International Journal of Theoretical Physics, Vol. 32, No. 11, 1993

An Axiomatization of General Relativity

G. M. Covarrubias[1]

Received September 17, 1992

An axiomatization of the general theory of relativity is proposed. The assumed philosophical background is critical realism. None of the "principles" commonly considered as founding the theory, such as (a) the equality of inertial and gravitational mass, (b) the principle of equivalence, (c) the principle of general covariance, (d) the geodesic postulate, and (e) Mach's principle, are taken as axioms in our system.

1. INTRODUCTION

I shall not apologize for proposing an axiomatization of such a relevant and profound physical theory as the general theory of relativity (GR). It is my belief that the axiomatization of a theory, whether physical or not, is an epistemological need, if only because it brings order and systemicity to the former. Of course, the theory must have attained a certain degree of maturity so that one may distinguish its basic postulates from the heuristic principles, analogies, etc., which are always present in its process of birth.

The latter remarks are not intended to provide an account of physical axiomatics. This has been dealt with in detail by many philosophers of science, notably by Bunge (1967, 1977, 1979). I shall structure this axiomatization according to the outlines of physical axiomatics proposed by this author. The philosophical background intended here is critical realism. Two of its theses are central to this work: the objective existence of nature and the nonobjective existence of concepts. The relation between these two levels of abstraction is dealt with in Bunge's semantics (Bunge, 1974*a*,*b*).

[1]Departamento de Física, Universidad Autónoma Metropolitana-Iztapalapa, Mexico D.F. 09340, Mexico.

0020-7748/93/1100-2135$07.00/0

Also, I shall freely use his terminology. Readers not familiar with the latter are referred to Bunge's works listed in the bibliography.

To my knowledge, the first explicit axiomatization of GR was given by Bunge (1967) (together with the axiomatization of four more theories). The present work leans so heavily on it that perhaps it should be called a reaxiomatization. The main new features are: (a) The axioms are formulated by means of the intrinsic geometry of the manifold referring to spacetime in GR in its modern spirit; (b) I do not restrict the gravitational fields to those associated with isolated physical systems, thereby allowing for cosmological models (actually, Bunge makes a serious mistake in restricting himself to finite regions of spacetime. In this way it would be hard, if possible at all, to appreciate the features of the gravitational field. An isolated system is finite three-dimensionally, but not in four dimensions); (c) the number of axioms is reduced from 16 to 10.

After a brief note on the heuristic components of the theory in Section 2, I provide the primitive base in Section 3. Section 4 contains a list of interesting defined concepts and in Section 5 I offer the body of the axioms or axiom base. In Section 6 a set of comments to the axioms are made which should help to understand the whole scheme. Finally, Section 7 contains a couple of very interesting theorems: the equivalence principle and the geodesic postulate. In order to keep this paper within a reasonable extension, I restrain from formulating and proving any more theorems which would enhance my belief that the system is both p- and d-complete. An Appendix is included at the end to specify the precise mathematical ideas used here, as they may vary in detail from author to author.

2. HEURISTIC COMPONENTS

Most of the treatises on GR lead one to believe that the theory is based on a number of "principles." The most common are: (a) the equality of inertial and gravitational mass, (b) the principle of equivalence, (c) the principle of general covariance, early in the development of the theory, (d) the geodesic postulate, and the highly ambiguous (e) Mach's principle.

Bunge (1967) shows that (a) cannot even be formulated in the language of GR; that (c) is a metanomological statement: that is, it refers to law statements and therefore is not to be placed on the same status as the latter; while, as we shall show, (b) and, as is widely recognized nowadays, (d) are theorems of GR.

All these principles, except perhaps (d), played a part in the construction of the theory. They were heuristic clues along with many others, but they do not belong in the foundations of the finished theory. Therefore none of them will be postulates of our axiom system.

I shall not pursue this matter further here. For a formulation and discussion of (a), (b), and (c) the reader is referred to Bunge (1967).

3. BASIC CONCEPTS

3-i) $\{(M^n, \mathbf{g}, \nabla)\}$ is a set of pseudo-Riemannian manifolds, with manifold M^n, where $n \in N$, metric tensor $\mathbf{g}$, and metric connection ∇ defined on the manifold structure.

3-i-a) M^n designates spacetime.

3-i-b) Some elements $x \in M^n$ designate events in spacetime.

3-i-c) The 2-covariant symmetric tensor field (metric tensor) $\mathbf{g}$ designates the gravitational potential.

3-i-d) The connection coefficients $\gamma^{\mu}_{\alpha\beta}$ in the dual moving frame $(\mathbf{e}_\mu), (\boldsymbol{\theta}^\alpha)$ designate the components of the gravitational field in such a frame.

3-ii) Σ is the collection of all (possible and actual) gravitational fields.

3-iii) $\{\mathbf{L}\}$ is a family of 2-covariant tensor field functionals of $\mathbf{g}$ and possibly of state variables. An element $\mathbf{L} \in \{\mathbf{L}\}$ is called an energy-momentum tensor.

3-iv) $\bar{\Sigma}$ is the collection of macroscopic physical systems other than gravitational fields.

3-v) K is the collection of reference frames.

3-vi) κ is a negative dimensional constant.

4. DEFINED CONCEPTS

4-i) If $\{(U_\alpha, \varphi_\alpha)\}_{\alpha \in I}$ is a C^k atlas in M^n, any other equivalent C^k atlas $\{(V_\beta, \psi_\beta)\}_{\beta \in J}$ is a global change of coordinates class C^k. If $x \in M^n$ and (U, φ) is a chart in the neighborhood U of x, i.e., $x \in U$, any other chart (V, ψ) in the neighborhood V of x defines a change of coordinates in $U \cap V$, a neighborhood of x. The C^k map $\psi \circ \varphi^{-1}$: $\varphi(U \cap V) \to R^n$ gives the "new" coordinates in terms of the "old" ones.

4-ii) Any mapping (arbitrary range) f defined on the manifold M^n, i.e., f: $M^n \to F$, represents an invariant. If $x \in M^n$ and (U, φ) is a chart in the neighborhood U of x, $f \circ \varphi^{-1}$ is "the invariant expressed in terms of the coordinates in that neighborhood."

4-iii) Any formula (such as an equation) or statement formulated in terms of mathematical concepts derived from the intrinsic geometry of $(M^n, \mathbf{g}, \nabla)$ (that is, which does not depend on the coordinates) is called "covariant." When an atlas is considered and the concepts above are realized using the coordinate system thus given, one says that the formula is expressed in terms of the coordinates. Different coordinate systems give different expressions, but one says that the statement is "satisfied when subject to arbitrary coordinate transformations."

4-iv) The curvature operation is defined by

$$\boldsymbol{\rho}(\mathbf{u})\mathbf{v} = \nabla_{\mathbf{u}}\nabla_{\mathbf{v}} + \nabla_{\mathbf{v}}\nabla_{\mathbf{u}} - \nabla_{[\mathbf{u},\mathbf{v}]}$$

where $\nabla_{\mathbf{u}}$ is the covariant derivative in the direction of $\mathbf{u}$, $[\mathbf{u}, \mathbf{v}]$ is the Lie bracket $[\mathbf{u}, \mathbf{v}] = \mathbf{uv} - \mathbf{vu}$, and $\mathbf{u}$ and $\mathbf{v}$ are vector fields over M^n.

4-v) The curvature or Riemann tensor is

$$\mathbf{R}(\mathbf{w}, \boldsymbol{\alpha}, \mathbf{u}, \mathbf{v}) = \boldsymbol{\alpha}(\boldsymbol{\rho}(\mathbf{u}, \mathbf{v}) \cdot \mathbf{w})$$

where $\boldsymbol{\alpha}$ is a 1-form field, $\boldsymbol{\alpha} \in \Lambda^1(M^n)$. In the dual moving frame $(\mathbf{e}_\alpha)$, $(\boldsymbol{\theta}^\beta)$, the components of $\mathbf{R}$ are

$$\begin{aligned} R_\alpha{}^\beta{}_{\gamma\delta} &= \mathbf{R}(\mathbf{e}_\alpha, \boldsymbol{\theta}^\beta, \mathbf{e}_\gamma, \mathbf{e}_\delta) \\ &= (\mathbf{e}_\gamma)(\gamma^\beta_{\delta\alpha}) - \mathbf{e}_\delta(\gamma^\beta_{\gamma\alpha}) + \gamma^\beta_{\gamma\mu}\gamma^\mu_{\delta\alpha} - \gamma^\beta_{\delta\mu}\gamma^\mu_{\gamma\alpha} - C^\mu_{\gamma\delta}\gamma^\beta_{\mu\alpha}, \end{aligned}$$

where $\gamma^\mu_{\alpha\beta}$ are the connection coefficients defined by

$$\nabla \mathbf{e}_\beta = \gamma^\mu_{\alpha\beta}\boldsymbol{\theta}^\alpha \otimes \mathbf{e}_\mu$$

and the $C^\mu_{\alpha\beta}$ are given by

$$[\mathbf{e}_\alpha, \mathbf{e}_\beta] = C^\mu_{\alpha\beta}\mathbf{e}_\mu$$

and are called "the structure coefficients of the moving frame." In the coordinate frame $(\partial/\partial x^\mu)$, (dx^ν) in U, defined naturally by the chart (U, φ), $\varphi(x) = x^\mu$, one has

$$R_\alpha{}^\beta{}_{\gamma\delta} = \partial_\beta \Gamma^\beta_{\delta\alpha} - \partial_\delta \Gamma^\beta_{\gamma\alpha} + \Gamma^\beta_{\gamma\mu}\Gamma^\mu_{\delta\alpha} - \Gamma^\beta_{\delta\mu}\Gamma^\mu_{\gamma\alpha}$$

Here ∂_α denotes $\partial/\partial x^\alpha$ and the connection coefficients are written as $\Gamma^\mu_{\alpha\beta}$. *Note*: The Einstein summation convention is being used, where summation is understood when in an expression an index occurs twice, once as superscript and once as subscript.

The following three definitions apply to a Riemannian connection (see the Appendix).

4-vi) The Ricci tensor is

$$R_{\alpha\beta} = R_\alpha{}^\gamma{}_{\beta\gamma}$$

4-vii) The curvature scalar is

$$R = R^\alpha{}_\alpha = g^{\alpha\beta}R_{\alpha\beta}$$

where

$$g^{\alpha\beta} = \frac{1}{\det(g_{\alpha\beta})}\frac{\partial \det(g_{\alpha\beta})}{\partial g_{\alpha\beta}}$$

is the inverse matrix of the matrix formed with the components of the metric tensor. It defines a 2-contravariant tensor field $\mathbf{g}^*$:

$$\mathbf{g}^* = g^{\alpha\beta}\mathbf{e}_\alpha \otimes \mathbf{e}_\beta$$

4-viii) The Einstein tensor is

$$G_{\alpha\beta} = R_{\alpha\beta} - \tfrac{1}{2}g_{\alpha\beta}R$$

4-ix) A reference frame in a set $U \subset M^n$ is a moving frame defined on it (tetrad, for $n = 4$).

4-x) Let $\mathbf{u} \in T_x$ (the tangent space of M^n at x). The signature used here is $+---$.

4-x-a) If $\mathbf{g}_x(\mathbf{u}, \mathbf{u}) > 0$, $\mathbf{u}$ is called "timelike."

4-x-b) If $\mathbf{g}_x(\mathbf{u}, \mathbf{u}) < 0$, $\mathbf{u}$ is called "spacelike."

4-x-c) If $\mathbf{g}_x(\mathbf{u}, \mathbf{u}) = 0$, $\mathbf{u}$ is called "null."

Where $\mathbf{g}_x$ is the tensor $\mathbf{g}_x \in \otimes^2 T_x^*$ associated by the tensor field $\mathbf{g}$ to the point x. T_x^* denotes the space dual to T_x.

5. AXIOMS

Axiom Group I: The Referents

5-i-a) Σ is a nonempty collection.

5-i-b) Every $\sigma \in \Sigma$ designates a gravitational field.

5-ii-a) $\bar{\Sigma}$ is nonempty, and $\Sigma \cap \bar{\Sigma} = \varnothing$ ($\varnothing$ = empty set). There is an element in $\bar{\Sigma}$, $\Box$, which represents the absence of a physical system.

5-ii-b) Every $\bar{\sigma} \in \bar{\Sigma}$ other than $\Box$ designates a macroscopic physical system other than a gravitational field.

Axiom Group II: Spacetime

5-iii) The dimension n of spacetime is 4. The pseudo-Riemannian manifold $(M^4, \mathbf{g})$ is thus called a "hyperbolic manifold."

5-iv) M^4 is a differentiable manifold of at least class C^4. (We want to have a metric tensor field $\mathbf{g}$ class C^3 [except possibly for a set of zero (Lebesgue) measure]. For that, M^4 must be at least class C^4.)

5-v) For every $x \in M^4$, there exists a chart (U, φ) in the equivalent class of C^4 atlases which define the differential structure of $(M^4, \mathbf{g})$, such that one of the coordinates, say x^0, is timelike; that is, $\partial/\partial x^0$ is a timelike vector at each point of U, and the remaining three coordinates are spacelike. This is equivalent to the following condition: in these coordinates, $g_{00} > 0$, and the three-dimensional quadratic form $g_{ij}v^iv^j$ is negative definite. If $g_{00} > 0$, then $(1/u)\,\partial/\partial x^0$, with $u^2 = g_{00}$, is timelike, and the second part of the condition is: the three-dimensional quadratic form defined by $(g_{ij} - g_{0i}g_{0j}/g_{00})\,dx^i\,dx^j$, $i = 1, 2, 3$, is negative definite.

Axiom Group III: The Metric

5-vi) $\mathbf{g}$ is a 2-covariant symmetric tensor field over M^4 of class C^3, except possibly in a set of (Lebesgue) measure zero.

Axiom Group IV: Frame and Physical Coordinates (Pointless for L = 0)

5-vii-a) $K \neq \varnothing$ and $K \subset \bar{\Sigma}$.

5-vii-b) Every $k \in K$ is a physical reference frame. For every such frame, there exists a point $x \in M^4$ and a coordinate neighborhood of x, (U, φ), over which a moving frame $(\mathbf{e}_\alpha)$, $\alpha = 0, 1, 2, 3$, can be defined and such that $(\mathbf{e}_\alpha) \triangleq K$.

Axiom Group V: The Energy-Momentum Tensor

5-viii-a) Each $\mathbf{L} \in \{\mathbf{L}\}$ is a symmetric 2-covariant tensor field, functional of $\mathbf{g}$ and possibly of state variables, over M^4, of class C^1 up to a set of (Lebesgue) measure zero.

5-viii-b) For every $\bar{\sigma} \in \bar{\Sigma}$ there exists an $\mathbf{L} \in \{\mathbf{L}\}$, which represents the energy, momentum, and stresses of the system. For $\bar{\sigma} = \square$ (the void), $\mathbf{L} = 0$.

Axiom Group VI: The Field Equations

5-ix) For every $\bar{\sigma} \in \bar{\Sigma}$ there exists a hyperbolic manifold $(M^4, \mathbf{g})$ and its corresponding connection ∇, defined over M^4 such that

$$\mathbf{G} = \kappa \mathbf{L}$$

where $\mathbf{G}$ is the Einstein tensor and $\mathbf{L}$ is the energy-momentum tensor functional corresponding to $\bar{\sigma}$ in accord with the previous axiom.

Axiom Group VII: The Coupling Constant

5-x) κ is a negative constant which represents the coupling of the gravitational field with nongravitational systems. It can be written in terms of the Newtonian gravitational constant G and the speed of light in vacuum c as

$$\kappa = -8\pi \frac{G}{c^4}$$

The dimensions of κ are $L^{-1}M^{-1}T^{-2}$ (L = length, M = mass, T = time).

6. COMMENTS

Axiom Group I

In GR, physical systems are partitioned into gravitational and nongravitational. Every macroscopic nongravitational physical system gives rise to a gravitational field. I should point out, however, that there are also gravitational fields in the absence of any physical systems. These are included in Σ.

These two axioms are semantical. They assert that there are physical systems in nature. Neither the elements of Σ nor of $\bar{\Sigma}$ are concepts nor are the gravitational or nongravitational fields themselves. They are mere symbols related to the physical systems (except for $\Box$, which is related to the void, a concept) by the semantic relation of designation (Bunge, 1974*a*).

Σ and $\bar{\Sigma}$ are collections, not sets. By a collection I mean a set of variable membership. Physical systems do not exist as such in nature. They are constituted by physical entities but defined by the physicist in the light of the physical theories that deal with them.

Physical systems are restricted to macroscopic ones because this has been so since GR was proposed by Einstein. It is true that there have been serious efforts to build a quantum field theory of gravitation, but, at present, the theory has not been properly born and it would therefore be out of place to consider microscopic systems (an electron, say) without knowing how to deal with them.

Axiom Group II

That the dimension of spacetime is four must seem today a natural fact. Since Einstein's proposal of special relativity in 1905 in which time loses its absolute character, one has become used to considering it, along with space, as relative to the inertial reference frame one cares to single out, all inertial reference frames being equivalent to each other. Perhaps one can trace the origin of spacetime as a single category to the suggestion of Minkowski (1908) of considering the four coordinates (t, x, y, z) as those of a four-dimensional space.

But this is by no means a trivial fact. As early as 1919 (four years after Einstein's publication of his general theory) F. E. Kaluza contacted Einstein proposing a five-dimensional spacetime theory. At that time only two of the four fundamental forces that we recognize nowadays were known: gravitation, described by GR, and electromagnetism, described by Maxwell's theory. Kaluza's idea was to give unified account of the two forces.

Today much effort is being devoted to building a unified field theory of the four forces. Three can be said to have been unified (the rebel is precisely the gravitational force) in a scheme called grand unification, although there are still problems (the theory predicts the decay of the proton and it has never been observed). And all theoretical formulations are made in more than four dimensions. Three main classes of modern theories can be mentioned: the so-called Kaluza–Klein theories, in which one of the main problems is to decide how many dimensions the spacetime must have, but it is certain that the number should be greater than five; supergravity, in which it is clear that the number of dimensions must be 11; and the modern string theories, in which mathematical consistency demands 10 dimensions. It is likely that if a unified field theory will be built (and it seems most certain that it will) the spacetime in which we live will be found to be more than 4-dimensional.

Axiom 5-iv) is about the mathematical structure of spacetime. Differential geometry is a field in which different authors differ in matters of detail. I give in the Appendix the precise definitions that are intended here. A manifold is locally Euclidean. This concept is not directly related to curvature. The idea is that every point of it belongs to a neighborhood which can be coordinated; that is, every point of the neighborhood can be assigned a four-tuple (we are assuming $n = 4$) of real numbers and the coordinates of different points in the neighborhood are related in a continuous way. This generalizes the concept of "parametric representation" of a surface of R^4. These neighborhoods are called charts and the "surface" is then covered with them. The concept of a differentiable manifold generalizes the concept of a differentiable surface of R^4; that is, a surface which admits a tangent plane at each of its points.

At every point of the differentiable manifold a huge algebraic structure is built using the local differentiable properties of the manifold. Thus, to define the algebraic entities (vectors, tensors, etc.) at least one degree of differentiability is used. These in turn can be differentiable if the manifold admits one degree of differentiability in addition to those used to define them.

Now, the Einstein equations (axiom 5-ix) involve the Einstein tensor, whose definition requries two differentiations, and moreover, they obey the Bianchi identities, which involve one more differentiation. Consequently the manifold which is to model a physical situation must be at least class C^4.

Nevertheless one cannot assume such a nice smooth manifold. According to the work of Hawking and Ellis (1977) on the structure of spacetime in the large, every physically acceptable model of the universe should contain singularities. These can be of two kinds. Singularities in our past, associated with the collapse of the universe as a whole, and singularities in the future, associated with the collapse of stars. In the first case we have an initial singularity in our past. This result has given strong support to the theories known as "big bang" theories in which the universe has an origin, evolving from such a singularity. In the second case the singularities are inside black holes. This is why there is strong confidence on the existence of black holes in nature, although this prediction has not had a direct empirical confirmation to date.

Moreover, even solutions of the Einstein field equation for relatively simple systems, as is the case of the Schwarzschild solution, exhibit singularities. We have therefore to allow for a breakdown of the smoothness of the manifold to make room for such singularities. Axiom 5-iv) asserts that they have to constitute a set of measure zero.

Axiom 5-v) concerns the much discussed principle of general covariance. We have pointed out that this principle is metanomological (see Section 2). In fact, the formulation of general relativity in terms of the abstract geometric mathematical structure of differential geometry grants the requirements of making its laws independent of the coordinate system we care to select. But this does not mean that every coordinate system is suited to physical interpretation. For example, we would not take as physical a coordinate system which makes simultaneous two events connected causally. This means that we have to take our "time" coordinate to be consistent with what we can measure with clocks, that is, with physical time.

This is a very important point. Since the proposal of the special theory of relativity we have become used to obliterating the distinction between space and time and considering spacetime as a four-dimensional continuum. Yet, physically (and philosophically) the distinction between these two structures remains. For example, the equations of motion privilege time.

It is possible, as Bunge (1967) does, to restrict GR's covariance to those coordinate systems which satisfy the requirements to be physically acceptable. I prefer to formulate general covariance as "general" as it can be, and to postulate that every manifold which is to model spacetime for a physical situation can always be coordinated in such a way that we can distinguish space from time. But this is a conventional matter.

Axiom Group III

As was mentioned previously, the definition of the Einstein tensor requires two differentiations, and the fulfillment of the Bianchi identities demands one more degree of differentiability. These differentiations are made on the metric tensor field and therefore this tensor field has to be class C^3.

Aside from the need to make room for the singularities associated with the manifold itself, we have to allow for discontinuities in the metric which can arise from considering bodies with sharp boundaries (gravitational shock waves, for example). We postulate that these discontinuities have to take place in a set of Lebesgue measure zero.

Axiom Group IV

The first part of axiom vii) asserts that the physical reference systems are nongravitational in nature. A gravitational field does not qualify as a reference frame.

In its second part, it is stated that the description of physical phenomena using reference frames should be made locally. As it happens, it is impossible to cover spacetime with a single chart. The simplest examples of a non-Euclidean manifold require more than one chart to be successfully coordinated. This corresponds to the fact that in nature a single reference system does not suffice for its description. So, for every chart, there must be in the conceptual scheme a moving frame (tetrad for $n=4$) and a physical reference frame such that the first mirrors the second. The mirroring relation ($\hat{=}$) is semantic and the reader is referred to Bunge's semantics (Bunge, 1974*a*) for details.

Axiom Group V

Every physical field theory is formulated by a relation between the sources of the field and the field itself. In GR this is done with Einstein's equations, which are postulated next.

Axiom 5-viii) refers to the sources of the gravitational field in GR. These are, as we have pointed out (see the comments to the axiom group I), all macroscopic nongravitational physical systems. The pertinent information is theoretically given in the energy-momentum tensor.

This specification is far from trivial. In the first place, a tensor field "lives" in a manifold, so the latter should be known in advance of specifying an energy-momentum tensor. This amounts to knowing in advance the spacetime structure, i.e., the field that the system determines, information we cannot have until we solve Einstein's equations. To avoid this circularity I postulate the existence, not of an energy-momentum tensor itself, but of a tensor field functional of the metric tensor and of whatever state variables are pertinent to the problem. I will deal with this point in connection with the axiom postulating the Einstein equations in the next section of comments.

The next point to be discussed is that GR makes no provision for computing the energy-momentum tensor associated to a physical system; it has to be postulated. The most common procedure for doing so is to find out first what is the association in the special theory of relativity. One then gives the tensor a covariant form, i.e., one finds the expression of the same form which can be formulated in terms of the abstract geometric concepts of GR background. For example, ordinary derivatives go into covariant derivatives, etc. What we mean by "form" is not very clear, but since the energy-momentum tensor is postulated and not derived, there is no need for elucidating such a concept.

And finally, we have to resort to other theories to postulate the energy-momentum tensor. Concepts like charge, spin, and even the concept of mass, which is central to gravity, are foreign to GR. General relativity is not an autonomous theory; it leans heavily on other physical theories, notably the special theory of relativity.

The mathematical properties of the energy-momentum tensor are not very demanding; it has to be class C^1. This is so because the Bianchi identities must hold (see next set of comments). We allow for a set of Lebesgue measure zero in which its differential properties could break down.

Axiom Group VI

This is the most important axiom in this scheme. As is well known, the aim of every physical theory is to describe exhaustively the behavior of every physical system in its referents [see this term in Bunge's semantics (Bunge, 1974*a*)]. Einstein's equations determine the gravitational field "produced" by a nongravitational system. By symmetry, they comprise ten differential equations for the ten components of the metric tensor $g_{\mu\nu}$. So it would seem that we have a consistent and well-posed mathematical problem. However, the situation is more complex than this.

To start discussing the status of the Einstein equations I shall refer first to the epistemological point touched on in the previous section. As was mentioned, the energy-momentum tensor cannot be known in advance of solving Einstein's equations. On the other hand, we need the energy-momentum tensor to formulate the equations at all. To avoid this problem relativists postulate the energy-momentum tensor not explicitly but as a functional of the metric tensor and of whatever state variables are pertinent in the problem. In this way, the unknowns of the problem (the ten components $g_{\mu\nu}$) occur on both sides of the equation. For example, if we are dealing with a perfect fluid, the energy-momentum tensor functional reads

$$\mathbf{L} = (p + \rho)\mathbf{u} \otimes \mathbf{u} + p\mathbf{g} \tag{6.1}$$

Here, the occurrence of the metric tensor is clear, and the state variables which are introduced ar the proper pressure p, the proper density ρ, and the four-velocity vector $\mathbf{u}$. Actually, some more state variables are introduced in this description, although they do not occur in the energy-momentum tensor.

Evidently by solving Einstein's equations we must come out, among other things, knowing the energy-momentum tensor explicitly, and this requires determining the state variables, which is equivalent to enlarging the set of unknowns of the problem. This calls for more equations if the mathematical problem is to be complete. The additional equations are indeed introduced into the problem and are called constitutive equations.

As is the case with many fundamental concepts in GR, the constitutive equations come from outside the theory. In the case of the perfect fluid (6.1), thermodynamic considerations are essential, and indeed one of the constitutive equations to be considered is the first law of thermodynamics.

With this equation counting the mathematical problem may appear complete. Yet it is not so. The Bianchi identities

$$\nabla \cdot \mathbf{G} = 0 \tag{6.2}$$

where $\mathbf{G}$ is the Einstein tensor (see Section 2) hold in the theory. These are not equations, but identities, which really are four differential relations among the ten unknowns $g_{\mu\nu}$ which hold independently of the ten algebraic-independent relations provided by the Einstein equations. This amounts to reducing the number of independent equations from ten to six. This means that there are four degrees of freedom which correspond exactly to the four degrees of freedom we have to select a coordinate system (four x^{μ}). Therefore, to solve Einstein's equations we must impose what are called coordinate conditions; that is, we select a class of coordinate systems in advance.

Moreover, we are sure that the Einstein tensor $\mathbf{G}$ obeys the Bianchi identities *a fortiori* and from this we demand that the energy-momentum tensor should satisfy what is called an energy-momentum conservation "law":

$$\nabla \cdot \mathbf{L} = \mathbf{0} \tag{6.3}$$

which generally imposes differential relations among the state variables. The constitutive equations supplemented by equations (6.3) should constitute an equal number of equations and unknowns when coordinate conditions have been selected.

The last equations above are usually considered as an energy and momentum conservation "law". Unfortunately we are in want of such a result in GR. In fact, the situation is even worse: we do not have a proper definition of energy, nor of momentum in our theory. By proper I mean essentially covariant. We only have partial results applicable to asymptotically flat spacetimes. The problem can be traced to the fact that the Einstein equations are (highly) nonlinear and this can be interpreted as the gravitational field becoming a source of itself in an infinite mechanism. That is to say: once we have a gravitational field produced by any source, it generates an additional gravitational field, which in turn is also a source of gravitational field, etc. But the feature of spacetime of being locally flat or, in more familiar terms, the principle of equivalence, allows us to cancel locally any gravitational field by a mere change of coordinates. On the other hand, any consideration of energy and momentum has to take into account the contribution of the gravitational field itself, which, as can be inferred from the above considerations, cannot be covariant. The energy and momentum of the gravitational field have been described in a nonunique way by quantities called "pseudotensors."

Also, the field equations are usually thought of as the mechanism by which matter determines the structure of spacetime. This is one of the versions of the highly ambiguous Mach principle. In fact, the above comments and the feature of GR of including spacetimes for $\mathbf{L}=0$ make this assertion as ambiguous as the principle itself. There are theories like the Jordan and Brans–Dicke theory (Brans and Dicke, 1961) which try to incorporate Mach's principle in a more satisfactory way.

A final consideration will be made on this axiom. Of the four forces of nature, the gravitational force is the only one which governs the dynamics of the universe in the large (astronomical dimensions). This is so because the so-called strong and weak forces, which are displayed by elementary particles, are very short-ranged. They act at distances of the order of 10^{-13} cm; at greater separations they are practically nil. So they do not play a part in the interaction of celestial objects. The electromagnetic force is long-ranged, but the world, as we know it, is remarkably neutral (equal amounts of positive and negative charge).

Any theory of gravitation, then, must describe cosmology, and the models of the universe provided by GR which are consistent with the empirical evidence (which is essentially that the latter is homogeneous and isotropic at length scales of 10^8 light-years and greater) are nonstatic. This fact disturbed Einstein so much that he was led to modify his field equations in a way that later, when Hubble (1929) in 1927 discovered the expansion of the universe, he called "the biggest blunder of my life." He proposed instead of axiom 5-ix) the field equations

$$\mathbf{G}+\Lambda\mathbf{g}=\kappa\mathbf{L} \tag{6.4}$$

where Λ is a number called the cosmological constant. In this way $\mathbf{G}$ would not be zero in the vacuum ($\mathbf{L}=0$).

Although at present there is no need for the cosmological constant, many relativists are unwilling to abandon it on cosmological grounds. One of the motivations to keep it is a reinterpretation of the field equations (6.4). If they are written as

$$\mathbf{G} = \kappa(\mathbf{L} + \mathbf{L}^{(\mathrm{vac})}) \tag{6.5}$$

with

$$\mathbf{L}^{(\mathrm{vac})} = -\frac{\Lambda}{\kappa}\,\mathbf{g} \tag{6.6}$$

$\mathbf{L}^{(\mathrm{vac})}$ is construed as the energy-momentum tensor of the vacuum, which would be completely unobservable, except for its gravitational effects. It is expected that a quantum theory of gravitation would give an estimation of such a quantity. But if this is so, $\mathbf{L}^{(\mathrm{vac})}$ would be an object belonging to another theory, not to GR. The notion of endowing the vacuum with energy, in the form of virtual matter, so common in quantum field theory, is foreign to GR. Besides, even if we concede the objective existence of possible agents producing a gravitational field in the vacuum, the empirical evidence forces the theories dealing with them to be consistent with the conspiracy of the latter as to not manifest collectively at all. In other words, what we empirically know about the universe is consistent with a zero cosmological constant, a fact which puzzles many researchers who believe in such agents, extrapolating the notion of vacuum surrounding microscopic systems like a hydrogen atom to the vacuum surrounding the most macroscopic systems, as are galaxies and clusters of galaxies. They find that this particular value is peculiar. Perhaps what is needed is an elucidation of the concept of vacuum.

What I am presenting in this article is an axiomatization of the general theory of relativity in a form which is consistent with all the available empirical evidence, and conspiracies or not, the latter indicates that the cosmological constant should be zero. So axiom 5-ix) is postulated as Einstein originally did.

There is an additional motivation for keeping the cosmological constant, which will be only briefly described here. The idea of a big-bang genesis of the universe was first proposed by George Gamow (1948). To account for a universe as we know it nowadays very special initial conditions must have held at the moment of creation. As someone put it, finding the present universe in this model would be as unlikely as finding a pencil balanced on its point after an earthquake. In 1980 Alan H. Guth (1981) proposed a model called the inflationary universe according to which the latter expanded a factor of 10^{30} in a fraction of 10^{-30} s. This scheme solved many of the original big-bang problems, but it is not altogether free from particular initial conditions needed to account for our present universe. These are broader and far more plausible, but still needed. Guth proposed that the early universe consisted of so-called scalar-field particles, which are not the furniture of the universe as we know it today, but that are very common in many theories. The density of this scalar field had to be approximately constant, and this requirement or initial condition is equivalent to postulating the cosmological constant for the young universe, at least for a brief time.

It is expected that many problems will be solved and a more accurate knowledge of our universe will be gathered with the so-called gravitational wave atronomy. The most optimistic researchers in the field trust that we will be able to detect (if they exist!) gravitational waves at the turn of the century.

Axiom Group VII

In physics, when we relate concepts with different referents, we have to couple them, if at all to make the relationship dimensionally consistent. It is said that in Einstein's field equations the geometry occurs on the left-hand side and that the dynamics on the right-hand side. We couple these different concepts with the coupling constant κ, which has to have the value, in terms of the fundamental constants G and c, given in the axiom, in order to retrieve Newton's gravitation theory in the circumstances where the latter should hold. In these, gravitation should be weak and all velocities in the system should be small compared with the speed of light c. Of course, these conditions can only hold in coordinate systems very nearly Galilean.

The metanomological statement called the correspondence principle, which states that if a physical theory is to supersede another one, it should reduce to the latter where it has proved to be successful, is correctly satisfied by GR in relation with Newton's theory. This requires the coupling constant to be postulated as we did.

With the aim of not being repetitive, the reader is referred to Bunge's axiomatization of GR (Bunge, 1967) for additional interesting comments, notably in relation to the issue of regarding the theory a geometrization of physics.

7. THEOREMS

Theorem 1. The equivalence principle. For every gravitational field and for each point of spacetime, there is a coordinate system in which the field vanishes at the point.

Proof. This is just the statement that for every point $q \in M^n$, there exists a neighborhood U of q in which we can choose normal coordinates and in which $\Gamma^{\mu}_{(\alpha\beta)}|_q = 0$, where $\Gamma^{\mu}_{(\alpha\beta)} = \frac{1}{2}(\Gamma^{\mu}_{\alpha\beta} - \Gamma^{\mu}_{\beta\alpha})$ is the antisymmetrization of the connection. For a Riemannian connection, which is what I assume, we have $\Gamma^{\mu}_{\alpha\beta} = 0$.

Remarks. For remarks about the connection see the comments in the Appendix. In GR the components of this quantity represent the gravitational field. The metric tensor is the gravitational potential, which does not qualify as the field because, among other things, it does not have the property of the connection enunciated in this theorem.

Strictly speaking, the field vanishes only at a point. The principle of equivalence is often stated as the fact that it is possible to make a static homogeneous gravitational field vanish (Bunge, 1967) by a mere change of coordinates.[2] If such fields (static homogeneous) would exist in GR, the statement would be true. However, this is not the case. Actually, there are no solutions (curved spacetimes) to the homogeneous and isotropic field equations (Robertson–Walker-like element) for the vacuum. The simplest solutions arise if we relax the condition of isotropy (Bianchi universes).

However, this theorem is in accordance with the heuristic equivalence principle in that the vanishing of the field at one point guarantees that the field will be arbitrarily close to zero in a region of spacetime sufficiently close to the point, on account of the continuity of the connection. In fact, the region can be chosen small enough so that no physical instrument can detect any field inside it.

Theorem 2. The geodesic postulate. If $\bar{\sigma} \in \bar{\Sigma}$ represents a (structureless) test particle, then its corresponding $\sigma \in \Sigma$ is such that in a chart with coordinates x^{μ}

$$\frac{d^2x^{\mu}}{ds^2} + \Gamma^{\mu}_{\alpha\beta}\frac{dx^{\alpha}}{ds}\frac{dx^{\beta}}{ds} = 0 \tag{7.1}$$

where s is the "proper time," $ds^2 = g_{\mu\nu}\,dx^{\mu}\,dx^{\nu}$. This is the equation of a geodesic.

Proof. The corresponding energy-momentum tensor for a point structureless particle is

$$L^{\alpha\beta} = m\delta(x^{\gamma} - x^{\gamma}(s))u^{\mu}u^{\nu} \tag{7.2}$$

[2]Some coordinate systems refer to reference systems, although not all of them. There are many coordinate systems suited to a reference system. In the equivalence principle the adequate reference system is what we look for.

where u^μ is the four-velocity of the particle, m is the mass of the particle, and the Dirac δ function is treated as a scalar. It is defined so that

$$\int f(x^\alpha)\delta(x^\delta - x_0^\delta)\sqrt{-g}\, d^4x = f(x_0^\alpha) \tag{7.3}$$

for any function f (of compact support[3]) of the coordinates. Now, equation (6.3) in tensor components can be written as

$$\partial_\nu(L^{\mu\nu}\sqrt{-g}) = -\Gamma^\mu_{\alpha\beta}L^{\alpha\beta}\sqrt{-g} \tag{7.4}$$

If we substitute (7.2) into (7.4) and integrate over any four-dimensional volume containing $x^\delta(s)$, equation (7.1) follows straightforwardly.

Remarks. This theorem applies to a test particle moving in the field $\Gamma^\mu_{\alpha\beta}$ produced by other massive objects. The corrections to the field due to the presence of the particle itself are ignored. This is precisely what is meant by a "test" particle. The energy-momentum tensor (7.2) is understood to be defined outside the region where the rest of the mass-energy resides, and the region of integration is taken so as to exclude the latter.

It must be warned that the use of Dirac δ functions to specify the energy-momentum tensor of test particles causes an infinite correction to the field at the position of the particles. However, any other method used to deal with the "test" feature of the particles is unsatisfactory in this and/or many other respects.

Our test particle is pointlike and structureless. If structure is taken into account, the geodesic postulate would fail to hold; that is, the particle would not move along a geodesic of the complete field.

I close with a very important final remark. In every classical field theory before GR the field equations were supplemented by the appropriate equations of motion. When GR was formulated it was believed to share this feature, and the geodesic postulate was thought to be necessary in this sense. Later it was discovered that in GR the equations of motion are included in the field equation through equation (6.3).

APPENDIX. DEFINITIONS

1. A topological manifold X^n of dimension n is a Hausdorff topological space such that every point has a neighborhood homeomorphic to R^n.

2. A chart (U, φ) of a manifold X^n is an open set U of X^n called the domain of the chart, together with an homeomorphism $\varphi: U \to V$ onto an open set $V \subset R^n$.

[3]The support of a function is the set where it does not vanish.

3. An atlas of class C^k on a manifold X^n is a set $\{(U_\alpha, \varphi_\alpha)\}_{\alpha \in I}$ of charts of X^n such that the domains $\{U_\alpha\}_{\alpha \in I}$ cover X^n and the homeomorphisms satisfy the following compatibility condition: the maps $\varphi_\beta \circ \varphi_\alpha^{-1}\colon \varphi_\alpha(U_\alpha \cap U_\beta) \to \varphi_\beta(U_\alpha \cap U_\beta)$ are maps of open sets of R^n into R^n of class C^k.

4. Two C^k atlases $\{(U_\alpha, \varphi_\alpha)\}_{\alpha \in I}$ and $\{(U_\beta, \varphi_\beta)\}_{\beta \in J}$ are equivalent if and only if their union (domains $\{U_\alpha\} \cup \{U_\beta\}$ and homeomorphisms $\{\varphi_\alpha\} \cup \{\varphi_\beta\}$) is again a C^k atlas.

5. A C^k manifold M^n of dimension n is a topological manifold together with an equivalence class of C^k atlases (a C^k structure).

6. T_x denotes the tangent vector space at $x \in M^n$, and T_x^* its dual. $\Lambda^p(M^n)$ denotes the set of all p-form fields defined over M^n.

7. A pseudo-Riemannian manifold is a C^1 manifold M^n, together with a continuous 2-covariant tensor field $\mathbf{g}$, called the metric tensor such that:

(i) $\mathbf{g}$ is symmetric.

(ii) For each $x \in M^n$, the bilineal form $g_x\colon T_x \times T_x \to R$ ($\mathbf{g}_x$ is the 2-covariant symmetric tensor associated by the field to the point x) is nondegenerate that is, $\mathbf{g}_x(\mathbf{u}, \mathbf{v}) = 0\ \forall \mathbf{v} \in T_x$ if and only if $\mathbf{u} = 0$].

8. Any set of n linearly independent differentiable vector fields $(\mathbf{e}_\alpha)$ which form a basis for the set $H^{k-1}(U)$ of all C^{k-1} vector fields on an open set U of a C^k manifold M^n is called a moving frame (tetrad, for $n = 4$) on U. Such a set may not exist globally on M^n. (On a C^k manifold, there can only exist C^r vector fields with $r \leq k - 1$.)

9. A linear connection on a C^k ($k \geq 2$) manifold M^n is a mapping ∇ from the set $H(M^n)$ of all differentiable vector fields on M^n to the set of all differentiable tensor fields of type (1, 1) on M^n: i.e., $\nabla\colon H(M^n) \to T_1^1(M^n)$, such that:

(i) $\nabla(\mathbf{u} + \mathbf{v}) = \nabla\mathbf{u} + \nabla\mathbf{v}$

(ii) $\nabla(f\mathbf{u}) = df \otimes \mathbf{u} + f\nabla\mathbf{u}$

f is a differentiable function on M^n. $\nabla\mathbf{u}$ is called the covariant derivative or absolute differential of $\mathbf{u}$.

10. The covariant derivative of $\mathbf{w}$ in the direction of $\mathbf{v}$ in an open set U where a moving frame $(\mathbf{e}_\alpha)$ exists with dual basis $(\boldsymbol{\theta}^\beta)$ is

$$\nabla_\mathbf{v}\mathbf{w} = (\nabla\mathbf{w})(\mathbf{v}, \boldsymbol{\theta}^\alpha)\mathbf{e}_\alpha$$

11. The covariant derivative can be extended to tensors of any type by requiring that:

(i) $\Omega_\mathbf{v} f = \mathbf{v}(f)$, f is a differentiable function on M^n.

(ii) $\nabla_\mathbf{v}(\mathbf{t} + \mathbf{s}) = \nabla_\mathbf{v}\mathbf{t} + \nabla_\mathbf{v}\mathbf{s}$.

(iii) $\nabla_\mathbf{v}(\mathbf{t} \otimes \mathbf{s}) = \nabla_\mathbf{v}\mathbf{t} \otimes \mathbf{s} + \mathbf{t} \otimes \nabla_\mathbf{v}\mathbf{s}$.

(iv) ∇ commutes with the operation of contracted multiplication.

Thus, if $\mathbf{t}$ is a tensor of type (p, q), $\nabla\mathbf{t}$ is the tensor type $(q+1, p)$ defined by

$$\nabla\mathbf{t}(\mathbf{v}, \mathbf{v}_1, \ldots, \mathbf{v}_q; \mathbf{w}_1, \ldots, \mathbf{w}_p) = (\nabla_{\mathbf{v}}\mathbf{t})(\mathbf{v}_1, \ldots, \mathbf{v}_q; \mathbf{w}_1, \ldots, \mathbf{w}_p)$$

12. The following theorem is relevant to the axiomatization:

Theorem. On a Riemannian manifold (including a pseudo-Riemannian manifold) there exists a unique connection such that:

(i) $\mathbf{T} = 0$.

(ii) $\nabla\mathbf{g} = 0$.

Here $\mathbf{T}$ is the torsion tensor defined in terms of the torsion operation

$$\tau(\mathbf{u}, \mathbf{v}) = \nabla_{\mathbf{u}}\mathbf{v} - \nabla_{\mathbf{v}}\mathbf{u} - [\mathbf{u}, \mathbf{v}]$$

as

$$\mathbf{T}(\boldsymbol{\alpha}, \mathbf{u}, \mathbf{v}) = \boldsymbol{\alpha}(\tau(\mathbf{u}, \mathbf{v}))$$

In a natural frame ($\mathbf{e}_\alpha = \partial/\partial x^\alpha$), the components of ∇ are denoted by $\Gamma^\mu_{\alpha\beta}$ and the components of $\mathbf{T}$ are $T^\mu_{\alpha\beta} = \Gamma^\mu_{\alpha\beta} - \Gamma^\mu_{\beta\alpha}$.

Remark. We have presupposed the conditions of this theorem to hold throughout this work. That is, we have assumed a symmetric connection which in natural coordinates takes the form, in terms of the metric,

$$\Gamma^\mu_{\alpha\beta} = \tfrac{1}{2}g^{\mu\nu}(\partial_\beta g_{\alpha\nu} + \partial_\alpha g_{\beta\nu} - \partial_\nu g_{\alpha\beta})$$

The so called Einstein–Cartan theory of gravity (Cartan, 1992; Trautman, 1972) does not use the condition (i) above and the motivation to have a nonvanishing torsion was to relate its antisymmetric part with the intrinsic spin of matter.

Note that on account of this theorem we have written $(M^n, \mathbf{g})$ instead of $(M^n, \mathbf{g}, \nabla)$ throughout this work except in the primitive and defined concepts.

ACKNOWLEDGMENTS

I am indebted to Profs. M. Bunge, C. Lämmerzahl, and M. P. Ryan for useful and illuminating discussions.

REFERENCES

Adler, R., Bazin, M., and Schiffer, M. (1965). *Introduction to General Relativity*, McGraw-Hill, New York.

Bergman, P. G. (1976). *Introduction to the Theory of Relativity*, Dover, New York.

Bishop, R. L., and Goldberg, S. I. (1968). *Tensor Analysis on Manifolds*, Macmillan, New York.

Bondi, H. (1961). *Cosmology*, Cambridge University Press, Cambridge.
Bowler, M. G. (1976). *Gravitation and Relativity*, Pergamon Press, Oxford.
Brans, C., and Dicke, R. H. (1961). *Physical Review*, **124**, 925–935.
Brillouin, L. (1970). *Relativity Reexamined*, Academic Press, New York.
Bunge, M. (1967). *Foundations of Physics*, Springer-Verlag, New York.
Bunge, M. (1974*a*). *Treatise on Basic Philosophy*, Vol. 1, Reidel, Dordrecht.
Bunge, M. (1974*b*). *Treatise on Basic Philosophy*, Vol. 2, Reidel, Dordrecht.
Bunge, M. (1977). *Treatise on Basic Philosophy*, Vol. 3, Reidel, Dordrecht.
Bunge, M. (1979). *Treatise on Basic Philosophy*, Vol. 4, Reidel, Dordrecht.
Burke, W. L. (1980). *Spacetime, Geometry, Cosmology*, University Science Books, Mill Valley, California.
Cartan, E. (1922). *Journal de Mathematiques Pures et Appliquées*, **1**, 141–203.
Choquet-Bruhat, Y., De Witt-Morette, C., and Dillard-Bleik, M. (1977). *Analysis, Manifolds and Physics*, North-Holland, Amsterdam.
Eddington, A. S. (1965). *The Mathematical Theory of Relativity*, Chelsea, New York.
Eisenhart, L. P. (1964). *Riemannian Geometry*, Princeton University Press, Princeton, New Jersey.
Fock, V. (1964). *The Theory of Space, Time and Gravitation*, Pergamon Press, Oxford.
Frankel, T. (1979). *Gravitational Curvature, an Introduction to Einstein's Theory*, Freeman, San Francisco.
Gamow, G. (1948). *Nature*, **162**, 680–682.
Guth, A. (1981). *Physical Review D*, **32**, 347.
Hawking, S. W., and Ellis, G. F. (1977). *The Large Scale Structure of Space-Time*, Cambridge University Press, Cambridge.
Held, A. (1980). *General Relativity and Gravitation*, Plenum Press, New York.
Hubble, E. P. (1929). *Proceedings of the National Academy of Sciences of the USA*, **15**, 169–173.
Kobayashi, S., and Nomizu, K. (1963). *Foundations of Differential Geometry*, Interscience, New York.
Landau, L. D., and Lifshitz, E. M. (1962). *The Classical Theory of Fields*, Addison-Wesley, Reading, Massachusetts.
McVittie, G. C. (1956). *General Relativity and Cosmology*, Chapman and Hall, London.
Minkowski, H. (1908). Address delivered at the 80th Assembly of German Natural Scientists and Physicians, Cologne [English translation, Space and time, in *The Principle of Relativity*, Dover, New York, 1923].
Misner, C. W., Thorne, K. S., and Wheeler, J. A. (1971). *Gravitation*, Freeman, San Franscisco.
Papapetrou, A. (1974). *Lectures on General Relativity*, Reidel, Dordrecht.
Raychaudhuri, A. K. (1979). *Theoretical Cosmology*, Clarendon Press, Oxford.
Schutz, B. F. (1985). *A First Course in General Relativity*, Cambridge University Press, Cambridge.
Stephani, H. (1982). *General Relativity, An Introduction to the Theory of the Gravitational Field*, Cambridge University Press, Cambridge.
Synge, J. L. (1960). *Relativity: The General Theory*, North-Holland, Amsterdam.
Tolman, R. C. (1987). *Relativity Thermodynamics and Cosmology*, Dover, New York.
Trautman, A. (1972). *Bulletin de l'Academie Polonaise des Sciences Serie des Sciences Mathematiques, Astronomiques et Physiques*, **205**, 185–190.
Wald, R. M. (1984). *General Relativity*, University of Chicago Press, Chicago.
Weinberg, S. (1972). *Gravitation and Cosmology: Principles and Applications of the General Theory of Relativity*, Wiley, New York.