Feynman Quantization of General Relativity*†

CHARLES W. MISNER

Palmer Physical Laboratory, Princeton University, Princeton, New Jersey

1. STRUCTURE OF QUANTUM THEORY OF GENERAL RELATIVITY

THIS paper reports the beginnings of the quantum theory of general relativity based on the Feynman[1,2] integral or "sum over histories." We have not seen a way to use the Feynman integral to solve immediately all the principal problems. We have to study the theory one piece at a time and to set each fragment in place when we are able to understand it. In this sort of approach we need not follow any logical order, but may study the easy parts first and hope to fill in the rest later. However, some over-all picture of what the completed puzzle may look like is necessary in order to recognize the pieces.‡

This paper describes in broad outline the principal features of a quantum theory of general relativity and fills in a few details. If rigor could be supplied we would have a theory rather than an approach to one.

Two fragmentary but concrete contributions to the theory based on the Feynman integral are made. (1) We formulate an *H principle* (Sec. 4), which specifies the relative weights to be given to different values of the fields in the Feynman integral. This principle—applied to the metrics of general relativity (Sec. 5)—leads directly to an operator form of the field equations (Sec. 6). (2) We make (Sec. 7) a partial evaluation of the Feynman propagator. In consequence we are able to write down immediately the state of the field on any of a rather wide class of hypersurfaces when we have specified the state on one hypersurface. This evaluation is possible because the answer is trivial. The state is essentially identical on all hypersurfaces of the same class. In other words, we prove the important result that *the Hamiltonian operator is zero.* This situation is peculiar to a theory in which the metric is quantized.

2. ALTERNATIVE APPROACHES TO THE QUANTIZATION OF GENERAL RELATIVITY

Four approaches have been suggested to discover the content of the quantum theory of general relativity: (1) One considers an ideal Lorentz space, and treats the departures of the actual metric from this flat condition as small—that is, one linearizes the gravitational field equations, applies standard methods of field quantization,[3,4] and then attempts to modify this basically linear theory by reinstating the nonlinear terms.[5,6] All other approaches consider the field equations of general relativity in their fully covariant but nonlinear form. (2) The canonical formalism endeavors to investigate the noncommutative algebra of the operators of the theory in general and the Hamiltonian operator in particular, with a view to finding ultimately in this way eigenvalues and transition probabilities.[7–9] One first sets up an appropriate algebraic structure on the field variables by defining Poisson or Dirac[10] brackets. Here "appropriate" is an abbreviation for two conditions: (a) the algebraic formulation of the classical theory is equivalent to the Lagrangian formulation, and (b) the algebraic structure admits a representation by linear operators in Hilbert space where the brackets are represented by commutators. The problem thus defined has next to be *solved* by finding a particular operator representation.

The other two approaches deal with propagators, $(f_2\sigma_2|f_1\sigma_1)$, that are natural generalizations of the propagator,

$$\langle x_2t_2|x_1t_1\rangle=[m/2\pi i(t_2-t_1)]^{\frac{1}{2}} \times \exp im(x_2-x_1)^2/2(t_2-t_1) \quad (1)$$

of a simple one particle problem.

The hypersurfaces, σ_2 and σ_1, take the place of the time variables, t_2 and t_1; and the field configurations on these surfaces, f_2 and f_1, take the place of the position coordinates, x_2 and x_1. (3) The action principle

---

* A part of this work (Sec. 5) was done while the author held a Charlotte Elizabeth Proctor Fellowship (1954–1955) which he gratefully acknowledges.

† Based in part on a thesis submitted to Princeton University, May 1957, in partial fulfillment of the requirements for Ph.D.

[1] R. P. Feynman, thesis, Princeton University, 1942.

[2] R. P. Feynman, Revs. Modern Phys. **20**, 367 (1948).

‡ *Note added in proof.*—Two other papers meant to be read in association with this paper, C. W. Misner and J. A. Wheeler, "Classical Physics as Geometry," and J. A. Wheeler, "On the Nature of Quantum Geometrodynamics," were completed too late for publication in this issue and have been submitted for publication elsewhere.

[3] L. Rosenfeld, Ann. Physik **5**, 113 (1930).

[4] W. Pauli and M. Fierz, Helv. Phys. Acta **12**, 297 (1939).

[5] A summary of such linearized quantum analyses has been given by F. Belinfante, Revista Mex. Fis. 4, 192 (1955).

[6] B. E. Laurent, Nuovo cimento (10) 4, 1445 (1956).

[7] This approach has been intensively studied by P. Bergmann Revs. Modern Phys. **29**, 352 (1957), Helv. Phys. Acta., Suppl. IV, 79–97 (1956).

[8] Belinfante, Caplan, and Kennedy, Revs. Modern Phys. **29**, 518 (1957).

[9] B. S. DeWitt, Revs. Modern Phys. **29**, 377 (1957).

[10] P. A. M. Dirac, Can. J. Math. **2**, 129 (1950); **3**, 1 (1951).

of Schwinger[11] deals with the infinitesimal changes in this propagator,

$$\delta\langle f_2\sigma_2|f_1\sigma_1\rangle=(+i/\hbar)\langle f_2\sigma_2|\delta I|f_1\sigma_1\rangle, \qquad (2)$$

which come about through changes in the action, I' either by way of alterations in the position of the two surfaces, in the configurations of the fields on the surfaces, or otherwise. In distinction to the canonical formalism with its use of Hamiltonian and momenta, Schwinger's method deals with the action and expresses itself in a manifestly covariant form. This method has not been applied independently to general relativity, but is used in conjunction with the Feynman method. (4) The Feynman method focuses, not on a differential equation to be solved for the propagator, but on a formula for the solution:

$$(f_2\sigma_2|f_1\sigma_1)=N^{-1}\sum_H \exp(iI_H). \qquad (3)$$

With this expression goes a formula for a matrix element which Dyson[12] writes in the form,

$$(f_2\sigma_2|\mathcal{O}|f_1\sigma_1)=N^{-1}\sum_H \mathcal{O}_H \exp(iI_H). \qquad (4)$$

Here H indicates a field history, that is, a definite specification of the field throughout the region between the hypersurfaces σ_2 and σ_1. The quantity $\mathcal{O}$ is a functional of the field and $\mathcal{O}_H$ is its value for the particular field history H. The sum extends only over field histories whose boundary values on σ_2 and σ_1 are f_2 and f_1. The normalization factor N depends on σ_2 and σ_1, not on f_2 and f_1, and is introduced to secure the unitarity of the propagator. The (dimensionless) action is taken in general relativity to have the value

$$I_H=(c^3/16\pi G\hbar)\int R(-g)^{\frac{1}{2}}d^4x. \qquad (5)$$

The Newtonian gravitation constant, G, and the quantum of angular momentum, $\hbar$, appear in the theory, never individually, but only in a combination with the dimensions of length,

$$L^*=(\hbar G/c^3)^{\frac{1}{2}}=1.62\times10^{-33}\text{ cm}. \qquad (6)$$

Like the Schwinger method, the Feynman method deals with manifestly covariant quantities. The operators that are associated with physical quantities are *defined* by integrals of the type (4), where every quantity on on the right-hand side of the equation is a c number. Another convenient quality of the Feynman method is that it is flexible and can be applied in a limited way. An explicit representation of an algebraa dequate to express the entire classical theory—such as the canonical method demands—contains a vast and indigestible amount of information. In the Feynman theory we need not consider all these operators at once; we may begin by constructing just one of them, or some other operator which we think might be simpler than the basic operators of the canonical theory. In this way the theory may be attacked in a succession of short skirmishes rather than in a single frontal assault.

What of the problems of the three methods (1), (2), and (4) that have received some detailed consideration? We discuss here (1) the inappropriateness of the linearized treatment, (2) the problem of observables in the true theory of general relativity, and its different immediate consequences for the (a) canonical and (b) Feynman type of quantization, and (3) the work up to now on the methodology of Feynman quantization.

In a linearized version of general relativity it is easy to make calculations by standard perturbation methods. However, the characteristic length L^* is so small compared to the distances relevant in any familiar experimental context that such computations are not of much interest. If gravity is to occupy a significant place in modern physics, it can do so only by being *qualitatively* different from other fields. As soon as we assume gravity behaves qualitatively like other fields, we find that it is quantitatively insignificant.

Classical theory contains no characteristic length L^*. There we do have the possibility to confine our attention —if we wish—to the realm of weak fields. Consider the classical field equations,

$$R_{\mu\nu}-\tfrac{1}{2}g_{\mu\nu}R=0, \qquad (7)$$

and an initial space-like surface, σ. On it admit as initial condition only such a metric as leaves σ nearly flat. Moreover, permit only small normal derivatives of the metric at σ. Then the same weak field conditions will prevail some time later. Therefore the linear approximation makes sense.

Not so in quantum theory. We can form a state which makes the initial surface σ nearly flat, even on a submicroscopic level. However, uncertainty relationships will be expected to prevent our simultaneously restricting the time derivative of the metric at σ to a small value. Consequently large curvatures over small regions are to be expected. Similarly, even if there were a "vacuum state" for the gravitational field, there is no reason to expect that it would be a state where the metric is nearly flat over *small* regions—a point that has been emphasized by Professor Wheeler. In other words, a linearized version of general relativity seems precisely adapted to throwing out just those features which are physically new and interesting. We therefore pass it by.

In the full nonlinear theory one meets the fundamental issue, what are the real physical observables? We always use the word observable in a classical sense: an *observable* $\mathcal{O}$ is a functional of field histories H for

[11] J. Schwinger, Phys. Rev. **82**, 914 (1951).
[12] F. J. Dyson, *Advanced Quantum Mechanics* (Cornell University, Ithaca, 1954, mimeographed), p. 54.

which there is in classical theory a conceptual experiment whose outcome is interpreted as providing the value $\mathcal{O}_H$ of $\mathcal{O}$ for the unique field history H which existed physically at the performance of the experiment. This definition of an observable avoids using the idea of an *event*, but the simplest examples of observables require this idea. In electromagnetic theory, for instance, where the field in question is the vector potential A, an event ξ determines a functional $\mathcal{A}_\xi$ whose value for a field history A is $\mathcal{A}_\xi(A)=A(x_\xi)$. Here x_ξ is the point corresponding to the event ξ. This functional, $\mathcal{A}_\xi$, is not an observable; however, $\mathfrak{F}_\xi(A)=F(x_\xi)=\text{curl}A(x_\xi)$ defines an observable functional $\mathfrak{F}_\xi$. In electromagnetic theory it is usual to speak of $F(x)$ as an observable when one really means the observable $\mathfrak{F}_\xi$. The fancy $\mathfrak{F}_\xi$ is in fact superfluous since every physicist knows what is and is not observable (ideally) in classical electromagnetic theory. There is nothing to be gained there by explicitly distinguishing all distinguishable ideas. In general relativity it is better to make the necessary distinctions explicitly.

A fundamental distinction is that between a point x and an event ξ. The possibility of making distinctions is the fundamental fact which makes logical thinking possible. A *mathematical object* is anything which is subject to logical discussion; as such it must satisfy one axiom—that it be distinct from every other mathematical object. A *point* is a mathematical object which is an element of a set $\mathfrak{M}$ satisfying further axioms; for present purposes we require $\mathfrak{M}$ to be a four-dimensional differentiable manifold.[13] These axioms completely exhaust the meaning of the word "point" as we use it.

In studying electromagnetic theory the distinction between points and events is not interesting since we may assume that there is a unique point x_ξ in the theory corresponding to each physical event ξ. The event may be thought of as a time and place where two neutral point particles collide, or where it could have been arranged that such a collision occur. The idea of an event, then, need not involve electromagnetism, and it may be assumed that the correspondence $\xi\leftrightarrow x_\xi$ has been established on the basis of a nonelectromagnetic theory. Thus, before we begin a study of electromagnetism, we already know that a point is a good mathematical model of an event.[14] The correspondence $\xi\leftrightarrow x_\xi$ may be called *the point theory of events*, and on the basis of this theory we may use the events ξ themselves as points in discussing electromagnetic theory.

In general relativity, a different situation prevails. Before studying general relativity we have no notion of a metric, no theory of distance and time. We are therefore unable to imagine any idealized physical theory which would provide a theory of events to use in discussing general relativity. The theory of events must spring up within general relativity, not logically precede it. When a scalar such as $R(x)$ is considered as an observable in classical general relativity, the point x represents an event and is in reality a complicated functional of other events which the observer used to define his location, and of the metric throughout a region containing both those reference events and the event associated with x. If we were prepared to discuss such an observable, we would use a notation that gave fair warning of the complexity of the computation it envisages, and of the large number of quantities on which it depends. When we write $R(x)$ we mean a comparatively simple mathematical object, the value of the curvature scalar at a point x for the metric under consideration. The point is not an observable in the classical theory. Consequently, $R(x)$ is not an observable functional of the metric, nor is $g_{\mu\nu}(x)$, nor is the value of *any* scalar or tensor function at x. If ϕ is any function of points defined by the metric then we expect no corresponding operator $\phi(x)$ to be constructible in the quantum theory of general relativity. For trivial functions like $\phi=0$ we can find corresponding trivial operators. One advantage of the Feynman approach is that it allows us to work with functions like $R(x)$ without having to assume the existence of a corresponding operator.

Bergmann[15] finds that to carry through the canonical quantization it may be necessary to find the "true observables" in general relativity and use them in place of more familiar field variables. Similarly, we expect the Feynman method will generally not provide constructions for operators corresponding to classical quantities which are not observables. As discussed earlier in this section, however, a start can be made in the Feynman theory with only one or two such observables in hand, while the canonical theory seems to require that a large number of them be expressed in manageable form.

Distinct from the problem of characterizing and finding observables (which we have defined as a classical problem) is the problem of describing the *measurability* of such observables in the light of quantum theory. This problem has been discussed on the basis of quantum limitations on the measuring instruments by Osborne,[16] by Anderson,[17] and by Saleker and Wigner.[18] In electromagnetic theory a satisfactory discussion from this point of view has been given by

[13] Georges de Rham, *Variétés Différentiables* (Hermann et Cie, Paris, 1955), p. 1; H. Whitney, Ann. Math. **37**, 645 (1936).

[14] Convincing arguments have been given which indicate that a point is not a good mathematical model for a quantum mechanical "event." See E. P. Wigner, Revs. Modern Phys. **29**, 255 (1957).

[15] P. G. Bergmann, Nuovo cimento (10) **3**, 1177 (1956); see also reference 7.

[16] M. F. M. Osborne, Phys. Rev. **75**, 1579 (1949).

[17] J. L. Anderson, Revista Mex. Fis. **3**, 176 (1954).

[18] H. Saleker and E. P. Wigner (to be published). See also reference 14.

Bohr and Rosenfeld,[19] but only after an analysis of the measurability had been provided by the quantum theory of the electromagnetic field. The two approaches —one from quantum theory of the measuring apparatus, the second from the quantum theory of the field—are complementary but should be consistent. A discussion of measurability based on the quantum field theory alone is more direct, since it requires no ingenuity in discovering an optimal experimental arrangement. A desire to understand the quantum limitations on measurements of the gravitational field is therefore one reason for investigating the quantum theory of general relativity.

The Feynman method is the basis of our approach to the quantum theory of general relativity, so we now consider the techniques which have been used to express the Feyman method mathematically. Feynman's technique[1,2] in defining the "sum over histories" was based on a use of Hamilton's principal function S. In the quantum mechanics of a particle, the propagator is

$$\langle x''t''|x't'\rangle = \int \exp\left\{\frac{i}{\hbar}\int_{t'}^{t''} L dt\right\}\delta x, \tag{8}$$

where L is the Lagrangian. To define this expression Feynman first considered the case of a small time interval Δt, and set

$$\langle x'', t+\Delta t|x't\rangle = A^{-1}\exp\{iS(x'', t+\Delta t; x',t)/\hbar\}, \tag{9}$$

where S is $\int Ldt$ evaluated for the classical path with the required end points,[20] and A is a normalization factor. An iteration of the infinitesimal propagator (9) together with a limit $\Delta t \rightarrow 0$ produces the "sum over paths" of Eq. (8). The normalization factor A is fixed by requirements inherent in this limiting process; it may depend on x', x'', t, and Δt, but is independent of x' and x'' if the coefficients of the $\dot{x}$ terms in L are independent of x, and is independent of t if L is not explicitly a function of t. This S method of defining the sum over paths has been applied by Choquard[21] to situations where the potential is anharmonic. It has been extended[22] to Lagrangians where the coefficients of the velocity terms are not constants. Anderson[23] has pointed out, however, that this S technique when applied to gauge-invariant theories such as general relativity leads to an unacceptable result: $A^{-1}=0$. The explanation of this failure of the method is easy to find: In a gauge-invariant theory there are infinitely many histories which describe the same physical situation; and if one attempts to normalize after integrating, the result is an unsatisfactory $A^{-1}=0$ normalization. This difficulty has been eliminated in electromagnetic theory by Wheeler[24] and by Laurent[6] who perform an average over gauges in place of an integration. There were, however, two problems which the factor A was to solve. One problem was to determine the absolute weight to be given to a field history in the sum over histories; it was this problem which led to the $A^{-1}=0$ failure, but which is solved by using an average over gauges. The second problem, which logically precedes the first, was to determine, through the dependence of A on x, the relative weights to be given different histories. In Sec. 3 we formulate the H principle which states the relative weights which are to be assigned different field histories. It is then meaningful to speak of an average so that a way is open to consider subsequently the normalization of the "sum over histories." The averaging process, and the remaining, field independent, normalization factor, are not investigated.

More recent treatments of Feynman quantization may be called non-S methods. Most of these non-S treatments are based on the work of Davison[25] who introduced the technique of Fourier analyzing the history and then integrating over the Fourier coefficients. This technique has been applied by Burton and De Borde[26] to several mechanical systems, and is used by Matthews and Salam,[27] and by Laurent[6] in field theory. Laurent's work is the first application of the Feynman method to general relativity. None of these techniques can be applied to a quantization of the full theory of general relativity, since Fourier analysis is applicable only to linear systems.

Wheeler[24] has used an essentially non-S technique to quantize the electromagnetic field. In common with the original Feynman approach, he approximates the space-time volume V between two hypersurfaces by a grid of points $\mathfrak{N}$. A field history is then approximated by a *skeleton history* where the value of the vector potential is prescribed only at the grid points. Next the action integral is also *skeletonized* in being approximated by a sum which is defined for every skeleton history. Finally the propagator

$$(A_2\sigma_2|A_1\sigma_1) = \int \exp\left\{(-i/16\pi\hbar c)\int F_{\mu\nu}F^{\mu\nu}d^4x\right\}\delta A \tag{10}$$

is defined by integrating over the values of A at the grid points $\Pi_{x\epsilon\mathfrak{N}} d^4A_x$ and passing to a limit as the mesh

[19] N. Bohr and L. Rosenfeld, Kgl. Danske. Viden selskab, Mat.-fys. Medd. **12**, No. 8 (1933).

[20] See for instance, H. Goldstein, *Classical Mechanics* (Addison-Wesley Press, Cambridge, 1950), p. 276.

[21] Philippe Choquard, thesis (École Polytechnique Fédérale, Zurich, 1955).

[22] Cécile Morette, Phys. Rev. **81**, 848 (1951).

[23] J. L. Anderson (private communication).

[24] J. A. Wheeler, *Fields and Particles* (unpublished). Lectures given at Princeton (1954–1955 and 1956–1957) and at Leiden (1956).

[25] B. Davison, Proc. Roy. Soc. (London) **A225**, 252 (1954).

[26] W. K. Burton and A. H. De Borde, Nuovo cimento (10) **2**, 197 (1955).

[27] P. T. Matthews and A. Salam, Nuovo cimento (10) **2**, 120 (1955).

of the grid tends to zero. The entire procedure may be abbreviated by the phrase

$$\delta A = N^{-1}(V) \prod_{x \epsilon V} d^4 A(x), \qquad (11)$$

where the normalization factor N indicates both a finite normalization factor $A(\mathfrak{N})$ which appears with each skeleton integral, and the procedure of averaging over gauges.

Polkinghorne[28] has shown how operator field equations and canonical commutation rules may be derived in a non-S Feynman field theory. He does not require any details of the construction of the Feynman integral, and his work provides a justification for the use of non-S definitions of the functional integral.

In applications of the Feynman method to field theory considered by the British group, the states are treated abstractly and denoted by Dirac kets[26–28]: $|f'\sigma'\rangle$ is a state for which the field f on σ' is characterized by eigenvalues f'. The Feynman propagator then is a matrix element of the unit operator $\langle f_2\sigma_2|1|f_1\sigma_1\rangle$. Wheeler[24] has introduced into his discussions of Feynman field theory the idea of representing the states explicitly as functionals $\psi_\sigma(f)$ of the *field configurations*. A field configuration is simply a specific value of the field given for every point of the surface σ. As an example of a state functional we quote from Wheeler the ground state of the electromagnetic field:

$$\psi_0 = N^{-1} \exp\left\{ -\frac{1}{16\pi^4 \hbar c} \int\int \frac{\mathbf{H}(x)\cdot\mathbf{H}(y)}{|\mathbf{x}-\mathbf{y}|^2} d^3x d^3y \right\}. \qquad (11a)$$

Here N is a normalization factor which is independent of the field, but which depends on the way the functional integration over field configurations is defined. This idea of a state functional is a key idea in the quantum theory of general relativity, for we may begin with state functionals and then later, when we know something about the inner product of two such functionals, find the relationship between these functionals and vectors in Hilbert space. In this way we avoid assuming that there exists a $g_{\mu\nu}(x)$ operator and a normalizable physical state which is an approximate eigenfunction of $g_{\mu\nu}(x)$. In a fully gauge-invariant theory neither of these assumptions is true.

The idea of a state functional is independent of any idea of an inner product. The idea of a state vector is that of an element of Hilbert space, and is therefore stated in terms of an inner product. To distinguish the two situations we use the notation ψ_a, ψ_b, for state functionals, and $|a\rangle$, $|b\rangle$ for the corresponding state vectors when they exist. A state vector $|a\rangle$ can be defined by a state functional ψ_a for which $\int|\psi_a(f)|^2\delta f$ is finite. This functional integration is over the field configurations f which are the arguments of ψ_a (not over field histories). It is easily possible that $|a\rangle = |b\rangle$ even though $\psi_a \neq \psi_b$. In (3) and (4) the quantities $(f_2\sigma_2|f_1\sigma_1)$ and $(f_2\sigma_2|\mathcal{O}|f_1\sigma_1)$ are functionals of the two field configurations f_2 and f_1. They serve to define matrix elements through the formula

$$\langle 2|\mathcal{O}|1\rangle = \int \psi_2^*(f_2)(f_2\sigma_2|\mathcal{O}|f_1\sigma_1)\psi_1(f_1)\delta f_2 \delta f_1, \qquad (12)$$

where $|1\rangle$ and $|2\rangle$ are state vectors defined by the state functionals ψ_1 on σ_1 and ψ_2 on σ_2. We assign no meaning to the isolated symbol $|f\sigma\rangle$, and would wait until the existence of an appropriate state were made probable before writing $|f\sigma\rangle$.

Both Laurent and Wheeler employ a gauge-invariant action, $\int F_{\mu\nu}F^{\mu\nu}d^4x$, in quantizing the electromagnetic field, and suggest using $\int R(-g)^{\frac{1}{2}}d^4x$ in general relativity. This does not seem to avoid the use of a subsidiary condition, but does allow one to use a gauge invariant subsidiary condition. In electromagnetism this condition is

$$\psi_\sigma(A) = \psi_\sigma(A + \text{grad}\lambda) \text{ for all } \lambda. \qquad (13)$$

In general relativity the subsidiary condition on the state functionals is similar in form: ψ_σ must be constant over every set of field configurations which differ only by gauge transformations.

3. NATURE OF QUANTUM THEORY OF GENERAL RELATIVITY

We now sketch an over-all view of what one is to mean by the quantization of general relativity with a minimum of explanation or justification:

(1) We introduce a *4-manifold*,[13] $\mathfrak{M}$, of points. These points have no physical significance in themselves. However, they serve as handles for stating the significant mathematical relationships. They act as formless building material; no metric has yet been introduced, nor will a definite choice of a metric for $\mathfrak{M}$ ever be made.

(2) We use subsets of these points called *hypersurfaces*, σ_1, σ_2, $\cdots$. They are 3-dimensional submanifolds[29] of $\mathfrak{M}$.

(3) We introduce the concept of a metric at a point x. The metric *is* the inner product operation $(\ ,\ )_x$. A particular metric is defined by giving a way of computing the inner product $u\cdot v \equiv (u,v)_x$ of every pair of tangent vectors u, v at x. It is most commonly defined by giving its components $g_{\mu\nu}(x)$ in a particular coordinate system so that the computation of $u\cdot v$ is

$$u\cdot v = (u,v) = g_{\mu\nu}u^\mu v^\nu.$$

[28] J. C. Polkinghorne, Proc. Roy. Soc. (London) **A230**, 272 (1955).

[29] C. Chevalley, *Theory of Lie Groups* (Princeton University Press, Princeton, 1946). Definition 1, p. 85, and Def. 2, p. 80, but read differentiable for analytic.

The notation ds_x^2 for the metric at x is more common[30] than $(\ ,\ )_x$ but has the same meaning. Thus we shall always mean by ds_x^2 a metric (an inner product operation) at x. In particular, the *operation* ds^2 may be used to compute distance intervals, but is *not* to be understood as itself being a distance interval.

(4) We introduce a *field history*, ds^2 for the metric, f for a more general field, by giving at each point of $\mathfrak{M}$ a definite value $ds^2(x)$ or $f(x)$ for the field there.

(5) If values $ds^2(x)$ or $f(x)$ are specified only over a single hypersurface σ_1, we call this a *field configuration* $(ds^2)_1$ or f_1 at σ_1.

(6) For any hypersurface σ or σ_1 we consider *state functionals* ψ_σ or ψ_1 (to be distinguished later from the narrower class of *physical state functionals*). A particular state functional ψ_1 at σ_1 is defined by giving for each field configuration f_1 at σ_1 a complex number $\psi_1(f_1)$, the value of ψ_1 for the field configuration f_1.

A sample state functional ψ_σ can be defined by

$$\psi_\sigma(ds^2)=\exp\left\{-\int_\sigma {}^3R_{ij}\,{}^3R^{ij}({}^3g)^{\frac{1}{2}}d^3x\right\} \tag{14}$$

if σ is space-like for the metric ds^2,

$$\psi_\sigma(ds^2)=0 \quad \text{otherwise.}$$

Here we assume coordinates around σ chosen so that σ is defined by $x^0(x)=0$. Having written the field configuration ds^2 on σ in the form

$$ds^2(x)=g_{00}(x)\,(dx^0)^2+2g_{0i}(x)dx^0dx^i+g_{ij}(x)dx^idx^j$$

we compute the three-dimensional Ricci curvature tensor ${}^3R_{ij}$ and volume element $({}^3g)^{\frac{1}{2}}d^3x$ from the metric ${}^3g_{ij}=g_{ij}$ defined on σ by ds^2. We do not know exactly what physical situation this sample state functional describes. It gives greatest probability to the field configurations ds^2 for which ${}^3R_{ij}=0$, i.e., for which the g_{ij} metric makes σ flat (assuming that the topology of σ allows a flat metric). It is assumed that ψ_1 gives a complete description of a physical state of the system, so that from ψ_1 *plus* the machinery of the theory one can compute the expectation value of every observable. In a particle theory $\mathfrak{M}$ is one dimensional, $-\infty<t<\infty$, and the σ are zero dimensional, $t=T$, while a state functional ψ_σ is an ordinary function $\psi_T(x)$ of the field $f=x=$particle coordinate.

(7) From all conceivable *families* $\psi=\{\psi_t\}^{\infty}{}_{t=-\infty}$ of state functionals in particle theory, the Schroedinger equation, $i\hbar\partial\psi_t/\partial t=H\psi_t$, selects certain *admissible* ones called *states*. The entire admissible family of state functionals is determined when we give one member, say ψ_0. When we represent ψ by ψ_0 we are in the Heisenberg picture, while when we represent ψ by the admissible family $\{\psi_t\}$ we are in the Schroedinger picture.

A similar situation prevails in field theory where a state ψ may be represented either by a state functional ψ_0 at a chosen surface σ_0 (Heisenberg) or by a family $\{\psi_\sigma\}$ of state functionals (Schroedinger) satisfying the dynamical principle.

(8) As *dynamical principle* we use, in place of the Schroedinger equation, the Feynman principle: Given a state functional ψ_0 at σ_0, we find the other members of the family $\{\psi_\sigma\}$ for the state determined by ψ_0 through the formula

$$\psi_\sigma(f_\sigma)=\int K(f_\sigma,\sigma;\,f_0\sigma_0)\psi_0(f_0)\delta f_0. \tag{15}$$

Here $\int\cdots\delta f_0$ is a functional integration over field configurations f_0 at σ_0, and K is the Feynman propagator which may also be written $(f_\sigma\sigma|f_0\sigma_0)$.

(9) The Feynman propagator is to be defined by a functional integration over field histories:

$$K(f_2,\sigma_2;\,f_1,\sigma_1)=\int \exp\left\{\frac{i}{\hbar}I_V(f)\right\}\delta f, \tag{16}$$

where $I_V(f)=\int_V\mathfrak{L}d^4x$ is the action integral for the classical theory of the field, evaluated for the particular field history f defined over the four-dimensional V which has σ_2 and σ_1 as boundaries. The functional integration extends over all field histories f with boundary values f_2 on σ_2 and f_1 on σ_1. We write

$$\delta f=N^{-1}(V)\prod_{x\epsilon V}df_x \tag{17}$$

to suggest that the definition of the functional integration might be approached by the Wheeler-Feynman skeletonizing process (Sec. 2). Here df_x is a measure on the space of values $f(x)=f_x$ of the field at a point. The normalization factor $N(V)$ suggests the limiting processes and normalizations which go into the definition of the functional integration process.

In Sec. 4 we give two axioms (H) and (L) which guide in the choice of the measure df_x to be used in constructing δf, and in Sec. 6 we see that the operator field equations can be derived on the basis of these axioms. In Sec. 5 we define the measure $\int\cdots df_x$ to be used in quantizing general relativity, $f=ds^2$, and the operator field equations for general relativity are given explicitly in Sec. 6. The problem of defining the functional integral has been temporarily side-stepped, not solved.

(10) Gauge-invariance brings with it a subsidiary condition so that not every state functional ψ_0 at σ_0

[30] Computation methods more efficient than classical tensor analysis focus attention on the quadratic differential form $ds^2=g_{\mu\nu}dx^\mu dx^\nu$ rather than upon its components $g_{\mu\nu}$ directly, [a] [b]. This viewpoint [c] leads to new insights into differential geometry. An intrinsic definition of a differential [b] [d] identifies the differential form ds^2 with the inner product operation $(\ ,\ )$. [a] É. Cartan, *Leçons sur la Géométrie des Espaces de Riemann* (Gauthiers-Villars, Paris, 1946), second edition. [b] H. Flanders, Trans. Am. Math. Soc. **75**, 311 (1953). [c] S. S. Chern, Bull. Am. Math. Soc. **52**, 1 (1946). [d] See reference 29, p. 81.

generates a usable state $\psi=\{\psi_\sigma\}$ by (15). Those ψ_0 which do define usable states ψ, and the ψ they define, are called *physical states*. The criterion which distinguishes physical states is called the *subsidiary condition*. The sample state function of (14) satisfies the subsidiary condition in general relativity.

(11) Gauge invariance in general relativity has its origins in the *topological invariance* of the theory. Topological invariance means that, given two manifolds $\mathfrak{M}$ and $\mathfrak{M}'$ which are topologically equivalent (differentiably homeomorphic), if we first perform all constructions required by theory and then choose a particular homeomorphism (differentiable 1-1 correspondence between the points x of $\mathfrak{M}$ and x' of $\mathfrak{M}'$) between $\mathfrak{M}$ and $\mathfrak{M}'$, it will always be possible to extend it to a 1-1 correspondence between the elements of the structure over $\mathfrak{M}$ and those of the structure over $\mathfrak{M}'$. Roughly, topological invariance means that nothing in the theory changes on replacing $\mathfrak{M}$ by a topologically equivalent manifold. Thus we might let $\mathfrak{M}$ be the set of quadruples (r, z^1, z^2, z^3) where $r>0$ is a real number and each z^i is a complex number with $|z^i|=1$. On the other hand we might just as well let $\mathfrak{M}$ be the set of quadruples $\{s,\theta^1,\theta^2,\theta^3\}$ where s is any real number and each θ^i any family $\{\theta+2\pi n\}_n$ of real numbers $(n=\cdots -1, 0, 1, 2, \cdots)$. This second manifold is customarily called $R\times T^3$.

The topological invariance of general relativity has other consequences besides the subsidiary condition. If a pair of hypersurfaces σ_2, σ_1 (e.g., $s=2$, 1 in the foregoing) are topologically indistinguishable from another pair σ_4, σ_3 (e.g., $s=40, 30$), then the corresponding Feynman propagators $K(2,1)$ and $K(4,3)$ should be equivalent. This fact we consider in detail in Sec. 7 and find that it implies that the Hamiltonian operator vanishes. Consequently, among the members of a state $\psi=\{\psi_\sigma\}$ defined by a state functional ψ_0 on σ_0, every time there is no topologically invariant criterion which distinguishes the relationship of σ_1 to $\mathfrak{M}$ and σ_0 from the relationship of σ_2 to $\mathfrak{M}$ and σ_0, ψ_1 and ψ_2 will be equivalent. That is ψ_1 puts the same physics on σ_1 as ψ_2 puts on σ_2—ψ_σ is essentially independent of σ and the Heisenberg and Schroedinger pictures essentially coincide. For example, (14) defines a ψ_σ not just for a particular surface σ, but as σ is allowed to vary it defines a family $\psi=\{\psi_\sigma\}$. This family satisfies the dynamical principle (15) for the appropriate class of surfaces σ, e.g. those surfaces defined in $R\times T^3$ by $s=f_\sigma(\theta^1,\theta^2,\theta^3)$ where f_σ is any differentiable (periodic) function.

(12) The idea of Hamiltonian, Heisenberg, and Schroedinger picture, and many other concepts familiar in other field theories can only be incorporated in quantized general relativity by analogy. They do not have exactly the same meaning, for we do not have at hand any metric, or meaningful idea of time, to use in defining them. The manifold $\mathfrak{M}$ started out without a metric, and never gets one. We cannot, for example, speak of space-like surfaces, but must replace this idea with that of a *space-like state functional*: a ψ_σ which satisfies $\psi_\sigma(ds^2)=0$ for every field configuration ds^2 which does not make σ space-like.

The dynamical principle (15) was likewise written by analogy, but it turns out to be very undynamic (ψ_σ independent of σ)—at least we hope it would be undynamic in this sense if the constructions could be carried out. With no metric on $\mathfrak{M}$ there is little physical significance to a choice of σ so probability amplitudes ought not to depend on this choice. The true dynamics is the dependence of observables on time, and therefore takes place at a point in the theory where there is a metric with which a meaningful idea of time can be stated. The true dynamics is an *inner dynamics* which is to be found within the Feynman propagator, rather than in its external action on state functionals. We give one example: The volume $V(\sigma,ds^2)=\int_\sigma({}^3g)^{\frac{1}{2}}d^3x$ is an observable (gauge invariant) functional of σ and the field configuration ds^2 on σ. For a state ψ [such as the one defined by the state functional ψ_0 of Eq. (14) with σ_0 the surface $s=0$ in $R\times T^3$] its expectation value is

$$\langle V(\sigma)\rangle_\psi=\int\psi_\sigma^*(f)V(\sigma,f)\psi_\sigma(f)\delta f, \qquad (18)$$

where the functional integration extends over all field configurations $f=ds^2$ at σ. Since the Hamiltonian vanishes, $\delta\langle V(\sigma)\rangle_\psi/\delta\sigma(x)=0$. This says not that $\langle V\rangle_\psi$ is a constant of motion, but that it is topologically invariant. To define the time derivative of V we need a field history, not just a field configuration. If ds^2 is given everywhere over V between σ_2 and σ_1 making σ_1 space-like, then we can define a function $t(x)$ as the length of that geodesic from x to σ_1 which is normal to σ_1. Then $t(x)$ is well defined for all x sufficiently near σ_1. Next define $\bar{V}(\sigma_1,ds^2,t)=V(\sigma_t,ds_t^2)$ where σ_T is the hypersurface $t(x)=T$, and ds_T^2 is the field configuration induced on σ_T by the field history ds^2. Then we have

$$\langle\dot{V}\rangle_\psi=\int\psi_2^*(f_2)\left[\frac{d}{dt}\bar{V}(\sigma_1,f,t)\right]_{t=0} \times\exp\left\{\frac{i}{\hbar}I_V(f)\right\}\psi_1(f_1)\delta f, \qquad (19)$$

where ψ_2, ψ_1 are the members of $\psi=\{\psi_\sigma\}$ corresponding to σ_2 and σ_1, and the functional integration extends over all field histories $f=ds^2$ defined over V, including an integration over the boundary values f_n on σ_n. This is an admittedly complicated definition, for t depends on ds^2. We do not know how to evaluate such formulas, nor how to give them a mathematically precise meaning. This example is intended merely as an indication of how dynamical questions may be formulated within the theory.

4. *H* PRINCIPLE

We view the Feynman method as a Huygen's principle[31]: In forming

$$\int \exp\left\{\frac{i}{\hbar}\text{ (action)}\right\} d\text{ (field histories)} \qquad (20)$$

the exponential by interference effects should select for us the fields that propagate significantly. In the integration all fields should be weighed equally, otherwise our arbitrary choice of what weight to put on each history in d (field histories) would compete with the exponential in selecting the significant fields. Of course we have yet to say what "weighed equally" means. The phrase suggests immediately, if we look about for suitable mathematical expression of it, the idea of invariant integration over a group.

Can we discover an invariant group integration in a simple problem? Consider a scalar field $\phi(x)$. An obvious group is the set of fields under addition. The integral over values of the field $\phi(x)=\phi_x$ at a point actually used here, $\int\cdots d\phi_x$, to generate the functional integration $\int\cdots\delta\phi$ is in fact the unique invariant integral over the additive group of real numbers ϕ_x. To construct $\int\cdots\delta\phi$ from $\int\cdots d\phi_x$, normalization factors $A(\mathfrak{N})$ appear depending on the grid of points $\mathfrak{N}$ used in the approximation of $\int\cdots\delta\phi$ by a multiple integral. The normalization factors are, however, independent of ϕ. In the Feynman propagator

$$K(\phi_2,\sigma_2;\phi_1,\sigma_1)=\int \exp\left\{\frac{i}{\hbar}I_V(\phi)\right\}\delta\phi, \qquad (16')$$

where V is the volume with σ_2 and σ_1 as boundaries and

$$I_V(\phi)=\int_V \mathfrak{L}d^4x$$

is the action integral, the functional integration $\int\cdots\delta\phi$ extends over all fields ϕ over V which reduce to ϕ_2 on σ_2 and ϕ_1 on σ_1. This set of fields, $F(V;\phi_2,\phi_1)$, does not form a group; however the fields ϕ over V which are zero on σ_2 and σ_1 do form a group, call it $G(V,bdyV)$, and the functional integration is invariant under this group. By this we mean

$$\int \mathcal{T}(\phi-\chi)\delta\phi=\int\mathcal{T}(\phi)\delta\phi \qquad (21)$$

for every functional $\mathcal{T}$ of ϕ [such as the exponential of the action in Eq. (16′)] when χ is in $G(V,bdyV)$. This invariance condition can be stated

$$\delta(\phi+\chi)=\delta\phi \quad \text{for} \quad \chi\epsilon G(V,bdyV)$$

as is seen by replacing ϕ by $\chi+\phi$ in the left-hand side of (21). The functional integral is thus invariant under a group $G(V,bdyV)$, even though it is an integral over $F(V;\phi_2,\phi_1)$. The relationship between G and F can be stated by saying that G is a transitive group of permutations[32] of F. That $\chi\epsilon G$ is a permutation of F means that the transformation χ of F defined by $\phi\rightarrow\phi+\chi$ is a 1-1 transformation of F onto F, while transitive means that if ϕ_0 is one field in F, every other $\phi\epsilon F$ is of the form $\phi=\phi_0+\chi$ for some $\chi\epsilon G$.

We now state these ideas in a still more abstract form in order to isolate certain aspects of the simple problem (quantizing a scalar field) and state them in a way which shows no traces of the problem in which they were recognized. Then the statements may be considered a discussion of the simpler aspects of more difficult problems. Were we to look immediately at general relativity the pertinent simplicities might get confused with difficult questions which we are not trying to answer immediately. Thus in the theory of an unspecified field f we will have a Feynman propagator

$$K(f_2,\sigma_2;f_1\sigma_1)=\int \exp\left\{\frac{i}{\hbar}I_V(f)\right\}\delta f \qquad (22)$$

defined by an action integral

$$I_V(f)=\int_V \mathfrak{L}d^4x \qquad (23)$$

and a functional integration $\int\cdots\delta f$. The integral extends over the set $F(V;f_2,f_1)$ of fields which reduce to f_2 on σ_2 and f_1 on σ_1. We want to have a transitive group of permutations of $F(V;f_2,f_1)$; we denote this group by $G(V,bdyV)$.

Definition: If G is a transitive group of permutations of F we say

F is an *homogeneous* set and G its *homogeneity*.

If $\pi\epsilon G$ is one of the transformations, the transform of $f\epsilon F$ will be written f^π in order not to suggest that we are necessarily dealing with a linear theory. (In the preceding example $f=\phi$, $\pi=\chi$, and $f^\pi=\phi+\chi$.) Successively applying two transformations, π then ω, we get $(f^\pi)^\omega=f^{\pi\omega}$ where $\pi\omega$ is the product in G. (In the example the group was written additively and $\pi\omega=\chi_1+\chi_2$ if $\pi=\chi_1$, $\omega=\chi_2$.) The invariance of the functional integration under G is expressed by

$$\delta(f^\pi)=\delta f \qquad \text{(H)}$$

which states that the volume element at any point f^π of F is just as big as it is at f, and means that for any

[31] R. P. Feynman, reference 2, Sec. 7, p. 377.

[32] This is identical with the idea of a transitive transformation group, but this name is more common when topological considerations also enter. D. Montgomery and L. Zippin, *Topological Transformation Groups* (Interscience Publishers, Inc., New York, 1955), Sec. 1.19, 1.26, 2.12.

functional $\mathcal{T}$

$$\int \mathcal{T}(f^{\pi^{-1}})\delta f = \int \mathcal{T}(f)\delta f \tag{24}$$

for all $\pi\epsilon G$.

This *H principle* is so named because of the Huygen's principle which suggested the "weighed equally" phase, the Homogeneity[32,33] of the fields F under the group G which gives precise meaning to the phrase, and the Hurwitz integral[34] or Haar measure[33,35] which will enter into the construction of an integration satisfying this postulate.

As a matter of convenience, we note that an invariant integral

$$\mathcal{J}(\mathcal{T}) = \int \mathcal{T}(f)\delta f \tag{25}$$

over $F(V; f_2,f_1)$ can be defined (formally) in terms of an invariant integral $\int\cdots\delta\pi$ over $G(V,bdyV)$. We define $\mathcal{T}^\pi$, the transform of a functional $\mathcal{T}$ under π, by the condition that $\mathcal{T}^\pi$ have the same value at f^π as $\mathcal{T}$ did at f, i.e.

$$\mathcal{T}^\pi(f^\pi) = \mathcal{T}(f). \tag{26}$$

Then the invariance requirement (24) can be stated

$$\mathcal{J}(\mathcal{T}^\pi) = \mathcal{J}(\mathcal{T}) \quad \text{for all } \pi\epsilon G(V,bdyV). \tag{27}$$

We define $\mathcal{J}$ by

$$\mathcal{J}(\mathcal{T}) = \int \mathcal{T}(f_0^\pi)\delta\pi, \tag{28}$$

where f_0 is some fixed field.
Then

$$\mathcal{J}(\mathcal{T}^\omega) = \int \mathcal{T}^\omega(f_0^\pi)\delta\pi = \int \mathcal{T}(f_0^{\pi/\omega})\delta\pi$$

$$= \int \mathcal{T}(f_0^\pi)\delta(\pi\omega) = \mathcal{J}(\mathcal{T})$$

provided $\delta(\pi\omega)=\delta\pi$, i.e., if $\int\cdots\delta\pi$ is a right invariant integral over $G(V,bdyV)$. This definition of $\mathcal{J}=\int\cdots\delta f$ in terms of $\int\cdots\delta\pi$ appears to depend on the choice of f_0; however if $\int\cdots\delta\pi$ is left invariant, $\delta(\omega\pi)=\delta\pi$, then

$$\mathcal{J}(\mathcal{T}) = \int \mathcal{T}(f_0^\pi)\delta\pi = \int \mathcal{T}(f_0^{\omega\pi})\delta\pi.$$

Thus f_0^ω, i.e., any field in $F(V; f_2,f_1)$, could have been used in place of f_0 without changing the value of $\mathcal{J}(\mathcal{T})$.

In view of the construction of $\int\cdots\delta f$ we imagine (17), the homogeneity group of the functional integral is to be generated by an homogeneity group $G(x)$ of the point integral $\int\cdots df_x$ over the space $F(x)$ of field values $f_x=f(x)$ at x. If $\pi(x)\epsilon G(x)$ we write the image of $f(x)$ under $\pi(x)$ as $f(x)^{\pi(x)}$. Let $G(x)$ be an homogeneity of $F(x)$ for each x, and π a field with values $\pi(x)\epsilon G(x)$ and for x on $bdyV$, $\pi(x)=1$. Then we define the homogeneity $G(V,bdyV)$ of the set of fields $F(V; f_2,f_1)$ as the group of all transformations $f\rightarrow f^\pi$ where f^π is defined by

$$f^\pi(x) = f(x)^{\pi(x)} \quad \text{for each } x\epsilon V. \tag{L}$$

If π, ω are two elements of $G(V,bdyV)$ their product is defined by

$$(\pi\omega)(x) = \pi(x)\omega(x)$$

where the product on the right is taken in $G(x)$ for each $x\epsilon V$. Eq. (L) says that the π transformations are Local. Thus we have reduced the problem defining $G(V,bdyV)$ to that of defining $G(x)$. Since $\int\cdots\delta f$ can be defined formally in terms of $\int\cdots\delta\pi$, a point integral $\int\cdots d\pi_x$ over $G(x)$ can be taken as the starting point for a construction of the Feynman propagator.

5. MEASURE ON METRICS

Applying the ideas of the preceding section, we must define a transitive group of permutations $G(x)$ on the space $F(x)$ of metrics at a point. If $g_{\mu\nu}(x)$ are the components of $ds^2(x)$ we define the components $g_{\mu\nu}{}^\pi(x)$ of $ds^2(x)^{\pi(x)}$ by

$$g_{\mu\nu}{}^\pi(x) = \pi_\mu{}^\alpha(x) g_{\alpha\beta}(x)\pi_\nu{}^\beta(x), \tag{29}$$

where $\pi(x)$ is a tensor satisfying

$$\det\pi = \frac{1}{4!}\delta_{\alpha\beta\gamma\delta}{}^{\mu\nu\rho\sigma}\pi_\mu{}^\alpha\pi_\nu{}^\beta\pi_\rho{}^\gamma\pi_\sigma{}^\delta > 0. \tag{30}$$

The point integral[36]

$$\int\cdots d\pi_x = \int\cdots(\det\pi)^{-4}\prod_{\mu,\nu=0}^{3} d\pi_\nu{}^\mu \tag{31}$$

is a left and right invariant integral over $G(x)$, unique to within a factor independent of π. We *assume* that a functional integral $\int\cdots\delta\pi$ will be generated by $\int\cdots d\pi_x$. Since in $\delta\pi = N^{-1}(V)\prod d\pi_x$ the normalization factor is independent of π, $\int\cdots\delta\pi$ will be a left and right invariant integral over $G(V,bdyV)$, and thus give an invariant functional integral over $F(V;f_2,f_1)$ by (28). This $\int\cdots\delta f$ is then used in (22) to define the Feynman propagator of general relativity. The action integral we use in general relativity is

$$I_V(ds^2) = \frac{c^3}{16\pi G}\int_V R(-g)^{\frac{1}{2}}d^4x. \tag{32}$$

---

[33] A. Weil, *L'Intégration dans les Groupes Topologiques* (Hermann et Cie, Paris, 1940), Chap. II.

[34] E. P. Wigner, *Gruppentheorie* (Edwards Brothers, Inc., Ann Arbor, 1944), pp. 103–108.

[35] P. R. Halmos, *Measure Theory* (D. Van Nostrand Company, Inc., New York, 1950), pp. 254, 263.

[36] F. D. Murnaghan, *The Theory of Group Representations* (Johns Hopkins Press, Baltimore, 1938), p. 204.

To what extent is the choice of $G(x)$ arbitrary? We do not know precisely,[37] however since we have gone to all this trouble to avoid problems raised by gauge-invariance, it is worth noting that $G(x)$ coincides with the gauge group at the single point x (consider $\pi_\nu{}^\mu=\partial y^\mu/\partial x^\nu$). This criterion on the choice of $G(x)$—let $G(x)$ be the gauge group at x—also gives the right answer in electromagnetic theory and determines $G(x)$ uniquely. The choice of $G(x)$ should be independent of the Lagrangian of the theory however. It is much the same question as "Is the field f to be a scalar, vector, tensor, or what?" which must be decided in a classical theory before field equations are written for the theory. If we erroneously thought of general relativity as a theory of simply a symmetric tensor field $g_{\mu\nu}$, then the space $\bar{F}(x)$ of field values $g_{\mu\nu}(x)$ at x consists of all symmetric tensors and is a vector space under addition. The obvious choice of an homogeneity $\bar{G}(x)$ of $(x\bar{F})$ is the group of translations $g_{\mu\nu}{}^\pi=g_{\mu\nu}+\pi_{\mu\nu}$ and linear integration

$$\int\cdots\prod_{\mu\leq\nu}dg_{\mu\nu} \tag{33}$$

is unique to within a factor independent of $g_{\mu\nu}$. Using $\int R(-g)^{\frac{1}{2}}d^4x$ as the action integral in the Feynman propagator, and $\bar{G}$ as the homogeneity we could probably construct a quantum theory with the Einstein equations

$$R_{\mu\nu}-\tfrac{1}{2}g_{\mu\nu}R=0 \tag{34}$$

as its classical limit. However, neither the action integral, nor the group $\bar{G}$, distinguishes between the (diagonal) tensors $\hat{g}_{\mu\nu}=[-1,1,1,1]$ and $g_{\mu\nu}{}^*=[1,1,1,1]$. Both are solutions of the (34), and neither would have a preferred status in the classical limit of the quantum theory based on $\bar{G}$, so this classical theory would not be general relativity.

The invariant integral (31) over $G(x)$ is unique to within a factor independent of π. We have chosen this factor in (31) in a definite way that is independent of the choice of coordinates used to represent π by its components $\pi^\mu{}_\nu$. The invariant integral (33) over $\bar{F}(x)$ is also unique to within a factor independent of $g_{\mu\nu}$. In this case, however, had we chosen different coordinates $\bar{x}^\mu(x^\nu)$ we would have written

$$\prod_{\mu\leq\nu}d\bar{g}_{\mu\nu}=r(A)\prod_{\mu\leq\nu}dg_{\mu\nu}, \tag{35}$$

where $A=\|\partial x^\mu/\partial\bar{x}^\nu\|$ and r is some function of the matrix A. It is easily seen that r defined in this way is a representation of the general linear group. As a one-dimensional representation it is necessarily of the form[38]

$$r(A)=(\det A)^k$$

and by counting the number of factors $A_\nu{}^\mu$ which enter in (35) from the transformation law for $g_{\mu\nu}$ we see that $k=5$. This observation allows us to write down immediately the integral

$$\int\cdots d(\text{metric})_x=\int\cdots(-g)^{-\frac{5}{2}}\prod_{\mu\leq\nu}{}^{\frac{1}{2}}\,dg_{\mu\nu} \tag{36}$$

over $F(x)$ invariant under $G(x)$, where again the choice of the arbitrary $g_{\mu\nu}$-independent factor is made independently of the choice of coordinates. The range of integration in (36) is the connected region of $g_{\mu\nu}$-space containing $g_{\mu\nu}=\text{diag }(-1,\,1,\,1,\,1)$ and bounded by $\det g_{\mu\nu}=0$.

6. OPERATOR FIELD EQUATIONS

Consider the integral (22) for the propagator in the form

$$K(f_2,\sigma_2;f_1,\sigma_1)=\int\exp\{iI_V(f^\pi)/\hbar\}\delta f. \tag{37}$$

Using the invariance (H) of the integration we have inserted a transformation $\pi\epsilon G(V,bdyV)$. This integral is a functional of π which is independent of π, therefore

$$0=\frac{\delta K}{\delta\pi(x)}=\frac{i}{\hbar}\int\frac{\delta I_V(f^\pi)}{\delta\pi(x)}\exp\left\{\frac{i}{\hbar}I_V(f^\pi)\right\}\delta f. \tag{38}$$

Here the requirement (L) that the π's be local transformations of the field f was used, since π must be a field $\pi(x)$ in order to define the variational derivatives in (38). Defining the Jacobian transformation $\partial f^\pi/\partial\pi$ by

$$\delta f^\pi(x)=\frac{\partial f_x{}^\pi}{\partial\pi_x}\delta\pi(x), \tag{39}$$

Eq. (38) can be written

$$\left\langle 2\left|\frac{\delta I_V(f)}{\delta f(x)}\left(\frac{\partial f_x{}^\pi}{\partial\pi_x}\right)_{\pi=1}\right|1\right\rangle=0 \tag{QF}$$

after setting $\pi=$ the identity transformation. These are the Quantum mechanical form of the Field equations. The equations

$$\frac{\delta I_V(f)}{\delta\pi(x)}=\frac{\delta I_V(f)}{\delta f(x)}\frac{\partial f_x{}^\pi}{\partial\pi_x}=0 \tag{40}$$

[37] The problem then is: Are there two distinct Lie groups G_1, G_2, which are effective homogeneities of the coset space $F=Gl(4)/L$ considered simply as a differentiable manifold, and which define essentially different invariant integrals on F. ($L=$Lorentz group.)

[38] We thank Professor V. Bargmann for pointing out this argument.

are equivalent to the Classical Field equations

$$\frac{\delta I_V(f)}{\delta f(x)}=0 \tag{CF}$$

for the Jacobian is of maximum rank: any desired change in f^π may be produced by a suitable change in π. (G is transitive.)

In linear field theories equations (40) are identical to (CF); it follows directly from $f^\pi=f+\pi$ (linearity) that $\partial f^\pi/\partial\pi=1$.

In general relativity we have from (29) a representation of the Jacobian transformation by a matrix

$$\frac{\partial g^\pi{}_{\mu\nu}}{\partial \pi_\beta{}^\alpha}=g_{\alpha\gamma}\pi_\nu{}^\gamma\delta_\mu{}^\beta+\pi_\mu{}^\nu g_{\gamma\alpha}\delta_\nu{}^\beta \tag{41}$$

when π is the identity, $\pi_\beta{}^\alpha=\delta_\beta{}^\alpha$ and

$$\left(\frac{\partial g_{\mu\nu}{}^\pi}{\partial \pi_\beta{}^\alpha}\right)_{\pi=1}=g_{\alpha\nu}\delta_\mu{}^\beta+g^\mu{}_\alpha\delta_\nu{}^\beta. \tag{42}$$

The field equations (CF) in this coordinate system are

$$\frac{\delta I(ds^2)}{\delta g_{\mu\nu}(x)}=\frac{-c^3}{16\pi G}(R^{\mu\nu}-\tfrac{1}{2}g^{\mu\nu}R)(-g)^{\frac{1}{2}} \tag{43}$$

so (QF) reads

$$\langle 2|(R_\alpha{}^\beta-\tfrac{1}{2}\delta_\alpha{}^\beta R)(-g)^{\frac{1}{2}}|1\rangle=0. \tag{44}$$

Obvious linear combinations of these equations give

$$\langle 2|R(-g)^{\frac{1}{2}}|1\rangle=0 \tag{45}$$

and

$$\langle 2|R_\alpha{}^\beta(-g)^{\frac{1}{2}}|1\rangle=0. \tag{46}$$

The form of (44) depends only upon the invariance of the integration with respect to the transformation π, not upon the fact that we concern ourselves with the $g_{\mu\nu}$ rather than $g^{\mu\nu}$ or some other equivalent tensor constructable algebraically from $g_{\mu\nu}$. A π transformation of ds^2 defines corresponding transformations of these tensors as well, and $\delta I_V/\delta\pi(x)$ is independent of the factorization (40) used to compute it.

The content of (44) is this: in the quantum theory of general relativity there exists an operator corresponding to the classical tensor $(R_\alpha{}^\beta-\frac{1}{2}\delta_\alpha{}^\beta R)(-g)^{\frac{1}{2}}$, and this operator is the zero operator. There is in (44) *no* statement about operators corresponding to $(-g)^{\frac{1}{2}}$ nor $R_\alpha{}^\beta-\frac{1}{2}\delta_\alpha{}^\beta R$; hence the order in which the factors are written in Eq. (44) is not significant.

7. TOPOLOGICAL INVARIANCE

Conventional field theories start with a manifold possessing a metric (usually the flat Minkowski metric) and describe the way in which some other field is to be assigned to the manifold. Such theories are metrically invariant (Lorentz invariant): they are unaffected by any transformation of the manifold which leaves the metric unchanged. General relativity starts with a manifold and describes the way in which a metric is to be assigned to the manifold. In general relativity, only the differentiable structure (frequently abbreviated "topology") is fixed at the start, and the theory of general relativity is unaffected by any transformation of the manifold which leaves the topology unchanged. We do not mean that every quantity in the theory remains absolutely fixed, but that to every permissible transformation of the manifold there corresponds a transformation of all the other quantities in the theory of such a nature as to leave the physical content unchanged. A transformation of a manifold which leaves the topology unchanged is called a homeomorphism. In this section we look at the transformation of the Feynman propagator corresponding to a homeomorphism, and see that if every homeomorphism is a permissible transformation (i.e., if the theory is topologically invariant) as in the theory of general relativity, then the Feynman propagator connecting equivalent hypersurfaces is trivial.

We say two hypersurfaces σ_2 and σ_3 are *equivalent* if (1) there exists a hypersurface σ_1, a V_{21} with σ_2 and σ_1 as boundaries, and a V_{31} with σ_3 and σ_1 as boundaries, and (2) there is a homeomorphism h of the 4-manifold $\mathfrak{M}$ which maps V_{31} onto V_{21}, σ_3 onto σ_2, and which restricts to the identity map of σ_1 onto itself.

A homeomorphism gives not only a correspondence between points by $y=\mathrm{h}x$, but can be used to construct a correspondence of all fields. If ϕ is a scalar function, the corresponding function under h is written ϕ^{h} and defined by $\phi^{\mathrm{h}}(x)=\phi(\mathrm{h}x)$. The metric corresponding under h to ds^2 will be designated by $(ds^2)^{\mathrm{h}}$. If x^μ, y^μ are local coordinates around x, and around $y=\mathrm{h}x$, respectively, then h can be represented locally by differentiable functions h^μ:

$$y^\mu=\mathrm{h}^\mu(x^0,x^1,x^2,x^3). \tag{47}$$

The components $g_{\mu\nu}{}^{\mathrm{h}}(x)$ of $(ds^2)^{\mathrm{h}}(x)$ are given in terms of the components $g_{\mu\nu}$ of ds^2 by

$$g_{\mu\nu}{}^{\mathrm{h}}(x)=g_{\alpha\beta}(\mathrm{h}x)\frac{\partial \mathrm{h}^\alpha}{\partial x^\mu}\frac{\partial \mathrm{h}^\beta}{\partial x^\nu}. \tag{48}$$

In a similar way a field f^{h} corresponding under h to a field f could be defined. A state functional $\mathrm{h}\psi$ at $\mathrm{h}\sigma$ corresponding to ψ at σ is then defined by requiring

$$\mathrm{h}\psi(f)=\psi(f^{\mathrm{h}}) \tag{49}$$

for every field configuration f at $\mathrm{h}\sigma$.

Let h be a homeomorphism which establishes the equivalence of two hypersurfaces σ_2 and σ_3. A state functional ψ_1 at σ_1 (see definition of equivalent) defines a state (family of state functionals, see Sec. 3)

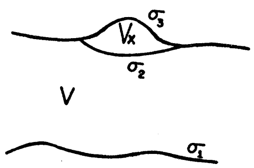

FIG. 1. The vanishing of the Hamiltonian. The hypersurface σ_3 is obtained from σ_2 by a change $\delta\sigma(x)$ which consists of sweeping σ_2 across V_x into σ_3. The corresponding change $\delta\psi_2$ in the state functional defines the Hamiltonian density operator $\mathcal{H}_{\rm op}(x)$ according to the Schwinger-Tomonaga equation

$$\delta\psi_2(f_2)=(i/\hbar)\int(\bar{f}_2|\int_{V_x}\mathcal{H}(x)d^4x|f_2)\psi_2(f_2)\delta f_2.$$

The state is defined by specifying a state functional ψ_1 on σ_1. Standing at σ_1 however, we are unable (in a topologically invariant theory) to formulate a criterion for choosing a second hypersurface σ which would permit the choice $\sigma=\sigma_2$ but which would not allow $\sigma=\sigma_3$. Thus ψ_2 and ψ_3 are equivalent, i.e. $\delta\psi_2=0$. In any topologically invariant theory we therefore must have $\mathcal{H}_{\rm op}=0$.

containing the state functionals ψ_2 at σ_2 and ψ_3 at σ_3 which are defined by ψ_1 and the Feynman propagator over V_{21} and V_{31}, respectively. In a "matrix" notation which suppresses arguments (f_n) and integrations (δf_1) we may write (compare (15))

$$\psi_2=K(\sigma_2,\sigma_1)\psi_1$$
$$\psi_3=K(\sigma_3,\sigma_1)\psi_1. \tag{50}$$

Since h is the identity on σ_1, $\mathrm{h}\sigma_1=\sigma_1$ and $\mathrm{h}\psi_1=\psi_1$. If h is a permissible transformation we must have

$$\mathrm{h}K(\sigma_3,\sigma_1)=K(\mathrm{h}\sigma_3,\mathrm{h}\sigma_1)\mathrm{h}, \tag{51}$$

or in the present case, since $\mathrm{h}\sigma_3=\sigma_2$,

$$\mathrm{h}\psi_3=\psi_2. \tag{52}$$

In summary, let $\psi=\{\psi_\beta\}$ be a state for a topologically invariant quantum theory, and let σ_2, σ_1 be equivalent under a homeomorphism h with $\mathrm{h}\sigma_2=\sigma_1$, then $\mathrm{h}\psi_2=\psi_1$. The equation $\psi_2=\psi_1$ is nonsense since ψ_2 and ψ_1 are functionals with different argument domains (the field configurations on σ_2 and σ_1, respectively). In a topologically invariant theory the only distinguished transformations which we could use to identify these two domains are the homeomorphisms. Consequently one does not misinterpret the equation $\mathrm{h}\psi_2=\psi_1$ by saying that ψ_2 is the same as ψ_1. When ψ_2 is given by an explicit formula such as (14), the equation $\mathrm{h}\psi_2=\psi_1$ says that a formula for ψ_1 may be obtained simply by replacing σ_2 by σ_1 in the formula for ψ_2. The general idea is that in a topologically invariant theory, a homeomorphism is the same as an identity transformation, modulo physically meaningless mathematical distinctions.

We have taken as the basic elements of quantum field theory the definition (16) of the Feynman propagator, the dynamical principle (15), and the formula (4) for computing matrix elements. We did not have any need for a Hamiltonian in the classical theory (in general relativity it is quite difficult to find one[7]) so we may define a Hamiltonian density operator $\mathcal{H}_{\rm op}(x)$ directly in the quantum theory. The Schwinger-Tomonaga equation,[39,40]

$$i\hbar\frac{\delta\psi_\sigma}{\delta\sigma(x)}=\mathcal{H}_{\rm op}(x)\psi_\sigma, \tag{53}$$

serves this purpose. Let a pair of hypersurfaces such as σ_2 and σ_3 in Fig. 1 differ by the boundary of a small (cellular) neighborhood V_x of a point x. A comparison of ψ_3 and ψ_2 is defined by

$$\delta\psi_2=\mathrm{h}\psi_3-\psi_2, \tag{54}$$

where h is any homeomorphism $\mathrm{h}:\sigma_3\to\sigma_2$ which reduces to the identity on the common submanifolds of σ_3 and σ_2. There is no structure available in a topologically invariant theory to make a more definite choice of h, but $\mathcal{H}_{\rm op}$ will be well defined without any further restrictions on h. Note that the definition of $\mathrm{h}\psi_3$ in terms of ψ_3 is independent of the equations of motion (15). Now suppose that both ψ_3 and ψ_2 are members of a state ψ defined by a state functional ψ_1 on σ_1, and that V is bounded by σ_2 and σ_1 so $V+V_x$ is bounded by σ_3 and σ_1. We may in this case interpret $\delta\psi_2$ as the change in the state functional produced by a change $\delta\sigma(x)$ displacing the surface σ_2 across V_x. Then since V_x is assumed to be a cell (a homeomorph of the unit cube E^4 in Euclidean 4-space) the homeomorphism $\mathrm{h}:\sigma_3\to\sigma_2$ can be extended to a homeomorphism $\mathrm{h}:V+V_x\to V$ which is the identity on σ_1. Therefore σ_3 and σ_2 are equivalent, $\mathrm{h}\psi_3=\psi_2$, $\delta\psi_2=0$, and consequently $\mathcal{H}_{\rm op}=0$. *In any topologically invariant theory, the Hamiltonian operator vanishes.* For the conclusion $\mathcal{H}_{\rm op}=0$ to hold it is necessary that the quantum theory, not just the classical theory, be topologically invariant.

Because of different methods of definition, our $\mathcal{H}_{\rm op}$ Hamiltonian is not necessarily the operator corresponding to the Hamiltonian defined in classical theory. However Professor J. L. Anderson at the Chapel Hill conference voiced suspicions that the classical Hamiltonian in general relativity would be zero.

Is the vanishing of the Hamiltonian embarrassing? Not at all. As defined here the Hamiltonian is, so to speak, the generator of infinitesimal homeomorphisms and is totally unrelated to the energy of the system or to the dynamic changes taking place in the observables. Its vanishing is merely the local statement of the topological invariance of the theory, and is therefore desirable. We postulate it as one of the characteristic properties of the Feynman propagator in quantizing general relativity or any other topologically invariant field theory.

SUMMARY OF CONCRETE RESULTS

It may reasonably be expected that whenever any construction of the Feynman integral is possible,

[39] J. Schwinger, Phys. Rev. **74**, 1449 (1948).
[40] S. Tomonaga, Progr. Theort. Phys. **1**, 34 (1946).

this construction can be done in conformity with the H principle. The H principle characterizes a property of the Feynman integral sufficient to guarantee that an operator form of the field equations will hold. The choice of the homogeneity of the field determines uniquely the relationship between the form of the operator field equations and the form of the classical field equations. The homogeneity of the metric field of general relativity has been defined, and the operator form of the Einstein field equations has been given. In a topologically invariant quantum field theory, such as quantized general relativity, the (Schroedinger) state functionals on equivalent hypersurfaces are equivalent, so that the Hamiltonian vanishes.

ACKNOWLEDGMENTS

Special appreciation is expressed to the sponsors of the International Conference on the Role of Gravitation in Physics held at Chapel Hill, North Carolina, January 18–22, 1957, and to Professor and Mrs. Bryce DeWitt for the stimulus which this conference provided. Among the many participants whose remarks helped to clarify the Feynman quantum theory of general relativity, we are particularly indebted to Professor J. L. Anderson, Dr. S. Deser, and Dr. B. Laurent. Above all, we thank Professor J. A. Wheeler. He originally suggested the problem of formulating the Feynman quantization of general relativity. We regret that our approach to this problem has not the conciseness of his solution to it, which was simply to write

$$\int \exp\{(i/\hbar)(\text{Einstein action})\}d\ (\text{field histories}).$$

He provided constant encouragement in pursuing the investigation; and his provoking ideas on the nature of the theory of general relativity led us to adopt the intrinsic or coordinate-free point of view which has been essential in the development of this theory. His help and advice in writing this paper are reflected in those sections which are most clearly presented.

REVIEWS OF MODERN PHYSICS VOLUME 29, NUMBER 3 JULY, 1957

Reality of the Cylindrical Gravitational Waves of Einstein and Rosen

JOSEPH WEBER, *Lorentz Institute, University of Leiden, Leiden, Netherlands, and University of Maryland, College Park, Maryland*

AND

JOHN A. WHEELER, *Lorentz Institute, University of Leiden, Leiden, Netherlands, and Palmer Physical Laboratory, Princeton University, Princeton, New Jersey*

1. INTRODUCTION

QUESTIONS have been raised whether gravitational radiation has any well-defined existence.[1] Supporting this skeptical position, Rosen has investigated further[2] the cylindrical gravitational waves first considered by him and Einstein[3] as an outgrowth of a suggestion by H. P. Robertson. A monochromatic wave, or a pulse, of cylindrical symmetry, moves inward in matter-free space, implodes on the axis, and moves out again. This is the only problem of gravitational radiation where one has an *accurate* solution of the field equations of general relativity. The problem is special enough not to illustrate all features of gravitational radiation. On the other hand, all correct general statements about gravitational radiation must obviously be compatible with this problem. This problem therefore occupies a special position in the theory of gravitational radiation.

Rosen finds an unexpected result. The pseudotensor that measures the density of gravitational energy and momentum in the cylindrical wave is everywhere zero. The significance of this finding is the subject of this paper. We conclude that many of the otherwise apparently paradoxical properties of this cylindrical wave can be understood by taking into account the analogy between gravitational waves and electromagnetic waves, and the special demands of the equivalence principle, which rules out a special role for any particular frame of reference.

Section 2 recapitulates the expressions of Einstein and Rosen and of Rosen for the metric of the cylindrical wave. Two kinds of solution are of interest: monochromatic waves and pulses. A pulse type of solution is constructed that is represented by particularly simple mathematical expressions. Section 3 reviews the proof that the pseudotensor density of gravitational

[1] A. E. Scheidigger, Revs. Modern Phys. **25**, 451 (1953).

[2] N. Rosen in *Jubilee of Relativity Theory*, edited by A. Mercier and M. Kervaire (Birkhäuser Verlag, Basel, 1956).

[3] A. Einstein and N. Rosen, J. Franklin Inst. **223**, 43 (1937); N. Rosen, Bull. Research Council Israel **3**, 328 (1953).