International Journal of Geometric Methods in Modern Physics
Vol. 5, No. 1 (2008) 101–156

AN INTRODUCTION TO QUANTUM GRAVITY

BRYCE S. DEWITT

Department of Physics, University of Texas at Austin
Austin, Texas 78712, USA

GIAMPIERO ESPOSITO

Istituto Nazionale di Fisica Nucleare, Sezione di Napoli
Complesso Universitario di Monte S. Angelo
Via Cintia, Edificio 6, 80126 Napoli, Italy
giampiero.esposito@na.infn.it

Received 13 November 2007
Accepted 28 November 2007

After an overview of the physical motivations for studying quantum gravity, we reprint **THE FORMAL STRUCTURE OF QUANTUM GRAVITY**, i.e. the 1978 Cargèse Lectures in Levy & Deser (1979): Recent Developments in Gravitation pp. 275–322 by Professor B. S. DeWitt, with kind permission of Springer Science and Business Media. The reader is therefore introduced, in a pedagogical way, to the functional integral quantization of gravitation and Yang–Mills theory. It is hoped that such a paper will remain useful for all lecturers or Ph.D. students who face the task of introducing (resp. learning) some basic concepts in quantum gravity in a relatively short time. In the second part, we outline selected topics such as the braneworld picture with the same covariant formalism of the first part, and spectral asymptotics of Euclidean quantum gravity with diffeomorphism-invariant boundary conditions. The latter might have implications for singularity avoidance in quantum cosmology.

Keywords: Quantum gravity; diffeomorphism group; functional integral; zeta-function regularization.

1. Motivations for and Approaches to Quantum Gravity

The aim of theoretical physics is to provide a clear conceptual framework for the wide variety of natural phenomena, so that not only are we able to make accurate predictions to be checked against observations, but the underlying mathematical structures of the world we live in can also become sufficiently well understood by the scientific community. What are therefore the key elements of a mathematical description of the physical world? Can we derive all basic equations of theoretical physics from a set of symmetry principles? What do they tell us about the origin and evolution of the universe? Why is gravitation so peculiar with respect to all other fundamental interactions?

The above questions have received careful consideration over the last decades, and have led, in particular, to several approaches to a theory aimed at achieving a synthesis of quantum physics on the one hand, and general relativity on the other hand. This remains, possibly, the most important task of theoretical physics. The need for a quantum theory of gravity is already suggested from singularity theorems in classical cosmology. Such theorems prove that the Einstein theory of general relativity leads to the occurrence of spacetime singularities in a generic way [1].

At first sight one might be tempted to conclude that a breakdown of all physical laws occurred in the past, or that general relativity is severely incomplete, being unable to predict what came out of a singularity. It has been therefore pointed out that all these pathological features result from the attempt of using the Einstein theory well beyond its limit of validity, i.e. at energy scales where the fundamental theory is definitely more involved. General relativity might be therefore viewed as a low-energy limit of a richer theory, which achieves the synthesis of both the **basic principles** of modern physics and the **fundamental interactions** in the form currently known. So far, no less than 16 major approaches to quantum gravity have been proposed in the literature. In alphabetical order (to avoid being affected by our own preference) they are as follows.

1. Affine quantum gravity [2]
2. Asymptotic quantization [3, 4]
3. Canonical quantum gravity [5–9]
4. Condensed-matter view: the universe in a helium droplet [10]
5. Manifestly covariant quantization [11–17]
6. Euclidean quantum gravity [18, 19]
7. Lattice formulation [20, 21]
8. Loop space representation [22, 23]
9. Non-commutative geometry [24]
10. Quantum topology [25], motivated by Wheeler's quantum geometrodynamics [26]
11. Renormalization group and Weinberg's asymptotic safety [27, 28]
12. R-squared gravity [29]
13. String and brane theory [30–32]
14. Supergravity [33, 34]
15. Triangulations [35–37] and null-strut calculus [38]
16. Twistor theory [39, 40]

After such a broad list of ideas, we hereafter focus on what can be taught in a Ph.D. course aimed at students with a field-theoretic background. We have therefore chosen to reprint, from Sec. 2 to Sec. 17, the 1978 DeWitt Lectures at Cargèse, with kind permission of Springer Science and Business Media and of Professor C. DeWitt–Morette. In the second part we select two related topics, i.e. the braneworld picture and the spectral asymptotics of Euclidean quantum gravity, since they have possibly a deep impact on the current attempts to describe a quantum origin of the physical universe. Relevant group-theoretical material is summarized in the Appendix.

2. Introduction and Notation

In 1956 Utiyama pointed out that the gravitational field can be regarded as a non-Abelian gauge field. In 1963 Feynman found that in order to construct a quantum perturbation theory for a non-Abelian gauge field he had to introduce new graphical rules not previously encountered in quantum field theory. He showed, in one-loop order, that to preserve unitarity one must add to every standard closed-loop graph another, involving a closed integral-spin fermion loop. In 1966 an explicitly gauge invariant functional-integral algorithm was found which extended Feynman's new rules to all orders (DeWitt (1967b)). A short time later it was shown that the algorithm could be obtained by a method of factoring out the gauge group (Fadde'ev and Popov (1967)).

Formally the gravitational field and the Yang–Mills field can be treated identically. In the computation of amplitudes for specific physical processes, however, the two differ by the fact that the Yang–Mills field yields a renormalizable theory while the gravitational field does not.

Some of the proposals that have been made for dealing with quantum gravity despite its nonrenormalizability will be discussed briefly later. But it must be admitted at the outset that we are dealing with an incomplete theory. The student may take comfort in the fact that every formal statement will be true for all field theories, even those, like supergravity, possessing supergauge groups, provided they are formulated in such a way that the action of the (super)gauge group on the field variables is expressible without use of field equations, and the group operations thus given are closed.

To emphasize the generality of the formalism we shall, most of the time, suppress field symbols such as $g_{\mu\nu}$ for the gravitational field or $A^{\alpha}_{\ \mu}$ for the Yang–Mills field. The index i (or j, k, l, etc.) will be understood to label not only a field component but also a spacetime point x. Thus, in the gravitational case, i will be understood to stand for the set $\{\mu, \nu, x\}$ and, in the Yang–Mills case, for the set $\{\alpha, \mu, x\}$. In a supergauge theory the set i may include spinor indices. When it does, i (or φ^i) is said to be fermionic; otherwise it is bosonic.

The reason for including the continuous label x in the set i is that much of the formalism of quantum field theory is purely combinatorial, with summation over dummy indices being accompanied by integration over spacetime. In order to avoid having to write a lot of integral signs we lump x and the field indices together and adopt the convention that the repetition of a lower case Latin index implies a combined summation-integration. Correspondingly, a comma followed by a lower case Latin index will denote **functional** differentiation:

$$A_{,i} \equiv A \frac{\overleftarrow{\delta}}{\delta \varphi^i}. \tag{1}$$

The change in a functional A (of the fields φ^i) resulting from an infinitesimal variation $\delta\varphi^i$ is then

$$\delta A = A_{,i}\delta\varphi^i. \tag{2}$$

If A, or any of the φ^i, is fermionic one must distinguish left from right differentiation:

$$_{i,}A \equiv \frac{\overrightarrow{\delta}}{\delta\varphi^i}A,$$
$$\delta A = \delta\varphi^i\ {}_{i,}A.$$

Evidently

$$_{i,}A = (-1)^{i(A+1)}A_{,i} \tag{3}$$

where we adopt the rule that when an index (such as i) or a dynamical quantity (such as A) appears as an exponent of -1 it is to be understood as assuming the value 0 or 1 according as it is bosonic or fermionic. The summation-integration convention is not to be understood as applying to indices appearing as exponents. Such indices may participate in summation-integrations induced by their appearance twice elsewhere in an expression, but they themselves may not induce summation-integrations. Note that we are here treating all quantities ($A, \varphi^i, \delta\varphi^i$, etc.) as **supernumbers**, i.e. as even or odd elements of an infinite-dimensional Grassmann algebra. Bosonic quantities commute with everything; fermionic quantities anticommute among themselves. In the quantum theory this perfect commutativity or anticommutativity is broken. The corresponding quantities will then be written in boldface.

The following notation will sometimes be convenient for expressing repeated functional differentiation:

$$_{\dots ij,}A_{,kl\dots} \equiv \frac{\overrightarrow{\delta}}{\delta\varphi^i}\frac{\overrightarrow{\delta}}{\delta\varphi^j}A\frac{\overleftarrow{\delta}}{\delta\varphi^k}\frac{\overleftarrow{\delta}}{\delta\varphi^l}\cdots \tag{4}$$

Note the particular examples:

$$\varphi^i{}_{,j} = \delta^i{}_j, \tag{5}$$

$$(\varphi^i\varphi^j)_{,kl} = (-1)^{ij}\delta^i{}_k\ \delta^j{}_l + \delta^i{}_l\ \delta^j{}_k. \tag{6}$$

If x belongs to the set i and x' belongs to the set j, the "generalized Kronecker delta" $\delta^i{}_j$ includes, as a factor, the spacetime delta function $\delta(x, x')$.

When we are displaying specific details of a given field theory lower case indices from the middle of the Greek alphabet will be used to label tensor components. Coordinates in a given local patch, or chart, will be denoted by x^μ, with μ running from 0 to $n-1$, n being the dimensionality of spacetime. (With an eye to the ultimate application of methods such as dimensional regularization and the renormalization group, we do not hold n fixed at 4 here.) Commas followed by lower-case mid-alphabet Greek indices will denote ordinary differentiation with respect to the coordinates.

The following abbreviations will be useful:

$$\frac{\delta g_{\mu\nu}}{\delta g_{\sigma'\tau'}} = \delta_{\mu\nu}{}^{\sigma'\tau'} \equiv \frac{1}{2}(\delta_\mu{}^\sigma \delta_\nu{}^\tau + \delta_\mu{}^\tau \delta_\nu{}^\sigma)\delta(x,x'), \tag{7}$$

$$\frac{\delta A^\alpha{}_\mu}{\delta A^{\beta'}{}_{\nu'}} = \delta^\alpha{}_{\mu\beta'}{}^{\nu'} \equiv \delta^\alpha_\beta \delta_\mu{}^\nu \delta(x,x'). \tag{8}$$

It is straightforward to verify that

$$\delta_{\mu\nu}{}^{\sigma'\tau'}{}_;{}^\nu = -\delta_\mu{}^{\sigma'}{}_;{}^{\tau'} - \delta_\mu{}^{\tau'}{}_;{}^{\sigma'}, \tag{9}$$

$$\delta^\alpha{}_{\mu\beta'}{}^{\nu'}{}_;{}^\mu = -\delta^\alpha{}_{\beta';}{}^{\nu'}, \tag{10}$$

where

$$\delta_\mu{}^{\sigma'} \equiv \delta_\mu{}^\sigma \delta(x,x'), \quad \delta^\alpha_{\beta'} \equiv \delta^\alpha_\beta \delta(x,x'),$$

the semicolons denote covariant differentiation, and tensor indices may be lowered and raised by the metric tensor $g_{\mu\nu}$ and its inverse $g^{\mu\nu}$ respectively. (Note that we always leave the semicolons in the lower position regardless of what happens to the indices.) The delta functions appearing above are 2-point tensor densities, or **bitensor densities**, of total weight unity. In Eqs. (9) and (10) the apportionment of the weight between the points x and x' is arbitrary; in Eqs. (7) and (8) all the weight is at x'. In Eqs. (9) and (10) the derivatives on the left are at x, on the right at x'.

In Eqs. (8) and (10) lower case Greek indices from the beginning of the alphabet appear. These are associated with the Yang–Mills group. The laws of covariant differentiation of tensors bearing various kinds of indices are determined as follows: let the symbol T represent a tensor field, with indices suppressed, in which we imagine all the components strung out in a single column. Let T also be coupled to the Yang–Mills field. Then

$$T_{;\mu} \equiv T_{,\mu} + G^\nu{}_\sigma \Gamma^\sigma{}_{\nu\mu} T + G_\alpha A^\alpha{}_\mu T, \tag{11}$$

where

$$\Gamma^\sigma{}_{\nu\mu} \equiv \frac{1}{2} g^{\sigma\tau}(g_{\tau\nu,\mu} + g_{\tau\mu,\nu} - g_{\nu\mu,\tau}), \tag{12}$$

and the G^μ_ν and G_α are respectively the matrix generators of the representations of the linear group and Yang–Mills Lie group to which T corresponds. These generators satisfy

$$[G^\mu{}_\nu, G^\sigma{}_\tau] = \delta^\mu{}_\tau G^\sigma{}_\nu - \delta^\sigma{}_\nu G^\mu{}_\tau, \tag{13}$$

$$[G_\alpha, G_\beta] = G_\gamma f^\gamma{}_{\alpha\beta}, \tag{14}$$

where $f^\gamma{}_{\alpha\beta}$ are the structure constants of the Yang–Mills Lie group. When the suppressed indices on T are restored their positions generally determine the particular representations involved. Thus a Yang–Mills index in the upper position indicates the adjoint representation of the Lie group and one in the lower position the contragredient representation, etc. For physical reasons (positive probability) the Yang–Mills Lie group is required to be compact. Therefore given representations and their contragredient forms are equivalent, and Yang–Mills indices may be lowered (and raised) by the matrix $\gamma_{\alpha\beta}$ (and its inverse $\gamma^{\alpha\beta}$) that connects the adjoint and co-adjoint representations. When the Lie group is simple $\gamma_{\alpha\beta}$ may be taken to be the Kronecker delta, and all Yang–Mills indices may be dropped to the lower position.

We make no attempt here to list the rather complicated additional structures that appear in supergravity theories. (The student should consult the published literature for those details.) We remark only that when spinors are present the **local frame group** and its spin representations must be introduced. The local frame group is completely analogous to the Yang–Mills group and makes a corresponding contribution to the covariant derivative on the right side of Eq. (11), with $A^{\alpha}_{\ \mu}$ replaced by the connection components in the local frame and the G_α replaced by the generators of the relevant spin representation of the Lorentz group. (See De Witt (1965) for details.)

In the next section we shall show how to place the Yang–Mills group and the diffeomorphism group (with which tensor indices are associated) on a common footing. Despite the analogies between the two groups, as displayed for example by the similarity of the second and third terms on the right of Eq. (11), the diffeomorphism group is much more complicated than the Yang–Mills group and much less is known about its structure. Moreover, there is a lack of symmetry between the two groups in the fact that when they are combined into a single group (as is necessary when both Yang–Mills and gravitational fields are present) they are united not in a direct product but in a semi-direct product based on the automorphisms of the Yang–Mills group under diffeomorphisms. The same is true for the combined diffeomorphism and local frame groups, when spinor fields are present.

Our list of notational conventions is completed with the following statements and definitions:

$$T_{;\mu\nu} - T_{;\nu\mu} = -(G^{\sigma}_{\ \tau} R^{\tau}_{\ \sigma\mu\nu} + G_\alpha F^{\alpha}_{\ \mu\nu})T, \tag{15}$$

$$R^{\tau}_{\ \sigma\mu\nu} \equiv \Gamma^{\tau}_{\ \sigma\nu,\mu} - \Gamma^{\tau}_{\ \sigma\mu,\nu} + \Gamma^{\tau}_{\ \mu\rho}\Gamma^{\rho}_{\ \sigma\nu} - \Gamma^{\tau}_{\ \nu\rho}\Gamma^{\rho}_{\ \sigma\mu}, \tag{16}$$

$$F^{\alpha}_{\ \mu\nu} \equiv A^{\alpha}_{\ \nu,\mu} - A^{\alpha}_{\ \mu,\nu} + f^{\alpha}_{\ \beta\gamma} A^{\beta}_{\ \mu} A^{\gamma}_{\ \nu}, \tag{17}$$

$$R_{\mu\nu} \equiv R^{\sigma}_{\ \mu\sigma\nu}, \quad R \equiv R_{\mu}^{\ \mu}. \tag{18}$$

Unless otherwise specified we shall assume that spacetime is globally hyperbolic and complete.[a] Without loss of generality x^0 may then be assumed to be a global time coordinate, in the sense that it defines a foliation of spacetime into smooth complete hypersurfaces x^0 = constant, arranged in a temporal order. These hypersurfaces need not be everywhere spacelike, although if they are noncompact they must be asymptotically spacelike. The signature of the metric tensor will be $-+++\cdots$, and units (when needed) will be chosen to be "absolute", with $\hbar = c = 32\pi G = 1$. The

[a]There is some evidence (although hardly overwhelming yet) that quantization suppresses the singularities in spacetime that often develop automatically in classical general relativity.

absolute units of length, time and mass respectively are 1.6×10^{-32} cm, 5×10^{-43} sec and 2×10^{-6} g, which gives an idea of the domains in which quantum gravity becomes relevant.

3. The Gauge Group

The gauge group of quantum gravity is the diffeomorphism group and that of Yang–Mills theory is the Yang–Mills group. We begin with the latter.

Elements of the Yang–Mills group are locally parametrized by a set of differentiable scalar functions $\xi^\alpha(x)$, with $\xi^\alpha = 0$ denoting the identity element. Elements infinitesimally close to the identity are parametrized by infinitesimal scalars. The action of such an element on the Yang–Mills potentials is given by

$$\delta A^\alpha_{\ \mu} = -\delta\xi^\alpha_{\ ,\mu} + f^\alpha_{\ \gamma\beta} A^\beta_{\ \mu} \delta\xi^\gamma = -\delta\xi^\alpha_{\ ;\mu}, \tag{19}$$

the covariant derivative being determined by noting that $\delta\xi^\alpha$ transforms (under inner automorphisms) according to the adjoint representation of the group. It will be convenient to rewrite Eq. (19) in the form

$$\delta A^\alpha_{\ \mu} = \int Q^\alpha_{\ \mu\beta'} \delta\xi^{\beta'} d^n x', \quad d^n x' \equiv \prod_{\mu=0}^{n-1} dx^\mu, \tag{20}$$

or, in the generic notation,

$$\delta\varphi^i = Q^i_{\ \alpha} \delta\xi^\alpha, \tag{21}$$

where

$$Q^\alpha_{\ \mu\beta'} \equiv -\delta^\alpha_{\ \beta';\mu}. \tag{22}$$

In passing from Eq. (20) to the generic form (21) one replaces the labels α, μ, x by the index i and the labels β', x' by the index α, and one understands that repetition of the latter index implies a combined summation-integration.

In quantum gravity the action of the diffeomorphism group can be expressed in identical generic form. The diffeomorphism group is the group of mappings $f : M \to M$ of the spacetime manifold M into itself such that f is one-to-one and both f and f^{-1} are differentiable. In practice one may require f and f^{-1} to be C^∞ and, if the sections $x^0 =$ constant are noncompact, to reduce asymptotically (i.e. "at spatial infinity") to the local identity mapping. Such mappings define a "dragging" of all tensor fields defined on M, and if the mapping is infinitesimally close to the identity the "dragging" may be viewed as a physical displacement of the fields through an infinitesimal vector $\delta\xi$. If all fields, including the metric (i.e. gravitational) field, are displaced by the same amount the physics remains unchanged. It is conventional in physics, therefore, to adopt an opposite viewpoint and to regard an infinitesimal diffeomorphism as leaving the "physical" points of the manifold untouched while dragging all coordinate patches (i.e. the complete atlas) through the negative vector $\delta\xi$. Locally this is expressed by the coordinate transformation $x^\mu \to \xi^\mu$ where

$$\xi^\mu = x^\mu + \delta\xi^\mu. \tag{23}$$

Let T be a tensor field and δT its change under dragging through $\delta\xi$. The Lie derivative of T with respect to $\delta\xi$ is defined by

$$\mathcal{L}_{\delta\xi} T = -\delta T. \tag{24}$$

Let p be a point of M and p' the point to which it is dragged under $\delta\xi$. Then the coordinates of p in the new coordinate system (23) are identical with those of p' in the old coordinate system. Moreover, the components of T at p in the new coordinate system are identical with those of $T + \delta T$ at p' in the old coordinate system. One has only to cast Eq. (24) into component language, therefore, to regard the Lie derivative as expressing the negative of the change in the **functional form** of the components of T, **viewed as functions of the local coordinates**, under the diffeomorphism. This enables one to compute

$$\begin{aligned}\mathcal{L}_{\delta\xi} T &= T_{,\mu}\delta\xi^{\mu} - G^{\nu}{}_{\mu} T \delta\xi^{\mu}{}_{,\nu} \\ &= T_{;\mu}\delta\xi^{\mu} - G^{\nu}{}_{\mu} T \delta\xi^{\mu}{}_{;\nu},\end{aligned} \tag{25}$$

which yields, in particular, the gauge transformation law for the metric tensor:

$$\delta g_{\mu\nu} = -(\mathcal{L}_{\delta\xi} g)_{\mu\nu} = -\delta\xi_{\mu;\nu} - \delta\xi_{\nu;\mu}. \tag{26}$$

Equation (26) may be rewritten

$$\delta g_{\mu\nu} = \int Q_{\mu\nu\sigma'} \delta\xi^{\sigma'} d^n x', \tag{27}$$

$$Q_{\mu\nu\sigma'} \equiv -\delta_{\mu\sigma';\nu} - \delta_{\nu\sigma';\mu}, \tag{28}$$

which, if the labels μ, ν, x are replaced by i and the labels σ', x' by α, takes the generic form (21).

Lower case Greek indices from the first part of the alphabet, as in Eq. (21), will from now on be called **group** indices. If the group indices are allowed to label fermionic as well as bosonic gauge parameters then the generic form (21) holds also for the **super**gauge transformations of supergravity theories. Although we shall not go into the specific details of such theories we shall, in all that follows, allow for their possible presence.

4. Structure Constants

By invoking the requirement that the commutator of two infinitesimal gauge group operations be itself a group operation (the closure property) one arrives at the functional differential identity

$$Q^{i}{}_{\alpha,j} Q^{j}{}_{\beta} - (-1)^{\alpha\beta} Q^{i}{}_{\beta,j} Q^{j}{}_{\alpha} = Q^{i}{}_{\gamma} C^{\gamma}{}_{\alpha\beta}, \tag{29}$$

where the C's are certain coefficients known as the **structure constants** of the gauge group. They possess the symmetry

$$C^{\gamma}{}_{\alpha\beta} = -(-1)^{\alpha\beta} C^{\gamma}{}_{\beta\alpha}. \tag{30}$$

The structure constants of the Yang–Mills group may be determined by straightforward computation from Eqs. (19), (20) and (22). They are the components of the following 3-point tensor density:

$$C^{\alpha}_{\beta'\gamma''} = f^{\alpha}{}_{\beta\gamma}\delta(x,x')\delta(x,x''). \tag{31}$$

The weights are at x' and x''.

The structure constants of the diffeomorphism group may be determined by recalling the commutation law for the Lie derivative:

$$[\mathcal{L}_X, \mathcal{L}_Y]T = \mathcal{L}_{[X,Y]}T. \tag{32}$$

Here $[X,Y]$ is the Lie bracket of the vectors X and Y:

$$[X,Y] = \mathcal{L}_X Y = -\mathcal{L}_Y X. \tag{33}$$

The structure constants are the components of the 3-point tensor density defined by

$$\begin{aligned}\int d^n x' \int d^n x'' C^{\mu}{}_{\nu'\sigma''} X^{\nu'} Y^{\sigma''} \\ = -[X,Y]^{\mu} = X^{\mu}{}_{,\nu} Y^{\nu} - Y^{\mu}{}_{,\nu} X^{\nu} = X^{\mu}{}_{;\nu} Y^{\nu} - Y^{\mu}{}_{;\nu} X^{\nu}.\end{aligned} \tag{34}$$

Evidently

$$C^{\mu}{}_{\nu'\sigma''} = \delta^{\mu}{}_{\nu',\tau}\delta^{\tau}{}_{\sigma''} - \delta^{\mu}{}_{\sigma'',\tau}\delta^{\tau}{}_{\nu'} = \delta^{\mu}{}_{\nu';\tau}\delta^{\tau}{}_{\sigma''} - \delta^{\mu}{}_{\sigma'';\tau}\delta^{\tau}{}_{\nu'}. \tag{35}$$

The weights are at x' and x''.

The action of the gauge group on the field variables φ^i, expressed by Eq. (21), is a **realization** of the group. This realization is always a **faithful** one, which implies that $Q^i{}_{\alpha}X^{\alpha} = 0$ for all i if and only if $X^{\alpha} = 0$ for all α. By functionally differentiating Eq. (29) with respect to φ^k, multiplying by $Q^k{}_{\gamma}$, judiciously permuting the indices α, β, γ, adding the results, and invoking the faithfulness of the realization, one obtains the following cyclic identity satisfied by the structure constants:

$$C^{\delta}{}_{\alpha\epsilon}C^{\epsilon}{}_{\beta\gamma} + (-1)^{\alpha(\beta+\gamma)}C^{\delta}{}_{\beta\epsilon}C^{\epsilon}{}_{\gamma\alpha} + (-1)^{\gamma(\alpha+\beta)}C^{\delta}{}_{\gamma\epsilon}C^{\epsilon}{}_{\alpha\beta} = 0. \tag{36}$$

In the case of the Yang–Mills group this identity reduces to the corresponding identity for the constants $f^{\alpha}_{\beta\gamma}$. In the case of the diffeomorphism group it is the Jacobi identity for Lie brackets.

5. Configuration Space. Orbits

For each point x in the spacetime manifold M, the field φ (index i suppressed) takes its "value" in a certain finite-dimensional differentiable manifold Φ_x, which may but need not be a vector space or subspace thereof. In pure gravity theory, for example, Φ_x is the subspace of $\mathrm{Sym}(T_x^* \otimes T_x^*)$ containing all local symmetric covariant second rank tensors at x having nonvanishing determinant and signature $-+++\cdots$. Here T_x^* is the dual of the tangent space to M at x, and "Sym" denotes the symmetric part of the tensor product $T_x^* \otimes T_x^*$.

The set of all Φ_x with x in M, may be regarded as forming a fiber bundle over M. Each Φ_x is a fiber, and each field φ is a cross section of the bundle.[b] The bundle may but need not be a simple product bundle.

The set of all cross sections, i.e. of all field configurations φ may be assembled into a space Φ called the **configuration space**. Because of differentiability requirements on the field configurations Φ is endowed naturally with a functional differentiable structure and may be viewed as an infinite dimensional differentiable manifold, or, if φ^i includes fermion fields, as a differentiable **super**manifold (also known as a Z_2-graded manifold. See Kostant (1977)). Since M is never compact (at least in the time direction) the fields φ are usually constrained to obey also special boundary conditions "at infinity". In both Yang–Mills and gravity theory these can be of considerable importance.

Let the gauge group be denoted by G and let ξ be an element of G. Denote by ${}^{\xi}\varphi$ the "point" of Φ to which φ is displaced under the action of ξ. The set of points ${}^{\xi}\varphi$ for all ξ in G is known as the **orbit** of φ and denoted by Orb.(φ). The set of all orbits can be assembled into a space called the **space of orbits**, denoted by the quotient symbol Φ/G. Since all fields on a given orbit describe the same physics it is the space of orbits that constitutes the real **physical configuration space** of the theory. In pure gravity theory Φ is the space, Lor(M), of Lorentzian (also called pseudo-Riemannian) metrics on M, G is the group, Diff(M), of diffeomorphisms of M, and the physical configuration space, Lor(M)/Diff(M), is the space of **Lorentzian geometries** on M.

Because gauge groups can be "coordinatized" by differentiable functions (i.e. the gauge parameters) G, like Φ, can be regarded as an infinite dimensional differentiable manifold (or supermanifold). In Yang–Mills theory G may have a simple product structure inherited from the associated Lie group of the theory (see Eq. (31)), or it may itself be a twisted bundle. The diffeomorphism group of gravity theory, by contrast, cannot be viewed as a bundle but has a structure that is much less well understood. Some, but only a little, of its complexity will emerge as we go along.

Since both Φ and G are differentiable (super) manifolds the quotient space Φ/G too is a differentiable (super) manifold, or rather it is a differentiable (super) manifold that may have a boundary.

To see how a boundary can arise consider a typical, i.e. **generic**, orbit. Modulo a possible discrete center it is a **copy** of G, because it provides a realization of G and has the same dimensionality. Not all orbits need to have this dimensionality. There is often a class of degenerate orbits having fewer dimensions. These are the orbits that remain invariant under the action of nontrivial continuous subgroups of G. They are the boundary points of Φ/G. To see this think of Φ as being $\mathbf{R}^3$ and G as being the group of rotations about a fixed axis. The orbits are then circles perpendicular to

[b]If the fiber bundle admits no global cross sections the field must be defined by introducing overlapping patches.

and centered on the axis, and the orbit space is a half-plane whose boundary points correspond to the points on the axis, which remain invariant under the group.

The greater the dimensionality of the subgroup that leaves a given orbit invariant, the smaller the dimensionality of the orbit. Fischer (1970) has shown that if the invariance group has only one dimension then the orbit is an ordinary boundary point of Φ/G. If the invariance group has two dimensions then the orbit lies on a boundary of the boundary and so on. The whole orbit manifold, with its boundary, and its boundaries of boundaries, etc., is known as a **stratified** manifold.

In gravity theory the boundary orbits are the symmetrical geometries, i.e. those that possess Killing vectors. The boundary structure of $\mathrm{Lor}(M)/\mathrm{Diff}(M)$ in general depends critically on M. Since there exists no complete classification of n-dimensional manifolds ($n > 3$) that can possess globally hyperbolic metric tensors, there exists also no complete classification of possible configuration spaces for the gravitational field. On some spacetimes there may be **no** Lorentzian geometries possessing Killing vectors. Such spacetimes are called **wild**. If M is wild $\mathrm{Lor}(M)/\mathrm{Diff}(M)$ has no boundary points. For technical reasons it is frequently necessary to regard certain familiar spacetimes as wild. For example, asymptotically flat spacetimes diffeomorphic to $\mathbf{R}^n$ are usually treated as wild. The reason for this is to keep Lorentz transformations distinct from gauge transformations, by requiring the gauge parameters $\delta\xi^\mu$ to vanish at infinity. Flat Minkowski spacetime is then not a boundary point of $\mathrm{Lor}(\mathbf{R}^n)/\mathrm{Diff}(\mathbf{R}^n)$ because the Poincaré isometries are not regarded as being contained in $\mathrm{Diff}(\mathbf{R}^n)$.

6. Metrics on Configuration Space

It turns out to be both possible and useful to regard Φ and Φ/G not merely as differentiable (super)manifolds but as pseudo-Riemannian (super)manifolds as well. Let $d\varphi^i$ be an infinitesimal displacement in Φ. We may associate with this displacement a (super) arc length ds, given by

$$ds^2 = d\varphi^i \; {}_i\gamma_j \; d\varphi^j, \tag{37}$$

where ${}_i\gamma_j$ are the components of a (super) metric tensor on Φ. The ${}_i\gamma_j$ are functionals of φ having the symmetry

$${}_i\gamma_j = (-1)^{i+j+ij} {}_j\gamma_i$$

and forming an invertible continuous matrix. The inverse, denoted by γ^{ij}, satisfies

$${}_i\gamma_k \gamma^{kj} = \delta_i{}^j, \quad \gamma^{ik} {}_k\gamma_j = \delta^i{}_j, \quad \gamma^{ij} = (-1)^{ij}\gamma^{ji}. \tag{38}$$

If the metric ${}_i\gamma_j$ is chosen in such a way that the actions of G on Φ are isometries then ${}_i\gamma_j$ induces also a metric on the orbit space Φ/G. One simply defines the distance between neighboring orbits in Φ/G to be the orthogonal distance between them in Φ. This requires selecting ${}_i\gamma_j$ in such a way that the continuous matrix $(-1)^{\alpha(i+1)} Q^i{}_\alpha \, {}_i\gamma_j Q^j{}_\beta$ is nonsingular, on all orbits so that a vector cannot be simultaneously tangent to and orthogonal to any of them.

Every element of G infinitesimally close to the identity generates a vector field $Q^i{}_\alpha \delta\xi^\alpha$ on Φ (see Eq. (21)). Each of these fields is a linear combination of the basic vector fields $Q^i{}_\alpha$, and the structure of G is determined by the (super) Lie bracket relations (29) that they satisfy. The condition that G acts isometrically on Φ may be translated into the statement that the (super) Lie derivatives of the metric ${}_i\gamma_j$ with respect to the $Q's$ all vanish:

$$0 = {}_i\gamma_j \overrightarrow{\mathcal{L}}_{Q_\alpha} = {}_i\gamma_{j,k}\, Q^k{}_\alpha + (-1)^{\alpha(j+k)}{}_{i,}Q^k{}_\alpha\, {}_k\gamma_j + (-1)^{\alpha j}{}_i\gamma_k\, Q^k{}_{\alpha,j}. \tag{39}$$

It is not difficult to verify that Eq. (29) is the integrability condition for (39). Equation (39) generally has an infinity of solutions differing nontrivially from one another. If it did not, i.e. if the solution were unique up to a constant factor, this would mean that G acts transitively on Φ and hence that Φ/G is trivial, the theory having no physical content.

In order to understand what Eq. (39) says in more familiar terms it is helpful to note that the fields φ^i encountered in practice usually provide **linear** realizations of their gauge groups. This is cetainly true for the Yang–Mills and gravitational fields. What it means is that the functional derivatives $Q^i{}_{\alpha,j}$ are independent of the φ^i and, when regarded as continuous matrices (in i and j), yield a matrix **representation** of the (graded) Lie algebra associated with the group.

Of course this simplicity is generally lost if the φ's are replaced by nonlinear functions of themselves. But it is remarkable that there is usually a "natural" set of field variables of which the Q's are linear functionals. In gravity theory, in fact, there is a **family** of "natural" fields, namely all tensor densities of the form

$$\mathcal{G}^{\mu\nu} \equiv g^r g^{\mu\nu} \text{ or } \mathcal{G}_{\mu\nu} \equiv g^{-r} g_{\mu\nu}, \quad r \neq 1/n, \tag{40}$$

where $g \equiv -\det(g_{\mu\nu})$.

Consider now the way in which the Q's themselves change under infinitesimal gauge transformations. Using Eq. (29) one finds

$$\begin{aligned} \delta Q^i{}_\alpha &= Q^i{}_{\alpha,j}\delta\varphi^j = Q^i{}_{\alpha,j} Q^j{}_\beta \delta\xi^\beta \\ &= (-1)^{\alpha\beta}(Q^i{}_{\beta,j} Q^j{}_\alpha - Q^i{}_\gamma C^\gamma{}_{\beta\alpha})\delta\xi^\beta, \end{aligned} \tag{41}$$

which says that $Q^i{}_\alpha$ is a two-point function that transforms at the point associated with i according to the representation generated by the matrices $(Q^i{}_{\alpha,j})$ and at the point associated with α contragrediently to the representation generated by the matrices $(C^\alpha_{\gamma\beta})$. In gravity theory this says that the function $Q_{\mu\nu\sigma'}$ of Eq. (28) transforms like a covariant tensor at x and like a covariant vector density of unit weight at x',[c] which indeed it does. Equation (41) yields an analogous statement about the transformation law of the two-point function $Q^\alpha_{\mu\beta'}$ of Eq. (22) under the Yang–Mills group.

[c]Under the diffeomorphism group a tensor density of weight w, having p covariant and q contravariant indices, transforms contragrediently to a tensor density of weight $1 - w$, having q covariant and p contravariant indices.

We are now ready to interpret Eq. (39). Under the gauge group the γ's change according to

$$\begin{aligned}\delta\ {}_i\gamma_j &= {}_i\gamma_{j,k}\ \delta\varphi^k = {}_i\gamma_{j,k}\ Q^k{}_\alpha \delta\xi^\alpha \\ &= -(-1)^{\alpha j}[(-1)^{\alpha k}{}_{i,}Q^k{}_{\alpha\ k}\gamma_j + {}_i\gamma_k\ Q^k{}_{\alpha,j}]\delta\xi^\alpha,\end{aligned} \tag{42}$$

which says that the γ's are two-point functions that transform at each point contragrediently to the representation generated by the matrices $(Q^i{}_{\alpha,j})$. In gravity theory this implies that when the φ^i are chosen to be the components of the covariant metric tensor, ${}_i\gamma_j$ must transform at each point like a symmetric contravariant tensor density of unit weight. Any γ's that transform in this way, and have an inverse γ^{ij}, provide an acceptable metric on $\mathrm{Lor}(M)$.

Among all such metrics on $\mathrm{Lor}(M)$ there is a unique (up to a constant factor) 1-parameter family of them that may be characterized as local. These are given by

$$\gamma^{\mu\nu\sigma'\tau'} = \gamma^{\mu\nu\sigma\tau}\delta(x,x'), \tag{43}$$

$$\gamma^{\mu\nu\sigma\tau} \equiv \tfrac{1}{2}g^{1/2}(g^{\mu\sigma}g^{\nu\tau} + g^{\mu\tau}g^{\nu\sigma} + \lambda g^{\mu\nu}g^{\sigma\tau}), \quad \lambda \neq -\frac{2}{n}. \tag{44}$$

For Yang–Mills theory in flat spacetime the corresponding metric is

$$\gamma_\alpha{}^\mu{}_{\beta'}{}^{\nu'} = \gamma_{\alpha\beta}\eta^{\mu\nu}\delta(x,x'), \tag{45}$$

where $\eta^{\mu\nu}$ is the Minkowski metric. The matrices $(-1)^{\alpha(i+1)}Q^i{}_{\alpha i}\gamma_j Q^j{}_\beta$ in the two cases are readily calculated to be respectively

$$-2g^{1/2}\left\{\delta_{\mu\nu';\sigma}{}^\sigma + R_\mu^{\ \sigma}\delta_{\sigma\nu'} - (1+\lambda)[\delta(x,x')]_{;\mu\nu'}\right\} \tag{46}$$

and

$$-\delta_{\alpha\beta';\mu}{}^\mu. \tag{47}$$

The continuous matrix (47) is effectively the negative of the Yang–Mills-invariant Laplace–Beltrami operator. If the Yang–Mills field is untwisted (see the lectures by Avis and Isham in this volume) it is a nonsingular operator having a unique Green's function for each choice of boundary conditions at infinity. If the Yang–Mills field is twisted, however it may have zero eigenvalues, which means that the choice (45) fails to yield a globally valid metric on the orbit manifold. Although this is an important and interesting situation we shall not attempt to deal with it in these lectures.

The continuous matrix (46) too may become singular. Its structure is simplest when $\lambda = -1 (n \neq 2)$; it is then effectively a slightly generalized form of the standard Laplace–Beltrami operator. Considerable evidence exists to indicate that it is nonsingular when the spacetime manifold is diffeomorphic to $\mathbf{R}^n$. But for other topologies it may have zero eigenvalues. Again we exclude this situation from consideration.

When the matrices (46) and (47) are nonsingular, expressions (43) and (45) constitute metrics on the space of fields which define, by orthogonal projection, globally valid nonsingular metrics on the space of orbits. It is possible to develop a theory of geodesics on these configuration spaces. The geometry defined by Eq. (45) is flat and the geodesics in the space of Yang–Mills potentials are trivial. The geometry defined by (43) and (44), on the other hand, is not flat, and the resulting theory of geodesics on the space of metric tensors $g_{\mu\nu}$ is not trivial. It can nevertheless be shown that any pair of points in this space can be connected by a unique geodesic. It can also be shown that if a geodesic intersects one orbit orthogonally then it intersects every orbit in its path orthogonally, and, moreover, traces out a geodesic curve in the space of orbits. Methods for proving these theorems can be found in DeWitt (1967a). Using these theorems together with the fact that a vector in the space of metric tensors cannot be simultaneously parallel and orthogonal to an orbit, one can then prove that any pair of orbits can be connected by a unique geodesic. It should be stated that all of these theorems depend upon the maintenance of fixed boundary conditions (on fields and diffeomorphisms) at infinity.

7. Volume Elements on Configuration Space

With a metric defined on the space of fields it is possible to introduce a formal volume element $\mu d\varphi$ ($d\varphi \equiv \prod_i d\varphi^i$) by choosing

$$\mu = \text{const.} \times |\det({}_i\gamma_j)|^{1/2}. \tag{48}$$

This volume element is gauge invariant and can be used to define gauge invariant functional integrals over configuration space. When fermionic fields are present the determinant in Eq. (48) is the **super** determinant (see Nath (1976)), which satisfies the variational law

$$\delta \ln \det({}_i\gamma_j) = (-1)^i \gamma^{ij} \delta\, {}_j\gamma_i. \tag{49}$$

This law, combined with Eq. (39), yields the following equation of "divergenceless flow" that could in principle be used to select a gauge invariant volume element independently of a metric:

$$(-1)^{i(\alpha+1)} (\mu Q^i{}_\alpha)_{,i} = 0. \tag{50}$$

The delta functions contained in the metrics (43) and (45) give these metrics a block structure that yields simple formal expressions for their determinants. In Yang–Mills theory the determinant is a constant; in gravity theory it is given by

$$\det(\gamma^{\mu\nu\sigma'\tau'}) = \prod_x \gamma(x), \tag{51}$$

where $\gamma(x)$ is the determinant of the $\frac{1}{2}n(n+1) \times \frac{1}{2}n(n+1)$ matrix $\gamma^{\mu\nu\sigma\tau}$. It is not a difficult computation to show that

$$\gamma = (-1)^{n-1} \left(1 + \frac{n\lambda}{2}\right) g^{\frac{1}{4}(n-4)(n+1)}. \tag{52}$$

In a 4-dimensional spacetime γ, and hence $\det(\gamma^{\mu\nu\sigma'\tau'})$, is seen to be a constant, independent of the $g_{\mu\nu}$. The functional μ, in the volume element over the configuration space of gravity theory, may therefore be taken to be a constant. Without loss of generality it may be chosen equal to 1. This will no longer be true in other dimensions, or when other fields are present in addition to the gravitational field, if we stick to the $g_{\mu\nu}$ as the basic field variables. However, we can in principle replace the $g_{\mu\nu}$ by one of the family of variables defined in Eq. (40) and choose r so that μ remains constant. In practice, as we shall see later, this is unnecessary. To set μ "effectively" equal to unity it turns out to be necessary only to choose basic fields that transform linearly under the gauge group.

8. Group Coordinates

The scalar functions $\xi^\alpha(x)$ that parameterize the elements of the Yang–Mills group may be regarded as "coordinates" in the group manifold. In the case of the diffeomorphism group the group coordinates may be taken to be the functions $\xi^\mu(x)$ that define the coordinate transformation $x^\mu \to \xi^\mu$ associated with each diffeomorphism in each coordinate chart or patch. Note that the functions $\xi^\mu(x)$ are neither scalars nor components of vectors. Note also that in both cases the coordinatization of the group cannot generally be achieved without bringing in the whole apparatus of charts, atlases and consistency conditions in the regions of intersection of overlapping charts.

Any group may be regarded as acting on itself through multiplication either on the left or on the right. Every group thus provides a realization of itself, and if it is a gauge group G, possesses a set of functionals $Q^\alpha_{\ \beta}[\xi]$ analogous to the functionals $Q^i_{\ \alpha}[\phi]$ over the configuration space Φ. The $Q^\alpha_{\ \beta}$ are then defined by

$$[(I+\delta\xi)\xi]^\alpha = \xi^\alpha + Q^\alpha_{\ \beta}[\xi]\delta\xi^\beta \quad \text{for all } \delta\xi^\alpha \text{ and all } \xi \text{ in } G, \tag{53}$$

where I denotes the identity element of G and $I+\delta\xi$ denotes an element of G whose coordinates differ by infinitesimal amounts $\delta\xi^\alpha$ from those of I. Using the fact that $I^\mu(x) = x^\mu$, and that $(\xi\xi')^\mu(x) = \xi^\mu(\xi(x))$ for all ξ and ξ' in G, it is not difficult to verify that the $Q^\alpha_{\ \beta}$ for the diffeomorphism group are given explicitly by

$$Q^\mu_{\ \nu'}[\xi] = \delta^\mu_{\ \nu}\, \delta(\xi(x), x'). \tag{54}$$

In the case of the Yang–Mills group the $Q^\alpha_{\ \beta}$ take the forms

$$Q^\alpha_{\ \beta'}[\xi] = g^\alpha_{\ \beta}(\xi(x))\delta(x, x'), \tag{55}$$

where the $g^\alpha_{\ \beta}$ are the corresponding quantities for the associated Lie group.

Now let $\delta\xi^\alpha = \xi^\alpha \delta t$ for some fixed ξ^α and consider the curve in G defined by

$$\xi(t) = \lim_{\delta t\to 0}(I+\delta\xi)^{t/\delta t}. \tag{56}$$

Evidently

$$\xi(s)\xi(t) = \xi(t)\xi(s) = \xi(s+t), \quad \xi(0) = I, \quad \xi^{-1}(t) = \xi(-t), \tag{57}$$

$$\frac{d\xi^\alpha(t)}{dt} = Q^\alpha_{\ \beta}[\xi(t)]\xi^\beta. \tag{58}$$

The points on the curve are seen to constitute a one-parameter Abelian subgroup of G.

If it could be proved that all the elements of G in a neighborhood N of the identity can be obtained by a process of exponentiation of the form (56) then it would follow that the one-parameter Abelian subgroups completely span the neighborhood N. A special set of coordinates ${\xi_c}^\alpha$, known as **canonical coordinates**, could be introduced in N for which the functions $\xi^\alpha(t)$ above take the simple form

$$ {\xi_c}^\alpha(t) = \xi^\alpha t. \tag{59} $$

Let us assume that our coordinates are already canonical, so that we may drop the subscript c. Then we have

$$ I^\alpha = 0, \quad {\xi^{-1}}^\alpha = -\xi^\alpha, \tag{60} $$

$$ \xi^\alpha = Q^\alpha_{\ \beta}[\xi]\xi^\beta = {Q^{-1}}_\beta^{\ \alpha}[\xi]\xi^\beta, \tag{61} $$

where $({Q^{-1}}^\alpha_{\ \beta})$ is the (continuous) matrix inverse to $(Q^\alpha_{\ \beta})$. By taking note of the fact that the $Q^\alpha_{\ \beta}$ must satisfy an identity analogous to (29), namely

$$ Q^\gamma_{\ \alpha,\delta}\, Q^\delta_{\ \beta} - (-1)^{\alpha\beta} Q^\gamma_{\ \beta,\delta}\, Q^\delta_{\ \alpha} = Q^\gamma_{\ \delta}\, C^\delta_{\ \alpha\beta}, \tag{62} $$

we may show that in a canonical coordinate system the $Q^\alpha_{\ \beta}$ are completely determined by the structure constants.

We begin by rewriting Eq. (62) in the equivalent form

$$ {Q^{-1}}^\alpha_{\ \beta,\gamma} - (-1)^{\beta\gamma} {Q^{-1}}^\alpha_{\ \gamma,\beta} + (-1)^{\epsilon(\delta+\beta)} C^\alpha_{\ \delta\epsilon} {Q^{-1}}^\delta_{\ \beta}\, {Q^{-1}}^\epsilon_{\ \gamma} = 0. \tag{63} $$

Multiplying this equation on the right by ξ^α and using Eq. (61) we get

$$ {Q^{-1}}^\alpha_{\ \beta,\gamma}\xi^\gamma - (-1)^{\beta\gamma} {Q^{-1}}^\alpha_{\ \gamma,\beta}\xi^\gamma + C^\alpha_{\ \delta\epsilon}\xi^\epsilon {Q^{-1}}^\delta_{\ \beta} = 0. \tag{64} $$

On the other hand, differentiating Eq. (61) with respect to ξ^β we find

$$ (-1)^{\beta\gamma} {Q^{-1}}^\alpha_{\ \gamma,\beta}\xi^\gamma + {Q^{-1}}^\alpha_{\ \beta} = \delta^\alpha_{\ \beta}. \tag{65} $$

Addition of Eqs. (64) and (65) yields

$$ {Q^{-1}}_{,\alpha}\xi^\alpha + Q^{-1} - C\cdot\xi Q^{-1} = 1, \tag{66} $$

where “1” denotes the unit matrix (delta function) and

$$ Q^{-1}[\xi] \equiv ({Q^{-1}}^\alpha_{\ \beta}[\xi]), \quad C\cdot\xi \equiv (-1)^{\beta\gamma} C^\alpha_{\ \gamma\beta}\xi^\gamma = -(C^\alpha_{\ \beta\gamma}\xi^\gamma). \tag{67} $$

The solution of Eq. (66) satisfying the necessary boundary condition

$$ Q^\alpha_{\ \beta}[I] = \delta^\alpha_{\ \beta} \tag{68} $$

(see Eq. (53)) is

$$ Q^{-1}[\xi] = \frac{e^{C\cdot\xi} - 1}{C\cdot\xi} \equiv 1 + \frac{1}{2!}C\cdot\xi + \frac{1}{3!}(C\cdot\xi)^2 + \cdots\ . \tag{69} $$

The series (69) converges for all values of the ξ^α. For certain values the (continuous) matrix Q^{-1} may have vanishing roots. For these values some of the $Q^\alpha_{\ \beta}$, and hence the canonical coordinate system itself, become singular. In the case of an untwisted Yang–Mills group it can be shown that the one-parameter Abelian subgroups do span a neighborhood of the identity. (This is, in fact, a corollary of the corresponding theorem for the associated Lie group.) Indeed they span the entire group — or, rather, that part of the group that is connected to the identity, i.e. the proper group. The whole group can therefore be parameterized by canonical coordinates (supplemented, perhaps, with some discrete labels).

Canonical coordinates for the Yang–Mills group have a periodic, or angular, nature. At a given point x of the spacetime manifold let the $\xi^\alpha(x)$ in Eq. (55) increase in magnitude but maintain fixed ratios to one another. Eventually all of the $g^\alpha_{\ \beta}$ will become singular at once. One has returned to the identity element of the associated Lie group. By allowing the canonical coordinates to range from $-\infty$ to ∞ one evidently covers the gauge group an infinity of times. Despite the fact that the $Q^\alpha_{\ \beta}$ become singular for certain values of the $\xi's$, the canonical coordinates are good in that, no matter what their values, they always define a unique element of the group.

9. No Canonical Coordinates for the Diffeomorphism Group

If canonical coordinates could be introduced into the diffeomorphism group one could dispense with the apparatus of charts, atlases, etc. in parametrizing the group. Every diffeomorphism could be characterized by a (finite) vector field just as those infinitesimally close to the identity can be characterized by an infinitesimal vector field. And a vector field has a meaning independent of charts and atlases.

Unfortunately the one-parameter Abelian subgroups of the diffeomorphism group do not span a neighborhood of the identity. If the dimensionality n of the spacetime manifold M is greater than or equal to 2 there are C^∞ diffeomorphisms arbitrarily close to the identity that cannot be obtained by exponentiation as in Eq. (56). The proof, which we now outline, was first given by Freifeld (1968).

It suffices to confine attention to $\mathbf{R}^2$ or, equivalently, to the complex plane $\mathbf{C}$. Let x be a point of $\mathbf{C}$. Instead of breaking x into its real and imaginary parts we may treat x and its complex conjugate x^* formally as independent variables. A C^∞ diffeomorphism $\xi : \mathbf{C} \to \mathbf{C}$ is then a one-to-one complex function $\xi(x, x^*)$, of class C^∞ in both x and x^*, whose inverse, $x(\xi, \xi^*)$ is C^∞ in ξ and ξ^*.

Let N be a positive integer and α a positive real number. Suppose ξ has the analytic form

$$\xi(x, x^*) = e^{\frac{2\pi i}{N}} x + \alpha x^{N+1} \tag{70}$$

in a finite neighborhood of the origin (e.g., in a circle of finite radius), and suppose that outside of this neighborhood ξ changes smoothly (C^∞) to the identity function $\xi(x, x^*) = x$. If N is chosen large and α is chosen small then ξ and all its derivatives may be made uniformly close to those of the identity. We shall show that ξ does not lie on a one-parameter subgroup of C^∞ diffeomorphisms $\xi(t) : \mathbf{C} \to \mathbf{C}$ with $\xi(0) = I$.

Suppose we assume that it does lie on such a subgroup. Without loss of generality we may also assume that $\xi(1) = \xi$, and then we have

$$\xi(0, x, x^*) = x,\ \xi(1, x, x^*) = \xi(x, x^*), \tag{71}$$

as well as

$$\xi(s, \xi(t, x, x^*), \xi^*(t, x, x^*)) = \xi(t, \xi(s, x, x^*), \xi^*(s, x, x^*)) = \xi(s + t, x, x^*). \tag{72}$$

Note that the diffeomorphism (70) leaves the origin fixed. Therefore

$$\xi(0, 0, 0) = 0, \quad \xi(1, 0, 0) = 0. \tag{73}$$

Define

$$z(t) \equiv \xi(t, 0, 0). \tag{74}$$

The function $z(t)$ describes a closed curve passing through the origin in the complex plane. Using Eqs. (72) and (73) we find

$$\begin{aligned} \xi(z(t), z^*(t)) &= \xi(1, z(t), z^*(t)) = \xi(1, \xi(t, 0, 0), \xi^*(t, 0, 0)) \\ &= \xi(t, \xi(1, 0, 0), \xi^*(1, 0, 0)) = \xi(t, 0, 0) = z(t), \end{aligned} \tag{75}$$

which implies that the diffeomorphism (70) leaves every point on this closed curve fixed. But the only curve passing through the origin that (70) leaves fixed is the degenerate curve consisting of the single point $x = 0$. Therefore every one of the diffeomorphisms $\xi(t)$ must leave the origin fixed:

$$\xi(t, 0, 0) = 0 \quad \text{for all } t. \tag{76}$$

Since ξ and the $\xi(t)$ are C^∞ we may consider their formal Taylor series at the origin. The formal Taylor series for ξ, which is just expression (70), must lie on the one-parameter group of formal Taylor series for the $\xi(t)$, which may be written in the form

$$\xi(t, x, x^*) = \sum_{m,n=0}^{\infty} a_{m,n}(t) x^m x^{*\,n}. \tag{77}$$

Furthermore, these formal Taylor series must satisfy (formally) Eqs. (72).

In view of Eqs. (71) and (76) it is evident that

$$\begin{aligned} & a_{0,0}(t) = 0 \quad \text{for all } t; \\ & a_{1,0}(0) = 1, \quad \text{all other } a_{m,n}(0)'s \text{ vanish}; \\ & a_{1,0}(1) = e^{\frac{2\pi i}{N}},\ a_{N+1,0}(1) = \alpha, \quad \text{all other } a_{m,n}(1)'s \text{ vanish}. \end{aligned} \tag{78}$$

Moreover, inserting (77) into (72) with $s = t = \frac{1}{2}$, one finds

$$\begin{aligned} e^{\frac{2\pi i}{N}} x + \alpha x^{N+1} &= \sum_{m,n=0}^{\infty} a_{m,n}\left(\frac{1}{2}\right)\left[\xi\left(\frac{1}{2}, x, x^*\right)\right]^m \left[\xi^*\left(\frac{1}{2}, x, x^*\right)\right]^n \\ &= a_{1,0}\left(\frac{1}{2}\right)\left[a_{1,0}\left(\frac{1}{2}\right) x + a_{0,1}\left(\frac{1}{2}\right) x^* + \cdots\right] \\ &\quad + a_{0,1}\left(\frac{1}{2}\right)\left[a_{1,0}^*\left(\frac{1}{2}\right) x^* + a_{0,1}^*\left(\frac{1}{2}\right) x + \cdots\right] + \cdots, \end{aligned}$$

whence

$$e^{\frac{2\pi i}{N}} = \left[a_{1,0}\left(\frac{1}{2}\right)\right]^2 + |a_{0,1}(1/2)|^2, \tag{79}$$

$$0 = a_{0,1}\left(\frac{1}{2}\right) \mathrm{Re} a_{1,0}\left(\frac{1}{2}\right). \tag{80}$$

Suppose $a_{0,1}\left(\frac{1}{2}\right) \neq 0$. Then $a_{1,0}\left(\frac{1}{2}\right)$ must be pure imaginary, and the right-hand side of Eq. (79) must be a real number, which contradicts the left hand side. Therefore

$$a_{0,1}\left(\frac{1}{2}\right) = 0, \quad a_{1,0}\left(\frac{1}{2}\right) = e^{\frac{1}{2}\left(\frac{2\pi i}{N} + 2\pi i K\right)},$$

for some integer K. Repeating this reasoning for $s = t = \frac{1}{4}, s = t = \frac{1}{8}$, etc., one obtains, by continuity,

$$a_{0,1}(t) = 0 \quad \text{for all } t, \quad a_{1,0}(t) = e^{\beta t}, \quad \beta = 2\pi i\left(\frac{1}{N} + K\right). \tag{81}$$

We now have

$$\xi(t, x, x^*) = e^{\beta t} x + \sum_{m+n \geq 2} a_{m,n}(t) x^m x^{*\,n}. \tag{82}$$

Insertion of this formal series into (72) yields

$$a_{m,n}(s+t) = e^{\beta s} a_{m,n}(t) + e^{(m-n)\beta t} a_{m,n}(s), \quad m + n = 2. \tag{83}$$

This functional equation can be solved by differentiating with respect to s and setting $s = 0$:

$$\left(\frac{d}{dt} - \beta\right) a_{m,n}(t) = \dot{a}_{m,n}(0) e^{(m-n)\beta t}, \quad m + n = 2. \tag{84}$$

(Here the dot denotes the derivative.) The Green's function for the operator $\frac{d}{dt} - \beta$ appropriate to the boundary conditions (78) is

$$[\theta(t - t')\theta(t') - \theta(t' - t)\theta(-t')] e^{\beta(t - t')}$$

where θ is the step function. Use of this Green's function yields

$$a_{m,n}(t) = \frac{\dot{a}_{m,n}(0)}{(m-n-1)\beta}[e^{(m-n)\beta t} - e^{\beta t}], \quad m+n=2, \tag{85}$$

which is easily verified to satisfy (83). Now if N is large $a_{m,n}(1)$ $(m+n=2)$ must vanish in virtue of the last of Eqs. (78). But the right-hand side of Eq. (85) does not vanish at $t=1$ unless $\dot{a}_{m,n}(0)=0$. Therefore

$$a_{m,n}(t) = 0 \quad \text{for all } t \text{ when } m+n=2, \tag{86}$$

and hence

$$\xi(t,x,x^*) = e^{\beta t}x + \sum_{m+n\geq 3} a_{m,n}(t)x^m x^{*\,n}. \tag{87}$$

Inserting **this** series into (72) one gets

$$a_{m,n}(s+t) = e^{\beta s}a_{m,n}(t) + e^{(m-n)\beta t}a_{m,n}(s), \quad m+n=3, \tag{88}$$

which is identical with Eq. (83) except that now $m+n=3$. The solution is the same as before:

$$a_{m,n}(t) = \frac{\dot{a}_{m,n}(0)}{(m-n-1)\beta}[e^{(m-n)\beta t} - e^{\beta t}], \quad m+n=3. \tag{89}$$

It is now possible for the factor $m-n-1$ in the denominator to vanish, in which case this solution is replaced by its limit as $m-n\to 1$:

$$a_{m,n}(t) = \dot{a}_{m,n}(0)te^{\beta t}, \quad m-n=1. \tag{90}$$

Once again, comparing (89) and (90) with the boundary condition $a_{m,n}(1)=0$ $(m+n=3)$, one must conclude that

$$a_{m,n}(t) = 0 \quad \text{for all } t \text{ when } m+n=3. \tag{91}$$

In fact, continuing in this way one finds

$$a_{m,n}(t) = 0 \quad \text{for all } t \text{ and all } m,n \text{ with } 2\leq m+n\leq N. \tag{92}$$

One arrives finally at the case $m+n=N+1$, where one obtains

$$a_{N+1,0}(t) = \frac{1}{N\beta}\dot{a}_{N+1,0}(0)e^{\beta t}(e^{N\beta t}-1). \tag{93}$$

This expression **vanishes** at $t=1$, precisely where we do not want it to! According to Eq. (78) we must have $a_{N+1,0}(1)=\alpha$. We have thus arrived at a contradiction. Q.E.D.

Since only a small (infinitesimal) neighborhood of the origin is really involved in the above analysis, it follows that the restriction to $\mathbf{R}^2$ is not essential. For any differentiable manifold of dimension greater than or equal to 2 there exist C^∞ diffeomorphisms arbitrarily close to the identity that do not lie on one-parameter subgroups of C^∞ diffeomorphisms.

10. Gauge Conditions

A **gauge condition** is a set of constraints that picks out a subspace in the configuration space Φ, of codimension equal to the dimension of the gauge group G. The gauge condition is said to be **globally valid** if this subspace intersects each orbit in precisely one point. Such a subspace exists in the Yang–Mills case only if the gauge group corresponds to an untwisted fiber bundle. For the diffeomorphism group it probably exists if spacetime is diffeomorphic to $\mathbf{R}^n$. We confine our attention to these cases. The subspace may then be regarded as **representing** the orbit manifold Φ/G. Each orbit is represented by the point at which it intersects the subspace.

To express this idea in equations one may think of the variables φ^i as being replaced by other variables I^A, P^α, where the I^A label individual orbits and are gauge invariant, and the P^α label corresponding points **in** each orbit. The point on each orbit that is selected by the given gauge condition may be chosen as the origin of the "coordinates" P^α in that orbit. The gauge condition is then simply $P^\alpha = 0$.

It will actually prove convenient to work with the continuum of gauge conditions

$$P^\alpha[\varphi] = \zeta^\alpha, \tag{94}$$

where the ζ^α are constants (i.e. independent of the φ^i) whose values range over some preselected domain. Explicit functional forms for the P^α in terms of the φ^i may be obtained (in principle) as follows. Remembering that each (generic) orbit is a copy of G, choose the P^α to be a set of group coordinates. Since the action of the gauge group on each (generic) orbit mimics its action on itself such P's must be solutions of the functional differential equations[d]

$$P^\alpha_{,i}[\varphi]Q^i{}_\beta[\varphi] = Q^\alpha{}_\beta[P[\varphi]]. \tag{95}$$

The domain over which the ζ^α in Eq. (94) range may then be taken to be the full domain of the group coordinates.

Equation (95) does not suffice completely to determine the P^α. Additional conditions are needed to "line up" corresponding points on adjacent orbits. One possible way to do the lining up is as follows. Introduce into the configuration space Φ one of the metrics ${}_i\gamma_j$ previously discussed. Choose a generic orbit and call it the **base orbit**. Call the identity element on that orbit the **base point**. Let V be the subspace of Φ generated by the set of all geodesics emanating from the base point in directions orthogonal to the base orbit. As previously noted, these geodesics intersect all orbits in their paths orthogonally. Using the fact that every pair of points in Φ can be connected by a unique geodesic (at least in the Yang–Mills and gravitational cases) and the fact that a geodesic cannot be simultaneously orthogonal to and tangent to an orbit, one can show that V ultimately intersects all orbits. To keep it from intersecting a given orbit more than once one may terminate each of

[d]These equations are readily verified to be integrable by virtue of the identities (29) and (62).

the generating geodesics as soon as it strikes a boundary point of Φ/G. V is then topologically (but not necessarily metrically) a copy of Φ/G.

To gain an appreciation of some of the metrical situations that can arise think of Φ as being $\mathbf{R}^3$ and G as being the group of screw motions with fixed nonvanishing pitch about some axis. The orbits are then helices and all, including the axis itself, are generic. If $\mathbf{R}^3$ bears the Cartesian metric then the orbit space Φ/G is topologically but not metrically a plane. Note that in this example there exist no surfaces that intersect all orbits orthogonally, although every plane not containing the axis is perpendicular to some orbit at its intersection point and is a surface like V, based on that orbit.

Returning now to the general problem, we may place the identity element on each orbit at the point where the orbit intersects V. If another subspace V' is constructed like V but starting from another point on the base orbit, it too will intersect all the orbits. Because the group operations are isometries of ${}_i\gamma_j$, the P^α will be constants over V'. That is, once the identity points are "lined up" all the other points are automatically lined up too. The gauge condition (94) is therefore globally valid for all ζ^α in the domain G.

Unfortunately in practice it is almost hopelessly difficult to implement constructions like this one, which are guaranteed to yield globally valid gauge conditions. In the present construction, because V is generally orthogonal to none but the base orbit, one is faced with the problem of solving global functional constraints rather than functional differential equations. Even the functional differential equations that one has, namely Eq. (95), are highly nontrivial. In the Yang–Mills and gravitational cases they take respectively the forms

$$[\delta P^\alpha(x)/\delta A^\beta{}_\mu(x')]_{;\mu'} = \mathcal{G}^\alpha_\beta(P(x))\delta(x,x'), \tag{96}$$

$$2[\delta P^\mu(x)/\delta g_{\nu\sigma}(x')]_{;\sigma'} = \delta^\mu_\nu\delta(P(x),x'), \tag{97}$$

(see Eqs. (22), (28), (54) and (55)).

By far the bulk of all work on Yang–Mills theory and quantum gravity has made use of **linear** gauge conditions, i.e. conditions (94) with P^α taken in the form

$$P^\alpha[\varphi] = P^\alpha_i[\varphi_B]\phi^i, \quad \phi^i = \varphi^i - \varphi^i_B \tag{98}$$

where φ^i_B is some fiducial field, often called a **background field**. To ensure that the subspaces defined by (94) do indeed intersect the orbits uniquely, at least in the vicinity of the background field and with the ζ^α close to zero, one often makes use of the orthogonality idea by choosing

$$P^\alpha_i[\varphi_B] = (-1)^{\alpha(j+1)}Q^i{}_\alpha[\varphi_B]\,{}_j\gamma_i[\varphi_B]. \tag{99}$$

For example, if ${}_i\gamma_j$ has the form (43) then, with the choice (99), the condition $P^\alpha = 0$ becomes

$$g_B^{1/2}(2\phi_\mu{}^\nu + \lambda\delta_\mu{}^\nu\phi_\sigma{}^\sigma)_{;\nu} = 0, \quad \phi_{\mu\nu} = g_{\mu\nu} - g_{B\mu\nu}. \tag{100}$$

Here indices are raised and lowered by means of the background metric $g_{B\mu\nu}$ and the covariant derivative is defined in terms of it. With λ set equal to -1 (see the comments following Eqs. (46) and (47)) this is a very popular gauge condition in quantum gravity. The corresponding condition in Yang–Mills theory, with ${}_i\gamma_j$ given by Eq. (45), is

$$\phi_\alpha{}^\mu{}_{;\mu} = 0, \tag{101}$$

$$\phi^\alpha_\mu = A^\alpha_\mu - A_B{}^\alpha{}_\mu. \tag{102}$$

Here indices are raised and lowered by means of the metrics $\gamma_{\alpha\beta}$ and $\eta_{\mu\nu}$ and the covariant derivative is defined in terms of the background field $A_B{}^\alpha{}_\mu$. Condition (101) is known as the **Lorenz condition**.

Linear gauge conditions are extremely convenient in perturbation theory, where the field φ^i is treated as if it never gets very far from the background φ^i_B. Covariant (with respect to the background) gauge conditions like (100) and (101) are usually the best, but for some purposes noncovariant gauges (e.g., the Coulomb gauge in Yang–Mills theory) are more useful. In non-perturbative studies, however, linear gauge conditions have to be used with great care (see Gribov (1977)). At least five things can go wrong with linear gauge conditions when applied globally:

(1) The subspace defined by a linear condition may or may not have a boundary, and if it does this boundary may not coincide with the boundary (if any) of Φ/G.
(2) The subspace defined by a linear condition may intersect some orbits more than once.
(3) There may be some orbits that it does not intersect at all.
(4) Even if it intersects all orbits when the ζ^α in Eq. (94) have certain values, it may not intersect all orbits when the ζ^α have other values. This means that there is no natural domain for the ζ^α.
(5) When G is "twisted" there are no globally valid gauge conditions at all, linear or otherwise.

If any of the above situations hold, the subspace defined by (94) will not represent Φ/G faithfully. It is possible in some cases to patch things up so that the advantages of linear gauge conditions can be maintained. This has been done in certain global studies in Yang–Mills theory. However, the diffeomorphism group, as we have repeatedly emphasized, is a much more complicated group than the Yang–Mills group and both the difficulties to which it gives rise globally and the opportunities that it presents for technical innovation are almost unknown at the present time. In order to keep all options open we shall first develop the formal theory using foolproof gauge conditions, such as those based on group coordinates, and then make some remarks about how things might go when other gauge conditions are used.

11. The Action. Vertex Functions. Renormalizability

The dynamical behavior of any field is determined by its action functional S. The action functionals of (pure) Yang–Mills and gravity theories are respectively

$$S_A = -\frac{1}{4}\int F_{\alpha\mu\nu}F^{\alpha\mu\nu}d^n x, \tag{103}$$

$$S_g = 2\int g^{1/2} R d^n x. \tag{104}$$

As long as the limits of integration are not specified these integrals must be regarded as purely formal expressions that serve merely to yield the dynamical equations:

$$0 = \delta S_A/\delta A^\alpha_{\ \mu} \equiv -F_\alpha^{\ \mu\nu}{}_{;\nu}, \tag{105}$$

$$0 = \delta S_g/\delta g_{\mu\nu} \equiv -2g^{1/2}\left(R^{\mu\nu} - \frac{1}{2}g^{\mu\nu}R\right). \tag{106}$$

For some purposes, however, values need to be assigned to the actions. Integration boundaries must then be specified and, in the case of the gravitational field, a surface integral must be split off from Eq. (104) so that the integrand involves derivatives of $g_{\mu\nu}$ of order no higher than the first.

We refer the student to standard references (e.g., Misner, Thorne and Wheeler (1973)) for analyses of the initial value problems associated with Eqs. (105) and (106). From these analyses it is readily deduced that in a spacetime of n dimensions the Yang–Mills field has $n-2$ degrees of freedom per spatial point and the gravitational field has $\frac{1}{2}n(n-3)$.

In the generic notation, Eqs. (103) and (104) are written

$$S_{,i} = 0. \tag{107}$$

Gauge invariance of the theory is guaranteed by the identity

$$S_{,i}Q^i_{\ \alpha} = 0, \tag{108}$$

or, more explicitly,

$$F_\alpha^{\ \mu\nu}{}_{;\nu\mu} = 0, \quad 4\left[g^{1/2}\left(R^{\mu\nu} - \frac{1}{2}g^{\mu\nu}R\right)\right]_{;\nu} = 0. \tag{109}$$

The left hand side of Eq. (107) transforms linearly under the gauge group and hence the gauge group leaves the field equations intact. This is most easily seen by functionally differentiating Eq. (108), which yields

$$\delta S_{,i} = S_{,ij}\delta\varphi^j = S_{,ij}Q^j_{\ \alpha}\delta\xi^\alpha = -(-1)^{i\alpha}S_{,j}Q^j_{\ \alpha,i}\delta\xi^\alpha. \tag{110}$$

The action functionals (103), (104) may be expanded in functional Taylor series about a background field. In generic notation one writes

$$S = S_B + (S_{,i})_B\phi^i + \frac{1}{2!}(S_{,ij})_B\phi^j\phi^i + \frac{1}{3!}(S_{,ijk})_B\phi^k\phi^j\phi^i + \cdots, \tag{111}$$

$$\phi^i = \varphi^i - \varphi^i_B.$$

If the background fields satisfy the classical field equations then the second term on the right may be omitted.

The functional derivatives $(S_{,i_1\cdots i_N})_B$ with $N \geq 3$ are known as (bare) **vertex functions**. In the case of the Yang–Mills field the vertex functions with $N > 4$ vanish and the Taylor series terminates. In the case of the gravitational field the series may or may not terminate depending on what choice is made for the basic field variables. By expressing inverse matrices in terms of minors and determinants, and by examining the number of determinants needed to yield unit total weight for the integrand of (104), one easily verifies that if the basic field variables are taken to be $\mathcal{G}^{\mu\nu} \equiv g^r g^{\mu\nu}$ and r is chosen to be $5/(4n+2)$ then the vertex functions with $N > 2n+1$ vanish. Alternatively, if $\mathcal{G}_{\mu\nu} = g^{-r} g_{\mu\nu}$ are chosen as the basic field variables, with $r = 5/(6n-2)$, then the vertex functions with $N > 3n-1$ vanish.[e] There are three reasons, however, why neither of these choices is useful. First, the vertex functions of gravity theory are exceedingly complicated, involving thousands of terms already for $N = 4$. Nobody is going to work out the vertex functions up to maximum order even with the aid of a computer. Second, any imagined advantage in these choices is lost as soon as one tries to introduce dimensional regularization into the quantum theory. A specific choice of field variables has then to be made, and it cannot vary continuously with the dimension. Third, although a series that terminates has an infinite radius of convergence, the range of the variables $\phi^{\mu\nu} \equiv \mathcal{G}^{\mu\nu} - \mathcal{G}_B^{\mu\nu}$ or $\phi_{\mu\nu} \equiv \mathcal{G}_{\mu\nu} - \mathcal{G}_{B\mu\nu}$ is in fact limited. These variables must avoid regions where the signature of the metric tensor changes.

The third reason is the most important, at least in perturbation theory. As is well known, the Feynman rules are obtained by inserting the expansion (111) into the Feynman functional integral (see the next section) and evaluating the integral as a sum (asymptotic series) of Gaussian integrals, with the ϕ^i ranging from $-\infty$ to ∞. Any constraint on the ϕ^i would make these integrals almost impossible to evaluate, and although one may for some purpose wish to extend the Feynman integrand into nonphysical regions, one never does this by naively removing constraints.

These remarks suggest, in fact, that none of the variables (40) is good to use in perturbation theory. A better choice would be something like

$$\begin{gathered} \phi \equiv [\ln(g\eta^{-1})]\eta, \quad g = e^{\phi\eta^{-1}}\eta, \quad \phi \equiv (\phi_{\mu\nu}), \quad g \equiv (g_{\mu\nu}), \\ \eta \equiv (\eta_{\mu\nu}), \quad \eta^{-1} \equiv (\eta^{\mu\nu}), \end{gathered} \tag{112}$$

which maintains the signature of the spacetime metric. With these variables the series (111), of course, does not terminate, and one speaks of gravity theory as being a **non-polynomial Lagrangian theory** (Isham, Strathdee and Salam (1971), (1972)). It will be noted that all such "safe" variables inevitably transform **non-linearly** under the diffeomorphism group.

Regardless of the choice of variables it is not difficult to draw preliminary conclusions about the renormalizability or nonrenormalizability (in perturbation theory)

[e]Both of these choices require $n \neq 2$. (See Eqs. (40).)

of a given quantum field theory. Although momentum space is not, in an absolute sense, appropriate for use in quantum gravity, conclusions about the high energy behaviour of amplitudes in perturbation theory may be safely drawn with its aid. Consider a Feynman graph with L_e external lines, L_i internal lines, and V_N Nth-order vertices ($N \geq 3$). L_e, L_i and V_N are related by the topological condition

$$L_e + 2L_i = \sum_N NV_N. \tag{113}$$

The number of independent closed loops, or momentum integrations, in the graph is given by

$$I = L_i - \sum_N V_N + 1. \tag{114}$$

In quantum gravity 2 powers of momentum are associated with each vertex, -2 powers with each internal line, and n powers with each momentum integration. The superficial degree of divergence of the graph is therefore

$$D = -2L_i + 2\sum_N V_N + nI = (n-2)I + 2, \tag{115}$$

which, for $n > 2$, increases without limit as the number of independent closed loops increases. This means that for $n > 2$, there is an infinite number of primitive divergences, and, if one attempts to compute order by order, an infinite number of experimentally determined coupling constants is needed to determine the theory. These conclusions are not altered if account is taken of the "ghost" contributions which, as we shall see in the following sections, must be included. The theory is said to be nonrenormalizable.

In Yang–Mills theory, in contrast, only one power of momentum is associated with each 3rd-order vertex, and the 4th-order vertices have no momentum dependence at all. This leads to

$$D = -2L_i + \sum_N (4-N)V_N + nI = 4 + (n-4)I - L_e. \tag{116}$$

For $n > 4$ this theory too is nonrenormalizable, but for $n = 4$ and an arbitrary background field there are only four primitive divergences (corresponding to $L = 1, 2, 3, 4$), and the theory is renormalizable. The proof of renormalizability is not trivial and depends crucially on gauge invariance as well as some of the formal developments to be discussed in the following sections. The primitive divergences turn out to be related in virtue of gauge invariance.

The nonrenormalizability of standard quantum gravity has stimulated investigations of alternative theories in which terms of the form $g^{1/2}(\alpha R^2 + \beta R_{\mu\nu}R^{\mu\nu})$ are added to the integrand of expression (104). Such theories generally suffer from "physical ghosts" with negative probabilities, but they do improve the convergence situation. Each vertex now carries 4 powers of momentum and each internal line carries -4 powers. This leads to

$$D = -4L_i + 4\sum_N V_N + nI = (n-4)I + 4. \tag{117}$$

When $n = 4$ all diagrams have the same superficial degree of divergence, namely 4. There is an infinity of primitive divergences, but they are all related by gauge invariance, and only three experimental coupling constants are required. The proof of renormalizability has been carried out by Stelle (1977) using methods similar to those applied to Yang–Mills theory.

It will be observed, by virtue of Eq. (115), that the same methods should work in the case of standard quantum gravity when $n = 2$, although since the number of degrees of freedom in the field is then **negative** it is not clear what such a theory means. Weinberg (1979) has studied the asymptotic stability of quantum gravity when $n = 2+\epsilon, \epsilon \ll 1$, and has given plausibility arguments concerning its relevance for trying to make sense out of the theory when $n = 4$.

In addition to its ultraviolet divergences quantum gravity also possesses infrared divergences. Gravitational field quanta **-gravitons-** are massless. In fact this itself need not lead to difficulties worse than those encountered in quantum electrodynamics where the divergences are completely understood and are removable by standard methods. Gravitons, however, are coupled to other massless quanta (photons, neutrinos, etc.) as well as to themselves. In Yang–Mills theory as well as in massless electrodynamics such a situation gives rise to infrared divergences of a new type that cannot be removed by standard techniques or argued away on physical grounds. In quantum gravity these new divergences are miraculously absent (Weinberg (1965) and DeWitt (1967c)). It appears therefore that the mysteries of Yang–Mills theory and gravity theory lie at opposite ends of the momentum spectrum. There is an increasing body of evidence that the Yang–Mills field solves its infrared dilemma by adopting a nonstandard behaviour at long wavelengths, which is intimately related to the phenomena of quark confinement and dynamical symmetry breaking. These phenomena may also bear a technical relation to the failure of gauge conditions to be globally valid in Yang–Mills theory (Gtibov (1977)). No analogous phenomena are known to exist for gravity, at least when spacetime is diffeomorphic to $\mathbf{R}^n$. The mysteries of gravitation theory thus appear to lie solely at the high end of the momentum spectrum.

12. The Feynman Functional Integral. Factoring Out the Gauge Group

Consider a transition amplitude of the form $\langle \text{out}|\text{in}\rangle$ where the vectors $|\text{in}\rangle$ and $|\text{out}\rangle$ refer to states in which the field is maximally specified (in the quantum mechanical sense, e.g., in terms of complete sets of commuting observables) in regions "in" and "out" respectively. These states need not be "vacuum" states and the regions "in" and "out" need not refer to the infinite past and future respectively. If the background field (which enters naturally in most calculations) has singularities in the past and/or future, $|\text{in}\rangle$ and $|\text{out}\rangle$ may be defined not in terms of observables at all but by some analytic continuation procedure (e.g., to the "Euclidean sector") that removes the singularities. It will be assumed only that the "in" and "out" regions lie respectively to the past and future of the region of dynamical interest.

There are many ways of showing that the amplitude $\langle\text{out}|\text{in}\rangle$ can be expressed as a formal functional integral:

$$\langle\text{out}|\text{in}\rangle = N \int e^{iS[\varphi]}\mu[\varphi]d\varphi, \; d\varphi \equiv \prod_i d\varphi^i. \tag{118}$$

Here N is a normalization constant, $S[\varphi]$ is the classical action functional, $\mu[\varphi]$ is chosen to make the "volume element" $\mu d\varphi$ gauge invariant (see Eq. (50)), and the integration is to be extended over all fields φ that satisfy the boundary conditions appropriate to the given "in" and "out" states. We have remarked earlier (and will show later) that μ may be set "effectively" equal to unity if the φ^i are chosen to transform linearly under the gauge group. We shall assume that such a choice has been made and henceforth drop μ from the theory. (It can always be restored if desired.)

Expression (118) was first derived by Feynman (1948) in ordinary quantum mechanics, without gauge groups, and later (1950) applied by him to field theory. The full extension to field theories with gauge groups is the work of many people, and the student is referred to the literature for details.[f] When the Feynman integral is applied to the gravitational field the only additional comment that needs to be made is that the integration may have to embrace as many topologies as can be reached by analytic continuation from the given background topology.

If any of the fields φ^i in (118) are fermionic the integration with respect to them is to be carried out according to the formal rules for integrating with respect to anticommuting variables that were first introduced by Berezin (1966). These rules are analogous in many ways to the well known rules for ordinary definite integrals from $-\infty$ to ∞ with integrands that vanish asymptotically. For example, integrals of total derivatives vanish, and the position of the zero point may be shifted. On the other hand, with Berezin rules, transformations of variables and evaluation of Gaussian integrals lead to determinants precisely inverse to those of standard theory. When both bosonic and fermionic fields are involved it is the **super** determinant that appears.

All physical amplitudes can be deduced from expression (118) by examining how $\langle\text{out}|\text{in}\rangle$ changes under variations in the action. Physical amplitudes can alternatively be obtained by judicious use of

$$\langle\text{out}|T(A[\varphi])|\text{in}\rangle = N \int A[\varphi]e^{iS[\varphi]}d\varphi, \tag{119}$$

[f]Useful modern references are Fadde'ev (1969), (1976) and Abers and Lee (1973).

where $A[\phi]$ is any functional of the field **operators** φ^i, and the T symbol removes ambiguities about ordering the φ^i by arranging them chronologically (with appropriate $\pm$ signs thrown in if any of the φ^i are fermionic).

When a gauge group is present the integration in Eq. (118) is redundant. Furthermore Eq. (119) is generally ambiguous, unless A is gauge invariant, in which case the integration in Eq. (119) too is redundant. This is because, owing to the gauge invariance of the classical action, the exponent in the integrands of Eqs. (118) and (119) remains constant as φ ranges over a group orbit in the configuration space Φ. One can remove this redundancy and/or ambiguity by adopting a gauge condition like Eq. (94). The details of the procedure were first given by Fadde'ev and Popov (1967).

Let ξ be an element of the gauge group G, with coordinates ξ^α, and let ${}^\xi\varphi$ be the field to which φ is displaced under the action of ξ. Define

$$\triangle[\zeta,\varphi] \equiv \int_G \delta[P[{}^\xi\phi] - \zeta]\mathrm{det}Q^{-1}[\xi]d\xi, \quad d\xi \equiv \prod_\alpha d\xi^\alpha, \tag{120}$$

where $\delta[\ ]$ is the delta **functional**, P^α are the functionals appearing in Eq. (94), Q^{-1} is the inverse of the matrix formed out of the Q^α_β of Eq. (53), and the integration extends over the entire gauge group! We shall assume that the gauge condition (94) is globally valid. The integrand in Eq. (120) then "switches on" at only one point in G, namely that point for which ${}^\xi\varphi$ is equal to the unique field φ_ζ lying on the orbit containing φ and picked out by the gauge condition:

$$P^\alpha[\varphi_\zeta] = \zeta^\alpha. \tag{121}$$

By building infinitesimal parallelepipeds in the group manifold G and making use of Eq. (53) one can verify that the combination $\mathrm{det}Q^{-1}[\xi]d\xi$ appearing in Eq. (120) is a right-invariant volume element, satisfying

$$\mathrm{det}Q^{-1}[\xi\xi']d(\xi\xi') = \mathrm{det}Q^{-1}[\xi]d\xi \quad \text{for all } \xi' \text{ in } G. \tag{122}$$

The presence of this volume element renders the functional gauge invariant:

$$\triangle[\zeta, {}^{\xi'}\varphi] = \triangle[\zeta,\varphi] \quad \text{for all } \xi' \text{ in } G. \tag{123}$$

Several comments must be made about its use, however. In the case of the diffeomorphism group, with the Q's given by Eq. (54), it is easily checked that

$$Q^{-1\mu}{}_{\nu'}[\xi] = \delta^\mu{}_\nu \delta(x,\xi(x'))\frac{\partial(\xi(x'))}{\partial(x')}. \tag{124}$$

No one has ever discovered how to evaluate or give a meaning to the determinant of this continuous matrix. The right-invariant volume element of the diffeomorphism group, therefore, can only be defined (and used) purely formally. The same is true for the invariance group of supergravity theory. It should be noted that when the group is a supergauge group, possessing anticommuting as well as commuting coordinates, det Q^{-1} is a superdeterminant, and the integral (120) involves the Berezin rules.

(Remark: the delta functional in Eq. (120) presents no difficulty. Delta functions of anticommuting variables turn out to be easy to define. They can even be given Fourier representations.)

The gauge invariance of $\triangle$ makes it an easy functional to evaluate. One has only to shift φ to φ_ζ so that the integrand in Eq. (120) switches on at the identity element I. All quantities can then be expanded in power series in $\xi^\alpha - I^\alpha$. For example, the argument of the delta functional becomes

$$\begin{aligned} P^\alpha[{}^\xi\varphi_\zeta] - \zeta^\alpha &= P^\alpha[\varphi_\zeta] - \zeta^\alpha + P^\alpha_{,i}[\varphi_\zeta]Q^i{}_\beta[\varphi_\zeta](\xi^\beta - I^\beta) + \cdots \\ &= F^\alpha_\beta[\varphi_\zeta](\xi^\beta + I^\beta) + \cdots \end{aligned} \tag{125}$$

where

$$F^\alpha_\beta[\varphi] \equiv P^\alpha_{,i}[\varphi]Q^i{}_\beta[\varphi]. \tag{126}$$

Similarly, making use of Eq. (68), we find

$$Q^{-1\alpha}{}_\beta[\xi] = \delta^\alpha_\beta - Q^\alpha_{\beta,\gamma}[I](\xi^\gamma - I^\gamma) + \cdots \tag{127}$$

and hence

$$\begin{aligned} \triangle[\zeta,\varphi] &= \int_G \delta[F[\varphi_\zeta](\xi - I) + \cdots][1 - (-1)^\alpha Q^\alpha_{\alpha,\beta}[I](\xi^\beta - I^\beta) + \cdots]d\xi \\ &= (\det F[\varphi_\zeta])^{-1}, \end{aligned} \tag{128}$$

F being the matrix with elements F^α_β. If any of the group indices is fermionic the determinant is again a superdeterminant. Note that if the P's are constructed according to Eq. (95), $F[\varphi]$ is identical to the matrix $Q[P[\varphi]]$. Although this construction will not be assumed in what follows, we shall, for simplicity and convenience, assume that the gauge condition (94) is globally valid for all ζ^α lying in the ranges of the functionals $P^\alpha[\varphi]$.

The next step is to insert unity into the integrand of Eq. (118), in the guise of

$$(\triangle[\zeta,\varphi])^{-1}\int_G \delta[P[{}^\xi\varphi] - \zeta]\det Q^{-1}[\xi]d\xi,$$

and interchange the order of integrations, obtaining

$$\langle \text{out}|\text{in}\rangle = N\int_G \det Q^{-1}[\xi]d\xi \int d\varphi e^{iS[\varphi]}(\triangle[\zeta,\varphi])^{-1}\delta[P[{}^\xi\varphi] - \zeta]. \tag{129}$$

We have assumed a choice of variables for which the volume element $d\varphi$ is gauge invariant ($\mu = 1$). $S[\varphi]$ and $\triangle[\zeta,\varphi]$ are also gauge invariant. Therefore a superscript ξ may be affixed to every φ in the integrand of (129) that does not already bear one. But every ${}^\xi\varphi$ is then a dummy, and hence all the ξ's may be removed. Making use of Eq. (128) one immediately obtains

$$\langle \text{out}|\text{in}\rangle = N' \int e^{iS[\varphi]} \det F[\varphi]\delta[P[\varphi] - \zeta], \tag{130}$$

where $F[\varphi_\zeta]$ has been replaced by $F[\varphi]$ in the integrand because of the presence of the delta functional, and where

$$N' \equiv N \int_G \det Q^{-1}[\xi] d\xi. \tag{131}$$

The gauge group has now been factored out, and its "volume" has been absorbed into the new normalization constant N'. The integration in Eq. (130) is restricted to the subspace $P^\alpha[\varphi] = \zeta^\alpha$.

The technique of confining the fields φ^i to a particular subspace can also be used to remove the ambiguity from the integral (119) when $A[\varphi]$ is not gauge invariant. Strictly speaking, matrix elements are definable only for gauge invariant operators. However, given a non-gauge-invariant operator $A[\underline{\varphi}]$, one can construct a gauge invariant operator out of it by the following definition:

$$T(A[\underline{\varphi}_\zeta]) \equiv T\left((\triangle[\zeta, \underline{\varphi}])^{-1} \int_G A[{}^\xi\underline{\varphi}]\delta[P[{}^\xi\underline{\varphi}] - \zeta] \det Q^{-1}[\xi] d\xi \right). \tag{132}$$

The chronological ordering symbol is used here so that the non-commutativity (or anti-commutativity) of $A[{}^\xi\underline{\varphi}]$ with both $(\triangle[\zeta, \underline{\varphi}])^{-1}$ and the delta functional can be effectively ignored. Note that because the gauge group acts linearly on the $\underline{\varphi}$'s there is no ambiguity about the symbol ${}^\xi\underline{\varphi}$. Note, however, that diffeomorphisms in gravity theory can drag the field in very complicated ways. The chronological operation, which orders field operators solely by the value of the coordinate x^0, rearranges the "physical" fields in correspondingly complicated ways as the variable ξ in the integral (132) ranges over the group.

Applying Eq. (119) to the operator $T(A[\underline{\varphi}_\zeta])$ and following the same reasoning as was used in passing from Eq. (129) to Eq. (130), one finds

$$\langle \text{out}|T(A[\underline{\varphi}_\zeta])|\text{in}\rangle = N' \int A[\varphi] e^{iS[\varphi]} \det F[\varphi]\delta[P[\varphi] - \zeta] d\varphi, \tag{133}$$

valid for any functional $A[\varphi]$.

13. Averaging Over Gauges

It is possible to develop a perturbation theory based on Eqs. (130) and (133), but it is usually more convenient to work with a formalism from which the delta functionals have been eliminated. Note that although the parameters ζ^α appear on the right side of Eq. (130), the amplitude $\langle \text{out}|\text{in}\rangle$ is actually independent of them. Therefore nothing changes if we integrate over these parameters, with a weight factor.

In practically all studies of non-Abelian gauge theories to date, Gaussian weight factors of the form

$$\exp\left(\frac{1}{2}i\zeta^\alpha\ {}_\alpha M_\beta\ \zeta^\beta\right),$$

where M is a nonsingular constant matrix having the symmetry

$${}_\alpha M_\beta = (-1)^{\alpha+\beta+\alpha\beta}\,{}_\beta M_\alpha,$$

have been used. From a fundamental standpoint a Gaussian weight factor can be used only if the bosonic ζ's can range from $-\infty$ to ∞ without the gauge condition (94) becoming globally invalid — for example, if the P's satisfy Eq. (95), with $Q^\alpha_{\ \beta}$'s based on canonical group coordinates (Eq. (69)). This condition, however, is almost universally violated. Indeed, the Gaussian weight function is most frequently employed in combination with linear gauge conditions like Eq. (98), where it almost certainly introduces errors globally (i.e. in non-perturbative analyses).

In the case of quantum gravity, where the gauge group has no canonical coordinates, it seems particularly inappropriate to confine our attention to Gaussian weight factors. We shall therefore introduce a more general weight factor, of the form $\exp(iU[\zeta])$, where we specify nothing about the functional $U[\zeta]$ except the following three conditions:

(1) U becomes infinite on the boundary of the allowable domain of the ζ's (which domain we are assuming coincides with the range of the $P[\varphi]$'s).
(2) U and all its first functional derivatives $U_{,\alpha}$ vanish at some chosen point (e.g., at $\zeta^\alpha = I^\alpha$ when the P's are group coordinates); its second functional derivatives $U_{,\alpha\beta}$, however, form a nonsingular continuous matrix at that point.
(3) U vanishes nowhere else, and its first derivatives all vanish simultaneously nowhere else.

The third condition is imposed mainly for convenience. Note that all three are satisfied by the Gaussian exponent $\frac{1}{2}\zeta^\alpha\ {}_\alpha M_\beta\ \zeta^\beta$ whenever it can be legitimately used.

Inserting $\exp(iU[\varphi])$ into the integrand of Eq. (130) and integrating over the ζ's, one obtains

$$\langle \mathrm{out}|\mathrm{in}\rangle = N''[U]\int e^{i(S[\varphi]+U[P[\varphi]])}\mathrm{det}F[\varphi]d\varphi, \tag{134}$$

$$N''[U] \equiv \frac{N'}{\int e^{iU[\zeta]}d\zeta}, \tag{135}$$

the integration domain in Eq. (135) being understood to be the allowable domain of the ζ's. Equation (133) too may be replaced by a weighted average. Defining

$$T(A[\varphi]) \equiv \left(\int e^{iU[\zeta]}d\zeta\right)^{-1}\int T(A[\varphi_\zeta])e^{iU[\zeta]}d\zeta, \tag{136}$$

one may write

$$\langle \text{out}|T(A[\varphi])|\text{in}\rangle = N''[U]\int A[\varphi]e^{i(S[\varphi]+U[P[\varphi]])}\det F[\varphi]d\varphi. \tag{137}$$

Equation (137), and generalizations of it, will be used frequently in the following sections. Definitions (132) and (136) reveal precisely what kind of averaged quantum operator is associated with each classical functional $A[\varphi]$ in this formalism. Note that if $f[\zeta]$ is any functional of the ζ's we have

$$T(f[P[\varphi_\zeta]]) = f[\zeta] \tag{138}$$

and

$$\langle \text{out}|T(f[P[\varphi]])|\text{in}\rangle = \langle f\rangle\langle \text{out}|\text{in}\rangle \tag{139}$$

where

$$\langle f\rangle \equiv \frac{\int f[\zeta]e^{iU[\zeta]}d\zeta}{\int e^{iU[\zeta]}d\zeta}. \tag{140}$$

Having so freely manipulated formal expressions we should now check that no inconsistencies have crept into our results, by verifying directly that the right side of Eq. (134), for example, is truly independent of the choices we have made for the functionals $P^\alpha[\varphi]$ and $U[\zeta]$. Obviously the right side **will** be affected if we naively switch to P's for which the gauge condition (94) is no longer globally valid for all ζ's in the range of the P's. Therefore we must assume that the changes δP^α (which, without loss of generality, may be taken infinitesimal) maintain global validity.

We also confine our attention to changes δU that leave the location of the zero of U, as well as the three conditions that we imposed upon U, intact. It is not difficult to see that δU may then always be expressed in the form

$$\delta U[\zeta] = U_{,\alpha}[\zeta]\delta V^\alpha[\zeta], \tag{141}$$

where the δV^α vanish at the zero of U. Note that under this change we have

$$\delta N''[U] = -iN''[U]\frac{\int e^{iU[\zeta]}U_{,\alpha}[\zeta]\delta V^\alpha[\zeta]d\zeta}{\int e^{iU[\zeta]}d\zeta} = N''[U]\langle(-1)^\alpha\delta V^\alpha_{\ ,\alpha}\rangle, \tag{142}$$

the final form being obtained by an integration by parts in which the boundary of the integration domain contributes nothing because $\exp(iU[\zeta])$ oscillates infinitely rapidly there.

Making use of Eqs. (126), (139), (141) and (142) we now have

$$\begin{aligned}\delta\langle \text{out}|\text{in}\rangle = N''[U]\int & e^{i(S[\varphi]+U[P[\varphi]])}\{(-1)^\alpha\delta V^\alpha_{\ ,\alpha}[P[\varphi]] \\ & + iU_{,\alpha}[P[\varphi]](\delta V^\alpha[P[\varphi]] + \delta P^\alpha[\varphi]) \\ & + (-1)^\alpha F^{-1\alpha}{}_\beta[\varphi]\delta P^\beta_{\ ,i}[\varphi]Q^i{}_\alpha[\varphi]\}\det F[\varphi]d\varphi,\end{aligned} \tag{143}$$

where the inverse F^{-1}, if it is a Green's function (as it often will be), must satisfy the boundary conditions appropriate to the "in" and "out" states. The integral

(143) does not obviously vanish. The way to show that it is nevertheless zero is as follows. Replace each φ^i in the integral (134) by $\overline{\varphi}^i$, where

$$\overline{\varphi}^i = \varphi^i + Q^i{}_\alpha[\varphi]\delta\xi^\alpha[\varphi], \tag{144}$$

$$\delta\xi^\alpha[\varphi] = F^{-1\alpha}{}_\beta[\varphi](\delta V^\beta[P[\varphi]] + \delta P^\beta[\varphi]). \tag{145}$$

Since the φ's are just dummies this replacement has no effect. However, it is not difficult to show that the net apparent change in the integral is given precisely by Eq. (143) **provided** one is entitled to make the identifications

$$(-1)^{i(\alpha+1)}Q^i{}_{\alpha,i} = 0, \quad (-1)^{\beta(\alpha+1)}C^\beta{}_{\alpha\beta} = 0. \tag{146}$$

We shall comment on these equations presently.

It is easy to see that the second term inside the curly brackets in Eq. (143) comes from the change that the replacement $\varphi \to \overline{\varphi}$ induces in the exponent of Eq. (134). That the first and third terms come from the change in the product $\det F[\varphi]d\varphi$ may be shown as follows. First compute

$$\begin{aligned}\overline{\phi}^i{}_{,j} &= \delta^i{}_j + (-1)^{j\alpha}Q^i{}_{\alpha,j}[\varphi]\delta\xi^\alpha[\varphi]\\ &\quad - Q^i{}_\alpha[\varphi]F^{-1\alpha}{}_\beta[\varphi]((-1)^{jk}P^\beta_{,kj}[\varphi]Q^k{}_\gamma[\varphi] + (-1)^{j\gamma}P^\beta_{,k}[\varphi]Q^k{}_{\gamma,j}[\varphi])\delta\xi^\gamma[\varphi]\\ &\quad + Q^i{}_\alpha[\varphi]F^{-1\alpha}{}_\beta[\varphi](\delta V^\beta_{,\gamma}[P[\varphi]]P^\gamma_{,j}[\varphi] + \delta P^\beta_{,j}[\varphi]),\end{aligned}$$

which, after rearrangement of some factors and use of Eq. (126), yields the (super) Jacobian

$$\begin{aligned}\det(\overline{\phi}^i{}_{,j}) &= 1 + (-1)^{i(\alpha+1)}Q^i{}_{\alpha,i}[\varphi]\delta\xi^\alpha[\varphi]\\ &\quad - F^{-1\alpha}{}_\beta[\varphi]((-1)^{\alpha(j+1)}P^\beta_{,ji}[\varphi]Q^i{}_\alpha[\varphi]Q^j{}_\gamma[\varphi]\\ &\quad + (-1)^{\alpha(\gamma+1)}P^\beta_{,j}[\varphi]Q^j{}_{\gamma,i}[\varphi]Q^i{}_\alpha[\varphi])\delta\xi^\gamma(\varphi)\\ &\quad + (-1)^\alpha\delta V^\alpha_{,\alpha}[P[\varphi]] + (-1)^\alpha F^{-1\alpha}{}_\beta[\varphi]\delta P^\beta_{,i}[\varphi]Q^i{}_\alpha[\varphi].\end{aligned} \tag{147}$$

Combining $\delta d\varphi \equiv d\overline{\varphi} - d\varphi = [\det(\overline{\varphi}^i{}_{,j}) - 1]$ with

$$\begin{aligned}\delta\det F[\varphi \equiv \det F[\overline{\varphi}] - \det F[\varphi] &= (-1)^\alpha\det F[\varphi]F^{-1\alpha}{}_\beta[\varphi]F^\beta{}_{\alpha,i}[\varphi]Q^i{}_\gamma[\varphi]\delta\xi^\gamma[\varphi]\\ &= (-1)^\alpha\det F[\varphi]F^{-1\alpha}{}_\beta[\varphi]((-1)^{\alpha i}P^\beta_{,ij}[\varphi]Q^j{}_\alpha[\varphi]Q^i{}_\gamma[\varphi]\\ &\quad + P^\beta_{,j}[\varphi]Q^j{}_{\alpha,i}[\varphi]Q^i{}_\gamma[\varphi])\delta\xi^\gamma[\varphi],\end{aligned}$$

and making use of Eq. (29), one finds for the change in $\det F[\varphi]d\varphi$ under the replacement $\varphi \to \overline{\varphi}$,

$$\begin{aligned}\delta(\det F[\varphi]d\varphi) &= \{[(-1)^{i(\alpha+1)}Q^i{}_{\alpha,i} - (-1)^{\beta(\alpha+1)}C^\beta{}_{\alpha\beta}]\delta\xi^\alpha[\varphi] + (-1)^\alpha\delta V^\alpha_{,\alpha}[P[\varphi]]\\ &\quad + (-1)^\alpha F^{-1\alpha}{}_\beta[\varphi]\delta P^\beta_{,i}[\varphi]Q^i{}_\alpha[\varphi]\}\det F[\varphi]d\varphi.\end{aligned} \tag{148}$$

If Eqs. (146) are assumed to hold one is left with precisely the first and third terms inside the curly brackets in Eq. (143).

Equations (146) were not needed in the derivation of Eq. (134). Why are they needed now? When the gauge group has no anticommuting coordinates the answer is that Eqs. (146) are forced on us by the procedure of factoring out the gauge group. Our interchanging the orders of integration in arriving at Eq. (129), and our use of Eq. (131), amount to adopting the rule that the gauge group is to be treated formally as if it were **compact**. For consistency the associated Lie algebra must likewise be treated as compact. The generators of real representations of compact Lie algebras all have vanishing trace. Hence Eqs. (146). (Remember, we are assuming that the φ's transform linearly under the gauge group.)

In Yang–Mills theories Eqs. (146) hold automatically because the generating group is always compact. In gravity theory the situation is more subtle. Both $Q^{i}{}_{\alpha,i}$ and $C^{\beta}_{\alpha\beta}$, if one tries to compute them from Eqs. (28) and (35), are meaningless expressions involving derivatives of delta functions with coincident arguments. However, both are metric-independent covariant vector densities of unit weight. Any sensible regularization scheme **must** assign them the value zero, for otherwise spacetime would be endowed with a preferred direction even before a metric is imposed on it.

If G is a **super**gauge group, with anticommuting coordinates, the formal compactness argument fails. But the notion of simplicity, or semisimplicity, survives. The invariance groups of all known supergauge theories are semisimple, and the generators of real representations of such groups satisfy the supertrace laws (146). The semisimplicity argument can also be invoked in the case of the local frame group, which enters when the gravitational field is expressed in terms of local frame components rather than directly in terms of the metric tensor (e.g., when spinor fields are present).

If we choose φ's that do not transform linearly under the group then the functional $\mu[\varphi]$ of Eq. (118) has to be reintroduced into the theory. It is easy to verify that consistency of the above formalism is maintained under these circumstances provided the first of Eqs. (146) is replaced by Eq. (50), which is just the condition that the product $\mu[\varphi]d\varphi$ be gauge invariant. Equation (50) is, of course, consistent with the first of Eqs. (146) when $\mu = 1$.

At this point the student may object that, in the case of quantum gravity at least, there appears to be an inconsistency in what we have done. Consider the sets of variables defined by Eqs. (40). **All** of these sets transform linearly under the diffeomorphism group. However, the Jacobian that arises in transforming from one set to another is not generally constant. How can one maintain $\mu =$ constant for all sets? The answer is that one **must**. If the Jacobian is replaced by the exponential of its logarithm, it contributes a formally divergent term of the form const. $\times\, \delta(0) \int \ln g d^n x$ to the exponents in the Feynman functional integrals. All terms of this kind must be suppressed by any viable regularization scheme. By this criterion the dimensional regularization method, for example, is a viable scheme.

14. Ghosts. The BRS Transformation. The Generating Functional

The perturbation rules to which Eqs. (134) and (137) lead may be summarized as follows. The exponent in the integrands is, as usual, expanded about a stationary background φ_B, and the integrals are evaluated as series of Gaussian integrals. If $\exp(iU)$ is a Gaussian weight factor and the P^α are chosen (unwisely) to have the linear form (98), then the vertex functions are just the functional derivatives $S_{,i_1\ldots i_N}$, $N \geq 3$. Otherwise the vertex functions include contributions from $U[P[\varphi]]$. In addition to the usual graphs that one can draw there is an infinite set of new graphs arising from the factor $\det F[\varphi]$, involving a new set of "formal particles" called **ghosts**. The inverse matrix $F^{-1\alpha}{}_\beta[\varphi]$ is the bare ghost propagator in an arbitrary field φ, i.e. with an arbitrary number of φ-lines attached. The ghost propagators always enter in closed loops, never as external lines.

The conditions previously imposed on the functional U ensure that the φ-propagator exists. The presence of $U[P[\varphi]]$ in the exponents of expressions (134) and (137) breaks the gauge symmetry and eliminates the redundancy that exists in the integration (118). An important symmetry nevertheless survives. It is most easily revealed by introducing two new fields, χ_α and ψ^α, that have the unusual property of being fermionic when the index α is bosonic and vice versa. Use of these fields together with the Berezin integration rules allows one to express $\det F[\varphi]$ in the form

$$\int e^{i\chi_\alpha F^\alpha_\beta[\varphi]\psi^\beta} d\chi\, d\psi = C \det F[\varphi], \tag{149}$$

where C is a (divergent) constant, and hence

$$\langle \text{out}|\text{in}\rangle = \overline{N}[U] \int e^{i(S[\varphi]+U[P[\varphi]]+\chi F[\varphi]\psi)} d\varphi\, d\chi\, d\psi \tag{150}$$

$$\overline{N}[U] \equiv N''[U]/C. \tag{151}$$

Equation (150) shows that the fields χ_α, ψ^α are associated with the ghost particles, which are now placed on a common footing with the φ-particles.

It was discovered by Becchi, Rouet and Stora (BRS) (1975) that both the exponent and the volume element $d\varphi\, d\chi\, d\psi$ in Eq. (150) are invariant under a set of transformations whose infinitesimal forms are given by

$$d\varphi^i = Q^i{}_\alpha[\varphi]\psi^\alpha\, \delta\lambda, \quad \delta\chi_\alpha = \delta\lambda\, U_{,\alpha}[P[\varphi]], \quad \delta\psi^\alpha = -\frac{1}{2}C^\alpha_{\beta\gamma}\psi^\gamma\, \delta\lambda\, \psi^\beta, \tag{152}$$

where $\delta\lambda$ is an arbitrary infinitesimal anticommuting constant. Using the special (anti)commutativity properties of the χ's and ψ's, together with the identity (29) and the definition (31), one readily verifies the invariance of the exponent. By computing the super-Jacobian of the BRS transformation one finds that the volume element $d\varphi\, d\chi\, d\psi$ is likewise invariant, provided Eqs. (146) are assumed to hold. It is also straightforward to verify that, if confined to the φ's and ψ's, the BRS transformations constitute an Abelian group. Inclusion of the χ's destroys the group property unless $F^\alpha_\beta[\varphi]\psi^\beta = 0$. Note that the BRS transformations do **not** constitute a local gauge group. The $\delta\lambda$'s are constants; they are not functions over spacetime. Thus the integral (150) contains no redundancy.

The BRS transformations play an important role in simplifying the derivation of the "Ward–Takahashi identity" satisfied by the so-called generating functional. What follows is a partial account, adapted to the case in which the P's are nonlinear and $\exp(iU)$ is non-Gaussian, of the theory of the generating functional given by B. W. Lee (1976) and originally due to Zinn–Justin.

One begins by replacing the exponent in Eq. (150) by

$$\widetilde{S}[\varphi, \chi, \psi, K, L, M] + J_i\varphi^i + \overline{J}^\alpha\chi_\alpha + \widehat{J}_\alpha\psi^\alpha,$$

where

$$\begin{aligned}\widetilde{S}[\varphi, \chi, \psi, K, L, M] \equiv\ & S[\varphi] + U[P[\varphi]] + \chi_\alpha F^\alpha_\beta[\varphi]\psi^\beta \\ & + \{K_i + M(U[P[\varphi]])_{,i}\}Q^i{}_\alpha[\varphi]\psi^\alpha \\ & - \frac{1}{2}(-1)^\beta L_\alpha C^\alpha{}_{\beta\gamma}\psi^\gamma\psi^\beta\end{aligned} \tag{153}$$

and by generalizing Eq. (137) to

$$\langle\text{out}|T(A[\varphi, \chi, \psi])|\text{in}\rangle \equiv \overline{N}[U]\int A[\varphi, \chi, \psi]e^{i(\widetilde{S}+J\varphi+\overline{J}\chi+\widehat{J}\psi)}d\varphi\, d\chi\, d\psi. \tag{154}$$

$J_i, \overline{J}^\alpha, \widehat{J}_\alpha, K_i, L_\alpha$ and M are **external sources**,[g] and the "matrix element" (154) is a functional of them. J_i and L_α are bosonic when their indices are bosonic and fermionic when their indices are fermionic. With $\overline{J}^\alpha, \widehat{J}_\alpha$ and K_i the association is just the opposite. M is fermionic.

If the functional A in (154) is replaced by unity one gets a generalization of the "in-out" amplitude:

$$\begin{aligned}e^{iW[J,\overline{J},\widehat{J},K,L,M]} &\equiv \langle\text{out}|\text{in}\rangle \\ &\equiv \overline{N}[U]\int e^{i(\widetilde{S}+J\varphi+\overline{J}\chi+\widehat{J}\psi)}d\varphi\, d\chi\, d\psi.\end{aligned} \tag{155}$$

This generalized amplitude is called the **generating functional**, because if it is expanded in a power series in the sources $J_i, \overline{J}^\alpha$ and $\widehat{J}_\alpha$ the coefficients are the matrix elements of chronological products of field operators. The coefficient of zero order reduces to the original amplitude (150) when K_i, L_α and M vanish.

The functional $\widetilde{S}$ may be viewed as a generalized action functional. With the aid of Eqs. (29) and (36) one may readily show that it is BRS invariant. Suppose the variables φ, χ, ψ in the integrand of Eq. (155), as well as in the volume element, are subjected to a BRS transformation. Since these variables are dummies the integral remains unaffected. Explicitly, however, the terms in $J, \overline{J}, \widehat{J}$ change. Therefore

$$\begin{aligned}0 = i\overline{N}[U]\int &\left\{J_iQ^i{}_\alpha[\varphi]\psi^\alpha + (-1)^\alpha\overline{J}^\alpha U_{,\alpha}[P[\varphi]] + \frac{1}{2}(-1)^\beta\widehat{J}_\alpha C^\alpha_{\beta\gamma}\psi^\gamma\psi^\beta\right\} \\ &\times e^{i(\widetilde{S}+J\varphi+\overline{J}\chi+\widehat{J}\psi)}d\varphi\, d\chi\, d\psi.\end{aligned} \tag{156}$$

[g] M is a constant. The others depend, through their indices, on position in spacetime.

This result can be expressed in an alternative form through use of

$$
\begin{aligned}
0 &= \int \frac{\delta}{\delta \chi_\alpha} \left\{ f[\varphi] e^{i(\widetilde{S}+J\varphi+\overline{J}\chi+\widehat{J}\psi)} \right\} d\varphi\, d\chi\, d\psi \\
&= i \int \left\{ F^\alpha_\beta[\varphi]\psi^\beta - (-1)^\alpha \overline{J}^\alpha \right\} f[\varphi] e^{i(\widetilde{S}+J\varphi+\overline{J}\chi+\widehat{J}\psi)} d\varphi\, d\chi\, d\psi, \qquad (157)
\end{aligned}
$$

where f is any functional of the φ's. One obtains

$$
\begin{aligned}
0 &= i\overline{N}[U] \int \left\{ J_i Q^i{}_\alpha[\varphi]\psi^\alpha + U_{,\alpha}[P[\varphi]] F^\alpha_\beta[\varphi]\psi^\beta + \frac{1}{2}(-1)^\beta \widehat{J}_\alpha C^\alpha_{\beta\gamma}\psi^\gamma\psi^\beta \right\} \\
&\quad \times e^{i(\widetilde{S}+J\varphi+\overline{J}\chi+\widehat{J}\psi)} d\varphi\, d\chi\, d\psi \\
&= \overline{N}[U] \int \left(J_i \frac{\delta}{\delta K_i} + \frac{\partial}{\partial M} - \widehat{J}_\alpha \frac{\delta}{\delta L_\alpha} \right) e^{i(\widetilde{S}+J\varphi+\overline{J}\chi+\widehat{J}\psi)} d\varphi\, d\chi\, d\psi, \qquad (158)
\end{aligned}
$$

in which use is made of the fact that the term containing M in $\widetilde{S}$ may be written in the form

$$
M U_{,\alpha}[P[\varphi]] F^\alpha_\beta[\varphi]\psi^\beta.
$$

Multiplying Eq. (158) by $-ie^{-iW}$, we finally get

$$
\begin{aligned}
0 &= -ie^{-iW} \left(J_i \frac{\delta}{\delta K_i} + \frac{\partial}{\partial M} - \widehat{J}_\alpha \frac{\delta}{\delta L_\alpha} \right) e^{iW} \\
&= J_i \frac{\delta W}{\delta K_i} + \frac{\partial W}{\partial M} - \widehat{J}_\alpha \frac{\delta W}{\delta L_\alpha}. \qquad (159)
\end{aligned}
$$

This relation expresses an important symmetry property of the generating functional, which leads directly to the Ward–Takahashi identities to be discussed presently. First we must review some standard material on the so-called effective action.

15. Many-Particle Green's Functions. The Effective Action

In this section important use will be made of the **Schwinger average**:

$$
\langle \underline{A} \rangle \equiv \frac{\langle \mathrm{out} | T(\underline{A}) | \mathrm{in} \rangle}{\langle \mathrm{out} | \mathrm{in} \rangle}. \qquad (160)
$$

Here $\underline{A}$ is an arbitrary functional of the operators $\underline{\varphi}^i, \underline{\chi}_\alpha, \underline{\psi}^\alpha$, and the numerator and denominator on the right are defined by Eqs. (154) and (155) respectively. It will be convenient to define

$$
\varphi^i \equiv \langle \underline{\varphi}^i \rangle, \quad \chi_\alpha \equiv \langle \underline{\chi}_\alpha \rangle, \quad \psi^\alpha \equiv \langle \underline{\psi}^\alpha \rangle. \qquad (161)
$$

When the sources $\overline{J}^\alpha, \widehat{J}_\alpha, K_i, L_\alpha, M$ vanish, the averages χ_α and ψ^α vanish. Note that although the symbols $\varphi^i, \chi_\alpha, \psi^\alpha$ have previously been used for integration variables, no confusion about their meaning will arise in practice.

It will also be convenient to denote the operators $\underline{\varphi}^i, \underline{\chi}_\alpha$, and $\underline{\psi}^\alpha$ collectively by $\underline{\varphi}^A$, their averages by φ^A, and the sources $J_i, \overline{J}^\alpha$ and $\widehat{J}_\alpha$ collectively by J_A. Let $\triangle J_A$ be arbitrary finite increments in the sources. Then we may write

$$\begin{aligned}
&\sum_{n=0}^{\infty} \frac{i^n}{n!} \triangle J_{A_n} \cdots \triangle J_{A_1} \langle \mathrm{out}|T(\underline{\phi}^{A_1} \cdots \underline{\phi}^{A_n})|\mathrm{in}\rangle \\
&\quad = \exp\left(\triangle J_A \frac{\delta}{\delta J_A}\right) \langle \mathrm{out}|\mathrm{in}\rangle = (e^{iW})_{J \to J + \triangle J} \\
&\quad = \exp\left(iW + i \triangle J_A \varphi^A + i \sum_{n=2}^{\infty} \frac{1}{n!} \triangle J_{A_n} \cdots \triangle J_{A_1} G^{A_1 \cdots A_n}\right),
\end{aligned} \tag{162}$$

where

$$\varphi^A = \langle \underline{\varphi}^A \rangle = e^{-iW} \frac{\delta}{i\delta J_A} e^{iW} = \frac{\delta W}{\delta J_A}, \tag{163}$$

$$G^{A_1 \cdots A_n} \equiv \frac{\delta}{\delta J_{A_1}} \cdots \frac{\delta}{\delta J_{A_n}} W. \tag{164}$$

Dividing both sides of Eq. (162) by e^{iW} and comparing like powers of $\triangle J_A$, one obtains an infinite sequence of relations:

$$\begin{aligned}
\langle \underline{\varphi}^A \underline{\varphi}^B \rangle &= \varphi^A \varphi^B - iG^{AB}, \\
\langle \underline{\varphi}^A \underline{\varphi}^B \underline{\varphi}^C \rangle &= \varphi^A \varphi^B \varphi^C - iP_3 \varphi^A G^{BC} + (-i)^2 G^{ABC}, \text{ etc.}
\end{aligned} \tag{165}$$

where P means "sum over the N distinct permutations of indices, with a plus sign or a minus sign according to whether the permutation of the indices associated with fermionic fields is even or odd". G^{AB} is known as the **one-particle** propagator, and the $G^{A_1 \cdots A_n}$, $n \geq 3$ are known as **many-particle Green's functions**. They satisfy the boundary conditions specified by the "in" and "out" states.

Any functional of the sources J_A may be alternatively regarded as a functional of the averages φ^A. From Eqs. (163) and (164) one sees that the one-particle propagator is the transformation matrix from one set of variables to the other:

$$G^{AB} = \frac{\delta \varphi^B}{\delta J_A}. \tag{166}$$

This fact may be used to establish an important relation between the functional W and the Schwinger average of the operator field equations. The latter is obtained from the formal functional identity

$$\begin{aligned}
0 &= -ie^{-iW} \overline{N}[U] \int e^{i(\widetilde{S} + J\varphi + \overline{J}\chi + \widehat{J}\psi)} \frac{\overleftarrow{\delta}}{\delta \varphi^A} d\varphi \, d\chi \, d\psi \\
&= \langle \underline{\widetilde{S}}_{,A} \rangle + J_A,
\end{aligned} \tag{167}$$

where $\underline{\widetilde{S}}_{,A}$ is the operator corresponding to the functional $\widetilde{S}_{,A}$.

If we differentiate Eq. (167) on the left with respect to J_B and make use of Eq. (166) we obtain

$$G^{BC}{}_{C,}\langle \underline{\widetilde{S}}_{,A}\rangle = -\delta^B_A. \tag{168}$$

Here the functional derivative inside the brackets $\langle\ \rangle$ is with respect to the field operator $\underline{\varphi}^A$, and the functional derivative outside the brackets is with respect to the field average φ^C. ${}_{C,}\langle \underline{\widetilde{S}}_{,A}\rangle$ is seen to be the operator of which the one-particle propagator G^{BC} is the Green's function. Because of its boundary conditions G^{BC} may be shown to be both a left Green's function, as in Eq. (168), and a right Green's function of ${}_{A,}\langle \underline{\widetilde{S}}_{,B}\rangle$ as well. It has the symmetry

$$G^{AB} = (-1)^{AB} G^{BA} \tag{169}$$

which implies

$${}_{A,}\langle \underline{\widetilde{S}}_{,B}\rangle = (-1)^{A+B+AB}{}_{B,}\langle \underline{\widetilde{S}}_{,A}\rangle. \tag{170}$$

But this is just the condition that there exists a functional $\widetilde{\Gamma}[\varphi,\chi,\psi,K,L,M]$ such that

$$\widetilde{\Gamma}_{,A} = \langle \underline{\widetilde{S}}_{,A}\rangle. \tag{171}$$

$\widetilde{\Gamma}$ is known as the **effective action**. It satisfies the equations

$$\widetilde{\Gamma}_{,A} = -J_A, \tag{172}$$

$${}_{A,}\widetilde{\Gamma}_{,C} G^{CB} = -\delta_A{}^B, \tag{173}$$

and is related to the functional W by a Legendre transformation:

$$W = \widetilde{\Gamma} + J_A \varphi^A. \tag{174}$$

This relation may be verified through differentiation with respect to J_B and use of Eq. (172) in the form

$${}_{A,}\widetilde{\Gamma} = -(-1)^A J_A.$$

Thus

$$\frac{\delta W}{\delta J_B} = \frac{\delta \varphi^A}{\delta J_B}[{}_{A,}\widetilde{\Gamma} + (-1)^A J_A] + \varphi^B = \varphi^B,$$

which is just Eq. (163). Since $\widetilde{\Gamma}$ is determined only up to an arbitrary constant of integration, Eq. (174) may be regarded as fixing it.

$\widetilde{\Gamma}$ is also known as the **generating functional for proper vertices**. This stems from its relation to the many-particle Green's functions. By differentiating

Eq. (173) one can relate functional derivatives of the one-particle propagator to derivatives of $\widetilde{\Gamma}$. These relations yield, for example,

$$\begin{aligned} G^{ABC} &= \frac{\delta}{\delta J_A} G^{BC} = G^{AD}{}_{D,}G^{BC} \\ &= (-1)^{(B+C)D+(C+D)E+(D+E)F} G^{AD}G^{BE}G^{CF}{}_{DEF,}\widetilde{\Gamma}. \end{aligned} \tag{175}$$

If the propagators are represented by lines and the third and higher derivatives of $\widetilde{\Gamma}$ are represented by vertices, one easily sees that each new differentiation with respect to a source inserts a new line in all possible ways into the previous diagram. Each Green's function of given order is thus representable as a sum of all the possible tree diagrams of that order.

Suppose the spatial sections of spacetime are noncompact. Then an S-matrix can be introduced, connecting states defined "at infinity". The S-matrix is expressible in terms of the chronological products appearing in Eq. (162). Because these products are expressible in terms of the Green's functions (Eqs. (165)), it follows that when $\widetilde{\Gamma}$ is used, only tree diagrams are needed in the construction of the S-matrix. No closed loops appear. The vertices generated by $\widetilde{\Gamma}$ are the **proper** vertices, already containing all quantum corrections. By noting that identical tree diagrams occur in classical perturbation theory, but with $\widetilde{\Gamma}$ replaced by $\widetilde{S}$, one can show that $\widetilde{\Gamma}$ describes the quantum-corrected dynamics of coherent large-amplitude fields. One must expect the same to be true also when the spatial sections are compact and there is no S-matrix.

16. The Ward–Takahashi Identity

We now resume use of the symbols $\varphi^i, \chi_\alpha, \psi^\alpha, J_i, \overline{J}^\alpha, \widehat{J}_\alpha$ and rewrite Eq. (174) in the more explicit form

$$W[J, \overline{J}, \widehat{J}, K, L, M] = \widetilde{\Gamma}[\varphi, \chi, \psi, K, L, M] + J_i\varphi^i + \overline{J}^\alpha \chi_\alpha + \widehat{J}_\alpha \psi^\alpha. \tag{176}$$

The averages $\varphi^i, \chi_\alpha, \psi^\alpha$ depend on all six sources, but because K_i, L_α and M do not participate in the Legendre transformation one may show that

$$\frac{\delta W}{\delta K_i} = \frac{\delta \widetilde{\Gamma}}{\delta K_i}, \quad \frac{\delta W}{\delta L_\alpha} = \frac{\delta \widetilde{\Gamma}}{\delta L_\alpha}, \quad \frac{\partial W}{\partial M} = \frac{\partial \widetilde{\Gamma}}{\partial M}, \tag{177}$$

where the derivatives on the right refer only to the explicit dependence of $\widetilde{\Gamma}$ on K_i, L_α and M. This result, combined with Eq. (172) in the form

$$_{A,}\widetilde{\Gamma} = -(-1)^A J_A,$$

allows Eq. (159) to be rewritten as

$$-(-1)^i \frac{\delta \widetilde{\Gamma}}{\delta \varphi^i} \frac{\delta \widetilde{\Gamma}}{\delta K_i} + \frac{\partial \widetilde{\Gamma}}{\partial M} - (-1)^\alpha \frac{\delta \widetilde{\Gamma}}{\delta \psi^\alpha} \frac{\delta \widetilde{\Gamma}}{\delta L_\alpha} = 0, \tag{178}$$

all derivatives being left derivatives. This is the **Ward–Takahashi** identity.

The Ward–Takahashi identity has important implications for the structure of $\widetilde{\Gamma}$. That it implies the existence of some sort of symmetry possessed by $\widetilde{\Gamma}$ becomes obvious when one notes that, because of its BRS invariance, $\widetilde{S}$ too satisfies the Ward–Takahashi identity. Unfortunately, to work from Eq. (178) **to** the symmetry possessed by $\widetilde{\Gamma}$ is a much harder task. In principle one might do the following. Assume that $\widetilde{\Gamma}$ can be expanded as a power series in $\chi_\alpha, \psi^\alpha, K_i, L_\alpha$ and M. Such an assumption has nothing ***a priori*** to do with perturbation theory since the expansion is to be carried out **after** the functional integration (155) has been performed. It is based on the reasonable belief that Eq. (155) varies smoothly (at least after appropriate renormalizations) as $K_i, L_\alpha, M, \overline{J}^\alpha$ and $\widehat{J}_\alpha$ (and hence χ_α and ψ^α) go to zero.

In determining the kinds of terms that can appear in the expansion it is useful to introduce the notion of "ghost number." If one assigns the ghost numbers 1 to ψ^α and $\overline{J}^\alpha$; 0 to φ^i and J_i; -1 to $\chi_\alpha, \widehat{J}_\alpha, K_i$ and M; and -2 to L_α; one easily sees that the integrand in Eq. (155) and the integral itself have total ghost number zero. Hence W and $\widetilde{\Gamma}$ have total ghost number zero, and all the terms in the expansion of $\widetilde{\Gamma}$ must have this property as well. Also the expansion can contain no terms in M of higher order than the first since M is an anticommuting constant.

If one inserts the expansion into (178) and groups together terms of like powers, one obtains an infinite sequence of subsidiary Ward–Takahashi identities relating the φ-dependent coefficients. Unfortunately there seems to be no easy way of drawing simple inferences from these identities **en gros**. So far Eq. (178) has been applied only to renormalizable models in perturbation theory. There it has proved to be of great service in the practical details of the renormalization program, as well as in the demonstration that the theory is indeed renormalizable to all orders and that unitarity is maintained. (The student is referred to the literature for details.) What role it is destined to play in quantum gravity remains to be seen.

17. References in the DeWitt Lectures

C. Becchi, A. Rouet and R. Stora, *Comm. Math. Phys.* **42** (1975) 127.

F. A. Berezin, *The Method of Second Quantization* (Academic Press, New York, 1966).

B. S. DeWitt, *Dynamical Theory of Groups and Fields* (Gordon and Breach, New York, 1965).

B. S. DeWitt, *Phys. Rev.* **160** (1967a) 1113.

B. S. DeWitt, *Phys. Rev.* **162** (1967b) 1195.

B. S. DeWitt, *Phys. Rev.* **162** (1967c) 1239.

L. D. Fadde'ev, in *Methods in Field Theory*, eds., R. Balian and J. Zinn–Justin, 1975 Les Houches Lectures (North Holland, Amsterdam, 1976).

R. P. Feynman, *Rev. Mod. Phys.* **20** (1948) 367.

R. P. Feynman, *Phys. Rev.* **80** (1950) 440.

R. P. Feynman, *Acta Phys. Polon.* **24** (1963) 697. See also in *Proceedings of the 1962 Warsaw Conference on the Theory of Gravitation*, PWN-Editions Scientifiques de Pologne, Warszawa (1964).
A. E. Fischer, in *Relativity: Proceedings of the Relativity Conference in the Midwest*, eds. Carmeli, Fickler and Witten (Plenum, New York, 1970).
C. Freifeld, in *Battelle Rencontres, 1967 Lectures in Mathematics and Physics*, eds. C. M. DeWitt and J. A. Wheeler (W. A. Benjamin, Inc., New York, 1968).
V. N. Gribov, Lecture at the Twelfth Winter School of the Leningrad Nuclear Physics Institute (1977) (SLAC translation No. 176).
C. J. Isham, J. Strathdee and A. Salam, *Phys. Rev. D* **3** (1971) 1805.
C. J. Isham, J. Strathdee and A. Salam, *Phys. Rev. D* **5** (1972) 2584.
B. Kostant, in *Differential Geometrical Methods in Mathematical Physics* (Proceedings of the July 1–4, 1975 Symposium in Bonn), eds. Bleuler and Reetz, Lecture Notes in Mathematics No. 570 (Springer, Berlin, 1977).
B. W. Lee, in *Methods in Field Theory*, eds. R. Balian and J. Zinn-Justin, 1975 Les Houches Lectures, (North Holland, Amsterdam, 1976).
C. W. Misner, K. S. Thorne and J. A. Wheeler, *Gravitation* (Freeman, San Francisco, 1973).
P. Nath, in *Proceedings of the Conference on Gauge Theories and Modern Field Theory*, Boston, 1975 (M.I.T. Press, Cambridge, Massachusetts, 1976).
K. S. Stelle, *Phys. Rev. D* **16** (1977) 953.
R. Utiyama, *Phys. Rev.* **101** (1956) 1597.
S. Weinberg, *Phys. Rev.* **140** (1965) B516.
S. Weinberg, in *General Relativity, an Einstein Centenary Survey*, eds. S. W. Hawking and W. Israel, (Cambridge University Press, 1979).

18. Old Unification

We can now outline how the space-of-histories formulation of Secs. 2–16 provides a common ground for describing the "old" and "new" unifications of fundamental theories. Quantum field theory begins once an action functional S is given, since the first and most fundamental assumption of quantum theory is that every isolated dynamical system can be described by a characteristic action functional [41]. The Feynman approach makes it necessary to consider an infinite-dimensional manifold such as the space Φ of all field histories φ^i. On this space there exist (in the case of gauge theories) vector fields

$$Q_\alpha = Q^i{}_\alpha \frac{\delta}{\delta \varphi^i} \tag{179}$$

that leave the action invariant, i.e. (see Eq. (108)),

$$Q_\alpha S = 0. \tag{180}$$

The Lie brackets of these vector fields lead to a classification of all gauge theories known so far.

18.1. *Type-I gauge theories*

The peculiar property of type-I gauge theories is that the Lie brackets $[Q_\alpha, Q_\beta]$ are equal to linear combinations of the vector fields themselves, with structure constants, i.e.

$$[Q_\alpha, Q_\beta] = C^{\gamma}_{\ \alpha\beta}\, Q_\gamma, \tag{181}$$

where $C^{\gamma}_{\ \alpha\beta,i} = 0$. The Maxwell, Yang–Mills, Einstein theories are all examples of type-I theories (this is the "unifying feature"). All of them, being gauge theories, need supplementary conditions, since the second functional derivative of S is not an invertible operator. After imposing such conditions, the theories are ruled by a differential operator of D'Alembert type (or Laplace type, if one deals instead with Euclidean field theory), or a non-minimal operator at the very worst (for arbitrary choices of gauge parameters). For example, when Maxwell theory is quantized via functional integrals in the Lorenz [42] gauge,[h] one deals with a gauge-fixing functional

$$\Phi(A) = \nabla^b A_b, \tag{182}$$

and the second-order differential operator acting on the potential reads as

$$P_a^{\ b} = -\delta_a^{\ b}\Box + R_a^{\ b} + \left(1 - \frac{1}{\alpha}\right)\nabla_a \nabla^b, \tag{183}$$

where α is an arbitrary gauge parameter. The Feynman choice $\alpha = 1$ leads to the minimal operator

$$\widetilde{P}_a^{\ b} = -\delta_a^{\ b}\Box + R_a^{\ b},$$

which is the standard wave operator on vectors in curved spacetime. Such operators play a leading role in the one-loop expansion of the Euclidean effective action.

18.2. *Type-II gauge theories*

For type-II gauge theories, Lie brackets of vector fields Q_α are as in Eq. (181) for type-I theories, but the structure constants are promoted to structure functions. An example is given by simple supergravity (a supersymmetric gauge theory of gravity, with a symmetry relating bosonic and fermionic fields) in four spacetime dimensions, with auxiliary fields [34].

[h]In [42] the author, L. Lorenz, built a set of retarded potentials which can be shown to satisfy the Lorenz gauge, although in 1867 no-one had thought of electrodynamics as a gauge theory. This author is not H. Lorentz, whose name is incorrectly associated to such a gauge in previous literature.

18.3. *Type-III gauge theories*

In this case, the Lie bracket (181) is generalized by

$$[Q_\alpha, Q_\beta] = C^{\gamma}_{\ \alpha\beta}\, Q_\gamma + U^{i}_{\ \alpha\beta}\, S_{,i}, \tag{184}$$

and it therefore reduces to Eq. (181) only on the *mass-shell*, i.e. for those field configurations satisfying the Euler–Lagrange equations. An example is given by theories with gravitons and gravitinos such as Bose–Fermi supermultiplets of both simple and extended supergravity in any number of spacetime dimensions, without auxiliary fields [34].

18.4. *From supergravity to general relativity*

It should be stressed that general relativity is naturally related to supersymmetry, since the requirement of gauge-invariant Rarita–Schwinger equations [43] implies Ricci-flatness in four dimensions [44], which is then equivalent to vacuum Einstein equations. The Dirac operator is more fundamental in this framework, since the m-dimensional spacetime metric is entirely re-constructed from the γ-matrices, in that

$$g^{ab} = \frac{1}{2m}\mathrm{tr}(\gamma^a\gamma^b + \gamma^b\gamma^a). \tag{185}$$

19. New Unification

In modern high energy physics, the emphasis is no longer on fields (sections of vector bundles in classical field theory, operator-valued distributions in quantum field theory), but rather on extended objects such as strings. In string theory, particles are not described as points, but instead as strings, i.e. one-dimensional extended objects. While a point particle sweeps out a one-dimensional worldline, the string sweeps out a worldsheet, i.e. a two-dimensional real surface. For a free string, the topology of the worldsheet is a cylinder in the case of a closed string, or a sheet for an open string. It is assumed that different elementary particles correspond to different vibration modes of the string, in much the same way as different minimal notes correspond to different vibrational modes of musical string instruments. The five different string theories are different aspects of a more fundamental theory, called M-theory [45]. In the latest developments, one deals with “branes”, which are classical solutions of the equations of motion of the low-energy string effective action, that correspond to new non-perturbative states of string theory, break half of the supersymmetry, and are required by T-duality in theories with open strings. They have the peculiar property that open strings have their end-points attached to them. With the language of pseudo-Riemannian geometry, branes are timelike surfaces embedded into bulk spacetime [32]. According to this picture, gravity lives on the bulk, while standard-model gauge fields are confined on the brane. For branes, the normal vector N is spacelike with respect to the bulk metric G_{AB}, i.e.

$$G_{AB}N^AN^B = N_CN^C > 0. \tag{186}$$

The action functional S splits into the sum [32] ($g_{\alpha\beta}(x)$ being the brane metric)

$$S = S_4[g_{\alpha\beta}(x)] + S_5[G_{AB}(X)], \tag{187}$$

while the effective action [55] Γ is formally given by

$$e^{i\Gamma} = \int DG_{AB}(X)\, e^{iS} \times \text{g.f. term}. \tag{188}$$

In the functional integral, the gauge-fixed action reads as (here there is summation as well as integration over repeated indices)

$$S_{\text{g.f.}} = S_4 + S_5 + \frac{1}{2} F^A \Omega_{AB} F^B + \frac{1}{2} \chi^\mu \omega_{\mu\nu} \chi^\nu, \tag{189}$$

where F^A and χ^μ are bulk and brane gauge-fixing functionals, respectively, while Ω_{AB} and $\omega_{\mu\nu}$ are non-singular "matrices" of gauge parameters. The gauge-invariance properties of bulk and brane action functionals can be expressed by saying that there exist vector fields on the space of histories such that (cf. Eq. (180))

$$R_B S_5 = 0, \quad R_\nu S_4 = 0, \tag{190}$$

whose Lie brackets obey a relation formally analogous to Eq. (181) for ordinary type-I theories, i.e.,

$$[R_B, R_D] = C^A_{\ BD}\, R_A, \quad [R_\mu, R_\nu] = C^\lambda_{\ \mu\nu}\, R_\lambda. \tag{191}$$

The bulk and brane ghost operators are therefore

$$Q^A_{\ B} = R_B F^A = F^A_{\ ,a}\, R^a_{\ B}, \tag{192}$$

$$J^\mu_{\ \nu} = R_\nu \chi^\mu = \chi^\mu_{\ ,i}\, R^i_{\ \nu}, \tag{193}$$

respectively. The full bulk integration means integrating first with respect to all bulk metrics G_{AB} inducing on the boundary ∂M the given brane metric $g_{\alpha\beta}(x)$, and then integrating with respect to all brane metrics. Thus, one first evaluates the cosmological wave function of the bulk spacetime [32], i.e.

$$\psi_{\text{Bulk}} = \int_{G_{AB}[\partial M] = g_{\alpha\beta}} \mu(G_{AB}, S_C, T^D) e^{i\widetilde{S}_5}, \tag{194}$$

where μ is taken to be a suitable measure, the S_C, T^D are ghost fields, and (of course, S_A here differs from the symbol for the action in Eq. (103))

$$\widetilde{S}_5 \equiv S_5[G_{AB}] + \frac{1}{2} F^A \Omega_{AB} F^B + S_A Q^A_{\ B} T^B. \tag{195}$$

Eventually, the effective action results from

$$e^{i\Gamma} = \int \widetilde{\mu}(g_{\alpha\beta}, \rho_\gamma, \sigma^\delta) e^{i\widetilde{S}_4} \psi_{\text{Bulk}}, \tag{196}$$

where $\widetilde{\mu}$ is another putative measure, ρ_γ and σ^δ are brane ghost fields, and

$$\widetilde{S}_4 \equiv S_4 + \frac{1}{2} \chi^\mu \omega_{\mu\nu} \chi^\nu + \rho_\mu J^\mu_{\ \nu} \sigma^\nu. \tag{197}$$

20. Bulk and Brane BRST Transformations

This scheme is invariant under infinitesimal BRST transformations on the bulk, given by

$$\delta G^a = R^a_{\ A}\, T^A\, \delta\Lambda, \tag{198}$$

$$\delta S_A = \Omega_{AB} F^B\, \delta\Lambda, \tag{199}$$

$$\delta T^A = -\frac{1}{2} C^A_{\ BD} T^B T^D\, \delta\Lambda, \tag{200}$$

where $T^A\, \delta\Lambda = -\delta\Lambda\, T^A$, $T^A T^B = -T^B T^A$, as well as under formally analogous BRST transformations on the brane, i.e.

$$\delta g^i = R^i_{\ \mu}\, \sigma^\mu\, \delta\lambda, \tag{201}$$

$$\delta\rho_\mu = \omega_{\mu\nu}\chi^\nu\, \delta\lambda, \tag{202}$$

$$\delta\sigma^\mu = -\frac{1}{2} C^\mu_{\ \nu\zeta}\sigma^\nu\sigma^\zeta\, \delta\lambda, \tag{203}$$

where $\sigma^\mu \delta\lambda = -\delta\lambda\, \sigma^\mu$, $\sigma^\nu\sigma^\zeta = -\sigma^\zeta\sigma^\nu$.

21. New Perspectives in the Spectral Asymptotics of Euclidean Quantum Gravity

Since the early eighties there has been a substantial revival of interest in quantum cosmology, motivated by the hope of obtaining a complete picture of how the universe could arise and evolve [46–49]. By complete we here mean a theoretical description where, by virtue of the guiding principles of physics and mathematics, both the differential equations of the theory and the associated boundary (and initial) conditions are fully specified. Even though modern theoretical cosmology deals with yet other deep issues such as dark matter, dark energy [50, 51] and cosmic strings [52], the effort of formulating the appropriate boundary conditions for the quantum state of the universe [53], or at least for its (one-loop) semiclassical approximation, plays again a key role, since the universe might have had a semiclassical origin [54], and the various orders in $\hbar$ in the loop expansion describe the departure from the underlying classical dynamics.

The physical motivations of our work result therefore from the following active areas of research:

(i) Functional integrals and space-time approach to quantum field theory [55].
(ii) Attempt to derive the whole set of physical laws from invariance principles [56].
(iii) How to derive the early universe evolution from quantum physics; how to make sense of a wave function of the universe and of Hartle–Hawking quantum cosmology [53, 57].
(iv) Spectral theory and its physical applications, including functional determinants in one-loop quantum theory and hence the first corrections to classical dynamics [58].

The boundary conditions that we study are part of a unified scheme for Maxwell, Yang–Mills and Quantized General Relativity at one loop, i.e. [59]

$$[\pi \mathcal{A}]_{\mathcal{B}} = 0, \tag{204}$$

$$[\Phi(A)]_{\mathcal{B}} = 0, \tag{205}$$

$$[\varphi]_{\mathcal{B}} = 0. \tag{206}$$

With our notation, π is a projector acting on the gauge field $\mathcal{A}$, Φ is the gauge-fixing functional, φ is the full set of ghost fields [60]. Both Eqs. (204) and (205) are preserved under infinitesimal gauge transformations provided that the ghost obeys homogeneous Dirichlet conditions as in Eq. (206). For gravity, we choose Φ so as to have an operator P of Laplace type on metric perturbations in the one-loop Euclidean theory.

22. Eigenvalue Conditions for Scalar Modes

On the Euclidean 4-ball, we expand metric perturbations $h_{\mu\nu}$ in terms of scalar, transverse vector, transverse-traceless tensor harmonics on S^3. For vector, tensor and ghost modes, boundary conditions reduce to Dirichlet or Robin [61]. For scalar modes, one finds eventually the eigenvalues $E = x^2$ from the roots x of [61]

$$J_n'(x) \pm \frac{n}{x} J_n(x) = 0, \tag{207}$$

$$J_n'(x) + \left(-\frac{x}{2} \pm \frac{n}{x} \right) J_n(x) = 0. \tag{208}$$

Note that both x and $-x$ solve the same equation. For example, at small n and large x, the roots of Eq. (208) with $+$ sign in front of $\frac{n}{x}$ read as (here $s = 0, 1, \ldots, \infty$)

$$x(s,n) \sim \beta(s,n) \left[1 + \frac{\gamma_1}{\beta^2(s,n)} + \frac{\gamma_2}{\beta^4(s,n)} + \frac{\gamma_3}{\beta^6(s,n)} + \mathrm{O}(\beta^{-8}) \right], \tag{209}$$

where

$$\beta(s,n) \equiv \pi \left(s + \frac{n}{2} + \frac{3}{4} \right), \tag{210}$$

and (having defined $m \equiv 4n^2$)

$$\gamma_1(m) \equiv -\frac{(m-17)}{8}, \tag{211}$$

$$\gamma_2(m) \equiv -\frac{3455}{384} + 2m^{1/2} + \frac{67}{192} m - \frac{7}{384} m^2, \tag{212}$$

$$\gamma_3(m) = \frac{1117523}{15360} - \frac{115}{4} m^{1/2} - \frac{5907}{5120} m + \frac{3}{4} m^{3/2} + \frac{421}{3072} m^2 - \frac{83}{15360} m^3, \tag{213}$$

as has been found in [62].

23. Four Generalized ζ-Functions for Scalar Modes

From Eqs. (207) and (208) we obtain the following integral representations of the resulting ζ-functions upon exploiting the Cauchy theorem and rotation of contour [61, 62]:

$$\zeta_{A,B}^{\pm}(s) \equiv \frac{(\sin \pi s)}{\pi} \sum_{n=3}^{\infty} n^{-(2s-2)} \int_0^{\infty} dz\, z^{-2s} \frac{\partial}{\partial z} \log F_{A,B}^{\pm}(zn), \tag{214}$$

where (here $\beta_+ \equiv n, \beta_- \equiv n+2$)

$$F_A^{\pm}(zn) \equiv z^{-\beta_\pm}(znI_n'(zn) \pm nI_n(zn)), \tag{215}$$

$$F_B^{\pm}(zn) \equiv z^{-\beta_\pm}\left(znI_n'(zn) + \left(\frac{(zn)^2}{2} \pm n\right) I_n(zn)\right), \tag{216}$$

I_n being the modified Bessel functions of first kind. Regularity at the origin is easily proved in the elliptic sectors, corresponding to $\zeta_A^{\pm}(s)$ and $\zeta_B^{-}(s)$.

24. Regularity of ζ_B^+ at $s = 0$

We now define $\tau \equiv (1+z^2)^{-1/2}$ and consider the uniform asymptotic expansion (away from $\tau = 1$, with notation as in [61, 62])

$$z^{\beta_+} F_B^+(zn) \sim \frac{e^{n\eta(\tau)}}{h(n)\sqrt{\tau}} \frac{(1-\tau^2)}{\tau} \left(1 + \sum_{j=1}^{\infty} \frac{r_{j,+}(\tau)}{n^j}\right), \tag{217}$$

the functions $r_{j,+}$ being obtained from the Olver polynomials for the uniform asymptotic expansion of I_n and I_n' [63]. On splitting $\int_0^1 d\tau = \int_0^\mu d\tau + \int_\mu^1 d\tau$ with μ small, we get an asymptotic expansion of the left-hand side by writing, *in the first interval* on the right-hand side,

$$\log\left(1 + \sum_{j=1}^{\infty} \frac{r_{j,+}(\tau)}{n^j}\right) \sim \sum_{j=1}^{\infty} \frac{R_{j,+}(\tau)}{n^j}, \tag{218}$$

and then computing

$$C_j(\tau) \equiv \frac{\partial R_{j,+}}{\partial \tau} = (1-\tau)^{-j-1} \sum_{a=j-1}^{4j} K_a^{(j)} \tau^a. \tag{219}$$

The integral $\int_\mu^1 d\tau$ is instead found to yield a vanishing contribution in the $\mu \to 1$ limit [62]. Remarkably, by virtue of the spectral identity

$$g(j) \equiv \sum_{a=j}^{4j} \frac{\Gamma(a+1)}{\Gamma(a-j+1)} K_a^{(j)} = 0, \tag{220}$$

which holds $\forall j = 1, \ldots, \infty$, we find

$$\lim_{s\to 0} s\zeta_B^+(s) = \frac{1}{6} \sum_{a=3}^{12} a(a-1)(a-2) K_a^{(3)} = 0, \tag{221}$$

and

$$\zeta_B^+(0) = \frac{5}{4} + \frac{1079}{240} - \frac{1}{2}\sum_{a=2}^{12}\omega(a)K_a^{(3)} + \sum_{j=1}^{\infty} f(j)g(j) = \frac{296}{45}, \tag{222}$$

where

$$\omega(a) \equiv \frac{1}{6}\frac{\Gamma(a+1)}{\Gamma(a-2)}[-\log(2) - \frac{(6a^2-9a+1)}{4}\frac{\Gamma(a-2)}{\Gamma(a+1)}$$
$$+ 2\psi(a+1) - \psi(a-2) - \psi(4)], \tag{223}$$

$$f(j) \equiv \frac{(-1)^j}{j!}[-1 - 2^{2-j} + \zeta_R(j-2)(1-\delta_{j,3}) + \gamma\delta_{j,3}]. \tag{224}$$

The spectral cancellation (220) achieves three goals: (i) Vanishing of $\log 2$ coefficient in Eq. (222); (ii) Vanishing of $\sum_{j=1}^{\infty} f(j)g(j)$ in Eq. (222); (iii) Regularity at the origin of ζ_B^+.

To cross-check our analysis, we evaluate $r_{j,+}(\tau) - r_{j,-}(\tau)$ and hence obtain $R_{j,+}(\tau) - R_{j,-}(\tau)$ for all j. Only $j = 3$ contributes to $\zeta_B^{\pm}(0)$, and we find

$$\begin{aligned}\zeta_B^+(0) &= \zeta_B^-(0) - \frac{1}{24}\sum_{l=1}^{4}\frac{\Gamma(l+1)}{\Gamma(l-2)}\left[\psi(l+2) - \frac{1}{(l+1)}\right]\kappa_{2l+1}^{(3)} \\ &= \frac{206}{45} + 2 = \frac{296}{45},\end{aligned} \tag{225}$$

in agreement with Eq. (222), where $\kappa_{2l+1}^{(3)}$ are the four coefficients on the right-hand side of

$$\frac{\partial}{\partial\tau}(R_{3,+} - R_{3,-}) = (1-\tau^2)^{-4}(80\tau^3 - 24\tau^5 + 32\tau^7 - 8\tau^9). \tag{226}$$

Within this framework, the spectral cancellation reads as

$$\sum_{l=1}^{4}\frac{\Gamma(l+1)}{\Gamma(l-2)}\kappa_{2l+1}^{(3)} = 0, \tag{227}$$

which is a particular case of

$$\sum_{a=a_{\min}(j)}^{a=a_{\max}(j)}\frac{\Gamma((a+1)/2)}{\Gamma((a+1)/2-j)}\kappa_a^{(j)} = 0. \tag{228}$$

Interestingly, the full $\zeta(0)$ value for pure gravity (i.e. including the contribution of tensor, vector, scalar and ghost modes) is then found to be positive: $\zeta(0) = \frac{142}{45}$ [62], which suggests a quantum avoidance of the cosmological singularity driven by full diffeomorphism invariance of the boundary-value problem for one-loop quantum theory [62].

25. Selected Open Problems

Several open problems should be brought to the attention of the reader, and are as follows.

(i) We have encountered in Secs. 21–24 a boundary-value problem where the generalized ζ-function remains well defined, even though the Mellin transform relating ζ-function to heat kernel does not exist (see further comments below), since strong ellipticity is violated [59] (see also [64]). Are the spectral cancellations (220) and (228) a peculiar property of the Euclidean 4-ball, or can they be extended to more general Riemannian manifolds with non-empty boundary?

(ii) What is the deeper underlying reason for finding $\zeta_B^{+}(0) - \zeta_B^{-}(0) = 2$? Is it possible to foresee a geometrical or topological or group-theoretical origin of this result?

(iii) Is it correct to say that our positive $\zeta(0)$ value for pure gravity engenders a quantum avoidance of the cosmological singularity at one-loop level? [54, 62] Does the result remain true in higher-loop calculations or on using other regularization techniques for the one-loop correction?

(iv) The whole scheme might be relevant for AdS/CFT in light of a profound link between AdS/CFT and the Hartle–Hawking wave function of the universe [65].

(v) What happens if one considers instead non-local boundary data, e.g., those giving rise to surface states for the Laplacian? [56, 66, 67]

As far as item (i) is concerned, we should add what follows. The integral representation (214) of the generalized ζ-function is legitimate because, for any fixed n, there is a countable infinity of roots x_j and $-x_j$ of Eqs. (207) and (208), and they grow approximately linearly with the integer j counting such roots. The functions $F_A^{\pm}$ and $F_B^{\pm}$ admit therefore a canonical-product representation [68] which ensures that the integral representation (214) reproduces the standard definition of generalized ζ-function [61]. Furthermore, even though the Mellin transform relating ζ-function to integrated heat kernel cannot be exploited when strong ellipticity is not fulfilled, it remains possible to define a generalized ζ-function. For this purpose, a weaker assumption provides a sufficient condition, i.e. the existence of a sector in the complex plane free of eigenvalues of the leading symbol of the differential operator under consideration [61, 62]. To make sure we have not overlooked some properties of the spectrum, we have been looking for negative eigenvalues or zero-modes, but finding none. Indeed, negative eigenvalues E would imply purely imaginary roots $x = iy$ of Eq. (208), but such roots do not exist, as one can check both numerically and analytically; zero-modes would be non-trivial eigenfunctions belonging to zero-eigenvalues, but all modes (tensor, vector, scalar and ghost modes) are combinations of regular Bessel functions [61] (since we require regularity at the origin of the left-hand side of Eqs. (204)–(206)) for which this is impossible. As far as we can see, we still find sources of singularities at the origin in the generalized ζ-function resulting from lack of strong ellipticity, but the particular symmetries of the Euclidean 4-ball background reduce them to the four terms in Eq. (227), which add up to zero despite two of them being non-vanishing.

In [62] we have proposed to interpret the result $\zeta(0) = \frac{142}{45}$ for pure gravity as an indication that **full diffeomorphism invariance of the boundary-value problem engenders a quantum avoidance of the cosmological singularity**. Indeed, on one hand, the work by Schleich [69] had found that, on restricting the functional integral to transverse-traceless perturbations, the one-loop semiclassical approximation to the wave function of the universe diverges at small volumes, at least for the boundary geometry of a three-sphere. The divergence of the wave functional does not imply, by itself, that the probability density of the wave functional diverges at small volumes, since the probability density $p[h]$ on the space of wave functionals $\psi[h]$ is given by $p[h] = m[h]|\psi|^2[h]$, where $m[h]$ is the measure on this space, the scaling of which is not known in general. On the other hand, in our manifestly covariant evaluation of the one-loop functional integral for the wave function of the universe, it seems incorrect to assume that the measure $m[h]$ scales in such a way as to cancel exactly the contribution of the squared modulus of ψ, which is proportional to the three-sphere radius raised to the power $2\zeta(0)$. Thus, we find that our one-loop wave function of the universe vanishes at small volume. The normalizability condition of the wave function in the limit of small three-geometry, which is weaker than requiring it should vanish in this limit, was instead formulated and studied in Ref. 70.

The years to come will hopefully tell us whether our calculations may be viewed as a first step towards finding under which conditions a quantum theory of gravity is singularity free in cosmology [71]. For this purpose, it might also be interesting to study diffeomorphism-invariant boundary conditions for $f(R)$ theories of gravity, recently studied at one-loop level on manifolds without boundary [72].

On the non-perturbative side, encouraging progress has been made towards finding cosmological applications of non-perturbative quantum gravity via renormalization-group methods [27], including, in particular, theoretical models that might account for the accelerated expansion of the universe [73] and for flat rotation curves of galaxies [74].

Appendix A. Lie Groups

A Lie group is a group G which is also a manifold with a C^∞ structure such that the maps

$$(x, y) \rightarrow xy$$
$$x \rightarrow x^{-1}$$

are C^∞ functions. It is indeed enough to assume that $(x, y) \rightarrow xy^{-1}$, or $(x, y) \rightarrow xy$, are C^∞. Relevant examples of Lie groups are as follows [75].

(1) The space $\mathbf{R}^n$ endowed with the addition $+$.
(2) The circle S^1 defined as the quotient space $\mathbf{R}/Z$.

(3) If G and H are Lie groups, their product $G \times H$ is also a Lie group.
(4) The torus $S^1 \times S^1$.
(5) The general linear group $GL(n, \mathbf{R})$, i.e. the group of all $n \times n$ nonsingular real matrices. This is a subset of $\mathbf{R}^{n^2}$.
(6) The orthogonal group

$$O(n) \equiv \left\{ A \in GL(n, \mathbf{R}) : \ AA^t = I \right\} . \tag{A.1}$$

Here, with respect to the usual base of $\mathbf{R}^n$, the matrix A represents a linear map which is an isometry, i.e. it preserves the norm and the inner product.
(7) If $H \subset G$ is a subgroup of G and also a submanifold of G, then H is a Lie group. Thus, for example, $S^1 \subset \mathbf{R}^2$ is a Lie group because

$$S^1 = \left\{ z \in \mathbf{C} : \ z\overline{z} = x^2 + y^2 = 1 \right\} . \tag{A.2}$$

(8) The three-sphere S^3 is the Lie group of unit norm quaternions (among all the S^n, only S^1 and S^3 admit a Lie-group structure). On introducing the three symbols i, j, k such that

$$i^2 = j^2 = k^2 = -1, \tag{A.3}$$

$$ij = -ji = k, \quad jk = -kj = i, \quad ki = -ik = j, \tag{A.4}$$

a quaternion can be expressed in the form

$$x = x_1 + x_2 i + x_3 j + x_4 k, \tag{A.5}$$

with $(x_1, x_2, x_3, x_4) \in \mathbf{R}^4$, so that the complex conjugate quaternion reads as

$$\overline{x} = x_1 - x_2 i - x_3 j - x_4 k, \tag{A.6}$$

and the equation defining S^3 can be indeed satisfied, i.e.

$$x\overline{x} = \sum_{i=1}^{4} x_i^2 = 1. \tag{A.7}$$

(9) The group $SO(n)$ of all orthogonal matrices with unit determinant, i.e.

$$SO(n) \equiv \{ A \in O(n) : \det A = 1 \} , \tag{A.8}$$

is a Lie group.
(10) The group $E(n)$ of all isometries of $\mathbf{R}^n$ is a Lie group. Every element of $E(n)$ can be written uniquely as $A \cdot \tau$ where $A \subset O(n)$, and τ is a translation

$$\tau(x) = \tau_a(x) = x + a. \tag{A.9}$$

Note that $E(n) \neq O(n) \times \mathbf{R}^n$ because translations and orthogonal transformations do not commute. It is however true that $O(n)$ is diffeomorphic to $O(n) \times \mathbf{R}^n$.

Given a Lie group G, its Lie algebra g is the tangent space to G at its identity element: $g = T_e G$. In general, a Lie algebra is a finite-dimensional vector space V, endowed with an antisymmetric bilinear map [,]

$$[X, Y] = -[Y, X] \quad \forall X, Y \in V \tag{A.10}$$

which satisfies the Jacobi identity

$$[[X, Y], Z] + [[Y, Z], X] + [[Z, X], Y] = 0 \quad \forall X, Y, Z \in V. \tag{A.11}$$

An important theorem asserts that, in the finite-dimensional case, a Lie algebra is always isomorphic to the Lie algebra g of a finite-dimensional Lie group G. With a standard notation, one writes $gl(n, \mathbf{R})$ for the Lie algebra of $GL(n, \mathbf{R})$ and $o(n)$ for the Lie algebra of $O(n)$. Every $A \in g$ is defined by its value at the unit element of G. Given a basis $\{e_i\}$ in g, one has the Lie-bracket relations (see Eq. (14))

$$[e_i, e_j] = f^k_{\ ij}\, e_k, \tag{A.12}$$

where $f^k_{\ ij}$ are the structure constants of G.

In quantum gravity and quantum Yang–Mills theories one deals however with infinite-dimensional Lie groups (also called pseudo-groups in the literature). The adjoint representation of the diffeomorphism group is provided by a contravariant vector field X^μ, as can be seen from the transformation law [76]

$$\delta X^\mu = \int d^n x' \int d^n x''\, C^\mu_{\ \nu'\sigma''} X^{\sigma''}\ \delta\xi^{\nu'} = -X^\mu_{\ ,\tau}\delta\xi^\tau + X^\tau \delta\xi^\mu_{\ ,\tau}. \tag{A.13}$$

The coadjoint representation is instead provided by a covariant vector density of unit weight according to [76]

$$\delta Y_\mu = -\int d^n x' \int d^n x''\, Y_{\sigma''} C^{\sigma''}_{\ \ \nu'\mu} \delta\xi^{\nu'} = -(Y_\mu \delta\xi^\tau)_{,\tau} - Y_\sigma \delta\xi^\sigma_{\ ,\mu}. \tag{A.14}$$

We refer the reader to the work in [77] for recent results on gravitation as gauge theory of the diffeomorphism group, while deformations of diffeomorphisms are studied in detail in Refs. 78, 79.

Acknowledgments

G. Esposito is grateful to the Dipartimento di Scienze Fisiche of Federico II University, Naples, for hospitality and support, and to Professor C. DeWitt–Morette for warm approval of the present project. His work has been partially supported by PRIN *SINTESI*. Springer Science and Business Media has kindly granted permission to reprint, from Secs. 2 to 17, the 1978 Cargèse Lectures by Professor B.S. DeWitt. Along the years, G. Esposito has learned a lot from many colleagues, e.g., I.G. Avramidi, P.D. D'Eath, A.Yu. Kamenshchik, K. Kirsten, and the interaction with many Ph.D. students has been very helpful in the course of preparing this teaching-oriented review.

References

[1] S. W. Hawking and R. Penrose, *Proc. R. Soc. Lond. A* **314** (1970) 529.
[2] J. R. Klauder, gr-qc/0612168.
[3] R. Geroch, *Ann. Phys. (N.Y.)* **62** (1971) 582.
[4] A. Ashtekar, *Phys. Rev. Lett.* **46** (1981) 573.
[5] B. S. DeWitt, *Phys. Rev.* **160** (1967) 1113.
[6] C. J. Isham and A. C. Kakas, *Class. Quantum Grav.* **1** (1984) 621.
[7] C. J. Isham and A. C. Kakas, *Class. Quantum Grav.* **1** (1984) 633.
[8] A. Ashtekar, *Phys. Rev. Lett.* **57** (1986) 2244.
[9] G. Esposito, G. Gionti and C. Stornaiolo, *Nuovo Cim. B* **110** (1995) 1137.
[10] G. E. Volovik, The universe in a helium droplet, *Int. Ser. Monogr. Phys.* **117** (2006) 1.
[11] C. W. Misner, *Rev. Mod. Phys.* **29** (1957) 497.
[12] B. S. DeWitt, *Phys. Rev.* **162** (1967) 1195.
[13] G. 't Hooft and M. Veltman, *Ann. Inst. H. Poincaré A* **20** (1974) 69.
[14] M. H. Goroff and A. Sagnotti, *Nucl. Phys. B* **266** (1986) 709.
[15] I. G. Avramidi, *Nucl. Phys. B* **355** (1991) 712.
[16] G. A. Vilkovisky, *Class. Quantum Grav.* **9** (1992) 895.
[17] F. Canfora, *Nucl. Phys. B* **731** (2005) 389.
[18] G. W. Gibbons and S. W. Hawking, *Phys. Rev. D* **15** (1977) 2752.
[19] S. W. Hawking, *Phys. Rev. D* **18** (1978) 1747.
[20] P. Menotti and A. Pelissetto, *Phys. Rev. D* **35** (1987) 1194.
[21] M. Carfora and A. Marzuoli, *Class. Quantum Grav.* **9** (1992) 595.
[22] C. Rovelli and L. Smolin, *Phys. Rev. Lett.* **61** (1988) 1155.
[23] C. Rovelli and L. Smolin, *Nucl. Phys. B* **331** (1990) 80.
[24] J. Gracia-Bondia, hep-th/0206006.
[25] C. J. Isham, *Class. Quantum Grav.* **6** (1989) 1509.
[26] J. A. Wheeler, *Ann. Phys. (N.Y.)* **2** (1957) 604.
[27] M. Reuter, *Phys. Rev. D* **57** (1998) 971.
[28] O. Lauscher and M. Reuter, *hep-th/0511260.*
[29] K. Stelle, *Phys. Rev. D* **16** (1977) 953.
[30] E. Witten, hep-th/0212247.
[31] G. T. Horowitz, *New J. Phys.* **7** (2005) 201.
[32] A. O. Barvinsky, *Phys. Rev. D* **74** (2006) 084033.
[33] D. Z. Freedman, P. van Nieuwenhuizen and S. Ferrara, *Phys. Rev. D* **13** (1976) 3214.
[34] P. van Nieuwenhuizen, *Phys. Rep.* **68** (1981) 189.
[35] G. Gionti, gr-qc/9812080.
[36] J. Ambjorn, M. Carfora, D. Gabrielli and A. Marzuoli, *Nucl. Phys. B* **542** (1999) 349.
[37] R. Loll, arXiv:0711.0273.
[38] A. Kheyfets, N. J. Lafave and W. A. Miller, *Class. Quantum Grav.* **6** (1989) 659.
[39] R. Penrose and M.A. H. MacCallum, *Phys. Rep.* **6** (1972) 241.
[40] R. Penrose, *Class. Quantum Grav.* **16** (1999) A113.
[41] B. S. DeWitt, *Dynamical Theory of Groups and Fields* (Gordon & Breach, New York, 1965).
[42] L. Lorenz, *Phil. Mag.* **34** (1867) 287.
[43] W. Rarita and J. Schwinger, *Phys. Rev.* **60** (1941) 61.
[44] S. Deser and B. Zumino, *Phys. Lett. B* **62** (1976) 335.
[45] K. Becker, M. Becker and J. H. Schwarz, *String Theory and M-Theory: A Modern Introduction* (Cambridge University Press, Cambridge, 2007).
[46] S. W. Hawking, *Pont. Acad. Sci. Scr. Varia* **48** (1982) 563.

[47] A. Vilenkin, *Phys. Rev. D* **30** (1984) 509.
[48] A. Vilenkin, *Phys. Rev. D* **33** (1986) 3560.
[49] A. Vilenkin, *Phys. Rev. D* **37** (1988) 888.
[50] A. Yu. Kamenshchik, U. Moschella and V. Pasquier, *Phys. Lett. B* **511** (2001) 265.
[51] S. Capozziello, S. Nojiri and S. D. Odintsov, *Phys. Lett. B* **632** (2006) 597.
[52] A. Vilenkin, hep-th/0508135.
[53] S. W. Hawking, *Nucl. Phys. B* **239** (1984) 257.
[54] S. W. Hawking, *Phys. Scripta T* **117** (2005) 49.
[55] B. S. DeWitt, The global approach to quantum field theory, *Int. Ser. Monogr. Phys.* **114** (2003) 1.
[56] G. Esposito, *Int. J. Mod. Phys. A* **15** (2000) 4539.
[57] J. B. Hartle and S. W. Hawking, *Phys. Rev. D* **28** (1983) 2960.
[58] G. Esposito, *Quantum Gravity, Quantum Cosmology and Lorentzian Geometries, Lecture Notes in Physics, New Series m*, Vol. **m12** (Springer, Berlin, 1994).
[59] I. G. Avramidi and G. Esposito, *Commun. Math. Phys.* **200** (1999) 495.
[60] G. Esposito, A.Yu. Kamenshchik and G. Pollifrone, *Euclidean Quantum Gravity on Manifolds with Boundary*, Vol. 85 of *Fundamental Theories of Physics* (Kluwer, Dordrecht, 1997).
[61] G. Esposito, G. Fucci, A.Yu. Kamenshchik and K. Kirsten, *Class. Quantum Grav.* **22** (2005) 957.
[62] G. Esposito, G. Fucci, A.Yu. Kamenshchik and K. Kirsten, *J. High Energy Phys.* 05**09**:063 (2005).
[63] F.W. J. Olver, *Phil. Trans. R. Soc. Lond. A* **247** (1954) 328.
[64] J. S. Dowker and K. Kirsten, *Class. Quantum Grav.* **14** (1997) L169.
[65] G. T. Horowitz and J. Maldacena, *J. High Energy Phys.* 04**02**:008 (2004).
[66] M. Schröder, *Rep. Math. Phys.* **27** (1989) 259.
[67] G. Esposito, *Class. Quantum Grav.* **16** (1999) 1113.
[68] L. V. Ahlfors, *Complex Analysis* (McGraw-Hill, New York, 1966).
[69] K. Schleich, *Phys. Rev. D* **32** (1985) 1889.
[70] A. O. Barvinsky and A. Yu. Kamenshchik, *Class. Quantum Grav.* **7** (1990) L181 .
[71] C. Kiefer, *Ann. der Phys.* **15** (2005) 129.
[72] G. Cognola, E. Elizalde, S. Nojiri, S. D. Odintsov and S. Zerbini, *J. Cosmol. Astropart. Phys.* **02**:010 (2005).
[73] A. Bonanno, G. Esposito, C. Rubano and P. Scudellaro, *Class. Quantum Grav.* **23** (2006) 3103.
[74] G. Esposito, C. Rubano and P. Scudellaro, *Class Quantum Grav.* **24** (2007) 6255.
[75] M. Spivak, *A Comprehensive Introduction to Differential Geometry*, Vol. 1 (Publish or Perish, Houston, 1970).
[76] B. S. DeWitt, in *Relativity, Groups and Topology II*, eds. B. S. DeWitt and R. Stora (North-Holland, Amsterdam, 1984).
[77] Y. M. Cho, K. S. Soh, J. H. Yoon, *Q-Han Park, Phys. Lett. B* **286** (1992) 251.
[78] P. Aschieri, C. Blohmann, M. Dimitrijevic, F. Meyer, P. Schupp and J. Wess, *Class. Quantum Grav.* **22** (2005) 3511.
[79] P. Aschieri, M. Dimitrijevic, F. Meyer and J. Wess, *Class. Quantum Grav.* **23** (2006) 1883.