Quantizing space-time

J. G. Taylor
Department of Mathematics, King's College, London
(Received 5 June 1978)

We outline a theory of space-time based on quantum mechanics and general covariance. The universe is assumed to be constructed from a non-affinely-connected differentiable manifold and acts as the arena for the dynamical forms which can be supported by it. There are only a restricted set of such forms, corresponding to particles of spin up to 2, and only a restricted set of particle symmetries [octonions combined with de Sitter symmetry broken to O(4)]. Quantum mechanics is fomulated covariantly by the functional integral method, and space and time as usually experienced is reconstructed by the classical limit of $h \to 0$. The conclusion is drawn that neither space nor time themselves can be regarded as fundamental.

I. QUANTUM MECHANICS VERSUS GRAVITY

One of the outstanding questions for present theoretical physics is to obtain a physically sensible combination of quantum mechanics and gravity. The problem of quantizing gravity has received a great deal of attention in the last decade[1] though without any resolution of various difficulties that have become apparent in the attempt. The most extreme difficulty has been that caused by the ultraviolet divergences which arise when the standard techniques of quantum field theory are applied to Einstein's gravitational theory.[2] It is possible that these problems are only symptoms of deeper diseases caused by the very attempt to quantize gravity. It is that question we wish to consider in this paper, though we will not consider if our analysis of the situation can necessarily alleviate the ultraviolet nonrenormalizability.

If we wish to quantize gravity then not only do we have to take gravity seriously by choosing, for example, a specific description of gravity, but also we have to do the same for quantum mechanics. We will here assume that *all* material phenomena are to be described in a quantum-mechanical framework, with its usual paraphernalia of state vectors, operators, and associated probability interpretation. This latter has well-known interpretational difficulties for closed cosmologies,[1] but we will find that there are even deeper questions to resolve when we try to construct a theory of gravity in such a framework.

Our first basic assumption, that (I) quantum mechanics is all embracing, may be false. We may find that it presents such insuperable difficulties to constructing a reasonable theory of gravity that we are forced to admit classical concepts at the start of such a program. However, we will attempt to discover how far we can go towards constructing a world in which $\hbar$ is supreme. We should approach such a task with the attitude of one who has only been taught quantum mechanics. At worst our ideal investigator should believe that classical mechanics is merely a useful approximation scheme to quantum mechanics valid only in certain situations.

If classical concepts are only approximations of quantum ones then no classical fields of any sort can exist *ab initio*. Classical fields can thus only be regarded as useful approximations to certain quantum fields, being obtained as expectation values of their quantized brethren in suitable states. As for ordinary quantum mechanics, Ehrenfest's theorem would apply to indicate under which conditions the classical fields would describe important dynamical features of a mechanical situation. But the classical fields could not be used to give a fundamental dynamical description of the system.

We meet here the basic difficulty of combining quantum mechanics with gravity. For the latter is presently regarded, with overwhelming experimental support, as a metric theory of space-time, in which the invariant length ds^2 of an infinitesimal line element between the four-vectors x and $x+dx$ is expressed in terms of the gravitational potentials $g_{\mu\nu}(x)$ as

$$ds^2 = g_{\mu\nu}\, dx^\mu dx^\nu \, . \tag{1}$$

It is not possible to interpret (1) in any sensible manner if $g_{\mu\nu}$ is a quantized field, since if the left-hand side is defined as an expectation value of the right-hand side there is no specification of the state by which the expectation value is to be defined.

One can avoid this difficulty by considering the dynamical development of the quantized gravitational field as consisting of fluctuations about some classical gravitational field. Equation (1) is thus regarded as the average result when this latter field is used for $g_{\mu\nu}$. But such an approach is unsatisfactory for a quantized metric theory of gravity since in any such theory the metric (1) cannot play a fundamental role, as we have remarked above. In other words the notion of space-time as a Riemannian manifold can only be a classical approximation to an underlying quantum reality. Thus, the vast superstructure of Riemannian geometry does not appear necessarily relevant to quantum gravity. It is certainly possible to make the simple extrapolation of Riemannian geometry from the classical to the quantum-field-theoretic context.

Such a simple extension of classical to quantal is the one which has been used over and over again since the earliest days of quantum mechanics. The successes achieved thereby lead one to expect that such an extension will also work in the gravitational case. But here one is also dealing with the structure of space-time. If $g_{\mu\nu}$ is a quantized field then, as was remarked, *a priori* there is no Riemannian structure for space-time. There appears to be no way of defining such a structure except by singling out some classical metric field, which possibility was excluded by postulate (I), as was remarked above.

In such a context it is not even clear that space-time is a differentiable manifold. But then how can the concepts of quantum field theory, such as an action principle, be defined in such a general setting? *A priori* there would appear to be no definition, for example, of an invariant space-time volume in terms of d^4x, since this latter would only be definable for space-time a differentiable manifold, but not an arbitrary topological space.

The usual approach to quantum gravity does not discuss this question. If no background field is assumed explicitly, then it is usually introduced implicitly by the magical formula

$$\langle g_{\mu\nu}\rangle = \eta_{\mu\nu} . \tag{2}$$

This brings in the background Minkowski metric, for which there is no problem of the above sort; perturbation theory around the Minkowski metric then allows the effects of fluctuations around that metric to be calculated explicitly.

But there may be many (even infinitely many) solutions of the appropriate field equations which could be substituted for the right-hand side of (2). There appears no *a priori* reason to choose among various such possibilities. On physical *a posteriori* grounds it may be argued that an expanding universe metric is more appropriate than the flat Minkowski value. However, our assumption (I) cannot support the *a priori* choice of any specific background field over any other. We note parenthetically that the most satisfactory theory of the forces of nature would be one with only one classical solution, so that such a choice of backgrounds is destroyed. We will not attempt to apply such a criterion here. Our further discussion will therefore only apply to theories which have a plethora of classical solutions, as is the case with present theories containing Einstein general relativity.

Our conclusion from the above arguments is that quantum mechanics is not compatible with a Riemannian metric space-time. Nor, on the same grounds, is there expected to exist any preferred classical affine connection, so that space-time is not even an affine manifold. It is necessary to consider a set of space-time events with more general structure in order to include quantum gravity in a consistent fashion.

Having lost any preferred metric, we have also lost the ability to define time either locally, by means of local inertial frames, or globally by means of a suitable globally defined variable. Traditional quantum mechanics has so far only been defined when some time variable has been present. We will therefore have to construct a modified form of quantum mechanics in which time is not indispensable. We will also have to see how the time of our physical experience can arise from some more general approach. In order to do that we will first try to develop a general approach to quantum space-time which will hopefully contain quantum gravity without the above-mentioned inconsistencies.

II. SPACE-TIME AS A DIFFERENTIABLE MANIFOLD

One of the most basic features of our experience of the material world is that events in it can be described by means of four real coordinates. We thus have good experiential grounding for the second basic assumption of this paper, that (II) space-time is a differentiable manifold M of dimension four. We are assuming thereby, that the set of all space-time events forms a topological space which is locally homeomorphic to Euclidean four-space with certain compatibility relations on the coordinate functions on overlapping neighborhoods. The relations are that if ϕ and ψ are coordinate functions which map their domains in M homeomorphically into open sets in R^4 and if domain $\phi \cap$ domain $\psi \neq \phi$ (ϕ =nul set), then $\psi \circ \phi^{-1}$ is a C^k map from ϕ (domain $\phi \cap$ domain ψ) into R^4; k is a non-negative integer, ∞, or ω (for analytic maps). The resulting manifold is called a C^k-differentiable manifold.[3] On such a manifold it is then possible to define s-times differentiable (or C^s) functions f on M by the condition that $f \circ \phi^{-1}$ is C^s on the image of each coordinate function $\phi (s \leq k)$.

Such a supposition seems in agreement with naive experience, though goes considerably beyond it in the compatibility condition on overlaps. Thus, (II) may be false. For example, there may be regions of space-time in which the overlap condition breaks down or which are not locally homeomorphic to Euclidean four-space R^4 or even to any finite-dimensional Euclidean space. It would seem difficult to give numerical predictions in such pathological regions. Loss of the compatibility conditions gives a corresponding loss of suitably differentiable functions defined on M. Real numbers representing events seem to be the *sine qua non* of scientific validation. Yet we may find that experience is better described by a different assumption than the one above. As in the previous section we will test our basic assumption by deducing all we can from it to see if it fits experience correctly or not. We have not yet specified the degree of differentiability which M should possess. In order that we may have at least twice-differentiable functions defined on M we will assume that M is a C^k-differentiable manifold with $k \geq 2$.

We have assumed the dimension of the space-time manifold to be 4. It might be possible to prove that to be the only value supporting a nontrivial dynamics by showing that for lower dimension the dynamics degenerates, while for higher dimension than 4 it is unrenormalizable from a quantum-field-theory viewpoint. Since we have not yet completed such a proof we have made assumption (II) contain the four-dimensionality. Some of our further results will not depend heavily on such dimensionality, though the details of the dynamical forms to be given shortly will so do.

We must now ascertain if our two assumptions I and II contradict each other. They would have done so if we had chosen for our second assumption that space-time was an affine or Riemannian metric differentiable manifold. At no time need there be any preferred classical field, be it the metric tensor or the affine connection, if (II) be valid. Yet there is still a preferred manifold structure on M, however, so that some contradiction might still be expected with (I). Indeed that will specifically be so if M were assumed to be paracompact, and so still possess a Riemannian structure.[4] But space-time must have some structure in order for differentiable functions to be defined on it, as was remarked above. How far does our corollary of (I), that there are no preferred classical fields on space-time, extend to the nonexistence of preferred topological structures of any sort?

We might conjecture that the quantized fields allowed on a general enough differentiable manifold M give dynamical consequences which are independent of the specific differentiable structure on M. We would thus have compatibility between (I) and (II). (I) and (II) are compatible provided that for (II) we require that M has neither Riemannian metric nor even affine structure. We will have to return to this question as to whether or not such compatibility does actually occur when the possible dynamical consequences of (I) and (II) have been more fully explored. In the process we will, however, try to single out those dynamical features independent of the particular differentiable structure on M.

Though we have apparently rejected the metrical foundations upon which Einstein's theory of general relativity was based there is one crucial feature of that theory which can be saved, the principle of general covariance. As stated succintly by Einstein[5] "As all our physical experience can be ultimately reduced to such coincidence" (of point events), "there is no immediate reason for preferring certain systems of coordinates to others, there is to say, we arrive at the requirement of general covariance." We will assume, with Einstein, that general covariance is valid. In other words, we will require that physically observable results are always expressible in coordinate independent fashion. We will further assume that dynamical equations themselves can also be expressed in a coordinate independent way; there are to be no coordinate-independent observables needed for specifying the dynamics.

It is accepted that general covariance cannot be used in classical general relativity to delimit physical laws.[6] Only through the additional requirement of the principle of equivalence can the detailed form of interactions of gravity and matter fields become specified. There are various forms of this latter principle, but all require the ability to choose local coordinates near a point which removes the gravitational field. This is not possible in general when gravity is regarded as a quantized field. We meet again the same difficulty of the previous section: The principle of equivalence can only apply to expectation values or classical gravitational fields. Since the existence of preferred cases of these latter is contradictory to our basic assumption (I), we have not only lost the Riemannian geometric interpretation of gravity given by Einstein, but also any natural way of expressing gravity as a universal quantized field of nature interacting in a very specific manner with matter fields.

It is important to realise that while general covariance does not have such direct physical import as the principle of equivalence in the classical regime this is no longer true when assumptions (I) and (II) are considered. In particular, the possible dynamical fields on a nonaffine differentiable manifold are extremely restricted. If that is so then we need not mourn too much the loss of the principle of equivalence. We might even attempt to prove it from the allowed dynamical structures. To see how this might be so let us turn to investigating the possible dynamics on a general manifold.

III. DYNAMICS ON ARBITRARY MANIFOLDS

We will analyze the dynamical properties which can be supported by M by determining the set of all the fields which can be regarded as developing dynamically on M. Such development will be assumed to be described by differential equations satisfied by certain of the fields on M, following the tradition of development of field theory since Newton's theory of gravitation. It is not possible to differentiate tensor fields on a coordinate-independent manner, since we have agreed that we cannot introduce an affine connection on M. We may, however, use the forms on M,[7] since for any form ω_r of degree r on M, the derivative $d\omega_r$ is defined on M in a coordinate-independent fashion, being of degree $(r+1)$. An equivalent and more useful form of this is that it is possible to differentiate the covariant antisymmetric tensor fields on M.

To proceed further we assume that dynamics on M is governed by an action principle in terms of a local action dA (essentially d^4xL in terms of the usual Lagrangian density L). dA will itself be assumed to be a form of degree 4 on M which is also a local function of the various dynamical forms of various degrees on M. If these are denoted by $\omega_0, \omega_0', \ldots, \omega_1, \omega_1', \ldots, \omega_r, \omega_r', \omega_r'', \ldots,$ then

$$dA = dA(\omega_0, d\omega_0; \omega_0', d\omega_0', \ldots; \omega_1, d\omega_1, \ldots)\,. \tag{3}$$

We may integrate A locally by means of the coordinate functions in the usual way,[8] so can give sense to the action $A(V)$ defined over some open set V in M:

$$A(V) = \int_v dA\,. \tag{4}$$

In this manner we may obtain the Euler-Lagrange variational equations from (4) by infinitesimal variations of the forms $\omega_0, \omega_0', \ldots,$ etc.:

$$\delta A(V) = 0\,. \tag{5}$$

We can attempt to justify such use of an action principle since the resulting dynamical system should then be quantized in a manner which will satisfy unitarity and other reasonable physical requirements. We have limited the local action only to be a function of the forms and their first derivatives due to the ghost difficulties associated with higher derivatives. In any case we cannot justify (3), (4), and (5) fully except by further analysis of their consequences.

We will next consider the possible contributions to dA. Since there are no preferred classical fields we can only construct dA from the exterior products of forms of suitable degree. There are only a limited set of such products, which we can tabulate explicitly. Denoting by ω_r the generic forms of degree r, we note first that forms of degree 4 containing two derivatives are, to within powers of ω_0

$$\begin{aligned}&\text{(i) } d\omega_i \wedge d\omega_{2-i}\,,\\ &\text{(ii) } d\omega_i \wedge d\omega_j \wedge \omega_{2-i-j}\,,\\ &\text{(iii) } d\omega_i \wedge d\omega_j \wedge \omega_k \wedge \omega_{2-i-j-k}\\ &(0 \le i,j \le 1;\ 1 \le k \le 4;\ i+j+k < 2)\,.\end{aligned} \tag{6}$$

Those containing one derivative are

$$\begin{aligned}&\text{(i) } d\omega_3, \quad \text{(ii) } d\omega_i \wedge \omega_{3-i}\,,\\ &\text{(iii) } d\omega_i \wedge \omega_j \wedge \omega_{3-i-j}\,,\\ &\text{(iv) } d\omega_i \wedge \omega_j \wedge \omega_k \wedge \omega_{3-i-j-k}\\ &(1 \le j,k \le 4,\ \ i+j+k < 3)\,,\end{aligned} \tag{7}$$

while those with no derivatives are

$$\begin{aligned}&\text{(i) } \omega_4, \quad \text{(ii) } \omega_j \wedge \omega_{4-j}\,,\\ &\text{(iii) } \omega_j \wedge \omega_k \wedge \omega_{4-j-k}\,,\\ &\text{(iv) } \omega_j \wedge \omega_k \wedge \omega_l \wedge \omega_{4-j-k-l}\,,\end{aligned} \tag{8}$$

where $i \le j,k,l \le 4,\ j+k+l < 4$.

Since $d\omega \wedge d\phi = d(\omega \wedge d\phi)$, where $d^2\phi = 0$, then the first term in (6) gives only boundary contributions to $A(V)$ plus a term of the form (6) (iii) so can be neglected. We can thus write the independent terms of (6), (7), and (8) in detail as in Table I, those from (6) being expressed as quadratics in $d\omega_r$. We note that the forms ω_r are only representatives of the class of all r forms, so that Table I denotes possibly an infinite set for each

TABLE I. The components of L.

Order in $d\omega_r$		Terms
Quadratic, (6)	(i)	$d\omega_1 \wedge d\omega_1$
	(ii)	$d\omega_0 \wedge d\omega_0 \wedge \omega_2$
	(iii)	$d\omega_0 \wedge d\omega_0 \wedge \omega_1 \wedge \omega_1$
Linear, (7)	(i)	$d\omega_3$
	(ii)	$d\omega_2 \wedge \omega_1$
	(ii)	$d\omega_1 \wedge \omega_2$
	(iii)	$d\omega_1 \wedge \omega_1 \wedge \omega_1$
	(iv)	$d\omega_0 \wedge \omega_1 \wedge \omega_1 \wedge \omega_1$
Zero, (8)	(i)	ω_4
	(ii)	$\omega_3 \wedge \omega_1$
	(ii)	$\omega_2 \wedge \omega_2$
	(iii)	$\omega_2 \wedge \omega_2 \wedge \omega_1$
	(iv)	$\omega_1 \wedge \omega_1 \wedge \omega_1 \wedge \omega_1$

entry. We would expect the quadratic contributions to correspond to Bose fields, the linear to Bose fields in first-order form or Fermi fields and the zero terms to describe interactions with no dynamical propagation in M. Thus, we will restrict our attention for the present to the linear and quadratic terms.

Let us consider the quadratic contributions in more detail. The various terms can be written in coordinate-dependent fashion as proportional (factors of powers of the scalar field A being omitted) to

$$\epsilon^{\mu\nu\lambda\sigma}\partial_\mu A_\nu \partial_\lambda A_\sigma , \tag{9}$$

$$\epsilon^{\mu\nu\lambda\sigma}\partial_\mu A\,\partial_\nu A\,A_{\lambda\sigma} , \tag{10}$$

$$\epsilon^{\mu\nu\lambda\sigma}\partial_\mu A\,\partial_\nu A\,A_\lambda A_\sigma , \tag{11}$$

where, locally in some coordinate representation $\omega_2 = A_{\mu\nu}dx^\mu \wedge dx^\nu$, $\omega_1 = A_\mu dx^\mu$, $\omega_0 = A$, and $\epsilon^{\mu\nu\lambda\sigma} = \pm 1$ ($\mu\nu\lambda\sigma = \pm$ permutation of 1234), and $=0$ otherwise. In none of (9), (10), or (11) is it possible to obtain a quadratic derivative term in A_ν or A, respectively, by integrating by parts, due to the presence of $\epsilon^{\mu\nu\lambda\sigma}$ (corresponding to $d^2\omega_r = 0$). Thus, in no classical background fields A, $A_{\mu\nu}$ or A_μ, respectively, will A_μ or A propagate as bosons in terms of the inverses of d'Alembertians. We take this as indicating that none of (9), (10), or (11) are suitable to propagate the forms ω_0 or ω_1 dynamically on M. It is to be noted that this criterion does not depend on the existence of any preferred classical field on M. Indeed, it is a criterion arising naturally from the functional integral definition of quantization, where a sum over all classical fields is used. We are thus left with the five classes of terms linear in $d\omega_r$ in Table I from which to obtain dynamics on M.

Let us now consider which expressions from these latter terms could serve as first-order terms to describe the propagation of bose fields on M. We will again require them to reduce to the usual quadratic second-order form contribution to the action density in an arbitrary classical background. The term $d\omega_3$, in coordinate-dependent form, is locally

$$f(A)\epsilon^{\mu\nu\lambda\sigma}\partial_\mu A_{\nu\lambda\sigma} ,$$

where $f(A)$ is an arbitrary local function of A. This can only serve for a first-order action density if there is also a term in L quadratic in ω_0 or ω_3. Only the first possibility could occur, according to Table I, being proportional to A^2 in a background field A_μ or $A_{\mu\nu}$. The resulting quadratic expression in L will be $(\epsilon^{\mu\nu\lambda\sigma}\partial_\mu A_{\nu\lambda\sigma})^2$ multiplied by background field terms; writing $\epsilon^{\mu\nu\lambda\sigma}A_{\nu\lambda} = B^\mu$ we see that the expression $(\partial_\mu B^\mu)^2$ would result. This has no propagation (the quadratic derivative operator is $\partial_\mu\partial_\nu$ with characteristic surface reduced to the origin) so does not seem an appropriate candidate for a dynamical form on M.

We are thus left with the last four linear terms in Table I, together with appropriately chosen terms from the zero-order terms, to describe in first-order form the dynamics of M. These terms appear to give suitable candidates. In order to consider them in detail we must clarify the range of values of the forms of ω_r and conditions related thereto.

IV. GAUGE INVARIANCE ON THE MANIFOLD

So far we have not specified the space in which the forms ω_r on M take their values. In order that the discussion of the last section be sensible we must require that this space be at least an algebra A, which we take to have a finite basis denoted by X_a. Thus, each form ω_r will (locally) have a representation (summation convention assumed)

$$\omega_r = \omega_r^a X_a . \tag{12}$$

We can now impose the further condition that all dynamics must be independent of any local change of basis. Thus, in addition to our assumption (II) of coordinate independence of results (space-time labels do not matter), we also require that a labels do not matter, even locally. If A were a Lie algebra, then this condition would be that of local gauge invariance under the corresponding Lie group. In the general case we will find that the allowed terms in the Lagrangian will be even more restricted than those we arrived at in the previous section.

An important reason why we should search for an algebra A larger than the trivial one-dimensional one is that we wish to obtain a dynamical structure on M including at least that of classical general relativity. Otherwise, neither space not time will be differentiated or even seem possible of being defined. A must therefore be chosen large enough to allow us to recognize spin-2 fields which might qualify for a metric tensor or affine connection on M. More positively, we might expect that M would have realized on it the largest possible A restricted by certain criteria. We will look for such an A with the further condition that there is only one spin-2 massless field contained in the theory. This latter requirement, as we will see allows us to find a simple maximal A.

Let us first remark that not all forms need be assumed to belong to the adjoint representation, as (12) dictates, but some may be chosen to transform as the fundamental (or some other) representation of A. Nor is it clear that the X_a's need to be chosen only as the generators of a Lie algebra. Thus, they could be taken as the basis elements of a finite-dimensional not necessarily associative algebra A. Local gauge invariance would now be invariance under the automorphisms of A. This allows inclusion of octonion-valued fields, which have been discussed recently with special reference to quark confinement.[9] We will not discuss such a modification here in detail.

We would develop our discussion in the language of fiber bundles,[10] but feel that the purely differential-geometric language of forms is more appropriate; we will persist therefore in such a framework. We first require, for the existence of a covariant derivative, a particular one-form which we denote by ω_1. This will transform under a general gauge transformation g of the Lie group G defined by the Lie algebra A (or the group G of automorphisms of A if A is not a Lie algebra) as

$$\omega_1 \to g\omega_1 g^{-1} - dg\,g^{-1} . \tag{13}$$

We may thus define the covariant derivative $D\omega_1$ of ω_1 by

$$D\omega_1 = d\omega_1 + \omega_1 \wedge \omega_1 , \tag{14}$$

with the appropriately covariant transformation law

$$D\omega_1 \to gD\omega_1 g^{-1} .$$

Any r-form ω_r transforming covariantly

$$\omega_r \to g\omega_r g^{-1}$$

will thus have a covariant derivative

$$D\omega_r = d\omega_r + \omega_1 \wedge \omega_r - (-1)^r \omega_r \wedge \omega_1 , \tag{15}$$

while if ω_r transforms as

$$\omega_r \to g\omega_r ,$$

then the corresponding covariant derivative is

$$D\omega_r = d\omega_r + \omega_1 \wedge \omega_r . \tag{16}$$

We note that singling out ω_1 from the other forms does not contradict our assumption (I). The existence of ω_1 does not require the existence of any preferred classical field, as would the existence of a metric or affine structure for M with associated covariant derivative of tensors on M. Thus, ω_1 can be a quantized field without any difficulty being caused thereby to the structure we have set up.

One solution has already been found to contain the Einstein Lagrangian and be maximal[11]. The generators X_a are the set $\{\sigma_{ab}, \gamma_a, \gamma_5\gamma_a X_i, \gamma_5\gamma_i, X_i\}$, where γ_a are the usual 4 Dirac matrices ($\sigma \leq a \leq 3$), $\sigma_{ab} = (1/2i)[\gamma_a, \gamma_b]_-$, $\gamma_5 = \gamma_0\gamma_1\gamma_2\gamma_3$, and X_i are the other generators of a (not necessarily associative) algebra with product rule

$$X_i X_j = -C_{ij}\underline{1} + f_{ijk} X_k , \tag{17}$$

where $C_{ij} = C_{ji}$ and $f_{ijk} = -f_{jik}$. It is the term involving C_{ij} which prevents extra gravitons from appearing when the algebra A is closed. If it is assumed that the f_{ijk} are the structure constants of a Lie group, then it is straightforward to see that the only possibility for $\{X_i\}$ is the set of 2 $\times$2 Pauli matrices. For let H be the Cartan subalgebra of the Lie group (assumed simple), the basis elements in a particular representation being H_i in H and E_α otherwise, and $|m\rangle$ a state in that representation. We have, from the commutation relations of A and (17),

$$\begin{aligned} H_i E_\alpha |m\rangle &= (m_i + \alpha_i) E_\alpha |m\rangle \\ &= -C_{i\alpha}|m\rangle + \tfrac{1}{2}[H_i, E_\alpha]_- |m\rangle \\ &= C_{i\alpha}|m\rangle + \tfrac{1}{2}\alpha_i E_\alpha |m\rangle \end{aligned}$$

or

$$(m + \tfrac{1}{2}\alpha_i) E_\alpha |m\rangle + C_{i\alpha}|m\rangle = 0 .$$

Since $E_\alpha|m\rangle$ and $|m\rangle$ are linearly independent, then $C_{i\alpha} = 0$ and either $(m_i + \frac{1}{2}\alpha_i) = 0$ or $E_\alpha|m\rangle = 0$. In the latter case the representation can only be the trivial one, while in the former only a two-dimensional representation is possible, with basis $|m\rangle, E_{-2m}|m\rangle$. It is straightforward to see that this is only possible for the Lie algebra SU(2). Furthermore, (17) excluded direct products of SU(2) with itself any number of times. In this case the gauging is straightforward, going along the usual lines. The other known solution to (17) is the octonion nonassociative algebra.[12] This does not allow such a simple gauging due to the nonassociative character of the octonion algebra[9]; we will not consider this question further here.

The components of the gauge form ω_1 were written in Eq. (13) of Ref. 11 as

$$\begin{aligned} \omega_1 = dx^\mu (&B_\mu^{ab}\sigma_{ab} + l_\mu^a \gamma_a + i a_\mu^{ai}\gamma_a\gamma_5 X_i \\ &+ p_\mu^i \gamma_5 X_i + S_\mu^i X_i) , \end{aligned} \tag{18}$$

where the associated local action is, in first-order form,

$$dA = \mathrm{Tr}(\gamma_5 D\omega_1 \wedge D\omega_1) . \tag{19}$$

We remark that although (19) appears to be one of the quadratic derivative terms which we discarded earlier the truly quadratic term $\mathrm{Tr}(\gamma_5 d\omega_1 \wedge d\omega_1)$ is indeed zero, leaving only terms linear and of zeroth order in the derivative (as shown in detail in Ref. 11).

The second-order form of dA in (19) was shown, independently of the existence of any preferred fields such as (2), to be that of a massless spin-2 field described by the vierbein l_μ^a and connection B_μ^{ab} in interaction with a multiplet of massless vector mesons S_μ^i with associated field strength proportional to a_μ^{ai}; the components p_μ^i do not propagate. Since the theory based on (19) is generally covariant, by construction, and the self-interacting Lagrangian for the spin-2 terms is that of Einstein, it is natural to suppose that this latter describes the gravitational field. Thus, (19) should be a suitable Lagrangian for interacting gravitational and Yang-Mills theories (the latter having only a restricted range of symmetries).

We see from Table I and our earlier discussion that it is possible to introduce other terms in the local action beyond (19). We will discuss their nature briefly now so as to understand better the range of possibilities allowed by assumptions (I) and (II) above. Fermion forms ω_r may contribute bilinearly in dA if $r=0$ or 1, the corresponding terms (for ω_r being a vector in the representation space of A and $\tilde{\omega}_r$ transforming contragrediently) being $D\tilde{\omega}_1'\gamma_5 \wedge \omega_1'' \wedge \omega_1'$ and $D\tilde{\omega}_0\gamma_5 \wedge \omega_1'' \wedge \omega_1'' \wedge \omega_1''\omega_0$, where ω_1'' are some suitably covariant one-forms. Choosing these latter to be ω_1, we obtain the usual Rarita-Schwinger and Dirac Lagrangian densities $\epsilon^{\mu\nu\lambda\sigma}D_\mu\bar{\psi}_\nu\gamma_5 l_\lambda\psi_\sigma$ and $\epsilon^{\mu\nu\lambda\sigma}D_\mu\bar{\psi}\gamma_5 l_\nu l_\lambda l_\sigma\psi$, respectively, along with other interaction terms. There will also be mass and self-interaction terms constructed from $\omega_0\omega_1 \wedge \omega_1 \wedge \omega_1 \wedge \omega_1\omega_0$ and $\overline{\omega}_1' \wedge \omega_1 \wedge \omega_1 \wedge \omega_1'$, possibly multiplied by powers of $\tilde{\omega}_0\omega_0 = \overline{\psi}\psi$. There are also terms involving the bosonic two-form ω_2, if that can be propagated on M; these would be $\tilde{\omega}_1'\gamma_5 \wedge \omega_2 \wedge \omega_1'$ and $\tilde{\omega}_0\gamma_5\omega_2 \wedge \omega_2\omega_0$. These give contributions to the Lagrangian density, in component form, proportional to $\epsilon^{\mu\nu\lambda\sigma}\overline{\Psi}_\mu\gamma_5 l_{\nu\lambda}\overline{\Psi}_\sigma$ and $\epsilon^{\mu\nu\lambda\sigma}\overline{\psi}[\gamma_5 l_{\mu\nu}l_{\lambda\sigma}]_-\psi$, where $\omega_2 = dx^\mu \wedge dx^\nu l_{\mu\nu}$.

Let us turn, then, to the further bosonic possibilities on M. Since all the terms listed in Table I are in first-order form, it seems presently necessary to go through the same exercise as in Ref. 11 and reduce the terms to second-order form before their physical content can be unravelled. We first consider the bosonic two-form ω_2, for which there is the derivative term $\mathrm{Tr}(\gamma_5 D\omega_2 \wedge \omega_1)$. In order to obtain a nontrivial result it is necessary to include the term $\omega_2 \wedge \omega_2$, so that the variational equation for ω_2 will express it as a derivative of ω_1 and the resulting local action as quadratic in the derivatives of ω_1. However, the form ω_2 will then be purely a function of ω_1, so not be propagating independently. We may choose another 1-form ω_1'', transforming under a local gauge transformation g as $\omega_1'' \to g\omega_1''g^{-1}$, and consider the joint effect of the terms $\mathrm{Tr}(\gamma_5 D\omega_2 \wedge \omega_1'')$ and $\mathrm{Tr}(\gamma_5\omega_2 \wedge \omega_1'' \wedge \omega_1'')$. It is possible to construct terms with suitable propagation characteristics for ω_1'', such as the Lagrangian density

$$\epsilon_{abcd}\epsilon^{\mu\nu\lambda\sigma}(\partial_\mu B_\nu''^{ab}B_{\lambda\sigma}^{cd} + B_\mu''^{aa_1}B_\nu''^{a_1 b}B_{\lambda\sigma}^{cd}) . \quad (20)$$

In a range of classical background fields (20) will allow propagation of the field $B_{\mu\nu}''^{ab}$ which reduces to a massless scalar. This is because the quadratic terms in (20) are invariant under $\delta B_{\mu\nu}''^{ab} = \partial_\mu\epsilon_\nu^{ab} - \partial_\nu\epsilon_\mu^{ab}$, with $\epsilon_\mu^{ab} = -\epsilon_\mu^{ba}$, so that $B_{\mu\nu}''^{ab}$ has only one independent component, and can describe at most a massless scalar. The other components of ω_1'' can describe at most spin-1 or -0 fields; we will give a detailed discussion of these fields elsewhere.

Another possibility for ω_1'' and ω_2 is from the terms $\mathrm{Tr}(\gamma_5 D\omega_2 \wedge \omega_1'')$ and $\mathrm{Tr}(\gamma_5\omega_2 \wedge \omega_2)$. However, these terms give the only satisfactory second-order term $\epsilon^{\mu\nu\lambda\sigma}\epsilon_{abcd}\partial_\mu B_\nu''^{ab}\partial_\lambda B_\sigma''^{cd}$, which in any case is zero to within a surface term. Hence, there are no ways of propagating the pair ω_2, ω_1'' than by the terms we discussed above involving $D\omega_2 \wedge \omega_1''$ and $\omega_2 \wedge \omega_1'' \wedge \omega_1''$, these containing fields of spin 0 and 1.

We turn to the possibility of propagating ω_1'' without use of ω_2. The only way of achieving purely quadratic expressions for derivatives of ω_1'' would be by use of the terms $\mathrm{Tr}(\gamma_5 D\omega_1'' \wedge \omega_1''' \wedge \omega_1)$ and $\mathrm{Tr}(\gamma_5\omega_1'' \wedge \omega_1'' \wedge \omega_1 \wedge \omega_1)$. In a locally constant background ω_1 the result of such terms would be the Einstein-restricted Yang-Mills theory described by (19), but with ω_1'' replacing ω_1, and in its linearized version. This would seem to produce a further graviton, though this can be offset by the term $\mathrm{Tr}(\gamma_5\omega_1''' \wedge \omega_1''' \wedge \omega_1 \wedge \omega_1)$. On a locally constant background ω_1 this gives the Pauli-Fierz mass contribution proportional to $(l_{\mu\mu}'''l_{\nu\nu}''' - l_{\mu\nu}'''l_{\mu\nu}''')$, so that the resulting spin-2 field need not compete with the graviton. The other contributions from these combined terms can have spin 1 in suitable locally constant background fields.

Finally, we mention the possibility of scalar fields arising directly from zero forms, using one- or two-forms in the manner described above to give a first-order local action whose second-order version will be quadratic in the scalar fields in suitable locally constant background fields. Here again we will not give any details, but reserve this for elsewhere.

We conclude that the terms in Table I can give contributions to the local action corresponding to fields of spins 0, $\frac{1}{2}$, 1, $\frac{3}{2}$, and 2. This conclusion is made modulo the existence of locally constant background fields of spin 1 and 2 (Einstein-Yang-Mills). This is not a statement in contradiction with assumptions (I) and (II), but will require discussion when the role of classical fields in the quantal framework has been explicated. We must turn, then, to the question of field quantization, but before doing so we should note that an alternative group approach is possible in which the whole algebra is not used for local gauge transformations. Thus, the SL(2, C) approach to unified Einstein-Yang-Mills theories in Ref. 11

followed this avenue by taking the group SL(2, C) as the unbroken invariance group. The algebra sl(2, C) was extended to that of su(2, 2). One-forms with values in $[\mathrm{su}(2,2) - \mathrm{sl}(2,C)]\times A$ for any Lie algebra A, transforming covariantly under SL(2, C), were then introduced and used to construct four-forms invariant under local SL(2, C) and A gauge transformations, giving the Einstein-Yang-Mills Lagrangian when reduced to second-order form. The extra internal symmetry is gained, however, at the expense of the rather artificial group extension process involved. It may indeed be necessary to describe the universe in this way, but presently the scheme based on local gauge invariance relative to this whole group is to be preferred.

V. QUANTIZATION ON THE MANIFOLD

We must now turn to the question of how to quantize the dynamical forms we introduced in the previous sections. This question is crucial to answer. For as we emphasized in the first two sections, we are regarding the quantum framework as superior to the classical one. The latter is only to be regarded as derived from the former as an approximation, valid in certain circumstances. But the problem of defining a quantization process without any preferred classical fields is not trivial. We do not have a definition of time in the manifold, so the traditional approach by way of Schrödinger or Heisenberg equations of motion does not exist. Nor can we use canonical or covariant techniques in any form, again due to lack of timelike surfaces. Even the more general methods of axiomatic field theory do not seem particularly helpful since we have to attempt to obtain a detailed dynamical description in a completely coordinate-free manner. Indeed, it seems that this is a problem which has not been discussed in the literature, at least as far as I know.

To proceed we will consider first the notion of a quantum-mechanical state. We may define the labels of a state in a coordinate and gauge-independent fashion by means of the eigenvalues of the dynamical forms [those for which $d\omega$ enters into dA of (3)] on submanifolds. It is not to be expected that these eigenvalues can be independently evaluated at all points of the whole manifold, and we will assume that they can at least be given on three-dimensional submanifolds. For one such, σ, we denote by $|\omega(\sigma)\rangle$ that state for which the dynamical forms ω take the values ω' when measured in σ. There may also be restrictions on the choice of σ, though any such might come into conflict with our first basic assumption. We will consider this further now when we define the inner product between two states.

We wish now to define this inner product or overlap $\langle\omega_1(\sigma_1)\,|\,\omega_2(\sigma_2)\rangle$ between two such states in a coordinate- and gauge-independent fashion. We will do that by means of the sum-over-paths formulation of quantum mechanics,[13] using the action density dA of Eq. (3). We consider first the case that the region $R(\sigma_1,\sigma_2)$ between σ_1 and σ_2 is enclosed by a coordinate neighborhood U, with associated coordinate functions. By this we mean that $R(\sigma_1,\sigma_2)\subset U$. We write, formally,

$$\langle\omega_1(\sigma_1)\,|\,\omega_2(\sigma_2)\rangle = \int \prod_{x\in R(\sigma_1,\sigma_2)} d\omega\, J(x) \exp\left[\frac{i}{h}\int_{R(\sigma_1,\sigma_2)} dA\right], \quad (21)$$

where the values of ω are fixed to be ω_i on σ_i. The integration measure $\Pi d\omega J(x)$ is formally defined in terms of the components $A_{\mu_1\cdots\mu_r}$ of each of the dynamical forms ω_r of degree r by

$$d\omega(x)=\prod_{r,\mu_1\cdots\mu_r} dA_{\mu_1\cdots\mu_r}(x), \quad J(x)=[\det l^a_\mu(x)]^{-\alpha}.$$

Here ϵ^a_μ are the γ^a components of the fundamental one-form ω_1 of Eq. (13), as defined by Eq. (18), and $\alpha=\sum r n_r$, where there are n_r separate components of the dynamical forms of degree r. It is the factor $J(x)$ which gives the coordinate invariance of the right-hand side of (21), since in an overlapping coordinate system U' with coordinates x' we have $J'(x')=[\partial(x)/\partial(x')]^{-\alpha}J(x)$, while $d\omega' =[\partial(x)/\partial(x')]^{\alpha}d\omega$.

We may extend (21) to the case that the region $R(\sigma_1,\sigma_2)$ between σ_1 and σ_2 is the union of a set of (possibly overlapping) coordinate neighborhoods $\{U_\alpha\}$. In each of these we choose the coordinate representations for the components of the forms, the different possible choices in overlapping neighborhoods not affecting the value of $d\omega J(x)$. We can thus define the overlap between the two states by the same formula (21), now extended to this general situation. This is not, in fact, the most general situation, since even though the manifold is covered by a suitable atlas defining the manifold structure, the region $R(\sigma_1,\sigma_2)$ between two manifolds need not be well defined in general. For this to be so it would seem necessary that the submanifolds σ_1 and σ_2 can be continuously deformed into each other with preservation of orientation. The integrations in (21) will then be over the points covered by this deformation. We will assume such a relation between σ_1 and σ_2 from now on, and say that σ_1 and σ_2 are measurably related.

There are two features of (21) which require further analysis. Firstly, the measure $\Pi d\omega$ is very poorly defined. There are various ways of making it more precise,[14] whose relevance to (21) we will consider elsewhere. Such delay is especially necessary since the rigorous definition of (21) would in any case require what is expected to be a lengthy analysis of ultraviolet and infrared divergences contained in the theory.

The second question is that of group invariance. Indeed the right-hand side of (21) is invariant under local group transformations if these are brought about by the groups considered in Sec. IV, since by (13) to (16) the changes in $d\omega$ brought about by the gauge transformation (13) will be at most proportional to $\det g$, which is one. We will have to consider the Faddeev-Popov ghost construction (15) in due course, but at this point we will turn to various formal features of (21) which seem of value to describe.

We note first that we have not called (21) a transition amplitude, which is its usual term. For we have no direction of time defined on the manifold, so no concept of transition. We will interpret (21) as a measure of the proportion of the state $|\omega'(\sigma_2)\rangle$ in the state $|\omega'(\sigma_1)\rangle$ (or vice versa). Such an interpretation would seem to require the ability to prepare the system in the state $|\omega'(\sigma_1)\rangle$ in the first place. But this latter phrase again brings time back into the discussion. It is difficult to describe in the usual way how anything at all happens in this present framework where time is absent, since we are so used to thinking in some sort of causal sequence. In the lack of the latter we can only consider events at given (locally defined) coordinates. In that sense we assume that somehow a state can be prepared over a three-dimensional submanifold. The similarity of this state to another prepared over another such submanifold is given by (21).

This similarity function has the following properties:

$$\text{(a)}\ \langle\omega'(\sigma_1)\,|\,\omega''(\sigma_2)\rangle\begin{cases}=0 & (\omega'\neq\omega'')\\ =\delta(\omega'-\omega'') & (\omega'=\omega'')\end{cases}\text{when } \sigma_1=\sigma_2 \tag{22}$$

since when $\sigma_1=\sigma_2$ the integrand of the right-hand side of (21) is the unit matrix in (ω', ω'') space, and there is no integration over internal variables,

$$\text{(b)}\ \sum_{\omega''}\langle\omega'(\sigma_1)\,|\,\omega''(\sigma_1)\rangle\langle\omega''(\sigma_2)\,|\,\omega'''(\sigma_3)\rangle =\langle\omega'(\sigma_1)\,|\,\omega'''(\sigma_3)\rangle\,, \tag{23}$$

where the summation over ω'' should more correctly be given as

$$\prod_{x\in\sigma_2}\int d\omega''(x)J''(x)\,.$$

(c) We may use (23) above to re-express (21) as a sum over contributions from similarities between neighboring submanifolds. Thus, if $\sigma(t)$, $0\leq t\leq 1$, is a family of manifolds defined by a family of continuous maps of a given manifold, say $\sigma(0)$, then we may write

$$\langle\omega'(\sigma(0))\,|\,\omega''(\sigma(1))\rangle=\lim_{n\to\infty}\sum_{\omega^{(1)},\ldots,\omega^{(n-1)}}\prod_{r=0}^{n-1}\langle\omega^{(r)}(\sigma(r/n))\,|\,\omega(\sigma^{(r+1)}((r+1)/n))\rangle\,, \tag{24}$$

where $\omega^{(0)}(\sigma(0))=\omega'(\sigma(0))$, $\omega^{(n)}(\sigma(1))=\omega''(\sigma(1))$. This formula (24) is that frequently used at various levels of rigor[13,14] to give a definition of the left-hand side of (24) by means of similarities between states on neighboring manifolds, $\sigma(t)$ and $\sigma(t+\epsilon)$, for arbitrarily small ϵ. We will attempt to use it shortly in evaluating the left-hand side of (24). We may now develop more general states than $|\omega'(\sigma)\rangle$ by taking sums of such eigenstates, of the form

$$|\Psi,\sigma\rangle=\int\prod_{x\in\sigma}d\omega'(x)\Psi(\omega'(\sigma))\,|\,\omega'(\sigma)\rangle\,, \tag{25}$$

where Ψ is a complex-valued functional of the eigenvalues ω' of ω on σ. The inner product between two such states $|\Psi_1,\sigma_1\rangle$ and $|\Psi_2,\sigma_2\rangle$ will thus be

$$\langle\Psi_1,\sigma_1\,|\,\Psi_2,\sigma_2\rangle=\int\prod_{x\in\sigma_1}d\omega_1(x)\prod_{y\in\sigma_2}d\omega_2(y)\Psi_1^*(\omega_1(\sigma_1))\Psi_2(\omega_2(\sigma_2))\langle\omega_1(\sigma_1)\,|\,\omega_2(\sigma_2)\rangle\,. \tag{26}$$

We will interpret (27) as being the similarity amplitude between the states $|\Psi_1,\sigma_1\rangle$ and $|\Psi_2,\sigma_2\rangle$. We may also define expectation values of observables by means of the resolution of the identity expressed by the right-hand side of (23). Thus, if $F(\omega(P))$ is a function of the dynamical forms at the point P, then we define its matrix element between the states $|\Psi_1\sigma_1\rangle$ and $|\Psi_2\sigma_2\rangle$ as

$$\langle\Psi_1\sigma_1\,|\,F(\omega(P))\,|\,\Psi_2(\sigma_2)\rangle=\int\prod_{x\in\sigma_1}d\omega_1(x)\prod_{y\in\sigma_2}d\omega_2(y)\Psi_1^*(\omega_1(\sigma_1))\Psi_2(\omega_2(\sigma_2))\langle\omega_1(\sigma_1)\,|\,F(\omega(P))\,|\,\omega_2(\sigma_2)\rangle\,, \tag{27}$$

where

$$\langle\omega_1(\sigma_1)\,|\,F(\omega(P))\,|\,\omega_1(\sigma_2)\rangle = \sum_{\omega'(\sigma)} \langle\omega_1(\sigma_1)\,|\,\omega_1(\sigma)\rangle F(\omega'(P))\langle\omega'(\sigma)\,|\,\omega_2(\sigma_2)\rangle\,, \tag{28}$$

with σ some three-dimensional submanifold through P and measurably related to σ_1 and σ_2. The right-hand side of (28) is independent of the choice of σ since if σ' is also measurably related to σ_1 and σ_2 it will also be so to σ, so

$$\sum_{\omega'(\sigma)} \langle\omega_1(\sigma_1)\,|\,\omega'(\sigma)\rangle F(\omega'(P))\langle\omega'(\sigma)\,|\,\omega_2(\sigma_2)\rangle$$
$$= \sum_{\omega'(\sigma),\,\omega''(\sigma'),\,\omega'''(\sigma')} \langle\omega_1(\sigma_1)\,|\,\omega''(\sigma')\rangle\langle\omega''(\sigma')\,|\,\omega'(\sigma)\rangle F(\omega'(P))\langle\omega'(\sigma)\,|\,\omega'''(\sigma')\rangle\langle\omega'''(\sigma')\,|\,\omega_2(\sigma_2)\rangle\,. \tag{29}$$

But

$$\sum_{\substack{\omega'(\sigma)\\ \omega'(P)=\omega''(P)=\omega'''(P)=\delta(\omega''-\omega''')}} \langle\omega''(\sigma')\,|\,\omega'(\sigma)\rangle\langle\omega'(\sigma)\,|\,\omega'''(\sigma')\rangle = \langle\omega''(\sigma')\,|\,\omega'''(\sigma')\rangle$$

by (22), so that the right-hand side of (29) reduces to the left-hand side of (29) with σ replaced by σ'. Thus, (27) is unambiguously defined.

We may also define the quantized form $\hat{\omega}_r$ of ω_r at P as the operator with matrix elements given by (27) when $F(\omega(P)) = \omega_r(P)$.

We have thus constructed a quantum field theory of the dynamical forms on the manifold with a (formally) satisfactory Hilbert space structure. The interpretation of the formalism is somewhat strange, however, since neither space nor time as they are usually considered are present. We could only call the quantities $\langle\Psi(\sigma_1)\,|\,\Psi(\sigma_2)\rangle$ "similarity amplitudes," the usual term "transition amplitude" not being allowable. Moreover, in a world without time the scientific method would itself seem to be lost, since the key notion of prediction cannot be made sense of. It might be claimed that the analysis presented so far has a claim to *a priori* validity since the theory being constructed is all that can be in such a general situation. That claim has some grounds for truth, but it does not seem satisfactory when compared to the usual scientific *"a posteriori"* approach. We may relate to the latter by finding under what conditions our theory reduces to that of the macroscopic world with its usual concepts of space and time. To do that we must attempt to find out how classical solutions can play any role in the theory. Hopefully such understanding will allow us to make testable predictions relevant to this macroscopic world.

VI. THE EMERGENCE OF SPACE AND TIME

The functional integral quantization scheme set up on an arbitrary manifold in the preceding section has been analyzed repeatedly in its more traditional version on a pseudo-Riemannian manifold. In particular, much investigation has been given in the case of the Minkowski metric, especially with the recent appreciation of the importance of Euclidean solutions termed instantons. Indeed, these latter have indicated the enlargement of the physical Hilbert space of states to include vacuum sectors described by pure gauge fields at infinity belonging to nontrivial homotopy classes of the gauge group under consideration. Since we are considering gauge groups which include the Lorentz group, we should expect to have vacuums described by at least two integers, since $\pi_3(O(4)) = Z \times Z$, where Z is the additive group of integers. There are, however, certain difficulties over the naive application of instanton physics to our quantum formula (21), and we will have to consider those carefully. In so doing we will appreciate certain features of (21) which clarify the problem of defining space and time by its means.

Firstly, it is usual to suppose that the local field theory being considered can be defined in a Euclidean (more generally Riemannian) space-time related to the physical Minkowski (pseudo-Riemannian) universe we inhabit by a supposed analytic continuation in time. The existence of such a continuation imposes severe constraints on the background space-time.

In the function formulation (21) this would be expressed by integrating only over suitable analytic forms ω. However, it is likely that these give a small contribution compared to more pathological ones, by analogy with the Wiener integral.[14] Thus, we expect to introduce instantons in a different manner.

Secondly, and more importantly, the multiplication of vacua arises only by certain requirements on the allowed gauge fields, in particular, that they be trivial at infinity in spatial directions. Since we have no sense of spatial or temporal directions given by (21) such conditions are meaningless. Thus the use of $\pi_3(O(4))$ may be incorrect.

We are clearly faced with the problem of defining space and time before these questions can be resolved. Since such a problem has great interest in its own right we turn to that now.

We only expect to be able to define space and time in our usual sense if we can introduce a pseudo-Riemannian metric. For only then will we have a field of cones defined on the manifold giving a separation of tangent vectors into spacelike or timelike. Thus, we have to discover how classical background fields can arise in (21). But that is easy (on the face of it): We use the method of stationary phase to write the right-hand side of (21) as a sum of contributions from classical forms $\omega^{(\mathrm{cl})}$, each contribution being of fluctuations around the associated form $\omega^{(\mathrm{cl})}$:

$$\langle \omega_1\sigma_1 | \omega_2\sigma_2\rangle \simeq \sum_{\omega^{(\mathrm{cl})}} \int \prod_{x\in R(\sigma_1\sigma_2)} d\chi(x) J^{\mathrm{cl}}(x) \exp\left[\frac{i}{h}\int_{R(\sigma_1,\sigma_2)} dA(\chi+\omega^{(\mathrm{cl})})\right]. \tag{30}$$

We have written (30) as an approximate formula, valid if the classical solutions $\omega^{(\mathrm{cl})}$ of

$$\delta dA/\delta\omega = 0 \tag{31}$$

are well separated in some suitable sense in the space of all forms on the manifold. We may then approximate each term in the sum in (30) by the Gaussian approximation, using only quadratic terms in the exponent and expanding the rest:

$$dA(\chi+\omega^{(\mathrm{cl})}) = dA(\omega^{(\mathrm{cl})}) + \chi dA''\chi + O(\chi^3)\,.$$

It is then possible to perform the resulting Gaussian integrals to obtain the standard results of quantum field theory. Each of the terms in the summation on the right-hand side of (30) corresponds to a quantum field theory in the given background $\omega^{(\mathrm{cl})}$. It will have the associated interpretational features as well as the expected ultraviolet-divergence problems especially serious for the gravitational interactions. It is possibly of value to remark here that these latter difficulties may only arise due to the approximation scheme itself. The correct expression describing quantum features of the universe is claimed to be Eq. (21); that has no immediate similarity to the separate terms of the background quantization approximation (30). We conclude that (21), not (30), is that requiring direct analysis.

For any choice of gauge group, which incorporates the Lorentz group in the manner described in the preceding sections, we may identify the classical gauge 1-form (ω_1^{cl}) coefficients of γ_a and σ_{ab} as the vierbein and connection fields of a classical gravitational field. We can then construct a traditional space-time description in terms of the associated metric.

We note that classical solutions ω^{cl} do exist. Using the above interpretation of ω_1^{cl}, as containing the vierbein and connection, all solutions of Einstein's matter-free equations qualify as possible ω^{cl}. It is clearly of interest to find the set of all ω^{cl}, but this is not yet even known for Einstein's equation.

Such a process allows space and time to be introduced for each of the possible classical solutions of (31) with the given values on the boundary of R. We seem now to have obtained a plethora of space-times, not the single one to which we are so accustomed. However, different space-times are to be expected to give different contributions to (30), so that not all such space-times need be of importance. The situation is complicated by the position dependence in the classical phase factor $A_R(\omega^{\mathrm{cl}})$. Thus in certain regions the manifold of classical solutions of importance may be different from that for other regions.

It is necessary to give a great deal of further analysis to the question of approximating classical space-time before the situation is suitably clearer. We can, however, give a general sketch of the features of the "big-bang" universe as seen in our above terms. In the early stages of the big bang (the first 10^{-43} sec) all associated regions R of the universe, being inside the Planck length, involve a great deal of quantum fluctuations. No preferred classical solution, or even family of them, would appear to be related to this part of the manifold M, so no classical notion of space or time is applicable there. We did call this region M_0 of M the "first 10^{-43} sec", above, but that is clearly just a name without any real descriptive power. Joined to M, is a sequence of regions $M_1, M_2, \ldots$ in which there should be a successively better description in terms of quantum fluctuations around a single classical metric. This means that there are regions in M_n, for large n, which can be described by quantizing on the classical background. However, even here small-enough regions could not be so described. In an ever-expanding universe the above description would be complete; if collapse ultimately occurs there must be a further region M_∞ which has the same intrinsically quantum character as M_0.

VII. CONCLUSION

We have attempted to construct a theory based on two principles, (I) that quantum mechanics is fundamental (so that there are no preferred classical fields) and (II) that of general covariance. The combination of these two necessitated the use of a nonaffinely connected four-dimensional differential manifold M as the bedrock of existence. Upon this a limited set of differential forms were allowed which could be regarded as dynamical in the sense that they satisfied nontrivial second-order differential equations in an arbitrary background field. In order to achieve such a structure it was found necessary for the forms to take values in a suitable algebra. This latter also had restrictions placed on it by the need to obtain nontrivial dynamics. That the only possibilities were quaternions or octonions in a suitable combination with elements of the de Sitter group [broken to O(4)] can be regarded as a prediction of the theory which could possibly be recognized in the particle symmetries. Identification of the octonionic structure with color and O(4) as flavor would require four "quarks" as fundamental particles. However, there has been no differentiation between electromagnetism, weak or strong interactions, nor of the introduction of coupling strengths. That must be done before firm connection with particle physics can be achieved. Such a program does not appear impossible since, as we remarked above, there are only a limited number of allowed possibilities of basic fields and interactions between them.

A very clear prediction of the theory is that there are only fundamental fields with spin up to 2. No higher is allowed, so that the discovery of a fundamental particle with higher spin would destroy the theory completely. It is necessary to be able to decide whether or not a particle is fundamental or not to make such a prediction of real value. One way of so doing is to require that it belongs to the fundamental representation of a particle symmetry which is itself fundamental (in that it cannot be constructed from underlying symmetries). Thus, the existence of a spin-$\frac{5}{2}$ quark would be very damaging to the theory.

Before that can be properly effected there are various questions to be answered. In order to use (I) properly, we had to construct a quantum mechanics for the dynamical forms. We did this in terms of states associated with the eigenvalues of the dynamical forms on given three-submanifolds of M. The resulting quantum theory had the same formal appearance as the functional integral definition of traditional quantum field theory. Its interpretation is decidedly different, however, since there is no notion of time present in the formalism. We found that quantity only arose by taking the classical limit, effectively $\hbar\rightarrow 0$. Thus, physically experienced time (and space) are to be regarded as constructs available only at a certain approximate level, and only then if the environment is correct. What, then, is the correct interpretation of the quantum-mechanical formalism in the absence of the usual notion of time? We have tried to set up a consistent answer to this question, but admit that when time is to be banished from discussion it is very difficult to make precise statements. We seem to be so thoroughly imbued with time as to find it almost impossible to think without it. Yet we have to do so if it is to be created from our formalism.

Another question of importance for our program is that of evaluation of the functional integral (21). It is different from the traditional one in being devoid of background fields. Can it be evaluated without their introduction? Even its definition may not be made sensible without the use of some symmetry breaking; since the usual problem of group invariance of the integrand has to be overcome by some such method, with the associated introduction of Faddeev-Popov ghosts. These are problems which require further analysis.

The final conclusion we reach, modulo the resolution of the above problems, is that space and time disappear as $\hbar\rightarrow\infty$ and that all events occur as manifestations of the allowed dynamics on a differentiable manifold. Thus, the question as to what happened in the first 10^{-43} sec of the universe or what will happen in the last 10^{-43} sec to matter at the center of a black hole is to be answered by the realization that time itself has lost its meaning for such events. Looked at from such small time units as 10^{-43} sec time itself is not meaningful. We must regard experience as determined by the underlying structure on M itself; the purpose of science is to discover the detailed nature of that structure. What determines such a structure on M is clearly the next question; it also clearly goes beyond the bound of the present analysis.

---

[1]For a review of progress in this area see, for example, *Quantum Gravity*, edited by C. J. Isham, R. Penrose, and D. W. Sciama (Oxford Univ. Press, London, 1975); S. Deser, P. van Nieuwenhuizen, and D. Boulware, in *Proceedings of the Seventh International Conference on General Relativity*, edited by G. Shaviv and J. Rosen (Wiley, New York, 1975).

[2]See references in the contribution of S. Deser, in Ref.

1, and also J. G. Taylor and M. Nouri-Moghadam, Proc. R. Soc. London A344, 87 (1975), A351, 197 (1976).

[3]See, for example, I. M. Singer and J. A. Thorpe, *Lecture Notes on Elementary Topology and Geometry*, CLV (Scott, Foresman and Co., Glenview, Illinois, 1967).

[4]See, for example, Ref. 3, Theorem 13, p. 134.

[5]A. Einstein, Ann. Phys. (Leipzig) 49, 769 (1916).

[6]W. Pauli, *Theory of Relativity* (Pergamon, New York, 1958), p. 150.

[7]See, for example, Ref. 3, §5.2 or S. I. Goldberg, *Curvature and Homology* (Academic, New York, 1962), Sec. 1.4.

[8]Ref. 7, S. I. Goldberg, Sec. 1.6.

[9]R. Casalbuoni, G. Domokos, and S. Kövesi-Domokos, Nuovo Cimento 31A, 423 (1976); M. J. Hayashi, SLAC Report No. SLAC-PUB-1936, 1977 (unpublished).

[10]See, for example, M. E. Meyer, *Fibre Bundle Techniques in Gauge Theories* (Springer, New York, 1977).

[11]J. G. Taylor, Phys. Rev. D 18, 3544 (1978).

[12]See, for example, R. D. Schafer, *An Introduction to Non-Associative Algebras* (Academic, New York, 1966).

[13]R. P. Feynman and A. R. Hibbs, *Quantum Mechanics and Path Integrals* (McGraw-Hill, New York, 1965).

[14]See various contributions in *Functional Integration and its Applications*, edited by A. M. Arthurs (Clarendon, London, 1975).

[15]L. D. Faddeev and V. N. Popov, Phys. Lett. 25B, 29 (1967).