LIVING REVIEWS
in relativity

The Hole Argument and Some Physical and Philosophical Implications

John Stachel

Center for Einstein Studies,
Boston University
745 Commonwealth Avenue
Boston, MA 02215
U.S.A.
email: john.stachel@gmail.com

Abstract

This is a historical-critical study of the hole argument, concentrating on the interface between historical, philosophical and physical issues. Although it includes a review of its history, its primary aim is a discussion of the contemporary implications of the hole argument for physical theories based on dynamical, background-independent space-time structures.

The historical review includes Einstein's formulations of the hole argument, Kretschmann's critique, as well as Hilbert's reformulation and Darmois' formulation of the general-relativistic Cauchy problem. The 1970s saw a revival of interest in the hole argument, growing out of attempts to answer the question: Why did three years elapse between Einstein's adoption of the metric tensor to represent the gravitational field and his adoption of the Einstein field equations?

The main part presents some modern mathematical versions of the hole argument, including both coordinate-dependent and coordinate-independent definitions of covariance and general covariance; and the fiber bundle formulation of both natural and gauge natural theories. By abstraction from continuity and differentiability, these formulations can be extended from differentiable manifolds to any set; and the concepts of permutability and general permutability applied to theories based on relations between the elements of a set, such as elementary particle theories.

We are closing with an overview of current discussions of philosophical and physical implications of the hole argument.

Keywords: General relativity, History of science, Philosophy of science

Imprint / Terms of Use

Because a *Living Reviews* article can evolve over time, we recommend to cite the article as follows:

The date given as <date> then uniquely identifies the version of the article you are referring to.

Article Revisions

Living Reviews supports two ways of keeping its articles up-to-date:

Fast-track revision. A fast-track revision provides the author with the opportunity to add short notices of current research results, trends and developments, or important publications to the article. A fast-track revision is refereed by the responsible subject editor. If an article has undergone a fast-track revision, a summary of changes will be listed here.

Major update. A major update will include substantial changes and additions and is subject to full external refereeing. It is published with a new publication number.

For detailed documentation of an article's evolution, please refer to the history document of the article's online version at http://www.livingreviews.org/lrr-2014-1.

Contents

1 Introduction

"Hole argument" is the English translation of the German phrase "Lochbetrachtung," used by Albert Einstein to describe his argument against the possibility of generally-covariant equations for the gravitational field, developed in late 1913 and accepted until late 1915. Einstein realized the desirability of general covariance, and showed that it was easily implemented for the rest of physics; but the hole argument purported to show why it could not be demanded of the gravitational field equations he was trying to formulate for the metric tensor.[1]

This article is a historical-critical study, in Ernst Mach's sense.[2] It includes a review of the literature on the hole argument that concentrates on the interface between historical, philosophical and physical approaches. Although recounting the history of the hole argument, the primary purpose is to discuss its contemporary significance – in both physics and philosophy – for the study of space-time structures. Like Mach, while presenting various other viewpoints, I have not hesitated to advocate my own. In physics, I believe the main lesson of the hole argument is that any future fundamental theory, such as some version of "quantum gravity," should be background independent, with basic elements obeying the principle of maximal permutability. In the philosophy of space-time, this leads me to advocate a "third way" that I call dynamic structural realism, which differs from both the traditional absolutist and relationalist positions.

1.1 Why should we care?

One of the most crucial developments in theoretical physics was the move from theories dependent on fixed, non-dynamical background space-time structures to background-independent theories, in which the space-time structures themselves are dynamical entities. This move began in 1915 when Einstein stated the case against his earlier hole argument. Even today, many physicists and philosophers do not fully understand the significance of this development, let alone accept it in practice.

So it is of more than historical interest for physicists and philosophers of science to understand what initially motivated this move, as well as the later developments stemming from it. Einstein's starting point was the search for a generalization of the special theory that would include gravitation. He quickly realized that the equivalence principle compelled the abandonment of the privileged role of inertial (i.e., non-accelerated) frames of reference, and started to investigate the widest class of accelerated frames that would be physically acceptable. His first impulse was to allow all possible frames of reference; since he identified frames of reference and coordinate systems, this choice corresponds mathematically to a generally-covariant theory. But he soon developed an argument – the hole argument – purporting to show that generally-covariant equations for the metric tensor are incompatible with his concept of causality for the gravitational field. The argument hinged on his tacit assumption that the points of space-time are inherently individuated, quite apart from the nature of the metric tensor field at these points. Only two years later, after other reasons compelled him to reconsider general covariance, did Einstein finally recognize the way out of his dilemma: One must assume that, in an empty region of space-time, the points have no inherent individuating properties – nor indeed are there any spatio-temporal relations between them – that do not depend on the presence of some metric tensor field.

Thus, general relativity became the first fully dynamical, background-independent space-time theory. Without some knowledge of this historical background, it is difficult to fully appreciate

---

[1] For a historical review of the disputes over general covariance, see Norton (1993). Throughout this paper, unless there is some indication to the contrary, "metric" is to be understood to mean "pseudo-metric tensor field with Lorentz signature;" and "connection" is the be understood to mean "linear affine connection with vanishing torsion."

[2] See Mach (1988, 1986).

either the modern significance of the hole argument, or the compelling physical motives for the requirement of background independence.

1.2   Summary: Where we are headed

Einstein's starting point in his search for a theory of gravitation was the theory we now call special relativity. From a contemporary viewpoint, its most important feature is that it has two fixed, kinematical space-time structures – the chrono-geometry embodied in the Minkowski metric tensor field and the inertial field embodied in the associated flat affine connection – both of which are invariant under the ten-parameter Lie group now called the Poincaré or inhomogenous Lorentz group.[3]

In Minkowski space-time, all dynamical theories must be based on geometric objects that form a representation (or more generally, a realization) of this group. There is a preferred class of spatial frames of reference in Minkowski space-time, the inertial frames. Einstein had shown how to define a class of physically preferred coordinate systems for each inertial frame of reference; in particular, he defined a clock synchronization procedure that provides a preferred global time for each frame. This enabled him to show how the principle of relativity of all inertial frames could be reconciled with the universal properties of light propagation in vacuum.

The lesson he drew was the need to find a physical interpretation of the coordinates associated with an inertial frame of reference – a lesson that had to be painfully unlearned in his search for a generalized theory of relativity. In large part, the history of the hole argument is the story of that unlearning process. The end result was the formulation of the general theory of relativity, the first background-independent physical theory turning all space-time structures into dynamical fields. This was such a revolutionary break with all previous physical theories, in which space-time structures constitute a fixed, non-dynamical background, that its ultimate significance is still debated by physicists.[4]

Understanding the hole argument in both its historical and contemporary aspects can help to clarify the issues at stake in this debate. The basic issue can be stated as follows: Given a physical theory, when should an equivalence class of mathematically distinct models of the theory be identified as corresponding to single, unique physical model? The hole argument shows that, for any theory defined by a set of generally-covariant field equations, the only way to make physical sense of the theory is to assume that the entire equivalence class of diffeomorphically-related solutions to the field equations represent a single physical solution. As will be seen later, mathematically this result can best be stated in the language of natural bundles.

But a similar result holds for the even broader class of all gauge-invariant field theories, notably Yang–Mills theories: an equivalence class of gauge-related models of any such theory must be physically identified. Mathematically, broadening the question in this way requires the language of gauge-natural bundles. General relativity itself may also be treated by the use of gauge-natural bundle techniques: its similarities to and differences from gauge theories of the Yang–Mills type will also be discussed.

This move to natural and gauge-natural formulations of field theories also has important implications for the philosophy of space and time. The old conflict between absolute and relational interpretations of space and then space-time has been renewed on this new ground. But I shall argue that this reformulation of the question suggests a third position, around which a consensus is forming. This position has been given various names, but I prefer dynamic structural realism.

---

[3] This the semi-direct product of the homogeneous Lorentz group $SO(3,1)$ and translation group T_4.

[4] Indeed, even the then-future quantum theories were based on such fixed backgrounds: Galilei–Newtonian space-time for non-relativistic quantum mechanics, and Minkowski space-time for quantum field theory. Much of the debate about the form of an acceptable theory of quantum gravity that somehow reconciles quantum field theory with general relativity revolves around this question. See Section 6.4.

1.3 Outline of the article

Sections 2.1 – 2.5 recount the developments leading up to Einstein's adoption of the hole argument against general covariance in 1913, how it misled him for over two years, the reasons for his rejection of it in late 1915, and its replacement by the point-coincidence argument for general covariance.

Section 2.6 discusses Kretschmann's 1917 critique of the concept of general covariance and Einstein's 1918 reply; decades later this debate led Komar to propose the use of what are now called Kretschmann–Komar coordinates as a way of resolving the hole argument.

Finally, Section 2.7 discusses Hilbert's 1917 reformulation of the hole argument: He replaced the four-dimensional hole in space-time with a space-like hypersurface, on which he posed an initial value problem for the field equations; this was the first step in a series of developments culminating a decade later in a fully satisfactory formulation by Darmois of the general-relativistic Cauchy problem.

Section 3 discusses the revival of interest in the hole argument in the 1970s, which grew out of an attempt to answer a historical question: Why did three years (1912 – 1915) elapse between Einstein's adoption of the metric tensor to represent the gravitational field and his adoption of what are now called the Einstein equations for this field? Some highlights of this discussion are recalled, from the post-World War II revival of interest in general relativity up to the present.

Section 4 presents a modern version of the hole argument in general relativity, and its generalization from metric theories of gravitation to gauge-natural field theories. By abstraction from continuity and differentiability, a formulation general enough to include theories with discrete fundamental elements is obtained. The concept of general covariance of a field theory is similarly extended to general permutability, a concept wide enough to include theories based on relations between the elements of any set.

Sections 5 and 6 focus on current discussions of philosophical and physical implications of the hole argument, respectively; no attempt is made to rigidly separate issues that overlap both areas.

Section 5 discusses such issues as: the range of applicability of the hole argument, the correct mathematical definition of general covariance and its physical significance, the controversy between relationalists and substantivalists in discussions of space-time structures. The arguments of Earman, Pooley and Stachel are reviewed, and their convergence on a third alternative, which I call dynamic structural realism, is stressed.

Section 6 discusses such issues as partially background-independent theories, including mini- and midi-solutions to the Einstein field equations; the reformulation of general relativity as a gauge natural theory; and some implications of the hole argument for attempts to formulate a quantum theory of gravity.

2 Early History

2.1 From the special theory to the search for a theory of gravity

Einstein attributed his success in formulating the special theory in 1905 in no small measure to his insistence on defining coordinate systems that allowed him to attach physical significance to spatial and temporal coordinate intervals:

> The theory to be developed – like every other electrodynamics – is based on the kinematics of rigid bodies, since the propositions of any such theory concern relations between rigid bodies (coordinate systems), clocks and electromagnetic processes. Not taking this into account insufficiently is the root of the difficulties, with which the electrodynamics of moving bodies currently has to contend (Einstein, 1905, my translation).

His subsequent attempt to include gravitation in his theory focused on the equality of gravitational and inertial mass, and led him to adopt the *equivalence principle*: Inertia and gravitation are "wesensgleich" (the same in essence), and must be represented by a single *inertio-gravitational field*.[5] The distinction between the two is not absolute (i.e., frame independent), but depends on the frame of reference adopted.

In particular, he noted that a linearly accelerated (rigid) frame of reference in a space-time without a gravitational field is physically equivalent to an inertial frame of reference, in which there is a uniform, constant gravitational field: both result in equal acceleration of bodies moving relative to their respective frames.

He concluded that, in order to include gravitation, one must go beyond the special theory, with its privileged role for inertial frames, and look for a generalized ("verallgemeinerte") theory of relativity. In the simplest case, linearly accelerated frames in Minkowski space, the usual time coordinate loses its direct physical significance; and in uniformly rotating frames, a global time cannot even be defined. In the latter case, the spatial coordinates also lose their direct significance: the measured spatial geometry is no longer flat.

> I soon saw that, according to the point of view about non-linear transformations required by the equivalence principle, the simple physical interpretation of the coordinates had to be abandoned. ... This recognition tormented me a great deal because for a long time I was not able to see just what are coordinates actually supposed to mean in physics? (Einstein 1933, translation from Stachel, 2007, p. 86).

The equivalence principle

> made it not only probable that the laws of nature must be invariant with respect to a more general group of transformations than the Lorentz group (extension of the principle of relativity), but also that this extension would lead to a more profound theory of the gravitational field. That this idea was correct in principle I never doubted in the least. But the difficulties in carrying it out seemed almost insuperable. First of all, elementary arguments showed that the transition to a wider group of transformations is incompatible with a direct physical interpretation of the space-time coordinates, which had paved the way for the special theory of relativity. Further, at the outset it was not clear how the enlarged group was to be chosen (Einstein, 1956, my translation).

---

[5] The use of this term here is anachronistic: It was introduced much later. Weyl was the first to recognize the fundamental significance of this concept; he called it a "Führungsfeld" (see Weyl, 1923), which was translated as "guiding field" (see Weyl, 1950).

2.2 From the equivalence principle to the metric tensor

Einstein first attempted to develop a theory of the gravitational field produced by a static source, still based on the idea of a scalar gravitational potential. His earlier work had led him to consider non-flat spaces; this work led him to consider non-flat space-times: He found that his equation of motion for a test particle in a static field can be derived from a variational principle:

$$\delta \int \left\{ [c(x,y,z)]^2 \, dt^2 - [dx^2 + dy^2 + dz^2] \right\}^{1/2} = 0 \,, \tag{1}$$

where he interpreted $c(x,y,z)$ as a spatially-variable speed of light.[6]

Already familiar with Minkowski's four-dimensional formulation of the special theory, he realized that this variational principle could be interpreted as the equation for a geodesic (i.e, an extremal) in a non-flat space-time with $ds = \left\{ [c(x,y,z)]^2 \, dt^2 - [dx^2 + dy^2 + dz^2] \right\}^{1/2}$ as its line element. By explicitly introducing the flat Minkowski pseudo-metric $\eta_{\mu\nu}$, this ds can be rewritten in an arbitrary coordinate system. Einstein then made a big leap: He generalized the geodesic equation using a non-flat Riemannian pseudo-metric $g_{\mu\nu}$, and assumed that it would still describe the path of a test particle in an arbitrary non-static gravitational field. The gravitational theory he was seeking must be based on such a non-flat metric, which should both:

a) determine the line element, $ds^2 = g_{\mu\nu} \, dx^\mu \, dx^\nu$ representing the *chrono-geometry* of space-time;

b) serve as the potentials for the *inertio-gravitational* field.

While a student at the Swiss Federal Polytechnic, Einstein had learned about Gauss' theory of non-flat surfaces, and realized he needed a four-dimensional generalization. His old classmate and new colleague, the mathematician Marcel Grossmann, told Einstein about Riemann's generalization of Gauss' theory and about the tensor calculus ("absolute differential calculus"), developed by Ricci and Levi-Civita to facilitate calculations in an arbitrary coordinate system.

Still identifying a coordinate system with a physical frame of reference, his goal of extending the principle of the relativity led Einstein to investigate the widest possible group of coordinate transformations. With Grossman's help, he succeeded in formulating the influence of the inertio-gravitational field on the rest of physics by putting these equations into a "generally covariant" form.

The one exception was the gravitational field equations, the problem to which they now turned. "General covariance" then meant covariance under arbitrary coordinate transformations,[7] so they turned to coordinate-invariant tensors formed from the metric. The concepts of covariant derivative and Riemann tensor were based on the theory of differential invariants, and lacked a simple geometrical interpretation.[8] Nevertheless, Einstein seriously considered the Ricci tensor, the only second rank contraction of the Riemann tensor, for use in the gravitational field equations. He tried, in linear approximation, setting it equal to the stress-energy-momentum tensor of the sources

---

[6] In coordinates adapted to the static situation. Again, these coordinates have no direct physical significance. The measured local speed of light does not vary.

[7] The modern definition of a differentiable manifold and the distinction between diffeomorphisms of the manifold and coordinate transformations in and between regions of $\mathbb{R}^n$ (the n-dimensional Cartesian space) had not been formulated (see Stachel, 2014).

[8] The concept of parallel transport was only developed after completion of the general theory and largely in response to it. The only available geometrical interpretation of the Riemann tensor was in terms of the Gaussian curvature of two-dimensional sections, which was of little help in relating the Riemann tensor to anything in gravitation theory. Its interpretation in terms of geodesic deviation, so easily related to gravitational tidal forces, again lay in the future.

of the gravitational field; and even realized that, in order to obtain consistency with the vanishing divergence of the source tensor, the Ricci tensor would have to be modified by a trace term.[9]

However, after coming so close to the final form of field equations of GR, he retreated. His earlier work on static fields led him to conclude that, in adapted coordinates, the spatial part of the metric tensor must remain flat [see Eq. (1)], which is easily shown to be incompatible with field equations based on the Ricci tensor. So, as he later put it, he abandoned these equations "with a heavy heart," and began to search for non-generally-covariant field equations.

2.3   From the metric tensor to the hole argument

Einstein soon developed a meta-argument against a generally covariant set of field equations for the metric tensor. Why did he formulate this argument in terms of a hole – a finite region of space-time devoid of all non-gravitational sources?[10] It was probably the influence of Mach's ideas. One of Einstein's main motivations in the search for a generalized theory of relativity was his interpretation of Mach's critique of Newtonian mechanics. According to Mach, space is not absolute: It does not have any inherent properties of its own, and its apparent influence on the motion of a body – as manifested in the law of inertia, for example – must result from an interaction between the moving body and all the rest of the matter in the universe. Mach suggested that the inertial behavior of matter in the region of empty space around us is the effect of all the matter in the universe that surrounds this region, or "hole."

When Einstein adopted the metric tensor as the representation of the potentials of the inertio-gravitational field, he interpreted Mach's idea as the requirement that metric field in such a "hole" be entirely determined by its sources – that is by the stress-energy tensor of the surrounding matter.[11]

The hole argument purports to show that, if the field equations are generally covariant, this requirement cannot be satisfied: Even if the field and all its sources outside of and on the boundary of the hole are fully specified, such equations cannot determine a unique field in the interior of the hole. Conclusion: the gravitational field equations cannot be generally covariant.

We present the argument here in Einstein's original coordinate-based formulation (see Section 4 for a modern, coordinate-free form).

Let the metric tensor be symbolized by the single letter G, and a given four-dimensional coordinate system by a single letter x. Suppose $G(x)$ is a solution to a set of field equations. If both the coordinates and components of the metric tensor are subject to the transformation

$$x \rightarrow x' = f(x) \tag{2}$$

from coordinates x to another coordinate system x', then $G'(x')$ represents the same solution in the new coordinate system. This presents no problem – Einstein is quite clear on this point.[12]

But if the field equations are *generally covariant*, Einstein noted, it follows that $G'(x)$ must also be a solution. He emphasized clearly that $G(x)$ and $G'(x)$ represent two mathematically distinct

---

[9] He would have to subtract one-half of the Ricci scalar times the metric tensor (see Einstein's Zurich Notebook, and the commentary on it in Renn, 2007).

[10] As we shall see later, it might have been more natural mathematically to formulate it in terms of an initial-value problem (see Section 2.7).

[11] In 1918 he named this requirement "Mach's principle": "The G-field is determined *completely* by the masses of bodies. Since according to results of the special theory of relativity, mass and energy are the same, and since energy is formally described by the symmetric energy tensor $(T_{\mu\nu})$, this therefore entails that the G-field be conditioned and determined by the energy tensor of matter." (Einstein, 1918)

[12] Contrary to the claim by Pais and others that he did not then understand coordinate transformations (see Section 3.1).

solutions in the *same* coordinate system.[13]

Now consider a "hole" H – a bounded, closed region of space-time, in which all non-gravitational sources of the field – represented by the stress-energy tensor $T_{\mu\nu}$ – vanish; and suppose the field $G(x)$ and all its sources $T_{\mu\nu}$ are specified everywhere outside the hole and on its boundary, together with any finite number of normal derivatives of field on the boundary. This still does not suffice to determine a unique mathematical solution inside the hole. For there are coordinate transformations $x \to x'$ that reduce to the identity outside of and on the boundary of H, together with all their derivatives up to any finite order; yet which differ from the identity inside H. Such coordinate transformations will leave $T_{\mu\nu}$ unchanged, and the resulting $G'(x)$ will still equal $G(x)$ outside of and on the boundary of H; but $G(x)$ will *differ* from $G'(x)$ inside H.

In short, if the field equations are generally covariant, then specification of the gravitational field together with its sources outside of and on the boundary of such a hole do not suffice to determine the field inside. Einstein concluded that generally-covariant field equations could not be used to describe the metric tensor field, and began a search for non-covariant field equations.

Now the question became: If Lorentz invariance is *too little* (equivalence principle) and general covariance is *too much* (hole argument), what is the widest possible group (if it is a group!) of coordinate transformations, under which one can demand the invariance of such equations? The hole argument is not valid for the inhomogeneous Lorentz (Poincaré) group; and Einstein concluded that the invariance group of the field equations should be extended only up to – but not including – the invariance group, for which the hole argument becomes valid in coordinates adapted to this group.

2.4 From the hole argument back to general covariance

Because of problems unrelated to the hole argument, in mid-1915 Einstein abandoned the search for a non-covariant theory of gravitation and returned to the Riemann tensor. After several intermediate steps, by November of that year he adopted the set of generally-covariant equations now known as the Einstein equations. His successful explanation of the anomalous perihelion advance of Mercury convinced him – and many others – of the profound significance of the resulting theory, known today as general relativity.

What about the hole argument? Einstein realized that, to avoid it, he had only to drop one of the premises that he had tacitly adopted: the assumption that the points of space-time in the hole are individuated independently of the metric field. If that assumption is dropped, it follows that when the metric is dragged-along by a coordinate transformation, all the individuating properties and relations of the points of space-time are dragged along too.[14] While $G'(x)$ does differ *mathematically* from $G(x)$ inside H, they are merely different representations of the same physical solution. Properly-specified conditions outside the hole *will suffice* to specify a unique *physical solution* inside the hole.

---

[13] To avoid future confusion, let me emphasize that the terminology introduced in Section 4 differs from that of Einstein, and indeed that of many contemporary accounts. In this review, the requirement that, if $G(x)$ is a solution to the field equations, then $G'(x)$ must also be a solution is called covariance. General covariance is the requirement that $G(x)$ and $G'(x)$ represent one and the same physical solution. Another difference is that Section 4 these conditions are formulated coordinate independently.

[14] See Iftime and Stachel (2006) for a modern treatment of the hole argument using the language of categories, functors and natural objects, and generalizing it to other covariant theories; and Section 4 of this review.

2.5   The point coincidence argument

In order to better illustrate the flaw in the hole argument, Einstein developed a counter-argument, the point-coincidence argument.[15] There are actually two versions of this argument, which have been called "the private" and "the public" one.[16] First I shall cite the private version. In letters to friends, Einstein explained why the argument no longer applies to general relativity:

> Everything in the hole argument was correct up to the final conclusion. It has no physical content if, with respect to the same coordinate system K, two different solutions $G(x)$ and $G'(x)$ exist [see Section 2.3]. To imagine two solutions simultaneously on the same manifold has no meaning, and indeed the system K has no physical reality. The hole argument is replaced by the following consideration. Nothing is physically real but the totality of space-time point coincidences. If, for example, all physical events were to be built up from the motions of material points alone, then the meetings of these points, i.e., the points of intersection of their world lines, would be the only real things, i.e., observable in principle. These points of intersection naturally are preserved during all [coordinate] transformations (and no new ones occur) if only certain uniqueness conditions are observed. It is therefore most natural to demand of the laws that they determine no more than the totality of space-time coincidences. From what has been said, this is already attained through the use of generally covariant equations (letter to Michele Besso, 3 January 1916, in Schulmann *et al.*, 1998, p. 235, translation cited from Stachel, 1989, p. 86).

Einstein's argument consists of three points; in modern language, they are:

1. If two metrics in their respective *different* coordinate systems differ only in that one is the carry-along of the other, then *physically* there is no distinction between them.

2. Generally covariant equations have the property that, given a solution, any *carry-along* of that solution in the *same* coordinate system is also a solution to these equations.

3. In the absence of a metric tensor field, a coordinate system on a differentiable manifold has no intrinsic significance.[17]

Note that points 1) and 2) were included in the original hole argument (see Section 2.3). Point 3) is the crucial new element. It follows from the three points that the entire equivalence class of carry-alongs of a given solution in the same coordinate system corresponds to one physical gravitational field. Thus, the hole argument fails.

As will be seen in Section 4, point 1) constitutes a coordinate-dependent version, applied to the metric tensor, of what I call the basic or trivial identity; point 2) constitutes the coordinate-dependent version, applied to the metric tensor, of my definition of covariant theories. I would apply the term generally covariant to the conclusion that an entire equivalence class of carry alongs corresponds to one physical solution. The coordinate-independent versions of all three concepts are obtained by substituting basis vectors for coordinates and diffeomorphisms for coordinate transformations.

Einstein's 1916 review paper presents the public version of the argument to justify the requirement that any physical theory be invariant under all coordinate transformations:

---

[15] See Stachel (1989), which introduced this name. There is evidence that discussions with Moritz Schlick influenced Einstein's formulation of this argument: See Engler and Renn (2013) and references therein.

[16] See Rynasiewicz (1999), Section 4, pp. 443–448.

[17] In a hole, that is. Naturally, if other physical fields are present in a region of the manifold, it is possible that they may be utilized to individuate the points.

> 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 measurements 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 at 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 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 systems of values are made to correspond 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 coordinates to others, that is to say, we arrive at the requirement of general covariance (Einstein, 1916, pp. 776–777, reprinted in Kox *et al.*, 1996, pp. 291–292).

In this version of the argument, there is no mention of dynamical equations or even of fields. Indeed, he proceeds to illustrate it with a version of the trivial identity applied to a system of particles, rather than fields, the model being any set of particle world lines, without any requirement that they satisfy equations of motion. Einstein also mentions "the requirement of general covariance;" but here it amounts basically to point 3) together with a generalization of point 1) to any objects used in a physical theory, whether or not they obey any field equations. It is essentially a coordinate-dependent version of the basic identity, extended from metrics to all geometric object fields of a certain type (see Section 4.2).

We see here the origins of two differing usages of the term "general covariance" – one involves the field equations, the other does not. This difference has occasioned some confusion in recent discussions of the hole argument. (see Section 5.2 for an example of their conflation). My use of the term always involves the field equations.

In later years, Einstein explicitly rejected any positivistic interpretation of the point coincidence argument. He wrote to Schlick:

> Generally considered, your presentation of the [point coincidence] argument does not correspond with my conception of it since I find your entire conception too positivistic, so to speak. Physics indeed provides relations between sense experiences, but only indirectly. For me, its essence is by no means exhaustively characterized by this assertion. Physics is an attempt at the conceptual construction of a model of the real world, as well as of its lawful structure. Indeed it must represent exactly the empirical relations between the sense experiences that are accessible to us; but only in this way is it linked to the latter (Einstein to Moritz Schlick, 28 November 1930; cited from Engler and Renn, 2013, p. 18, my translation).

2.6   From general covariance to Kretschmann's critique

In 1915, even before Einstein completed the general theory of relativity, Erich Kretschmann (1915a,b) had undertaken an investigation that led him to a version of the trivial identity. Kretschmann (1917) uses Einstein's public point coincidence argument to conclude that any theory could be put into a form satisfying Einstein's principle of general covariance. Einstein (1918) concedes the point, but argued, not very successfully,[18] that an added criterion of simplicity gives the principle a heuristic significance. Evidently, he was not himself clear on the difference between his two arguments: While the public point coincidence argument does not provide a criterion for singling out theories, the criterion of general covariance in the private argument does.[19]

Of greater future significance was Kretschmann's suggestion: use four invariants of the Riemann tensor to fix a unique coordinate system (an individuating field in my terminology). Section III.1 of Kretschmann (1917) discusses the use of the principal directions of the Riemann tensor to fix the coordinate directions;[20] and Section III.2 proposes the use of four mutually-independent invariants of the Riemann tensor and metric as the space-time coordinates.[21]

Apparently unaware of Kretschmann (1917), Arthur Komar (1958) also suggested the use of four invariants of the Riemann tensor as coordinates. In subsequent discussions of the problem of true observables in general relativity, they are often referred to as Kretschmann–Komar coordinates. Stachel (1989, 1993) noted their use as a way of individuating the points of space-time, and they have subsequently figured in many discussions of the hole argument.

Kretschmann (1917) notes that: "This system of [principal] directions . . . naturally may be indeterminate or otherwise degenerate;" and that the four invariants may be used as coordinates "only by postulating that in no finite four-dimensional region are [they] mutually dependent."

Section 6.1 discusses the treatment of such cases, in which the symmetry or isometry group of an equivalence class of metrics is non-trivial.

2.7   The Cauchy problem for the Einstein equations: from Hilbert to Lichnerowicz

As we have seen, in 1913 Einstein formulated his argument against generally covariant equations in terms of the non-uniqueness of the field in a "hole" in space-time. David Hilbert, the renowned mathematician, became interested in the problem of a unified gravitational-electromagnetic theory and followed Einstein in arguing against generally covariant field equations. Instead of a hole, however, he formulated the argument in a mathematically more sophisticated way, using a space-like hypersurface.[22] He showed that there is no well-posed Cauchy problem for generally covariant equations; i.e., no finite set of initial values on such a hypersurface can determine a unique solution

---

[18] Einstein argued that: "If one were to put Newtonian gravitational mechanics into the form of absolutely covariant (four-dimensional) equations, one would undoubtedly be convinced that [the] principle . . . rules them out, granted not theoretically, but practically" (Einstein, 1918, p. 242; transl. from Rynasiewicz, 1999, p. 458). Within a few years, Cartan and Friedrichs proved Einstein wrong by providing simple, four-dimensional formulations of Newtonian theory that incorporate the equivalence principle (see, e.g., Ehlers, 1973).

[19] The four-dimensional form of Newton's theory (see the previous note) satisfies the first criterion, but not the second: its degenerate spatial and temporal metrics restrict the symmetry group of the theory (see, e.g., Ehlers, 1973).

[20] Kretschmann states: "G[ustav] Mie suggested this possibility of fixating the coordinate directions in an absolute way just by using the general theory of relativity in a letter of February 1916."

[21] Kretschmann's discussion of the bearing of these results on the relativity group of the theory fails to make a distinction between the group of transformations under which the theory is invariant (i.e., the nature of the equivalence classes of physically equivalent solutions), and the symmetry or isometry group of a particular solution to the equations.

[22] But see Section 6.4 and Section 7 for some further comments on this matter.

---

to these equations off the initial hypersurface.[23]

After Einstein returned to generally covariant field equations, Hilbert dropped this argument against them, and Hilbert (1917) is the first discussion of the Cauchy problem in general relativity; but the analysis is far from complete.[24] It was not until 1927 that Georges Darmois gave a reasonably complete treatment.[25] His discussion included the role of null hypersurfaces as characteristics, the use of the first and second fundamental forms on a space-like hypersurface as initial data, and the division of the ten field equations into four constraints on the initial data and six evolution equations. Most post World-War II discussions of the Cauchy problem in general relativity are based on the work of André Lichnerowicz,[26] but he acknowledges his debt to Darmois:

> In 1926 in Belgium, Darmois gave a course of four lectures on "the equations of Einsteinian gravitation" in the presence of De Donder. The monograph version (Darmois, 1927) ... became my bedside reading. In this book ... is the first rigorous analysis of the hyperbolic nature of the Einstein equations, i.e., the foundation of the relativistic theory of gravitation as a theory of wave propagation, With profound understanding, the splitting of the Einstein equations relative to the Cauchy problem into two sets is clearly discussed: one treats the initial conditions, and the other deals with time evolution (Lichnerowicz, 1992, p. 104).

Many current discussions of the non-uniqueness problem in general relativity are formulated in terms of the Cauchy problem rather than the original hole argument (see, e.g., Belot and Earman, 2001; Rickles, 2005; Lusanna and Pauri, 2006).

---

[23] For details, with references to Hilbert's lectures and papers, see Renn and Stachel (2007).

[24] For example, while stating that physically meaningful results should not depend on the coordinate system, Hilbert actually adopted Gaussian normal coordinates in order to get uniqueness. An even more serious fault: he failed to mention the constraints on the initial data.

[25] See Stachel (1992) for a review of the early history of the Cauchy problem in general relativity.

[26] Lichnerowicz (1955) was particularly influential. See, e.g., Bruhat (1962).

3   Modern Revival of the Argument

Discussion of the hole argument lay dormant for sixty years. Its modern revival came about as the result of debates about the reason for the delay of over two years between Einstein's adoption of the metric tensor in 1913 and his formulation of the generally-covariant field equations for the metric at the end of 1915 (see Section 2).

3.1   Did Einstein misunderstand coordinate transformations?

The answer to this question hinges on the answer given to the question of why Einstein formulated the hole argument and held to it during this entire period. In 1982, Pais summarized the generally-accepted view:

> In 1914 not only did he [Einstein] have some wrong physical ideas about causality but in addition he did not yet understand some elementary mathematical notions about tensors (Pais, 1982, p. 224). [27]

> Einstein still had to understand that this freedom [to make "an arbitrary coordinate transformation"] expresses the fact that the choice of coordinates is a matter of convention without physical content (ibid., p. 222). [28]

In 1979, Stachel started to study this question. Stachel (1979) presented a version of the standard account, but by the following year it had become evident that this account was incorrect. At the 1980 Jena meeting of the GRG Society, he presented a detailed analysis of the hole argument and its refutation by the point coincidence argument; it circulated as a preprint, but was not published until 1989 (Stachel, 1989). However, Torretti (1983, Chapter 5.6) gives a detailed account of the hole argument based on it;[29] and Norton (1984), the first detailed analysis of Einstein's 1913 Zurich Notebook,[30] also summarizes Stachel's account.

Stachel (1987) contains a historical-critical account of the hole argument, and Stachel (1986) uses the fiber bundle formalism to generalize the argument to any geometric object field obeying generally-covariant equations. These two talks helped to stimulate renewed interest in the meaning of diffeomorphism invariance among relativists, especially those working on quantum gravity (see, e.g., Rovelli, 1991).

Earman and Norton's presentations of the hole argument Earman and Norton (1987); Earman (1989) provoked renewed discussion of absolute versus relational theories of space-time among philosophers of science, a discussion that continues to this day.[31] Section 5 shows how several

---

[27] Even as astute a student of relativity as Cornelius Lanczos, who had worked with Einstein, misunderstood the issue. Explaining Einstein's mistake, he writes: "It would be fatal to obtain an infinity of solutions of the field equations in one and the same reference system. In fact, however, we have merely the freedom of introducing curvilinear coordinates at will. In every given coordinate system the solution is determined, under the proper initial and boundary conditions" (Lanczos, 1970, p. 236).

[28] For a list of some other accounts based on this assumption, and a critique of them, see Norton (1984, especially Section 3).

[29] "John Stachel, who ... was the first to understand the significance of Einstein's reappraisal of the 'hole' argument ... conjectures that Einstein's self-criticism ... rests on the tacit assumption that the points of spacetime where no matter is present cannot be physically distinguished except by the properties and relations induced by the spacetime metric" (Torretti (1983), p. 167)

[30] The Zurich notebook contains Einstein's earliest preserved research notes on general relativity. Renn (2007), a critical edition of the Zurich Notebook with extensive commentary, includes Janssen (2007), which presents new historical evidence on the origins of the hole argument.

[31] "The issues raised in this chapter [on "General Relativity and Substantivalism: A Very Holey Story"] have already touched off a lively debate. For a sampling of opinions, the reader is referred to the articles by J. Butterfield, T. Maudlin, J. Norton and J. Stachel in Fine and Leplin (1989) ..." (Earman (1989), p. 219, note 28). Actually, Stachel's contribution to this debate first appeared in Stachel (1993).

initially-different positions have converged on an approach that gives precise meaning to Einstein's vision of general relativity, and Section 6 reviews some physical topics related to the hole argument.

3.2   Einstein's vision and fiber bundles

As we have seen already, Einstein often posed a problem, the solution to which required mathematical tools that went far beyond his current knowledge. Another example is the vision of the nature of general relativity that replaced his earlier faith in Mach's principle (see Section 2.3). As we shall see, this new vision requires the theory of fiber bundles for its appropriate mathematical formulation.

When asked by a reporter to sum up the theory of relativity in a sentence, Einstein said, half jokingly:

> Before my theory, people thought that if you removed all the matter from the universe, you would be left with empty space. My theory says that if you remove all the matter, space disappears, too! (Einstein, 1931). [32]

In 1952, he developed the same idea at greater length:

> On the basis of the general theory of relativity . . . space as opposed to "what fills space" . . . has no separate existence. . . . 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 space-time], but absolutely nothing, and also no "topological space". . . . There is no such thing as an empty space, i.e., a space without field. Space-time does not claim existence on its own, but only as a structural quality of the field (Einstein, 1952, p. 155).

It is evident that this new approach to general relativity completely reverses his original Machian vision. Now the field is primary, and matter – like everything else – must be treated as an aspect of the field.

Einstein's comment occurs in the course of a discussion of the age-old conflict between absolute[33] and relational interpretations of space, which relativity theory metamorphosed into a conflict between interpretations of space-time.[34]

The quotation above stresses the role of the metric tensor, but elsewhere Einstein emphasizes the role of the affine connection, which he calls a displacement field:

> It is the essential achievement of the general theory of relativity that it freed physics from the necessity of introducing the "inertial system" (or inertial systems) . . . The development . . . of the mathematical theories essential for the setting up of general relativity had the result that at first the Riemannian metric was considered the fundamental concept on which the general theory of relativity and thus the avoidance of the inertial system were based. Later, however, Levi-Civita rightly pointed out that the element of the theory that makes it possible to avoid the inertial system is rather the infinitesimal displacement field Γ^i_{jk}. The metric or the symmetric tensor field g_{ik} which defines it is only indirectly connected with the avoidance of the inertial system in so far as it determines a displacement field (Einstein, 1955, pp. 139 and 141).

---

[32] Clearly, he meant to include time, too. When asked in 1921 "If matter were destroyed, then, what would happen to time and space?" Einstein replied "Then there would be no time and no space" (cited in Illy (2006), p. 222).

[33] The term "substantivalism" has largely replaced "absolutism" in recent philosophical discussions (see Section 5); it seems to have been introduced in Sklar (1974).

[34] For a brief historical review of these debates, see Stachel (2006b).

Einstein's vision can be summed in the sentence: "Space-time does not claim existence on its own, but only as a structural quality of the field." The two main elements of this vision are:

1. If there is no field, there can be no space-time manifold.

2. The spatio-temporal "structural qualities" of the field include the affine connection, which is actually of primary significance as compared to the metric tensor field.

Up until quite recently, the standard formulations of general relativity did not incorporate this vision. They start by postulating a four-dimensional differentiable manifold, which is described as a space-time before any metric tensor field is defined on it; and all other space-time structures, such as the Levi-Civita connection, are defined in terms of this one field.[35] But the concepts of fiber bundles and sheaves enable a mathematical formulation of general relativity consistent with Einstein's vision[36] (see Section 4.3):

1. If there is no total space for the fields, then there is no base manifold.

2. The conceptual distinction between the roles of the metric and connection becomes evident: The metric lives on the vertical fibers of the total space; while the connection lives on the horizontal directions of the total space, connecting the fibers with each other.

Clearly, Einstein's vision favors a non-absolutist view of space-time.[37] While no formalism can resolve a philosophical issue, the traditional approach that starts from a manifold M and defines various geometric object fields over it gives manifold substantivalists an initial advantage: Opponents must explain away somehow the apparent priority of M. The modern approach starts from a principal fiber bundle P with total space E and structure group G, and defines M as the quotient E/G; this gives non-substantivalists an initial advantage: The whole bundle (pun intended), which includes some geometric object field, a connection and a manifold, is there from the start; a manifold substantivalist must justify giving priority to M.

---

[35] This is the approach followed by most current textbooks. See, e.g., Hawking and Ellis (1973), Wald (1984), Stephani (2004), and Goenner (1996).

[36] See, e.g., Trautman (1970, 1980); Göckeler and Schücker (1987). Prugovečki (1992) even entitles his second chapter (pp. 30–65): "The Fiber Bundle Framework for Classical General Relativity." For a modern treatment, see Fatibene and Francaviglia (2003); for a historical account, see Varadarajan (2003).

[37] I use this negative formulation because I believe Einstein's vision is best represented by dynamic structural realism rather than by traditional forms of relativism (see Section 5).

4 The Hole Argument and Some Extensions [38]

After defining geometric and algebraic structures, a space is defined as a set of points with a geometric structure that is invariant under some group of transfomations of its points. Then I discuss product and quotient spaces, fibered spaces, and theories based on these spaces, in particular permutable and generally permutable theories (Section 4.1).

Up to this point, the only structures discussed have been relations between the elements of a set. But all definitions are still applicable – appropriately modified, of course – when additional structures are introduced. In particular, the case of most physical interest is that of geometric object fields defined on a differentiable manifold (Section 4.2). They provide the framework for coordinate-independent definitions of covariant and generally covariant theories, followed by a precise formulation of the of the original hole argument against general covariance and of the way to avoid its conclusion, discussed informally in Section 2.

Then I discuss fiber bundles, which consist of a total space, a base space, and a projection operator. Under certain circumstances, the base space may be defined as the quotient of the total space divided by an equivalence relation defining its fibers (Section 4.3) This approach allows a more precise formulation of Einstein's vision of general relativity, discussed informally in Section 3.2.

Finally, the distinction between natural and gauge natural bundles is discussed, and between the concepts of covariance and general covariance when applied to theories defined on each type of bundle (Section 4.4).

A number of philosophical concepts used but not defined in this section, such as: intrinsic and extrinsic properties, internal and external relations, and quiddity and haecceity, are discussed in the Appendix B.

4.1 Structures, algebraic and geometric, permutability and general permutability

Consider a set S of elements x, y, z, etc., together with a set of relations R between its elements.[39] There is a major distinction between a geometry and an algebra:

Geometry: In a *geometry*, the elements of S (hereafter called *points* and symbolized by p, q, etc.)[40] all are of the *same quiddity* (i.e., of the same nature) but lack *haecceity* (i.e., are not inherently individuated): The only distinctions between the points arise from a set of *internal relations* R between them. If we abstract from these relations, the set S is invariant under $\mathrm{Perm}(S)$, the group of *all permutations* of the points of S.[41]

The set of relations defining a *geometry structure* or *geometry* G on S will be symbolized by $\boldsymbol{R}_G$. The maximal subgroup Aut $\boldsymbol{R}_G(S)$ of $\mathrm{Perm}(S)$ that preserves all these relations between the

---

[38] See Stachel and Iftime (2005).

[39] In this subsection, the nature of S is not further specified. The number of its elements may be finite or infinite. And if infinite, the number may be denumerable or non-denumerable; and if non-denumerable, they may form a discretum or a continuum.

[40] An individual relation R, element of a set a, and point of a geometry p, etc., will be denoted by italic symbols; and the corresponding sets of relations, elements and points will be denoted by boldface italic symbols: $\boldsymbol{R}$, $\boldsymbol{a}$, and $\boldsymbol{p}$, respectively. By so designating the points of a geometry, we are negating their homogeneity; so all meaningful geometric assertions must be invariant under any permutation of these names. Naming is just a particular case of coordinatization, discussed below.

[41] For the present, this notation will be used although it is usually reserved for sets with a finite number of elements. Later in this section, differentiable manifolds M are discussed, for which $\mathrm{Diff}(M)$ is the symbol for the group that is analogous to $\mathrm{Perm}(S)$. As in the case of $\mathrm{Perm}(S)$, invariance under $\mathrm{Diff}(M)$ holds in the absence of any other geometrical structures on M. A study of the subgroups of $\mathrm{Diff}(M)$ and their relation to each other is thus equivalent to a study of all possible geometries on M and their relation to each other.

points of S is called the symmetry or automorphism group of this geometry, and could just as well be used to define it.[42]

Obviously, Perm(S) is the maximal possible automorphism group; so a study of its subgroups and their relation to each other is equivalent to a study of all possible geometries on S and their relation to each other.

Algebra: In contrast to a geometry, in an *algebra* each element (symbolized by a, b, etc.) in addition to having the same quiddity also has an *intrinsic haecceity* (individuality). An *algebraic structure* or *algebra* A on a set is also defined by a set of relations $\boldsymbol{R}_A$ between its elements; but these are *external relations*, which do not affect the intrinsic individuality of each element.[43]

Coordinatization: Since Descartes introduced analytic geometry, it has proved convenient and often necessary to apply algebraic methods in the solution of geometrical problems. This is done by a coordinatization of the geometry (see Weyl, 1946, for this term): A one-one correspondence is set up between the points of the geometry and certain elements of an appropriately chosen algebra. This coordinatization assigns to each point p of the geometry an element a of the algebra, called its *coordinate* and symbolized by $a(p)$. Now certain algebraic operations can be given a geometrical interpretation, and vice versa. But, by individuating the points of a geometry, a coordinatization *negates* their homogeneity, turning the geometry into an algebra. The only way to restore their homogeneity is to *negate* the coordinatization as follows: Introduce the class of all admissible coordinatizations of the geometry[44] based on the given algebra, so that each point of the geometry will have every admissible element of the algebra as its coordinate in at least one admissible coordinate system. Transformations between two admissible coordinate systems are called admissible coordinate transformations; they usually form a group that includes a subgroup isomorphic to the automorphism group of the geometry.

Permutations and the basic identity.   There are two distinct ways in which the assignment of all admissible coordinates to each point of a geometry may be accomplished.

1. Active point transformations: Keep the coordinate system fixed, $a \to a$, and permute the points of the geometry: $p \to q$, $a(p) \to a(q)$.

2. Passive coordinate transformations: Keep the points of the geometry fixed, $p \to p$, and carry out an admissible coordinate transformation of the elements of the algebra: $a \to b$, $a(p) \to b(p)$.

The terms "active" and "passive" refer to the effects of a transformation on the points of the geometry. A passive coordinate transformation is an active transformation of the elements of the algebra.[45]

---

[42] Plane Euclidean geometry provides a simple example. The identity of a triangle, for example, is preserved under the group of automorphisms consisting of all translations and rotations of the points of a two-dimensional manifold homeomorphic to $\boldsymbol{R}^2$.

[43] The real numbers provide a simple example: Each real number is uniquely defined, and collectively they form an algebraic field under the operations of addition and multiplication. The identity elements for addition ("zero") and multiplication ("one") are uniquely fixed, and all rational numbers may be generated from them by iterating the operations of addition, multiplication and their inverses subtraction and division. Real numbers may then be defined by Dedekind cuts between sets of rational numbers.

[44] In the continuous case, complications arise in going from local coordinates patches to a global coordinatization. This is especially so for theories like general relativity, in which the global topology is not fixed in advance, but depends on the solution to the field equations.

[45] From the standpoint of category theory, "The sharp distinction between point transformations and coordinate transformations has disappeared: coordinate systems are simply local diffeomorphisms into $\mathbb{R}^m$" (Nijenhuis, 1994). This is because category theory includes both algebras and geometries in the same category of sets.

Two active permutation groups of any geometry have already been introduced:

1. $\mathrm{Perm}(S)$, the group of all permutations of the elements of S;

2. Aut $\boldsymbol{R}_G(S)$, the subgroup of $\mathrm{Perm}(S)$ consisting of the permutations that belong to the automorphism group $\boldsymbol{R}_G$ of a particular geometry G.

Relations may also be permuted. Let $R(\boldsymbol{p})$ symbolize a relation between the set x of all points of S.[46]

Consider a permutation $P\boldsymbol{x}$ of the elements of S. Define the permuted relation PR as follows: $PR(\boldsymbol{x})$ holds iff $R(P\boldsymbol{x})$ does. When a permutation $P\boldsymbol{x}$ is carried out, the relation R will be said to be "carried along" if it is also permuted into the relation PR.[47]

It follows that, if $R(\boldsymbol{x})$ is valid, then so is $PR(P\boldsymbol{x})$.

The basic identity for the group of all permutations Perm(S). By virtue of the intrinsic homogeneity of its points, a geometry G remains unchanged if, for each Permutation P of $\mathrm{Perm}(S)$, the corresponding permutation $P\boldsymbol{R}_G$ of the set of relations $\boldsymbol{R}_G$ defining $\boldsymbol{G}$ is also carried out. For any relation $R_G \in \boldsymbol{R}_G$, it is clear that $PR_G(P\boldsymbol{p})$ holds if and only if $R_G(\boldsymbol{p})$ holds, so $PR_G \in \boldsymbol{R}_G$; thus $\boldsymbol{R}_G$ and $P\boldsymbol{R}_G$ describe the same geometry. I shall refer to this result as the basic or trivial identity for the group $\mathrm{Perm}(S)$: it holds for *any* geometry based on a subgroup of $\mathrm{Perm}(S)$.

Equivalence relations and quotient maps.[48] An *equivalence relation* R_{eq} on any set S is a two-place relation having the following three properties: For all x, y, z in S, it is *reflexive:* $R_{eq}(x, x)$ holds; *symmetric:* if $R_{eq}(x, y)$ holds, then so does $R_{eq}(y, x)$; and *transitive:* if $R_{eq}(x, y)$ and $R_{eq}(y, z)$ both hold, then so does $R_{eq}(x, z)$. If the context is clear, one often abbreviates $R_{eq}(x, y)$ by $x \approx y$. An equivalence relation divides S into *equivalence classes* S_R, often also called its *orbits* (see Neumann *et al.*, 1994, Chapter V). Every element of S belongs to one and only one such equivalence class. The quotient set of S by R_{eq}, often called the *orbit space* and abbreviated $S_Q = S/R_{eq}$, is defined by the condition that each element of the quotient set corresponds to one and only one such equivalence class.

Product sets and quotient sets. Given two sets A and B, we can form the product set $A \times B$, consisting of all pairs of elements (x, y), with $x \in A$ and $y \in B$. A mapping φ from the domain A to the range or codomain B (see Lawvere and Schanuel, 1997, pp. 13–14), often symbolized $\varphi : A \to B$, is defined as a subset of $A \times B$, such that for each x in the domain there is one and only one y in its range.

> In various contexts, "mappings" might also be called *functions, transformations*, or *operators*. *Homomorphisms, isomorphisms, homeomorphisms, diffeomorphisms, continuous* or *differentiable maps* will be more attached to certain classes of mappings, which "preserve" certain structures on the sets which are their domains and ranges (Hermann, 1973, p. 3).

---

[46] This is the most general type of relation, and includes all special cases. Any relation between a lesser number of elements of S is equivalent to an $R(\boldsymbol{p})$ that is always identically satisfied by all the remaining elements of the set.

[47] Of course, if *all* the relations $\boldsymbol{R}_G$ were totally symmetric, i.e., invariant under all permutations of the relata, there would be no need to introduce permutations of the relations. But most geometries involve some order relations that are not symmetric. Suppose, for example, two points on a straight line obey the relation: $p_1 < p_2$. Then permuting the two points will violate the relation "$<$", so we must introduce the permuted relation "$>$," thus preserving the validity of the permuted order relation $p_2 > p_1$.

[48] For this subsection, Lawvere and Schanuel (1997) is particularly useful, especially Part II.

The mapping is *surjective* if, for every element $y \in B$, there is at least-one element $x \in A$ that maps onto y. If both mappings $A \to B$ and $B \to A$ are surjective, the mapping is called *bijective*. If the bijective map preserves all structures on A and B, A and B are said to be isomorphic. If the mapping $\varphi : A \to B$ is surjective, the set B is isomorphic to the quotient set A/φ; so B can actually be defined as the quotient set.

> This "passing to the quotient" is a way of defining new spaces and mappings that is very important in all of mathematics, particularly in algebra and differential geometry. (Hermann, 1973, p. 5).

This possibility allows us to realize Einstein's vision of general relativity (see Section 3.2), which in this context is simply: If no A, then no B. We can define the mapping or morphism φ from S to S_Q, $\varphi : S \to S_Q$, which projects each element x of S into the element $\varphi(x)$ of S_Q corresponding to the equivalence class that includes x. Conversely a section of S is an inverse mapping φ^{-1} from the point $\varphi(x)$ of S_Q to a unique point y in the equivalence class of S that maps into that point of $S_Q : \varphi^{-1}(x) = y$, with $x \approx y$.

So far, these concepts can be applied to any set. If S is a geometry with automorphism group G, it is referred to as a G-space (see Neumann *et al.*, 1994, pp. 74ff). The equivalence relation is said to be *G-invariant* if, whenever $p \approx q$ holds for two points of S, then it follows that $g(p) \approx g(q)$ for all $g \in G$. Consequently, the action of an element $g \in G$ on S preserves the orbits of S but *permutes* them; so all orbits, henceforth called *fibers*, must be isomorphic to what is called a *typical fiber*. The quotient set $S_Q = S/R_{eq}$ is itself a G-space called the *quotient space*.

Fibered or fiber spaces. A *fibered space* consists of a *total space* E; a *base space* B; and a *projection operator* $\pi : E \to B$ that is a surjective mapping, as defined above. The fiber F_b over each point $b \in B$ is the set of all inverse elements $\pi^{-1}(b) \in E$; that is, all elements $p \in E$ such that $\pi(p) = b$. A section σ of a fiber space is a choice of one element on each fiber F_b for every $b \in B$.

To convert a homogeneous set S with an equivalence relation ρ into a fibered space, let S constitute the total space E; then S_Q forms the base space B, and the mapping φ becomes the projection operator π. If S is a geometry with automorphism group G, then G preserves the equivalence classes; so all the fibers are isomorphic, resulting in a fiber bundle:

A *fiber bundle* (E, B, π) consists of: 1) a *total space* E, divided into fibers by an *equivalence relation* ρ, all of these fibers being isomorphic to a typical fiber φ; 2) a *base space* B that is isomorphic to the quotient E/ρ; and 3) a *projection operator* $\pi : E \to B$ that takes each point of its domain E into the point p of its range B that corresponds to the fiber including p. A *section* of the bundle is a mapping σ that takes each point p of its domain B into a unique point of its range, consisting of the set of fibers of E. The point of σ on the fiber φ_p over p is symbolized by $\varphi_p(\sigma)$. If E has the automorphism group G, the action of an element $g \in G$ on the points of any section σ will result in a new section $'\sigma$; symbolically: $g(\sigma) = {'\sigma}$. So, given one section σ, the action of the elements of G produces a whole equivalence class of sections $\boldsymbol{g}(\sigma)$.

Theories, permutable and generally permutable. A *theory of a certain type* is a procedure for producing models of that type. A *particular theory* of that type is a rule for selecting a subset of these models. One type of theory is defined by the choice of a fiber bundle with automorphism group G; its models are the sections of this bundle. A particular theory is a rule for choosing a subset of sections of the bundle as models. If the rule is such that, when σ is a model, then so is the entire equivalence class of sections $\boldsymbol{g}(\sigma)$, the theory is *permutable*. It is *generally permutable* if this entire equivalence class is interpreted as a single model of the theory. In terms of the distinction between syntax and semantics, one may say: While each section of a theory is always syntactically distinguishable from the others, in a permutable theory they may also be semantically

distinguishable. However, in a generally permutable theory they are not; only an entire equivalence class of sections has a unique semantic interpretation.

Take Euclidean plane geometry, for example. All assertions about geometric figures, such as right triangles, rectangles, circles, etc., are invariant under its automorphism group, which consists of translations and rotations; so it is certainly a permutable theory. But these assertions actually apply to the whole equivalence class of geometric figures satisfying any of these definitions; so it is a generally permutable theory.

On the other hand, plane analytic geometry includes a choice of origin, unit of length, and a pair of rectangular axes. So all of its assertions are still permutable; but some of them include references to the origin, axes, etc. We can distinguish, for example, between a circle of radius r centered at the origin, and one of the same radius centered at some other point. So plane analytic geometry is a permutable theory, but not generally permutable. The reason, of course, is that the choice of a *unique* preferred coordinate system converts the Euclidean plane from a geometry into an algebra.

4.2 Differentiable manifolds and diffeomorphisms, covariance and general covariance[49]

For the space-time theories forming the main topic of this review, S is often a four-dimensional differentiable manifold M; and the analogue of $\mathrm{Perm}(S)$ is $\mathrm{Diff}(M)$, the diffeomorphism group consisting of all differentiable point transformations of the points x of M. Any given, fixed geometric structures defined on M, such as a metric tensor field, will be symbolized by $\Lambda(x)$; they represent the analogue of the relations $\boldsymbol{R}_G$. The Λ-geometry of M is also defined by the invariance of these $\Lambda(x)$ under the action of some Lie subgroup G of $\mathrm{Diff}(M)$. In other words, G plays the role, analogous to that of $\mathrm{Aut}(\boldsymbol{R}_G)$, of the automorphism group $\mathrm{Aut}(M, \Lambda)$ of the Λ-geometry of M. And just as in that case, here every G-space can also be defined as a quotient or orbit space:

> Every G-space can be expressed as in just one way as a disjoint union of a family of orbits. (Neumann *et al.*, 1994, p. 51)

Just as in analytic geometry, one may set up ordinary and partial differential equations for various particles and fields on M. Denote a set of such geometric object fields on M collectively by the symbol $\Phi(x)$, and consider the effect of an element $g(x) \in G$ on $\Phi(x)$.[50] From the definition of a geometric object (see Schouten, 1954, pp. 67–68) it follows that if $x \to g(x) = {}'x$, then $\Phi(x) \to {}'\Phi({}'x)$.

Geometric quantities form an important subclass of geometric objects which includes all tensor fields. Their transformation law under $g(x)$ is linear and homogeneous in the components of $\Phi(x)$ and homogeneous in the derivatives of $g(x)$.

In both Galilean space-time (see Yaglom, 1979) and in special relativistic space-time (Minkowski space), G is a ten parameter Lie group. Four of these parameters generate spatial and temporal translations of the points, making these space-time geometries *homogeneous*. And in both, the six remaining parameters act at each point of space-time: three generate spatial rotations and three generate "boosts", making both space-times non-isotropic. But they do so in different ways because their "boosts" differ: For Galilean space time, they are Galilei transformations that preserve the

---

[49] The reader unfamiliar with the concepts of differentiable manifold, geometric object, and Lie group can consult Schouten (1954) or Isham (1999), for example.

[50] I adopt the Schouten (1954) kernel-index convention for symbols: Any addition to the symbol for a geometrical object placed to the *left* of the symbol denotes a *different object* of the *same geometric type*; while any addition placed to the *right* denotes the *same* object in a *different* coordinate system. Thus, if x denotes a point of a differentiable manifold, ${}'x$ denotes a different point of the same manifold, while x' denotes the same point in a different coordinate system.

invariance of the absolute time. For Minkowski space-time, they are Lorentz transformations that combine spatial and temporal intervals into an invariant, truly four-dimensional space-time interval. Both of these groups are subgroups of $SL(4, R)$, the group of four-volume-preserving transformations.[51] And both theories have a homogeneous, flat affine connection in common that is the mathematical expression of the Newton's first law of inertia. Its invariance group is $AL(4, R)$, which is a subgroup of $SL(4, R)$.

Newtonian gravitational theory, in the form which incorporates the equivalence principle, preserves the global space-time structure of Galilean space-time, but abandons the homogeneous flatness of the affine connection in favor of a non-flat affine connection that is the mathematical expression of the dynamical inertia-gravitational field. This field is non-homogeneous, but its compatibility with the space-time structure requires that locally it remain invariant under $AL(4, R)$, which means that its globally automorphism group must be $S\mathrm{Diff}(M)$, the group of unimodular diffeomorphisms.

General relativity similarly abandons the homogeneous flatness of the affine connection in favor of a non-flat affine connection that is the mathematical expression of the dynamical inertia-gravitational field. But, in order to preserve its compatibility with the special-relativistic chrono-geometry expressed by the metric tensor, the latter must also become a dynamical field. It preserves the local space-time structure of the special theory at each point. But globally both dynamical fields must have automorphism groups consisting of diffeomorphisms of M, the space-tme manifold, now itself no longer globally fixed. Traditionally, $\mathrm{Diff}(M)$, the full diffeomorphism group, has been assumed to be the correct automorphism group for general-relativistic theories. However, there are good arguments for restricting this group to $S\mathrm{Diff}(M)$, the group of unimodular diffeomorphisms with determinant one.

But first some definitions are needed (see, e.g., Wikipedia: Group action). The action of G is said to be *effective* if its identity element is the only one that takes each point into itself: That is, if $g \in G$, $x \in M$ and $\pi : x \to g(x)$ is such that $g(x) = x$ for all x, then $g = e$, the identity element of G. The action is *transitive* if π is a map onto M that connects any two of its points: That is for any two points $x, y \in M$, there is always a $g \in G$ for which $g(x) = y$.

The *stabilizer group* H_x at a point x of M is the subgroup of transformations of G that leave the point x invariant: That is, $g \in H_x$ if and only if $g(x) = x$.[52] Since G is a Lie group, H_x is a closed subgroup at each point of G and these stabilizer groups are conjugate subgroups of G. Indeed, M is isomorphic to G/H_x; so one may actually define a geometry by the pair (G, H), where H is a closed subgroup of G.

The action of G is *free* or semiregular if its stabilizer group is the identity: That is, if $gx = x$ for some point x, then $g = e$, the identity element of G; equivalently, if $gx = hx$ for some x, then $g = h$. For example, the translation groups discussed above act freely on Galilean and special relativistic space-times.

Now we are ready to return to the question of automorphism groups for general relativistic theories. The action of the stabilizer of $\mathrm{Diff}(M)$ on the tangent space at each point x of M is $L_x = GL(n)$, the group of all linear transformations at x. But consider the objects defining the geometry of a general-relativistic space-time with $n = 4$: Again, if one wants to preserve the four-volumes of space-time, which are needed to formulate meaningful physical averages, one must restrict these transformations to $SL(4)$, the group of 23 special linear transformations with unit determinant. The linear affine connection at x, which represents the inertio-gravitational field, is only invariant under the subgroup $ASL(4)$, the group of affine transformations with unit determinant. And the invariance group of the metric tensor, which represents the chrono-geometry, is even further restricted to the pseudo-orthogonal subgroup $SO(3, 1)$ of $SL(4)$. In short, globally

---

[51] This account of the relation between various G-structures is highly simplified mathematically. For a more complete account, based on jet prolongations of fiber bundles, see Sánchez-Rodríguez (2008)

[52] It is also called the stability group, the isotropy group and the little group (see Isham, 1999, p. 180).

physical considerations suggest the need to start from $S\mathrm{Diff}(M)$ as the automorphism group of general-relativistic theories.

So physically, $\mathrm{Diff}(M)$ overshoots the mark by allowing non-unimodular transformations, i.e, transformations with any value of their determinant at a point. Geometrically, they correspond to similarity transformations, which preserve the shape but not the size of four-volumes in space-time. Usually, one "compensates" for this unwanted change of size by introducing tensor densities: When appropriate weights are introduced for various tensors, these densities can undo the effects of the size changes produced by non-unimodular transformations.

However, one may simply start from $S\mathrm{Diff}(M)$ as the automorphism group. The action of its stabilizer on the tangent space at each point of M is SL_x, the maximal symmetry group that preserves the size of four-volumes, thus avoiding the need to introduce densities, among its many other advantages (see Stachel, 2011; Bradonjić and Stachel, 2012). For much of the following discussion, however, the distinction between $\mathrm{Diff}(M)$ and $S\mathrm{Diff}(M)$ is inessential, so I shall continue to discuss $\mathrm{Diff}(M)$, and only point out the distinction at some places where it is really important.

Symmetry transformations: By definition, Aut $\boldsymbol{R}_G(S)$, the group of permutations of the points of S defining the geometry G, leaves the relations $\boldsymbol{R}_G$ (which equally well define the geometry of S) unchanged; so the $\boldsymbol{R}_G$ do not need to be permuted when the points of S are. Whichever Lie subgroup G of $\mathrm{Diff}(M)$ is chosen as the automorphism group defining the geometry of a differentiable manifold M, similar comments hold for it. As we shall see, the important difference for the hole argument is that between geometries based on finite-parameter Lie groups and those based on Lie groups depending on one or more functions (functional Lie groups).

Passive coordinate transformations and the trivial identity: Since it is no more than a relabeling of its points, any admissible passive coordinate transformation has no effect on a geometry (see Section 4.1). However, if one *restricts* the group of coordinate transformations to a subgroup of those corresponding elements of the automorphism group of the geometry, then there is an *isomorphism* between this *subgroup* of *passive* coordinate transformations and the group of active point transformations defining the geometry. Hence, it is possible to reformulate any statement about the geometry in terms of relations between the coordinate components of the geometric object fields Λ that are invariant under this subgroup of *restricted coordinate transformations*. In the past, this is how coordinate-dependent techniques were used to arrive at geometric results; and many contemporary treatments still utilize this technique. If one permutes the points of M by a diffeomorphism, carries along the Λ fields defining its geometry and the Φ fields defining the theory, and also carries out the corresponding coordinate transformations, then clearly the *new* fields at the *new* points will have the same coordinate components in the *new* coordinate system as the *old* fields at the *old* points in the old coordinate system. This observation is another, coordinate-dependent variant of the basic or trivial identity. It holds for any Φ fields, quite independently of any theory, or any field equations that these fields may obey.

Basis vectors and a coordinate-independent formulation of the trivial identity: Geometrically, a coordinate system corresponds to the choice of a holonomic basis $\boldsymbol{e}_i$ at each point of M: That is, there is a local coordinate system such that $\boldsymbol{e}_i = \partial/\partial x^i$. But the essential element geometrically is the choice of a basis, not its holonomicity. So, introduce an ordered set of basis vectors $\boldsymbol{e}_i(x)(i = 1, 2, \ldots, n)$, holonomic or not, at each point x of M together with the associated dual basis of covectors or one-forms $\boldsymbol{e}^j(x)$, such that $\langle \boldsymbol{e}_i \boldsymbol{e}^j \rangle = \delta_i^j$.[53] Associated with the geometric object fields Φ and Λ on M are their components with respect to such a pair of basis vectors,

---

[53] Note that i, j are not vectorial component indices, but merely serve to enumerate the order of the sequence of vectors and of convectors.

which will be symbolized by $\Phi[\boldsymbol{e}(x)]$ and $\Lambda[\boldsymbol{e}(x)]$: This is a set of *coordinate-independent* scalars that result from saturating all the free covariant and contravariant indices of Φ and Λ with the $\boldsymbol{e}_i$ and $\boldsymbol{e}^j$ respectively.

Of course, under a change of basis: $\boldsymbol{e}(x) \to '\boldsymbol{e}(x)$ these scalars transform appropriately. A diffeomorphism $D : x \to 'x$ induces such a change of basis: $e(x) \to '\boldsymbol{e}_D(x)$, and corresponding changes in the geometric object fields $\Lambda(\boldsymbol{e}) \to '\Lambda_D(x)$ and $\Phi(x) \to '\Phi_D(x)$. However, the values of these scalars remain unchanged if one carries out the associated push forwards and pull backs of Φ and Λ, as well of the basis vectors and covectors. That is, if we take the *new* basis vectors at the *new* point: $'\boldsymbol{e}_D('x)$; then the *new* components with respect to the *new* basis vectors at the *new* points will equal the *old* components with respect to the *old* basis vectors at the *old* points: $\Phi[\boldsymbol{e}(x)] = '\Phi_D['\boldsymbol{e}_D('x)]$ and $\Lambda[\boldsymbol{e}(x)] = '\Lambda_D['\boldsymbol{e}_D('x)]$. This observation is a coordinate-independent formulation of the basic identity. Since any model of a physical theory can only fix the values of such coordinate-independent scalars with respect to some basis for all geometric objects in that model, this identity cannot fail to hold for any theory based on the Λ-geometry of M.

Covariance and general covariance: Suppose we perform the push forwards and pull packs on the geometric object fields Φ, but *not* on the Λ-geometry or the basis vectors and convectors. That is, let $\boldsymbol{e}(x) \to \boldsymbol{e}(x)$ and $\Lambda[\boldsymbol{e}(x)] \to \Lambda[\boldsymbol{e}(x)]$, but $\Phi[\boldsymbol{e}(x)] \to '\Phi_D[\boldsymbol{e}(x)]$. In general, $\Phi[\boldsymbol{e}(x)] \neq '\Phi_D[\boldsymbol{e}(x)]$, so this results in a set of scalars that is distinct from $\Phi[\boldsymbol{e}(x)]$ at each point x of M.

A theory is covariant under the the Λ-geometry's automorphism group if, whenever $\Phi[\boldsymbol{e}(x)]$ is a model of the theory, then so is $'\Phi_D[\boldsymbol{e}(x)]$. Covariance clearly defines an equivalence relation between models of the theory; so covariance divides all models of a theory into equivalence classes.[54]

A covariant theory is *generally covariant under the Λ-geometry's automorphism group* if an *entire equivalence class* of its mathematically distinct models corresponds to a *single* physical model of the theory.

4.3   Fiber bundles: principal bundles, associated bundles, frame bundles, natural and gauge-natural bundles [55]

An ordered set of basis vectors $\boldsymbol{e}_i(x)$ at a point of M is called a linear frame, and the set of all such linear frames at a point of M constitutes one fiber of the bundle of linear frames over M. As Kobayashi explains, the bundle concept can be used to formulate any geometry on M as a G-structure:[56]

> Let M be a differentiable manifold of dimension n and $L(M)$ the bundle of linear frames over M. Then $L(M)$ is the principal fibre bundle over M with group $GL(n; \boldsymbol{R})$. Let G be a Lie subgroup of $GL(n; \boldsymbol{R})$. By a G-*structure* on M we shall mean a differentiable subbundle P of $L(M)$ with structure group G. (Kobayashi, 1972, p. 1).

Such a fiber bundle formulation of geometries has several crucial advantages:

It makes evident the fundamental distinction between vertical geometrical quantities, such as metric tensors, that live on the fibers of the bundle; and horizontal geometrical objects, such as linear affine connections, that serve to connect these fibers. This is the case whether the metric

---

[54] Of course, we could perform the opposite kind of dragging: Let $\boldsymbol{e}(x) \to' \boldsymbol{e}_D('x)$ and $\Lambda[\boldsymbol{e}(x)] \to' \Lambda_D['\boldsymbol{e}_D('x)]$, but $\Phi[\boldsymbol{e}(x)] \to \Phi['\boldsymbol{e}_D('x)]$. This leads to the same equivalence classes of models of a covariant theory: It is only the *relative dragging* that matters.

[55] Useful references for this subsection are Fatibene and Francaviglia (2003), Michor (2008), Olver (1995) and Sharpe (1997).

[56] Note the difference in the meaning of G: Kobayashi denotes by G, what is denoted above by H, the stabilizer group of what is denoted above by G: the subgroup of $\mathrm{Diff}(M)$ defining the Λ-geometry of M.

and/or connection are fixed and given components of Λ; or are components of Φ, themselves subject to dynamical field equations.

It enables us to go from global to local formulations of background-independent theories, such as general relativity, in which the global topology of the base manifold M cannot be specified *a priori*, because it differs for different solutions to the field equations.[57]

Fibered manifolds

The concept of fibered spaces for a set, discussed in Section 4.1, can now be applied to differentiable manifolds (see Section 4.2). After a fibered manifold is defined, the important cases of principal bundles, vector bundles, natural bundles and gauge-natural bundles and their physical applications are discussed, stressing the importance for general relativity of quotient bundles and local considerations.

A *fibered manifold* (E, M, π) consists of a *total manifold E*, a *base manifold M*, and a *projection operator* $\pi : E \to M$. E is a differentiable manifold, the points of which, u, v, etc., are grouped into *fibers* by an equivalence relation ρ between its points. M is also a differentiable manifold, the points of which are symbolized by x, y, etc. The fiber over x is symbolized by F_x. Note that, if the relation ρ is given initially, sometimes the base manifold M may be defined as the quotient of the total manifold E by $\rho : M = E/\rho$; in other words as the orbit space of G (see Sections 4.1 and 4.2). But the situation is generally somewhat more complicated:

> Usually, when symmetries and invariance groups are considered, a problem reduces to the corresponding orbit space, and therefore the structure of these spaces has to be investigated. This structure theory is quite complicated in general, since these spaces usually are singular spaces and not again manifolds. In fact, only if the action of the Lie group is free (i.e., all isotropy subgroups of single points are trivial), the resulting orbit space bears a manifold structure and forms together with the manifold and the quotient map a principal fiber bundle, whose structure is well known. More often, the orbit space admits a stratification into smooth manifolds with an open and dense largest stratum, the set of principal orbits.... This stratified space can then be treated almost like a manifold when taking special care. The existence of such a stratification is usually shown by proving the existence of slices at every point for the group action (Schichl, 1997, p. 1).

I shall assume that – as in general relativity – in any theory considered, the quotient space is either a manifold or a stratified manifold; and that any local solution to its field equations can be extended to a global solution.[58]

A *fiber bundle* is a fibered manifold in which all its fibers are isomorphic to a *typical fiber F*, itself a manifold; that is, for all x, $F_x \approx F$. Suppose F is q-dimensional and M is p-dimensional One can always introduce a *local trivialization* of the bundle: Let X be an open subspace of M. Locally, the total space E is a product space $(F \times X)$, and one can introduce p variables $(x^1, x^2 \ldots, x^p)$ as local coordinates of a point x of X, and q variables $(u^1, u^2, \ldots, u^q)$ as coordinates of a point u of F. So $(F \times X)$ is coordinatized by the $(p+q)$ coordinates (x, u) of a point u_x of E lying on the fiber F_x over the point x. Let G be a Lie group of *diffeomorphisms* that acts on E.[59] The action of an element $g \in G$ on the point (x, u) of E is symbolized by $(x, u) \to g(x, u) = ('x,' u) = [\chi(x, u), \psi(x, u)]$.

---

[57] The full significance of this observation involves a discussion of sheaves of cross sections defined over a germ (see Stachel and Iftime, 2005).

[58] More carefully formulated: Some criteria must be given that define the *maximal* extension of a local solution of the field equations. For example, the *Kruskal manifold* is the maximal *analytic* extension of the Schwarzschild solution to the Einstein equations. For some results relevant to general relativity, see Isenberg and Marsden (1982).

[59] I shall not discuss the order of differentiability postulated, which varies with the theory being defined. In the mathematical literature, *smoothness* is often postulated, i.e., differentiability to all orders. But in physical theories

Two subgroups of G are especially important:

- The *base transformations* (diffeomorphisms of X) that do not affect the fibers:
  $x \to {}'x = \chi(x), \quad {}'u = u.$

- The *pure fiber* or *pure gauge transformations* on the fiber at each point:
  ${}'x = x, \quad {}'u = \psi(x, u).$

Both of these are included in a third subgroup:

- The *fiber-preserving transformations*:
  $(x, u) \to ({}'x, {}'u) = [\chi(x), \psi(x, u)].$

If G is a connected Lie group, all of its actions can be constructed from iterations of the action of its Lie algebra, composed of its *infinitesimal generators*: The vector fields $\boldsymbol{V}$ on E, each of which generates a one-parameter group of point transformations, or *flow*, on E. Locally $\boldsymbol{V}$ may be written in terms of the coordinates (x, u):

$$\boldsymbol{V} = \sum_i \xi^i(x, u)\, \partial/\partial x^i + \sum_\alpha \varphi^\alpha(x, u)\, \partial/\partial u^\alpha \qquad i = 1, \ldots, p \quad \alpha = 1, \ldots, q. \tag{3}$$

The generator $\boldsymbol{V}$ is called:
Horizontal if $\varphi^\alpha = 0$, i.e., it generates only base transformations.
Vertical if $\xi^i = 0$, i.e., if it generates only pure fiber or pure gauge transformations.
The flow generated by $\boldsymbol{V}$ will be *fiber preserving* if and only if $\xi^i = \xi^i(x)$.

Natural bundles: A fiber-preserving diffeomorphism projects naturally into a *unique* diffeomorphism of the base manifold M; but generally the converse does not hold. If it does, i.e., if a base diffeomorphism of M lifts *uniquely* to a fiber-preserving diffeomorphism of E, then the bundle is a *natural bundle*. A geometric object defined on such a bundle is called a *natural object*.[60] This is the fiber bundle version of the definition of geometric objects in Section 4.2.

Principal bundles, associated vector bundles: If the the typical fiber F is isomorphic to the *structure group* $G : F \approx G$, then the bundle is a *principal fiber bundle* P with *structure group* $G : P = (E, M, \pi; G)$. Corresponding to any P with *structure group* G, there is a class of *associated vector bundles*. In such an associated bundle, each fiber forms a *vector representation* of G. This vector representation need not be irreducible, so the class of associated vector bundles includes all tensor fields.

4.4 Covariance and general covariance for natural and gauge-natural bundles

The use of fibered manifolds allows a precise formulation of the concepts of covariance and general covariance for any physical theory; and of the hole argument for background-independent theories, and even – with appropriate modifications – for some partially-background-dependent theories.

---

defined by hyperbolic systems of partial differential equations, it is precisely the existence of *non-smooth* solutions that allows for the transmission of information. Indeed, the *characteristic hypersurfaces* of such a system may be defined as those hypersurfaces, along which *discontinuities* may propagate.

[60] Note that this is the coordinate-independent analogue of the traditional definition of the law of transformation of a geometric or natural object under a coordinate transformation. See, e.g., Stachel (1986). For a category-theoretical treatment of natural bundles, see the Appendix A.

4.4.1   Fiber bundles needed in physics

Every *natural* physical theory can be formulated in terms of some *natural geometric object(s)*[61] that lives on an appropriate fibered differentiable manifold,[62] the nature of which depends on these geometric object(s). If the theory is defined on a differentiable manifold M that is the quotient of the fibered manifold E divided by the equivalence relation ρ defining the fibration, $M = E/\rho$; then there is an operator π, projecting each fiber onto the corresponding point of M: $\pi : E \to M$. Since the fibered manifold represents a natural object, there is a one-one correspondence between fiber-preserving diffeomorphisms of E and diffeomorphisms of M.

A number of most important *gauge natural* theories cannot be so formulated, but require the broader concept of *gauge natural bundles* for their precise formulation. Indeed, every classical physical theory can be reformulated as the jet prolongation of some gauge natural bundle by adjoining the derivatives of the geometric object fields to the original bundle.[63]

4.4.2   Background-dependent theories

A theory based on such fixed Λ-fields on M is called *background-dependent*, with $\mathrm{Aut}(M, \Lambda)$ as its symmetry group.[64] Any geometric object fields $\Phi(x)$ can then be introduced on this fixed-background space-time together with a set of field equations governing their dynamics, which generally involve some or all of the $\Lambda(x)$.

In many theories, the fixed geometric object fields on M consist of a vertical chrono-geometric *metric tensor* on each fiber and the corresponding horizontal inertio-gravitational *linear connection*. Any non-gravitational theory can be formulated on a fiber bundle associated with the principal bundle determined by the metric and connection: The $\Phi(x)$ break up into two subclasses: The fields of massive objects (such as charged bodies) are represented by geometric quantities living on the vertical fibers; and the gauge fields transmitting the forces between these objects (such as the electromagnetic field) are represented by verical connections along the fibers; these connections are only fixed up to some group of *gauge transformations*.

4.4.3   Background-independent theories

In the case of general relativity and other background-independent theories (such as the coupled Einstein–Maxwell equations), Λ reduces to the identity and there are no fixed background space-time structures on M. $\mathrm{Diff}(M)$ is chosen as $\mathrm{Aut}(M)$ in the usual formulations; but, as suggested in Section 4.2, $S\mathrm{Diff}(M)$, the unimodular subgroup, may be chosen. In that case, the space-time structures subdivide further: The pseudo-metric splits into a *conformal metric* with determinant -1 and a *scalar field*, both of which live on the vertical fibers; while the linear affine connection splits into a trace-free *projective connection* and a *one form*, both of which serve to connect the fibers.

---

[61] In order to treat theories involving more than one field, the object is not required to be irreducible.

[62] Again, we put aside the complications that arise when these objects are only defined locally and the global topology of the manifold is not defined *a priori*, but depends on the dynamics; i.e., on the solution to the field equations.

[63] "Indeed, every classical field theory can be regarded as taking place on some jet prolongation of some gauge natural (vector or affine) bundle associated with some principle bundle over some base manifold" (Matteucci, 2003, p. 115).

[64] It is possible to lift the fixed space-time structures from the base manifold into the fibers; but this does not change the essentially background-dependent nature of these theories (see Section 5).

4.4.4   Gauge symmetries

To define the gauge symmetries of a certain type of theory, one must consider the sections of the corresponding fiber bundle. A *local section* $\sigma : M \to X \in E$ or *global cross section* $\sigma : M \to E$ is a map taking each point p of X or E, respectively, into a unique point of the fiber F_p over p.[65]

For each type of physical theory, a section represents a particular configuration of the corresponding physical field. However, in theories of the gauge-field type, this representation is not unique. There is a group of *gauge transformations*, each element of which maps one mathematical representative of some field configuration into another representative of the *same* configuration. A *gauge symmetry* is an equivalence relation on the set of sections: Two sections σ and $'\sigma$ are *gauge equivalent* if there is a gauge transformation taking one into the other. This equivalence relation divides the set of all sections into equivalence classes, the *gauge orbits*; each section belongs to one and only one such orbit. If the gauge group of some type of theory consists entirely of fiber-preserving transformations, then the theory can be formulated on a natural bundle. But if its gauge group includes *non-fiber-preserving* transformations, then a *gauge-natural bundle* is needed to formulate this type of theory correctly.

The field equations of a particular gauge field theory serve to pick out a class of *preferred sections* consisting of the solutions to these equations. For a gauge theory, these equations must be of such a form that, if one section is a solution, then so are all members of the entire gauge orbit of that section. In other words, the gauge transformations must form a *symmetry group of the field equations.* This group is the *automorphism group* $\mathrm{Aut}(P)$ of the principal gauge-natural bundle P corresponding to the theory (see, e.g., Fatibene and Francaviglia, 2003, p. 223). Fatibene and his collaborators explain the distinction between the two types of theory well:

> The main technical difference between natural and gauge natural theories is that [base] diffeomorphisms are completely replaced by gauge transformations. In gauge natural theories spacetime diffeomorphisms do not act at all on fields, since the only action one can define in general is that of gauge transformations. This is due to the fact that although pure gauge transformations are canonically embedded into the group of generalized gauge transformations, there is no canonical "horizontal" complement to be identified with $\mathrm{Diff}(M)$. ... "Horizontal" symmetries, in fact, are generally associated to physically relevant conservation laws, such as energy, momentum and angular momentum. The definition of such quantities is almost trivial in natural theories; on the contrary, in gauge natural theories pure gauge transformations are easily associated to gauge charges (e.g., the electric charge in electromagnetism), while the absence of "horizontal" gauge transformations is a problem to be solved to appropriately define energy, momentum and angular momentum. For this reason, in gauge natural theories the dynamical connection plays an extra role in determining horizontal infinitesimal symmetries as the gauge generators which are horizontal with respect to the principal connection (Fatibene *et al.*, 2001, pp. 3–4).

While bundle formulations of the hole argument originally dealt only with natural bundles, Lyre (1999) develops a generalized version that can be applied to gauge-natural bundles:

> The generalized hole argument is motivated and extended from the spacetime hole argument .... [It] rules out fiber bundle substantivalism and, thus, a relationalistic interpretation of the geometry of fiber bundles is favored (Lyre, 1999, p. 1).

Healey (2001) also argues that "fiber bundle substantivalism ... is subject to an analogue of the 'hole' argument against space-time substantivalism."

---

[65] We shall not here discuss the important question of local versus global sections (see, e.g., Fatibene and Francaviglia, 2003). Unless the contrary is specifically stated, any mention of sections should be interpreted as a reference to local sections. For some further remarks on this question, see Section 6.

4.4.5 Four-geometries and stratified manifolds

To recapitulate: The choice of a bundle (E, M, π) selects a certain type of physical theory but does not picked out a particular theory of that type, nor introduce any space-time structures on E or M. The points of M form a geometry (see Section 4.1). As points of the space-time manifold, they have quiddity but they lack haecceity: *a priori* there is nothing to distinguish one such point from the others. Their *automorphism group* is the *diffeomorphism group* of M or some appropriate subgroup, such as the unimodular group (see Stachel, 2011; Bradonjić and Stachel, 2012). For example, a metric-free formulation of electromagnetic theory can be based on a bundle of one-forms.

A *particular theory* is a *rule* for choosing a preferred class of cross sections of the fiber bundle. This rule generally includes specification of some space-time structures on M. For example, in addition to a bundle of one-forms, source-free Maxwell electromagnetic theory, requires the specification of a conformal structure on M.

In general relativity, an *equivalence class* of diffeomorphically-equivalent pseudo-metrics on a four-dimensional manifold,[66] often referred to as a *four-geometry*, is regarded as corresponding to *a single* inertia-gravitational field. While the fiber space consisting of all four-metrics over a given manifold forms a manifold,[67] the space of all four-geometries does not form a simple manifold, but a *stratified manifold*. That is, it is partitioned into *slices*, each of which is itself a manifold, consisting of all four-geometries having the same *symmetry* or *isometry group*. The largest slice is the manifold of *generic geometries* having no nontrivial symmetries; it contains the vast majority of geometries. Thence one descends slice by slice down to the slice consisting of all four-geometries having the *maximal* symmetry group (see Stachel, 2009 and Section 6.1).

The rule specifying the choice of a preferred class of space time structures may or may not include some restriction on $\mathrm{Diff}(M)$, the maximal possible automorphism group of M. Obviously, diffeomorphisms always remain unrestricted in the sense of the trivial identity. A true restriction on the theory arises with the imposition of a finite-parameter Lie group as the symmetry group of the class of space-time structures picked out by the rule.[68] If there are no such restrictions, the theory is *background independent*. If the rule includes a Lie group involving some functions as well as parameters, the theory is *partially background dependent*. If the Lie group is *maximal* (ten-parameters in four dimensions), then the theory is *totally background dependent*. If the rule restricts the preferred class maximally, i.e., to the identity, then the theory specifies an *individuating field* on the space-time, turning it into an algebra. In this formulation, the symmetry group is included in the rule defining a physical theory, rather than being imposed a priori on the space-time structures defined on M. This change enlarges the class of physical solutions: For example, not fixing the global topology of M allows several possibilities for the global topology associated with a given local metric. But this does not alter the fact that the symmetry group of the space-time structures must be preserved by all such solutions. To what extent the hole argument applies to a non background-independent theory depends on the degree of background dependence that has been imposed (see Stachel, 2009); but if a theory is background independent, the hole argument certainly applies.

---

[66] This is clearly an equivalence relation (i.e., it is reflexive, symmetric and transitive), so it divides the set of all pseudo-metrics into equivalence classes.

[67] Again, here I sidestep the question of how one proceeds from the local to the global point of view.

[68] If v is a vector field generating a one-parameter subgroup of such a symmetry group and Λ is any geometric object involved in defining the space-time structure, then $\pounds_{\boldsymbol{v}}\Lambda = 0$.

5   Current Discussions: Philosophical Issues

On the basis of the analysis developed in the previous sections,[69] I shall re-examine some of the issues currently being discussed in the philosophy of science. Rather than attempting to cover the vast literature on this subject, the discussion is limited to a few representative samples of what I consider to be the most important trends, and will show that they are converging towards variations around a common denominator.

5.1   Relationalism versus substantivalism: Is that all there is?

Since Earman and Norton (1987) (see Section 3), philosophical discussion of the hole argument has centered largely around the issue of space-time absolutism – now often called substantivalism[70] – versus the opposing viewpoint, usually denominated relationalism or relationism. Einstein summarized an earlier version, the age-old controversy over the nature of space:

> Two concepts of space may be contrasted as follows:
>
> (a) space as positional quality of the world of material objects;
> (b) space as container of all material objects.
>
> In case (a), space without a material object is inconceivable. In case (b), a material object can only be conceived as existing in space; space then appears as a reality which in a certain sense is superior to the material world ("Foreword" to Jammer, 1954).

Relativity theory metamorphosed the object of controversy from space to space-time, and Einstein made is his own viewpoint quite clear:

> On the basis of the general theory of relativity ... space as opposed to 'what fills space' ... has no separate existence. If we imagine the gravitational field ... to be removed, there does not remain a space of the type [of the Minkowski space of SR], but absolutely nothing, not even a 'topological space' [i.e., a manifold] ... There is no such thing as an empty space, i.e., a space without field. Space-time does not claim existence on its own, but only as a structural quality of the field (Einstein, "Relativity and the Problem of Space," in Einstein, 1952).

Here are a couple of recent statements on the nature of the controversy:

> Substantivalists understand the existence of spacetime in terms of the existence of its pointlike parts, and gloss spatiotemporal relations between material events in terms of the spatiotemporal relations between points at which they occur. Relationists will deny that spacetime points enjoy this robust sort of existence, and will accept spatiotemporal relations between events as primitive (Belot and Earman, 2001, p. 227).

> A modern-day substantivalist thinks that spacetime is a kind of thing which can, in consistency with the laws of nature, exist independently of material things (ordinary matter, light, and so on) and which is properly described as having its own properties, over and above the properties of any material things that may occupy parts of it (Hoefer, 1996).

---

[69] For the benefit of those who did not read Section 4 and the Appendix B, this section reviews some concepts introduced there.

[70] "Substantivalism" is a neologism, not found it in any current dictionary of English or philosophy. Most dictionaries define "substantialism": For example, the *Oxford English Dictionary* defines it as "The doctrine that there are substantial realities underlying phenomena." It also defines "substantival" as "existing substantially."

> What is space? What is time? Do they exist independently of the things and processes
> in them? Or is their existence parasitic on these things and processes? Are they like
> a canvas onto which an artist paints; they exist whether or not the artist paints on
> them? Or are they akin to parenthood; there is no parenthood until there are parents
> and children? That is, is there no space and time until there are things with spatial
> properties and processes with temporal durations? These questions have long been
> debated and continue to be debated. The hole argument arose when these questions
> were asked in the context of modern spacetime physics. In that context, space and time
> are fused into a single entity, spacetime, and we inquire into its status. One view is that
> spacetime is a substance, a thing that exists independently of the processes occurring
> within spacetime. This is spacetime substantivalism (Norton, 2011).

In the light of the hole argument, I find it more fruitful to frame discussion in terms of two other distinctions, leading to a point of view about space-time distinct from either substantivalism or relationalism as traditionally defined. These are the distinctions between:

1. internal and external relations, and between

2. quiddity and haecceity.

These concepts are discussed in Appendix B and briefly reviewed in Section 5.5. When applied to mathematical structures, they lead to succinct discussions of algebraic and geometric structures and the nature of coordinatization in Section 5.6, which establishes a correspondence between the two (for a fuller discussion, see Section 4.1).

These concepts lead to a viewpoint on the nature of space-time that has been given various names, such as *structural spacetime realism*[71] and *sophisticated substantivalism* (see Pooley, 2000, summarized in Section 5.3). I have called it *dynamic structural realism* (see Stachel, 2006a), which has several advantages. It avoids use of the words "substantivalism" and "relationalism," fraught with so many unwanted implications; it places emphasis on diachronic aspects of structure; and its application is not confined to theories of space-time structure (see Stachel, 2005). The fiber bundle approach, motivated in Section 3 and treated in more detail in Section 4.3, allows a rigorous formulation of this viewpoint in Section 5.4.

But first, I shall give a brief account of the controversy between relationalists and substantivalists provoked by the hole argument and how it has led a number of participants from each camp to adopt this new viewpoint. Rather than attempting a (necessarily superficial) review of the vast philosophical literature on the controversy, I shall focus on an account of the views of one important relationalist and one important substantivalist.

---

[71] "[T]he debate can be solved by advancing a tertium quid, a third option between classical substantivalism and relationism. Ths option, which I call structural spacetime realism, sides with the latter doctrine in defending the relational nature of spacetime, but argues with the former that spacetime exists, at least in part, independently of particular physical objects and events . . . " (Dorato, 2000, pp. 1607–1608).

5.2 Evolution of Earman's relationism

Earman (1989) is a standard reference, so I shall employ its terminology and notation in discussing his views. Discussing a "modified form" of absolutism, he states that

> The only plausible candidate for the role of supporting the nonrelational structures [of a physical theory] is the space-time manifold M ... (ibid. p. 125).

Calling "manifold substantivalism" the view that "M is a basic object of predication," he sets out to show that this view "lays itself open to Leibniz's argument" (p. 126). In his formulation of the problem, Earman uses the standard pre-bundle approach to theories (see Sections 3.2 and 4.3): A model $\mathscr{M}$ of a theory consists of the manifold M, together with "[geometric] object fields on M", which he denotes by A_i and P_j, "characterizing respectively the space-time structure and the physical contents of space-time." Symbolically, $\mathscr{M} = \langle M, A_i, P_j \rangle$, with i and j each running through a finite sequence of integers. A manifold diffeomorphism $d : M \to M$ then results in a different model $\mathscr{M}_d = \langle M, d * A_i, d * P_j \rangle$, where $d * A_i, d * P_j$ denote the pull backs or the push forwards of A_i, P_j.

It is important to note that Earman (1989) defines "general covariance" in a way that is equivalent to my definition of "covariance" (see Section 4.4):

> Let us say that the laws of [a theory] T are *generally covariant* just in case whenever $\mathscr{M}$ [is a model of the theory] then also $\mathscr{M}^{\Phi}$ [is a model] for any manifold diffeomorphism ... (ibid., p. 47).

In his treatment of the hole argument (ibid., Chapter 9, pp. 175–208), Earman applies this concept of model to the formulation of general relativity in terms of the metric tensor and its first and second derivatives. Thus his A_i is restricted to "the metric field g_{ik};" while the P_j correspond to components of the stress energy tensor T_{ik}. He presents a version of "Einstein's Hole Argument," involving a diffeomorphism d, such that

> $d = $ id outside the hole H but $\neq$ id inside H and such that the two pieces join smoothly on the boundary ... The upshot is that we have produced two solutions, $\langle M, g, T \rangle$ and $\langle M, d * g, T \rangle$, which have identical T fields but different g fields – an apparent violation of the *Kausalgesetz* that the T field determines the g field (ibid., p. 176).

He then presents a version of Hilbert's Cauchy problem argument (see Section 2.7). Assuming the existence of a Cauchy surface, parameterized by $t = 0$, he considers:

> a diffeomorphism d such that $d = $ id for all $t \leq 0$ and $\neq$ id for $t > 0$ and such that there is a smooth join at $t = 0$ (ibid., p. 179).

One can then construct two solutions, $\langle M, g, T \rangle$ and $\langle M, d * g, T \rangle$, that do not differ for $t \leq 0$ and sharing the same initial data to any finite order of differentiability on $t = 0$.

> [This provides] a seeming violation of the weakest form of Laplacian determinism ... indeed, any nontrivial form of determinism suffers equally (ibid., p. 179).

The discussion of the range of applicability of the hole argument in Earman (1989) differs significantly from that in Earman and Norton (1987), which maintained that the hole argument applied to "every classical spacetime field theory [that] can be formulated as a local space-time theory." They included all special-relativistic field theories, which they maintained were made local by adjoining the Riemann tensor $R_{abcd}(g)$ to the set of geometric objects included in any model, and adding the equation $R_{abcd} = 0$ to the set of field equations defining acceptable special-relativistic models.

5.2.1   Confusion between the trivial identity and general covariance

I must linger a bit longer on Earman and Norton (1987), because their discussion of general covariance has led to much confusion. They divide geometric object fields $O_1, \ldots O_n$ into two classes: $O_1, \ldots, O_{k-1}$ and $O_k, \ldots, O_n$. The second, dynamical class is assumed to obey field equations

$$O_k = 0, \qquad O_{k+1} = 0 \ldots, \qquad O_n = 0, \tag{4}$$

while the first class may include non-dynamical, fixed background fields.[72] When applied to special relativistic field theories, the Minkowski metric and the associated flat affine connection are included among these non-dynamical fields. In order to demonstrate that the hole argument applies to all such theories (and in contrast to the definition of general covariance in Earman (1989), cited above), Earman and Norton (1987) prove a "Gauge Theorem (General covariance)":

> If $\langle M, O_1, \ldots, O_n \rangle$ is a model of a local spacetime theory and h is a diffeomorphism from M to M, then the carried along n-tuple $\langle M, h * O_1, \ldots, h * O_n \rangle$ is also a model of the theory (ibid., p. 520).

Notice that all the $O_1, \ldots, O_n$ fields of both classes are subject to arbitrary diffeomorphisms, which need not be symmetries of the non-dynamical fields. Thus, this "theorem" is essentially the trivial identity (see Sections 2.5, 4.2 and 5.7). Indeed, we can reformulate the trivial identity in their notation:

If $\langle M, O_1, \ldots, O_n \rangle$ is a model of a local spacetime theory and h is a diffeomorphism from M to M, taking a point $p \in M$ into a point $p' \in M : p \to p'$, then the carried along n-tuple is $\langle M, h * O_1, \ldots, h * O_n \rangle$. If we now carry out a coordinate transformation $x \to x'$, such that $x(p) = x'(p')$, i.e., the new coordinates of the new point equal the old coordinates of the old point; then $[h * O_i]' = O_i$, i.e, the new components of the new geometric objects are numerically equal to the old components of the old objects, then clearly nothing has changed. So it is not clear why the authors feel any:

> need to establish that the vanishing of the field equations $O_k = 0, O_{k+1} = 0 \ldots, O_n = 0$ is preserved under a diffeomorphism (ibid., p. 520).

While Earman (1989) avoids this confusion by silently renouncing this position, much of the later literature on the hole argument still falls into this error.

5.2.2   Back to Earman's evolution

Earman now argues that one cannot simply

> take a special relativistic theory of motion and rewrite the equations using covariant derivatives with respect to an undetermined Lorentz metric g. Then write the "field equation" for g, namely, $R^i_{jkl}(g) = 0$, where R^i_{jkl} is the Riemann curvature tensor (Earman, 1989, p. 183),

and then apply the hole argument to show that the theory is non-deterministic. He introduces

> a distinction between absolute and dynamical objects … [T]his distinction corresponds to the distinction between the [geometric] object fields (A_i) that characterize the structure of space-time and those (P_j) that characterize the physical contents of space-time (ibid., p. 184).

---

[72] The first class will also include dynamical fields, certain derivatives of which will constitute the objects of the second class. For example, the electromagnetic potential vector would be in the first class, while the vectorial wave operator acting on it would produce a object in the second class.

He then requires that,

> for any two dynamically possible models of the theory $\mathcal{M} = \langle M, A_i, \ldots P_j \ldots \rangle$ and $\mathcal{M}' = \langle M', A_i', \ldots P_j' \ldots \rangle$ there is a diffeomorphism $d : M \to M'$ such that $d * A_i = A_i'$ for all i (ibid., p. 184)

Letting $M = M'$, ones sees that the condition $d * A_i = A_i$ singles out those diffeomorphisms d of the manifold that are symmetries of the A_i fields.[73]

Without going into further detail (see ibid., p. 184), Earman essentially argues that the hole argument does not apply if the symmetry group makes the absolute-space time structures sufficiently rigid. It is also clear (although Earman does not make the point) that the trivial identity (Earman and Norton's "Gauge Theorem") is of no help in an attempt to apply the hole argument to such cases.

Earman (2004) continues the line of reasoning in Earman (1989), but with some further evolution: It emphases from the start the difference between finite-parameter Lie symmetry groups, covered by Noether's first theorem; and symmetry groups that are function groups, covered by Noether's second theorem. Earman (2006) starts off in a way reminiscent of Earman and Norton (1987):

> It will be assumed that the spacetime theories to be discusses have been formulated in such a way that (a) their models have the form $(\mathcal{M}, O_1, O_2, \ldots, O_N)$, where $\mathcal{M}$ is a differentiable manifold and the O_n are geometric object fields that live on M, and (b) their laws of motion/field equations have the form $F(\tilde{O}_1, \tilde{O}_2, \ldots, \tilde{O}_K) = 0$, where F is some functional and the $\tilde{O}_k$ are geometric object fields constructed from the O_n (ibid., p. 446).

The difference is that now he separates the field equations $F = 0$ from the quantities defining the model, and relaxes the demand that they be tensorial equations. He introduces the concept of gauge symmetry as a transformation, in which:

> the physical situation is not being changed; rather different but equivalent descriptions of one and the same physical situation are being generated. This is the characteristic feature of a gauge symmetry (ibid., p. 447).

He then defines substantive general covariance (SGC):

> The equations of motion/field equations of the theory display diffeomorphism invariance; that is, if $(\mathcal{M}, O_1, O_2, \ldots, O_N)$ is a solution, then so is $(\mathcal{M}, d * O_1, d * O_2, \ldots, O_N)$ for any $d \in \text{diff}(\mathcal{M})$. *And* this diffeomorphism is a gauge symmetry (ibid., p. 447).

In other words, the two solutions are mathematically distinct descriptions of the same physical solution.

He now states explicitly that "formal general covariance" is "a rather trivial gauge symmetry" (ibid., p. 447).

Earman's "formal general covariance" is what Bergmann (1957) calls "weak covariance", Stachel (1993), following a suggestion by Bergmann,[74] calls "trivial general covariance;" and Stachel and Iftime (2005) call "covariance" *tout court*. Earman's definition of "substantive general covariance"

---

[73] Note that the infinitesimal form of $d * A_i = A_i$ is $\pounds_{\boldsymbol{v}} A_i = 0$ where $\boldsymbol{v}$ is a vector field generating a one-parameter subgroup of the symmetry group.

[74] Bergmann (1957) speaks of the "fundamentally trivial nature" of "weak covariance."

now corresponds to the Stachel and Iftime (2005) definition of "general covariance" – but what's in a name?[75] The two positions are now "substantively" the same.

The remaining difference is mathematical: Rather than using fiber bundles, Earman still works with fields on a manifold, so his formalism is still vulnerable to the substantivalists' attack. As he noted in another context, "Formalism generated the problem and formalism is needed to resolve it" (Earman, 1989, p. 184). Or perhaps it would be better to say: "If you adopt a certain philosophical stance, you should adopt the formalism best suited to it."

5.3 Pooley's position: sophisticated substantivalism

Pooley describes "sophisticated substantivalism" succinctly as a "combination of anti-haecceitism and realism about spacetime points" (Pooley, 2006, p. 103).

> A frequent response [to the argument from Leibniz equivalence] is that one can regard all isomorphic models of general relativity as representing the same physical possibility (*Leibniz Equvalence*) AND regard spacetime as a basic, substantival and concrete entity.
>
> . . .
>
> Sophisticated substantivalism:
>
> Isomorphic models $\mathcal{M}$ and $\mathcal{M}_0$ represent the same physical possibility (= L[eibniz] E[quivalence]) AND spacetime points exist as fundamental entities.
>
> - LE accords with the practice of physics
> - the metric (plus manifold) gets its natural interpretation as spacetime
> - $\mathcal{M}$ and $\mathcal{M}_0$ can only be regarded as representing distinct possible worlds if spacetime points have primitive identity. Denying that they do is good metaphysics independently of the hole argument (Pooley, 2000).

> Sophisticated substantivalism may be compatible with taking seriously physicists' concerns, but does it have a coherent motivation? The obvious thing to be said for the position is that one thereby avoids the indeterminism of the hole argument. This motivation is, of course, rather ad hoc. A less ad hoc motivation would involve a metaphysics of individual substances that does not sanction haecceitistic differences, perhaps because the individuals are individuated by – their numerical distinctness is grounded by – their positions in a structure. . . . Stachel has recently sought to embed his response to the hole argument in exactly this type of more general framework. I hope enough has been said . . . to indicate the coherence of such a point of view; it is perhaps a modest structuralism about spacetime points, but it is a far cry from the objectless ontology of the ontic structural realist (Pooley, 2006, p. 102).

Again, there is "sophisticated substantial" agreement between Pooley's viewpoint and those of Earman and Stachel (see Pooley, 2013, for a more recent account of his position).

5.4 Stachel and dynamic structural realism

My earliest discussions of the hole argument were based on a purely relationalist approach to space time, which denied any physical significance to points of the four-dimensional manifold M; they only became elements of space-time after a metric tensor field was specified. This was largely in response to mathematical formulations of physical field theories in terms of geometric object fields

---

[75] "The position that GTR is not 'just another generally covariant theory' was first clearly articulated in the philosophical literature by Stachel (1993)" (Earman, 2006, p. 464, note 5).

on a given M. If one conceded that the points of this manifold represented elements of space-time, this seemed to hand victory to the absolutists (subsequently metamorphosed into substantivalists). When I realized the full implications of the fiber bundle approach, which allows the definition of M as the quotient of the total manifold of the bundle by the equivalence relation defining the fibration (see Section 4.2); and of Schouten's (1951) observation that, in contrast to mathematical tensor fields, physical tensor fields have physical dimensions; I came to recognize that the points of M, so defined, do have the physical character of elements of space-time even before the choice of a particular field (cross section of the bundle).[76] What they lack is individuality, or haecceity as I put it after adopting Teller's (1998) terminology (see Section 5.5). This led me to a structuralist account of physical theories, but not the kind of structuralism espoused by Ladyman and French (see, e.g., Ladyman, 1998; French and Ladyman, 2003) which they call "ontic structural realism;" but which is really a kind of hyper-relationalism.[77] Stachel (2006a) espouses a form of traditional realism as a philosophical position, and also stresses the priority of processes over states, hence it names this position "dynamic structural realism."

To summarize the last three sections, starting from various relationalist or substantivalist positions, Earman, Pooley and Stachel have been led to a third position, which Earman calls "substantive general covariance," Pooley calls "sophisticated substantivalism," and Stachel calls "dynamic structural realism"; but all three positions are essentially the same. The major difference is Stachel's emphasis on the utility of the fiber bundle approach for the mathematical expression of this position.

After this lengthy historical-critical excursus, I shall turn to some philosophical arguments for this approach, starting with the definition of some terms already given in Section 4 and Appendix B, but repeated here for the benefit of those who did not read that section.

5.5   Relations, internal and external, quiddity and haecceity

A relation is said to be *internal* if one or more essential properties of the relata[78] depend on the relation. It is said to be *external* if no essential property so depends.[79] This distinction is in turn based on the distinction between *intrinsic* and *extrinsic* properties of an entity. Some of its intrinsic properties serve to characterize what has been variously called the essence, nature or natural kind of the entity; if any of these essential intrinsic properties depend on its relation(s) to other entities, then these relations are internal. No extrinsic property can depend on an internal relation.

Whether a relation is internal or external is theory-dependent, and hence may depend on the theoretical level at which the objects are treated. In any physical theory, for example, a set of units must be adopted before a mathematical form can be given to any physical quantity. Its numerical expression is actually a relation – the ratio of the quantity to its unit. At this level, it is an external relation based on the properties of the quantity and its unit. Whether these properties themselves are intrinsic or extrinsic may depend on the theoretical level considered. In the system of units adopted, is the unit of this quantity "basic" or "derived"?

A second important distinction is that between *quiddity* and *haecceity*. Quiddity is what characterizes all entities of the same nature. Haecceity refers to those properties of the relata that enable us to individuate entities of the same quiddity. Up until the last century, it was assumed that entities of the same quiddity could always be individuated by some of their intrinsic properties,

---

[76] I thank Tian Yu Cao for discussions of this point.

[77] Pooley speaks of "the objectless ontology of the ontic structural realist" (Pooley, 2006, p. 102; see Section 5.3). See also Stachel (2006a) for a critique of the concept of "relations without relata."

[78] One distinguishes between a relation and the relata, the entities that it relates.

[79] In the former case one may speak of "things between relations," in the latter more familiar case of "relations between things" (see Stachel, 2002).

independently of any relations, into which they entered. This is Leibniz's principle: the identity of indiscernibles.[80]

Any further individuation due to such relations was supposed to supervene on this basic individuation.[81]

With the advent of quantum statistics, it was argued that there are entities – the elementary particles – that have quiddity (any particle with charge e, mass m_e and spin $1/2$ is an electron) but no inherent haecceity (one cannot distinguish one electron from another by any intrinsic property). And the refutation of hole argument can be similarly formulated: The points of space-time have quiddity but no inherent haecceity. So theoretical physics led to the introduction of a new category: entities having quiddity but no inherent haecceity.

5.6   Structures, algebraic and geometric

An important example of the utility of this category in mathematics is the fundamental distinction between geometric and algebraic structures.

Geometry deals with elements that *have* (the same) *quiddity but lack inherent haecceity*; a set of *internal relations* between these elements then defines a particular *geometric structure*. The *group of permutations* of these elements preserving the defining internal relations is the *symmetry* or *automorphism group* of the geometry. Each geometry has such a group of transformations of its elements, under which all geometrical relations of that geometry remain invariant. [82]

Algebra deals with elements that possess both *quiddity* and *haecceity*; a set of *external relations* between these elements defines a particular *algebraic structure*. [83]

Coordinatization of a geometry by an (appropriate) algebra is the assignment of a unique element of this algebra to each point of the geometry; one can carry out certain algebraic operations and then give the result a geometric interpretation.

Coordinate transformations: Any coordinatization of a geometry gives each of its elements a haecceity, thus negating their homogeneity. This is restored by negating in turn any individual coordinatization: A *group of coordinate transformations* between *all* admissible coordinate systems is introduced.

An *admissible coordinate transformation* is one that corresponds to an element of the automorphism group of the geometry. It follow that *each point* of the geometry will have *every element* of the algebra as its coordinate in (at least) *one* admissible coordinate system.

5.7   Does "general relativity" extend the principle of relativity?

To talk about a *principle of relativity* only makes sense if one has first defined a *frame of reference*. One then asserts that the laws of physics take the same form in all members of some class of frames of reference. In special relativity, this class of frames (actually a group in the case) consists of the *inertial frames of reference*. Given the Minkowski metric and its associated flat inertial

---

[80] "For there are never two things in nature that are perfectly alike and in which it is impossible to find a difference that is *internal*, or founded on *an intrinsic denomination*." (G.W. Leibniz, "The Monadology," cited from Loemker, 1969, p. 643; our emphasis). See Henry (2006) for an extensive discussion of haecceity.

[81] Leibniz's principle is often cited in discussions of the hole argument (see, e.g., Earman, 1989; Norton, 2011).

[82] Euclidean plane geometry provides a simple example. The relations characterizing any triangle are preserved under the automorphism group of the Euclidean plane, consisting of all translations and rotations.

[83] The real numbers provide a simple example, each element of which is uniquely defined. They form a field under the operations of addition and multiplication.

connection, such a frame may be defined by taking any time-like autoparallel ("straight") line, and constructing the family of such lines, one through each point of the manifold (i.e., a *fibration* of the space-time), each of which is *parallel* to the initial line.[84] One may then pick a fiduciary point on each such line, and use the proper time τ along this world line, counted forwards and backwards starting from that point $\tau = 0$, to individuate the points along the line. Assuming that each line is itself physically individuated (given *haecceity*) in some way, all the points of the space-time are now individuated. It is customary to choose all the fiduciary points to lie on the same space-like hyperplane orthogonal to the time-like fibration (Einstein convention for defining distant simultaneity). Then the entire group of inertial frames may be generated from the initial one by the action of the Poincaré group on the points of that inertial frame.

Of course, the trivial identity holds: if we move everything together with some diffeomorphism of the manifold, nothing has changed. But given that we move only the world lines with respect to the metric and connection, the Poincaré group is the automorphism group of the inertial frames. The inertial frames thus form a rigid structure, individuating the points of Minkowski space-time, and the hole argument fails, as it will for any finite-parameter Lie group.

In general relativity, a spatial frame of reference also corresponds to a *fibration* of the four-dimensional manifold M with the stipulation that, when a metric tensor field $\boldsymbol{g}$ is introduced, the fibration consist of curves with a *unit time-like vector field* $\boldsymbol{v}$ tangent to the fibration: $\boldsymbol{v} = dx/d\tau$.[85] We may then define projection operators $'P$ along the foliation and $''P$ orthogonal to it. The vector field $\boldsymbol{v}$ represents the four-velocities of observers in the chosen reference frame and the orthogonal projection of the metric field $^3\boldsymbol{g}$ represents the instantaneous spatial rest-frame of each observer. Again, one may pick a fiduciary point on each time-like world line and use the proper time, forwards and backwards starting from that point, to individuate the points along the line.

Evolution of any geometric object field $\boldsymbol{\Phi}$ along the congruence will be represented by its Lie derivatives with respect to $\boldsymbol{v}, \mathcal{L}_{\boldsymbol{v}}\boldsymbol{\Phi}$. One will usually pick the fiduciary points so that they fit together smoothly to form a space-like hypersurface that transvects the fibration. Now there are two possibilities:

Holonomic case: If one has chosen a congruence, the tangent field $\boldsymbol{v}$ of which has *vanishing rotation* ω, there will be a *foliation of space-time* consisting of a one-parameter family hypersurfaces *orthogonal* to the fibration. The fiduciary points can be chosen to lie on one hypersurface of the foliation, and the local spatial rest-frames of each observer will fit together to form a *global spatial rest-frame*; so that the local spatial rest-frames of each observer fit together to form a one-parameter family of global spatial rest-frames. This is the geometric basis of the traditional approach to the Cauchy problem in general relativity (see Section 2.7).[86]

Non-holonomic case: But there is no need to impose this requirement. It is customary to introduce a triad of orthonormal space-like vectors $\boldsymbol{e}_i (i = 1, 2, 3)$ that, together with $\boldsymbol{v}$, span the tangent space at each point of the manifold. Then, the components with respect to this tetrad of any geometric object field $\boldsymbol{\Phi}$, called the physical components by Pirani, are assumed to be the physically measurable quantities by an observer in that frame at that point. On the assumption that each curve in the three-parameter fibration is physically individuated (given haecceity) in some

---

[84] They will of course be parallel to each other – parallelism in an affine-flat manifold is an equivalence relation.

[85] One might be tempted, by analogy with the special-relativistic case, to require that the curves be autoparallels of the inertia-gravitational connection (geodesics of the metric), but this is not necessary in theory, and in practice often not helpful: autoparallels in a gravitational field tend to form caustics. Einstein's term for a similar construction was "mollusk of reference" (Einstein, 1952); but it was based on a coordinate system rather than a geometric construction.

[86] Of course, there is no need to confine evolution to this fibration. By the introduction of lapse functions and shift vectors, the evolution may be developed along any other foliation and fibration, respectively. But this does not change the circumstance that choice of an initial labeling of the world lines of the fibration with the associated proper times has destroyed diffeomorphism invariance in any but the trivial sense (see next footnote). The lapse and shift functions merely serve to situate points of space-time with respect to this initial individuating field.

way, and that some foliation is introduced to provides the fourth individuating quantity, the hole argument still fails, because a fibration and foliation provide an individuating field (see Section 4.4), whether or not the rotation ω of the foliation vanishes. Indeed, one does not even need a foliation: Just as in the case of SR, if one hypersurface intersecting all the fibers is chosen as the origin for the proper time τ on each fiber (i.e., $\tau = 0$ on this hypersurface); then the proper time on each fiber provides the fourth individuating quantity. The transformation from one fibration with associated proper times to another is merely a change of labeling of the individuation.[87] This individuation evades the hole argument and allows the formulation of the Cauchy problem for the Einstein field equations in terms of Lie derivatives of the tetrad components of the appropriate quantities with respect to any time-like congruence, holonomic (see Stachel, 1969) or non-holonomic (see Stachel, 1980).[88]

Thus, the principle of relativity has been extended beyond inertial frames in Minkowski space-time: the laws of any physical theory based on a geometric object field $\mathbf{\Phi}$, or indeed the laws governing any particle world-lines introduced into the theory, can be formulated with respect to any reference frame based on any such fibration and the associated proper times.

The automorphisms of these reference frames now form a function group, which can be defined by its action on the orthonormal tetrad field (v, e_i) $(i = 1, 2, 3)$ characterizing some initial frame. At any point x, an element of the group $SO(1, 3)$ will take one such tetrad into another $('v, 'e_i)$; such an element depends on six position-dependent parameters (three rotations and three pseudo-rotations). Since any smooth vector field is holonomic, the resulting $'v$ field will generate a new fibration. Four "translations" (i.e., diffeomorphisms of the manifold M that are transitive and effective) will take the "origin" of the first reference frame into the origin of the second.[89]

In this sense, the general theory does extend the principle of relativity from inertial frames in Minkowski space-time to arbitrary orthonormal tetrad frames in pseudo-Riemannian space-times, either given a priori (background-dependent theories) or constructed from a solution to the Einstein equations (background-independent theories such as general relativity).

---

[87] Again, the trivial identity still holds, of course. If we carry along all the fibrations with some diffeomorphism of the manifold, nothing has changed.

[88] Except in the trivial sense, diffeomorphism invariance is destroyed by the introduction of such a frame of reference. Often, an attempt is made to partially restore it by demanding invariance under all permutations of the labeling of the world-lines of the fibration. When a coordinate system is adapted to the fibration, this is described as three-dimensional diffeomorphism invariance or invariance under spatial diffeomorphisms, and the distinction is sometimes elided. "What are the quanta of the gravitational field? Or, since the gravitational field is the same entity as spacetime, what are the quanta of space?" (Rovelli, 2004, p. 18; see also Section 4.1, pp. 145–153).

[89] This is quite similar to Cartan's method of generalizing a homogeneous Klein geometry to what is now called a Cartan geometry (see, e.g., Sharpe, 1997, "Preface" and Chapters 4 and 5).

6   Current Discussions: Physical Issues[90]

6.1   Space-time symmetry groups and partially background-independent space-times

As noted in Section 2.6, in the case of a generic metric (i.e., one having no symmetries), the Kretschmann–Komar coordinates may be used to individuate the points of space-time. While non-generic metrics constitute a subset of measure zero, all known solutions to the Einstein field equations belong to this subset. Only the *a priori* imposition of some fixed, background symmetry group on the pseudo-metric tensor enables construction of such solutions (see, e.g., Stephani *et al.*, 2003). The symmetry group, also called the *isometry group* of the metric, determines a portion of the metric field non-dynamically; the remaining portion obeys a reduced set of dynamical Einstein equations. One must examine each symmetry group to see how much freedom remains in the class of solutions to these reduced equations; in particular, whether enough freedom remains for a restricted version of the hole argument to apply to these solutions.

The possible isometries of a four-dimensional pseudo-Riemannian manifold have been classified: they are characterized by two integers (m, o) (see, e.g., Stephani *et al.*, 2003; Hall, 2004):[91]

1. The dimension $m \leq 10$ of the isometry group.

2. The dimension $o \leq \min(4, m)$ of the highest-dimensional orbits of this group.

The two extreme cases are:

The *maximal symmetry group* $(m = 10, o = 4)$. Minkowski S-T is the unique Ricci-flat space-time in this class. Its isometry group is the Poincaré or inhomogeneous Lorentz group, which acts transitively on the entire space-time manifold. The field equations of special-relativistic field theories must be invariant under this group; they provide the most important physical example of background-dependent theories.

At the other extreme $(m = 0, o = 0)$ is the class of generic metrics already mentioned. Their isometry group is trivial, i.e., consists of only the identity element. Field theories that are covariant under the group of all diffeomorphisms of the underlying four-dimensional differentiable manifold (see Section 4), will include a subclass of the generic metrics among their models. If such a theory is generally covariant, so that all diffeomorphically-related models represent one physical model, it is called a *background-independent theory*.

6.1.1   Non-maximal symmetry groups and partially-fixed backgrounds

In between these two extremes lie all models of a background-independent theory that are restricted by the further requirement that they have some fixed *non-maximal, non trivial* symmetry group. If this symmetry group is a function group, we shall say the theory has a *partially-fixed* background. As noted above, all known exact models of general relativity fall into this category.[92] Considerable work has been done on two classes of such models:

1. The *mini-superspace* cosmological solutions, in which so much symmetry is imposed that only functions of one parameter (the "time") are subject to dynamical equations (see, e.g., Ashtekar *et al.*, 1993a,b; Ashtekar, 2010). Quantization here resembles more the quantization of a system of particles than of waves, and does not seem likely to shed too much light on the generic case.

---

[90] I have discussed a number of these issues at greater length in Stachel (2006a) and Stachel (2009).

[91] The isometry group is also referred to as the group of automorphisms or motions.

[92] For a review of the technique of finding group-invariant solutions to systems of partial differential equations, with applications to general relativity, see Anderson *et al.* (2000).

2. The *midi-superspace* (see Torre, 1999) solutions, notably the cylindrical wave metrics (see Ashtekar and Pierri, 1996; Ashtekar *et al.*, 1997a,b; Bičák, 2000). Here, sufficient freedom remains to include both degrees of freedom of the gravitational field. Diffeomorphisms of a two-dimensional manifold with pseudo-metric are still possible, so a two-dimensional version of the hole argument can be formulated.

Taking advantage of fact that any two-dimensional pseudo-metric is conformally related to two-dimensional Minkowski space, one can adopt a coordinate system in which the two degrees of freedom are represented by a pair of scalar fields obeying non-linear, coupled wave equations in this flat space-time (Stachel, 1966). In addition to static and stationary fields, the solutions include gravitational radiation fields having both states of polarization. Their quantization can be carried out formally as if they were two interacting two-dimensional fields (see Kouletsis *et al.*, 2003). But, of course, the invariance of any results under the remaining diffeomorphism freedom must be carefully examined, as well as possible implications for the generic case.

6.1.2 Small perturbations and the return of diffeomorphism invariance

As noted in Section 4.3, the space of all four-geometries forms a stratified manifold, partitioned into slices. Each slice consists of all geometries having the same isometry group. If a metric in one such slice is then perturbed, unless the perturbation is restricted to lie within the same (or some other) symmetry group, it will move the geometry into the generic slice of the stratified manifold. This observation is often neglected when perturbation-theoretic quantization techniques, developed for special relativistic field theories, are applied to perturbations of the Minkowski metric. Diffeomorphisms of such perturbations cannot be treated as pure gauge transformations on a fixed background Minkowki S-T; they modify the entire causal and inertio-gravitational structure of space-time (see, e.g., Doughty, 1990, Chapter 21). This seems to be the fundamental reason behind the problems that plague the application of special-relativistic quantization techniques to such perturbations.

6.1.3 Asymptotic symmetries

An important class of solutions to the field equations lacks global symmetries, but does have asymptotic symmetries as infinity is approached in null directions. This results in the possibility of separating kinematics from dynamics, and the asymptotic quantization of such solutions (see Komar, 1973, Section VI, and Ashtekar, 1987). The imposition of certain conditions on the behavior of the Weyl (conformal curvature) tensor in the future (past) null limit allows conformal compactification of a large class of space-times (Penrose, 1963) by adjoining boundary null hypersurfaces scri plus $\mathscr{I}_+$ (scri minus $\mathscr{I}_-$) to the space-time manifold. Both boundaries have a symmetry group that is independent of particular dynamical solutions to the field equations in this class.

Thus, there is a separation of kinematics and dynamics on $\mathscr{I}_\pm$, and a quantization based on this asymptotic symmetry group can be carried out in more or less conventional fashion.

"More or less" because the asymptotic symmetry group, the Bondi–Metzner–Sachs (BMS) group, is not a finite-parameter Lie group, as is the Poincaré group; it includes four so-called *supertranslation* functions, which depend on two "angular" variables. Nevertheless, asymptotic gravitons may be defined as representations of the BMS group, independently of how strong the gravitational field may be in the interior region (Ashtekar, 1987).

6.1.4   Generalization of this classification to other geometric object fields

This classification of metrics can be extended to other fields. Consider a natural or gauge-natural
fiber bundle involving a geometric object field $\mathbf{\Phi}$ with base manifold M, and a covariant theory T
that picks out a valid class of sections $\boldsymbol{\sigma}$ of the $\mathbf{\Phi}$-bundle. Since the theory is assumed covariant,
if σ is a model of the theory, then so is $d * \sigma$, where $d*$ is the unique fiber-preserving bundle
diffeomorphism corresponding to a base diffeomorphism d_M if the bundle is natural; or to any
member of the class of such diffeomorphisms if the bundle is gauge-natural.

Consider the equation $\mathcal{L}_{\boldsymbol{v}}\mathbf{\Phi} = 0$, where $\boldsymbol{v}$ is a vector field generating a one-parameter family
of base diffeomorphisms. We say that such a family of diffeomorphisms is a *symmetry of* $\mathbf{\Phi}$
generated by $\boldsymbol{v}$. The class of all models of the theory is divided into equivalence classes by the
following equivalence relation: Two models σ and σ' are equivalent if any $\boldsymbol{v}$ that generates a
symmetry of σ also generates a symmetry of σ' and vice versa.

If there is more than one generator of the symmetries in such an equivalence class, they form a
group under addition with constant coefficients. The Poisson bracket of two symmetry generators:
$[\boldsymbol{v}, \boldsymbol{v}']$ is always a generator of a symmetry; so the generators form a Lie algebra.[93]

If we assume the theory based on $\mathbf{\Phi}$ to be background-independent, i.e., generally covariant,
we can impose a further condition on models of the theory: that they all have the symmetries
generated by some particular Lie sub-algebra. This results in a class of partially background-
dependent theories, each based on the background independent theory.

6.2   General relativity as a gauge theory

There are a number of ways to treat general relativity as a gauge theory (see Sections 4.3 and 4.4;
for a survey that includes some generalizations of general relativity, see Hehl *et al.*, 1995). We
shall follow the discussion in Trautman (1980):

> For me, a gauge theory is any physical theory of a dynamic variable which, at the clas-
> sical level, may be identified with a connection on a principal bundle. The structure
> group G of the bundle P is the group of gauge transformations of the first kind; the
> group $\mathscr{G}$ of gauge transformations of the second kind may be identified with a sub-
> group of the group $\operatorname{Aut} P$ of all automorphisms of P. In this sense, gravitation is a
> gauge theory. The basic gauge field is a linear connection ω (or a connection closely
> related to the linear connection). In addition to ω, there is a metric tensor g which
> plays the role of a Higgs field. The most important difference between gravitation and
> other gauge theories is due to the soldering of the bundle of frames LM to the base
> manifold M. The bundle LM is constructed in a natural and unique way from M,
> whereas a noncontractible M may be the base of inequivalent bundles with the same
> structure group. ... The soldering form θ leads to a torsion which has no analogue in
> nongravitational theories. Moreover, it affects the group $\mathscr{G}$, which now consists of the
> automorphisms of LM preserving θ. This group contains no vertical automorphism
> other than the identity; it is isomorphic to the group $\operatorname{Diff} M$ of all diffeomorphisms of
> M (ibid., p. 306).

Trautman contrasts such theories with a "pure gauge" theory:

> In a gauge theory of the Yang–Mills type over Minkowski space-time, the group $\mathscr{G}$ is
> isomorphic to the semi-direct product of the Poincaré group by the group $\mathscr{G}_0$ of vertical
> automorphisms of P .... In other words, in the theory of gravitation, the group $\mathscr{G}$ of

---

[93] More precisely, it is a Lie algebra over the field of real numbers. The Lie bracket of three generators is bilinear
and satisfies the Jacobi identity, while the Lie bracket of a generator with itself vanishes.

"pure gauge" transformations reduces to the identity; all elements of $\mathscr{G}$ correspond to diffeomorphisms of M (ibid., p. 306).

Trautman points out that, even for a given theory, the choice of structure group G is not unique.

Since space-time M is four-dimensional, if $P = LM$ then $G = GL(4, \mathbf{R})$. But one can equally well take for P the bundle AM of affine frames; in this case G is the affine group. There is a simple correspondence between affine and linear connections which makes it really immaterial whether one works with LM or AM. If one assumes – as usually one does – that ω and g are compatible, then the structure group of LM or AM can be restricted to the Lorentz or Poincaré group, respectively. It is also possible to take, as the underlying bundle for a theory of gravitation, another bundle attached in a natural way to space-time, such as the bundle of projective frames or the first jet extension of LM. The corresponding structure groups are natural extensions of $GL(4, \mathbf{R}), O(1, 3)$ or the Poincaré group (ibid., p. 306).

As indicated earlier (see Section 4.2), for all relevant physical theories there are good arguments for considering $SL(4, \mathbf{R})$ as the frame automorphism group, with the unimodular diffeomorphism group as the corresponding structural group (see Stachel, 2011; Bradonjić and Stachel, 2012).

6.3   The hole argument for elementary particles

To recapitulate the main results of Section 4.1: The essence of the hole argument is independent of the differentiability or even the continuity properties of a manifold. When one abstracts from these properties,[94] a manifold diffeomorphism becomes a permutation of the members of the set; so one can use fibered sets to formulate the hole argument for permutable theories. The *covariance* of a theory defined on a fibered manifold – any valid model of it is turned into another valid model by a diffeomorphism acting on its base manifold – becomes the *permutability* of a theory defined on a fibered set: A theory is *permutable* if any valid model of it is turned into another valid model by a permutation of the elements of its base set. The theory is *generally permutable* if an equivalence class of such mathematically distinct models corresponds to a single model of the theory. These concepts can be applied to elementary particles.

Consider a system consisting of N so-called *identical*, or better, *indistinguishable* elementary particles, sharing a common quiddity. By nature, these particles lack an inherent haecceity; but in order to formulate a dynamical theory for the system (e.g., in order to write down a Lagrangian or Hamiltonian for it in non-relativistic quantum mechanics), one needs to *enumerate* them (i.e., assign a *number* from 1 to N to each of the particles). This is a discrete example of *coordinatization* (see Section 4.1), and one must 'undo' this individuation by requiring invariance under all possible permutations of the initial enumeration.

As discussed in Section 4.1 for sets in general, such permutations can be done in either of two ways:

- Active: Fix the enumeration and permute the particles; or

- Passive: Fix the particles and permute the enumeration.

Either way, it is obvious that *each* of the N particles will have *each* of the integers from 1 to N assigned to it in some of these permutations.

Like space-time points, particles of the same natural kind (quiddity) can only be individuated (to the extent that they are) by their position in some relational structure in a theory that is

---

[94] In category theory, the generalization from the category of manifolds to the category of sets can be carried out by the use of *forgetful functors*.

permutable in both the active and passive senses: If some state of affairs is possible for a system that includes N indistinguishable particles, then the state of affairs resulting from the action of any element of the permutation group Perm(N) on these N particles must be an identical state of affairs. In other words, the theory must be generally permutable.

What is a "possible state of affairs" in quantum mechanics? As discussed in the Appendix B, one may adopt a synchronic or diachronic approach to such questions. The term "state of affairs" is often interpreted synchronically, as referring to the instantaneous state of the system (e.g., its state vector or wave function). But quantum theory can only treat open systems (again see the Appendix B); and its task is *diachronic*: to compute the *probability amplitude* for a *complete process* (or "phenomenon" in Bohr's terminology).

In short, only a complete process constitutes a possible state of affairs for a quantum-mechanical system (see Stachel, 1997). If the system is generally permutable and a certain value of the probability for such a process is calculated, then the *same* value must be calculated for any process that results from this one by a permutation of the indistinguishable particles in either the initial act of *preparation* or (inclusive "or") the final act of *registration*. In order to verify these probabilities, some macroscopic device must register the results of the preparation and registration.

> The macroscopic counter is assumed to be inherently individuated. It seems that, for such individuation of an object, a level of structural complexity must be reached, at which it can be uniquely and irreversibly "marked" in a way that distinguishes it from other objects of the same nature (quiddity). My argument is based on an approach, according to which quantum mechanics does not deal with quantum systems in isolation, but only with processes that such a system can undergo. .... A process (Feynman uses process, but Bohr uses "phenomenon" to describe the same thing) starts with the preparation of the system, which then undergoes some interaction(s), and ends with the registration of some result (a "measurement"). In this approach, a quantum system is defined by certain essential properties (its quiddity); but manifests other, non-essential properties (its haecceity) only at the beginning (preparation) and end (registration) of some process. (Note that the initially-prepared properties need not be the same as the finally-registered ones.) The basic task of quantum mechanics is to calculate a probability amplitude for the process leading from the initially prepared values to the finally-registered ones. (I assume a maximal preparation and registration – the complications of the non-maximal cases are easily handled) (Stachel, 2006a).

Consider a scattering process for a system of indistinguishable particles, for example. The process consists of the transition from some initial free in-state of the particles to some final free out-state of the particles after their interaction with the target producing the scattering. The cross section for this process depends on the choice of initial in- and out-states.[95] For a permutable theory, if some value for such a cross section is a possible result, then the same value must result for the processes that arise from separate permutations of the particles in both the in-state and the out-state.

6.3.1 The analogue of the hole argument for a permutable theory

Suppose the theory were permutable but not generally permutable. Then specification of a unique initial preparation for some process (an experiment) could never result in a unique prediction of the outcome of the registration: Every possible permutation of the particles in the registration result would be equally probable.

---

[95] Of course in a relativistic version of scattering theory, the number of particles of the same kind in the incoming and outgoing states need not be the same. But maximal permutability is still required of both the in- and out-states.

Conversely, registration of a unique result (an observation) could never produce a unique retrodiction of the preparation that led to this outcome: Every possible permutation of the particles in the act of preparation would be equally probable.

The way out of this non-uniqueness paradox, of course, is to require that any permutable theory involving indistinguishable particles be *generally permutable*. Then all such permutations belong to the *same* process.

Similar arguments apply to special-relativistic quantum field theory. Here the basic entities are elementary particles (fermions) and field quanta (bosons).

> I reserve the term "elementary particles" for fermions and "field quanta" for bosons, although both are treated as field quanta in quantum field theory. I aim thereby to recall the important difference between the two in the classical limit: classical particles for fermions and classical fields for bosons. . . . [I]n the special-relativistic theories, a preparation or registration may involve either gauge-invariant field quantities or particle numbers.

> At the level of non-relativistic quantum mechanics for a system consisting of a fixed number of particles of the same type, this [difference] is seen in the need to take into account the bosonic or fermionic nature of the particle in question by the appropriate symmetrization or anti-symmetrization procedure on the product of the one-particle Hilbert spaces . . . . At the level of special-relativistic quantum field theory, in which interactions may change particle numbers, it is seen in the notion of field quanta, represented by occupation numbers (arbitrary for bosons, either zero or one for fermions) in the appropriately constructed Fock space; these quanta clearly lack individuality (Stachel, 2006a).

As discussed in Section 4.3, in background-independent theories such as general relativity, an analogous principle of general covariance holds for the points of space-time. This common lack of haecceity suggests that, whatever their nature, the fundamental entities of any theory purporting to explain a deeper physical level should satisfy the principle of maximum permutability.

6.3.2 Principle of maximal permutability

Thiemann has pointed out that

> The concept of a smooth space-time should not have any meaning in a quantum theory of the gravitational field where probing distances beyond the Planck length must result in black hole creation which then evaporate in Planck time, that is, spacetime should be fundamentally discrete. But clearly smooth diffeomorphisms have no room in such a discrete spacetime. The fundamental symmetry is probably something else, maybe a combinatorial one, that looks like a diffeomorphism group at large scales (Thiemann, 2001).

It is hard to believe that, having had to renounce intrinsic individuality at the level of field quanta in QFT and at the level of events in GR, that it will reemerge as we go to a deeper level, from which both QFT and GR will emerge in the appropriate limits.

> [T]he way to assure the inherent indistinguishability of the fundamental entities of the theory is to require the theory to be formulated in such a way that physical results are invariant under all possible permutations of the basic entities of the same kind . . . I have named this requirement the principle of maximal permutability. . . . The exact content of the principle depends on the nature of the fundamental entities. For theories, such as non-relativistic quantum mechanics, that are based on a finite number of

discrete fundamental entities, the permutations will also be finite in number, and maximal permutability becomes invariance under the full symmetric group. For theories such as general relativity, that are based on fundamental entities that are continuously, and even differentiably related to each other, so that they form a differentiable manifold, permutations become diffeomorphisms. For a diffeomorphism of a manifold is nothing but a continuous and differentiable permutation of the points of that manifold. So, maximal permutability becomes invariance under the full diffeomorphism group. Further extensions to an infinite number of discrete entities or mixed cases of discrete-continuous entities, if needed, are obviously possible (Stachel, 2006a).

Current versions of string theory, for example, do not meet this criterion, and it has been suggested that an ultimately satisfactory version of that theory will have to be background independent (see the discussions in Stachel, 2006a,b; Greene, 2004). On the other hand, various discretized space-time theories, such as causal set theory, do seem to meet this criterion; but have other problems (see the discussion in Stachel, 2006a).

The next Section 6.4 deals with a much more modest problem: attempts to quantize the field equations of general relativity based on the application of some variant of the standard techniques for the quantization of field theories.

6.4   The problem of quantum gravity

There is a well known tension between the methods of quantum field theory and the nature of general relativity. The methods of quantization of pre-general-relativistic theories are based on the existence of some fixed, background space-time structure(s) with a given symmetry group.[96] The space-time structure is needed both for the development of the formalism and – equally importantly – for its physical interpretation (see Dosch *et al.*, 2005). It provides a fixed kinematical background for the dynamical theory to be quantized: The dynamical equations for particle or fields must be invariant under all automorphisms of the symmetry group.

As we have seen, general relativity does not fit this pattern. It is a background-independent theory, with no fixed, non-dynamical structures. To recapitulate, its field equations are generally covariant under all differentiable automorphisms (diffeomorphisms) of the underlying manifold, the points of which have no haecceity (and hence are indistinguishable) until the dynamical fields are specified. In a background-independent theory, there is no kinematics independent of the dynamics. This applies to both the homogeneous Einstein equations, and to the inhomogeneous Einstein equations coupled to the dynamical equations for any non-gravitational fields. And this is still the case if the automorphism group is restricted to the unimodular diffeomorphisms (see Section 4.2).

However general relativity and (special-relativistic) quantum field theory do share one fundamental feature that often is not sufficiently stressed: the primacy of processes over states. The four-dimensional approach, emphasizing processes in regions of space-time, is basic to both (see, e.g., Stachel, 2006a; Reisenberger and Rovelli, 2002; DeWitt, 2003). The ideal approach to quantum gravity would be a diffeomorphism-invariant method of quantization that takes process as primary. However, the most successful approach so far, loop quantum gravity, only goes a certain way in this direction: It singles out a preferred fibration and foliation (see Section 5.7); and by adopting a $(3 + 1)$ Hamiltonian approach and a set of redundant variables subject to constraints on the initial hypersurface, it effects a certain separation between kinematics and dynamics. But given these limitations, it has produced a mathematically rigorous – and surprisingly a unique (but see Nicolai and Peeters, 2007) – kinematic Hilbert space (see Ashtekar, 2010, for a brief review).

---

[96] That is, the group of automorphisms of the space-time manifold preserving the background structures.

Even in non-relativistic quantum mechanics, the basic goal is to calculate a probability amplitude for a process connecting some initial preparation to some final registration. However, the existence of an absolute time allows one to choose a temporal slice of space-time so thin that it is meaningful to speak of "instantaneous measurements" of the initial and final states of the system (see Micanek and Hartle, 1996). But this is not so for measurements in (special-relativistic) quantum field theory, nor in general relativity (see, e.g., Bohr and Rosenfeld, 1933, 1979; Bergmann and Smith, 1982; Reisenberger and Rovelli, 2002; Stachel and Bradonjić, 2013). The breakup of a four-dimensional space-time region into lower-dimensional sub-regions – in particular, into a one parameter family of three-dimensional hypersurfaces – raises a problem for measurements in both quantum field theory (see Bohr and Rosenfeld, 1933, 1979; DeWitt, 2003) and general relativity (see Bergmann and Smith, 1982). A breakup of a process into the evolution of instantaneous states on a family of space-like hypersurfaces is a useful, perhaps sometimes indispensable, calculational tool. But no fundamental significance should be attached to such breakups, and mathematical results obtained from them should be carefully examined for their physical significance from the four-dimensional, process standpoint (see, e.g., Nicolai and Peeters, 2007).

An important criterion, called "measurability analysis" (see Bergmann and Smith, 1982), based on "the relation between formalism and observation" (see Reisenberger and Rovelli, 2002), sheds light on the physical implications of any formalism. The possibility of the definition of some physically significant quantity within a theoretical framework should coincide with the possibility of its measurement in principle; i.e., by means of idealized measurement procedures consistent with that theoretical framework. This is true at the classical level, at which any complete classical set of physical observables should be measurable in principle by a single compound procedure.[97] This criterion is easily met by unconstrained systems, such as a set of non-relativistic particles interacting via two-body potentials, or a scalar field obeying the classical Klein–Gordon equation. But delicate problems arise in applying it to constrained dynamical systems, and in particular to gauge field theories, including general relativity (see Section 6.2).

---

[97] Under what conditions such a measurement process, which even ideally always occupies a four-dimensional region of space-time, can be modeled as occupying a lower-dimensional region is a delicate one. See, e.g., Micanek and Hartle (1996) for a discussion of non-relativistic quantum mechanics.

7   Conclusion: The Hole Argument Redivivus

The problem becomes even more delicate for quantum systems, in which the existence of the quantum of action h is taken into account. The finite value of h precludes the measurement of a complete set of classical observables by a single compound procedure. It becomes important to show that a complete set of quantum observables, as defined by the theory, can indeed be so measured in principle. Non-relativistic quantum mechanics and quantum electrodynamics, have been show to meet this criterion; and it has been employed as a test of proposals for what should be the fundamental physical quantities defined in quantum gravity (see Bergmann and Smith, 1982; Amelino-Camelia and Stachel, 2009). Rovelli (2004) and Oeckl (2008, 2013) have shown how to define such measurements on the hypersurface bounding a four-dimensional region of space-time, even in a background-independent theory.

> In field theory, the analog of the data set (x, t, x', t') is the couple $[\Sigma, \varphi]$, where Σ is a 3d surface bounding a finite spacetime region, and φ is a field configuration on Σ. ... The data from a local experiment (measurements, preparation, or just assumptions) must in fact refer to the state of the system on the entire boundary of a finite spacetime region. The field theoretical space $\mathcal{G}$ is therefore the space of surfaces Σ and field configurations φ on Σ. Quantum dynamics can be expressed in terms of an [probability] amplitude $W[\Sigma, \varphi]$. Following Feynman's intuition, we can formally define $W[\Sigma, \varphi]$ in terms of a sum over bulk field configurations that take the value φ on the boundary Σ. ... Notice that the dependence of $W[\Sigma, \varphi]$ on the geometry of Σ codes the spacetime position of the measuring apparatus. In fact, the relative position of the components of the apparatus is determined by their physical distance and the physical time elapsed between measurements, and these data are contained in the metric of Σ. ... What is happening is that in background-dependent QFT we have two kinds of measurements: those that determine the distances of the parts of the apparatus and the time elapsed between measurements, and the actual measurements of the fields' dynamical variables. In quantum gravity, instead, distances and time separations are on an equal footing with the dynamical fields. This is the core of the general relativistic revolution, and the key for background-independent QFT (Rovelli, 2004, p. 23).

In this sense, Einstein's hole, as a symbol of process, has reasserted its physical primacy over Hilbert's Cauchy surface, as a symbol of instantaneous state (see Section 2.7).

A   Nijenhuis on the Formal Definition of Natural Bundles

This appendix contains excerpts from Nijenhuis (1994, pp. 109–110).

Natural bundles ... are defined through *functors* that, for each type of geometric object, associate a fiber bundle with each manifold. To formalize this, we need two categories.

First, consider the category $\mathscr{M}f_m$ of m-dimensional (smooth) manifolds. Its morphisms are diffeomorphisms (into). Every open set of an m-manifold belongs to $\mathscr{M}f_m$: the theory is fundamentally a local one.

Second, consider the category $\mathscr{FM}$ of fibered manifolds (N, p, M), where M, N are manifolds and $p : N \to M$ is a surjective submersion. The inverse images $p^{-1}(x)$, for $x \in M$, are the fibers, N is the total space, and M the base space. The morphisms of $\mathscr{FM}$ are the fiber-preserving (smooth) maps. The base functor $B : \mathscr{FM} \to \mathscr{M}f$ assigns to each fibered manifold (N, p, M) its base manifold M and to each morphism in $\mathscr{FM}$ the induced map on the base spaces.

With these definitions, a *bundle functor* on $\mathscr{M}f_m$, or a *natural bundle* over m-manifolds, is a covariant functor $F : \mathscr{M}f_m \to \mathscr{FM}$ with these simple properties:

(1) (Prolongation) The base space of the fibered manifold FM is M itself.

(2) (Locality) If U is an open subset of M, then the total space of FU is $p^{-1}(U)$, the part of N above U.

...

The concept of natural bundle was first formalized by the reviewer. It was little more than a definition, however, until some real theorems were proved. Work by Epstein and Thurston shows that natural bundles are of finite order (i.e., the structure group is a homomorphic image of finite-order jets of diffeomorphisms of $\mathbb{R}^m$ fixing the origin). In addition, their work shows that natural bundles have a natural smooth structure that automatically satisfies a regularity condition (not stated here) in the original definition of natural bundle. Basic to all of this is Peetre's Theorem, with a number of refinements. Other fundamental work, initiated by Palais and Terng, deals with the classification of natural vector bundles.

...

The pursuit of D[ifferntial]G[eometry] consists to a large extent of performing operations on sections of natural bundles. Connections are constructed from Riemann metrics, covariant derivatives are taken, Lie brackets of vector fields are formed, etc. These operations are natural; they commute with point transformations, yield smooth sections in natural bundles from the same, and have non-increasing supports. All such natural operations are, as implied by Peetre-like theorems, of finite order and so induce natural transformations between corresponding jet bundles.

B   Appendix to Section 4: Some Philosophical Concepts

This appendix discusses some concepts that may be unfamiliar to some readers, non-philosophers in particular; and some concepts that may be familiar, but are used here in unfamiliar ways.

B.1   Things: Synchonic states and diachronic processes

Every theory involves things,[98] their properties, and the relations between them. A physical thing occupies some region of space. As for time, there are two main approaches. The diachronic approach treats a thing as a process, developing over time.

> [A] thing, insofar as it is more than an instantaneous occurrence and has duration through time, is a process. This introduces some odd results in our ways of talking. For example, talking would be a process but we would hardly talk of it as a "thing"; similarly, it is not usual to talk of a rock or a human being as a process (Wartofsky, 1968, p. 332).

The synchronic approach characterizes a thing as it exists at one moment of time, often called its state at time t. A diachronic process is treated as a time-ordered sequence of such synchronic states.

B.2   Open versus closed systems: Determinism versus causality

A closed system is a collection of things not subject to any external influence. Determinism is associated with closed systems and implies fatalism: Nothing external can alter the course of a process in a closed system. An open system is a collection of things that are subject to external influences. Causality is associated with open systems and implies control: By manipulating these external factors, one can affect the course of a process occurring in an open system.

> Most systems in the world are open systems; closed systems are the rare exception in nature. ...If all systems were naturally closed, there would be no need for – nor indeed any possibility of – experiment. ...[E]xperiments help us to formulate and test hypotheses about things, structures, and mechanisms (hereafter structures for short) that underlie the world of events.... [T]he laws that express the nature of these structures operate as tendential laws...For example, Galileo's law of falling bodies asserts that gravity produces a tendency of all objects to fall to earth with the same acceleration. Due to the effect on it of air resistance, which varies with velocity, an object actually falls to earth with varying acceleration – or even none, if it reaches terminal velocity.... A body subject to gravity may even move upward if it is immersed in a liquid, the buoyant force of which on it is greater than the gravitational force.... [I]f one does not distinguish between events and underlying mechanisms, and hence fails to recognize that laws are tendential, it is hard to recognize ...that the upward motion of an air bubble in water is just as much an expression of the law of gravity as the downward motion of a stone in air (Stachel, 2003, pp. 144–145).

We almost always construct models of bounded systems undergoing finite processes. Such a system is confined to a finite, bounded region of space-time. Its surroundings, or environment, consist of the contents of those parts of space-time external to this region that are capable of influencing the system. At the boundary of such a finite process, new data can be used to initiate

---

[98] One often refers to things as objects or subjects, depending on the role that they play in the theory.

a new process (preparation), fed into the system as it evolves (information), and extracted from it to record the final outcome of the process (registration).

Example: Skipping rope: If each end of a (finite) rope is held by one person, they can determine its shape at some initial time; their movement of the two ends will then help to change the evolution of the rope's shape; and a photograph can be taken of its final shape at a later time just before they let go.

It is often assumed that, if an open system is enlarged sufficiently by adjoining to it enough of its environment, this will ultimately result in a closed system. But this is a philosophical assumption, not a physical result, even if the enlargement is cosmological in scope. Open cosmological models have been proposed; and the rejection of one model for physical reasons, such as the steady state universe, does not imply that all cosmological models must be closed.

From this discussion, it is clear that the concepts of things and processes are intertwined with those of space and time. As our views on the nature of space and time change, so must our concepts of things and processes.

B.3 Properties: Intrinsic and extrinsic

As mentioned above, each thing is characterized by its properties. Property is often defined as a one-place relation:

> Properties will have to be counted among the relations, just as 1 is taken to be a natural number (Weyl, 1949, p. 4)

However, the word is used here without such a restriction. Rather, one must distinguish between intrinsic properties, which are indeed one-place relations; and extrinsic properties, which are not.

Intrinsic properties: Those one place relations of an entity that serve to characterize its nature, essence or natural kind. An entity that does not possess these properties cannot be of the same kind as an entity that does.

Example: Electrons and protons differ in their charge and mass, which are intrinsic properties. Hence, electrons and protons in the same atom are of different natural kinds, while an electron on Earth and an electron on Mars are both of the same kind.

Extrinsic properties: Those multi-place relations of an entity that do not serve to characterize its nature. Entities of the same natural kind will generally differ in their extrinsic properties.

Example: Its position, energy, and momentum are extrinsic properties of an elementary particle. They relate it to some independently-defined external frame of reference.

In other words, the intrinsic properties of a thing are independent of its relation to any other thing, while its extrinsic properties do depend on its relation(s) with other things.[99]

B.4 Relations: Internal and external

An N-place relation R involves N things ϑ_i in an ordered sequence: $R(\vartheta_1, \vartheta_1, \ldots, \vartheta_N)$. Although $N = 1$ is not excluded (see the discussion of properties), relations proper always involve two or more things, often called the relata.

Example: Parent and child are the relata in the two-place relation of parenthood. In general the order of the N things is important. Invariance under some permutation(s) of its relata is a special property of a relation.

---

[99] Note that this distinction is theory dependent. For, example, a property that is intrinsic at one level, may prove to be extrinsic at a deeper level.

One must distinguish between internal and external relations:[100]

Internal relations: There are relations, in which the relata are primary and their relation is secondary: no essential property of the relata depends on the particular relation under consideration. In this case, we customarily speak of the relations between things.

External relations: However, there are also relations, in which the relation is primary and the relata are secondary: at least one essential property of each of the relata must depend on the relation. In this case, one may speak of the things between relations (Stachel, 2002).

In order to make sense of the distinction between these two types of relations, one must assume that there is a distinction between the essential and non-essential or accidental properties of each thing. Such a distinction can be traced back to Aristotle:

> [I]t is clear that each individual thing is one and the same with its essence, and not accidentally so, but because to understand anything is to understand its essence (Metaphysics, Book IV).

Until recently, it was assumed that all relations – internal and external – supervened on intrinsic properties that uniquely defined each of the relata. Hermann Weyl warned against such an assumption:

> The (explicit or implicit) assumption that every relation must be based on [intrinsic] properties has given rise to much confusion in philosophy [the original German text is much stronger: "viel Unheil angerichtet" = "has wrought much havoc"]. A statement asserting, say, that one rose is differently colored from a second is indeed founded on the fact that one is red, the other yellow. But the relation 'the point A lies on the left of B' is not based on a qualitatively describable position of A alone and of B alone. The same holds for kinship relations among people (Weyl, 1949, p. 4).

B.5   Quiddity and haecceity

Two terms useful in discussing this question are quiddity and haecceity.

Quiddity: What characterizes all entities having the same nature, i.e., the same intrinsic properties.
   Example: All electrons have the same quiddity, as do all goats.

Haecceity: What individuates entities of the same quiddity.
   Example: An electron bound to an atom on Earth and an electron bound to an atom on Mars differ in their haecceity.

We can now reformulate the earlier observation: Up until the last century, it was assumed that each entity of some quiddity always had an intrinsic haecceity; i.e., it could always be individuated independently of any relations, into which it entered. This assumption is often referred to as Leibniz's principle:

> For there are never two things in nature that are perfectly alike and in which it is impossible to find a difference that is internal, or founded on an intrinsic denomination (G.W. Leibniz, "The Monadology" in Loemker, 1969).

---

[100] Rorty (1967) gives an excellent account of discussions about internal and external relations.

Any further individuation of a thing through some of its relations was assumed to supervene on its intrinsic haecceity.

Example: As a biological organism, Bill Gates has the same quiddity as all human beings, but a haecceity based on his individuating his physical characteristics (e.g., fingerprints). He is further individuated socially by his relationship to other members of the IT community as the Chairman of Microsoft.

However, the advent of quantum statistics led to the recognition that elementary particles indeed have quiddity but no inherent haecceity. As noted above, every electron has the same mass, charge and spin, which fix its quiddity; but an individual electron is only singled out (to the extent that can be) by its relations to other entities, such as its effect on the emulsion of a photographic plate.

Einstein's way of avoiding the hole argument can be similarly formulated: The points of space-time have the same quiddity, but no intrinsic haecceity. A space-time point is only singled out (to the extent that can be) by the unique physical properties of the fields at that point.

So both relativity and quantum theory lead to the same conclusion: Leibniz's principle is not universally applicable. There is a category of entities with quiddity but no inherent haecceity. Given that both general relativity and quantum mechanics are based on such entities, it is difficult to believe that, in any theory purporting to underlie both relativity and quantum theory, inherent individuality would re-emerge in its fundamental entities, whatever they are (see Section 6.4).

References

Amelino-Camelia, G. and Stachel, J., 2009, "Measurement of the space-time interval between two events using the retarded and advanced times of each event with respect to a time-like world-line", *Gen. Relativ. Gravit.*, **41**, 1107–1124. [DOI], [ADS], [arXiv:0710.5608 [gr-qc]].   (Cited on page 50.)

Anderson, I.M., Fels, M.E. and Torre, C.G., 2000, "Group Invariant Solutions Without Transversality", *Commun. Math. Phys.*, **212**, 653–686. [DOI], [ADS], [arXiv:math-ph/9910015].   (Cited on page 42.)

Ashtekar, A., 1987, *Asymptotic Quantization*, Bibliopolis, Napoli, Italy.   (Cited on page 43.)

Ashtekar, A., 2010, "The Issue of the Beginning in Quantum Cosmology", in *Einstein and the Changing Worldviews of Physics*, Proceedings of the 7th Conference on the History of General Relativity, La Orotava, Tenerife, March 2005, (Eds.) Lehner, C., Renn, J., Schemmel, M., Einstein Studies, 12, pp. 347–363, Birkhäuser, Boston; Basel.   (Cited on pages 42 and 48.)

Ashtekar, A. and Pierri, M., 1996, "Probing quantum gravity through exactly soluble midisuperspaces I", *J. Math. Phys.*, **37**, 6250–6270. [DOI], [ADS], [arXiv:gr-qc/9606085].   (Cited on page 43.)

Ashtekar, A., Tate, R. and Uggla, C., 1993a, "Minisuperspaces: Observables and Quantization", *Int. J. Mod. Phys. D*, **2**, 15–50. [DOI], [ADS], [arXiv:gr-qc/9302027].   (Cited on page 42.)

Ashtekar, A., Tate, R.S. and Uggla, C., 1993b, "Minisuperspaces: Symmetries and Quantization", in *Directions in General Relativity, Vol. 1*, Proceedings of the 1993 International Symposium, Maryland: Papers in honor of Charles Misner, (Eds.) Hu, B.L., Ryan Jr, M.P., Vishveswara, C.V., pp. 29–42, Cambridge University Press, Cambridge; New York.   (Cited on page 42.)

Ashtekar, A., Bičák, J. and Schmidt, B.G., 1997a, "Asymptotic structure of symmetry-reduced general relativity", *Phys. Rev. D*, **55**, 669–686. [DOI], [ADS], [arXiv:gr-qc/9608042].   (Cited on page 43.)

Ashtekar, A., Bičák, J. and Schmidt, B.G., 1997b, "Behavior of Einstein-Rosen waves at null infinity", *Phys. Rev. D*, **55**, 687–694. [DOI], [ADS], [arXiv:gr-qc/9608041].   (Cited on page 43.)

Belot, G. and Earman, J., 2001, "Pre-Socratic quantum gravity", in *Physics Meets Philosophy at the Planck Scale: Contemporary Theories in Quantum Gravity*, (Eds.) Callender, C., Huggett, N., chap. 10, pp. 213–255, Cambridge University Press, Cambridge; New York. [DOI].   (Cited on pages 15 and 32.)

Bergmann, P.G., 1957, "Topics in the Theory of General Relativity", in *Lectures in Theoretical Physics: Brandeis Summer Institute, 1957*, Brandeis Summer School 1957, pp. 1–44, W.A. Benjamin, New York.   (Cited on page 36.)

Bergmann, P.G. and Smith, G.J., 1982, "Measurability Analysis for the Linearized Gravitational Field", *Gen. Relativ. Gravit.*, **14**, 1131–1166. [DOI], [ADS].   (Cited on pages 49 and 50.)

Bičák, J., 2000, "Exact radiative spacetimes: some recent developments", *Ann. Phys. (Leipzig)*, **9**, 207–216. [DOI], [ADS], [gr-qc/0004031].   (Cited on page 43.)

Bohr, N. and Rosenfeld, L., 1933, "Zur Frage der Messbarkeit der elektromagnetischen Feldgrössen", *Mat.-Fys. Medd. K. Dan. Vid. Selsk.*, **12**(8), 3–65. Online version (accessed 15 May 2013):

http://www.sdu.dk/en/bibliotek/materiale+efter+type/hostede+ressourcer/matfys.
(Cited on page 49.)

Bohr, N. and Rosenfeld, L., 1979, "On the Question of the Measurability of Electromagnetic
Field Quantities", in *Selected Papers of Léon Rosenfeld*, (Eds.) Cohen, R.S., Stachel, J., Boston
Studies in the Philosophy of Science, 21, pp. 357–400, D. Reidel, Dordrecht; Boston. [DOI].
(Cited on page 49.)

Bradonjić, K. and Stachel, J., 2012, "Unimodular conformal and projective relativity", *Europhys.
Lett.*, **97**, 10001. [DOI], [ADS], [arXiv:1110.2159 [gr-qc]]. (Cited on pages 25, 31, and 45.)

Bruhat, Y., 1962, "The Cauchy Problem", in *Gravitation: An Introduction to Current Research*,
(Ed.) Witten, L., pp. 130–168, Wiley, New York; London. (Cited on page 15.)

Darmois, G., 1927, *Les équations de la gravitation einsteinienne*, Mémorial des Sciences
Mathématiques, 25, Gauthier-Villars, Paris. Online version (accessed 14 May 2013):
http://www.numdam.org/item?id=MSM_1927__25__1_0. (Cited on page 15.)

DeWitt, B.S., 2003, *The Global Approach to Quantum Field Theory, 2 vols.*, International Series of
Monographs on Physics, 114, Clarendon Press, Oxford; New York. (Cited on pages 48 and 49.)

Dorato, M., 2000, "Substantivalism, Relationism, and Structural Spacetime Realism", *Found.
Phys.*, **30**, 1605–1628. [DOI]. (Cited on page 33.)

Dosch, H.G., Müller, V.F. and Sieroka, N., 2005, *Quantum Field Theory in a Semiotic Perspective*,
Math.-Phys. Kl. Heidelberger Akad. Wiss., 17, Springer, Berlin; New York. [Google Books]. Online
version (accessed 16 November 2013):
http://philsci-archive.pitt.edu/1624/. (Cited on page 48.)

Doughty, N.A., 1990, *Lagrangian Interaction: An Introduction to Relativistic Symmetry in Elec-
trodynamics and Gravitation*, Addison-Wesley, Reading, MA. (Cited on page 43.)

Earman, J., 1989, *World Enough and Space-Time: Absolute Versus Relational Theories of Space
and Time*, MIT Press Classics, MIT Press, Cambridge, MA; London. (Cited on pages 16, 34,
35, 36, 37, and 39.)

Earman, J., 2004, "Laws, Symmetry, and Symmetry Breaking: Invariance, Conservation Principles,
and Objectivity", *Philos. Sci.*, **71**, 1227–1241. [DOI]. (Cited on page 36.)

Earman, J., 2006, "Two Challenges to the Requirement of Substantive General Covariance", *Syn-
these*, **148**, 443–468. [DOI]. (Cited on pages 36 and 37.)

Earman, J. and Norton, J.D., 1987, "What Price Spacetime Substantivalism? The Hole Story",
Brit. J. Phil. Sci., **38**, 515–525. [DOI]. (Cited on pages 16, 32, 34, 35, and 36.)

Ehlers, J., 1973, "The Nature and Structure of Spacetime", in *The Physicist's Conception of
Nature*, Symposium on the Development of the Physicist's Conception of Nature in the Twentieth
Century, held in Trieste, Italy, 18–25 September 1972, (Ed.) Mehra, J., pp. 71–91, D. Reidel,
Dordrecht; Boston. [Google Books]. (Cited on page 14.)

Einstein, A., 1905, "Zur Elektrodynamik bewegter Körper", *Ann. Phys. (Leipzig)*, **17**, 891–921.
[DOI]. Online version (accessed 15 May 2013):
http://echo.mpiwg-berlin.mpg.de/MPIWG:FDQUK2HX. (Cited on page 8.)

Einstein, A., 1916, "Die Grundlage der allgemeinen Relativitätstheorie", *Ann. Phys. (Leipzig)*, **49**, 769–822. [DOI]. Online version (accessed 15 May 2013): http://echo.mpiwg-berlin.mpg.de/MPIWG:ACAKHYZX. (Cited on page 13.)

Einstein, A., 1918, "Prinzipielles zur allgemeinen Relativitätstheorie", *Ann. Phys. (Leipzig)*, **55**, 241–244. [DOI]. Online version (accessed 15 May 2013): http://echo.mpiwg-berlin.mpg.de/MPIWG:T4OG38NP. (Cited on pages 10 and 14.)

Einstein, A., 1931, "Einstein Says 'Several' Here Understand Relativity Theory", *New York Times*, (March 5). (Cited on page 17.)

Einstein, A., 1952, *Relativity: The Special and the General Theory*, Crown, New York, 15th edn. [Google Books]. (Cited on pages 17, 32, and 40.)

Einstein, A., 1955, *The Meaning of Relativity*, Stafford Little Lectures, 1921, Princeton University Press, Princeton, NJ, 5th edn. [Google Books]. (Cited on page 17.)

Einstein, A., 1956, "Autobiographische Skizze", in *Helle Zeit–Dunkle Zeit: In memoriam Albert Einstein*, (Ed.) Seelig, C., pp. 9–17, Europa Verlag, Zürich; Stuttgart. (Cited on page 8.)

Engler, F.O. and Renn, J., 2013, "Hume, Einstein und Schlick über die Objektivität der Wissenschaft", in *Moritz Schlick – Die Rostocker Jahre und ihr Einfluss auf die Wiener Zeit*, 3. Internationales Rostocker Moritz-Schlick-Symposium, November 2011, (Eds.) Engler, F.O., Iven, M., Schlickiana, 6, pp. 123–156, Leipziger Universitätsverlag, Leipzig. (Cited on pages 12 and 13.)

Fatibene, L. and Francaviglia, M., 2003, *Natural and Gauge Natural Formalism for Classical Field Theories: A Geometric Perspective including Spinors and Gauge Theories*, Kluwer Academic, Dordrecht; Norwell, MA. (Cited on pages 18, 26, and 30.)

Fatibene, L., Francaviglia, M. and Raiteri, M., 2001, "Gauge natural field theories and applications to conservation laws", in *Differential geometry and its applications*, 8th International conference on Differential Geometry and its Applications, Opava, Czech Republic, August 27–31, 2001, (Eds.) Kowalski, O., Krupka, D., Slovák, J., 3, pp. 401–413, Silesian University, Opava, Czech Republic. URL (accessed 15 May 2013): http://conferences.math.slu.cz/8icdga/proceedings.html. (Cited on page 30.)

Fine, A. and Leplin, J. (Eds.), 1989, *PSA 1988, Volume Two*, Proceedings of the 1988 Biennial Meeting of the Philosophy of Science Association, held in Evanston, IL, USA, Philosophy of Science Association, East Lansing, MI. (Cited on page 16.)

French, S. and Ladyman, J., 2003, "Remodelling Structural Realism: Quantum Physics and the Metaphysics of Structure", *Synthese*, **136**, 31–56. [DOI]. (Cited on page 38.)

Göckeler, M. and Schücker, T., 1987, *Differential geometry, gauge theories, and gravity*, Cambridge Monographs on Mathematical Physics, Cambridge University Press, Cambridge; New York. [Google Books]. (Cited on page 18.)

Goenner, H., 1996, *Einführung in die spezielle und allgemeine Relativitätstheorie*, Spektrum, Heidelberg; Oxford. (Cited on page 18.)

Greene, B., 2004, *The Fabric of the Cosmos: Space, Time, and the Texture of Reality*, Alfred A. Knopf, New York. (Cited on page 48.)

Hall, G.S., 2004, *Symmetries and Curvature Structure in General Relativity*, World Scientific Lecture Notes in Physics, 46, World Scientific, Singapore; River Edge, NJ. [Google Books]. (Cited on page 42.)

Hawking, S.W. and Ellis, G.F.R., 1973, *The Large Scale Structure of Space-Time*, Cambridge Monographs on Mathematical Physics, Cambridge University Press, Cambridge. [Google Books]. (Cited on page 18.)

Healey, R., 2001, "On the Reality of Gauge Potentials", *Philos. Sci.*, **68**, 432–455. [DOI], [PhilSci:328]. (Cited on page 30.)

Hehl, F.W., McCrea, J.D., Mielke, E.W. and Ne'Eman, Y., 1995, "Metric-affine gauge theory of gravity: field equations, Noether identities, world spinors, and breaking of dilation invariance", *Phys. Rep.*, **258**, 1–171. [DOI], [ADS], [arXiv:gr-qc/9402012]. (Cited on page 44.)

Henry, S., 2006, "Metaphysical Disputation on Haecceitism and the Principle of the Identity of Indiscernibles", *Ex Nihilo*, **6**, 19–34. Online version (accessed 7 May 2013): http://hdl.handle.net/2152/13598. (Cited on page 39.)

Hermann, R., 1973, *Geometry, Physics and Systems*, Pure and Applied Mathematics, 18, Dekker, New York. (Cited on pages 21 and 22.)

Hilbert, D., 1917, "Die Grundlagen der Physik (Zweite Mitteilung.)", *Nachr. Koenigl. Gesellsch. Wiss. Goettingen, Math.-Phys. Kl.*, **1917**, 53–76. Online version (accessed 14 January 2014): http://www.digizeitschriften.de/dms/resolveppn/?PPN=GDZPPN002504561. (Cited on page 15.)

Hoefer, C., 1996, "The metaphysics of space-time substantivalism", *J. Philos.*, **93**, 5–27. [DOI]. (Cited on page 32.)

Iftime, M. and Stachel, J., 2006, "The hole argument for covariant theories", *Gen. Relativ. Gravit.*, **38**, 1241–1252. [DOI], [ADS], [arXiv:gr-qc/0512021]. (Cited on page 11.)

Illy, J. (Ed.), 2006, *Albert Meets America: How Journalists Treated Genius During Einstein's 1921 Travels*, Johns Hopkins University Press, Baltimore. (Cited on page 17.)

Isenberg, J.A. and Marsden, J.E., 1982, "A slice theorem for the space of solutions of Einstein's equations", *Phys. Rep.*, **89**, 179–222. [DOI], [ADS]. (Cited on page 27.)

Isham, C.J., 1999, *Modern Differential Geometry for Physicists*, World Scientific Lecture Notes in Physics, 61, World Scientific, Singapore; River Edge, NJ, 2nd edn. [Google Books]. (Cited on pages 23 and 24.)

Jammer, M., 1954, *Concepts of Space: The History of Theories of Space in Physics*, Harvard University Press, Cambridge, MA. [Google Books]. (Cited on page 32.)

Janssen, M., 2007, "What Did Einstein Know and When Did He Know It? A Besso Memo Dated August 1913", in *The Genesis of General Relativity, Vol. 2: Einstein's Zurich Notebook: Commentary and Essays*, (Ed.) Renn, J., Boston Studies in the Philosophy of Science, 250, pp. 785–838, Springer, Dordrecht. [DOI]. (Cited on page 16.)

Kobayashi, S., 1972, *Transformation Groups in Differential Geometry*, Springer, Berlin; New York. [DOI], [Google Books]. (Cited on page 26.)

Komar, A., 1958, "Construction of a Complete Set of Observables in the General Theory of Relativity", *Phys. Rev.*, **111**, 1182–1187. [DOI], [ADS]. (Cited on page 14.)

Komar, A., 1973, "The General Relativistic Quantization Program", in *Contemporary Research in the Foundations and Philosophy of Quantum Theory*, Proceedings of a conference held at the University of Western Ontario, London, Canada, (Ed.) Hooker, C.A., The Western Ontario Series in Philosophy of Science, 2, pp. 305–327, D. Reidel, Dordrecht; Boston. [DOI]. (Cited on page 43.)

Kouletsis, I., Hájíček, P. and Bičák, J., 2003, "Gauge-invariant Hamiltonian dynamics of cylindrical gravitational waves", *Phys. Rev. D*, **68**, 104013. [DOI], [ADS], [arXiv:gr-qc/0308032]. (Cited on page 43.)

Kox, A.J., Klein, M.J. and Schulmann, R. (Eds.), 1996, *The Collected Papers of Albert Einstein, Vol. 6: The Berlin Years: Writings, 1914–1917*, Princeton University Press, Princeton, NJ. (Cited on page 13.)

Kretschmann, E., 1915a, "Über die prinzipielle Bestimmbarkeit der berechtigten Bezugssysteme beliebiger Relativitätstheorien (I)", *Ann. Phys. (Leipzig)*, **48**, 907–942. [DOI]. (Cited on page 14.)

Kretschmann, E., 1915b, "Über die prinzipielle Bestimmbarkeit der berechtigten Bezugssysteme beliebiger Relativitätstheorien (II)", *Ann. Phys. (Leipzig)*, **48**, 943–982. [DOI]. (Cited on page 14.)

Kretschmann, E., 1917, "Über den physikalischen Sinn der Relativitätspostulate, A. Einsteins neue und seine ursprüngliche Relativitätstheorie", *Ann. Phys. (Leipzig)*, **53**, 575–614. [DOI]. (Cited on page 14.)

Ladyman, J., 1998, "What is Structural Realism?", *Stud. Hist. Philos. Sci.*, **29**, 409–424. [DOI]. (Cited on page 38.)

Lanczos, C., 1970, *Space Through The Ages: The evolution of geometrical ideas from Pythagoras to Hilbert and Einstein*, Academic Press, London; New York. (Cited on page 16.)

Lawvere, F.W. and Schanuel, S.H., 1997, *Conceptual Mathematics: A first introduction to categories*, Cambridge University Press, Cambridge; New York. [Google Books]. (Cited on page 21.)

Lichnerowicz, A., 1955, *Théories relativistes de la gravitation et de l'électromagnétisme: Relativité générale et théories unitaires*, Masson, Paris. (Cited on page 15.)

Lichnerowicz, A., 1992, "Mathematics and General Relativity: A Recollection", in *Studies in the History of General Relativity*, Proceedings of the 2nd Conference on the History of General Relativity, Luminy, Marseille, France, 6–9 September 1988, (Eds.) Eisenstaedt, J., Kox, A.J., Einstein Studies, 3, pp. 103–108, Birkhäuser, Boston; Basel. [Google Books]. (Cited on page 15.)

Loemker, L.E. (Ed.), 1969, *Gottfried Wilhelm Leibniz: Philosophical Papers and Letters*, D. Reidel, Dordrecht. [Google Books]. (Cited on pages 39 and 54.)

Lusanna, L. and Pauri, M., 2006, "Explaining Leibniz equivalence as difference of non-inertial appearances: Dis-solution of the Hole Argument and physical individuation of point-events", *Stud. Hist. Phil. Mod. Phys.*, **37**, 692–725. [DOI], [PhilSci:2714]. (Cited on page 15.)

Lyre, H., 1999, "Gauges, Holes, and their 'Connections'", Fifth International Conference on the History and Foundations of General Relativity, July 8–11, 1999, University of Notre Dame, Notre Dame, Indiana, conference paper. [ADS], [arXiv:gr-qc/9904036]. (Cited on page 30.)

Mach, E., 1986, *Principles of the Theory of Heat: Historically and Critically Elucidated*, Vienna Circle Collection, 17, D. Reidel, Dordrecht; Boston. [DOI]. This translation from the 2nd edition, 1900. (Cited on page 5.)

Mach, E., 1988, *Die Mechanik in ihrer Entwicklung: Historisch-kritisch dargestellt*, Philosophiehistorische Texte, Akademie-Verlag, Berlin. (Cited on page 5.)

Matteucci, P., 2003, "Einstein-Dirac theory on gauge-natural bundles", *Rep. Math. Phys.*, **52**, 115–139. [DOI], [ADS], [arXiv:gr-qc/0201079]. (Cited on page 29.)

Micanek, R.J. and Hartle, J.B., 1996, "Nearly instantaneous alternatives in quantum mechanics", *Phys. Rev. A*, **54**, 3795–3800. [DOI], [ADS], [arXiv:quant-ph/9602023]. (Cited on page 49.)

Michor, P.W., 2008, *Topics in Differential Geometry*, Graduate Studies in Mathematics, 93, American Mathematical Society, Providence, RI. (Cited on page 26.)

Neumann, P.M., Stoy, G.A. and Thompson, E.C., 1994, *Groups and Geometry*, Oxford University Press, Oxford; New York. [Google Books]. (Cited on pages 21, 22, and 23.)

Nicolai, H. and Peeters, K., 2007, "Loop and Spin Foam Quantum Gravity: A Brief Guide for Beginners", in *Approaches to Fundamental Physics: An Assessment of Current Theoretical Ideas*, (Eds.) Stamatescu, I.-O., Seiler, E., Lecture Notes in Physics, 721, pp. 151–184, Springer, Berlin; New York. [DOI], [ADS], [hep-th/0601129]. (Cited on pages 48 and 49.)

Nijenhuis, A., 1994, "Book Review: 'Natural operations in differential geometry', by Ivan Kolář, Peter W. Michor, and Jan Slovák, Springer-Verlag, Berlin et al., 1993", *Bull. Am. Math. Soc.*, **31**, 108–112. [DOI]. (Cited on pages 20 and 51.)

Norton, J.D., 1984, "How Einstein found his field equations: 1912–1915", *Hist. Stud. Phys. Sci.*, **14**, 253–316. [DOI]. (Cited on page 16.)

Norton, J.D., 1993, "General covariance and the foundations of general relativity: eight decades of dispute", *Rep. Prog. Phys.*, **56**, 791–858. [DOI]. (Cited on page 5.)

Norton, J.D., 2011, "The Hole Argument", in *The Stanford Encyclopedia of Philosophy*, (Ed.) Zalta, E.N., Stanford University, Stanford, CA. URL (accessed 16 November 2013): http://plato.stanford.edu/archives/fall2011/entries/spacetime-holearg/. (Cited on pages 33 and 39.)

Oeckl, R., 2008, "General boundary quantum field theory: Foundations and probability interpretation", *Adv. Theor. Math. Phys.*, **12**, 319–352. [ADS], [arXiv:hep-th/0509122]. (Cited on page 50.)

Oeckl, R., 2013, "A Positive Formalism for Quantum Theory in the General Boundary Formulation", *Found. Phys.*, **43**, 1206–1232. [DOI], [ADS], [arXiv:1212.5571 [quant-ph]]. (Cited on page 50.)

Olver, P.J., 1995, *Equivalence, Invariants, and Symmetry*, Cambridge University Press, Cambridge; New York. [Google Books]. (Cited on page 26.)

Pais, A., 1982, *'Subtle is the Lord ...': The Science and the Life of Albert Einstein*, Oxford University Press, Oxford; New York. [Google Books]. (Cited on page 16.)

Penrose, R., 1963, "Asymptotic properties of fields and space-times", *Phys. Rev. Lett.*, **10**, 66–68. [DOI], [ADS]. (Cited on page 43.)

Pooley, O., 2000, "Spacetime Realism and Quantum Gravity", conference paper. Online version (accessed 15 May 2013): http://users.ox.ac.uk/~ball0402/research/. (Cited on pages 33 and 37.)

Pooley, O., 2006, "Points, Particles, and Structural Realism", in *The Structural Foundations of Quantum Gravity*, (Eds.) Rickles, D., French, S., Saatsi, J., pp. 83–120, Clarendon Press, Oxford; New York. [PhilSci:2939].   (Cited on pages 37 and 38.)

Pooley, O., 2013, "Substantivalist and Relationalist Approaches to Spacetime", in *The Oxford Handbook of Philosophy of Physics*, (Ed.) Batterman, R., Oxford Handbooks in Philosophy, Oxford University Press, Oxford; New York. [PhilSci:9055].   (Cited on page 37.)

Prugovečki, E., 1992, *Quantum Geometry: A Framework for Quantum General Relativity*, Fundamental Theories of Physics, 48, Kluwer Academic, Dordrecht; Boston. [DOI].   (Cited on page 18.)

Reisenberger, M.P. and Rovelli, C., 2002, "Spacetime states and covariant quantum theory", *Phys. Rev. D*, **65**, 125016. [DOI], [arXiv:gr-qc/0111016].   (Cited on pages 48 and 49.)

Renn, J. (Ed.), 2007, *The Genesis of General Relativity, Vol. 1: Einsteins's Zurich Notebook: Introduction and Source*, Boston Studies in the Philosophy of Science, 250, Springer, Dordrecht. [DOI].   (Cited on pages 10 and 16.)

Renn, J. and Stachel, J., 2007, "Hilbert's Foundations of Physics: From a Theory of Everything to a Constituent of General Relativity", in *The Genesis of General Relativity, Vol. 4: Gravitation in the Twilight of Classical Physics: The Promise of Mathematics*, (Ed.) Renn, J., Boston Studies in the Philosophy of Science, 250, pp. 1778–1895, Springer, Dordrecht. [DOI].   (Cited on page 15.)

Rickles, D.P., 2005, "A new spin on the hole argument", *Stud. Hist. Phil. Mod. Phys.*, **36**, 415–434. [DOI], [PhilSci:1859].   (Cited on page 15.)

Rorty, R., 1967, "Relations, Internal and External", in *The Encylopedia of Philosophy, vol. 7*, (Ed.) Edwards, P., pp. 125–133, Macmillan, New York.   (Cited on page 54.)

Rovelli, C., 1991, "What is observable in classical and quantum gravity?", *Class. Quantum Grav.*, **8**, 297–316. [DOI].   (Cited on page 16.)

Rovelli, C., 2004, *Quantum Gravity*, Cambridge Monographs on Mathematical Physics, Cambridge University Press, Cambridge; New York. [Google Books].   (Cited on pages 41 and 50.)

Rynasiewicz, R., 1999, "Kretschmann's Analysis of Covariance and Relativity Principles", in *The Expanding Worlds of General Relativity*, Proceedings of the 4th Conference on the History of General Relativity, Berlin 1995, (Eds.) Goenner, H., Renn, J., Ritter, J., Sauer, T., Einstein Studies, 7, pp. 431–462, Birkhäuser, Boston; Basel. [Google Books].   (Cited on pages 12 and 14.)

Sánchez-Rodríguez, I., 2008, "Geometrical Structures of Space-Time in General Relativity", in *Geometry and Physics*, XVI International Fall Workshop, Lisbon, Portugal, 5 – 8 September 2007, (Eds.) Fernandes, R.L., Picken, R., AIP Conference Proceedings, 1023, pp. 202–206, American Institute of Physics, Melville, NY. [DOI], [ADS], [arXiv:0803.1929 [gr-qc]].   (Cited on page 24.)

Schichl, H., 1997, *On the existence of slice theorems for moduli spaces on fiber bundles*, Ph.D. thesis, Universität Wien, Vienna. Online version (accessed 10 January 2014): http://ubdata.univie.ac.at/AC01930824.   (Cited on page 27.)

Schouten, J.A., 1951, *Tensor Analysis for Physicists*, Clarendon Press, Oxford.   (Cited on page 38.)

Schouten, J.A., 1954, *Ricci-Calculus: An Introduction to Tensor Analysis and Its Geometrical Applications*, Die Grundlehren der Mathematischen Wissenschaften, X, Springer, Berlin; Heidelberg, 2nd edn. [DOI].   (Cited on page 23.)

Schulmann, R., Kox, A.J. and Janssen, M. (Eds.), 1998, *The Collected Papers of Albert Einstein, Vol. 8 Bd. 2, The Berlin Years: Correspondence, 1914-1917*, Princeton University Press, Princeton, NJ. (Cited on page 12.)

Sharpe, R.W., 1997, *Differential Geometry: Cartan's Generalization of Klein's Erlangen Program*, Graduate Texts in Mathematics, 166, Springer, New York; Berlin. [Google Books]. (Cited on pages 26 and 41.)

Sklar, L., 1974, *Space, Time, and Spacetime*, University of California Press, Berkeley. [Google Books]. (Cited on page 17.)

Stachel, J., 1966, "Cylindrical Gravitational News", *J. Math. Phys.*, **7**, 1321–1331. [DOI]. (Cited on page 43.)

Stachel, J., 1969, "Covariant Formulation of the Cauchy Problem in Generalized Electrodynamics and General Relativity", *Acta Phys. Pol.*, **35**, 689–709. (Cited on page 41.)

Stachel, J., 1979, "The genesis of general relativity", in *Einstein Symposion Berlin*, Aus Anlaß der 100. Wiederkehr seines Geburtstages 25. bis 30. März 1979, (Eds.) Nelkowski, H., Hermann, A., Poser, H., Schrader, R., Seiler, R., Lecture Notes in Physics, 100, pp. 428–442, Springer, Berlin; New York. [DOI], [ADS]. Reprinted in Stachel, J., Einstein from 'B' to 'Z', Einstein Studies, vol. 9, pp. 233–244, Birkhäuser, Boston; Basel. (Cited on page 16.)

Stachel, J., 1980, "The anholonomic Cauchy problem in general relativity", *J. Math. Phys.*, **21**, 1776–1782. [DOI]. (Cited on page 41.)

Stachel, J., 1986, "What a Physicist Can Learn From the Discovery of General Relativity", in *The Fourth Marcel Grossmann Meeting on recent developments in theoretical and experimental general relativity, gravitation and relativistic field theories*, Proceedings of the meeting held at the University of Rome 'La Sapienza', 17 – 21 June, 1985, (Ed.) Ruffini, R., pp. 1857–1862, North-Holland; Elsevier, Amsterdam; New York. (Cited on pages 16 and 28.)

Stachel, J., 1987, "How Einstein discovered general relativity: A historical tale with some contemporary morals", in *General Relativity and Gravitation*, Proceedings of the 11th International Conference on General Relativity and Gravitation, Stockholm, July 6 – 12, 1986, (Ed.) MacCallum, M.A.H., pp. 200–208, Cambridge University Press, Cambridge; New York. (Cited on page 16.)

Stachel, J., 1989, "Einstein's Search for General Covariance 1912–1915", in *Einstein and the History of General Relativity*, Based on the proceedings of the 1986 Osgood Hill Conference, North Andover, Massachusetts, 8 – 11 May, (Eds.) Howard, D., Stachel, J., Einstein Studies, 1, pp. 63–100, Birkhäuser, Boston; Basel. [ADS]. (Cited on pages 12, 14, and 16.)

Stachel, J., 1992, "The Early History of the Cauchy Problem in General Relativity, 1916-1937", in *Studies in the History of General Relativity*, Proceedings of the 2nd Conference on the History of General Relativity, Luminy, Marseille, France, 6 – 9 September 1988, (Eds.) Eisenstaedt, J., Kox, A.J., Einstein Studies, 3, pp. 407–418, Birkhäuser, Boston; Basel. (Cited on page 15.)

Stachel, J., 1993, "The Meaning of General Covariance: The Hole Story", in *Philosophical Problems of the Internal and External Worlds: Essays on the Philosophy of Adolf Grünbaum*, (Eds.) Earman, J., Janis, A.I., Massey, G.J., Rescher, N., Pittsburgh-Konstanz Series in the Philosophy and History of Science, pp. 129–160, University of Pittsburgh Press / Universitätsverlag Konstanz, Pittsburgh; Konstanz. [Google Books]. Online version (accessed 25 Novemver 2013):

http://digital.library.pitt.edu/cgi-bin/t/text/text-idx?idno=31735062135235;
view=toc;c=pittpress. (Cited on pages 14, 16, 36, and 37.)

Stachel, J., 1997, "Feynman Paths and Quantum Entanglement: Is There Any More to the Mystery?", in *Potentiality, Entanglement and Passion-at-a-Distance: Quantum Mechanical Studies for Abner Shimony, Vol. 2*, (Eds.) Cohen, R.S., Horne, M., Stachel, J., Boston Studies in the Philosophy of Science, 194, pp. 245–256, Kluwer Academic, Dordrecht; Boston. [DOI]. (Cited on page 46.)

Stachel, J., 2002, "'The Relations between Things' versus 'The Things Between Relations': The Deeper Meaning of the Hole Argument", in *Reading Natural Philosophy: Essays in the History and Philosophy of Science and Mathematics*, (Ed.) Malament, D., pp. 231–266, Open Court, Chicago; LaSalle. [Google Books]. (Cited on pages 38 and 54.)

Stachel, J., 2003, "'Critical Realism': Wartofsky and Bhaskar", in *Constructivism and Practice: Towards a Historical Epistemology*, (Ed.) Gould, C.C., pp. 137–150, Rowman & Littlefield, Lanham, MD; Oxford. (Cited on page 52.)

Stachel, J., 2005, "Structural Realism and Contextual Individuality", in *Hilary Putnam*, (Ed.) Ben-Menahem, Y., Contemporary Philosophy in Focus, pp. 203–219, Cambridge University Press, Cambridge; New York. [Google Books]. (Cited on page 33.)

Stachel, J., 2006a, "Structure, Individuality, and Quantum Gravity", in *Structural Foundations of Quantum Gravity*, (Eds.) Rickles, D., French, S., Saatsi, J., pp. 53–82, Oxford University Press, Oxford; New York. (Cited on pages 33, 38, 42, 46, 47, and 48.)

Stachel, J., 2006b, "Albert Einstein: A Man for the Millenium?", in *A Century of Relativity Physics: XXVIII Spanish Relativity Meeting (ERE 2005)*, Oviedo, Asturias, Spain, 6 – 10 September 2005, (Eds.) Mornas, L., Diaz Alonso, J., AIP Conference Proceedings, 841, pp. 195–227, American Institute of Physics, Melville, NY. [DOI]. (Cited on pages 17 and 48.)

Stachel, J., 2007, "The First Two Acts", in *The Genesis of General Relativity, Vol. 2: Einstein's Zurich Notebook: Commentary and Essays*, (Ed.) Renn, J., Boston Studies in the Philosophy of Science, 250, pp. 81–111, Springer, Dordrecht. [DOI]. (Cited on page 8.)

Stachel, J., 2009, "Prolegomena to any future Quantum Gravity", in *Approaches to Quantum Gravity: Toward a New Understanding of Space, Time and Matter*, (Ed.) Oriti, D., pp. 44–67, Cambridge University Press, Cambridge; New York. (Cited on pages 31 and 42.)

Stachel, J., 2011, "Conformal and projective structures in general relativity", *Gen. Relativ. Gravit.*, **43**, 3399–3409. [DOI], [ADS]. (Cited on pages 25, 31, and 45.)

Stachel, J., 2014, "General Relativity and Differential Geometry: The Einstein Connection", in *Beyond Einstein*, Based upon the conference, Johannes Gutenberg University, Mainz Germany, 22 – 26 September 2008, (Ed.) Rowe, D., Einstein Studies, Birkhäuser, Boston; Basel. Forthcoming. (Cited on page 9.)

Stachel, J. and Bradonjić, K., 2013, "Quantum Gravity: Meaning and Measurement", *arXiv*, e-print. [ADS], [arXiv:1302.2285 [gr-qc]]. (Cited on page 49.)

Stachel, J. and Iftime, M., 2005, "Fibered Manifolds, Natural Bundles, Structured Sets, G-Sets and all that: The Hole Story from Space Time to Elementary Particles", *arXiv*, e-print. [ADS], [arXiv:gr-qc/0505138]. (Cited on pages 19, 27, 36, and 37.)

Stephani, H., 2004, *General Relativity: An introduction to the theory of the gravitational field*, Cambridge University Press, Cambridge; New York, 3rd edn. [Google Books]. (Cited on page 18.)

Stephani, H., Kramer, D., MacCallum, M.A.H., Hoenselaers, C. and Herlt, E., 2003, *Exact Solutions to Einstein's Field Equations*, Cambridge Monographs on Mathematical Physics, Cambridge University Press, Cambridge; New York, 2nd edn. [Google Books]. (Cited on page 42.)

Teller, P., 1998, "Quantum Mechanics and Haecceities", in *Interpreting Bodies: Classical and Quantum Objects in Modern Physics*, (Ed.) Castellani, E., pp. 114–141, Princeton University Press, Princeton, NJ. [Google Books]. (Cited on page 38.)

Thiemann, T., 2001, "Introduction to Modern Canonical Quantum General Relativity", *arXiv*, e-print. [arXiv:gr-qc/0110034]. (Cited on page 47.)

Torre, C.G., 1999, "Midisuperspace Models of Canonical Quantum Gravity", *Int. J. Theor. Phys.*, **38**, 1081–1102. [DOI]. (Cited on page 43.)

Torretti, R., 1983, *Relativity and Geometry*, Foundations and philosophy of science and technology, Pergamon Press, Oxford; New York. [Google Books]. (Cited on page 16.)

Trautman, A., 1970, "Fibre bundles associated with space-time", *Rep. Math. Phys.*, **1**, 29–62. [DOI]. (Cited on page 18.)

Trautman, A., 1980, "Fiber Bundles, Gauge Fields, and Gravitation", in *General Relativity and Gravitation: One Hundred Years After the Birth of Albert Einstein, Vol. 1*, (Ed.) Held, A., chap. 9, pp. 287–308, Plenum Press, New York; London. (Cited on pages 18 and 44.)

Varadarajan, V.S., 2003, "Vector Bundles and Connections in Physics and Mathematics: Some Historical Remarks", in *A Tribute to C.S. Seshadri: A Collection of Articles on Geometry and Representation Theory*, Symposium to felicitate C.S. Seshadri's 70th birthday, held in Chennai, India, March 1, 2002, (Eds.) Lakshmibai, V., Balaji, V., Mehta, V.B., Nagarajan, K.R., Paranjape, K., Sankaran, P., Sridharan, R., Trends in Mathematics, pp. 502–541, Birkhäuser, Basel; Boston. (Cited on page 18.)

Wald, R.M., 1984, *General Relativity*, University of Chicago Press, Chicago. [ADS], [Google Books]. (Cited on page 18.)

Wartofsky, M.W., 1968, *Conceptual Foundations of Scientific Thought: An Introduction to the Philosophy of Science*, Macmillan, New York. (Cited on page 52.)

Weyl, H., 1923, *Raum, Zeit, Materie: Vorlesungen über allgemeine Relativitätstheorie*, Julius Springer, Berlin, 5th edn. [DOI]. (Cited on page 8.)

Weyl, H., 1946, *The Classical Groups: Their Invariants and Representations*, Princeton Landmarks in Mathematics, Princeton University Press, Princeton, NJ, 2nd edn. [Google Books]. (Cited on page 20.)

Weyl, H., 1949, *Philosophy of Mathematics and Natural Science*, Princeton University Press, Princeton, NJ. [Google Books]. (Cited on pages 53 and 54.)

Weyl, H., 1950, *Space, Time, Matter*, Dover Publications, New York. [Google Books]. (Cited on page 8.)

Wikipedia contributors, "Group action", online resource, Wikipedia Foundation. URL (accessed 13 December 2013):
http://en.wikipedia.org/wiki/Group_action. (Cited on page 24.)

Yaglom, I.M., 1979, *Simple Non-Euclidean Geometry and Its Physical Basis: An Elementary Account of Galilean Geometry and the Galilean Principle of Relativity*, Heidelberg Science Library, Springer, New York; Berlin. [DOI].   (Cited on page 23.)