are both coordinate systems of a Minkowski spacetime. Because of this, we would now be inclined to picture the entire principle as operating within special relativity. This seems not to have been Einstein's view. He seems to have regarded special relativity supplemented with the principle of equivalence as having more physical content than special relativity alone. The supplemented theory had a wider covariance and it dealt with a new phenomenon, homogeneous gravitational fields.

- Einstein's principle of equivalence was not a prescription for transforming away *arbitrary* gravitational fields; it was just a recipe for creating a *special type* of gravitational field.

Einstein's principle of equivalence gave a recipe for creating a *homogeneous* gravitational field by transforming to a uniformly accelerated coordinate system. The infinitesimal principle gives a recipe for transforming away an arbitrary gravitational field: one first

† The case for these differences between Einstein's version and the common infinitesimal version of the principle is laid out in some detail in Norton (1985).

homogenizes it by considering an infinitesimal region of spacetime and then transforms it away by the reverse transformation of Einstein's principle. Einstein repeatedly insisted that *his* principle of equivalence did not allow one to transform away an arbitrary gravitational field, but only gravitational fields of a quite special type, those produced by acceleration of the coordinate system. (Einstein devotes a paragraph of near page length to this point (1916a, pp 640–1). See Norton (1985, section 2).)†

4.2. The early years: 1916–1930

Einstein had named Mach's principle as one of the three fundamental principles of general relativity. However, the principle or its precursor, the relativity of inertia, has played the least role in accounts of the foundations of general relativity. Typically the principle does not appear in the discussion of the foundations of the theory. If it appears in an exposition, it arises most commonly later in the context of the cosmological problem and not always in a favourable light, even in expositions otherwise well disposed to Einstein's viewpoint.

This pattern was set at the earliest moments. In 1916 and 1917 the Dutch astronomer de Sitter took up the task of allowing the Germans and British to exchange more than artillery shells. He presented a three part report to the British Royal Astronomical Society on Einstein's new theory of gravitation (de Sitter 1916). Whilst otherwise favourable to Einstein, its second part concluded with criticism of Einstein's notion of the relativity of inertia. Development of this criticism continued in the third part. Einstein's 1917 work on the cosmological problem and his 1918 formulation of Mach's principle did not improve the reception of his ideas on the origin of inertia. Laue's (1921, pp 179–80) early general relativity text mentions them only in passing as incompatible with Minkowskian boundary conditions at spatial infinity. He finds the whole question physically too unclarified to warrant further discussion. Pauli (1921) does give the question more coverage, but only in a later, closing section (section 62). Einstein's ideas on the relativity of inertia figured more prominently in more popular expositions of general relativity. For example Freundlich (1919, section 4), Thirring (1922, section XV), Born (1924, ch VII, section 1) and Kopff (1923, pp 1–5, 191–5) treat the relativity of inertia. Indeed, the more popular the text, the more likely we are to find these ideas used to explain the foundations of general relativity.

The literature on Mach's principle has become enormous and is flourishing today. However its concerns have come to diverge from the concerns of this article, general covariance and the foundations of general relativity. The interested reader is referred to Reinhart (1973) and Torretti (1983, pp 194–202) for further discussion.

What is most important for our concerns is that the majority of expositions of relativity theory from this period emphasize the general covariance of general relativity as especially important. Of course this emphasis was justified if only for the novelty of general covariance. However the achievement of general covariance was also routinely assumed to ensure automatic satisfaction of a generalized principle of relativity. In some expositions this assumption was discussed in detail; in others it was merely suggested by labelling the requirement of general covariance, a principle of relativity. Accounts that emphasize general covariance and presume an automatic connection to a generalized principle of relativity include: De Sitter (1916, pp 700–02), Freundlich (119, p 28), Carmichael (1920, ch VII),

† Einstein himself never employed the trick of homogenizing an arbitrary gravitational field by considering infinitesimal regions of spacetime. In 1912, when his principle still dealt only with homogenous gravitational fields, he was forced to restrict it to infinitesimal regions of space to overcome certain technical difficulties with his theory of static gravitational fields. When they were overcome, the restriction disappeared. See Norton (1985, section 4.3).

Page (1920, p387), Schlick (1920, pp52–3), Cunningham (1921, ch VII), De Donder (1921, pp10–14), Laue (1921, p21), Pauli (1921, section 52), Weyl (1921, section 27), Becquerel (1922), Kottler (1922, pp 188–9), Thirring (1922, p151), Kopff (1923), Born (1924, ch VII), Reichenbach (1924, p141), Levi-Civita (1926, p294), Levinson and Zeisler (1929, p70). Some of these accounts explicitly invoke Einstein's point coincidence argument to establish general covariance. They include: De Sitter (1916, p700). Carmichael (1920, ch VII), Schlick (1920, pp52–3).

Many of these expositions also place great emphasis on the principle of equivalence. A few from the very earliest years state the principle in exactly Einstein's fashion: Thirring (1922, p109), Kopff (1923, p110) (also Carmichael (1920, p80), although critically). Others employ the now familiar infinitesimal principle of equivalence, other variant formulations of the principles or give vague characterizations of the principle that defy clear classification. The following at least name a principle of equivalence in the foundations of general relativity: Freundlich (1919, section 5), Laue (1921, pp18–21), Pauli (1921, p145), Becquerel (1922, section 55), Kottler (1922, p 192), Born (1924, ch VII), Reichenbach (1924, pp141–2).

4.3. The lean years: 1930–1960

During these three lean decades for general relativity, the volume of publication fell to the merest trickle. Within that trickle, Einstein's view of general covariance remained a dominant theme. Accounts of general relativity which emphasized the general covariance of the theory and either explicitly or tacitly took this general covariance to extend the principle of relativity include: Bergmann (1942, ch X), Schrödinger (1950, p2), Moller (1952, ch VII), Jordan (1955, section 14), Kratzer (1956, section 15), Bargmann (1957, p162), Tonnelat (1959, ch XI). All but Schrödinger and Jordan introduce a principle of equivalence by name.

Moller (1952, pp219–20) introduces general relativity with a discussion of the relativity of inertia. Tolman (1934, p3 and ch VI) is exceptional in offering Einstein's three principles of 1918—the principle of covariance, the principle of equivalence and Mach's principle—as the foundations of general relativity. However his version of the principle of equivalence is the infinitesimal version never endorsed by Einstein and he accepts Kretschmann's view of the physical vacuity of the principle of covariance, while insisting with Einstein on its heuristic value.

4.4. Rebirth: 1960–1980

The renaissance of general relativity in the 1960s brought clearer divisions in the literature on the foundations of general relativity. As we shall see below, one increasingly important strand either simply ignored Einstein's view of the foundations of the theory or became quite strident in its denunciation of Einstein's view. Another sought to repair Einstein's account in the face of such assaults. A major part of the literature, however, continued in simple assent with Einstein's view, only making smaller adjustment according to taste.

Most commonly, accounts in this last category found both an infinitesimal principle of equivalence and the principle of general covariance in the foundations of general relativity. Such accounts include: Weber (1961, sections 1.3, 2.4), Bergmann (1961, 1962), Lawden (1962, ch 6), Rosser (1964, sections 12.1, 12.2), McVittie (1965, ch 4), Yilmaz (1965, ch 15, 16), Skinner (1969, ch 3), Davis (1970, 5.I.2), Prasanna (1971, preface, ch 1), Mavridès (1973, sections III.4, III.5), Papapetrou (1974, Introduction, section 18), Pathria (1974, ch 6, 7), Bowler (1976, ch 9), Adler, Bazin and Schiffer (1977, p60 and section 5.1), Stephani (1977, section 8.1), Treder *et al.* (1980, Introduction). Most of these accounts explicitly connected general covariance with a generalized principle of relativity, either in name or by explicit discussion. These include: Bergmann (1961, 1962), Lawden (1962, ch 6), Rosser (1964, section 12.1), Yilmaz (1965, ch 15), Prasanna (1971), Mavridès (1973, section III.4), Papapetrou (1974, Introduction), Pathria (1974, ch 6), Bowler (1976, ch 9), Adler, Bazin and Schiffer (1977, section 5.1), Stephani (1977, section 8.1), Treder *et al.* (1980, Introduction). Skinner (1969, section 3.3.1) reported that the principle of general relativity required something beyond the principle of covariance: 'the laws of physics must determine the geometry of space-time appropriate for a particular physical circumstance'. Two accounts portrayed general covariance as a generalized principle of relativity but did not place the principle of equivalence by name in the foundations of general relativity: Charon (1963, Leçon 8), Atwater (1974). Mach's principle is mentioned by Lawden (1962, p 133).

Work on general relativity in this period also gave rise to a variant form of the principle of general covariance. Weinberg (1972, pp 91–2) defined his principle of general covariance as:

> It states that a physical equation holds in a general gravitational field, if two conditions are met:
>
> 1. The equation holds in the absence of gravitation; that is, it agrees with the laws of special relativity when the metric field $g_{\alpha\beta}$ equals the Minkowski tensor $\eta_{\alpha\beta}$ and when the affine connection $\Gamma^{\alpha}_{\beta\gamma}$ vanishes.
>
> 2. The equation is generally covariant; that is, it preserves its form under a general coordinate transformation $x \rightarrow x'$.

The novelty, of course, is that the second condition alone is usually taken as the principle of general covariance, whereas the first looks like of form of the infinitesimal principle of equivalence. Indeed Weinberg presents the principle as an alternate form of the infinitesimal principle of equivalence and shows how it follows from the principle of equivalence. He insists that it is *not* a relativity principle like the Lorentz invariance of special relativity. Bose (1980, ch 1) locates the foundations of general relativity in a local principle of equivalence and its re-expression in a two condition principle of general covariance equivalent to Weinberg's. Similarly Foster and Nightingale (1979, pp xi–xiii) locate the foundations of general relativity in an infinitesimal principle of equivalence and a version of the principle of general covariance essentially the same as Weinberg's. They strengthen Weinberg's condition 2. to read

> [2'.] the equation is a tensor equation (i.e. it preserves its form under general coordinate transformation).

The strengthening lies in the fact that not only tensor equations are covariant under arbitrary coordinate transformations. See also Treder *et al* (1980).

4.5. Recent years since 1980

The years since 1980 have seen no resolution of the disagreements over the foundations of general relativity. As we shall see later, the literatures that reject Einstein's account or seek major repairs continue to flourish. At the same time, a significant literature retains a viewpiont almost as close to Einstein's as the favourable reception of the 1920s. Broadly, in this latter literature, the foundations of general relativity are still located within an infinitesimal principle of equivalence and a principle of general covariance.

Two accounts offer essentially Weinberg's view. Both Straumann (1984, ch 2) and Kenyon (1990, ch 1) base general relativity on an infinitesimal principle of equivalence. (Kenyon discusses both Dicke's weak and strong version, with the latter amounting to an infinitesimal principle.) Kenyon (1990, section 6.4) gives a formulation of the principle of general covariance which is essentially Weinberg's as strengthened by Foster and Nightingale (see above). Without explicitly introducing the name, principle of general covariance, Straumann (1984, section 1.3) provides two requirements which are 'a mathematical formulation of the principle of equivalence'. The first is actually the principle of minimal coupling, a version of the principle of equivalence (Trautman 1965, Anderson 1967, p337, Anderson and Gautreau 1969). The second requirement is essentially Weinberg's version of the principle of general covariance.

De Felice and Clarke (1990, pp 7–13) locate the foundations of general relativity in the familiar infinitesimal principle of equivalence and principle of general covariance. Carmeli (1982, section 1.4, 1.5) locates the foundation of the theory in these same two principles. He does, however, delineate three versions of the principle of general covariance which, he notes, are 'not quite equivalent'.

1. All coordinate systems are equally good for stating the laws of physics. Hence all coordinate systems should be treated on the same footing, too.
2. The equations that describe the laws of physics should have tensorial forms and be expressed in a four-dimensional Riemannian spacetime.
3. The equations describing the laws of physics should have the same form in all coordinate systems.

Ellis and Williams (1988, section 5.2) locate the foundations of the theory in an infinitesimal principle of equivalence and what they call an extension of the principle of relativity: 'the laws of physics are the same for all observers, no matter what their state of motion'. The term principle of general covariance is not mentioned. Sexl and Urbantke (1983) treat all three of Einstein's principles of 1918. The principle of equivalence (section 1.2) is given most emphasis, although in its infinitesimal form. Mach's principle and the principle of general covariance are mentioned only apparently for historical interest (section 4.5), with the latter offered as Einstein's attempt to satisfy the former.

Finally, d'Iverno (1992, ch 9), in a chapter entitled 'The Principles of General Relativity', acknowledges that these principles have been a source of much controversy. However, as principles fundamental to general relativity or at least serious candidates for them, he presents Einstein's three principles of 1918, the Anderson and Gautreau principle of minimal coupling and a principle of correspondence (with Newtonian gravitation theory and special relativity in the limiting cases). The infinitesimal principle of equivalence is presented as the 'key principle'. Mach's principle is given three formulations, all closely connected with Einstein's cosmological ideas of 1917 and 1918. d'Inverno finds the 'full import' of the principle of general relativity ('All observers are equivalent') contained in the principle of general covariance ('the equations of physics should have tensorial form'). And, the hole argument, which figured so prominently in Einstein's early thinking about general covariance, is discussed in section 13.6. To my knowledge, this is the first time the hole argument has been discussed in a general relativity text in over half a century. The hole argument has also recently reappeared in the physics journal literature. See, for example, Rovelli (1991).

5. Is general covariance physically vacuous?

5.1. Kretschmann's objection: the point-coincidence argument turned against Einstein

In the tradition that is skeptical of Einstein's account of the foundations of general relativity, the best known of all objections is due to Kretschmann (1917, pp 575–6). He began his paper with the remarks†.

> The forms in which different authors have expressed the postulate of the Lorentz–Einstein theory of relativity—and especially the forms in which Einstein has recently expressed his postulate of general relativity—admit the following interpretation (in the case of Einstein, it is required explicitly): A system of physical laws satisfies a relativity postulate if the equations by means of which it is represented are covariant with respect to the group of spatio-temporal coordinate transformations associated with that postulate. If one accepts this interpretation and recalls that, in the final analysis, all physical observations consist in the determination of purely topological relations ('coincidences') between objects of spatio-temporal perception, from which it follows that no coordinate system is privileged by these observations, then one is forced to the following conclusion: By means of a purely mathematical reformulation of the equations representing the theory, and with, at most, mathematical complications connected with that reformulation, any physical theory can be brought into agreement with any, arbitrary relativity postulate, even the most general one, and this without modifying any of its content that can be tested by observation.

Kretschmann's point is that there must be something more to a relativity principle than covariance. For he argues that we can take any theory and reformulate it so that it is covariant under any group of transformations we pick; the problem is not physical, it is merely a challenge to our mathematical ingenuity. In brief, general covariance is physically vacuous.

This, at least, is how Kretschmann's point has been understood almost universally and it is almost what he actually argued. His real objection was a little more subtle. It depended on a non-trivial assumption that virtually all later commentators fail to report‡

> All physical observations consist in the determination of purely topological relations ('coincidences') between objects and spatio-temporal perception.

This assumption is clearly recognizable to us as the basic premise of Einstein's own point coincidence argument (see section 3.5 above). There can be no question of the importance of this assumption to Kretschmann's point, even though it is buried in the grammar of his statement. A little later, he repeats it (p 579).

> . . . according to the investigations of Ricci and Levi-Civita [1901] it may scarcely be doubted that one can bring any physical system of equations into a generally covariant form without alteration of its observationally testable content. This is obvious from the beginning, if one once again recalls that strictly only purely topological facts of natural phenomena or, according to Einstein, coincides are observable.

Thus, allowing that Kretschmann's mention of 'topological facts' alludes to his own version of the point-coincidence argument (see Howard and Norton, forthcoming), we find that Kretschmann's real objection is this: *if we accept the point coincidence argument, then any theory can be given a formulation of arbitrary covariance.*

This is a most striking reversal of fortunes. The point-coincidence argument had been Einstein's salvation from the hole argument and permitted his return to general covariance. However, in advocating the point-coincidence argument, Einstein had in effect already agreed to virtually everything in Kretschmann's objection. To establish the admissibility of

† I have suppressed Kretschmann's footnotes in this passage to other literature. For further discussion see Norton (1922, section 8). See also Howard and Norton (forthcoming) for speculation that these footnotes direct readers to Einstein's unacknowledged source for his point coincidence argument, Kretschmann (1915)!

‡ I cannot resist speculating that this misreading is at least in part due to the bewildering complexity of his German prose, which has been disentangled considerably in the above translation. This translation also slightly corrects the translation of Norton (1992, section 8.1).

general covariance for his own theory, Einstein had allowed that the physical content of a theory resides solely in the observable coincidences it sanctions. Since these coincidences are preserved under arbitrary coordinate transformation, the physical content of a theory is unaffected by the adoption of a generally covariant formulation. What Kretschmann noticed was that this argument depended on nothing peculiar to general relativity, so it could equally be used to establish the admissibility of a generally covariant formulation of any theory. Again it did not depend on the fact that the covariance group was the general group, so the same argument established the admissibility of formulations of any theory of arbitrary covariance.

5.2. *Einstein's reply*

Einstein (1918) responded to Kretschmann's objection. Having laid out the three principles upon which he believed general relativity to be based, he turned to Kretschmann's objection, which he restated correctly with its now lost premise (p 242):

> Concernign (a) [principle of relativity], Herr Kretschmann observes that a principle of relativity, formulated in this way, makes no assertions over physical reality, i.e. over the *content* of the laws of nature; rather, it is only a requirement on their mathematical *formulation*. That is, since all physical experience relates only to coincidences, it must always be possible to represent experiences of the lawful connections of these coincidences by generally covariant equations. Therefore he believes it necessary to connect another meaning with the requirement of relativity.

Einstein had little choice but to accept Kretschmann's point. The alternative was to renounce the point coincidence argument that he had advertised so widely. However he tried to salvage something of the special connection between general covariance and general relativity in the heuristics of theory choice. He continued:

> I believe Herr Kretschmann's argument to be correct, but the innovation proposed by him not to be commendable. That is, if it is correct that one can bring any empirical law into generally covariant form, the principle (a) still possesses a significant heuristic force, which has already proved itself brilliantly in the problem of gravitation and rests on the following. Of two theoretical systems compatible with experience, the one is to be preferred that is the simpler and more transparent from the standpoint of the absolute differential calculus. Let one bring Newtonian gravitational mechanics into the form of absolutely covariant equations (four-dimensional) and one will certainly be convinced that principle (a) excludes this theory, not theoretically, but practically!

Thus Einstein seems to accept Kretschmann's objection, begrudgingly, with a qualification on the role of general covariance in theory choice and with the reservation that general covariance in all theories would be impractical. Indeed it is ironic that the version of the principle of relativity given in this same paper by Einstein (quoted in section 3.7 above) essentially just restates Kretschmann's point†.

Whatever concession Einstein made to Kretschmann seems to have had a lesser effect on Einstein's later writings. He does occasionally allow that general covariance is 'more characteristic of the mathematical form of this theory [of general relativity] than its physical content' (1924, p90–1). Or that the 'requirement [of general covariance] (combined with

† The only difference is that Kretschmann allows the point coincidence argument to justify a formulation of any convariance, whereas Einstein sees it forcing a generally covariant formulation as the 'unique, natural expression' of the theory. Presumably this is because a generally covariant formulation adds the least to the catalog of coincidences. See Einstein to Besso, January 3, 1916, as quoted in Norton (1992, p 298).

that of the greatest possible logical simplicity of the laws) limits the natural laws concerned incomparably more strongly than the special principle of relativity' (1952, p153). The heuristic role of simplicity in connection with general covariance was emphasized in his *Autobiographical Notes* (1949, p65). But this emphasis seemed to be forgotten by p73, where he recalled: 'We have already given *physical* reasons for the fact that in physics invariance under the wider [general] group has to be required'. (Einstein's emphasis) More commonly, however, the qualification over simplicity is simply not mentioned. It does not appear at the relevant point in his text, Einstein (1922a, p61). Again, Einstein (1950, p352) insists, without explicit mention of simplicity considerations that

> . . . the principle of general relativity imposes exceedingly strong restrictions on the theoretical possibilities. Without this restrictive principle it would be practically impossible for anybody to hit on the gravitational equations . . .

How can we reconcile Einstein's concession to Kretschmann and his continuing emphasis on the importance of general covariance? The answer may well lie in Einstein's famous proclamation of his 1933 Herbert Spencer lecture, which revealed a metaphysics not present explicitly in Einstein's writings of 1918:

> Our experience hitherto justifies us in believing that nature is the realization of the simplest conceivable mathematical ideas. I am convinced that we can discover by means of purely mathematical constructions the concepts and laws connecting them with each other, which furnish the key to the understanding of natural phenomena . . . the creative principle resides in mathematics.

When Einstein replied to Kretschmann that one ought to pick of two empirically viable systems the simpler and more transparent within the absolute differential calculus, he may have been urging something more than merely a matter of practical convenience. It is not just that the simpler is more convenient, so that generally covariant formulations of Newtonian gravitational mechanics are (he believed) practical impossibilities. We can recognize the truth of a theory in its mathematical simplicity. And instead of being physically vacuous, general covariance is the right language in which to seek this simplicity. Later writers who endorsed Einstein's 1918 reply to Kretschmann may well have affirmed a more extreme metaphysics than they realized!

5.3. Generally covariant formulations of Newtonian mechanics

In 1918 Einstein sought to protect the special connection between general covariance and his general theory of relativity by issuing a challenge: find a generally covariant formulation of Newtonian gravitational mechanics. He had confidently predicted that should anyone try the result would be unworkable practically.

Einstein was shortly proved wrong. Cartan (1923) and Friedrichs (1927) found serviceable, generally covariant formulations of Newtonian gravitation theory. Einstein was right in so far as these generally covariant formulations were more complex than general relativity. However Einstein was quite wrong in predicting that such formulations would not be usable practically. Although they are not as attractive a host for routine calculation as the far simpler Galilean covariant formulation, they are of the same order of complexity as other theories routinely examined in physics. However there are certain circumstances in which their use is preferable if not mandatory. In an article comparing Newtonian and relativistic theories of gravitation, Trautman (1966, p413) pointed out such comparison can really only be effected reliably if the two theories under comparison are formulated in the same mathematical language. Otherwise it is hard to ascertain which differences are physical and which are accidents of the differences in formulation. Since general relativity is known only in a generally covariant formulation, this means we ought to compare it only with the generally covariant formulation of Newtonian theory. (For similar sentiments, see also Havas (1964, p 939) and Malament (1986, p 181).)

For this reason, a few expositions of relativity include a treatment of Newtonian spacetime theory in a generally covariant formulation, although the practice is not common. See for example Trautman (1964, ch 5) and Misner *et al* (1973, ch 12). In the philosophy space and time literature, however, the use of the general covariant formulation of Newtonian theory is becoming standard, even at the introductory level, see Earman and Friedman (1973), Earman (1974, pp 276–7), Friedman (1983, ch III), Malament (1986) and Norton (1992a).

Although both Cartan and Friedrichs were very much concerned with the relationship between their work and Einstein's general theory of relativity, it is striking that neither made the obvious point that their work had seriously weakened Einstein's 1918 reply to Kretschmann and raised very serious doubts over Einstein's claim to have generalized the principle of relativity to acceleration†. It is only later that this obvious point about generally covariant formulations of Newtonian theory is made: they provide an instantiation of Kretschmann's claim that any theory can be made generally covariant. See Havas (1964, p 939) and Misner *et al* (1973, p 302).

5.4. Automatic general covariance: coordinate free geometric formulation

It did not need the labours of Cartan and Friedrichs to show that theories other than general relativity admitted generally covariant formulations. In a sense this possibility had been known for a long time. As Painlevé pointed out as early as 1921 in his discussion of general relativity (1921, p 877), Lagrangian mechanics has always been invariant under arbitrary *spatial* transformation. Again, the moment Einstein applied the absolute differential calculus of Ricci and Levi-Civita to relativity theory in 1913, it was obvious that special relativity could be given generally covariant formulation. In this form, special relativity is simply the theory of a spacetime with line element (4), where $g_{\mu\nu}$ is symmetric with Lorentz signature and whose Riemann–Christoffel curvature tensor vanishes. That Einstein never embraced this obvious possibility suggests that his understanding of general covariance was a little more complex than the simple one supposed in Kretschmann's objection‡. Perhaps for this reason or perhaps just for its simplicity, the Lorentz covariant formulation of special relativity remains popular today. The possibility of formulating special relativity in arbitrary coordinates, however, was explicitly recognized in the literature quite early (see for example Kretschmann (1917, p 579), De Donder (1925, ch 1), Fock (1959, ch IV, p 350).

A number of commentators have observed that Ricci and Levi-Civita's calculus vindicates Kretschmann's objection in the sense that it provides the necessary mathematical apparatus for finding generally covariance formulation of 'practically *any* assumed law' (Whittaker 1951, vol II, p 159) or 'almost any law' (North 1965, p 58). This possibility has not really been exploited widely in the relativity literature until the 1960s and 1970s with the introduction of what Misner, Thorne and Wheeler (1973) label as the 'geometric' or

† Thus Hoffmann (1932, p 177) makes no mention of Cartan's and Friedrich's work when he remarks that the general principle of relativity 'holds in exactly the same words for the Newtonian theory [as for general relativity]. Rather the remark is supported merely be observing that the principle requires only that the mathematical expression of a theory be independent of the coordinate system and does not restrict the theory's content.

‡ Indeed, as he made clear through his principle of equivalence, he held that an extension of the covariance of special relativity beyond Lorentz covariance was a physical extension of the theory; his principle of equivalence tells us that extending the covariance to uniformly accelerated coordinates now allows the theory to embrace the phenomenon of gravitation in a special case.

'coordinate free' approach. This approach is based on Ricci and Levi-Civita's calculus. However, as was pointed out in section 3.2 above, the calculus was created quite explicitly as an abstract calculus, as independent as possible from geometric notions. The calculus was significantly altered to arrive at its modern geometric incarnation. It is now augmented with geometric ideas from topology. The most significant augmentations are the modern ideas of a differential manifold and of a geometric object of Veblen and Whitehead (1932), as well as an abstract, algebraic approach to vectors, tensors and the like, attributed to Cartan (Misner *et al* 1973, ch 8 and 9).

These methods became standard in the 1960s and 1970s through such expositions of relativity theory as Trautman (1965), Hawking and Ellis (1973), Misner, Thorne and Wheeler (1973), Sachs and Wu (1977). Following their methods, we would characterize special relativity as a theory of Minkowski spacetimes. That is, the theory has models

$$\langle M, g_{ab} \rangle$$

where M is a connected, four-dimensional, differentiable manifold and g_{ab} is a symmetric, second rank tensor of Lorentz signature which is flat, so that is satifies the equation

$$R^a{}_{bcd} = 0$$

where $R^a{}_{bcd}$ is the Riemann–Christoffel curvature tensor. There are obvious extensions if one wishes to include further fields, such as a Maxwell field and charge flux. Similarly, general relativity is the theory with models

$$\langle M, g_{ab}, T_{ab} \rangle$$

where now g_{ab} need not be flat. T_{ab} is the second rank, symmetric stress-energy tensor, which may be required to satisfy further 'energy conditions' (Hawking and Ellis, 1973, section 4.3). The metric tensor g_{ab} and T_{ab} are related by the gravitational field equation

$$G_{ab} = \kappa T_{ab}$$

where G_{ab} is the Einstein tensor and κ a constant.

A typical geometric formulation of Newtonian spacetime theory without absolute rest (after Malament 1986) has models

$$\langle M, t_a, h^{ab}, \nabla_a \rangle$$

The theory's temporal metric is t_a, is a smooth, non-vanishing co-vector field. The spatial metric is second rank, symmetric, smooth non-vanishing contravariant tensor, h^{ab}, which is degenerate through its signature $(0, 1, 1, 1)$. ∇_a is a smooth derivative operator, conferring affine structure on the spacetime. These structures satisfy orthogonality and compatibility conditions

$$h^{ab} t_a = 0 \qquad \nabla_a t_b = \nabla_a h^{bc} = 0$$

Many alternative, further conditions can be imposed upon this basic spacetime structure, for example, according to whether we wish to add gravitation as a distinct scalar field and leave the background spacetime flat or whether we wish to incorporate gravitation into the spacetime as curvature after the model of general relativity (see Friedman 1983, Ch. III).

These are all instances of a general, geometric formulation of spacetime theories. All such theories have models

$$\langle M, O_1, O_2, \ldots, O_n \rangle \tag{6}$$

where $O_1, O_2, \ldots, O_n$ are now just n geometric object fields subject to certain constraining field equations. Virtually all theories of space and time now given serious consideration can be formulated in this way†. Such theories are automatically generally covariant in a sense that actually follows from the *definitions* of the mathematical structures used in the formulation.

Following Standard definitions (e.g. Bishop and Goldberg 1968, ch 1, Hawking and Ellis 1973, ch 2, Torretti 1983, appendix), an n-dimensional differentiable manifold is a connected, topological space with a set of coordinate charts, such that every point of the topological space lies in the domain of a coordinate chart, which is a homeomorphism of an open set of the space with $\mathbb{R}^n$. The set of coordinate charts form a maximal or complete atlas in so far as the atlas contains every coordinate chart that can be constructed in the usual way from its coordinate charts by C^k-transformations on $\mathbb{R}^n$. k is some positive integer or, most commonly, infinity.

The next step is complicated by the vagueness of the definition of 'geometric object'. It is given by Veblen and Whitehead (1932, p46) as 'an invariant which is related to the space [under consideration]' where an invariant is 'anything which is unaltered by transformations of coordinates'.† Thus for our purposes, it is prudent to assume that our geometric object fields are like Anderson's (1967, p 15) 'local geometrical objects'. They are represented by a finite set of numbers for each point in the manifold in each coordinate charts and which transform under coordinate transformation in a way that respects transitivity, identity and inversion. These numbers are the geometric object's components in the coordinate charts. Let us say that a geometric object field O has components $O_{ik\ldots}$ where the integer valued i, k, ... represents a suitable set of index labels.

Combining, we now arrive at the sense in which any theory with models (6) is generally covariant. If N is any 'local coordinate neighbourhood' of M, an open set of N that is the domain of some coordinate chart x^i, then the restriction of the model (6) to N will be represented by

$$\langle A, (O_1)_{ik\ldots}, \ldots, (O_n)_{ik\ldots} \rangle \tag{7}$$

where A is the range of x^i and the remaining structures are the components of the objects $O_1, \ldots, O_n$ in the coordinate chart x^i. The theory is generally covariant in the sense that if (7) is a coordinate representation of the model (6), then so is any representation derivable from (7) by arbitrary C^k transformation. This is sometimes known as '*passive general covariance*'.

Put more briefly, once we have formulated a theory as having models of the form (6), then, built into the definitions of the structures used is the possibility of representing the models in coordinate systems that are related by the arbitrary transformations of Einstein's general covariance. (More precisely, they are related by C^k transformations if the manifold has a C^k maximal atlas of coordinate charts.) These coordinate representations behave exactly like the components of the generally covariant formulation of theories used by Einstein and others in the early years of general

† That is not to say that all intelligible theories of space and time must admit such a formulation. With a precise definition of geometric object in hand, it is just a matter of mathematical patience to construct a spacetime theory without such a formulation. One could begin, for example, by considering spacetimes whose event sets are very large but finite and do not admit smooth coordinate charts.

† The still vague 'related to the space' clause is an attempt to avoid the problem that '. . . strictly speaking, anything, such as a plant or an animal, which is unrelated to the space which we are talking about, is an invariant'.

relativity.

It is to this automatic general covariance that Thirring (1979, p 166) referred when he wrote

> At the time of the birth of gravitation theory, the requirement of general covariance provided some relief from the labor pains, but later on it was more often a source of confusion. The concept of a manifold incorporates it automatically when the definition uses equivalence classes of atlases, and hence only chart independent statements are regarded as meaningful. This program is by no means unique to gravitation theory—we have also followed it in classical mechanics and electrodynamics. The big difference [in general relativity] is that the metric g on M is now not determined a priori.

While the use of these geometric methods has become standard in modern work on general relativity, it should be noted that their dominance is not viewed universally with unmixed joy. Weinberg (1972, preface) notes that an emphasis on these methods tends to obscure the importance of the principle of equivalence within the theory and the natural connections to quantum theory.

Finally, there is a notion that is loosely dual to the notion of passive general covariance described above. It is the notion of '*active general covariance*'. The main mathematical difference is that the active version employs maps on the manifold M of the models (6) rather than transformations between coordinate charts. It can be defined as follows. Let h be an arbitrary diffeomorphism† from M to M. Then a theory with models of the form (6) is generally covariant in the active sense if every structure

$$\langle hM, h^*O_1, h^*O_2, \ldots, h^*O_n\rangle \tag{6'}$$

is a model whenever

$$\langle M, O_1, O_2, \ldots, O_n\rangle \tag{6}$$

is a model. In addition, it is routinely assumed that the structure (6) and (6′) represent the same physical circumstance (e.g., in the case of general relativity, see Hawking and Ellis 1973, p 56). This assumption has been called 'Leibniz equivalence' (Earman and Norton 1987).

Many theories are generally covariant in the active sense. A sufficient condition for active general covariance is that the object fields $O_1, O_2, \ldots, O_n$ that can be included in the models (6) are determined solely by tensor equations. Thus general relativity is covariant in this sense as are versions of special relativity and Newtonian spacetime theory.

Passive general covariance involves no physically contingent principles. Once models of the form of (6) are selected, passive general covariance follows as a matter of mathematical definition, no matter what the physical content of the theory. This is not the case with active general covariance/Leibniz equivalence. Structures (6) and (6′) are mathematically independent structures. That they represent the same physical circumstance is an assumption dependent on the properties of the physical circumstance and our methods of coordinating the structures to it. The differences between such pairs of structures as (6) and (6′) are generally of a nature that make it uninteresting to suppose anything other than Leibniz equivalence. However, it has been argued (Earman and

† For example, if M is a C^k manifold, then h might be any C^k diffeomorphism in the sense of Hawking and Ellis (1973, p 23).

Norton 1987, Norton 1988) that at least one doctrine, spacetime substantivalism, must deny Leibniz equivalence†.

Since the assumption of active general covariance/Leibniz equivalence is a physical assumption albeit weak, it does require physical arguments to support it. It turns out that Einstein's two celebrated arguments—the point-coincidence argument and the hole argument—can be put into modern forms that support active general covariance/Leibniz equivalence. According to the modernized point-coincidence argument, the two diffeomorphic models (6) and (6′) would agree on all observables, for all that is observable are coincidences that are preserved by the diffeomorphism. Therefore, if we deny Leibniz equivalence, we would have to insist that the two diffeomorphic models represent distinct physical circumstances, even though no possible observation could pick between them.

To construct the modernized hole argument, we consider some neighbourhood H of the manifold M in models (6) and (6′) and pick a diffeomorphism h that is the identity outside H but comes smoothly to differ from it within H. Then the two diffeomorphic models will be the same outside H but will come smoothly to differ within H. We now have a mathematical indeterminism, in the sense that the fullest specification of the model outside H will fail to determnine how it is to be extended into H according to the theory. This indeterminism is usually dismissed as a purely mathematical gauge freedom associated with active general covariance. If we deny Leibniz equivalence and insist that the two models represent distinct physical circumstnaces, then we convert this gauge freedom into a physical indeterminism. The differences between the models within H must now represent a difference of physical circumstances. Which will obtain within H cannot be determined by the fullest specification of the physical circumstances outside H, no matter how small H is in spatial and temporal extension.

For further discussion of the differences between active and passive general covariance, see Norton (1989, section 1, 2).

5.5. *Later responses to Kretschmann's objection*

Kretschmann's objections is probably the single most frequently mentioned of all objections to Einstein's views on the foundations of general relativity. As I have already indicated above, however, the objection which appears universally under Kretschmann's name in the literature is actually a considerably reduced version of what Kretschmann really said. It is commonly reported as the assertion that general covariance is physically vacuous, since it is merely a challenge to our mathematical ingenuity to bring any theory into generally covariant form. For the purposes of this section, which reviews later resposnes to the objection, I will take 'Kretschmann's objection' to be this reduced version, for that is the one that was responded to. Essentially no one other than Einstein seemed to realize that Kretschmann had based his objection on a contingent assumption, the premise of the point-coincidence argument. That assumption—that 'the laws of nature are only assertions of timespace coincidences'—is so non-trivial that Einstein actually made it the statement of his 1918 version of the principle of relativity.

In later literature, Kretschmann's objection is commonly accepted. Instances in which Kretschmann is cited by name include Havas (1964, p939), Rindler (1969, p196), Earman

† At present, however, there is no consensus in the philosophy of space and time literature over the connection between spacetime substantivalism, Leibniz equivalence and the hole argument, with virtually every conceivable position being defended. See Bartels (1993), Butterfield (1987, 1988, 1989), Earman (1989, ch 9), Norton (1992a, section 5.12), Cartwright and Hoefer (forthcoming, Maudlin (1988, 1990), Rynasiewicz (forthcoming (a), (b)), Stachel (forthcoming), Teller (forthcoming), Mundy (1992).

(1974, p 271), Friedman (1973, p 55), Ray (1987, p 70). Again Kretschmann's assertion of the physical vacuity of general covariance may be made without naming Kretschmann. Instances include: Silberstein (1922, pp 22–3), Szekeres (1955, p 212), Fock (1959, p 370, but see p xvi), Thirring (1979, p 166).

Einstein's 1918 response to Kretschmann also commands considerable assent. Einstein's response is encapsulated in the simple remark that general covariance is physically vacuous alone; however it achieves physical content and significant heuristic force when it is supplemented by the requirement that the laws of nature take simple forms. This viewpoint is advocated by: Painlevé (1921, p 877), Tolman (1934, pp 3, 166–67)†, Bridgman (1949, pp 339–40, 345), Whittaker (1951, vol. II, p 159), Weber (1961, pp 15–16), Skinner (1969, p 314), Adler, Bazin and Schiffer (1977, p 145). Ohanian (1976, pp 253–4) states Kretschmann's objection and quotes Einstein's 1918 reply at length, but he proceeds to elucidate Einstein's response in terms of the requirement of general invariance of the absolute object tradition (see section 8 below). In his 1918 reply to Kretschmann, Einstein urged the heuristic power of general covariance on the basis of his brilliant success with general relativity. d'Inverno (1992, p 131) comes closest to this viewpoint when he suggests that we cannot ignore general covariance, even if it is vacuous, precisely because it *was* of such importance to Einstein, rather than because of some as yet unrealized heuristic power. But perhaps Misner *et al* (1973, section 12.5) capture Einstein's metaphysics most clearly when they recapitulate Kretschmann's objection and retort

> But another viewpoint is cogent. It constructs a powerful sieve in the form of a slightly altered and slightly more nebulous principle: 'Nature likes theories that are simple when stated in coordinate-free, geometric language'. . . . According to this principle, Nature must love general relativity, and it must hate Newtonian theory. Of all theories ever conceived by physicists, general relativity has the simplest, most elegant geometric foundations. . . . By contrast, what diabolically clever physicist would ever foist on man a theory with such a complicated geometric foundation as Newtonian theory?

There are obvious problems with this view. To begin, it would seem that the view is plainly false. The very simplest laws, which nature ought to love the most, are just incompatible with experience. For example, it would be very simple if all of space, time and the distribution of matter were homogenous; but they are not homogeneous. So Nature's preferences can only be exercised among the more complicated dregs that remain after experience has drained off the truly simple—Nature's preference here is a rather contrived one. Next, it is not clear by what rules we are to judge which of two theories is the simpler. It cannot just be a matter of intuitive impressions, since then we have no way of adjudicating disagreements. But even a basic count of the number of mathematical structures in a theory is hard to do unambiguously‡. Or we might judge that general covariance implemented by *tensor* equations is simpler. Bondi (1959, p 108), however, endorses the view that general covariance is physically vacuous and points out that conservation laws explicitly involving gravitational energy – momentum in general relativity are not tensorial, but pseudo-tensorial. Finally, it is not obvious why nature should be so kind as to prefer laws that we humans deem simple. Thus North (1965, p 58) muses that the virtue of simplicity for covariant laws might merely be that they are more likely to be accepted by others.

My own view is that one should not look on simplicity as resulting from the emotional

† Tolman gives Kretschmann's objection in its full form insofar as the possibility of generally covariant formulation is taken to follow necessarily from the point-coincidence argument.

‡ Is the stress-energy tensor of pressureless dust, $T^{ab} = \rho U^a U^b$, counted as one structure T^{ab} or as two, the matter density ρ and the four-velocity field U^a?

attachments of Nature. Rather it arises from the labours of theorists who have constructed languages in which Nature's choices appear simple. Whether Nature's further choices will continue to appear simple in some language seems to me an entirely contingent matter and one takes a great risk elevating any language to the status of Nature's own. As we explore new domains of physical law, the one thing that is most clear is Nature's surprising versatility in frustrating our natural expectations. However this does not mean that there is no value in simplicity. Apart from its pragmatic value, it has an epistemic value. The more complicated a theory, the more likely we are to have introduced structures with no correlations in reality; and the more complicated the theory, the harder it will be to test for these physically irrelevant structures. We should prefer the simpler theory and seek languages that make our theories simple, but not because Nature is simple. Rather, if we restrict ourselves to simpler theories, we are more likely to know the truth when we find it.

There is a variation of Einstein's response to Kretschmann that avoids the difficult questions over simplicity. Its overall effect is to direct us towards simpler theories by restricting the structures we can employ in our formulations. It focuses on the process of finding generally covariant formulations of arbitrary laws. If we restrict the number of additional mathematical structures that can be introduced in this process, it may no longer be possible to construct a generall covariant formulations for some laws, so that we once again have an interesting division between generally covariant and other theories. Fock (1959, p xvi) describes the idea in its most general form

> . . . the requirement of covariance of equations has great heuristic value because it limits the variety of possible forms of equations and thereby makes it easier to choose the correct ones. However, one should stress that the equations can so be limited only under the necessary condition that the number of functions introduced is also limited; if one permits the introduction of an arbitrary number of new auxiliary functions, practically any equation can be given covariant form.

Trautman (1964, pp 122–3) illustrates how unrestricted admission of new structures allows construction of a generally covariant formulation of equations that clearly are coordinate dependent. He considers the equation

$$A_1 = 0$$

the vanishing of the first component of a covector A_a in some coordinate system. If u^a is the coordinate basis vector field associated with the x^1 coordinate, then this law admits generally covariant formulation as

$$u^a A_a = 0$$

The villain is the vector field u^a, since (p 123)

> one should not introduce such additional structures in addition to those already present in the axioms of the theory (e.g. the metric tensor, affine connection) and to those that are necessary to describe the physical system.

If we now apply this thinking to general relativity, we arrive at a popular means of injecting content into the general covariance of general relativity. In a Lorentz covariant version of special relativity, the metrical properties of spacetime are not represented explicitly. In the transition to the generally covariant, general theory of relativity, these properties become explicit as a new structure, the metric tensor g_{ab}. It is required that this new structure represent some definite physical element of reality and not just be a mathematical contrivance introduced to force through general covariance. The metric tensor satisfies this requirement in so far as it represents the gravitational field as well as the metrical properties of spacetime. Pauli (1921, p 150) describes this outcome

> . . . Kretschmann . . . took the view that the postulate of general covariance does not

> make any assertions about the physical *content* of the physical laws, but only about their mathematical *formulation*; and Einstein . . . entirely concurred with this view. The generally covariant formulation of the physical laws acquires a physical content only through the principle of equivalence, in consequence of which gravitation is described *solely* by the g_{ik} and these latter are not given independently from matter, but are themselves determined by field equations.

We find a similar view in Borel (1926, pp 172–3), Weyl (1921, pp 226–7), Reichenbach (1924, p 141), Anderson (1967, 1971—see section 8.1 below), Graves (1971, p 138) and even as recently as Wald (1984, p 57) who formulates the principle of general covariance as

> The principle of general covariance in this context [pre-relativistic and relativistic physics] states that the metric of space is the only quantity pertaining to space that can appear in the laws of physics. Specifically there are no preferred vector fields or preferred bases of vector fields pertaining only to the structure of space which appear in any law of physics.

(He cautions that 'the phrase "pertaining to space" does not have a precise meaning'.)

Both Pauli and Weyl stress a special aspect of the physical character of the metric in their discussions: the metric is not given *a priori* but is influenced or determined by the matter distribution via invariant field equations. This would, of course, rule out generally covariant formulations of special relativity. Weyl, in particular, sees this as the decisive property of general relativity. 'Only this fact justifies us in assigning the name "general theory of relativity" to our reasoning . . .' he wrote (p 226). Further, he emphasized the result that 'gravitation is a mode of expression of the metrical field' and that 'this assumption, rather than the postulate of general invariance, seems to the author to be the real pivot of the general theory of relativity' (pp 226–7). We shall see that this theme will be incorporated into the absolute object approach (see section 8 below).

A practical difficulty still remains. At the most fundamental level, the general principle is clearly correct: we should deny admission to theories or structures that do not represent elements of reality. The hope is that this restriction will preserve a unique association between general covariance and the general theory of relativity. However the principle may well not be sufficiently precisely formulated to have any force in realistic examples. Consider the structures dt_a, h^{ab} and ∇_a, introduced in constructing a generally covariant formulation of Newtonian theory. Are they admissible or not? Notice that Pauli and Weyl's emphsis on the dynamic character of the metric may not help us here. In versions of Newtonian gravitation theory, the gravitational field is incorporated into the affine structure ∇_a which then has similar dynamical properties to the metric of general relativity.

The strategy so far has been to augment the requirement of general covariance with additional requirements that make it non-trivial. It turns out that there is an extremely simple way of augmenting the principle of general covariance so that we cannot render generally covariant such theories as special relativity and versions of Newtonian theory that do not incorporate the gravitational field into affine structure. In both these cases, the associated generally covariant formulations have the property that they can be simplified by reintroducing restricted coordinate systems. This is not so in the case of general relativity, so we can pick between these cases by insisting that the generally covariant formulation not admit simplification. Bergmann (1942, p 159) explicitly incorporates this requirement into the statement of the principle of general covariance:

> The hypothesis that the geometry of physical space is represented best by a formalism which is covariant with respect to general coordinate transformations, and that a restriction to a less general group of transformations would not simplify that formalism, is called *the principle of general covariance*.

At first this seems like a purely *ad hoc* contrivance. However Bergmann's proposal connects directly with the view that relativity principles are geometric symmetry principles, as we shall see in section 6.2 below. Alternately, Bondi (1959, p 108) calls the proposal into question by recalling Foek's use of harmonic coordinates to reduce the covariance of general relativity (see section 9 below).

There have been other studies of the relationship between a theory and its generally covariant reformulation and these studies arrive at conclusions uncomfortable for Kretschmann's objection. Scheibe (1991, 1981) has considered the relationship within a more precise formal setting. He concludes that it is simply not obvious that any geometry of restricted covariance can always be recast in a generally covariant formulation. Post (1967) concludes that the process of rendering theories generally covariant is far from automatic triviality and must be treated with some care. In the case of electromagnetic theory, he shows how different ways of rendering the theory generally covariant actually lead to distinct theories. Mashoon (1986) similarly emphasizes that, while any theory can be rendered generally covariant, the manner in which it is done can have physical consequences, in particular, in the measurements of accelerated observers.

Many authors are prepared to accept Kretschmann's objection but feel that it has to be qualified in significant ways if the true significance of general covariance is to be appreciated. While accepting Kretschmann's objection and that a requirement of general covariance is not a relativity principle like that of special relativity, Weinberg (1972, pp 92, 111–3) characterizes general covariance as akin to the gauge invariance of electromagnetic fields. Accepting Kretschmann's objection, Bunge (1967, section 3.1.3) observes that if general covariance is understood as simply requiring form invariance of laws, then it does become a purely mathematical requirement. Therefore he concludes that general covariance is to be understood as a regulative rather than constitutive principle. Mavridès (1973, p 66) also accepts Kretschmann's objection but sees the significance of the principle in absorption of acceleration into the non-Euclidean structure of spacetime.

Zahar (1989, section 8.1) approaches the problem with a distinction introduced by logicians between an object language and its metalanguage. In this context, the object language contains the assertions about physics systems and the metalanguage contains assertions about the object language. Whether a body of object language assertions, such as Newtonian theory, is generally covariant is not itself an object language assertion. It belongs to the metalanguage. We may be able to find a generally covariant formulation of Newtonian theory which is logically equivalent to the original Galilean covariant version. However the meta-level property of general covariance is not inherited by the original formulation, for meta-level properties are not transmitted by logical equivalence. Therefore we cannot say that Newtonian theory itself is generally covariant. Several other authors have approached general covariance as a principle of operating a meta-level of language. See Graves (1971, pp 143–7). In particular, Törnebohm (1952, section 41) characterizes the principle of general covariance as a closure rule operating on a meta-level in which one quantifies over coordinate systems employed in statements of physical laws.

Finally, see Kuchar (1988) for a reincarnation of the issues raised by the debate of Kretschmann's objection in Hamiltonian dynamics and canonical quantization of generally covariant systems.

6. Is the requirement of general covariance a relativity principle?

6.1. Disanalogies with the principle of relativity of special relativity

In addition to accusations that his principle of general covariance is physically vacuous, Einstein's treatment of general covariance has been besieged by continuing complaints that the achievement of general covariance does not amount to a generalization of the principle of relativity to acceleration. These complaints have come in many different forms. Some of the earliest make the obvious point that such an extension of the principle of relativity to accelerated motion seems to be flatly contradicted by the simplest observations. The principle of relativity of inertial motion fits the experiences of a traveler in a train moving uniformly on smooth tracks; nothing within the carriage reveals the train's motion. However, the same is not so if the train accelerates, as was pointed out acerbically by Lenard (1921, p 15), whose involvement in the persecution of Einstein in Germany in the 1920s is well known:

> *Let the train in consideration undertake a distinct, non-uniform motion.* ... If, as a result, everything in the train is wrecked through the effects of inertia, while outside everything remains undamaged, then, I believe, no sound mind would want to draw any other conclusion than that the train had altered its motion with a jolt and not the *surroundings*.

For Einstein's reply to this exact passage, see Einstein (1918a).

It was only in the 1950s and 1960s that such long-standing worries took a prominent though still disputed place in the mainstream literature. This dissident view drew strength from such eminent relativists as Fock and Synge, who dared to proclaim what few would admit: they just could not see how Einstein's theory generalizes the principle of relativity—and they even suspected that Einstein could not see it either. So Synge (1966, p 7) wrote:

> ... the general theory of relativity. The name is repellent. Relativity? I have never been able to understand what that word means in this connection. I used to think that this was my fault, some flaw in my intelligence, but it is now apparent that nobody ever understood it, probably not even Einstein himself. So let it go. What is before us is Einstein's theory of gravitation.

See also Synge (1964, p 3) and (1960, p ix), where he wrote

> ... the geometric way of looking at space-time comes directly from Minkowski. He protested against the use of the word 'relativity' to describe a theory based on an 'absolute' (spacetime), and, had he lived to see the general theory of relativity, I believe he would have repeated his protest in even stronger terms.

In similar vein, Fock (1959, pp xvi–xviii, 367–8, 375–6) treated a relativity principle as stating a uniformity of spacetime. Thus special relativity admits a relativity principle because of the uniformity of a Minkowski spacetime. The spacetimes of general relativity, however, manifest this uniformity only in the infinitesimal, so that the naming of the theory 'general relativity' or 'general theory of relativity' is simply incorrect, betraying Einstein's failure to understand his own theory. Fock continued (p 368)

> The fact that the theory of gravitation, a theory of such amazing depth, beauty and cogency, was not correctly understood by its author, should not surprise us. We should also not be surprised at the gaps in logic, and even errors, which the author permitted himself when he derived the basic equations of the theory. In the history of physics we have many examples in which the underlying significance of a fundamentally new physical theory was realized not by its author but by somebody else and in which the derivation of the basic equations proposed by the author proved to be logically *inconsistent. It is sufficient to point to* Maxwell's *theory of electromagnetic field* ...

Allowing in addition that the only admissible sense of 'general relativity' is as the purely formal property of general covariance, Fock (1974, p 5) concluded

> Thus we can sum up: general relativity can not be physical, and physical relativity cannot be general.

These confessions were engagingly candid and their iconoclastic sentiments found receptive audiences. The heresy of disbelief in Einstein became respectable.

Fock and Synge are, of course, not alone in divorcing general covariance from a generalization of the principle of relativity and announcing the failure of Einstein's effort in this regard. See for example Landau and Lifshitz (1951, p 229), Davis (1970, p 219), Raine and Heller (1981, p 135) and Bondi (1979, p 129).

6.2. Relativity principles as symmetry principles

If covariance principles are not relativity principles, then what are relativity principles? New answers to this question have come repeatedly within the tradition that proposes the divorce of general covariance from a generalization of principle of relativity. We shall see that they eventually stabilize on the view that a relativity principle expresses a symmetry of the spacetime structure.

One of the earliest proposals comes from Kretschmann. His famous objection to general covariance actually only occupies a small part of his lengthy paper (1917). The bulk of it is devoted to developing an alternate interpretation of relativity principles. HIs proposals are embedded within extended calculations and circuitous verbiage. They appear to reduce to the following. The key idea in identifying the relativity principle of some given theory lies not in extending its covariance, but in reducing it to the minimum group possible. This reduction must be done in a way that identifies a group associated with the theory's physical content rather than some particular formulation of it.

In the case of special relativity, his general proposal leads to the expected result: the Lorentz group expresses the theory's relativity principle. Consider the bundle of all light-like world lines in the theory. In the Lorentz covariant formulation, this bundle is described by the equation

$$(x_1 - x_1^0)^2 + \ldots + (x_4 - x_4^0)^2 = 0 \quad (8)$$

where $x_1 = x, \ldots, x_4 = ict$ are the usual spacetime coordinates and $(x_1^0, \ldots, x_4^0)$ some arbitrary origin event. This bundle is mapped back into itself by any Lorentz transformation that preserves the origin. Kretschmann allowed that we could extend the usual Lorentz covariant formulation of the theory even as far as a generally covariant formulation, using the methods of Ricci and Levi-Civita. However, in a formulation of extended covariance, an allowed transformation will, in general, not map this bundle back into itself. Rather, such a transformation will alter the coordinate image of the bundle. Again, one could consider a formulation whose covariance is restricted to a group smaller than the Lorentz group. However this formulation could only be constructed at the expense of altering the physical content of the theory†. The Lorentz transformation is the formulation of minimal covariance faithful to the theory's physical content. Therefore the Lorentz transformation is the group associated with the theory's relativity principle.

A similar analysis in the case of general relativity leads to a quite different result. In effect Kretschmann finds that the single membered identity group plays the same role in general relativity as does the Lorentz group in special relativity. As a result, he can arrive at a conclusion that directly contradicts Einstein's (p 610)

† How Kretschmann arrived at this crucial conclusion is a little unclear to me. Such a formulation would need to replace (8) by another formula or formulae of more restricted covariance and presumably Kretschmann held that any such formulae would have to alter the physical content of (8). For example, to violate Lorentz covariance, the new formula might pick out one or other spatial direction as preferred, whereas equation (8) describing the bundle admits no such preferred directions.

> Therefore Einstein's theory satisfies no relativity principle at all in the sense developed [earlier in the paper]; on the basis of its content, it is a completely absolute theory.

To arrive at this result, Kretschmann considered the bundle of light-like worldlines and of free material particles within the theory. He found the former fixed the components of the metric tensor $g_{\mu\nu}$ up to a multiplicative factor λ and the latter forced λ to be a constant. (Notice that these are now familiar results. In modern language: conformal structure fixes the metric up to a conformal factor and specifying affine structure forces the factor to be constant.) Finally consideration of spacetime curvature rules out any value of λ other than unity. Thus the physical content of the theory fixes the metrical components. But once these components are fixed, the coordinate system is fixed and no covariance transformation remains; in effect the covariance group has become the identity group and one has no relativity principle. Kretschmann also showed that the same result could be arrived at in another way. As long as the spacetime metric is sufficiently non-uniform, it is possible to define a unique spacetime coordinate system for each metric by setting the four coordinates equal to unique curvature invariants. This once again reduces the covariance group to the identity.

Finally Kretschmann could extract one final blow from his calculations. In effect he could conclude that the Lorentz group was the largest group possible for any relativity principle in a spacetime theory of the type of special and general relativity (p 610):

> A physical theory, which accords an observationally accessible meaning to the external principle
>
> $$\left[\delta \int \mathrm{d}s = 0 \text{ where } \mathrm{d}s^2 = g_{\mu\nu}\mathrm{d}x_\mu \mathrm{d}x_\nu\right]$$
>
> of a space-time manifold with Minkowski normal form of the line element or posits that the invariant metrical character of the manifold is in some other way in principle observable to the same extent, can satisfy no broader relativity postulate in the sense [developed earlier in the paper] than that of the original Einsteinian theory of relativity.

Kretschmann's proposal has been criticized at length by Anderson (1966). He argues that the proposal fails since one can too readily reduce the covariance of a theory to the identity. His examples include electrodynamics and special relativity, provided that we add some other structure, such as a scalar field, to the Minkowski spacetime.

Cartan (1927) gave a less bellicose and mathematically more perspicacious characterization of the difference between the general covariance of general relativity and the Lorentz covariance of special relativity.

> General relativity threw into physics and philosophy the antagonism that existed between the two principle directors of geometry, Riemann and Klein. The space-times of classical mechanics and or special relativity are of the type of Klein, those of general relativity are of the type of Riemann.

Under Klein's Erlangen program a wide range of geometries were all characterized by their associated groups and the geometric entities they studied were the invariants of those groups. The key aspect of these Erlangen program geometries—the Euclidean, the projective, the affine, the conformal and others—was that all the spaces were homogeneous. In the Riemann tradition, one considered a space and a group of transformations. But the geometric entities investigated are no longer the invariants of the transformations, for in this case there are essentially none. Instead one is interested in the invariants of a quadratic differential form, the fundamental or metrical form, that is adjoined to the space. As result, the groups associated with geometries in the two traditions have very different significance. The spacetime geometry of special relativity, as introduced by Minkowski, is in the tradition of Klein. As a result its characteristic group, the Lorentz group, is associated with the homogeneity of the spacetime. General relativity lies in the Riemann tradition and, as a result, its general group of transformations is associated with no such homogeneity.

Sesmat (1937, pp 382–3) gave a more algebraic characterization of why he felt the general covariance of general relativity had failed to implement a generalization of the principle of relativity. What was needed was a theory whose laws would remain unchanged in form under transformations between all frames of reference including accelerated ones, in the same way that the laws of special relativity remained invariant under Lorentz transformation. The general covariance of general relativity just did not do this. Under the transformations of general covariance, such as a transformation between Cartesian and polar coordinates, the expressions for basic tensors do change. What general covariance does allow, however, is that a tensor, such as the Einstein tensor, can retain its zero value in empty space under these transformations, even though its expression changes.

Sesmat's point seems to be precisely the point that Weinberg (1972, p 92) is making when he explains the difference between the Lorentz invariance of special relativity and general covariance. One could, he notes, expand the covariance of Newton's second law by transforming it under Lorentz transformation. However, a new quantity, the velocity of the coordinate frame would appear in the transformed equation.

> The requirement that this velocity does *not* appear in the transformed equation is what we call the Principle of Special Relativity, or 'Lorentz invariance' for short, and this requirement places very powerful restrictions on the original equation. Similarly, when we make an equation generally covariant, new ingredients will enter, that is, the metric tensor $g_{\mu\nu}$ and the affine connection $\Gamma^{\lambda}_{\nu\mu}$. The difference is that we do not require that these quantities drop out at the end, and hence we do not obtain any restrictions on the equations we start with; rather, we exploit the presence of $g_{\mu\nu}$ and $\Gamma^{\lambda}_{\mu\nu}$ to represent gravitational fields.

Fock (1957) (see also Fock 1959, p xiii–xiv, 166) gave a synthesis of all these ideas: the homogeneity of spaces in the Klein tradition, the mapping back into themselves of Kretschmann's bundle of lightlike and inertial worldlines and he gave it in an algebraic form indicated by Sesmat and Weinberg. In considering the uniform or homogeneous spacetime of special relativity, he explained (p 325):

> The property of spacetime being homogeneous means that (a) there are no privileged points in space and in time; (b) there are no privileged directions, and (c) there are no privileged inertial frames (that all frames moving uniformly and in a straight line with respect to one another are on the same footing).
>
> The uniformity of space and time manifests itself in the existence of the Lorentz group. In particular, the equality of points in space and time corresponds to the possibility of a displacement, the equality of directions corresponds to that of spatial rotations, and the equality of inertial frames corresponds to a special Lorentz transformation.

Fock then gave this condition mathematical expression. The Lorentz transformation leaves unchanged the Minkowski line element

$$ds^2 = dx_0^2 - dx_1^2 - dx_2^2 - dx_3^2 = \eta_{\mu\nu} dx_\mu dx_\nu \tag{9}$$

where the $x_0, \ldots, x_3$ are the usual spacetime coordinates of the Lorentz covariant formulation. This same condition can be stated in arbitrary coordinates in which the line element (9) becomes

$$ds^2 = g_{\mu\nu}\,dx_\mu\,dx_\mu$$

The mathematical expression of the homogeneity of the Minkowski spacetime is now stated as the preservation of the functional form of the components of the metric in some class of coordinate systems. That is, if the metric has components $g_{\mu\nu}$ in some arbitrary coordinate system x_σ, then it will be possible to transform to a new coordinate system x'_σ in which the new components of the metric $g'_{\mu\nu}$ are the same functions of x'_σ as the $g_{\mu\nu}$ are of the x_σ. That is,

$$g'_{\mu\nu}(x'_\sigma) = g_{\mu\nu}(x_\sigma) \tag{10}$$

where the equality must be read as holding for equal numerical values of the quadruples x_σ and x'_σ. This condition is considerably more restrictive than merely requiring that the components $g_{\mu\nu}$ transform into $g'_{\mu\nu}$ under the usual tensor transformation rule. And it expresses a homogeneity of the spacetime since both coordinate systems x_σ and x'_σ relate in indistinguishable fashion to the metric tensor. The set of coordinate systems with this property are related by a ten parameter group which corresponds to the Lorentz group.

Notice that the algebraic expression for the transformations from x_σ to x'_σ in the Lorentz group can no longer be the familiar formulae (1) of Einstein's original 1905 paper. For example, in generalizing the coordinates, the coordinate system of (9) may remain inertial but with the Cartesian spatial coordinates replaced by polar coordinates, in which case the expression for the Lorentz transformation would have to be altered correspondingly. However, whatever may be their altered form, the transformation equations must leave unchanged the functional form of components of the metric tensor. Otherwise the spacetime would distinguish between two inertial coordinate systems, in violation of its uniformity. That is the condition expressed in (10).

The distinction between simple covariance and transformations of form (10) seems to be distinction between Buchdahl's (1981, p 29) 'improper' and 'proper form invariance'. In his example, the equation $g^{ij}S_{,i}S_{,j} = 0$ (where S is a scalar field and commas denote differentiation) is improperly form invariant if the transformed equation just retains this form as, say, $g^{i'j'}S_{,i'}S_{,j'} = 0$. It is properly form invariant if the $g^{i'j'}$ of the transformed equation also remain the same functions of the new coordinates as the untransformed g^{ij} were of the old.

Fock's proposal now relates directly to Bergmann's (1942, p 159) statement of the principle of general covariance as given in section 5.5 above. According to (10), a generally covariant formulation of special relativity will admit a ten parameter subgroup of transformation—the Lorentz transformation—that preserves the functional form of the components of the metric tensor $g_{\mu\nu}$. It can do so in many different ways. One merely selects some arbitary coordinate system in which the Minkowski metric has components $g_{\mu\nu}$ and allows condition (10) to generate the subgroup. If one begins with the usual diagonal form of the metric, $\eta_{\mu\nu}$, one arrives at the usual form of the Lorentz transformation (1). Each of these subgroups is associated with a formulation of special relativity of reduced covariance and the particular functional form of the metrical components that remain unaltered according to (10) will be built into its laws. Therefore Bergmann's statement of the principle of general covariance will judge the generally covariant formulation of special relativity to be inadmissible and thus preserves a distinction between the covariance of general relativity and of special relativity.

Notice also that the formulations of special relativity of reduced covariance are now of a form compatible with Klein's Erlangen program, since the Riemannian quadratic differential form are no longer transformed merely covariantly within the theory. Thus, in accord with Cartan's observations, the transformation groups of the formulations are now associated with the homogeneity of the spacetime.

Fock's condition (10) has an immediate expression in the geometric approach to spacetime theories. Let h be the dual manifold diffeomorphism of the coordinate transformation defined on a Minkowski $\langle M, g_{ab}\rangle$. Then Fock's condition (10) becomes

$$h^* g_{ab} = g_{ab} \tag{11}$$

and the group of transformations satisfying this condition is the Lorentz group†. That is, the Lorentz group is the group of diffeomorphisms that are the symmetry transformations or isometries of the Minkowski metric (Wald, 1984, pp58, 60, 438). The existence of this group expresses the uniformity of the Minkowski spacetime.

With this terminology, we can summarize why Fock and others believe that the transition from special to general relativity has failed to generalize the principle of relativity. Two groups are associated with the formulation of a theory: its covariance group characterizes purely formal aspects of its formulation; its symmetry group characterizes a physical fact, the degree of uniformity of the spacetime and this uniformity allows the theory to satisfy a relativity principle. In the transition from a Lorentz covariant formulation of special relativity to a generally covariant formulation of general relativity, the covariance group is expanded. This is, however, merely an accident of formulation. The symmetry group is actually reduced from the Lorentz group to the identity group, for the general case. The identity group is associated with no relativity principle at all. Therefore the transition from special to general relativity does not generalize the relativity principle. It eradicates it.

6.3. Coordinate systems versus frames of reference

Fock took it as immediate that his condition (10) automatically realized the equivalence of inertial frames of reference whereas general relativity embodies no such equivalence. That this is correct may not be immediately clear given that such formulations of the principle of general covariance as Bergmann's do preserve a sense in which the natural covariance of special relativity differs from that of general relativity. To give a precise statement of this result we require a clearer statement of what is a frame of reference.

In traditional developments of special and general relativity it has been customary not to distinguish between two quite distinct ideas. The first is the notion of a coordinate system, understood simply as the smooth, invertible assignment of four numbers to events in

† To see the transition, let the metric g_{ab} have components $g_{\mu\nu}$ in some coordinate system and let the transformation from coordinate systems x_σ to x'_σ satisfy condition (10). To generate the dual diffeomorphism h, we now just consider the functional relation between x_σ and x'_σ as a map from quadruples of reals x_σ to quadruples of reals $x'_\sigma(x_\sigma)$. In one of the coordinate systems allowed under (10), the diffeomorphism h maps an event p with the four coordinates x_σ to an hevent hp with coordinates $x'_\sigma(x_\sigma)$ in the same coordinate system. Consider the metric h^*g_{ab} carried along to hp from p under h. If the metric at p has components $g_{\mu\nu}$, then the carried along metric at hp will have the same components $g_{\mu\nu}$ in the *carried along* coordinate system and the carried along coordinate system will assign coordinates x_σ to hp. We now see that this carried along metric is the same as the original metric at hp, as (11) demands, by comparing their components in the original coordinate system. We transform the carried along metric back from the carried along coordinate system to the original by means of the coordinate transformation of (10) and find that the carried along metric has components $g'_{\mu\nu}$ at hp, which has coordinates x'_σ. Therefore the carried along metric agrees with the original metric since the functional forms of $g_{\mu\nu}$ and $g'_{\mu\nu}$ are the same. For further discussion of the duality of coordinate transformation and manifold diffeomorphism, see Norton (1989, section 2.3).

spacetime neighbourhoods. The second, the frame of reference, refers to an idealized physical system used to assign such numbers. More precisely, since the physical systems tend to be space filling, one is concerned with how such hypothetical systems would behave were they to be constructed. Many such systems are possible. For example one can imagine space full of similarly constituted clocks and all of them attached to a rigid frame of small rods. The clock readings give us the time coordinates and the counting of rods gives us spatial coordinates. To avoid unnecessary restrictions, we can divorce this arrangement from metrical notions. Following Kopczynski and Trautman (1992, pp 24–5), we could require only that the space-filling family of clocks bear three smoothly assigned indices (which could function as spatial coordinates), that the clocks tick smoothly, although not necessarily in proper time, and that time readings vary smoothly from clock to clock. Of special importance for our purposes is that each frame of reference has a definite state of motion at each event of spacetime.

Within the context of special relativity and as long as we restrict ourselves to frames of reference in inertial motion, then little of importance depends on the difference between an inertial frame of reference and the inertial coordinate system it induces. This comfortable circumstance ceases immediately once we begin to consider frames of reference in non-uniform motion even within special relativity. This became a major problem for Einstein to negotiate as early as 1907, when he began to consider uniformly accelerated frames of reference in his new gravitation theory. He found (1907, section 18) the need to introduce coordinate times which could not be read directly from clock measurements. Similarly, due to the Lorentz contraction of rods oriented in the direction of motion, the geometry associated with a uniformly rotating frame of reference ceased to be Euclidean. As a result, spatial coordinates can no longer be assigned by the usual methods with measuring rods. The point of Einstein's rotating disk thought experiment (first published in Einstein (1912, section 1) and best known from Einstein (1916, section 3)) is that spacetime coordinates will lose this direct metrical significance once we stray from the familiar inertial coordinate systems of special relativity†.

With the advent of general relativity, Einstein wished to consider frames of reference with arbitrary states of motion. However he deemed it impractical to retain even a vestige of the idealized physical system of the frame of reference. In their place he simply used arbitrary coordinate systems. The association of an arbitrary coordinate system with an arbitrary frame of reference became standard in the literature for many decades. Thus, for example Bergmann (1962, p 207) explains

> In all that follows we shall use the terms 'curvilinear four-dimensional coordinate system' and 'frame of reference' interchangeably.

Thus, in Einstein's writings, whatever equivalence is established by general covariance for arbitrary coordinate systems is also conferred upon arbitrary frames of reference and, if we recall the connection between a frame of reference and a state of motion, the powerful suggestion is that this is all that is needed to extend the principle of relativity to arbitrary motions. The connection is complicated slightly by the fact that some coordinate

† The problem is even more complicated than Einstein indicated. An inertial frame of reference in a Minkowski spacetime is naturally associated with Euclidean spaces, which are the spatial hypersurfaces everywhere orthogonal to the world lines of the frame's elements. The wordlines of the elements of a rotating disk admit no such orthogonal hypersurfaces. Since the spacetime of special relativity remains flat, we may well ask in what space does the geometry become non-Euclidean. The most direct answer is that this geometry is induced onto the 'relative space' formed by the worldliness of the elements of the disk. This space can be defined precisely as in Norton (1985, section 3). For further discussion of the role of the rotating disk thought experiment in Einstein's thought, see Stachel (1980a).

transformations clearly do not relate different states of motion, such as the transformation between spatial Cartesian and polar coordinates. However some subgroup of the general group of coordinate transformations is the appropriate one, as Einstein (1916, section 3) makes clear when he writes

> It is clear that a physical theory which satisfies this postulate [of general covariance] will also be suitable for the general postulate of relativity. For the sum of *all* substitutions in any case includes those which correspond to all relative motions of three-dimensional systems of co-ordinates.

More recently, to negotiate the obvious ambiguities of Einstein's treatment, the notion of frame of reference has reappeared as a structure distinct from a coordinate system. If one conceives of a frame of reference as a space filling system of hypothetical instruments moving with arbitrary velocities, then the minimum information needed to pick out the frame is the specification of the world lines of its elements. As a result, the simplest workable definition of an arbitrary frame of reference—and the one I shall use here—is that it is a congruence of curves, that is, a set of curves such that every event in the spacetime manifold lies on exactly one of its curves (Torretti 1983, p28, Norton 1985, section 3, Vladimirov *et al* 1987, p95). If the notion of timelike is defined, we would also require the curves be timelike to ensure that they are the worldlines of physical elements. In the case of the semi-Riemannian spacetimes of relativity theory, whatever further information one might need is supplied by the theory's metrical structure. From it we can read the time elapsed as read by proper clocks moving with the frame, or changes in the directions and spatial distances of neighbouring elements of the frame.

Various alternative definitions of frame of reference are possible. Since a smooth congruence of curves can be specified as the unique set of integral curves of a smooth, non-vanishing, timelike vector field, one could take a frame of reference to be such a timelike vector field (Earman 1974, p270, Jones 1981, p163). Again, one can employ richer structures. The timelike vector field could be supplemented by a triad of spacelike vectors pointing to the worldlines of neighbouring elements of the frame. A frame of reference then becomes the specification of an orthonormal tetrad of vectors over the spacetime manifold. (Synge 1960, ch III.5, Vladimirov *et al* 1987, p95). Finally a coordinate system is adapted to a frame of reference if the curves of the frame coincide with the curves of constant spatial coordinates. Therefore we could take a frame to be the equivalence class of all coordinate systems adapted to some congruence (Earman 1974, p270). This definition has the advantage of bringing us closest to the traditional correspondence between frames of reference and coordinate systems.

In special relativity, an inertial frame of reference is a congruence of timelike geodesics. An inertial coordinate system is a coordinate system adapted to an inertial frame of reference.

6.4. Relativity principles and the equivalence of frames

With the notion of frame of reference clarified, it proves possible to give a more precise treatment of the principle of relativity in so far as it asserts an equivalence of various states of motion, that is, of various frames of reference. Einstein's original treatment of the principle of relativity in special relativity amounted to requiring that the laws of physics adopt the same form when expressed in any inertial coordinate system. This type of formulation of the principle was quite serviceable in the context of a Lorentz covariant special theory of relativity. As we have seen, however, there have been significant challenges to the idea that form invariance of laws can capture any physical principle when we are prepared to employ mathematical techniques powerful enough to render virtually any theory generally covariant.

A precise formulation of the relevant notion of equivalence of frame has been developed within work that includes Earman (1974), Friedman (1983, especially ch. IV.5) and Jones (1981). Their proposals explore many variant definitions and do so within the context of a wide range of theories, including variants of Newtonian spacetime theory. The essential ideas they share can be illustrated by the following treatment of special and general relativity.

The essence of the principle of relativity in the special theory is the indistinguishability of all the inertial states of motion. Thus Einstein's 1905 special relativity paper had been motivated by the realization that no experiment in mechanics, optics or electrodynamics could reveal the uniform motion of the earth through the aether. That is, space and time 'look the same' experimentally to observers in any state of inertial motion. Einstein's task was to devise a theory in which they looked the same theoretically as well.

This condition can broken up into a kind of pseudo-experiment. We begin with an inertial observer, who performs a range of experiments in kinematics and other branches of physics. The observer is then boosted into uniform motion with respect to his original state of motion and carries along with him a complete record of all the experiments and their outcomes. These experiments are now repeated and the outcomes compared with those of the original set. The principle of relativity requires that both sets of outcomes must be the same and a theory satisfying the principle of relativity must predict that this will be so. (For a comparison of this sense of the principle and the one that requires form invariance of laws, see Anderson (1964, pp 176–82).)

This pseudo-experimental condition can be translated into a theoretical condition that amounts to the principle of relativity in special relativity. The theoretical analog of the *inertial observer is the inertial frame of reference. The analog of the setting of the observer* into uniform motion is a Lorentz transformation of the frame of reference. The setting up and outcome of all experiments performed by the observer will be determined fully by the spacetime structures of the theory. Therefore the carrying along of the complete description of the observer's experiments and outcomes translates into the carrying along under Lorentz transformation of the spacetime structures of the theory†. The principle of relativity now simply amounts to the requirement that the Lorentz transformation map spacetime structures allowed by the theory into spacetime structures allowed by the theory.

Without further assumption it follows that special relativity satisfies the principle of relativity as far as all kinematical experiments are concerned. These are idealized experiments in which the frame of reference directly 'sees' the metrical structure of the spacetime without assistance from further material systems. Their outcome is determined solely by that metrical structure. The satisfaction of the principle of relativity follows immediately from the fact that an arbitrary Lorentz transformation h is a symmetry of the Minkowski metric g_{ab}, that is, it satisfies Fock's condition (11). Therefore, if h transforms an inertial frame F_1 into an inertial frame F_2, then the metric seen by F_1 and carried along to F_2, h^*g_{ab}, is the same as the metric g_{ab} seen by F_2.

In the more realistic case, the experiments will involve further spacetime structures, such as electromagnetic fields and charges. The principle of relativity will be satisfied only if these further spacetime structures satisfy the following condition, which is the geometric

† This treatment assumes that there are no spacetime structures that elude experimental test, such as the absolute spacetime rigging of a Newtonian spacetime, which introduces a state of rest that cannot be revealed in any experiment (see Friedman 1983, ch III).

statement of the Lorentz covariance of the theories of these further structures. Let the theory have models

$$\langle M, g_{ab}, (O_1)_{ab...}, (O_2)_{ab...}, \ldots \rangle \tag{12}$$

where M is an $\mathbb{R}^4$ differentiable manifold, g_{ab} a Minkowski metric and $(O_1)_{ab...}$, $(O_2)_{ab...}$, . . . the extra spacetime structures. If h is any Lorentz transformation and (12) a model of the theory, then

$$\langle M, g_{ab}, h^*(O_1)_{ab...}, h^*(O_2)_{ab...}, \ldots \rangle$$

must also be a model of the theory. The satisfaction of the principle of relativity now follows. Let F_1 be an inertial frame of reference in which are conducted experiments associated with structures $(O_1)_{ab...}$, $(O_2)_{ab...}$, . . . If we transform via Lorentz transformation h to any other inertial frame F_2, we require that the theory admit precisely the same experiments and outcomes. That is, we require that the theory allow structures $h^*(O_1)_{ab...}$, $h^*(O_2)_{ab...}$, . . . This is precisely what the geometric version of Lorentz covariance allows.

This analysis gives us a precise sense in which the equivalence of inertial frames of reference is realized within the special theory of relativity. The basic moral of the work of Earman, Friedman and Jones is that there is no natural sense in which this equivalence obtains in the spacetimes of general relativity and that there is certainly no extension of it to accelerated frames of reference. In this sense, there is no principle of relativity in the general theory of relativity. This moral follows immediately from the fact that special relativity admits a non-trivial symmetry group, the Lorentz group, which maps inertial frames of reference into one another. The spacetimes of general relativity in general admit no symmetries. In general relativity, the closest analog of an inertial frame of reference is a frame in free fall. It is represented by a congruence of timelike geodesics. In general, a transformation that maps one freely falling frame or reference into another will not be a symmetry of the metrical structure. Therefore spacetime observers of the first frame will see different metrical properties in spacetime than will those of the second. The indistinguishability required for the equivalence of frames does not obtain. Considering arbitrary frames of reference rather than those in free fall clearly does not change this result.

That this sense of equivalence of frames fails to obtain in general relativity is not so surprising and it is difficult to imagine that Einstein ever expected that it would. The real puzzle, then, is to determine the sense in which Einstein believed the equivalence to be extended by general relativity. There is one reading in this geometric language that does allow a general equivalence of frames (Norton 1985, section 5). So far it has been assumed that the background spacetime is represented by the combination of manifold and metric. If instead one takes the manifold alone as the background spacetime, then one immediately has an equivalence of all frames of reference. For, considering just $\mathbb{R}^4$ manifolds for simplicity, an arbitrary automorphism is a symmetry of the manifold. Since any frame of reference can be mapped into any other by an automorphism, it follows that each frame 'sees' the same spacetime background so that they are equivalent in at least that sense.

If this equivalence is to be extended to the sort of equivalence of the principle of relativity of special relativity, then the metric tensor field of general relativity must be treated in a similar fashion to the structures $(O_1)_{ab...}$, $(O_2)_{ab...}$ of the above discussion of special relativity. Then a similar sense of equivalence of *arbitrary* frames follows directly from the active general covariance of general relativity. Let F_1 be any frame of reference which sees a metrical field g_{ab} and other fields $(O_1)_{ab...}$, $(O_2)_{ab...}$, . . . That is, the theory has a model

$$\langle M, g_{ab}, (O_1)_{ab...}, (O_2)_{ab...} \rangle.$$

Then, if F_2 is any other frame of reference, the theory must allow a model in which F_2 sees an identically configured set of fields. That is, if h is an automorphism that maps F_1 into F_2, then F_2 must see the fields h^*g_{ab}, $h^*(O_1)_{ab...}$, $h^*(O_2)_{ab...}$ so that theory must also have a model

$$\langle M, h^*g_{ab}, h^*(O_1)_{ab...}, h^*(O_2)_{ab...} \rangle.$$

That it does follows directly from its active general covariance (section 5.4 above).

The difficulty with this proposal is that it allows an equivalence of arbitrary frames of reference in all theories that are actively generally covariant. Such theories include versions of special relativity and Newtonian spacetime theory. Thus, if this generalized equivalence of frames is to be distinctive to general relativity, there must be some principled way of relegating the metric tensor to the contents of spacetime in general relativity, whereas in other spacetime theories, such as special relativity, this metrical structure is to be part of the background spacetime. What makes such a division plausible is the fact that the metric tensor of general relativity incorporates the gravitational field. Thus its state is affected by the disposition of masses in the same way as a Maxwell field is affected by the disposition of charges.

The analogy can be pressed further. In special relativity one can conduct an electrical experiment with some configuration of charges in an inertial frame of reference. The principle of relativity requires that, if we were to recreate that same configuration of charges in another inertial frame, then we would produce the identical fields and experimental outcomes. This is the sense in which all inertial frames of reference are equivalent. Similarly, one could consider some configuration of masses and the metric field they produce in relation to an arbitrary frame of reference in general relativity as a kind of gravitational experiment in that frame. The active general covariance of general relativity then tells us that we could have laid out the same configuration of masses and fields in any other frame of reference, so that the gravitational experiment would have proceeded identically in any frame of reference. This gives us a sense in which arbitrary frames of reference are equivalent in general relativity.

The success of this generalized equivalence depends fully on our being able to conceive of the metric field as a part of the contents of spacetime in general relativity but not in other theories like special relativity. Einstein's 1918 version of Mach's Principle allowed this conception since it required that the metric field be fully determined by the matter distribution, so that this field would have the same sort of status as the matter distribution. Since Mach's Principle in this form fails in many of the spacetimes of general relativity, it cannot be used to justify a generalized equivalence of frames in that theory. The only other well developed analysis that allows this conception of the metric field concerns the distinction between absolute and dynamic objects to be discussed in section 8 below. As a dynamical object, the metric of general relativity is naturally classified as part of the content of spacetime. As an absolute object, the Minkowski metric of special relativity is naturally classified as part of the background spacetime.

7. General relativity without principles

7.1. General relativity without general relativity

Einstein's own developments and discussion of the general theory of relativity place so much importance on general covariance and the extension of the principle of relativity that most accounts of the theory seems compelled to take a position on their importance. Many essentially agree with Einstein as we have seen in section 4. Many others, as we have seen in sections 5 and 6, disagree with Einstein's views; they develop general relativity without claiming general covariance as a fundamental physical postulate and they explain why they do so.

There is a third category of exposition of general relativity. These are the expositions that take no special notice of general covariance at all. Of course they develop general relativity in a generally covariant formalism, as is the standard practice. However the expositions are conspicuous for the absence of any statement of fundamental principle concerning covariance or relativity. There is no 'principle of general covariance', no 'general principle of relativity' and no pronouncement that the theory has extended the equivalence of frames of reference to accelerated frames. And there is no explanation of why these principles are not discussed.

It is difficult to know what significance to read into such formulations of general relativity without general relativity. Many of these expositions are mathematically oriented. So we might suppose that their authors simply decided not to contend with the question of the physical foundations in favor of other more mathematical aspects of the theory. It is hard to imagine, however, that an author writing on general relativity can be completely unaware of Einstein's views, if not also the disputes over them. Therefore when that author writes a textbook length exposition of general relativity which fails to include such phrases as 'general principle of relativity' or 'principle of general covariance', one must suppose that the author is making a statement by omission. (The omissions are typically so complete that, if the text has an index, these terms will not be listed in it.) We have already seen that Synge and Fock object to 'general relativity' as a misnomer. Thus it seems obvious that similar sentiments drive such authors as that of *Time and Space, Weight and Inertia: A Chronogeometrical Introduction to Einstein's Theory* (Fokker 1965) who display remarkable ingenuity in avoiding the term 'general relativity.'

Finally, even if no statement is being made by omission, the very possibility and frequency of such accounts of general relativity do indicate that the place of these principles in the theory might not be so straightforward. If the principles are fundamental physical axioms, they would be hard to avoid, even as consequences in an alternate axiomatization. One is hard pressed to imagine a formulation of thermodynamics without the law of conservation of energy as a fundamental axiom or one of the earliest and most important theorems! The subtlety of the situation is captured by Trautman, who observed well into his exposition (1964, p 122) of general relativity

> . . . we have managed to obtain general relativity by a (we hope) fairly convincing chain of reasoning without ever mentioning such a principle [of general covariance].

He did proceed, however, to list several senses of the principle and their non-trivial relationships to the theory. Thus one can find general covariance relevant without mentioning it in a development of the theory.

With these interpretative cautions, we can proceed to note that the tradition of exposition of general relativity without general relativity extends back to the earliest decades of the theory. There are many exposition of relativity theory with this character from the 1920s.

They include Bauer (1922), Birkhoff (1927), Darmois (1927), Chazy (1928) and De Donder (1925) (but De Donder (1921, pp 10–15) had emphasized the arbitrariness of coordinates in general relativity and the invariance of its fundamental equations). Eddington (1924, ch I, section 1) labours in detail the notion that one can use arbitrary 'space-time frames' for describing phenomena, but without ever mentioning a principle of covariance or a generalized principle of relativity. His earlier Eddington (1920, p 20) had allowed that a generalization of the principle of relativity in the theory in so far as he conceded 'it will be seen that this principle of equivalence is a natural generalization of the principle of relativity'. This remark was not repeated in Eddington (1924).

The lean years after the 1920s saw several exposition of general relativity without general relativity: Rainich (1950) and the synopsis of general relativity by Zatzkis (1955). The revival of interest in general relativity in the 1960s brought more such expositions and they have included some of the most important expositions of the theory: Fokker (1965), Schild (1967) (although he mentions (p 20) that general relativity 'shows there are no inertial frames as all'), Robertson and Noonan (1968), Ehlers (1971), Hawking and Ellis (1973), Dirac (1975), Falk and Ruppel (1975) (although the notion of a generalized principle of relativity is alluded to briefly, e.g., p 323), Sachs and Wu (1977), Clarke (1979) (although section 3.1.3 does emphasize the loss of global inertial systems and the novelty of arbitrary coordinate systems in general relativity), Frankel (1979), Shutz (1985) (although it is allowed (p 3) that general relativity is more general in allowing both inertial and accelerated observers), Martin (1988), Hughston and Tod (1990), Stewart (1990).

7.2. The principle of equivalence as the fundamental principle

While many of these accounts of general relativity avoid mention of principles of general covariance and of generalized relativity, many of them do find a special place for just one of the three fundamental principles listed by Einstein in 1918, the principle of equivalence. Of course the version used is typically not Einstein's but one or other variant of an infinitesimal principle of equivalence. The principle is not used in Einstein's manner as a stepping stone to a generalized the principle of relativity. Rather it is used to establish a notion claimed as a fundamental principle of general relativity, that special relativity holds infinitesimally in the theory; or, less commonly, it is just taken to be as much of the generalized principle of relativity as general relativity will admit.

Such treatments, which employ only the principle of equivalence as a fundamental principle, include: Silberstein (1922, p 12), Eddington (1924, section 17) (although emphasizing (p 41) that the principle is to be derived rather than postulated in the exposition), Birkhoff (1927, pp 140–4), Landau and Lifshitz (1951, ch 10), Fokker (1965, section V.6) (with the principle in Einstein's original form), Robertson and Noonan (1968, section 6.9), Schild (1967), Falk and Ruppel (1975, section 32), Clarke (1979, ch 3), Frankel (1979, ch 2), Raine and Heller (1981, ch 6,8), Schutz (1985, p 184), Martin (1988, sections 1.6, 5.11), Stewart (1990, section 1.13). We have the expositions of Tonnelat (1959), who takes the principle of equivalence to be a 'principle of generalized relativity' (p 327) and Wasserman (1992), who also remarks briefly (p 342) that the principle of equivalence extends the principle of relativity to include accelerated frames of reference.

7.3. Challenges to the principle of equivalence

One might well wonder if we have not at last found the uncontroversial core of Einstein's accounts of the foundational principles of general relativity in these expositions. That core would now just be the principle of equivalence, even if it is in an altered form Einstein never endorsed. However not even the popular versions of the principle of equivalence have escaped telling attack.

The best known challenge has been stated most clearly by Synge. His concern is that the presence or absence of a gravitational field must be characterized geometrically, that is, in invariant terms. He asserts that the presence of a gravitational field corresponds just with non-vanishing curvature of the spacetime. Such an invariant criterion is unaffected by coordinate transformation, by change of frame of reference or by a change of the state of motion of the observer. Therefore none of these changes will be able to transform away a gravitational field or bring one into existence, contrary to many versions of the principle of equivalence. He is unimpressed with the requirement that the spacetime metric be such that we can always find a coordinate system in which the components of the metric become diag $(1,1,1,-1)$ at some nominated event, thereby mimicking special relativity at least in some infinitesimal sense. Synge deems this trivial since it merely amounts to the requirement that the metric have Lorentz signature. Thus he wrote his famous lament (1960, p ix) about relativists who

> . . . speak of the Principle of Equivalence. If so, it is my turn to have a blank mind, for I have never been able to understand this Principle. 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 does vanish. This is an absolute property; it has nothing to do with any observer's world-line. Spacetime is either flat or curved, and in several places in the book I have been at considerable pains to separate truly gravitational effects due to curvature of space-time from those due to curvature of the observer's world-line (in most cases the latter predominate). The Principle of Equivalence performed the essential office of midwife at the birth of general relativity, but, as Einstein remarked, the infant would never get beyond its long-clothes had it not been for Minkowski's concept. I suggest that the midwife be now buried with appropriate honors and the facts of absolute space-time faced.

The idea that the presence of a gravitational field is associated with the invariant property of curvature can be translated into observational terms. The non-vanishing of the Riemann curvature tensor entails the existence of tidal forces acting on bodies in free fall. The goal of restricting versions of the principle of relativity to infinitesimal regions of spacetime is to eliminate these tidal forces. However they cannot be so eliminated; for example, the tidal bulges on a freely falling droplet remain as the droplet becomes arbitrarily small, ignoring such effects as surface tension; see Ohanian (1976, ch 1, 1977) and Bondi (1979). See also Norton (1985, section 10) for an attempt to characterize the imprecise restriction to infinitesimal regions as a restriction on access to certain orders of quantities defined at a point. Following a suggestion from Einstein, it turns out that an infinitesimal principle of equivalence can hold only at the expense of a restriction so severe that it trivializes the principle. See also Norton (1985, section 11) for Einstein's response to the idea that vanishing spacetime curvature is to be associated with the absence of a gravitational field.

8. Eliminating the absolute

8.1. Anderson's absolute and dynamical objects

However else he may have changed his viewpoint, we have seen (section 3.9) that Einstein maintained throughout the lifetime of this writings on general relativity that it was distinguished from earlier theories by a single achievement: it had eliminated a causal absolute, the inertial system. If we are to have an account that truly captures Einstein's understanding of general covariance, then we should expect this rather imprecisely formulated notion to play a prominent role. This notion surely lies behind Pauli and Weyl's emphasizing that the metric tensor is determined by the matter distribution through field equations and that this justifies (Weyl) the name 'general theory of relativity' (see section 5.5 above).

Einstein's notion surfaces more clearly in Bergmann's (1957, pp 11–12) conception of weak and strong covariance. Weak covariance is the type we see in when we use many different coordinate systems to describe the one phenomenon in Lagrangian mechanics.

> The fundamentally trivial nature of this 'weak covariance' derives from the rigidity of the classical metric.

This is quite distinct from the strong covariance of general relativity where†

> it is one's task to *calculate the metric* . . . as a dynamical variable. We can take one coordinate system or another for this job, but all that we can know is the relation of one frame to the other: we do not know the relation of *either* to the world. 'Strong covariance', therefore, contains not only a reference to the structural similarity of an equation and its transform; it implies as well that one frame is as good a starting point as another — that we do not *need* prior knowledge of its physical meaning . . . which is generated at the *end*.

Many important themes are touched on here, as has been indicated by Stachel (forthcoming, footnote 3). The distinction between weak and strong covariance amount to that between passive and active covariance. What concerns us here, however, is the contrasting of the 'rigidity of the classical metric' with the metric of general relativity 'as a dynamical variable'.

The most precise context so far for the statement of Einstein's causal concerns has been provided by Anderson (1964, 1967, ch 4, 1971) (but see also Anderson (1962) for a definition of absolute change within general relativity). In laying out his system, Anderson uses a somewhat idiosyncratic nomenclature. He labels the set of all possible values of the geometric objects of a theory the 'kinematically possible trajectories'. Those sanctioned by the 'dynamical laws' or 'equations of motion' of the theory, he calls the 'dynamically possible trajectories'. The principal novelty of Anderson's development is the distinction between 'absolute' and 'dynamical' objects. That distinction will be used to strengthen the principle of general covariance into a more restrictive 'principle of general invariance'.

Although allowing for a time that both special and general covariance principles are devoid of physical content (1964, p 182), Anderson (1967, section 4.2, 1971, pp 162–65) then came to urge that the requirement of general covariance is not physically vacuous. He allowed that one can take a physical theory and generate successive formulations of wider and wider covariance. However there is a point in the hierarchy at which we are forced to introduce elements which are unobservable or transcend measurement. Since we are prohibited from proceeding to this point in the hierarchy, covariance requirements have physical force. (This strategy for injecting physical content into covariance principles is essentially the one used by Pauli and others in section 5.5 above.)

The absolute objects of a spacetime theory are distinguished by precisely the causal criterion that allowed Einstein to designate the inertial systems of special relativity as absolute. Anderson and Gautreau (1969, p 1657) summarize:

> Roughly speaking, an absolute object affects the behaviour of other objects but is not affected by these objects in turn.

† The two ellipses '. . .' and emphasis are Bergmann's.

The remaining objects are dynamical. Thus the Minkowski metric of special relativity is an absolute object. In special relativistic electrodynamics, the Minkowski metric affects the Maxwell field and charge flux in determining, for example, which are the inertial trajectories of charges. However neither Maxwell field nor charge flux, the dynamical objects of the theory, affect the Minkowski metric. Whatever their form, the Minkowski metric stays the same. This is the sense in which it affects without being affected. Since the Minkowski metric induces the inertial frames on spacetime, Anderson's identification of the Minkowski metric as an absolute object fits exactly with Einstein's identification of inertial frames as absolutes.

This loose definition must be made more precise and Anderson (1967, pp 83–4) (see also Anderson 1971, p 166) gives a more precise definition. Having eliminated irrelevant objects from the set of geometric objects y_A allowed in the theory:

> We now proceed to divide the components of y_A into two sets, ϕ_α and z_a where the ϕ_α have the following two properties:
>
> (1) The ϕ_α constitute the basis of a faithful realization of the covariance group of the theory.
>
> (2) Any ϕ_α that satisfies the equations of motion of the theory appears, together will all its transforms under the covariance group, in every equivalence class of dpt [dynamically possible trajectories].
>
> The ϕ_α, if they exist, are the components of the absolute objects of the theory. The remaining part of y_A, the z_a, are then the components of the dynamical objects of the theory.

Condition (1) is an important but essentially technical condition that the transformation behaviour of the ϕ_α respect the group structure of the theory's covariance group (e.g. the ϕ_α ought to transform back into themselves under an identity transformation of the covariance group). Condition (2) essentially just says that the absolute objects ϕ_α are the same in every dynamically possible trajectory (i.e. model) of the theory. The condition, however, must allow that an absolute object, such as a Minkowski metric $g_{\mu\nu}$, can be manifested in many different forms as it transforms under the members of the covariance group. Therefore the second condition collects the dynamically possible trajectories into equivalence classes of intertransformable members. Since each class is closed under transformations of the covariance group, the one set of absolute objects and all their transforms will appear in each class. Thus condition (2) requires, in effect, that the absolute objects that appear in all models are the same up to a transformation of the theory's covariance group.

With this distinction in place, Anderson now defines the symmetry group or 'invariance group of a physical theory' (Anderson 1971, p 166) as

> that subgroup of the covariance group of the theory which leaves invariant the absolute objects of the theory. In particular, if there are no absolute objects, the invariance group and the covariance group are the same group.

The 'leaves invariant' is to be understood in the sense of a symmetry transformation such as given in (10) and (11) above. There is an analogous definition for the 'symmetry group of a physical system' (Anderson 1967, p 87).

Anderson's central claim (e.g. Anderson 1967, p 338) is that this symmetry group is what Einstein really had in mind when he associated the Lorentz group with special relativity and the general group with general relativity. For a requirement on a symmetry group, not a covariance group, is the correct way to express a relativity principle. Even if we formulate our theories in generally covariant fashion, they continue to be characterized by the groups expected if we look to their symmetry groups. The symmetry group of a generally covariant special relativity is the Lorentz group. Again, consider a generally covariant formulation of Newtonian spacetime theory with spacetime structures t_a, h^{ab} and ∇_a, where the gravitational field is not incorporated into ∇_a. Then these three objects are the absolute objects of the theory and their symmetry group is the Galilean group. Finally, general relativity has no absolute objects. Its symmetry group is the general group.

One can grasp the picture urged if one imagines that the background spacetime of a theory is the spacetime manifold together with the theory's absolute objects—although 'background spacetime' is not a notion discussed by Anderson. In the cases of special relativity and the above version of Newtonian spacetime theory, both admit a family of preferred inertial frames of reference which remain unchanged under the Lorentz group or Galilean group respectively. In the case of general relativity, the background spacetime is just the manifold whose symmetry group is the group of arbitrary transformations.

According to Anderson, what Einstein really intended with his principle of general covariance is what Anderson calls the 'principle of general invariance'. This principle requires that the symmetry group of a theory be the general group of transformations or, as Anderson calls them, the 'manifold mapping group'. This principle rules out the possibility of any non-trivial absolute objects in the theory, that is, those which have more than merely topological properties. In this sense, the principle of general invariance amounts to a no-absolute-objects requirement and offers a precise reading for Einstein's claim that general covariance has eliminated an absolute from spacetime.

8.2. Responses to Anderson's viewpoint

Anderson's ideas on absolute and dynamical objects have found a limited but favorable response in the literature. Misner *et al* (1973, section 17.6) present a requirement of no absolute objects in terms of the requirement of 'no prior geometry' where:

> By 'prior geometry' one means any aspect of the geometry of spacetime that is fixed immutably, i.e., that cannot be changed by changing the distribution of gravitating sources.

They describe Einstein as seeking both this requirement as well as a 'geometric, coordinate independent formulation of physics' when he required general covariance—and that this has been responsible for a half century of confusion.

Anderson's principle of general invariance appears in Trautman (1973), as does the distinction between absolute and dynamical objects in Kopczynski and Trautman (1992, ch 13). Ohanian (1976, pp 252–4) uses Anderson's principle of general invariance to respond to Kretschmann's objection that general covariance is physically vacuous. He does insist, however, that the principle is not a relativity principle and that the general theory of relativity is no more relativistic than the special theory (p 257). Anderson's ideas seem also to inform Buchdahl's (1981, Lecture 6) notion of 'absolute form invariance'.

The distinction between absolute and dynamical objects has been received and developed most warmly by philosophers of space and time, so that in place of (6), the general model of a spacetime theory is given as

$$\langle M, A_1, A_2, \ldots, D_1, D_2, \ldots \rangle$$

where $A_1, A_2, \ldots$ are the absolute objects and $D_1, D_2, \ldots$ the dynamical. However they do not generally allow that Anderson's reasoning has vindicated Einstein's claim that the general theory of relativity extends the principle of relativity of special relativity. See Earman (1974, 1989, ch 3), Friedman (1973, 1983) and Hiskes (1984). Earman (1989, section 3.4) investigates the possibilities for the symmetry group of the absolute objects of a theory differing from the symmetry group of the dynamical objects.

8.3. No gravitational field—no spacetime points

Stachel (1986, sections 5, 6) has provided an interesting extension of the viewpoint advanced by Anderson. Stachel's concern is that our formulations of general relativity are still not in a position to explicate Einstein's idea that spacetime cannot exist without the gravitational field (see section 3.5 above). Stachel faults our representing or physical spacetime events by the mathematical points of the spacetime manifold. Read naively, this definition tells us that a manifold without metrical field represents a physical spacetime of events with topological properties but with no metrical relations.

Stachel's proposal applies to spacetime theories without absolute objects, which he calls 'generally covariant', and can be reviewed only informally here. To form the models of such theories one assigns various geometric objects—tensor fields, for example—to each point of the manifold in the usual way. In principle, many different such fields could be assigned. In the case of general relativity, we have a host of possible metrical fields of all sorts of different curvature. The loose notion of the space of all such possible fields is given precise formulation by Stachel as a fibre bundle E over the manifold M. The particular fields that are chosen for inclusion in the theory's models are picked out through cross-sections of the fibre bundle E. Loosely speaking, a cross-section σ amounts to an association of a point of the manifold M with the geometric objects assigned to it in some model of the theory. (More precisely, a cross-section σ is a map that goes from a point p of the manifold M to a member $\sigma(p)$ of the fibre bundle E, where $\sigma(p)$ must be associated with p by the bundle's projection map π, so that $\pi\sigma(p) = p$.)

The core of Stachel's proposal is that the physical events of spacetime are represented by the inverse of this map σ. That is—loosely speaking—the physical events are not represented directly by the points of the spacetime manifold; rather, in their place, we use the association of the points of the manifold with the geometric structures defined on them. We now automatically have the property of spacetime that Einstein announced. If we take away the gravitational field, that is the metric field, from a spacetime in general relativity, then we have taken away the fibre bundle and with it the map that represents the physical spacetime events. In a theory with absolute objects, however, physical events are represented directly by points of the base manifold. Therefore their behaviour is quite different. See Stachel (1986) for further details of how theories with absolute objects are treated and of the machinery needed to allow that one physical situation is represented by an equivalence class of diffeomorphic models.

8.4. What are absolute objects and why should we despise them?

There are two areas of difficulty associated with the general theory of absolute and dynamical objects. The first is the question of how we define absolute objects. Anderson's definition was that an object was absolute if the same object (up to coordinate transformation) appeared in all the theory's models. In the coordinate free, geometric language, how are we to understand the 'same'? The obvious candidate is that two objects are the same if they are isomorphic. Global isomorphism is the criterion used in Earman's (1974, p282) definition of absolute objects to pick out when one has the same object in all models. Friedman (1973, p308–9, 1983, p58–60) uses only the requirement that the objects by locally diffeomorphic†.

The first difficulty with this criterion of diffeomorphic equivalence as sameness was pointed out by Geroch (Friedman 1983, p59). The criterion deems as the same all timelike, non-vanishing vector fields, so that however such a field arises in a theory, it will be one of its absolute objects. Thus, in standard 'dust' cosmologies, the velocity field U^a of the dust becomes an absolute object. To avoid the problem, Friedman suggests a rather contrived escape: formulate the theory of dust with the dust flux ρU^a, where ρ is the mass density, instead of ρ and U^a separately. (Friedman is relying here on the possibility that ρ vanishes somewhere. A better choice would have been the stress-energy tensor for pressureless dust $\rho U^a U^b$.)

More seriously, modifying slightly an example of Torretti (1984, p285), we could imagine the following hybrid classical relativistic cosmology. The spacetime structure is given *exactly* by any of the Robertson–Walker spacetime metrics. The metrics are posited *a priori* and not governed by the presumed inhomogeneous matter distribution through gravitational field equations. Therefore the curvature of the metric is unaltered in the vicinity of massive bodies. In this case, we would judge the metrical spacetime structure to act on the matter distribution without the matter distribution acting back on it. However, since models of the theory would allow metrics of different curvature, we cannot use existing definitions to identify the spacetime metric as an absolute object. Torretti's counterexample shows us that the basic notion of 'sameness' does not fully capture the notion of things that act but are not acted upon.

The second area of difficulty associated with the general theory of absolute and dynamical objects is a presumption of Anderson and Einstein (assuming that he is correctly interpreted by the theory). They presume that there is some compulsion to eliminate absolute objects. Of course they are right in the sense that our best theory of space and time happens not to employ absolute objects. Thus several of Anderson's arguments for the principle of general invariance and therefore against absolute objects essentially tell us that this assumption can form a premise of arguments that lead to empirically confirmed results (Anderson 1967, section 10.3, 1971, p169). However absolutes are supposed to be defective in a deeper sense. It is not just that we happen not to see absolutes in nature; Nature is somehow supposed to abhor things that act but are not acted upon. The difficulty is to clarify and justify this deeper sense.

Anderson (1967, p339, 1971, p169) sees in nature a 'generalized law [principle in 197] of action and reaction'. But the principle is so vague that it is unclear what the principle really says and where it can be applied. Does Planck's constant h or the gravitation constant G 'act' on matter without suffering 'reaction'? With this vagueness how can we tell if the law is true or even whether we should hope for it to be true? Is it, perhaps, a dubious guilt by association with Aristotle's Unmoved Mover? Einstein comes closer to an explanation with his analogy (section 3.9 above) to pots of water, one boiling, one not. There has to be a sufficient reason for the difference. Analogously, the difficulty with absolute objects is that there is no sufficient reason for them to be one way rather than another. Now we might allow that such a principle of sufficient reason applies to temporally successive states of systems, although quantum theory calls even that into doubt. But why should we require this sort of principle to hold for aspects of the universe as a whole? In answer, we might take

† More precisely, in the 1983 version of the definition, what Friedman calls '*d*-equivalence' is this: If a theory has models $\langle M, \Phi_1, \ldots, \Phi_n\rangle$ and $\langle M, \Psi_1, \ldots, \Psi_n\rangle$, then Ψ_i and Φ_i are *d*-equivalent if, for every $p \in M$, there are neighbourhoods A and B of p and a diffeomorphism $h : A \rightarrow B$ such that $\Psi_i = h^*\Phi_i$.

Born expansion of Einstein's (1916, section 2) denunciation of an absolute, inertial space as an *ad hoc* cause. Born (1924, p 311) explains

> If, however, we ask what absolute space is and in what other way it expresses itself, no one can furnish an answer other than that absolute space is the cause of centrifugal forces but has no further properties. This consideration shows that space as the cause of physical occurrences must be eliminated from the world picture.

It is hard to sympathize with Born's complaint. The absolute Minkowski metric of a special relativistic world has an extremely rich collection of properties all of which can be confirmed by possible experiences. It is difficult not to see the very objection of Born and Einstein as *ad hoc*. They seek to use vague and speculative metaphysics to convert something that happens to be false into something that has to be false. These seem to be Schlick's (1920, p 40) sentiments when he observes

> . . . we can . . . consider the expression 'absolute space' to be a paraphrase of the mere fact that these [centrifugal] forces exist. They would then simply be immediate data; and the question why they arise in certain bodies and are wanting in others would be on the same level with the question why a body is present at one place in the world and not at another. . . . I believe Newton's dynamics is quite in order as regards the principle of causality.

Special relativity has suffered too long from the crank myth that it not just happens to be true but it has to be true and that proper meditation on clocks and light signaling reveals it. Let us not create a similar myth for general relativity.

9. Boundaries and puzzles

9.1. Is general covariance too general? Or not general enough?

While most have been satisfied with general relativity as a generally covariant theory, Fock (1957, 1959, pp xv–xvi, section 93) has proposed that the four coordinate degrees of freedom of the generally covariant theory be reduced by application of a coordinate condition. Fock's 'harmonic coordinates' are picked out by the condition

$$\frac{\partial}{\partial x_\mu}\left(\sqrt{-g}\; g^{\mu\nu}\right) = 0.$$

Fock applies this restriction to the case of spacetimes which are Minkowskian at spatial infinity and finds that the resulting equations are the natural generalization of the standard Galilean coordinates of special relativity and are fixed up to a Lorentz transformation. Fock sees the physical importance of harmonic coordinates in such problems as the justifying of Copernican over the Ptolemaic cosmology. In harmonic coordinates, the earth orbits the sun and not vice versa.

Fock's proposal proved controversial. Criticism of Fock's proposal was aired at a conference in Berne in July 1955 for the jubilee of relativity theory (Fock 1956). Infeld argued that a restriction to harmonic coordinates is acceptable as a convenience. 'But to add it always (or almost always) to the gravitational equation and to claim that its virtue lies in the fact that the system is only Lorentz invariant, means to contradict the principle idea of relativity theory.' Trautman (1964, p 123) and Kopczynski and Trautman (1992, p 124) have also objected that Fock's proposal amounts to the postulation of new spacetime structures for which no physical interpretation can be given.

In so far as Fock intended to reduce permanently the covariance of general relativity and introduce further structure, then these critical attacks are warranted. The harmonic coordinate condition is unacceptable as a new physical principle. But Fock (1959) seems to hold a milder position. He emphasized (pp 350–1) that the introduction of harmonic coordinates is intended in a spirit no different from that which introduces preferred Galilean coordinates into a generally covariant formulation of special relativity. Thus 'the existence of a preferred set of coordinates . . . is by no means trivial, but reflects intrinsic properties of space-time'. In the case of a spacetime Minkowskian at spatial infinity, harmonic coordinates simply reveal a structure already assumed as part of the boundary condition. Their use does not amount to an unwarranted postulation of new structure—unless one deems the boundary conditions themselves unwarranted. For further discussion see Gorelik (forthcoming).

The issue surrounding Fock's proposal was whether a restriction of the covariance of general relativity could be justified. Arzeliès (1961, pp xliv–il, 5–7 ch XIV) has proposed a modification of general relativity which amounts to a kind of expansion of its covariance. He urges that Einstein's theory has still not satisfied the requirements of the generalized principle of relativity and that the transformations it allows should be extended in the following sense. If we start with is a coordinate system X^i, then, under a coordinate transformation, the coordinate differential dX^k transform into new coordinate differentials dx^i. It is customarily assumed that the coordinate differentials dx^i are exact, so that they can be integrated into the new coordinate systems x^i. Arzeliès proposed that this restriction be dropped. This would certainly generalize the group of transformations since the functions Λ^i_k of the equations $dx^i = \Lambda^i_k dX^k$ need no longer be restricted by the requirement of exactness. The modification is extremely far reaching, however, in so far as it leads to the loss of many familiar theorems. For example, it will now be possible to transform the line elements of non-flat metrics to the form

$$ds^2 = (dx^1)^2 + (dx^2)^2 + (dx^3)^2 + (dx^4)^2$$

over a neighbourhood (not just at a point), where this was formerly only possible if the metric was flat.

9.2. *The Einstein puzzle*

There is a presumption in much modern interpretation of Einstein's pronouncements on the foundations of the general theory of relativity. It is that much of what he says cannot be taken at face value. (Why does Einstein make such a fuss about introducing arbitrary spacetime coordinates? We have always been able to label spacetime events any way we please!) Thus we are either to translate what he really meant into some more precise context, as does Anderson, or to dismiss it as confused. The proposal of Norton (1989, 1992) is that our modern difficulty in reading Einstein literally actually stems from a change of context. (For related concerns see Norton (1993).)

The relevant change lies in the mathematical tools used to represent physically possible spacetimes. In recent work in spacetime theories, we begin with a very refined mathematical entity, an abstract differentiable manifold, which usually contains the minimum structure to be attributed to the physical spacetimes. We then judiciously add further geometric objects only as the physical content of the theory warrants. Moreover, we have two levels of representation. We first represent the physically possible spacetimes by the geometric models of form (6) and then these geometric models are represented by the coordinate based structures (7). General covariance is usually understood as passive general covariance and therefore arises as a mathematical definition, as we have seen.

In the 1910s, mathematical practices in physics were quite different. The two levels of representation were not used. When one represented a general space or spacetime, one used number manifolds—$\mathbb{R}^n$ or $\mathbb{C}^n$, for example. Thus Minkowski's 'world' was not a differentiable manifold that was merely topologically $\mathbb{R}^4$. It was literally $\mathbb{R}^4$; that is, it was the set of all quadruples of real numbers.

Now anyone seeking to build a spacetime theory with these mathematical tools of the 1910s faces very different problems from the ones we see now. Modern differentiable manifolds have too little structure and we must add to them. Number manifolds have far too much structure. They are fully inhomogeneous and anisotropic. The origin $\langle 0,0,0,0 \rangle$ is quite different from every other point, for examples. And all this structure had canonical physical interpretation. If one took the x_4 axis as the time axis, then x_4 coordinate differences were physically interpreted as differences of clock readings. Timelike straights would be the inertial trajectories of force free particles. The problem was not how to add structure to the manifolds, but how to deny physical significance to existing parts of the number manifolds. How do we rule out the idea that $\langle 0,0,0,0 \rangle$ represents the preferred center of the universe and that the x_4 coordinate axis a preferred state of rest?

Felix Klein's *Erlangen* program provided precisely the tool that was needed. One assigns a characteristic group to the theory. In Minkowski's case, it is the Lorentz group. Only those aspects of the number manifold that remain invariant under this group are allowed physical significance. Thus there is no physical significance in the preferred origin $\langle 0,0,0,0 \rangle$ of the number manifold since it is not invariant under the transformation. But the collection of timelike straights of the manifold are invariant; they represent the physically real collection of all inertial states of motion. As one increases the size of the group, one strips more and more physical significance out of the number manifold.

We can put this in another way. A spacetime theory coordinates a physically possible spacetime with the number manifold. The characteristic group of the theory tells us that many different such coordinations are allowed and equally good. What is physically significant is read off as that part of each coordination common to all of them. This coordination of physical events with quadruples of numbers in $\mathbb{R}^4$ is what was meant by 'coordinate system' and the equivalence of two such systems was far from a mathematical triviality. It was the essence of the physical content of the theory.

It is in this tradition that Einstein worked in the 1910s. His project was to expand the group of his theory as far as possible. But he had to proceed carefully since such expansions came with a stripping of physical significance from the number manifold. Thus Einstein (1916, section 3) needed to proceed very cautiously in explaining how the general covariance of his new theory had stripped the coordinates of their direct relationship to the results of measurement by rod and clock. The project is clearly also a project of relativization of motion. The imposition of the Lorentz group stripped the x_4 axis of the physical significance as a state of rest, implementing a principle of relativity for inertial motion. The transition to the general group stripped the set of timelike straights of physical significance as inertial motion, extending the principle to accelerated motion.

If this was all that Einstein had done, then his whole project would have remained within the *Erlangen* program tradition and there would be no debates today over whether Einstein succeeded in extending the principle of relativity. But, in the transition from the Lorentz to the general group, Einstein added an element that carried him out of the tradition of the Erlangen program. He associated a Riemannian quadratic differential form with the spacetime. (Thus Cartan (section 6.2 above) captures precisely the crucial point.) While

Einstein could correctly say that he had generalized the principle of relativity insofar as he had stripped physical significance from the timelike straights of the number manifold, what remained to be seen was whether he had reintroduced essentially this same structure by means of the quadratic differential form. In effect this question has become the focus of the debate over the generalized principle of relativity.

Finally, it is helpful to bear in mind that what Einstein meant by 'coordinate system' is not the same as the modern 'coordinate charts' of a differentiable manifold. The latter relate structures of (6) and (7) and the equivalence of each representation is a matter of mathematical definition. Einstein's coordinate systems are actually akin the representation relation between physically possible spacetimes and the models of form (6). That two models represent the one physically possible spacetime is a physical assumption that amounts to assuming that their mathematical differences have no physical significance. Correspondingly, within the context of Einstein's formulation of spacetime theories, that two coordinate system represent a physically possible spacetime is once again a physical assumption and for the same reason. That is, Einstein's covariance principles are most akin to modern active covariance principles.

In sum, there is no real puzzle in much that of what Einstein said. Rather it now only seems puzzling since he is solving problems we no longer have because of the greater sophistication of our mathematical tools. Indeed, in good measure we owe to Einstein's inspiration the development and widespread use of mathematical tools that automatically solve problems over which he laboured so hard.

10. Conclusion

The debate over the significance of general covariance in Einstein's general theory of relativity is far from settled. There are essentially three view points now current.

First is the viewpoint routinely attributed to Einstein. It holds that the achievement of general covariance automatically implements a generalized principle of relativity. In view of the considerable weight of criticism, this view is no longer tenable. Relativity principles are symmetry principles; the requirement of general covariance is not a symmetry principle. The requirement of general covariance, taken by itself, is even devoid of physical content. It can be salvaged as a physical principle by supplementing it with further requirements. The most popular are a restriction to simple law forms and a restriction on the additional structures that may be used to achieve general covariance. However neither supplementary condition has been developed systematically beyond the stage of fairly casual remarks.

The second viewpoint has been developed by Anderson and is based on his distinction between absolute and dynamical objects. His 'principle of general invariance' entails that a spacetime theory can have no non-trivial absolute objects. Anderson argues that the principle is a relativity principle, since it is a symmetry principle, and that it is what Einstein really intended with his principle of general covariance. In this approach, general relativity is able to extend the symmetry group of special relativity from the Lorentz group to the general group. This extension depends on the metric being a dynamical object, which is no longer required to be preserved by the symmetry transformations of the theory's relativity principle.

The third viewpoint holds that the dynamical character of the metric is irrelevant in this context and that the metric must be preserved under the theory's symmetry group, if that group is to be associated with a relativity principle. Since the metrics of general relativistic spacetimes have, in general, no non-trivial symmetries, there is no non-trivial relativity principle in general relativity. Whatever may have been its role and place historically, general covariance is now automatically achieved by routine methods in the formulation of all seriously considered spacetime theories. The foundations of general relativity do not lie in one or other principle advanced by Einstein. Rather, they lie in the simple assertion that spacetime is semi-Riemannian, with gravity represented by its curvature and its metric tensor governed by the Einstein field equations.

Acknowledgments

The researching and writing of this review was supported by the National Science Foundation under Grant No. SBE-9121326. I thank the foundation for its support and also Jean Eisenstaedt, Don Howard, Al Janis and Carlo Rovelli for helpful discussion. I am also very grateful to Don Howard for assistance with German translations.

References

Adler R, Bazin M and Schiffer M 1977 *Introduction to General Relativity* 2nd ed (Tokyo: McGraw-Hill Kogakusha)
Aguirre M and Krause J 1991 *International Journal of Theoretical Physics* **30** 495–509
Anderson J L 1962 *Recent Developments in General Relativity* ed 'Editorial Committee' (New York: MacMillan)
—— 1964 *Gravitation and Relativity* eds H Y Chiu and W F Hoffmann (New York: Benjamin) ch 9
—— 1966 *Perspectives in Geometry and Relativity* ed B Hoffmann (Bloomington In: Indiana University Press) pp 16–27
—— 1967 *Principles of Relativity Physics* (New York: Academic)
—— 1971 *General Relativity and Gravitation* **2** 161–72
—— and Gautreau R 1969 *Phys. Rev.* **185** 1656–661
Arzeliès H 1961 *Relativité Généralisée Gravitation* (Paris: Gauthier-Villars)
Atwater H A 1974 *Introduction to General Relativity* (Oxford: Pergamon)
Bargmann V 1957 *Reviews of Modern Physics* **29** 161–174
Bartels A 1993 *Semantical Aspects of Spacetime Theories* ed U Majer and H J Schmidt (Mannheim: BI)
Bauer H 1922 *Mathematische Einführung in die Gravitationstheorie Einsteins* (Leipzig: Franz Deuticke)
Becquerel J 1922 *La Principe de la Relativité et la Théorie de la Gravitation* (Paris: Gauthier-Villars)
Bergmann P G 1942 *Introduction to the Theory of Relativity* (New York: Dover)
—— 1957 *Topics in the Theory of General Relativity, Lectures in Theoretical Physics* Notes by N A Wheeler (Waltham, MA: Brandeis University, Summer School, Institute of Physics)
—— 1961 *Rev. Mod. Phys.* **33** 510–4
—— 1962 The General Theory of Relativity *Encyclopedia of Physics* Vol. IV ed S. Flügge (Berlin: Springer)
Birkhoff G D 1927 *Relativity and Modern Physics* (Cambridge, MA: Harvard University Press)
Bishop R L and Goldberg S I 1968 *Tensor Analysis on Manifolds* (New York: Dover)
Bondi H 1959 *Rep. Prog. Phys.* **22** 97–120
—— 1979 *Einstein: A Centenary Volume* ed A P French (Cambridge, MA: Harvard University Press) pp 113–29
Borel E 1926 *Space and Time* transl A S Rappaport and J Dougall (New York: Dover)
Born M 1924 *Einstein's Theory of Relativity* revised edn 1962, ed G. Leibfried and W Biem (New York: Dover)
Bose S K 1980 *An Introduction to General Relativity* (New York: Wiley)
Bowler M G 1976 *Gravitation and Relativity* (Oxford: Pergamon)
Bridgman P 1949 *Albert Einstein: Philosopher Scientist* 2nd edn, ed P A Schilpp (New York: Tudor) pp 315–54
Buchdahl H A 1981 *Seventeen Simple Lectures on General Relativity* (New York: Wiley)
Bunge M 1967 *Foundations of Physics* (New York: Springer)
Butterfield J 1987 *International Studies in the Philosophy of Science* **2** 10–32
—— 1988 *Proceedings of the 1988 Biennial Meeting of the Philosophy of Science Association* vol. 2 ed A Fine and J Leplin (East Lansing, MI: Philosophy of Science Association) pp 65–81
—— 1989 *British J. Phil. Sci.* **40** 1–28

Carmeli M 1982 *Classical Fields: General Relativity and Gauge Theory* (New York: Wiley)
Carmichael R 1920 *The Theory of Relativity* 2nd edn (New York: Wiley)
Cartan E 1923 *Ann. Sci. ENS* **40** 325–412; 1924 *Ann. Sci. ENS* **41** 1–25; transl 1982 A Magnon and A Ashtekhar in *On Manifolds with an Affine connection and the Theory of General Relativity* (Naples: Bibliopolis)
—— 1927 *L'Ensignment Mathématique* **26** 200–225; 1952 *Oevres Complètes* Part 1, vol II (Paris: Gauthier-Villars) pp 841–66
Cartwright N and Hoefer C (forthcoming) *Philosophical Problems of the Internal and External Worlds: Essays on the Philosophy of Adolf Grünbaum* ed J Earman, A Janis, G Massey, N Rescher (Pittsburgh PA: University of Pittsburgh/University of Konstanz)
Charon J E 1963 *15 Leçons sur la Relativité Générale avec une Introduction au Calcul Tensoriel* (Geneva: Éditions René Kister)
Chazy J 1928 *La Théorie de la Relativité et la Mécanique Céleste* (Paris: Gauthier-Villars)
Clarke C 1979 *Elementary General Relativity* (London: Edward Arnold)
Cunningham E 1921 *Relativity: The Electron Theory and Gravitation* (London: Longmans Green & Co.)
d'Inverno R 1992 *Introducing Einstein's Relativity* (Oxford: Clarendon)
Darmois G 1927 *Les Équations de la Gravitation Einsteinienne: Mémorial des Sciences Mathématique, L'Académie des Sciences de Paris. Fasc. XXV* (Paris: Gauthier-Villars)
Davis W R 1970 *Classical Fields, Particles and the Theory of Relativity* (New York: Gordon and Breach)
De Donder T 1921 *La Gravifique Einsteinienne* (Paris: Gauthier-Villars)
—— 1925 *Introduction à la Gravifiqué Einsteinienne: Mémorial des Sciences Mathémeatique, L'Académie des Sciences de Paris Fasc. VIII* (Paris: Gauthier-Villars)
De Felice F and Clarke C J S 1990 *Relativity on Curved Manifolds* (Cambridge: Cambridge University Press)
De Sitter W (1916) *Mon. Not. R. Astron. Soc.* **76** 699–728; 1916 **77** 155–184; 1917 **78** 3–28
Dirac P A M 1975 *General Theory of Relativity* (New York: Wiley)
Earman J 1974 *Found. Phys.* **4** 267–89
—— 1989 *World Enough and Spacetime: Absolute versus Relational Theories of Space and Time* (Cambridge, MA: Bradford/MIT Press)
Earman J and Friedman M 1973 *Phil. Sci.* **40** 329–59
Earman J and Norton J D 1987 *British J. Phil. Sci.* **38** 515–25
Eddington A S 1920 *Report on the Relativity Theory of Gravitation* 2nd ed (London: Fleetway)
—— 1924 *The Mathematical Theory of Relativity* 2nd edn (1952) (Cambridge: Cambridge University Press)
Ehlers J 1971 *Proc. Int. School of Physics 'Enrico Fermi' Course XLVII* ed B K Sachs (New York: Academic) pp 1–70
Einstein A 1905 *Ann. Phys.* **17** 891–921 (transl Lorentz 1923 pp 37–65)
—— 1907 *Jahrbuch der Radioaktivität und Elektronik* **4** 411–462; 1908 **5** 98–99
—— 1911 *Ann. Phys.* **35** 898–908 (transl Lorentz 1923, pp 99–108)
—— 1912a *Ann. Phys.* **38** 355–69
—— 1912b *Ann. Phys.* **38** 443–58
—— 1912c *Vierteljahrsschrift für gerichtliche Medizin und öffentliches Sanitätswesen* **44** 37–40
—— 1912d *Ann. Phys.* **38** 1059–64
—— 1913 *Physikalische Zeitschrift* **14** 1249–62
—— 1914 *Preussische Akademie der Wissenschaften (Berlin), Sitzungsberichte* 1030–1085
—— 1914a *Scientia* **15** 337–48.
—— 1915 *Preussische Akademie der Wissenschaften (Berlin), Sitzungsberichte* 778–86, 799–86, 831–9, 844–7
—— 1916 *Ann. Phys.* **49** 769–822 (translated 1923 Lorentz *et al*, pp 111–164 (without p 769))
—— 1916a *Ann. Phys.* **51** 639–642
—— 1917 *Preussische Akademie der Wissenschaften (Berlin), Sitzungsberichte* 142–152 (transl 1923, Lorentz *et al* pp 175–88)
—— 1917a *Relativity: the Special and the General Theory* (transl 1977 R W Lawson (London: Methuen))
—— 1918 *Ann. Phys.* **55** 240–44
—— 1918a *Naturwissenschaften* **6** 697–702
—— 1919 *Preussische Akademie der Wissenschaften, Sitzungsberichte* 349–56 (transl 1923 Lorentz *et al* pp 191–8)
—— 1922 *Sidelights on Relativity* transl G B Jeffrey and W Perrett (New York: Dutton) (Reprinted 1983 New York: Dover)
—— 1922a *The Meaning of Relativity* 5th edn (Princeton NJ: Princeton University Press)
—— 1924 *Schweizerische Naturforschende Gesellschaft, Verhandlungen* **105** 85–93
—— 1927 'The Mechanics of Newton and Their Influence on the Development of Theoretical Physics *Ideas and Opinions* (New York: Bonanza) pp 253–61

—— 1933 'On the Methods of Theoretical Physics' *Ideas and Opinions* (New York: Bonanza) pp 270–6
—— 1940 The Fundaments of Theoretical Physics *Ideas and Opinions* (New York: Bonanza) pp 323–35
—— 1949 *Autobiographical Notes* (La Salle, IL: Open Court)
—— 1950 'On the Generalized Theory of Gravitation' *Ideas and Opinions* (New York: Bonanza) pp 341–356
—— 1952 *Relativity: the Special and the General Theory* transl R W Lawson. (London: Methuen) Appendix V
Einstein A and de Sitter W 1932 *Proc. Nat. Acad. Sci.* (Washington) **18** 213–14
Einstein A and Grossmann M 1913 *Entwurf einer verallgemeinerten Relativitätstheorie und einer Theorie der Gravitation* (Leibzig: Teubner, separatum); addendum 1913 Einstein *Zeitschrift für Mathematik und Physik* **63** 225–61
Eisenstaedt J 1986 *Archive for History of Exact Sciences* **35** 115–185
—— 1989 *Einstein and the History of General Relativity: Einstein Studies*, vol 1 eds D Howard and J Stachel (Boston: Birkhäuser) pp 277–92
Ellis G F R and Williams R M 1988 *Flat and Curved Space-Times* (Oxford: Clarendon)
Falk G and Ruppel W 1975 *Mechanik Relativität Gravitation: Ein Lehrbuch* 2nd edn (Berlin: Springer)
Fock V 1956 *Helv. Phys. Acta*. Supplementum IV, pp 239–243
—— 1957 *Reviews of Modern Physics* **29** 325–333
—— 1959 *The Theory of Space Time and Gravitation* transl N Kemmer (New York: Pergamon)
—— 1974 *Die Stellung des Copernicanischen Systems im Ideenkreis der Einsteinschen Gravitationstheorie* (Leipzig: Karl Marx Universität)
Fokker A D 1965 *Time and Space, Weight and Inertia: A Chronogeometrical Introduction to Einstein's Theory* transl D Bijl; transl and ed D Field (Oxford: Pergamon)
Foster J and Nightingale J D 1979 *A Short Course in General Relativity* (London: Longman)
Frankel T 1979 *Gravitational Curvature: An Introduction to Einstein's Theory* (San Francisco: Freeman)
Freundlich E 1919 *The Foundations of Einstein's Theory of Gravitation* transl H L Brose (London: Methuen)
Friedman M 1973 *Space, Time and Geometry* ed P Suppes (Dordrecht: Reidel) pp 296–320
—— 1983 *Foundations of Space-Time Theories: Relativistic Physics and the Philosophy of Science* (Princeton, NJ: Princeton University Press)
Friedrichs K 1927 *Mathematische Annalen* **98** 566–75
Gorelik G (forthcoming) *The Attraction of Gravitation: New Studies in History of General Relativity* ed J Earman, M Janssen and J D Norton (Boston: Birkhäuser)
Graves J C 1971 *The Conceptual Foundations of Contemporary Relativity Theory* (Cambridge, MA: MIT Press)
Havas P 1964 *Rev. Mod. Phys.* **36** 938–65
Hawking S and Ellis G F R 1973 *The Large Scale Structure of Space-time* (Cambridge: Cambridge University Press)
Hilbert D 1915 *Akademie der Wissenschaften, Göttingen, Nachrichten* pp 395–407; 1917 *Akademie der Wissenschaften, Göttingen, Nachrichten* pp 53–76
Hiskes A L D 1984 *Found. Phys.* **14** 307–32
Hoffmann B 1932 *Rev. Mod. Phys.* **4** 173–204
Howard D 1992 *Studies in the History of General Relativity* ed J Eisenstaedt and A Kox (Boston: Birkhäuser) pp 154–243
Howard D and Norton J D (forthcoming) *The Attraction of Gravitation: New Studies in History of General Relativity* ed J Earman, M Janssen and J D Norton (Boston: Birkhäuser)
Hughston L P and Tod K P 1990 *An Introduction to General Relativity* (Cambridge: Cambridge University Press)
Jones R 1981 *After Einstein* ed P Barker and C G Shugart (Memphis, TN: Memphis State University Press)
Jordan P 1955 *Schwerkraft und Weltall* (Braunschweig: Vieweg)
Kenyon I R 1990 *General Relativity* (Oxford: Oxford University Press)
Kerszberg P 1989 *Einstein and the History of General Relativity* ed D Howard and J Stachel (Boston: Birkhäuser) pp 325–66
Klein F 1872 *Gesammelte Mathematische Abhnadlungen*. vol 1, ed R Fricke and A Ostrowski (Berlin: Springer) pp 460–497
Kopczynski W and Trautman A 1992 *Spacetime and Gravitation* (Chichester: Wiley)
Kopff A 1923 *Grundzüge der Einsteinschen Relativitätstheorie* (Leipzig: Hirzel)
Kottler F 1922 *Encyklopädie der mathematischen wissenschaften mit Einschluss an ihrer Auwendung* vol VI, part 2, 22a *Astronomie* ed K Schwertzschild, S Oppenheim and W v Dyke (Leipzig: Teubner) pp 159–237
Kratzer A 1956 *Relativitätstheorie* (Münster: Aschendorfsche)
Kretschmann E 1915 *Ann. Phys.* **48** 907–982
—— 1917 *Ann. Phys.* **53** 575–614
Kuchar K V 1988 *Highlights in Gravitation and Cosmology* (Cambridge: Cambridge University Press)

Landau L D and Lifshitz E M 1951 *The Classical Theory of Fields* transl M Hamermesh (Oxford: Pergamon)
Laue M 1911 *Das Relativitätsprinzip* (Braunschweig: Vieweg)
—— 1921 *Die Relativitätstheorie. Vol 2: Die allgemeine Relativitätstheorie und Einsteins Lehre von der Schwerkraft* (Braunschweig: Vieweg)
Lawden D F 1962 *An Introduction to Tensor Calculus and Relativity* (London: Methuen)
Lenard P 1921 *Über Relativitätsprinzip, Äther, Gravitation* 3rd edn (Leipzig: Hirzel)
Levi-Civita T 1926 *The Absolute Differential Calculus* transl M Long (New York: Dover)
Levinson H C and Zeisler E B 1929 *The Law of Gravitation in Relativity* (Chicago: University of Chicago Press) 2nd edn
Lorentz H A *et al* 1923 *Principle of Relativity* (London: Methuen); (1952 New York: Dover)
Mach E 1893 *Science of Mechanics* transl 1960 T J McCormack (La Salle, IL: Open Court)
Malament D 1986 *From Quarks to Quasars: Philosophical Problems of Modern Physics* ed R G Colodny (Pittsburgh, PA: University of Pittsburgh Press) pp 181–201
Martin J L 1988 *General Relativity: A Guide to its Consequences for Gravity and Cosmology* (Chichester: Ellis Horwood)
Mashoon B 1986 *Found. Phys.* **16** 619–35.
Maudlin T 1988 *Proc. 1988 Biennial Meeting of the Philosophy of Science Association vol. 2*, ed A Fine and J Leplin (East Lansing, MI: Philosophy of Science Association) pp 82–91
—— 1990 *Stud. Hist. Philos. Sci.* **21** 531–61
Mavridès S 1973 *L'Univers Relativiste* (Paris: Masson)
McVittie G C 1965 *General Relativity and Cosmology* (Urbana IL: University of Illinois Press)
Mehra J 1974 *Einstein, Hilbert and The Theory of Gravitation* (Dordrecht: Reidel)
Minkowski H 1908 *Königlichen Gesellschaft der Wissenschaften zu Göttingen, Mathematische-Physikalische Klasse, Nachrichten*, 53–111
—— 1909 *Physikalische Zeitschrift* **10** 104–111 (transl Lorentz 1923, pp 75–91)
Misner C W, Thorne K S and Wheeler J A 1973 *Gravitation* (San Francisco: Freeman)
Moller C 1952 *The Theory of Relativity* (Oxford: Clarendon)
Mundy B 1992 *Proc. 1992 Biennial Meeting of the Philosophy of Science Association* vol. 1 ed D Hull, M Forbes and K Okruhlik (East Lansing, MI: Philosophy of Science Association) pp 515–527
Newton I 1687 *Mathematical Principle of Natural Philosophy* transl 1934 F Cajori (Berkeley, CA: University of Los Angeles Press)
North J D 1965 *The Measure of the Universe: A History of Modern Cosmology* (Oxford: Clarendon)
Norton J D 1984 *Historical Studies in the Physical Sciences* **14** 253–316; reprinted in *Einstein and the History of General Relativity: Einstein Studies*, Vol 1 eds D Howard and J Stachel (Boston: Birkhäuser, 1989) pp 101–159
—— 1985 *Studies in History and Philosophy of Science* **16** 203–246; reprinted in *Einstein and the History of General Relativity: Einstein Studies*, Vol 1 eds D Howard and J Stachel (Boston: Birkhäuser, 1989) pp 3–47.
—— 1987 *Measurement, Realism and Objectivity* ed J Forge (Dordrecht: Reidel) pp 153–188
—— 1988 *Proceedings of the 1988 Biennial Meeting of the Philosophy of Science Association Vol. 2* eds A Fine and J Leplin (East Lansing, MI: Philosophy of Science Assoc., 1989) pp 56–64
—— 1989 *Found Phys* **19** 1215–1263
—— 1992 *Studies in the History of General Relativity: Einstein Studies, Vol. 3* J Eisenstaedt and A Kox eds (Boston: Birkhauser) pp 281–315
—— 1992a *Introduction to the Philosophy of Science* M H Salmon *et al* (Englewood Cliffs, NJ: Prentice-Hall) pp 179–231
—— 1993 *Semantical Aspects of Spacetime Theories* ed U Majer and H J Schmidt (Mannheim: BI)
Ohanian H C 1976 *Gravitation and Spacetime* (New York: Norton)
Page L 1920 *The Principle of General Relativity and Einstein's Theory of Gravitation* (New Haven, CT: Connecticut Academy of Arts and Sciences)
Painlevé P 1921 *C. R. Acad. Sci. Paris* **173** 873–887
Papapetrou A 1974 *Lectures on General Relativity* (Dordrecht: Reidel)
Pathria R K 1974 *The Theory of Relativity* 2nd edn (Oxford: Pergamon)
Pauli W 1921 *Encyklopädie der mathematischen Wissenschaften, mit Einschluss an ihrer Anwendung*. vol 5 *Physik*, Part 2 ed A Sommerfeld (Leipzig: Teubner) pp 539–775; 1958 *Theory of Relativity* transl G Field (London: Pergamon)
Post E J 1967 *Delaware Seminar in the Foundations of Physics* ed M Bunge (New York: Springer) pp 103–123
Prasanna A R 1971 *Lectures on General Relativity and Cosmology, Matscience Report No 71* (Madras: Institute of Mathematical Sciences)
Raine D J and Heller M 1981 *The Science of Space-Time* (Tucson, AZ: Pachart Publishing House)

Rainich G Y 1950 *Mathematics of Relativity* (New York: Wiley)
Ray C 1987 *The Evolution of Relativity* (Bristol and Philadelphia: Adam Hilger)
Reichenbach H 1924 *Axiomatization of the Theory of Relativity* transl 1969 M Reichenbach (Berkeley, CA: University of California Press, 1969)
Reinhardt M *Z Naturf* **28a** 529–37
Ricci G and Levi-Civita T 1901 *Math. Ann.* **54** 125–201; reprinted in Levi-Civita 1954 *Opere Matematiche*, vol. 1 (Bologna) pp 479–559
Rindler W 1969 *Essential Relativity: Special, General, and Cosmological* (New York: Van Nostrand-Reinhold)
Robertson H P and Noonan T W 1968 *Relativity and Cosmology* (Philadelphia, PA: Saunders)
Roll P G, Krotkov R and Dicke R H 1964 *Ann. Phys.* **26** 442–517
Rosser W G V 1964 *An Introduction to the Theory of Relativity* (London: Butterworths)
Rovelli C 1991 *Class. Quantum Grav.* **8** 297–316
Rynasiewicz R (forthcoming (a)) 'Rings, Holes and Substantivalism: On the Program of Leibniz Algebra' *Phil. Sci.*
Rynasiewicz R (forthcoming (b)) The Lessons of the Hole Argument
Ryckman T A 1992 *Stud. History Phil. Sci.* **23** 471–497
Sachs R K and Wu H 1977 *General Relativity for Mathematicians* (New York: Springer)
Scheibe E 1981 *Scientific Philosophy Today* eds J Agassi and R S Cohen (Dordrecht: Reidel) pp 311–31; reprinted 1983 *Space, Time, and Mechanics* ed D Mayr and G Süssman (Dordrecht: Reidel) pp 125–47
—— 1991 *Causality, Method, and Modality* ed G G Brittan Jr (Kluwer-Academic) pp 23–40
Schild A 1967 *Relativity Theory and Astrophysics Vol. 1 Relativity and Cosmology* ed J Ehlers (Providence, RI: American Mathematical Society) pp 1–116
Schlick M 1920 *Space and Time in Contemporary Physics* transl H L Brose (New York: Oxford University Press)
Schrödinger E 1950 *Space-Time Structure* (Cambridge: Cambridge University Press)
Schutz B F 1985 *A First Course in General Relativity* (Cambridge: Cambridge University Press)
Sesmat A 1937 *Systèmes de Référence et Mouvements (Physique Relativiste)* (Paris: Hermann)
Sexl R U and Urbantke H K 1983 *Gravitation und Kosmologie: Eine Einführung in die Allgemeine Relativitätstheorie* (Mannheim: Bibliographisches Institut)
Silberstein L 1922 *The Theory of General Relativity and Gravitation* (University of Toronto Press)
Skinner R 1969 *Relativity* (Waltham, MA: Blaisdell Publishing Co)
Sommerfeld A 1910 *Ann. Phys.* **32** 749–776; *Ann. Phys.* **33** 649–89
Stachel J (forthcoming) *Philosophical Problems of the Internal and External Worlds: Essays on the Philosophy of Adolf Grünbaum* eds J Earman, A Janis, G Massey and N Rescher (Pittsburgh, PA: University of Pittsburgh/University of Konstanz)
—— 1980 'Einstein's Search for General Covariance' *9th Int. Conf. on General Relativity and Gravitation, Jena*; reprinted in *Einstein and the History of General Relativity: Einstein Studies*, vol. 1 eds D Howard and J Stachel (Boston: Birkhäuser, 1989) pp 63–100
—— 1980a *General Relativity and Gravitation: A Hundred Years After the Birth of Einstein* ed A Held (New York: Plenum); reprinted in *Einstein and the History of General Relativity: Einstein Studies*, Vol. 1 D Howard and J Stachel eds (Boston: Birkhäuser, 1989) pp 48–62
—— 1986 *Proc. 4th Marcel Grossmann Meeting on General Relativity* ed R Ruffini (Amsterdam: Elsevier) pp 1857–1862
—— 1992 *Studies in the History of General Relativity: Einstein Studies*, Vol. 3 eds J Eisenstaedt and A Kox (Boston: Birkhäuser) pp 407–418
Stephani H 1977 *General Relativity: An Introduction for the Theory of the Gravitational Field* ed J Stewart, transl M Pollock and J Stewart (Cambridge: Cambridge University Press)
Stewart J 1990 *Advanced General Relativity* (Cambridge: Cambridge University Press)
Straumann N 1984 *General Relativity and Astrophysics* (Berlin: Springer)
Synge J L 1960 *Relativity: The General Theory* (Amsterdam: North-Holland)
—— 1964 *Relativity, Groups and Topology* eds C De Witt and B De Witt (New York: Gordon and Breach) pp 4–90
—— 1966 *Perspectives in Geometry and Relativity* ed B Hoffmann (Bloomington, IN: Indiana University Press) pp 7–15
Szekeres G 1955 *Phys. Rev.* **97** 212
Teller P (forthcoming) Substance, Relations and Arguments about the Nature of Space-Time *Phil. Rev.*
Thirring H 1922 *Die Idee der Relativitätstheorie* (Berlin: Springer)
Thirring W 1979 *A Course in Mathematical Physics Vol. 2* transl E M Harrell (New York: Springer)
Tolman R C 1934 *Relativity, Thermodynamics and Cosmology* (New York: Dover)
Tonnelat M 1959 *The Principles of Electromagnetic Theory and of Relativity* revised edn 1966, transl A J Knoedel (Dordrecht: Reidel)

Törnebohm H 1952 *A Logical Analysis of the Theory of Relativity* (Stockholm: Almqvist & Wiksell)
Toretti R 1983 *Relativity and Geometry* (Oxford: Pergamon)
—— 1984 *British J. Phil. Sci.* 35 280–292
Trautman A 1964 *Lectures on General Relativity: Brandeis Summer Institute in Theoretical Physics Vol. 1 1964* (Englewood Cliffs, NJ: Prentice Hall, 1965) pp 1–248
—— 1966 *Perspectives in Geometry and Relativity: Essays in Honor of Václav Hlavaty* ed B Hoffmann (Bloomington, IN: Indiana University Press) pp 413–25
—— 1973 *The Physicist's Conception of Nature* ed J Mehra (Dordrecht: Reidel) pp 178–98
Treder H-J, von Borzeszkowski H-H, van der Merwe A and Yourgrau W 1980 *Fundamental Principles of General Relativity Theories: Local and Global Aspects of Gravitation and Cosmology* (New York: Plenum, 1980)
Veblen O and Whitehead J H C 1932 *The Foundations of Differential Geometry* (Cambridge: Cambridge University Press)
Vladimirov Yu, Mitskievisch N and Horsky J 1987 *Space, Time and Gravitation* transl A G Zilberman, ed F I Fedorov (Moscow: Mir)
Wald R 1984 *General Relativity* (Chicago, IL: University of Chicago Press)
Wasserman R H 1992 *Tensors and Manifolds with Applications to Mechanics and Relativity* (New York: Oxford University Press)
Weber J 1961 *General Relativity and Gravitational Waves* (New York: Interscience)
Weinberg S 1972 *Gravitation and Cosmology: Principles and Applications of the General Theory of Relativity* (New York: Wiley)
Weyl H 1921 *Space Time Matter* transl 1952 H L Brose (New York: Dover)
Whitakker E T 1951 *A History of Theories of Aether and Electricity: Vol. 1 The Classical Theories, Vol. 2 The Modern Theories 1900–1926* (New York: Dover)
Yilmaz H 1965 *Introduction to the Theory of Relativity and Principles of Modern Physics* (New York: Blaidsdell)
Zahar E 1989 *Einstein's Revolution: A Study in Heuristic* (La Salle, IL: Open Court)
Zatzkis H 1955 in *Fundamental Formulas of Physics* ed D H Menzel (New York: Dover, 1960) Vol. 2, ch 7