Academic Press, 1962), pp. 69–115. The relevant pages are 110–115. On the general problem of the alleged superiority of one method of formalizing general relativity over another see the remarks of J. Stachel in his review of the O'Raifeartaigh volume in which the Ehler, Pirani, and Schild article appears—*Science* 180 (1973): 292–293.

34. Hawking and Ellis, *Large Scale Structure*, p. 192, pp. 196–197.
35. Woodhouse, "Differentiable and Causal Structures," pp. 496–497.
36. David Malament has shown (Ph.D. Thesis, Rockefeller University and his forthcoming "The Class of Continuous Timelike Curves Determines the Topology of Spacetime") that we need assume only past- and future-distinguishingness in order to be sure that the manifold topology is at least implicitly causally definable. Of course if the space-time is not strongly causal, the Alexandroff topology and the manifold topology will still not coincide. (For relevant definitions see Hawking and Ellis, *Large Scale Structure*, p. 192.) Malament has also shown that the class of continuous causal curves "defines" the topology in the sense that given two space-times equipped with the usual manifold topologies, if there is a one-to-one function between them taking continuous causal curves into continuous causal curves, this function will also be a homeomorphism. This result holds generally in space-times. Again we see that while causal connectibility will not, in general, "define" the topology in even the weakest possible sense, the notion of continuity along causal paths will come closer to doing the job. But even in this case much caution is called for if one is to decide in just what senses even the class of continuous causal curves does and does not "define" the topology. For more on these matters see L. Sklar, "What Might be Right About the Causal Theory of Time," forthcoming in *Synthèse*.

Geometry and Observables

These remarks are intentionally brief and semioutline in form. The opinions expressed are expected to be controversial, at least in part.

1. The Concept of Observable

The notion of observable is to be understood within both classical (i.e., nonquantum) and quantum physical theories. As long as the dynamical laws obey a strict and formal Cauchy principle—i.e., as long as Cauchy data at one time are sufficient to predict (or retrodict) unequivocally (expectation) values of all dynamical variables—the notions of dynamical variable and observable are coincident. It is assumed, of course, that within the conceptual framework of the theory, predicted quantities are in principle observable as well.

The data given usually provide information concerning not only the overall (hypothesized) physical situation, but also the (Galilean or Lorentz) frame of reference. By implication, instruments used for observation can be placed within this frame. This remark is nontrivial in that it makes sense only for invariance groups that are Lie groups (finite-dimensional groups) or, more generally, groups that permit the fixation of the frame of reference once and for all by means of information provided on a Cauchy surface (at one time). Groups such as the Bondi-Metzner-Sachs group, although not Lie groups, are included.

With respect to so-called gauge groups, the situation is fundamentally different. For the purpose of this discussion I define a gauge group as a group that involves arbitrary functions of the time (ordinarily as well as of the spatial) coordinates, so that no Cauchy data given at one time provide sufficient information to fix the dynamical variables at any other time a finite distance away. For instance, with ordinary gauge transformations of the second kind, the value of an electromagnetic potential off the Cauchy

hypersurface is in principle unpredictable, although the value of a field strength is not, nor is the value of an integral of the form $\oint \mathbf{A} \cdot \mathbf{dx}$ (Bohm-Aharonov).

If some of the dynamical variables are not predictable from Cauchy data, one might conclude that such a theory is noncausal. This conclusion appears unpalatable because (a) some dynamical variables remain predictable (all those that are gauge-invariant), and (b) the imposition of gauge conditions, which by assumption do not modify the physical characteristics of the theory, render it formally causal with respect to all dynamical variables. A way out, and the one that I have adopted, is to say that in a theory with a gauge group no Cauchy data fix the frame of reference (the gauge frame), but that in those theories that we are concerned with, Cauchy data do fix the physical situation. This formulation implies that only gauge-invariant dynamical variables are physically significant; the inference is that only gauge-invariant quantities are susceptible to observation and measurement by physical instruments. This conclusion is warranted if all physical interactions, including those with physical instruments, are necessarily gauge-invariant. Admittedly, this point is susceptible to further exploration and discussion; it will be adopted for what is to follow.

If the dynamical laws of a gauge-invariant theory are to be obtained from an action principle, then the action itself must be gauge-invariant, at least with respect to gauge transformations confined to the interior of the domain that supports the action functional. If the dynamics is local, i.e., if the action functional is an integral over a local integrand, then the generators of infinitesimal gauge transformations vanish (weakly, not identically, but modulo the dynamical laws) or equal (again weakly) exact divergences. This is a consequence of Noether's theorem, which states that the generator of an infinitesimal invariant canonical transformation is a constant of the motion.

Whether or not the gauge group is Abelian, the generators of its infinitesimal elements must be first-class in the sense of Dirac, just so as to mirror the assumed group properties. That is to say, in its Hamiltonian version a gauge-invariant theory involves first-class constraints, which generate the infinitesimal gauge group. These first-class constraints will appear in the Hamiltonian, with arbitrary coefficients, which formally reflect the lack of uniqueness in the prediction of variables that are not gauge-invariant, in spite of the fact that in their canonical version the

dynamical laws relate Cauchy data (the canonical variables) on one Cauchy hypersurface to those on another Cauchy hypersurface once a Hamiltonian has been fixed.

In a quantum theory one may visualize a linear vector space that permits the formulation of state vectors corresponding to both physical and nonphysical states. If the constraints are considered to be operators satisfying commutation relations akin to those of the infinitesimal gauge group, then those state vectors satisfying the constraints form a subspace. Only gauge-invariant variables map the subspace of physical states on itself. If a Hilbert metric, and hence self-adjointness and expectation values, are defined only on that subspace, and for gauge-invariant variables, then the resulting physical implications of that theory do not depend on the accident of a particular choice of (non-gauge-invariant) dynamical variables, nor will the physical statements of that formalism be affected by the adoption of gauge conditions. For all these reasons I reserve the term *observables* for gauge-invariant dynamical variables.

2. Application to Coordinate Invariance

Electrodynamics and Yang-Mills type theories are prime examples of physical theories whose complete invariance group consists of gauge and Poincaré transformations. Uniformly, the gauge group is an invariant subgroup, and the Poincaré group is the corresponding factor group. Hence there exist variables whose transformation properties are a faithful representation of the factor group only: precisely the gauge-invariant variables. If the gauge group is the group of curvilinear coordinate transformations (or of all diffeomorphic mappings of space-time on itself), there exists no such homomorphism, with the exception of theories that admit as physically meaningful solutions only those satisfying stated conditions of asymptotic flatness at infinity. In those latter theories groups resembling the BMS-group assume the role of factor group, with the details depending on the precise statement of the condition of asymptotic flatness—cf. treatments by R. K. Sachs, R. Geroch, R. Penrose, and others.

What makes the group of diffeomorphisms peculiar is that the mapping of one Cauchy hypersurface on another is not separable from the other gauge transformations, and hence the Hamiltonian of any general relativistic theory is necessarily a linear combination of gauge constraints. Dirac's treatment of Einstein's theory in the fifties is but a specific example of a very general state of affairs. If, for instance, one were to cast the

Brans-Dicke theory into canonical form, there would again be four first-class constraints at each point of a spacelike Cauchy hypersurface satisfying commutation relations isomorphic to Dirac's.

If in such a formalism the definition of observable introduced in the preceding section is adopted, then all observables are constants of the motion, i.e., their values are the same on all conceivable Cauchy surfaces. This result has been dubbed "the frozen formalism". Its adoption is unpalatable to many, as it appears to eliminate from the formalism all semblance of dynamical development.

In its defense I would make two points. First, problems in ordinary mechanics can be restated in terms of a frozen formalism. One has only to parametrize the theory, making the time variable the $(n + 1)$st configuration coordinate, with its canonical conjugate, p_{n+1}, being equal to $-H$. $H + p_{n+1} = 0$ is indeed the generating Hamiltonian constraint, and all the observables are constants of the motion. The transformation to these new canonical coordinates is well known to be generated by any solution of the Hamilton-Jacobi equation, and no one has taken offense at the Hamilton-Jacobi theory as a normal part of classical mechanics.

But this argument brings me to the second point, which is far more subtle. Even in classical mechanics the transformation generated by $S(q, P)$ is generally not global; in fact the requisite number of constants of the motion usually does not exist. By the same token, the observables of a general relativistic theory are probably not defined globally. But they are almost certainly not locally definable variables either, unless they are defined as coincidences, as by Bergmann and Komar, or in terms of asymptotic quantities (Bondi's news functions, Newman-Penrose constants) in theories admitting such quantities. It seems to me that there are conceptually incompletely understood problems here, whose analysis might teach us something fundamental concerning the nature of general relativistic theories.

In recent years B. DeWitt and G. Smith have published attempts to analyze the observability of variables from a nuts-and-bolts point of view, in terms of at least conceptually possible instrumental procedures.

3. Nonmetric Fields

John Stachel has asked me about the notion of observables in a theory with mixed variables, such as occur in Einstein-Maxwell theory. My response is partly implied by the fact that in the preceding section I have not confined myself to any particular kind of dynamical variables, and certainly not to a pure metric.

Suppose we deal with a theory that is orthodox Einstein 1916 but contains a number of additional field variables describing "matter." The Lagrangian of such a theory will consist, additively, of the standard Einstein term and one or several terms introducing the additional variables. The Brans-Dicke theory, including the electromagnetic field, may be put into this form. Presumably the number of independent constraints does not depend on the number of field variables introduced but only on the structure of the gauge group. If the gauge group consists of space-time diffeomorphisms plus electrodynamic gauge transformations, for instance, there will be five constraints, and their commutation relations, or Poisson brackets, will be predictable without reference to the details of the Lagrangian.

Presumably it is possible to construct observables only from the metric, for instance by the Komar-Bergmann method of intrinsic coordinates. Once this has been done, the remaining field variables can be considered observables in that intrinsic coordinate system. This approach is anything but aesthetic, or even practical, and in my opinion will serve at best only as an example guaranteeing existence. At any rate, all fields in addition to the metric will give rise to an equal number of additional observables.

4. Quantization of the Metric Field

This topic too has been suggested by John Stachel. Surely the components of the metric tensor play a dual role. They form the metric backbone required for the introduction of all other fields, but they are dynamical variables in their own right, and hence ought to be quantized in a full quantum theory of the physical universe. I am aware of the fact that C. Møller has pointed out that quantization of the metric field is not required *logically*. He has been supported in this view by the late L. Rosenfeld, who has thought about the quantization of the metric field earlier and longer than anyone else (his first paper on the subject known to me is dated 1930). I assume, however, that there are strong intuitive grounds for attempting its quantization; after all, gravitation is a physical field like all others that we know.

It seems to me that a quantized metric field should not be thought of as a local quantum field defined on a space-time whose world-points possess individual and classical identity. To me the physical universe is a function

space defined on a four-dimensional manifold; what has physical significance is the quotient space of the constraint hypersurface within this function space over the mappings associated with the full gauge group of the theory. Even on a three-dimensional Cauchy hypersurface it appears risky to think primarily of a diagonalized configuration space (i.e., of a sharply defined three-dimensional metric g_{mn} on a three-dimensional spacelike hypersurface). Although the constraints restrict to some extent the range of the canonically conjugate variables, their uncertainty is unbounded, sufficiently so that the assumed sharpness of the 3-metric does not propagate at all.

Perhaps it is irrelevant whether we think of well-defined world-points with a fuzzy light cone, or conversely, of a sharp light cone, with considerable uncertainty as to which world-points lie on it. Most likely, both of these viewpoints are too naïve. Suppose we attempt a physical measurement, by means of instruments that had better not intrude too crudely on the physical situation, lest their large masses and stresses (if they are to contain any rigid components) modify the gravitational field far beyond the minimal effects required by the uncertainty relations. In elaborating what such an instrument measures we must discuss in detail not only which components of the fields are to be observed, but also in which space-time region these observations are to take place.

Perhaps it is just as well if I conclude my introductory remarks on this uncertain note, with all the technical and nontechnical connotations of "uncertain" you can imagine. It is this uncertainty that makes the whole field of quantum gravitation attractive to me.

The Curvature of Physical Space

If one were seriously to entertain, even in a highly programmatic fashion, the thesis "there is nothing in the world except empty curved space. Matter, charge, electromagnetism, and other fields are only manifestations of the bending of space," [1] it would seem highly germane to examine the nature of this curvature, which is to serve as "a kind of magic building material out of which everything in the physical world is made." [2] Such an examination has been carried out in depth by Adolf Grünbaum in "General Relativity, Geometrodynamics, and Ontology," a chapter that appears for the first time in the new edition of his *Philosophical Problems of Space and Time*. [3] The present discussion is intended primarily as an addendum to that chapter [4]—although, I should hasten to add, not necessarily one that he would endorse.

1. Metrical Amorphousness

The question I shall be addressing can be phrased, "Does physical space possess intrinsic curvature?" This way of putting it is liable to serious misunderstanding on account of the term "intrinsic," for it would be natural to call the Gaussian curvature of a surface "intrinsic" because, as Gauss showed in his *theorema egregium*, it can be defined on the basis of the metric of the surface itself, without reference to any kind of embedding space. The mean curvature, in contrast, is not intrinsic in this sense. It is entirely uncontroversial to state that, *in this sense*, any Riemannian space—not just a two-dimensional surface—possesses an intrinsic curvature (possibly identically zero) which is given by the type (0, 4) covariant