# Observables in General Relativity*

PETER G. BERGMANN

*Syracuse University, Syracuse 10, New York*

THE general theory of relativity is the result of efforts to make gravitation a part of the relativistic physics that had in turn been necessitated by the fact that electromagnetic waves propagate without regard to the motion of our earth relative to the bulk of the surrounding universe. The so-called special theory of relativity, as is well known, led to modifications in the Newtonian concept of space and time that rendered instantaneous action at a distance unacceptable, not only for electromagnetic interactions, but for any forces whatsoever. Henceforth, a relativistic theory of gravitation would have to be a relativistic field theory, that is to say, a theory of the gravitational field. Because the inverse-square law holds for gravitational forces just as it does for electrostatic forces, it was a foregone conclusion that the gravitational field would propagate with the speed of light also. We know that relativistic fields do so if they obey a D'Alembert-type of equation, with vanishing mass term. So far, there seemed to be no need for a further modification of the space-time concepts of Einstein's 1905 paper on the electrodynamics of moving bodies.

But the gravitational forces differ from all other forces known in one respect: The sources of the field, that is to say the gravitational masses, equal the inertial masses of the bodies that gravitate. This is the classical field theorist's way of stating the law that under the influence of gravity all material bodies suffer the same acceleration. That this is so has been established with a very high degree of accuracy (at least $10^{-8}$) by several experimenters in the 20th century, first by Eötvös,[1] and most recently Dicke.[2] Thus, a motley collection of bodies exhibit under gravity a behavior that is analogous to their behavior under the influence of so-called inertial forces, such as centrifugal, or Coriolis forces. Thus, at least by local experiment, we are unable to separate clearly gravitational from inertial forces. This circumstance, which is known as the *principle of equivalence*, defeats any attempt to determine, again by local experiment, an inertial frame of reference; this "principle of impotence" goes far beyond the restricted principle of relativity, which postulates that we cannot determine a frame of absolute rest. According to the principle of equivalence, we cannot even determine a frame of uniform unaccelerated motion. From this principle of impotence, Einstein moved forward to the postulate that for the description of the laws of gravitation all frames of reference, that is to say, all four-dimensional space-time coordinate systems, are equivalent. In the presence of several different force fields, which, however, include the forces of gravity, this principle is to remain valid. A theory of nature that satisfies this postulate, the principle of *covariance*, is called a general-relativistic theory. The general theory of relativity proper is Einstein's theory of the gravitational field[3] of 1916. In what follows we shall deal with this theory, as the archetype of a general-relativistic theory.

Before launching into the subject proper, one more remark may be made: The transition from the narrow statement that in the presence of gravitational fields we cannot by local experiment determine an inertial frame of reference to the principle of covariance, represents a logical hiatus. Einstein felt compelled to adopt this extreme postulate after he had made a number of attempts to get along with more conservative approaches. Today, like any other physical theory, the principle of covariance derives its validity from physical reality; in this case, that it forms the foundation of the best available theory of the gravitational field. It is perfectly feasible that the principle will be modified by subsequent developments. We shall return to this point in the last section. For the present, we consider the principle of covariance as established.

## CONCEPT OF OBSERVABLES

Given the validity of the principle of covariance, the identification of a location-plus-instant-in-time by means of four real numbers becomes a matter of almost unrestricted choice, one that has to be exerted not only once in the course of the history of a physical situation but again and again. Suppose we adopt at some particular instant in time a three-dimensional coordinate system (curvilinear, to be sure), an instantaneous local rate of increase for the time coordinate, and also agree on the instantaneous rate of motion of the space coordinates. Even after all these choices are made, nothing in the theory tells us how to choose coordinates some finite time later. Accordingly, in a general-relativistic theory—and this is in sharp contrast to a Lorentz-covariant theory—no initial data enable us to predict the values of local field variables at some later time, simply because the identification of the future time and location by means of four coordinate numbers is meaningless (Fig. 1).

* Talk delivered February 2, 1961, before the American Physical Society, in New York City. This work was supported by the National Science Foundation, the U. S. Air Force Office of Scientific Research, and the Aeronautical Research Laboratories of Wright Air Development Division.

[1] R. V. Eötvös, Math. u. naturw. Ber. Ungar. **8**, 65 (1890); Ann. Physik **68**, 11 (1922).

[2] R. H. Dicke, Science **129**, 621 (1959).

[3] A. Einstein, Ann. Physik **49**, 769 (1916).

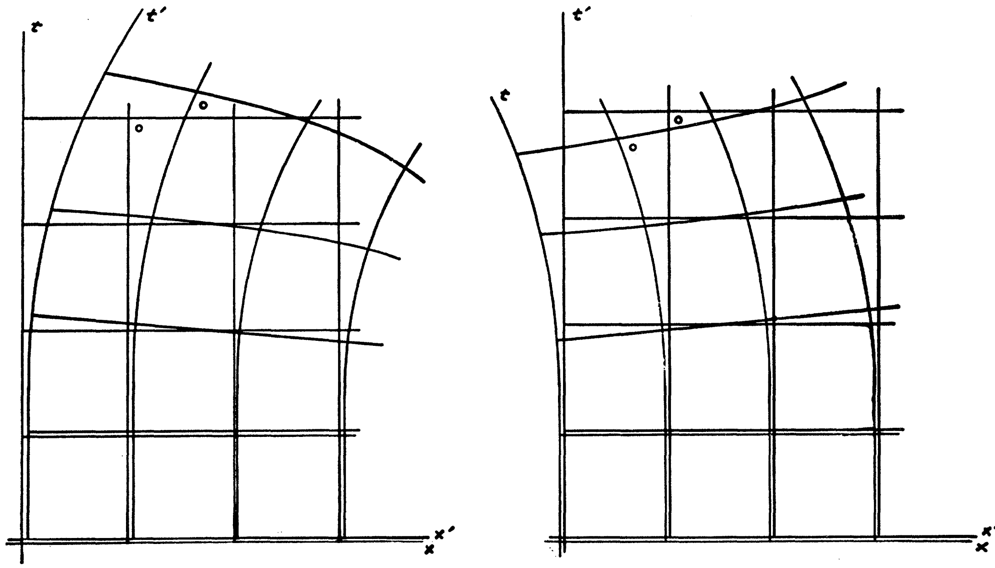


FIG. 1. Two different coordinate systems drawn in alternative ways. The two points are identical in the two sketches, but different points occupy the same position.

Inability to predict the values of locally defined field variables is a fundamental property of general-relativistic field theories, regardless of their detailed aspects. Nevertheless, such a theory can be completely causal, in a slightly modified sense. In view of the fact that a general-relativistic quantum theory has not yet been constructed in complete detail, we shall discuss principally *c*-number theory. It is to be assumed that the quantum theory will be related to nonquantum theory pretty much as it is outside general relativity.

Causality may be redefined in general relativity as follows: It is possible to give sufficient data at one time—on a spacelike hypersurface within the four-dimensional space-time manifold—to identify completely a solution of Einstein's field equations. Given such a set of complete initial data, all solutions of the field equations compatible with the given data can be transformed into each other simply by means of appropriate coordinate transformations.[4,5] The set of all solutions that can be transformed into each other by coordinate transformations is called an *equivalence class* of solutions.[6] All members of an equivalence class are formally different statements of the same physical situation in space-time. Identification of an equivalence class enables us to answer all questions that have been properly constructed, that is to say, questions that refer to the outcome of physical experiments rather than to the functional dependence of field variables on an (arbitrary) set of coordinates. As an example of a proper question, we may consider one concerning "coincidences": Suppose our theory enables us to construct from the field variables several scalar fields; we might then identify a "world point" (location-plus-instant-in-time) by the values assumed by four of these scalars and ask for the value, there and then, of a fifth field. Such a question can, we are assured, always be answered from a sufficient set of initial data, though the performance of this task may call for considerable mathematical agility.

Because the ordinary field variables taken singly contain almost no physical reality, because they cannot be predicted by an otherwise causal theory, we have introduced the concept of *observable*, a physical quantity that represents more nearly the stuff of which physical reality is made.[7] The name is of course borrowed from standard quantum terminology, and the expectation is that the quantum observables in a covariant quantum field theory will be quantities of the kind about to be described. We shall call a quantity an observable if it can be predicted uniquely from initial data.

This definition is complete, but not always the most useful. A somewhat modified definition would be to call an observable a quantity that is invariant under a coordinate transformation that leaves the initial data unchanged. If the initial data are given on a surface that is a nonzero distance away from the region on which our observable is defined, then, because of the arbitrary choice of coordinates, the observable must be an invariant, without further qualifications.

If we employ a Hamiltonian formalism to describe our theory, then there is a certain set of generators that produce infinitesimal coordinate transformations. An observable is then a dynamical variable that has vanishing Poisson brackets with all the generators of

[4] P. A. M. Dirac, Proc. Roy. Soc. (London) **A246**, 333 (1958); Phys. Rev. **114**, 924 (1959).

[5] A. Lichnerowicz, *Théories relativistes de la gravitation et de l'électromagnétisme* (Masson et Cie, Paris, 1955).

[6] P. G. Bergmann and A. B. Komar (to be published).

[7] P. Bergmann, Nuovo cimento **3**, 1177 (1956).

infinitesimal coordinate transformations. This last definition is rather useful: The generators of infinitesimal coordinate transformations are known in closed form. One can construct observables by systematic approximation, starting from crude but very intuitive expressions that are related to "plane waves" of the gravitational field.

We mentioned the concept of equivalence classes of solutions. With its help one can give the observables yet another interpretation. Suppose we construct a Hamiltonian formalism, with configuration-type field variables and momentum-type field variables. Just as in electrodynamics, some of these canonical variables are related to each other by equalities that contain no time derivatives of any canonical variables. Such equalities are called *constraints*.[8] Initial data in a Hamiltonian formalism must be chosen so as to be consistent with these constraints at some initial time $t_0$, but thereafter the constraints maintain themselves by virtue of the Hamiltonian equations of motion. Now it turns out that just these constraints are the generators of infinitesimal transformations.[9] The precise expressions are

$$\Gamma = -\int \xi^\rho H_\rho d^3x, \tag{1}$$

where

$$\xi^\rho \equiv \delta x^\rho \tag{2}$$

and

$$H_\rho \equiv e_\rho{}^s \mathcal{H}_s + l_\rho \mathcal{H}_L, \quad e_\nu{}^\mu \equiv \delta_\nu{}^\mu - l^\mu l_\nu, \quad e_\nu{}^0 \equiv 0. \tag{3}$$

The symbols $\mathcal{H}_s$ and $\mathcal{H}_L$ were introduced by Dirac and represent constraints. Given an expression that can be constructed from canonical field variables, its infinitesimal transformation law is represented by its Poisson bracket with the generator (1). For instance, the three spatial components of a covariant vector transform according to the law

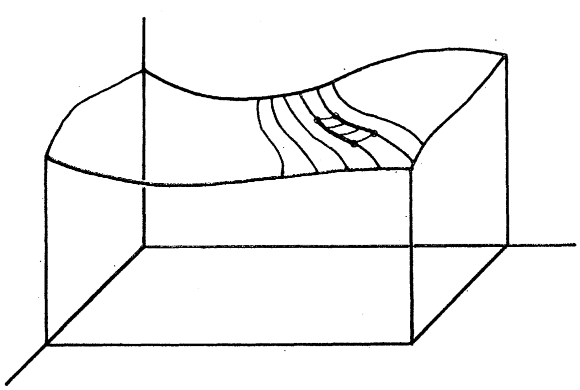

FIG. 2. Constraint hypersurface, equivalence classes, and a canonical mapping generated by an observable.

[8] P. Bergmann, Phys. Rev. **75**, 680 (1949).
[9] J. L. Anderson and P. G. Bergmann, Phys. Rev. **83**, 1081 (1951).

$$\bar{\delta} A_s = -(\partial \xi^\rho / \partial x^s) A_\rho - (\xi \cdot \nabla) A_s - \xi^0 (\partial A_s / \partial t) \equiv [A_s, \Gamma]. \tag{4}$$

The generators (1) enable us to navigate about within one equivalence class of solutions, that is to say, they map any solution on another one belonging to the same equivalence class. It follows that an observable generates a mapping that maps all members of one equivalence class on the members of one other equivalence class.[6] Figure 2 shows a highly simplified sketch of the phase space of a general-relativistic field theory. Because of the constraints, only representative points lying on the constraint hypersurface represent physically possible situations. But the constraint hypersurface is further subdivided into equivalence classes, which in Fig. 2 appear as curves. In reality, both the original phase space and the final equivalence classes are infinite-dimensional domains. Because constraints commute with observables, it is immaterial as to which of the two mappings, one representing a coordinate transformation and the other being generated by an observable, is performed first. The little parallelogram now indicates the course of the proof: Two points connected by a coordinate transformation must be mapped by the same observable into two points again related to each other by a coordinate transformation. Hence our conclusion that observables generate mappings of equivalence classes on each other.

Observables taken as generators map whole equivalence classes on each other. It is also clear that within any equivalence class all observables are constant. Two distinct equivalence classes must differ from each other with respect to at least one observable. We shall call a set of observables *complete* if knowledge of the whole set is sufficient to identify an equivalence class and distinguish it from all others. A complete set might be redundant. We shall call it a *minimal complete set* if knowledge of the numerical values of the whole set is required to identify an equivalence class.

We may use a minimal complete set of observables as a coordinate system in a new space, a space whose points represent equivalence classes. In this space the observables (not only those belonging to the minimal complete set) generate mappings. Thus, if we define commutator brackets properly, this new space may be interpreted as a new *reduced phase space*,[6] whose canonical transformations are generated by all the functions we can define on its points. Observables are simply these functions. It seems likely that an eventual quantum theory of the gravitational field will be approximated by the classical canonical transformation theory of the reduced phase space.

## CONSTRUCTION OF OBSERVABLES

So much for the general theory. How can we construct observables, and how, more particularly, a complete but minimal set of them? There have been several approaches to this problem, and we describe one that was initiated by Géhéniau and Debever,[10] of Brussels, and further developed by Komar and Bergmann, at Syracuse.[11,12] This approach is based on the idea of "coincidences" between various events, as described previously. One first chooses four scalars characteristic of local properties of the space-time manifold to identify world points; then any additional fields, described in terms of the new "intrinsic" coordinates, are observables. For a set of intrinsic coordinates it is essential that world points are identified completely in terms of intrinsic geometric properties of the space-time manifold. There cannot be a remaining set of transformations compatible with the definition chosen for our intrinsic coordinates.

There are exactly four independent scalars that can be formed from the curvature of a four-dimensional manifold whose metric satisfies Einstein's field equations. One form in which they may be written is

$$\begin{aligned} A^1 &= \mathrm{Tr}\,(gCgC), \\ A^2 &= \mathrm{Tr}\,(gC\epsilon C), \\ A^3 &= \mathrm{Tr}\,(gCgCgC), \\ A^4 &= \mathrm{Tr}\,(gCgC\epsilon C). \end{aligned} \tag{5}$$

These four expressions have been rendered in a compressed notation, which is usually credited to Petrov.[13] It involves the introduction of indices that run from 1 to 6, each representing an antisymmetric pair of indices of the four-dimensional manifold. Thus, Weyl's tensor, the remainder of Riemann's curvature tensor after half of its components have been fixed by the field equations, becomes a symmetric form of rank 2, symbolized in (5) by $C$. The symbol $g$ represents a similar matrix, formed from the components of the metric tensor, whereas $\epsilon$ stands for Levi-Civita's alternating tensor. The four expressions $A^1\cdots A^4$ exhaust the scalars that can be formed from the metric tensor at this differential level. Thus, they recommend themselves by their relative simplicity, though later developments may lead to a preference for other scalars. Any set of four functions $f^\rho(A^1\cdots A^4)$ may now serve as intrinsic coordinates.

The components of the metric tensor in the chosen intrinsic coordinates are observables. Moreover, they form a complete set, as knowledge of the metric is obviously sufficient for determination of all properties of our manifold. They are, however, not minimal: The metric must satisfy the field equations, as well as the conditions imposed on the coordinates, that they equal the $f^\rho$. If these conditions are all taken into account, it turns out that a minimal complete set of observables should number four pieces of numerical data at every point of a three-dimensional hypersurface. All other observables are then determined by the field equations and the coordinate conditions. We have not succeeded in isolating this minimal set, and there is some reason for believing that this is not possible in closed form.

Given some arbitrary fields, which have been turned into observables with the help of an intrinsic coordinate system, how do we construct proper commutator brackets between them, so that the structure of the reduced phase space may be made apparent? This question is not entirely trivial, as the value of an observable remains unchanged if we add to it an arbitrary combination of the constraints of the formalism, the expressions $\mathcal{H}_s$ and $\mathcal{H}_L$, Dirac's Hamiltonian constraints, and the expressions $(f^\rho - x^\rho)$, the coordinate conditions. However, the naively conceived Poisson brackets between two field variables are changed by such an addition. The answer was worked out by Dirac[14] about ten years ago: We must modify the Poisson brackets in such a manner as to assure ourselves that our canonical transformations map equivalence classes intact on equivalence classes. There are several ways in which we can visualize Dirac's prescription. One good way is this: As an observable must commute with all the constraints of the theory, we add to an arbitrarily defined field variable just that combination of constraints that make the Poisson brackets of the redefined variable with all constraints of the theory vanish.

Whatever the visualization, in our problem Dirac's technique leads us to the following new commutator brackets:

$$\{A,B\} = [A,B] + [[A,\mathcal{G}],[B,\mathcal{G}]], \tag{6}$$

where

$$\mathcal{G} = \int H_\rho (f^\rho - x^\rho) d^3x, \tag{7}$$

and where the square brackets denote ordinary Poisson brackets. The performance of the calculations indicated on the right-hand side is straightforward but rather laborious. No completed example can be shown as yet, though it may be only a short while away.

## GENERAL COVARIANCE VS LORENTZ COVARIANCE

The approach sketched here has been criticized[15] on the grounds that it leads to a Hamiltonian which not only vanishes but which forms vanishing commutator brackets with all observables. This state of affairs can also be produced in classical mechanics when we go from ordinary ("natural") canonical variables to those introduced by the Hamilton-Jacobi transformation. Thus, it would seem that the technique described here corresponds to a Hamilton-Jacobi formulation, whereas a physically intuitive description should lead to a

[10] J. Géhéniau and R. Debever, Bull. classe sci. Acad. roy. Belg. **42**, 114, 252, 313, 608 (1956); Helv. Phys. Acta Suppl. **4**, 101 (1956).
[11] A. B. Komar, Phys. Rev. **111**, 1182 (1958).
[12] P. Bergmann and A. Komar, Phys. Rev. Letters **4**, 432 (1960).
[13] A. Z. Petrov, Sci. Not. Kazan State Univ. **114**, 55 (1954).
[14] P. A. M. Dirac, Can. J. Math. **2**, 129 (1950).
[15] R. Arnowitt, S. Deser, and C. Misner, Phys. Rev. **118**, 1100 (1960).

nonvanishing Hamiltonian, which describes in the accustomed manner the dynamics of the gravitational field.

The analogy is not quite correct, insofar as the observables obtained with the help of an intrinsic coordinate system look like ordinary field variables and are, as such, functions of all four intrinsic coordinates. Their time derivatives can be obtained in a particular manner: We must determine that linear combination of the original field variable (observable) with all the constraints whose Poisson bracket with all constraints vanishes. The time derivatives of the original variable is then equal to the partial derivative of the new equal-valued function of the dynamical variables and the coordinates with respect to $x^0$. If $A$ was the original variable, and if $A^*$ is the new variable that commutes with all the constraints, then we have, as our new "dynamical" law,

$$\dot{A}=\dot{A}^*\equiv\partial A^*/\partial x^0, \quad A^*=A. \tag{8}$$

The symbol on the right-hand side denotes partial differentiation with respect to $x^0$, to the extent that it appears as an explicit argument of $A^*$. Analogous relations hold, incidentally, concerning derivatives with respect to the spatial coordinates.

Nevertheless, there is one part of the criticisms that deserves a more thorough exploration. It would certainly be a great help for an intuitive grasp of Einstein's theory of gravitation if we could divide all the coordinate transformations into two large classes: those that correspond to the gauge transformations of electrodynamics and which have no effect on the properties of a bounded gravitational field as viewed from afar, and those coordinate transformations that change the frame of reference of an outside observer. We might call the latter Lorentz transformations. If such a separation of the whole group of coordinate transformations were possible, then we should be able to construct quantities, geometric objects, which are invariant with respect to the gauge-like transformations but which are covariants with respect to Lorentz transformations.

Such a possibility exists if, and only if, the gauge-like transformations form an invariant subgroup of all our coordinate transformations. This formal requirement is equivalent to the requirement that the Lorentz transform of a gauge-like transformation is again a gauge-like transformation. If the subset of gauge-like transformations forms a group, and moreover an invariant subgroup of the group of all coordinate transformations, or of all coordinate transformations that maintain an asymptotic behavior of the fields at spatial infinity, then there is the possibility that we can construct quantities that are "gauge-invariant" but not Lorentz-invariant and, more particularly, not constants of the motion. If not, such quantities assuredly do not exist.

It is well known that the unrestricted group of coordinate transformations possesses no invariant subgroup. The situation is not quite as clear-cut with respect to the restricted group of coordinate transformations that maintain certain asymptotic boundary conditions. It is to be hoped that this question can be cleared up with not too much difficulty. Needless to say, this question, whether we can construct general relativity in the image of a Lorentz-covariant conventional field theory, is bound up with the question of a gauge-invariant energy concept. At present, opinions on that score are divided; I believe that if the invariant gauge-like group exists, we shall also be able to construct the Lorentz-covariant energy-momentum free-vector; otherwise I am very pessimistic.

## CRITIQUE OF GENERAL COVARIANCE

In closing, we shall return once more to the question of general covariance. It appears as if general relativity contained within itself the seeds of its own conceptual destruction, because we can construct "preferred" coordinate systems. These preferred coordinate systems are not "flat," but they are determined by the intrinsic conditions of the physical situation. It is easy to point out in which respects intrinsic coordinates differ from, say, inertial coordinates. Whereas the latter can be determined experimentally by the observation of the trajectories of force-free bodies, intrinsic coordinates can be determined only by much more elaborate experiments; they depend, at least, on the inhomogeneities of the ambient gravitational fields, that is to say, on derivatives two orders higher than Cartesian coordinates. It may be that the principle of covariance tells us no more than that in fixing an absolute frame of reference we must go to much more elaborate experiments than the reading of rigid scales and stable clocks. I consider that at this time that question is open. It is clearly one of the fundamental questions that affects our whole understanding of the properties of space and time.