International Journal of Theoretical Physics, Vol. 22, No. 7, 1983

Topological Geometrodynamics

Matti Juhani Pitkänen

Helsinki University of Technology, Laboratory of Physics, SF-02150, Otaniemi, Finland

Received July 30, 1982

An elementary particle model is proposed drawn from the string model and Yang–Mills theory. Instead of describing a particle as a mathematical point, we identify it as three-dimensional submanifold of some metric space H. This generalization leads to a topological classification of particles and their interaction vertices. A topological explanation for the generation degeneracy is proposed. The dynamics of the theory is based on the following assumptions. First, the theory should have the formal structure of the Einstein–Yang–Mills theory defined on a 4-surface describing the "orbits" of particles. Second, the boundary components of the 3-manifold should carry various elementary particle characteristics. Finally, only the natural geometric structures associated with space H should be used in the construction of the dynamics. It is found that the choice $H = V^4 \times CP_2$, where V^4 denotes either Minkowski space or its light cone (favored by cosmological considerations), produces all the basic predictions of the standard model ($\sin^2\theta_W = 9/26$). The isometry group of CP_2 is $SU(3)$ and is interpreted as a color group. Since color is related to CP_2 translational degrees of freedom and CP_2 has "radius" of order $G^{1/2}$, the uncertainty principle suggests the mass scale $G^{-1/2}$ for colored states. A unified semiclassical description of hadrons as stringlike objects is proposed and gluons are identified as topologically nontrivial excitations of these objects. Rather general arguments suggest that QCD describes the interaction between gluons and quarks, which turns out to be only one aspect of strong interactions. In fact, the so-called planar diagrams have nothing to do with QCD in the proposed scheme. Finally, semiclassical considerations suggest that the theory is also capable of describing gravitational phenomena and a topological mechanism generating the classical space-time is proposed.

1. INTRODUCTION

It is generally believed that gauge invariance plays a central role in particle physics. The successes of the standard model (Weinberg, 1967;

Salam, 1968; Abers and Lee, 1973; Bailin, 1977) in the description of the electroweak interactions have motivated the application of the gauge ideas in the problem of the strong interactions also and indeed, this QCD approach has been rather successful in describing certain aspects of strong interaction physics (Politzer, 1974; Close, 1979; Reya, 1981). The confinement problem, however, presents an outstanding difficulty for QCD approach. On the other hand, certain phenomenological models, e.g., string and bag models (Nambu, 1970; Jacob, 1974; Johnson, 1975) have provided considerable qualitative understanding about hadronic phenomena. Again it is widely believed that the QCD approach should be and actually is capable of producing these phenomenological descriptions in certain approximations. Also the attempts to unify electroweak and strong interactions via gauge group extension (Georgi and Glashow, 1974; Mahantappa and Randa, 1980) suffer from difficulties (generation puzzle, mass problem).

In this work a different approach is adopted to the description of strong interactions and, more generally, to the unification of the basic interactions. We believe that the confinement problem is not of a purely technical nature and hence the unification of the basic interactions does not reduce to the problem of finding the correct gauge group. Instead, it is the field theory approach itself which should be appropriately generalized in order to overcome the above-mentioned difficulties. We base the proposed generalization on a unification of gauge theory with topological ideas abstracted from the string model.

To make our goal concrete let us consider the basic topological aspects of the string model. First, the model affords a purely topological description of quark confinement, e.g., the string has either two ends or is closed. Second, the model provides a topological classification of the basic interaction vertices: Strings either merge together or, provided they are open, join along their ends. The idea is to generalize this approach. Instead of 2-manifolds in the metric space M^4 we consider 4-manifolds in the metric space $H = M^4 \times S$, where S is some compact space with spacelike metric. We interpret the four surfaces X^4 as "orbits" of 3-manifolds X^3 having particle interpretation. The generalization of the particle concept has immediate, highly nontrivial consequences:

(a) It becomes possible to classify particles using the topology of the representative 3-manifold. A rough classification is obtained using only the boundary topology of X^3: the number of the boundary components and the topology of the individual boundary component serve as classificational tools. The simplest working hypothesis is that different particle generations (Fritzsch and Minkowski, 1981) correspond to different orientable boundary topologies and that leptons, mesons, baryons, etc. correspond to 3-manifolds with 1, 2, 3, etc. boundary components.

(b) The topology of H can play an important role in particle classification. To see this, consider the choice $H = M^4 \times S^2$ or more generally $H = M^4 \times CP_n$. The point is that the space H has a nontrivial second homology group: $H_2(H) = Z$ (Hilton and Wylie, 1966; Eguchi et al., 1980) and therefore the boundary components of a submanifold X^3 can be classified by their homology equivalence classes in $H_2(H)$, e.g., we can associate to each boundary component an integer, which we shall call homology charge. The total homology charge however vanishes for orientable 3-manifolds by the very definition of the homology concept. This suggests an obvious explanation for the color degeneracy. One might even go further and argue that quarks carry nonvanishing homology charges so that the existence of a free quark is a topological impossibility. It turns out, however, that this identification is not correct although homology charge turns out to play an important role in hadronic dynamics.

(c) The basic interaction vertices can be classified topologically. The basic vertices changing particle number (the connected sum and boundary connected sum vertices) are obtained by a direct generalization from the corresponding vertices in string model. In addition, one new vertex analogous to a 3-particle vertex in field theories is obtained. Moreover, there are reactions changing the internal state of the particle: either the purely internal topology of the representative manifold changes or the topologies of the boundary components change (Cabibbo mixing) or even the number of boundary components changes.

The construction of the dynamics of the theory is based on the following principles. First, the boundary components of the 3-manifold should carry also dynamical elementary particle characteristics besides the topological ones (the ends of the string carry fermion number). A highly nontrivial consequence is that the spinor fields of the theory should be restricted on the boundary of X^4. Second, the dynamics should have the basic structure of the Einstein–Yang–Mills theory. Finally, only the natural geometric structures associated with the space H (such as metric, vielbein structure, and spinor structure) should be used in the construction of the dynamics. The basic mathematical device used to attain this goal is the so-called induction procedure, which makes it possible to define metric and Yang–Mills structures on X^4 using the geometric structures of H.

The choice of the space H might be considered to be a highly arbitrary element of the theory. It turns, however, that some rather general physical requirements fix the choice of H essentially uniquely. First, the compactness requirement for the gauge group suggests strongly the decomposition $H = V^4 \times S$, where V^4 denotes either Minkowski space M^4 or the light cone of M^4 (the latter alternative is favored by cosmological considerations). Second, the requirement that the gauge structure of the theory is that of the standard model (only the quantum numbers associated with the standard model and with the isometries of H and the topological characteristics are needed in the description of the particle spectrum) fixes the choice of S uniquely to $S = CP_2$. The isometries of CP_2 form the group $SU(3)$ and the identification as color group turns out to be possible. The color degeneracy results from the translational degrees of freedom of CP_2, e.g., the colored states correspond to nontrivial partial waves of CP_2 in a somewhat generalized sense.

Having described the basic ideas involved in the construction of the proposed elementary particle theory we now turn to the question: What is the relation between this theory and ordinary quantum field theories? The answer is suggested by the observation that the free field propagators $G(x, y)$ allow a purely geometric representation as a path integral over paths from x to y with a certain weight associated with an individual path (Symanzik, 1969; Schwinger, 1957). Therefore also the Feynman diagrams of the interacting theory allow a geometric representation as a sum over one-dimensional singular manifolds (presence of vertices). This observation, besides revealing particularly clearly the fact that field theory describes interacting pointlike particles, also shows that the suggested theory is obtained by thickening the one-dimensional singular manifolds in space M^4 to four-dimensional manifolds in space H or equivalently by generalizing the concept of the elementary particle (zero-dimensional submanifold to three-dimensional one). Indeed this generalization has far reaching consequences: consider only the topological classification of particles and their interaction vertices.

A second natural question is: Is it possible to obtain the classical spacetime of general relativity as some kind of idealization in the proposed theory? It is suggested that the emergence of the classical spacetime is kind of a topological many-particle phenomenon. The point is that the classical equations of motion allow vacuum solution, which are typically surfaces representable as a graph for a map from V^4 to S and can have macroscopic size. When the dimension of H is smaller than 8, the transition particle $\cup$ vacuum $\rightarrow$ particle $\sharp$ vacuum takes place with high probability in the presence of a vacuum solution. The resulting connected 4-surface or $\sharp$ condensate, as we call it, is identified as a classical space-time.

The plan of the paper is the following: In Section 2 the basic ideas are introduced. In Section 3 the topological aspects of the theory are considered: topological particle classification and generalization of the dual diagrams of string model are performed. Section 4 is devoted to the construction of the dynamics of the theory. In Section 5 the problems related to the quantization of the theory are considered and a particular emphasis is devoted to the study of the correspondence between the ordinary field theory and the proposed one. Section 6 concerns the choice of H and it is shown that the choice $H = V^4 \times CP_2$ leads essentially uniquely to the gauge structure of the standard model. Section 7 is devoted to the symmetries of the CP_2 theory. In particular, it is demonstrated how the color degeneracy is related to the isometries of CP_2. Section 8 deals with the strong interaction aspects of the theory. In Section 9 the problem of gravitation is discussed in the proposed theoretical framework, and Section 10 is devoted to cosmological considerations. In Appendix A the basic properties of the manifold CP_2 are reviewed. Appendix B concerns the realization of the isometries of the space H. In Appendixes C and D the classical aspects of the theory, in particular the solutions to the classical equations of motion, are studied.

Notation. The basic ingredient of the theory is the metric space H having the product decomposition $H = V^4 \times S$, where V^4 denotes either Minkowski space M^4 or the light cone M^4_+ of Minkowski space and S some compact space having spacelike metric. The coordinates of H, $M^4_{(+)}$ and S will be denoted by h^k, m^k, and s^k, respectively. For the components of the metric the analogous notations h_{kl}, m_{kl}, and s_{kl} will be used.

The gamma matrices Γ_k of the space H can be related to the flat space gamma matrices γ_A using the so-called vielbein coefficients e^A_k:

$$\Gamma_k = \gamma_A e^A_k \tag{1}$$

The covariant constancy requirement for the gamma matrices determines the so-called vielbein connection apart from a rotation in the tangent space of H:

$$V_k = -D_k e^A_l e^l_B \Sigma^B{}_A \tag{2}$$

where D_k denotes the usual covariant derivative and sigma matrices are defined as

$$\Sigma^A{}_B = 1/4[\gamma^A, \gamma_B] \tag{3}$$

N-dimensional submanifolds of H will be denoted by the symbol X^n: usually $n = 2, 3, 4$. The boundary of X^n will be denoted by the symbol δX^n. The following notations/definitions will be needed:

Coordinates of X^n: x^α.

Metric of X^n: $g_{\alpha\beta} = h_{kl} h^k{}_\alpha h^l{}_\beta$, where the shorthand notation $h^k{}_\alpha$ is used for the partial derivatives of the coordinate variables of H.

Gamma matrices of X^n: $\Gamma_\alpha = \Gamma_k h^k{}_\alpha$.

Yang-Mills connection of X^n: $A_\alpha = A_k h^k{}_\alpha$, where A_k denotes a connection defined on H (for instance, vielbein connection).

The projection of the Riemannian connection of H to X^n: $A^k{}_{l\alpha} = \{{k \atop lm}\} h^m{}_\alpha$ defines a covariant derivative for quantities which are tensors in H. In particular, the so-called second fundamental form $H^k_{\alpha\beta}$ is defined by the covariant derivatives of the tangent vectors $h^k{}_\alpha$:

$$H^k_{\alpha\beta} = D_\beta h^k{}_\alpha = \partial_\beta h^k{}_\alpha - \left\{ {\gamma \atop \alpha\beta} \right\} h^k{}_\gamma + \left\{ {k \atop lm} \right\} h^l{}_\alpha h^m{}_\beta \tag{4}$$

The Einstein tensor of X^4 has the following representation in terms of the coordinate variables h^k:

$$G^{\alpha\beta} = \left(g^{\alpha\gamma} g^{\beta\delta} - g^{\alpha\beta} g^{\gamma\delta}/2 \right) X_{\gamma\delta} \tag{5}$$

where the tensor $X_{\alpha\beta}$ is given by

$$\begin{aligned} X_{\alpha\beta} = {} & R_{ijkl} h^i{}_\gamma h^{l\gamma} h^j{}_\alpha h^k{}_\beta \\ & + h_{kl} \left(H^k_{\alpha\gamma} H^{l\gamma}_\beta - H^{k\gamma}_\gamma H^l_{\alpha\beta} \right) \end{aligned} \tag{6}$$

Here R_{ijkl} denotes the curvature tensor of H. The corresponding representation for the curvature scalar $R = -G^\alpha_\alpha$ is easy to obtain from this expression.

The trace, taken with respect to the metric of X^4, of the second fundamental form associated with the imbedding $\delta X^4 \subset X^4$ will be needed in sequel. The trace has expression

$$X^n = g_3^{\alpha\beta} H^\mu_{\alpha\beta} n_\mu \tag{7}$$

Here the subscript 3 refers to the metric of the boundary and n_μ denotes a unit vector orthogonal to δX^4. If the coordinates of the boundary are chosen so that the coordinate x^n is constant on the boundary, the vector n_μ is given by

$$n_\mu = \delta_{n,\mu} (g_3/g_4)^{1/2} \tag{8}$$

2. BASIC ELEMENTS OF THE THEORY

The string model (Nambu, 1970; Jacob, 1974) describes the meson as a string moving in Minkowski space M^4. The dynamics of the model is defined by the action, which is the area of the orbit of the string measured in the metric induced to the 2-surface from M^4. The ends of the string are interpreted as quarks or rather carriers of quantum numbers associated with quarks and the confinement is a topological phenomenon: the string has either two ends or is closed.

Many-particle states are represented as sets of strings (or rather 1-manifolds) in some spacelike 3-surface of M^4 and the transitions between different physical states are mediated by 2-manifolds having the corresponding 1-manifolds as their spacelike boundaries. There are two basic vertices for the transitions illustrated in Figures 1a and 1b. Strings either merge together or join along the boundaries (observe that the 2-surface mediating the topology change is completely smooth).

At quantum level the state is specified by a state functional in the set of 1-surfaces in a spacelike 3-surface and the transition amplitudes are obtained by summing over all 2-surfaces having the prescribed spacelike boundaries attaching the phase factor $\exp(iS)$, where S is the area of the surface, to an individual surface. The interpretation of these diagrams as a generalization of the ordinary Feynmann diagrams is highly suggestive. We postpone the discussion about the precise form of this correspondence in the context of the quantization of the proposed theory.

We conclude that the nicest features of the string model—the description of confinement and of interactions—are related to its nontrivial topological structure. On the other hand, also the basic difficulties of the model are met already at the topological level. Baryons cannot be described in any natural way and the different quantum numbers associated with quarks find no natural explanation in the model (boundary components are structureless points). At the dynamical level the basic difficulty is that there exists no convincing way to formulate the dynamics so that the ends of the string become charge carriers. This state of affairs suggests the generalization of the string model both at the topological and the dynamical level.

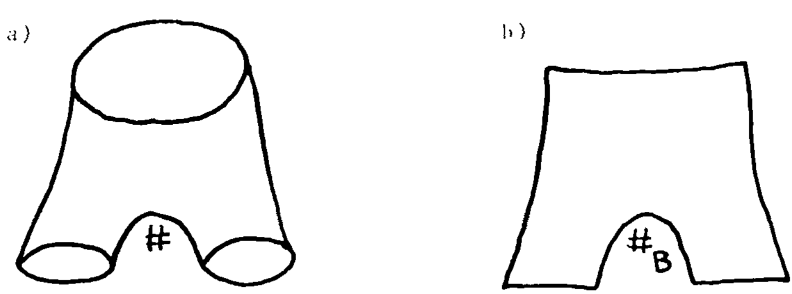

Fig. 1. The basic vertices of the string model. (a) "Trouser vertex": $\sharp$, and (b) join along boundaries: $\sharp_B$.

2.1. Generalization of the Topological Structure of String Model. The most obvious generalization of the string model is to increase the dimension of the basic dynamical entity, e.g., to make it an n-dimensional manifold of some metric space, not necessarily M^4. For $n = 2$ one obtains 2-manifolds with an arbitrary number of boundary components but the individual boundary component can have only the topology of 1-sphere S^1 and there is no hope of it describing the internal degrees of freedom of quarks topologically. The dimension $n = 3$ seems much more interesting. The topology of an orientable boundary component is characterized by its genus expressing the number of handles, which one must attach to two sphere in order to obtain a topologically equivalent 2-surface (Wallace, 1968, Chap. 7).

The tentative interpretation is that different orientable boundary topologies correspond to different particle generations (Fritzsch and Minkowski, 1981). We will speak shortly about generation-genus correspondence, e.g., (e, ν_e) and (u, d) doublets correspond to S^2 topology with $g = 0$, (μ, ν_μ) and (c, s) correspond to the torus topology with $g = 1$, etc. Observe that the dimension $n = 3$ seems to be also maximal because for $n = 4$ the boundary components are 3-manifolds and too numerous to allow any simple physical interpretation.

The obvious question is: how to choose the space H? The requirement of Poincaré invariance at laboratory scale dictates the decomposition $H = V^4 \times S$, where V^4 denotes either M^4 or M^4_+, the light cone of M^4, and S is some compact metric space with spacelike metric. The choice M^4_+ is favored by the fact that it leads naturally to the big bang cosmology provided the natural assumption that nothing enters to M^4_+ from "outside" is made. Assuming the scale of S to have the order of a typical elementary particle Compton length or even that of Planck length, we can expect that the transversal dimensions of the space H can be neglected, e.g., the space S can be effectively contracted to a point in the macroscopic limit. The dimensionality 3 of the observed world is assumed to reflect the dimension of the observers themselves, not the dimension of the space, where they "live."

How then to fix the choice of the space S? There are two alternative guidelines to follow depending on whether one tries to explain color degrees of freedom group theoretically or topologically. If one accepts that color is associated with an exact symmetry of Nature a natural identification of $SU(3)_c$ would be as the isometry group of S. Probably the simplest choice would be then $S = CP_2$. On the other hand, one might try to find a topological explanation also for color. Indeed, the choice $S = CP_n$ is characterized by the fact that the second homology group of S is nontrivial and isomorphic to that of integers (Eguchi et al., 1980, p. 240). We can thus associate with each boundary component of a given 3-manifold an integer expressing its homology equivalence class. We call this integer homology charge. By the very definition of the homology concept the sum of these integers however vanishes for any orientable 3-manifold (Hilton and Wylie, 1966; Eguchi et al., 1980)

$$\sum h_k = 0$$

Hence, postulating that quarks correspond to boundary components with homology charges h_i, $i = 1, 2, 3$, so that $h_i \neq h_j$, one obtains an explanation for the color degeneracy. A more stringent assumption $h_i \neq 0$ would explain the nonobservability of the free quark assuming that the 3-manifolds involved are orientable.

It is ironic that baryons indeed turn out to correspond to 3-manifolds having three "valence boundary components" carrying homology charges $\{1, -1, 0\}$ but that it is the group-theoretical explanation of color, which turns to be the correct one in the framework of the proposed theory.

2.2. Generalization at the Dynamical Level. Concerning the formulation of the theory one has rather obvious guidelines. First, the formulation should be based only on the use of the natural geometric structures associated with the space H, such as metric, vielbein, and spinor structure.

Second, the construction of the dynamics should be based on the generalization of the phenomenological string picture. In particular, the boundary components of the 3-manifold should be carriers of dynamical charges such as fermion number, spin, etc.

Thirdly, the successes of the standard model suggest strongly the use of gauge concepts in the formulation of the dynamics. Actually, accepting that the topological explanation for the generation degeneracy and either of the two alternative explanations for color, the only additional quantum numbers needed for the classification of the observed particles are those associated with the standard model. Therefore, an attractive working hypothesis is that the dynamics should be constructed and the space H chosen so that the gauge structure of the standard model results.

3. TOPOLOGICAL ASPECTS OF THE THEORY

This section is devoted to the purely topological aspects of the theory. Section 3.1 classifies the topological particles, and Section 3.2, the interaction vertices, so that the question, How do the dual diagrams of string model generalize? is answered.

3.1. Topological Particle Classification. The coarsest topological particle classification uses only the topology of the boundary of the 3-manifold. The boundary is specified topologically by the number of the boundary components N_c and the topology of the individual boundary components.

For orientable 3-manifolds only orientable boundary components are allowed and therefore the topology of a single boundary component is specified by its genus (Wallace, 1968, Chap. 7). In the nonorientable category also nonorientable boundary components are possible. Any nonorientable boundaryless 2-manifold is a connected sum of n projective spheres P^2 (Wallace, 1968, Chap. 7)

$$X^2 = P^2 \# P^2 \# \cdot \# P^2 = nP^2$$

The connected sum of two n manifolds is obtained by drilling holes D^n to the composites and joining the resulting boundary components S^{n-1} by a tube $D^1 \times S^{n-1}$. It is a result of the two-dimensional cobordism (Wallace, 1968) that the boundary of any 3-manifold can be expressed in the form

$$\delta X^3 = B \cup \left(\bigcup_k n_k P^2 \right)$$

where B is a disjoint union of an arbitrary number of orientable manifolds and in the latter disjoint union the constraint

$$\sum n_k = 0 \bmod 2$$

is satisfied. Loosely speaking, the boundary contains an even number of projective spheres.

Concerning the physical interpretation we adopt the hypothesis about generation genus correspondence. The physical identification of the nonorientable boundary topologies must be left open. Of course, the finer classification would take into account also the internal topology of X^3 (Wallace, 1968). We shall adopt the working hypothesis that these degrees of freedom have not yet revealed themselves experimentally.

As already found, also the topology of H can contribute to the particle classification. The boundary components of X^3 can be labeled by the elements h_k of the second homology group $H_2(H)$. By the definition of the homology concept (Hilton and Wylie, 1966) the total homology charge however vanishes for orientable 3-manifolds and is even for nonorientable manifolds (in nonorientable case one cannot associate any definite sign to the homology charge of an individual boundary component). Although the attempt to explain color topologically fails, it turns out that baryonic and mesic quarks correspond to boundary components with homology charges $\{1,-1,0\}$ and $\{1,-1\}$, respectively. Moreover, homology charge plays the role of coupling constant in planar dual diagrams.

Some closing remarks are in order. One might consider as very unsatisfying the fact that the proposed generalization of the string model suggests the existence of so many new particle states (generality of the "topological confinement condition," possibility of nonorientable manifolds, degrees of freedom associated with the internal topology of X^3, etc.). We, however, believe that the presently observed particle spectrum presents only the tip of an iceberg. Moreover, the semiclassical considerations suggest that the observed particles are exceptional in the sense that their mass scale is not given by Planck mass as that associated with a generic particle.

3.2. Topological Description of Particle Reactions. We are interested in the "basic vertices" for the particle reactions interpreted as topology changes of the representative 3-manifolds and in the possible topological selection rules. We base our approach to an intuitive idea that a many-particle state corresponds to a set of spacelike 3-manifolds imbedded in some spacelike $n-1$ submanifold of H ($\dim H = n$). In the context of quantization it will be seen that the imbeddability assumption is not necessary. The problem can be stated in more mathematical terms as follows.

Given two 3-dimensional spacelike submanifolds X_i^3 and X_f^3 in $(n-1)$-dimensional spacelike submanifolds H_i and H_f, respectively, is it possible to find a causal (having locally Minkowskian metric) submanifold X_{if}^4 having X_i^3 and X_f^3 as its spacelike boundaries so that the 4-manifold in question mediates the transition between initial and final states? Can we decompose a general transition into more elementary ones and which are the basic "vertices"? Are there any selection rules of topological origin?

This kind of problem is known as cobordism problem in topology (Wallace, 1968; Milnor, 1965; Thom, 1954).

It is useful to divide the possible particle reactions to the following basic types:

(a) The changes in the purely internal topology of the 3-manifold: the number of components of X^3 and the boundary topology remain unaltered.

(b) The reactions changing the particle number defined as the number of the boundary components of X^3.

(c) The transitions where the change in boundary topology is involved. Either the topology of an individual boundary component changes or even the number of the boundary components changes.

We shall proceed to describe shortly these different reaction types.

3.2.1. Changes in Purely Internal Topology. Since the topology of the boundary is unchanged in these reactions it is reasonable to restrict ourselves to the cobordism of the closed (boundaryless) manifolds. One obtains a rough idea about what is involved by observing that the problem reduces to a homology problem (Hilton and Wylie, 1966) if one gives up the requirement that the surfaces are manifolds, e.g., they can, for instance, intersect themselves. The selection rules for the homology problem result from the nontriviality of the third homology group $H_3(H)$, which is trivial, for example, for $H = M^4 \times CP_n$.

Therefore the possible selection rules must result from the requirement that the mediating 4-surface be manifold: both the internal topology of the 3-manifolds involved and the finite dimension of the imbedding space can lead to selection rules. It is however known that the so-called abstract cobordism (no imbedding assumed) is trivial for 3-manifolds (Wallace, 1968; Milnor, 1965; Thom, 1954). The conclusion is that the possible selection rules result from the finite dimension of H and possibly from the requirement of causality.

The problem of constructing the basic vertices for these changes is solved and we refer the reader to the literature (Wallace, 1968; Milnor, 1965). The characteristic property of these vertices is the localizability of the topology change, i.e., the change happens via a 3-manifold, which is singular in a single point. The transition from torus to sphere topology serves as a simple two-dimensional illustration of this property.

3.2.2. Reactions Changing Particle Number. As an immediate generalization of the string model vertices we obtain two kinds of vertices changing the component number of 3-manifold. We call these vertices connected sum ($\sharp$) and boundary connected sum ($\sharp_B$) vertices, respectively. There is also a third vertex not obtained in dual-model context: we will refer to this vertex as fusion ($\natural$) vertex.

The vertex is a generalization of the "trouser vertex" of the string model and is illustrated in Figure 2a for 2-manifolds. The reactants merge together in a point common to their interiors. Note that no selection rules are involved with this vertex. It can be shown that this vertex is the only vertex leading to a change in the component number of n-manifolds in the cobordism of closed n-manifolds (Wallace, 1968). This vertex makes possible the quark exchange and recombination processes used in low p_T phenomenology of hadrons and described by the so-called planar dual diagrams in dual models (Chew and Rosenzweig, 1978; Veneziano, 1974, 1976). In these kinds of processes homologically charged boundary components rearrange in combinations having vanishing total homology charges. If one accepts that gravitons correspond to closed 3-manifolds then the $\sharp$ vertex is the only possible vertex describing emission or absorption of graviton.

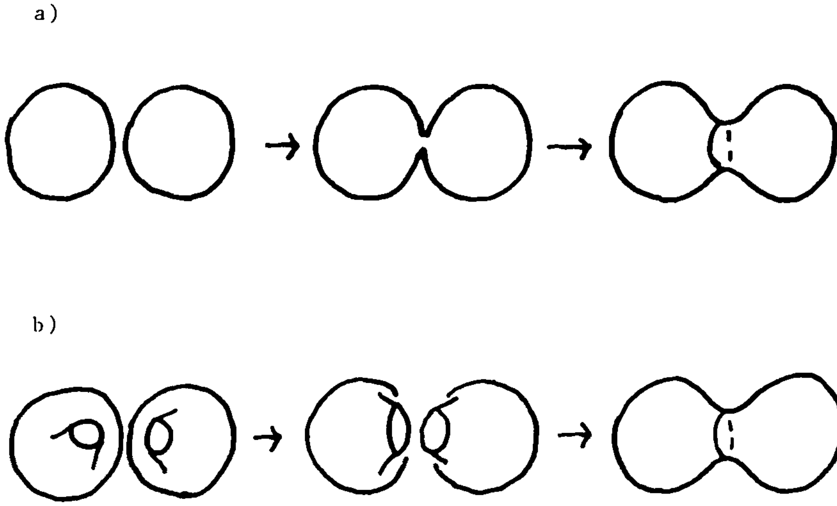

Fig. 2. (a) ♯ vertex and (b) $\sharp_B$ vertex for two-dimensional manifolds.

The $\sharp_B$ vertex presents the join of two 3-manifolds along their boundary components (Figure 2b). There are obvious selection rules associated with this vertex. The internal topologies of the boundary components must be the same and homology charges as well as various dynamical charges must have opposite values. The diagrams obtained using ♯ and $\sharp_B$ vertices correspond to the quark diagrams of the DTU approach (Veneziano, 1974, 1976). Observe that the homological triviality of the lepton states explains naturally why they do not participate in those strong interactions, which are mediated by this vertex. Figure 3a illustrates a typical quark diagram.

The fusion of two 3-manifolds at a point common to their boundaries (interiors in the ♯ vertex) is a vertex having no counterpart in dual models. We will denote this vertex by the symbol ♮. The fusion of two liquid droplets interpreted as manifolds with boundaries proceeds via this vertex. Clearly, the handle number and homology charge are conserved in this vertex. It is attractive to identify this vertex as the vertex responsible for the emission and absorption of gauge bosons and possible Higgs-type particles (observe however that we do not obtain any topological counterpart for the quartic vertices of gauge theories). Observe that the ability of the charged particle to emit photons could be simply understood as the instability of the charged boundary component against the decay resulting from the Coulombic self-repulsion of the charged boundary component.

3.2.3. Reactions Changing Boundary Topology. The reactions changing the boundary topology can be described by suitably generalizing the vertices already found. The vertex $\sharp_B$ for the boundary components belonging to the

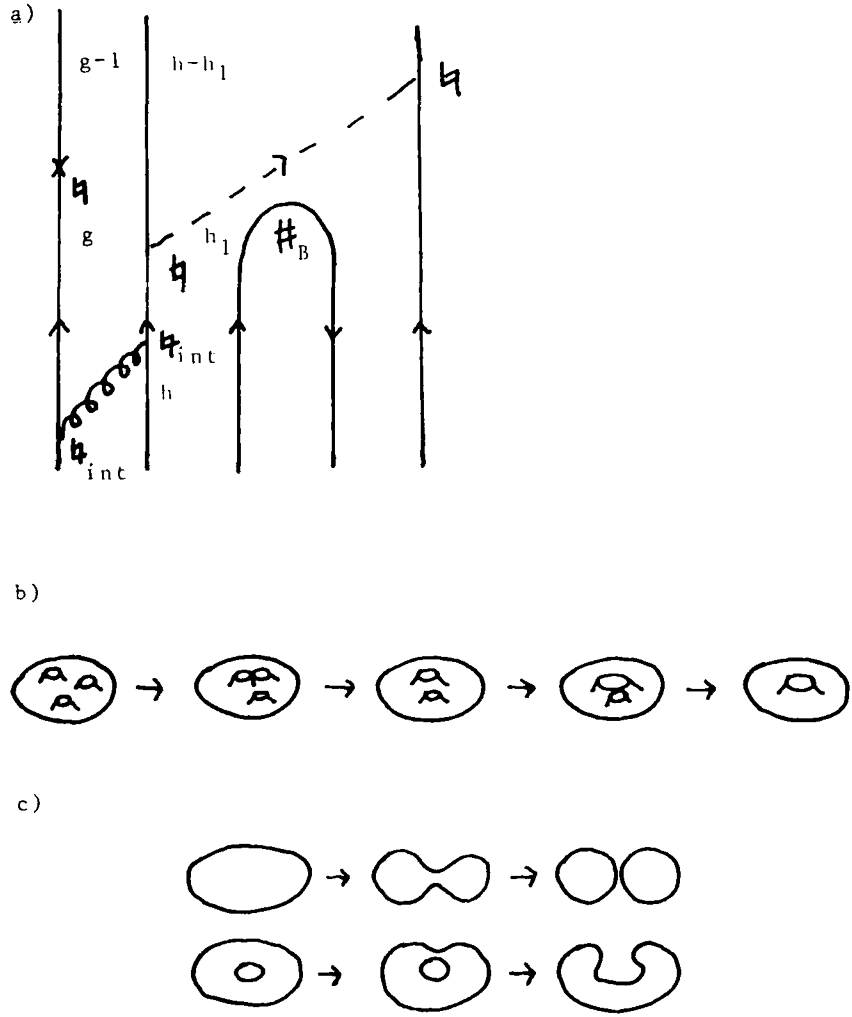

Fig. 3. (a) A "typical" quark diagram, (b) homological depolarization via ♮ vertex, (c) local equivalence of the ordinary and "internal" ♮ vertices.

same 3-manifold leads to the quark annihilation diagram in Figure 3a. We can expect this vertex to be effective for sea quarks in hadrons. The well-known OZI rule forbidding this kind of transition for mesons must be of dynamical origin (Chew and Rosenzweig, 1976, 1978; Veneziano, 1974, 1976) (also the causality requirement for the boundary of the associated 4-manifold might be involved).

The ♮ vertex generalized so that the two boundary components belong to the same 3-manifold leads to a change in the number of boundary components. This vertex is illustrated in Figure 3b). Same selection rules are associated with this vertex as with the corresponding vertex changing the component number of 3-manifold. As Figure 3c illustrates these two vertices look locally the same. Semiclassical considerations give strong support for the idea that gluons correspond to "holes in hadrons" and therefore their interactions with quarklike boundary components must take place via the ordinary ♮ vertex (quark and the absorbed gluon belong to separate hadrons) and the vertex just described ("$\natural_{int}$") (quark and the absorbed gluon belong to same hadron). Figures 4a and 4b illustrate these two, locally equivalent, gluon quark vertices.

Also the reactions changing the internal topology of a single boundary component are possible. A natural expectation is that it is this transition, which is involved in the various mixing phenomena [Cabibbo and neutrino mixing (Kobayashi and Maskawa, 1973; Bilenki and Pontecorvo, 1978)]. In fact this vertex can also be regarded as a special case of the ♮ vertex. Now the regions belonging to the same boundary component merge together at a common point.

The overall conclusion is that our approach might make it possible to imbed both the graphical rules of the dual models and those of field theories into a single mathematically well-defined scheme: the cobordism of four-dimensional manifolds in space $H = V^4 \times CP_2$. The remaining task is the construction of the dynamics so that we can associate transition amplitudes to various diagrams having a definite topological interpretation.

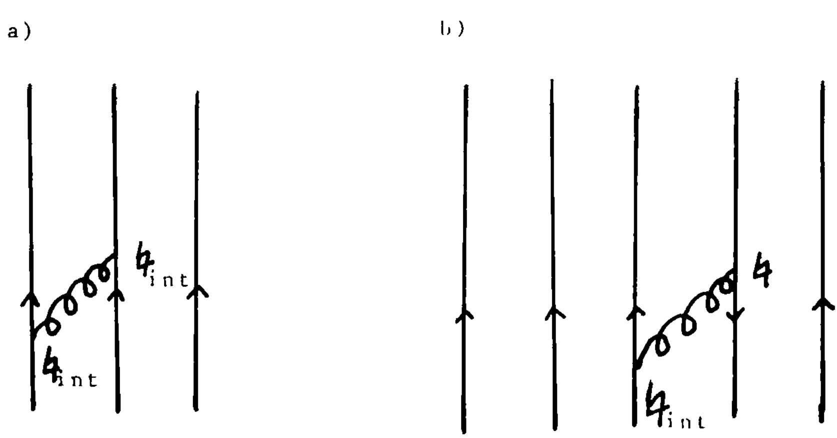

Fig. 4. Topological analog of gluon ("hole in hadron") exchange between two quarks in (a) same hadron and (b) in separate hadrons.

4. CONSTRUCTION OF THE DYNAMICS

In this section the basic principles used in the construction of the theory will be stated, the so-called induction procedure will be described, and finally, the action defining the theory will be introduced.

4.1. The Principles of the Construction. The first basic assumption used in the construction of the dynamics is that the boundary components of the 3-manifold are carriers, not only of topological but also of dynamical elementary particle characteristics. This means that the state of a single boundary component is representable symbolically by a ket $|g, h, Q, F, s, \ldots)$ (F and s denote the fermion number and spin of the boundary component, respectively). This assumption is motivated mainly by the idea that quarks indeed correspond to the ends of the string in the ordinary string model. A highly nontrivial consequence of this hypothesis is that the spinors of the theory should be restricted to the boundary of 4-manifold. This in turn gives hope that the theory is free of infinities as far as spinorial degrees of freedom are considered, because the dimensional regularization procedure needs no subtractions of infinite quantities in dimensions $N < 4$ ('t Hooft and Veltman, 1973).

The second hypothesis states that the dynamics of the theory should have the formal structure of the Einstein–Yang–Mills theory defined on the surface X^4, e.g., the bosonic part of the action is the standard Einstein–Yang–Mills action for some appropriate connection defined on X^4 and the spinorial part of the action, although restricted to the boundary δX^4, is of the standard form. The idea that spinorial charges are surface charges is consistent with the fact that one can define gauge charges for boundary components as fluxes of the appropriate gauge field components through the boundary component.

The third hypothesis states that the Einstein–Yang–Mills structure of X^4 should be obtained using only the natural geometric structures of the space H, such as metric, vielbein, and spinor structure. This hypothesis implies that the primary dynamical variables besides spinors are not the components of the Yang–Mills connection but the coordinate variables of H just as in the string model. Second, since the spinor structure in δX^4 is defined using the spinor structure of H, the so-called internal degrees of freedom, appearing in somewhat ad hoc manner in the conventional theories, find their natural place in our geometric setting. The mathematical device used to define the Yang–Mills structure on X^4 is the so-called induction procedure obtained as a direct generalization from the corresponding procedure for the metric: one obtains Yang–Mills connection of

X^4 simply by projecting the connection defined on H to X^4. The same procedure makes it possible to define also the spinor structure on δX^4.

4.2. Induction Procedure. The basic idea of the induction procedure is simply stated: one only projects the various tensor quantities defined on H to the submanifold X^n of H. In case of the metric this means the restriction of the line element of H to X^n so that lengths are measured using the units of H. The component representation for the induced metric is

$$g_{\alpha\beta} = h_{kl} h^k{}_\alpha h^k{}_\beta \tag{9}$$

In order to induce the spinor structure assume that H itself allows spinor structure so that there exists globally defined gamma matrices satisfying the anticommutation relations

$$\{\Gamma_k, \Gamma_l\} = 2h_{kl} \tag{10}$$

The gamma structure on δX^4 (or on any submanifold of H) is defined by projecting the gamma matrices of H to the surface

$$\Gamma_\alpha = \Gamma_k h^k{}_\alpha \tag{11}$$

The obvious requirement

$$\{\Gamma_\alpha, \Gamma_\beta\} = 2g_{\alpha\beta} \tag{12}$$

stating that the induced gamma matrices form Clifford algebras is satisfied and therefore the metric structure is obtained as a by-product because (12) can be regarded as a definition of the metric of δX^4.

For the spinors the induction procedure means simply the restriction to the submanifold in question. The conjugation operation: $\Psi \to \overline{\Psi}$ is defined as the corresponding operation in H. The handedness concept is generalized: the spinors are either left or right handed in the space H, not in X^4. Note, however, that this concept of handedness is defined only when H is even dimensional (Shanahan, 1978). For the spaces containing M^4 as a factor also the definition of the ordinary M^4 handedness is possible. In fact, it will turn out that the constructed theory is chirally invariant in the sense that there are two separately conserved fermion numbers corresponding to the two possible H chiralities. Physically these correspond to the conserved lepton and baryon numbers. An important feature of the induced spinor structure is that it is defined for all topologies of the submanifold unlike the ordinary spinor structure, which fails to be defined, for instance, for nonorientable manifolds (Lichnerowicz, 1968). The spinors have $2^{\dim H/2}$ components when H has even dimension.

Assuming that the quantities A_k are components of a Yang–Mills connection defined on H, the induced connection on submanifold X^n is defined in the manner already obvious:

$$A_\alpha = A_k h^k{}_\alpha \tag{13}$$

The most promising candidate for a connection to be induced is the vielbein connection of H, which is determined modulo a position-dependent rotation in the tangent space of H from the requirement that the gamma matrices of H are covariantly constant matrices (Lichnerowicz, 1968). The vielbein connection has the representation

$$A_k = A_k{}^{mn}\Sigma_{mn} \tag{14a}$$

$$A_k{}^{mn} = 1/4\{\Gamma^m, D^n\Gamma_k\} \tag{14b}$$

where D^n denotes the covariant derivative with respect to the usual Riemannian connection of H and the spin matrices are defined as the commutators of the gamma matrices of H ($N3$). The curvature form of the vielbein connection is essentially the curvature tensor of H

$$F_{kl} = 1/2 R_{klmn}\Sigma^{mn} \tag{14c}$$

Therefore the induced Yang-Mills field is the projection of the curvature tensor of H to X^4.

The gauge group associated with the vielbein connection is for the generic H with one time like direction the noncompact group $SO(n-1,1)$ but reduces to the group $SO(n-4)$ for the product decomposition $H = V^4 \times S$, where one has either $V^4 = M^4$ or $V^4 = M^4_+$. Hence the pathologies associated with the noncompact gauge groups afford the 'reason why' for the otherwise rather ad hoc choice $H = V^4 \times S$.

It should be noted that the physically most interesting choice $S = CP_2$ doesn't allow spinor structure in the ordinary sense. However, a respectable spinor structure is obtained by coupling the spinor field to a $U(1)$ gauge potential naturally associated with the geometry of CP_2. Therefore an additional $U(1)$ factor is introduced to the gauge group, the presence of which turns out to be necessary for the correctness of the gauge coupling structure.

4.3. Action Principle. Using the induced structures in X^4 and in its boundary we define the dynamics of the theory by the action

$$S = S_m + S_{gr} \tag{14d}$$

where the matter part S_m and the gravitational part S_{gr} of the action are given by the expressions

$$\begin{aligned} S_m = &-(1/4g^2)\int_{X^4} \mathrm{Tr}(F^{\alpha\beta}F_{\alpha\beta})(-g_4)^{1/2}\, d^4x \\ &+ i\int_{\delta X^4} \bar{\Psi}(\Gamma^\alpha \vec{D}_\alpha - \overleftarrow{D}_\alpha \Gamma^\alpha)\Psi(g_3)^{1/2}\, d^3x + \text{g.c.} \end{aligned} \tag{15}$$

and

$$\begin{aligned} S_{gr} = &-(1/16\pi G)\int_{\delta X^4} R(-g_4)^{1/2}\, d^4x \\ &-(1/18\pi G)\int_{\delta X^4} X^n(-g_3)^{1/2}\, d^3x \end{aligned} \tag{16}$$

In equation (15) $F_{\alpha\beta}$ denotes the projection of the curvature form of the connection of H to X^4. The spinor variables are regarded as Grassmann valued so that Ψ and $\bar{\Psi}$ are independent variables. The symbol g.c. denotes the quantity obtained by performing complex and Grassmannian conjugation for the spinorial quantity appearing in (15). The chiral invariance of the action makes it possible to apply the handedness condition to the spinor variables inducing a rather obvious change in the form of the action.

In equation (16) the symbol R denotes the curvature scalar of X^4 and the quantity X^n is the trace of the second fundamental form for the imbedding of δX^4 into X^4 [see equations (4) and (7) for the definition of the quantity X^n]. The boundary part is added in order to make S_{gr} additive in boundary-connected sum operation for 4-manifolds (Eguchi et al., 1980, p. 36). It should be stressed that concerning the precise form of the boundary part of the action, the situation is not completely settled. For instance, the addition of the term proportional to the "instanton density" $\mathrm{Tr}(F^{\alpha\beta}F^*_{\alpha\beta})$ to the interior part of the action density introduces a change in the boundary part of the action.

The action contains three constants: the gauge coupling g, the "fundamental length," say, R, given by the scale associated with the compact space S and gravitational constant G. The study of the classical aspects of $S = CP_2$ theory reveals that the addition of S_{gr} to the action is necessary: otherwise one would obtain a value for the gravitational constant differing by a factor 10^{38} from Newton's constant. Furthermore, the scale of S is given by Planck's length.

In this context some remarks concerning the concept of action, as it is used in our theory on the one hand in the ordinary field theory on the other hand, are in order. First, the action defining the ordinary quantum field theory should correspond in our framework to an effective action giving a convenient shorthand description for some of the lowest-order Green's functions (defined as appropriate generalizations of the Green's functions of the ordinary field theory) such as propagators and vertices, which, as we hope, can be used to generate at least approximatively the higher n-point functions appearing in the theory using appropriate "Feynmann rules." Second, the fact that the action defining our theory has the same form as the action of Einstein–Yang–Mills theory with constraints, can be understood as a consequence of the gauge invariance, which dictates to a high degree the form of the action, whether effective or not. Finally, the correctness of the action can be tested by studying the resulting classical equations of motion, which should provide reasonable models for elementary particles (Appendixes C and D).

5. QUANTIZATION OF THE THEORY

In this section some problems related to the formulation of the quantized theory are discussed. Section 5.1 gives the correspondence principle governing the transition from quantum field theory (QFT) to a more general theoretical framework having the suggested theory (to which we shall refer as quantum geometrodynamics, or QGD, hereafter) as a particular representative. In Section 5.2 the basic ideas of the suggested quantization procedure will be spelled out. In Sections 5.3 and 5.4 we consider the algebra of the so-called one-particle-state functionals and the definition of the transition amplitudes and probabilities, respectively.

5.1. Transition from QFT to QGD. A necessary prerequisite for the formulation of the quantized theory is a correspondence principle governing the transition from QFT to QGD. The key to this correspondence, we think, is provided by the well-known geometric representation for the propagators of the free scalar field theory (Symanzik, 1969; Schwinger, 1957) defined by the action

$$S = \Phi^{+}(1-\Gamma)\Phi \tag{17}$$

where ϕ is understood as an infinite-component vector having the space-time point as index and summation over space-time points (corresponding to integration in the continuum limit) is performed. The matrix Γ has in the Euclidian lattice the representation

$$\Gamma_{x,y} = C\sum_{\mu} (\delta x, y + \mu + \delta x, y - \mu) \equiv \Box/m^2\Delta^2 + 1 \tag{18}$$

where the sum is performed over the nearest neighbors of the lattice point, $\Box$ denotes the d'Alambert operator, and Δ is the lattice spacing.

The propagator $G(x, y)$ is defined as the inverse of the matrix $1 - \Gamma$,

$$G(x, y) = (1 - \Gamma)^{-1} \equiv \sum_{N} (\Gamma)^{N}_{x,y} \tag{19}$$

This expression can be interpreted as a sum over all possible paths leading from x to y with a weight factor proportional to the exponential of the path length attached to an individual path

$$\sum \Gamma^{N}_{x,y} = \sum_{N} [\exp(\ln CN)] \equiv \int D\gamma \exp[-kL(\gamma)] \tag{20}$$

where the latter summation is performed over all paths with fixed path length of N units and in the symbolic integral notation L denotes the path length. The representation suggests the interpretation of the field theory propagator associated with the pointlike particle described by a relativistic action proportional to the invariant path length.

Since the n point functions of an interacting scalar field theory decompose into propagators and vertices, we conclude that they have a purely geometric representation as sums over one-dimensional singular (because of the presence of the vertices) manifolds having the points $x_1, \ldots, x_n$ as boundary. This result holds true also for the spinor propagator: the presence of spin introduces only some additional factors besides the exponential factor to the functional integrand.

Preceding observations suggest that QGD should be regarded as a representative for a class of theories obtained by generalizing the field theory so that the manifolds appearing in the defining functional integrals are thickened from one- to n-dimensional manifolds and the theory is defined by specifying the factor attached to an individual "path" X^n contributing to the functional integral (one might even argue that the singularities of the field theories might reflect the singular nature if the 1-manifolds contribute to the defining functional integral).

The most straightforward test for the proposed correspondence principle is to look at what kind of a theory results, when the 4-manifolds contributing to the functional integral are contracted to one-dimensional manifolds. Clearly, the bosonic part of the action vanishes for the theory obtained using this, perhaps too naive, limiting procedure. If the bosonic action is taken to be proportional to the 4-volume of X^4, the situation of course changes: the limiting theory should describe scalar particles interacting with fermion. We think however that this limiting procedure is too rough, when applied to the YM-like theory and therefore fails to reveal the bosonic spectrum of the limiting theory.

For the spinor propagators this kind of limiting procedure is well defined and obviously one obtains for the fermion propagator the following expression:

$$G(x, y) = \int \exp[iS(\gamma)]\bar{\Psi}(x)\Psi(y)\, D\gamma D\Psi D\bar{\Psi} \tag{21a}$$

where S is the one-dimensional counterpart of the spinorial action defining the theory. Since the induced gauge field on the world line is pure gauge, we can by a suitable gauge transformation eliminate it all together and thus effectively contract S to a point. Therefore the summation in (21) is understood to be done over the paths of M^4. Taking the path length as a coordinate variable for the path contributing to the functional integral one obtains for the fermionic propagator the expression

$$G^{-1}(x, y) = C\gamma_k\left(\partial_0 m^k \partial_0 + \partial_0^2 m^k\right) = 1 + \Gamma \tag{21b}$$

Going to the Euclidian lattice and taking the coordinate condition into account one obtains for the derivative part of $1 + \Gamma$

$$C\gamma_k\left(\partial_0 m^k \partial_0\right)x, y = \sum_{\mu}\left(\delta x, y + \mu - \delta x, y - \mu\right) \tag{21c}$$

e.g., the same result as obtained in the discretization of the ordinary Dirac operator. There is however an additional term proportional to the curvature of the path having no field-theoretic counterpart. The interpretation of this additional term as a Thomas precision term resulting from accelerated motion might be possible (Jackson, 1970).

It should be noticed, that the pointlike limit of QGD makes sense only when the 4-manifolds contributing significantly to the transition amplitude in question have a definite upper bound for their scale. The fact that one cannot associate a definite-length scale to massless gauge bosons suggest an explanation for the failure of the pointlike limit for them. Furthermore, theory allows classical vacuum solutions which can have an arbitrary size, which suggests that also the limit when the size of 4-manifolds approaches infinity is physically relevant. We will postpone the discussion of this limit to the section where gravitation is discussed.

5.2. Heurjstics of the Quantization. On basis of the preceding considerations we base the quantization on the idea that QQD is obtained from the ordinary field theory by thickening the lines of Feynman diagrams to 4-manifolds, and in concordance with this, particles correspond to 3-manifolds in H.

With this in mind it is natural to generalize the Green's functions $G(x_1,\ldots,x_n)$ of the ordinary field theory to Green's functionals $A[\sigma_1,\ldots,\sigma_n]$ defined by the functional integral expression

$$A[\sigma_3{}^1,\ldots,\sigma_3{}^n]A[{}_1,\ldots,{}_n]=N\int_{\delta\sigma_4=\sigma_3}\exp(iS[\sigma_4]\,D\sigma_4 \tag{22}$$

where we have used the shorthand notation σ_n to denote the manifold X_n with a prescribed spinor configuration on its boundary and the symbol $D\sigma_4$ to denote the integration measure $DX^4\,D\Psi\,D\overline{\Psi}$. The normalization factor N is defined as the inverse of the integral appearing in (22) with $\delta\sigma_4$ empty (N divides away the vacuum bubbles). The functional integral is performed over all four surfaces and spinor configuration with prescribed spacelike boundary and associated spinor configuration. Since the spinorial variables are Grassmann valued the Green's functional as such is a rather formal object.

As in the ordinary field theory, the quantities of physical interest are obtained by contracting the amplitude A with the quantities, which we call one-particle-state functionals and which correspond to the Fourier components of the "classical fields" in field theory:

$$A^{p_1,\ldots,p_n}=\int\prod_k D\sigma_3^k\,\Omega^{p}k[\sigma_k]A[\sigma_3{}^1,\ldots,\sigma_3{}^n] \tag{23}$$

(here the label p_k denotes momentum and other quantum numbers). Because the one-particle-state functionals in general depend on spinor variables they are in general Grassmann valued. The transition amplitudes are, however, complex valued because the functional integral over the Grassmann variables gives an ordinary complex number, when performed according to the usual rules of Grassmann integration (Berezin, 1966).

5.3. The Algebra of State Functionals. The central quantities in the quantized theory are one-particle-state functionals, which should have at least the following basic properties.

(a) They are functionals of spacelike submanifolds X^3 of H and spinor configurations defined on their boundaries and are nonvanishing only for connected 3-manifolds.

(b) They transform irreducibly under the symmetries of the theory, e.g., they have for instance a definite momentum, spin, and $SU(3)$ quantum numbers.

(c) It should be possible to choose the basis for the state functionals so that the propagators $A^{\alpha,\beta}$ (where the indices refer to various quantum numbers associated with particles) are diagonal.

(d) One-particle-state functionals should correspond to the Fourier components of the ordinary fields when the 3-surface appearing as the argument of the functional has a small size in the scale defined by the wavelength associated with the functional.

Some remarks concerning the above-listed properties are in order. First, the diagonalizability property (c), which provides a nice way to define one-particle states, is not self-evident. Observe, however, that in the Euclidian formulation (which should be obtainable in a way analogous to that of the ordinary field theories) the propagator matrix is formally Hermitian and therefore is expected to be diagonalizable. So, the diagonalizability of the propagator matrix is probably ensured provided the analytic continuation to the Minkowskian regime preserves the diagonal form of the Euclidian propagator. Second, no separate scalar product for the state functionals is needed since one can cancel the dependence of the transition amplitudes on the normalization of one-particle-state functionals by multiplying the "legs" in n-point functionals by the quantity $G^{-1/2}$, where G denotes the propagator matrix. Finally, the state functionals generate in a natural way Grassmann algebra structure. The product of the state functionals provides an algebraic description for forming the disjoint union $\sigma_1 \cup \sigma_2$.

Next we turn to the construction of the one-particle-state. A rather natural basis consists of the separable functionals $\Omega = \alpha \times \beta$, where the functionals α and β depend only on X^3 and on spinor configuration, respectively. In principle, also the state functionals associating a fermion number larger than 1 to an individual boundary component are possible. The following example illustrates what the state functionals should look like:

$$\Omega = \int_{\sigma_3} \exp(ip \cdot m) Y^n X(g_3)^{1/2} d^3x \tag{24}$$

Here the functions Y^n form a complete basis in the space S being most naturally the "spherical harmonics" of S transforming irreducibly under the isometries of S. Their appearance demonstrates the degrees of freedom associated with the additional dimensions of S. As it turns out they give rise to the color degrees of freedom. For $X=1$ one obtains a state functional describing a scalar particle. The choice $X = g^{\alpha\beta}m^k{}_\alpha m^l{}_\beta - m^{kl}$ together with assumption that the state functional is restricted to the surfaces with the S^3 topology leads to a candidate for a graviton state functional. Observe the possibility to interpret this state functional as a generalization of the Fourier component of a gravitational field understood as a departure of the metric from the flat one. The choice $X = A_\alpha m^k{}_\beta g^{\alpha\beta}$ leads to a candidate for the interior part of a gauge boson state functional. The boundary part of the state functional is expected to contain a term quadratic in the spinor field.

As a second example we take the description of quark states. The functional

$$\Omega = \int_{(g,h)} \exp(ip\cdot m)\bar{\Psi}\chi^{p,r}(-g_2)^{1/2}\,d^2x \tag{25}$$

where r is a label describing the spin, charge, and color quantum numbers of the quarklike spinor χ. The label (g,h) tells that the state functional is restricted to the boundary components with prescribed genus and homology charge. We expect that the boundary part of the baryonic state functional can be obtained to a good approximation from products of the state functionals of form (25) by completely symmetrizing both in spin flavor labels and in homology charge label and by performing a complete antisymmetrization in color label.

5.4. Transition Amplitudes and Probabilities. Clearly the Green's functions $A^{p_1,\ldots,p_n}$ defined via the functional integral expression (22) should contain the physical prediction of the theory. The dependence of these quantities on the normalization of the one-particle-state functionals can be canceled by multiplying with the matrix $A^{-1/2}\equiv B$, where A is the matrix formed by the one-particle propagators, which is expected to be diagonalizable and which certainly is diagonal with this respect to four-momentum indices. So, for the states $\prod_k \Omega^{p_k}$ and $\prod_k \Omega^{q_k}$ with positive and negative energies, respectively, we define the transition amplitude as

$$T(p_1,\ldots,p_n \to q_1,\ldots,q_m) = \prod_k B^{p_k,s_k}\prod_k B^{q_k,r_k}A_{r_1,\ldots,s_m} \tag{26}$$

This amplitude is proportional to a momentum-conserving delta function and can be regarded as a transition amplitude from the state $\prod \Omega^{p_k}$ to the state $\prod \Omega^{*q_k}$.

The square of the T-matrix element has interpretation as an unrenormalized transition probability. The normalization is found by dividing with the quantity $X_n = \Sigma_r |T^{n \to r}|^2$ to obtain

$$P^{m \to n} = |T^{n \to m}|^2 / X_n \tag{27}$$

It should be emphasized that the proposed quantization procedure differs from the conventional one in certain respects. First, we do not assume sharp mass spectrum for the physical states. Instead the probability $P^{a \to a}$ serves as a measure for the lifetime of a single- (and also many-)particle state and thus for observability of the state. Second, the unitarity of the T matrix or any matrix simply related to it, is by no means necessary for the physical interpretation of the theory in the proposed sense. Of course, it is an important problem to find whether there exists a counterpart for the unitary S matrix in the proposed theory.

6. CHOICE OF THE SPACE H.

In this section the choice of the space H is considered. In Section 6.1 we will consider that the general requirements H should satisfy and in Section 6.2 we will show that the choice $H = V^4 \times CP_2$ indeed satisfies them.

6.1. Why the choice $H = V^4 \times CP_2$? The arbitrariness associated with the choice of the space H seems to give a rather ad hoc nature for the proposed elementary particle model. Quite surprisingly, it appears that some rather general physical inputs make the choice of H essentially unique.

First, the decomposition into a product $H = V^4 \times CP_2$, where one has either $V^4 = M^4$ or $V^4 = M^4_+$, is necessary in order to obtain a compact gauge group. Secondly, the suggested explanations for the color degeneracy (group-theoretic and topological) make the spaces CP_n appear as promising candidates for the space S. Of course, these spaces are favored also, because they are highly symmetric and allow additional structures (Kähler structure and associated symplectic structure) (Eguchi et al., 1980).

Thirdly, accepting the suggested topological explanation for the generation degeneracy, one can conclude that the only additional quantum numbers needed for the classification of the known elementary particles are those appearing in the standard model for electroweak interactions. Also accepting the existence of the right-handed neutrinos, one is led to conclude that the spinors of the space H should have 16 components corresponding to two weak isospin degrees of freedom or integer-charged leptons and fractionally charged quarks. Therefore one could have either $\dim S = 4$ or $\dim S = 6$ provided one imposes a chirality condition on the spinors in the latter case.

Finally, when S is four dimensional the gauge group associated with the vierbein connection decomposes into a product $SO(4) = SU(2)_L \times SU(2)_r$, where the factors act nontrivially only on the spinors with definite S chirality specified by the label $1, r$. Obviously the identification of the $SU(2)_L$ as the group $SU(2)_L$ of the standard model is suggestive and therefore the space $S = CP_2$ having $SU(3)/Z_3$ as its isometry group appears to be a particularly promising candidate. There are, however, two problems. First, one can wonder where is the $U(1)$ factor of the standard gauge group $SU(2)_L \times U(1)$. Second, CP_2 does not allow spinor structure (Gibbons and Pope, 1978)!

It is remarkable that the geometric structure of CP_2 provides a nice solution to both of these problems. The point is that CP_2 is a Kähler manifold (Gibbons and Pope, 1978; Eguchi et al., 1980) and therefore allows a closed, covariantly constant 2-form J, which is half-integer valued and defines symplectic structure in CP_2. Because $2J$ is integer valued and covariantly constant, it satisfies Maxwells equations and is interpreted as a $U(1)$ gauge field associated with a magnetic monopole of unit charge. In particular, there is a $U(1)$ connection B so that locally the equation

$$2J = dB \tag{28}$$

is satisfied. By coupling an *odd* multiple of $B/2$ (which itself defines only a gauge potential but not a connection) one obtains a respectable spinor structure as shown by Hawking (Hawking and Pope, 1978). More generally, in the case $H = M^4 \times CP_2$ it is possible to couple the components of the spinor field with different H chiralities $(\pm)$ independently to $B/2$, e.g., the odd integers n_+ and n_- need not be same. The identification of Ψ_+ and Ψ_- as quark- and leptonlike spinors, respectively, gives us a hope of obtaining correct electroweak couplings for the fundamental fermion fields. Of course, the introduction of the gauge potential $B/2$ is hoped to introduce the missing $U(1)$ factor to the gauge group. So, we are in the position to state the exciting question: Do we obtain the gauge structure characteristic to the standard model with a suitable choice of the integers n_+ and n_- and does the isometry group of CP_2 have the interpretation as color group?

6.2. Gauge Structure of the CP_2 Theory. In the preceding section it was found that the dimension of CP_2 and the intricacies associated with its spinor structure single it out as a unique candidate for the space S. In this section we show that the associated gauge group indeed reduces to that of the standard model, in particular, the right-handed neutrinos decouple totally from gauge interactions.

We begin by observing that the space H allows us to define three different chiralities for the spinors. First, there is the chirality defined in H: $i\Gamma_9\Psi_\pm = \pm\Psi_\pm$. We will interpret Ψ_+ and Ψ_- as quark- and leptonlike spinor degrees of freedom. This identification is motivated by the fact that the action is chirally symmetric in the sense that one can associate conserved fermion numbers to both chiralities. Secondly, we can define CP_2 chirality: $1 \times \gamma_5 \Psi_{l(r)} =_{(-)}^{+} \Psi_{l(r)}$. Finally, we can speak of M^4 chirality or rather handedness defines as $i\gamma_5 \times l\Psi_{L(R)} =_{(-)}^{+} \Psi_{L(R)}$.

It is a rather trivial but important observation that for a fixed H chirality there is complete correlation between CP_2 and M^4 chiralities: for Ψ_+ and Ψ_- these chiralities are the same and opposite, respectively. Hence, identifying Ψ_+ as the quark field q_- and Ψ_- as the charge conjugate L^c of the lepton field, one can perform the identification $SO(4) = SU(2)_L \times SU(2)_R$ for the vierbein part of the gauge group.

We now turn to the task of showing that the couplings of the standard model result from the covariant derivative defined via the spinor connection

$$A = V + B/2(n_+ 1_+ + n_- 1_-) \tag{29a}$$

where V and B denote the vierbein and Kähler connections, respectively, and 1_+ and 1_- project to the quark- and leptonlike subspaces, respectively. The integers n_+ and n_- are odd from the requirement of a respectable spinor structure. The explicit representations for V and B are (Eguchi et al., 1980, p. 257)

$$V = V_{AB}\Sigma^{AB} \tag{29b}$$

where the quantities V_{AB} are given by

$$V_{01} = -V_{23} = -e^1/r$$

$$V_{02} = -V_{13} = -e^2/r$$

$$V_{03} = (r - 1/r)e^3$$

$$V_{12} = (2r + 1/r)e^3 \tag{29c}$$

and

$$B = 2re^3 \tag{30}$$

respectively. For the definition of the coordinate variables and the vierbein forms e^k we refer to Appendix A, where the basic facts about the geometry of CP_2 are reviewed.

Identifying Σ_{03} and Σ_{12} as the diagonal Lie algebra generators for $SO(4)$ one finds that the charged part of the connection is given by

$$A_{\text{ch}} = V_{23} I_L^1 + V_{13} I_L^2 \tag{31}$$

where we have defined

$$I_L^1 = (\Sigma_{01} - \Sigma_{23})/2$$

$$I_L^2 = (\Sigma_{02} - \Sigma_{13})/2 \tag{32}$$

and is indeed purely left handed so that we can perform the identifications

$$W^{+(-)} = 2^{1/2}/r(e^1 \pm e^2) \tag{33}$$

where $W^{+(-)}$ denotes intermediate charged vector boson (of course it should be noticed that it is the projection of $W^{+(-)}$ to the 4-surface which represents the physical gauge boson field).

Next we turn to the identification of the gauge bosons γ and Z^0 as appropriate linear combinations of the functionally independent quantities $X = re^3$ and $Y = e^3/r$ appearing in the neutral part of the connection. The identification is found by imposing two rather obvious requirements. First, photons couple purely vectorially, and second, the "free" part of the YM action, when expressed in terms of γ and Z^0, should not contain nondiagonal terms of form γZ^0.

Having these objectives in mind we define

$$\begin{pmatrix} \bar{\gamma} \\ \bar{Z}^0 \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} X \\ Y \end{pmatrix} \tag{34}$$

imposing the normalization condition

$$ad - bc = 1 \tag{35}$$

The physical fields γ and Z^0 will be related to $\bar{\gamma}$ and $\bar{Z}^0$ via simple normalizations. For the neutral part of the connection one obtains

$$\begin{aligned} A_{\text{nc}} = {} & [(c+d)2\Sigma_{03} + (2d-c)2\Sigma_{12} + d(n_+ 1_+ + n_- 1_-)]\bar{\gamma} \\ & + [-(a+b)2\Sigma_{03} + (a-2b)2\Sigma_{12} - b(n_+ 1_+ + n_- 1_-)]\bar{Z}^0 \end{aligned} \tag{36}$$

Identifying Σ_{12} and $\Sigma_{03} = \gamma_5 \Sigma_{12}$ as vectorially and axially coupled Lie algebra generators, respectively, the requirement that photons couple vectorially leads to the condition

$$c + d = 0 \tag{37}$$

Combining this with (36) one obtains for the photonic part of A_{nc}

$$A^{\gamma}_{\mathrm{nc}} = \gamma(6\Sigma_{12} + n_{+}1_{+} + n_{-}1_{-})/6 \tag{38}$$

where we have defined

$$\gamma = 6d\bar{\gamma} \tag{39}$$

Already at this stage the electromagnetic couplings can be read from (38) and indeed the choice $n_{+} = 1$ and $n_{-} = 3$ leads to the fractional and integer charge spectrum for the quarks and leptons, respectively.

To fix the remaining parameters a and b we apply the diagonality condition to the free (or Maxwellian) part of the YM action, e.g., we pose the requirement that no terms of form $\overline{\gamma Z^0}$ appear. The free part of the action density in terms of the original fields reads

$$L^F = -(1/4g^2)16\left[(dV_{03})^2 + (dV_{12})^2 + (dB/2)^2 + 2(dV_{01})^2 + 2(dV_{02})^2\right] \tag{40}$$

where the shorthand notation $X \cdot Y$ for the tensor contraction of type $X^{\alpha\beta}Y_{\alpha\beta}$ is used. Numerical factors result from various traces, such as $\mathrm{Tr}(\Sigma_{ij}^{\ 2}) = 4$ and $\mathrm{Tr}(1_{+(-)}) = 8$. Expressing L^F in terms of the physical fields one obtains

$$L^F = -4/g^2\left[13d^2(d\bar{\gamma})^2 + (9b^2 + 2a^2 - 2ab)(d\bar{Z}^0)^2 \right.$$
$$\left. + (6a - 20b)\, d\bar{\gamma} \cdot d\bar{Z}^0)\right] + \text{charged part} \tag{41}$$

The coefficient of the nondiagonal term vanishes provided one has

$$a = 10b/3 \tag{42}$$

This equation, when combined with the normalization condition (24) leads to the result

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} 10/13d & 3/13d \\ -d & d \end{pmatrix} \tag{43}$$

For the neutral part of the connection one obtains

$$A_{\mathrm{nc}} = \gamma Q_{em} + Z^0\left(I_L^3 - \sin^2\theta_W Q_{em}\right) \tag{44}$$

where we have defined

$$I_L^3 = (\Sigma_{12} - \Sigma_{03})/2$$

$$Z^0 = 4\bar{Z}^0/d$$

$$\sin^2\theta_W = 9/26 \tag{45}$$

From these expressions we conclude that same coupling structure results as in the standard model. Some remarks concerning the coupling structure are in order. First, the right-handed neutrino decouples completely from the interactions mediated by the gauge bosons (as can be seen directly also from the original representation of the covariant derivative). Since this decoupling phenomenon is not generic we can regard the experimental absence of the right-handed neutrinos as a unique signature singling out CP_2 among the other alternatives. Secondly, the value $\sin^2\theta_W = 9/26$ [to be compared with the value 9/24 of the grand unifications (Georgi and Glashow, 1974; Mahantappa and Randa, 1980)] is acceptable provided we identify it as the unrenormalized value of the quantity. The bare value of the Weinberg angle should reveal itself at high energies, perhaps of order of Planck mass, as suggested by the grand unification schemes.

7. SYMMETRIES OF THE THEORY

In this section the symmetries of the theory will be considered. Section 7.1 is devoted to the problem of realizing the isometries of H as spinor transformations. Section 7.2 deals with the realization of discrete symmetries and some comments concerning the chiral invariance in a generalized sense are represented. In Section 7.3 a geometric analog for the Higgs mechanism is proposed.

7.1. Isometries of *H*. A quite natural expectation is that the isometries of the space H should be symmetries of the theory. However, the realization of CP_2 isometries as spinor transformations so that the spinor part of the action remains invariant is not quite straightforward. We shall approach the problem only from the infinitesimal point of view. So let $h^k \to h^k + \varepsilon j^k$ define the infinitesimal isometry. The first result needed and shown to be true in Appendix B is specific to CP_n.

Lemma. The isometries of CP_n are representable as Hamiltonian flows with respect to the symplectic form defined by the Kähler form J ($J^k{}_l J^l{}_r = -\delta^k{}_r$). The infinitesimal generators have the representation

$$j^k = J^{kl}\partial_l H \tag{46}$$

where H is the Hamiltonian associated with the flow.

The following theorem is proved in Appendix B:

Theorem. Let $H = M^4 \times CP_2$ and define the covariant derivative in X^4 using the induced spinor connection. Then the quantity $\bar{\Psi}\Gamma^{\alpha}D_{\alpha}\Psi$ is invariant under the isometries of H realized according to

$$\delta\Psi = i\varepsilon\left[\tfrac{1}{4}\partial_l j_k \Sigma^{kl} + j^k V_k + \left(j^k B_k/2 + H\right)(1_+ + 31_-)\right]\Psi \tag{47}$$

The transformation formula (27) has a simple geometric interpretation. The isometry is interpreted as flow and the spinor field is parallel translated along the flow lines: besides the usual rotation [the first term in (47)] the spinor field suffers a transformation [the second term in (47)] which has its origin on the nontrivial curvature properties of the space H. The last term is obviously absent when H has ordinary spinor structure. An important general feature of this representation is that the infinitesimal isometries are represented as infinitesimal gauge transformations. This means that the representation is not integrable in the ordinary sense but one has a representation modulo gauge transformation, which can be thought of as a generalization of the ordinary projective representation, say, for the rotation group $SO(3)$. A remarkable feature of the representation is that only a single Abelian subgroup is representable in the ordinary sense.

In order to elucidate the physical role of the $SU(3)$ isometries it is useful to study the action of the isometries using complex coordinates (ξ_1, ξ_2) for CP_2 defined in Appendix A. The action of $SU(3)$ on these coordinates can be deduced from its action on coordinate variables z^k, $k = 1,2,3$ of $\mathbb{C}^3$ (ordinary matrix multiplication). Clearly the action is nonlinear, the maximal linearly realized subgroup being $SU(2)\times U(1)$ representable as matrices of the form

$$\begin{pmatrix} U & 0 \\ 0 & \det(U^{-1}) \end{pmatrix}$$

[remember that $(\xi_1, \xi_2) = (z^1/z^3, z^2/z^3)$]. Observe that the center of $SU(3)$ is represented trivially, which implies that the CP_2 spherical harmonics must have triality zero.

How then to construct the spherical harmonics of CP_2? The construction is easily performed using the triplets $\{f^k,\ k = 1,2,3\}$ and $\{\bar{f}^k\}$, where the functions f^k are defined as

$$f^k = \xi^k/(1+r^2)^{1/2} \tag{48}$$

where one has $\xi^3 = 1$ and the phase factor U defined as

$$U = \xi^1\xi^2/|\xi^1\xi^2| \tag{49}$$

The triplets behave like $SU(3)$ triplets 3 and $\bar{3}$ apart from carrying an anomalous hypercharge $Y_A = -2/3$ and $Y_A = 2/3$, respectively. The phase factor carries an anomalous hypercharge $Y_A = -2$. One can construct irreducible representations of $SU(3)$ by forming products of the basic triplets and by canceling the anomalous hypercharge by a suitable power of the phase factor U. This corresponds to the construction of the completely symmetrized product $(3_A^m \times \bar{3}_A^n)_S$ carrying the anomalous hypercharge $Y_A = (n-m)2/3$ and the compensation of Y_A using the phase factor $U^{(m-n)/3}$. The triality zero rule follows from the one valuedness requirement for the phase factor. It should be emphasized, however, that one obtains also representations, which differ from triality nonzero representations by the associated anomalous hypercharge.

The infinitesimal action of the subgroup $SU(2)\times U(1)$ on the components of the connection is found directly from the representation of the vierbein in complex coordinates (Appendix A). Rather surprisingly, the connection and therefore also the spinor field are invariant under the action of the group $SU(2)$. Under the group $U(1)$ the field quantities transform as objects having an anomalous hypercharge $Y_A = 2Q_{em}$.

The fact that one can associate definite anomalous hypercharges both with the field quantities and with the triplets $\{f^k\}$ and $\{\bar{f}^k\}$ and the phase factor U suggests an obvious generalization of the ordinary CP_2 partial wave analysis. One simply associates with the field quantity having $Y_A = 2Q_{em}$ a generalized partial wave carrying an opposite anomalous hypercharge so that an irreducible multiplet results. As a consequence, leptons and gauge bosons necessarily correspond to the triality zero partial waves and for quarks only the partial waves with triality $t = 1, 2$ are allowed.

On basis of the observations just made the identification of $SU(3)$ as color group seems compelling. Moreover, the fact that the color degrees of freedom correspond to CP_2 translational degrees of freedom suggests a surprisingly simple explanation for color confinement or, more generally, for the experimental absence of the colored states. The point is that the scale of CP_2 turns out to be given by Planck length and therefore the uncertainty principle suggests a mass scale defined by Planck mass for the colored states. Also the fact that different members of the electroweak multiplets carry different anomalous hypercharges suggests that "symmetry breaking" indeed happens (in fact, the standard group is not a symmetry group in the proposed theory).

7.2. Discrete Symmetries and Generalized Chiral Invariance. We base our approach to the discrete symmetries C, P, and T on the following requirements. First, the symmetries should be realized as purely geometric transformations. Second, the transformation properties of the fieldlike variables should be essentially the same as in the conventional field theories (Björken and Drell, 1965). Finally, the assumption about the Grassmann valuedness of the spinor field is made so that Ψ and $\overline{\Psi}=\Psi^{+}\Gamma^{0}$ are regarded as independent dynamical variables.

The realization of the reflection P corresponding to the intuitive picture about parity breaking is

$$m^{k} \rightarrow P(m^{k})$$

$$\Psi \rightarrow \gamma^{0} \times \gamma^{0} \Psi \tag{50}$$

in the representation chosen for the gamma matrices (Appendix A). Indeed, the gauge bosons W and Z^0 break the symmetry because they do not commute with the matrix $\gamma^0 \times \gamma^0$. It is amusing that for $n_+ = n_-$ parity would be an exact symmetry, but now realized according to $\Psi \rightarrow \gamma^0 \times 1\Psi$ and transforming leptons and quarks to each other.

In case of time reflection the above-mentioned principles lead to nontrivial consequences already in the conventional theory. The purely geometric transformation formula for the $U(1)$ connection is $A_0(x) \rightarrow -A_0(Tx)$, $A_i(x) \rightarrow A_i(Tx)$ and is in variance with the physical transformation formula differing from the purely geometric one by an overall sign. Therefore the T of the physicist cannot reduce geometrically to a pure time reflection in M^4. The naive guess that also a complex conjugation in CP_2 is associated with the T transformation turns out to be correct. One can verify by a direct calculation that the action is invariant under T realized according to

$$m^{k} \rightarrow T(m^{k})$$

$$\xi^{k} \rightarrow (\xi^{k})^{*}$$

$$\Psi \rightarrow \gamma^{1}\gamma^{3} \times 1\Psi^{*} \tag{51}$$

where the symbol * denotes complex conjugation.

The operation bearing closest resemblance to the ordinary charge conjugation corresponds geometrically to complex conjugation in CP_2:

$$\xi^k \to (\xi^k)^*$$

$$\Psi \to \gamma^2 \times \gamma^0 \Psi^+ \tag{52}$$

Concerning the composite transformations one can conclude that CP and CPT are exact symmetries of the action as expected for the choice $V^4 = M^4$. The observed CP breaking (Lee, 1979; Weinberg, 1976) might result from the noninvariance of vacuum state under CP. Indeed, the classical equations of motion allow vacuum solutions, which in general are noninvariant under CP (Appendix D).

The theory allows as symmetries also the transformations

$$\Psi \to \exp(i\alpha X)\Psi \tag{53}$$

where the quantity X is a linear combination of unit matrix 1 and the matrix Γ_9. Thus we can speak about chiral invariance in a generalized sense. It is evident that the corresponding conserved currents give rise to the separately conserved lepton and baryon numbers. As a consequence the proton is absolutely stable against spontaneous decay provided the quarks are massive enough irrespective of confinement mechanism. Of course, we do not have the ordinary chiral invariance producing troubles in the ordinary gauge theories because the compactness of CP_2 introduces a natural length scale into the theory.

7.3. Higgs Mechanism Geometrically. The standard model prediction for the ratio m_W/m_{Z^0} and for m_γ are both in agreement with experiment. Owing to the sensitivity of the ratio m_W/m_Z to the symmetry-breaking mechanism it is important to try to identify the possible symmetry-breaking mechanism in our geometric setting and to look whether the standard model predictions result. The strategy is to identify the possible Higgs multiplet and then perform the transition to the unitary gauge (Abers and Lee, 1973) in order to obtain information about fermion masses.

The lack of an explicit Higgs term in the action leads to suspect that symmetry breaking is realized in some sense dynamically and hence it is natural to look whether the classical equations of motion, which certainly differ from their counterparts in the ordinary gauge theory, reveal any Higgs term. Of course, one can worry about the meaning of the spinorial equations of motion because spinors are Grassmann valued. However, the classical equations of motion are satisfied in the sense that the functional integral expectation values $\int X\exp(iS)\, D\sigma^4$ vanish for $X = \delta S/\delta y^k$ with y^k denoting either the coordinate variables of H or the spinor variables. This holds true because the functional integrand is a functional gradient.

The spinorial equations of motion read

$$i\Gamma^\alpha D_\alpha \Psi = -H/2\Psi \tag{54}$$

where the quantity H is defined as

$$H = ig^{\alpha\beta}H^k_{\alpha\beta}\Gamma_k \equiv H^k\Gamma_k \tag{55}$$

The components $H^k_{\alpha\beta}$ define the second fundamental form for δX^4 [equation (4)]. The term in (52) behaving like a Higgs term is given by the expression

$$M = ig^{\alpha\beta}S^k_{\alpha\beta}\gamma_5 \times \gamma_A e^A_k \tag{56}$$

Constructing the quantity $\bar{\Psi}M\Psi$ one can verify that M behaves like a Hermitian scalar field. Moreover, M behaves like a vector under the $SO(4)$ transformations. Hence it is attractive to interpret M as a classical counterpart of a Higgs field.

Next we address ourselves to the question, what kind of symmetry breaking might result from the nonvanishing vacuum expectation value for the quantity M. Clearly, the action does not contain any Higgs term, but if one accepts that the interactions between gauge bosons and ours might be Higgs are describable using ordinary field theory in M^4 using the effective action

$$L_{\text{eff}} = L_{\text{YM}} + \text{Tr}(D^\alpha M D_\alpha M) \tag{57}$$

where $D_\alpha M = \partial_\alpha M + [A_\alpha, M]$ is the covariant derivative dictated by the vector nature of M and A_α and M are understood as ordinary fields in M^4, one obtains predictions for the crucial quantities m_W/m_Z and m_γ. A nonvanishing diagonal expectation value for M (we shall show that diagonalization is always possible),

$$M = M^0\gamma_5 \times \gamma_0 + M^3\gamma_5 \times \gamma_3 \tag{58}$$

leads to effective mass terms for the gauge bosons in (35). The photon remains massless because M commutes with γ and for the above-mentioned ratio one obtains

$$m_W/m_Z = g_Z/g_W \tag{59}$$

just as in the standard model. The form of the mass matrix M also suggests the mass formula $M = M_0 + M_I I_3$ for elementary fermions.

In the preceding argument we assumed that the diagonalization of the matrix M is possible. Since this diagonalization or the transition to the unitary gauge, as it is called in the standard model (Abers and Lee, 1973), besides revealing the mass spectrum of the theory particularly clearly, also shows how the charged degrees of freedom associated with the Higgs field transform to the longitudinal degrees of freedom of the massive gauge bosons, it is important to show that this procedure is indeed possible always. Clearly the diagonalizing transformation is a certain element of the gauge group $SO(4)\times U(1)$ so that seven parameters are associated with it. There are rather natural conditions for the transformation, call it g, to obey. First, the transformation should leave the charged gauge bosons left handed. This is certainly true if g belongs to the subgroup $SU(2)_L \times U(1)_R \times U(1)$ (five parameters), where $U(1)_R$ is diagonal $U(1)$ subgroup of $SU(2)_R$. Second, one must require that the quantities $X = (V_{12} + V_{03})/3$ and $B/2$, which are equal, transform in the same way under g, in order to retain the meaning of the definition of the neutral gauge bosons. Since X is purely right handed this implies that $U(1)_R$ and $U(1)$ parts of g are simply related, e.g., $U(1)$ part is simply some power of $U(1)_R$ part. So, we are left with four remaining parameters. Thirdly, the photon should remain invariant under g. This requirement fixes two parameters in g since both the inhomogenous and homogenous part in the transformation formula of the photon fix one parameter. Finally, the remaining two parameters are fixed by the requirement that the vector M lies in the (0,3) plane so that the components M_1 and M_2 vanish.

Concerning the detailed nature of the Higgs mechanism one can only state some obvious questions and propose clues, which might lead one to answer to these questions. The basic questions are the following: What do we exactly mean when we say that the quantity M develops a nonvanishing vacuum expectation value? Is the Higgs mechanism a single particle or a many-body phenomenon in the proposed theoretical framework? The appearance of the Higgs term in the spinorial equations of motion and the fact that spinors are restricted on the boundaries suggest that the Higgs mechanism is a single-particle phenomenon as far as elementary fermions are considered. On the other hand, the classical equations of motion allow vacuum solutions with a nonvanishing matrix M (Appendix D, Section D2) and one might argue that the so-called ♯ condensation around these vacuons (see Section 10) giving rise to the classical space-time, as we propose, gives rise also to the Higgs phenomenon.

8. STRONG INTERACTIONS AND QCD

This section is concerned with the description of the strong interactions in the framework of the CP_2 theory. Section 8.1 is devoted to the description of the hadron spectroscopy. In Section 8.2 a unified semiclassical description of the hadrons as stringlike objects is proposed. In Section 8.3 the relation between the proposed theoretical framework and the QCD approach is discussed.

8.1. Classification of the Strongly Interacting Particles. CP_2 theory suggests a particle classification based on the use of (a) the quantum numbers associated with a single-particle generation in the standard model of the electroweak interactions, (b) the topological charges g and h, and (c) the labels associated with the partial waves of CP_2 transforming irreducibly under the isometry group $SU(3)$.

Concerning the explanation of the color there are two possible approaches, e.g., the group-theoretical and the topological approach. As was found in the context of symmetries, the group-theoretical explanation turns out to be the correct one since quarks necessarily correspond to the irreps of $SU(3)$ with a nonvanishing triality. Moreover, since the color is related to the translational degrees of freedom of CP_2 and since the scale of CP_2 is given by the Planck length, the uncertainty principle suggests that the mass scale of the colored states is given by the Planck mass. Of course, the group-theoretic explanation of the color is favored also by the fact that it makes room also for the gluons as boundary components having spin 1 and moving in octet partial wave of CP_2.

The approximate spin flavor symmetry (Close, 1979, Chap. 4) revealing itself via the multiplet structure and magnetic moments associated with the low-lying hadrons is easy to understand, when it is realized that the low-lying hadrons correspond to states completely symmetric with respect to the one-particle label (g, I_3,spin) and as such may be regarded as a basis for an irreducible representation for the appropriate spin flavor group [$SU(6)$ for the noncharmed hadrons]. Of course, this group has a purely formal meaning. It is the transformation properties of the individual states under permutations of the labels g, I_3, spin, which lead to an illusion about the presence of a real symmetry.

The suggested explanation for the generation phenomenon explains also the well-known $\Delta I_s = 1/2$ rule observed in the decay of the strange particles (Abers and Lee, 1973, p. 28). From the preceding it is clear that the concept of the strong isospin is somewhat artificial and indeed it is possible to label the hadron states using the weak isospin having a genuine group-theoretical interpretation. In general the multiplet assignments obtained using I_s^3 and I^3 differ, but in such a way that the above-mentioned nonconservation rule becomes a conservation law for the weak isospin! For instance, the kaon doublets decompose into a triplet and singlet with respect to the weak isospin and therefore in the decay $K^0 \rightarrow \pi^+\pi^-$ the final state with $I = 2$ is suppressed relative to that with $I = I_s = 0$.

Cabibbo mixing is another peculiar phenomenon, where both the weak and the strong interactions are involved (Kobayashi and Maskawa, 1973). As already noticed, our theory predicts a topological transition changing the genus of the boundary component by one unit. Therefore we expect mixing between the different particle generations to take place. Because of the symmetry breaking the mixings are expected to be different for the fermions with different weak isospin and thus to be observable in weak transitions mediated by the charged bosons $W^{+(-)}$. Stating this more explicitly, we can describe the mixings as

$$U_i \rightarrow U_i{}^j U_j = \overline{U}_i$$

$$D_i \rightarrow D_i{}^j D_j = \overline{D}_i \tag{60}$$

where the matrices $U_i{}^j$ and $D_i{}^j$ are unitary. Since the emission of the weak boson does not change the boundary topology the amplitude for the process $\overline{U}_i \rightarrow \overline{D}_j$ must be proportional to the quantity $U_{ir}^* D^r{}_j$, which is nonvanishing unless the mixings suffered by U_i and D_i differ only by phase multiplications performed for "initial" and "final" states. Of course, the mixings need not be the same for $h = 0$ and $|h| = 1$ quarks. The assumption that only the boundary components with $|h| = 1$ mix noticeably would explain nicely the radically different behavior of quarks and leptons as far as generation changing transitions are considered.

The absence of the neutral strangeness changing currents (Abers and Lee, 1973, p. 28) is also a peculiar phenomenon having a simple explanation as a manifestation of the $g = 0$ nature of the weak gauge bosons.

8.2. Semiclassical Description of Hadrons. The proposed theory leads to a unified semiclassical description of the hadrons as stringlike objects as shown in Sections C4 and C5 of Appendix C. The classical equations of motion allow solutions of the form $X^4 = X^2 \times S^2$, where X^2 is a minimal surface in M^4 having interpretation as an open or closed string and S^2 is a geodesic sphere in CP_2 (there are two nonequivalent, e.g., not isometrically related, geodesic spheres in CP_2). In the case of an open string one can add spinors to the boundaries and thus obtain a model for pion (ρ-meson). It is remarkable that one can build all the known hadrons from the pionic string. First, on can generate nucleons simply by drilling a hole in pionic string (having necessarily $h = 0$). Second, one can generate hadrons containing higher generation quarks by simply drilling "wormholes" starting from and ending at the boundary component in question and by adding the spinors.

The 4-momentum associated with the interior part of this universal hadronic solution can be expressed in the same form as that associated with the ordinary string Regge slope α_R has the expression

$$\alpha_R = \left(l_i/g^2R^2 - 1/16\pi G\right)8\pi \tag{61}$$

where the parameter l_i obtains the values $l_{\mathrm{I}} = 9$ and $l_{\mathrm{II}} = 1$ corresponding to the two possible geodesic spheres of S^2. Thus the theory is fixed when one chooses either of the two alternatives fixing the magnitude of the parameter gR. The positivity requirement for the energy favors the alternative II (the strings of type I have Regge slope of order $1/G$ and probably also their ground state masses are of the order of the Planck mass).

It is rather surprising that the energy of the string results from cancellations between enormous magnetic and gravitational energies, the magnetic energy being associated with a diagonal and a nondiagonal Abelian gauge field in the cases I and II, respectively. Moreover, the energy of the string has nothing to do with gluons as one might expect.

In the proposed picture the homology charge clearly plays the role of a coupling constant in the planar dual diagrams since only the ends of the strings are involved in these diagrams. So, as far as the strong interactions mediated by the dual diagrams are considered, only the $|h| = 1$ quarks are active and a diquark picture of hadrons, where the homologically nontrivial quarks behave like a single mesonlike unit, is suggested. Since homology charge endows quarks with additional dynamical degrees of freedom, one might argue that it equips quarks with an additional degeneracy, which should be seen in the famous ratio $R = \sigma(e^+e^- \to \text{hadrons})/\sigma(e^+e^- \to \text{muons})$. One can, however, represent the following counterarguments. First, the degeneracy associated with the sign of the homology charge is present only if one assigns a definite orientation to the various boundary components. This is not even possible, when the hadronic 3-manifold is nonorientable (a "twist" in the dual diagram makes the associated two surface X^2 in M^4 nonorientable). Second, the degeneracy associated with the $h = 0$ quarks is present only if the closed strings with two homologically trivial boundary components contribute to the annihilation vertex $\gamma \to q\bar{q}$. These kinds of decays might be suppressed simply by the massiveness of these kinds of hadrons. Also the requirement of causality might deny this decay: the closed string must contract to a point in the annihilation vertex so that the interior of the corresponding 4-manifold must become totally spacelike near the annihilation vertex.

Certainly the picture just outlined is an oversimplification. In particular, one might wonder, where are the gluons and sea quarks? It is rather amusing that semiclassical considerations give a strong support also for the existence of these objects. Indeed, the study of the small perturbations around the static string shows the existence of topologically nontrivial perturbations. These arise from the transversal perturbations of the string, which can be expressed as superpositions of the separable perturbations satisfying the equations

$$\Box_X 2\delta m_T^k = \kappa\lambda m_T^k$$

$$\Box_S 2\delta m_T^k = \lambda m_T^k \tag{62}$$

Here the subscript T expresses the transversality of the perturbations.

In the special case $\lambda = 0$ a general solution to these equations can be written in the form

$$m_T^k = f_+^k(z, m_+) + f_-^k(z, m_-) + \text{c.c.} \tag{63}$$

where $f^k_{+(-)}$ is an analytic function of the complex coordinate z labeling the points of the geodesic sphere S^2 and $m_\pm = m^0 \pm m^3$ are the light cone coordinates for the static string. The crux of the argument is that any nonconstant analytic function has singularities, e.g., poles and cuts. To make the perturbation finite one must eliminate the polelike singularities by cutting a hole around the pole. This in turn introduces a "hole" in the hadron. In the cut the perturbation is already discontinuous, which means the presence of the "hole." Now, the parameters describing the position of the pole or cut in general depend on the coordinate m_+ or m_- and as a consequence these singularities and therefore also the associated boundary components move with the velocity of light or are stationary (the dependence on m_+ of m_- can be also trivial).

The interpretation of the nonstationary boundary components as the classical counterparts of gluons is suggestive, since they are associated with the transversal perturbation δm_T^k, which is a 4-vector with two physical components and satisfies a massless wave equation. These are indeed properties characterizing a massless spin-1 particle in the conventional field theory. It should be stressed, however, that also the more general perturbations are allowed, for which a massive wave equation is satisfied. The singularities associated with these perturbations might be interpreted as "off mass shell gluons" (it is the analyticity, which might give a preferred role for the massless perturbations). By putting spinors on the stationary boundary component one obtains a candidate for a sea quark.

8.3. QCD Aspect of Strong Interactions. Quarks must necessarily move in triality nonzero CP_2 partial waves and, of course, the triplet partial waves are the simplest possible. We have also found evidence for the existence of "holes" in hadrons behaving like massless, spin-1 particles. Furthermore, the interaction between the quarklike and bosonic boundary components must take via $\natural_{\text{int}}$ or $\natural$ vertex, depending on whether the boundary components in question belong to same hadron or not (Figure 4 illustrates the interaction vertices). It is worth emphasizing that the two vertices are locally equivalent and so the associated couplings should be the same. Obviously, there is no reason why the quarks could not exchange their color quantum numbers via the emission of absorption of the bosonic boundary components so that they are good candidates for the carriers of CP_2 octet partial waves. Of course, one cannot exclude the presence of other partial waves too, in particular $SU(3)$ singlets. It is to be expected, however, that the bondary components carrying higher partial waves are unstable against decay to the boundary components carrying lower partial waves.

Why, then, should QCD describe (at least approximately) the interaction between gluons and quarks? There are rather general arguments in favor of this idea. First, the interaction between quarks and gluons proceeds via the $\natural$ vertex as do proceed also the interactions between other gauge bosons and fermions. Second, the gauge theory is the only known renormalizable theory describing the interactions between massless spin-1 particles and spin-1/2 fermions.

We close with some remarks concerning the differences between QCD and the proposed approach to strong interactions. First, the gluons in QGD are topological excitations in hadrons or in intermediate states formed from several hadrons using the operations $\sharp$, $\sharp_B$, and $\natural$. This, of course, does not prevent us from constructing the analogs for the QCD diagrams used in the modeling of e^+e^- annihilation to hadrons, of deep inelastic scattering, and of high p_T scattering. Second, the duality approach to strong interactions seems to have nothing to do with the nonperturbative QCD. Rather the dual diagrams and QCD-like diagrams both contribute to the hadronic reactions In fact, it is the homology charge (or equivalently the magnetic charg(associated with the Abelian gauge field), which plays the role of th(coupling constant in planar dual diagrams. Moreover, the "confining poten tial," e.g., the energy of the string has nothing to do with gluons. Finally, th confinement is expected to be a purely kinematical phenomenon: colore states should have mass of order $1/G^{1/2}$ on the basis of the uncertainty principle. Observe, however, that the energy of the string is proportional to its length. As a consequence one has another type of confinement phenomenon: the length of the string cannot become arbitrarily large.

Summarizing, in QGD framework the QCD is expected to describe only a certain aspect of the strong interactions, which is seen clearly in the processes, where the duality diagrams do not contribute, such as high p_T reactions and Zweig forbidden processes, for instance, upsilon decay. QCD is expected to be just the perturbative QCD!

9. GRAVITATIONAL INTERACTION AND QGD

QGD seems to offer two approaches to the description of the gravitational phenomena. We shall introduce first these approaches in Sections 9.1 and 9.2 and in Section 9.3 we shall propose a unification of these approaches based on the idea that the classical space-time is in a certain sense a topological many-particle phenomenon in the QGD framework.

9.1. Quantum Approach to Gravitation. The first approach to the description of the gravitational phenomena is based on the observation that the classical equations of motion allow massless solutions with S^3 topology and having rotational symmetry about the direction of the wave vector (Section C6, Appendix C). The last property makes it impossible to attach any vector polarization to the solution. It is however possible to assign a nontrivial tensor quantity to the solution, given for instance by the deviation of the metric from the flat metric. Therefore the interpretation of these solutions as classical gravitons is suggestive.

These particles can interact with the ordinary matter only via the $\sharp$ vertex and therefore they couple to all matter. A remarkable property possessed by the vertex is that it becomes impossible for generic 3-manifolds in the limit when H is contracted to M^4 (e.g., the "radius" of S goes to zero). Thus the continuity argument suggests that the associated coupling constant should be proportional to a positive power R^n of the "radius" of S so as to vanish in this limit. Since the simplest couplings of spin-0 and spin-1 particles to matter fields are dimensionless (Yukawa and gauge couplings, respectively) it seems that also the dimensional argument favors the spin-2 nature of our might-be-gravitons.

The effective action describing the interaction of massless spin-2 particles with matter is that of GRT as Weinberg has shown using rather general arguments (Papini and Valluri, 1977). It goes without saying that the that the neutral test particle in a ♯ condensate moves along a geodesic line of the condensate, and (c) that the Poincaré energy of the neutral test particle in a stationary ♯ condensate is given by $P^0 = E/g_{00}^{1/2}$, where E is the conserved energy associated with the particle in GRT so that in the Newtonian limit the Poincaré energy corresponds to the kinetic energy of the particle, one obtains the classical predictions of GTR for the gravitational phenomena in planetary scale (gravitational red shift, delay of the radar echo, the motion of the perihelion) (Misner et al., 1973; Adler et al., 1975).

Of course, the assumptions (a), (b), and (c) imply that the test particle changes Poincaré energy and momentum with the ♯ condensate and this, we propose, takes place via the emission of gravitons, which of course themselves suffer ♯ condensation (it should be noticed that gravitons differ from the other elementary particles in that their presence does not change the topology of the ♯ condensate).

Summarizing, the semiclassical and topological considerations suggest the following picture. First, the gravitational as all the other interactions are mediated by the exchange of particles understood as appropriate submanifolds of H. Second, the classical space-time is a topological many-particle phenomenon: particles form a ♯ condensate around a vacuon, which can have a macroscopic scale. The only difference between the interactions of free and ♯ condensed particles is that the various exchanged particles are free and move in the condensate, respectively.

10. COSMOLOGY AND QGD

The decomposition $H = V^4 \times CP_2$, where V^4 is flat is favored by the compactness requirement for the gauge group of the theory. V^4 could therefore be either Minkowski space M^4 or its light cone M^4_+. The latter choice is however favored for various reasons.

First, this choice implies the big bang cosmology (Weinberg, 1977; Dolgov and Zeldovich, 1981) provided the rather natural assumption that nothing enters from "outside" to M^4_+ is made. Indeed, the assumption implies that the particle at point $\{m^k\}$ has the average 4-velocity proportional to the 4-vector m^k and therefore the cosmological red shift results. Also the gross features of the cosmic thermal history follow from this assumption.

Second, certain symmetry considerations favor this choice. Clearly, the only exact isometries of the light cone are the Lorentz transformations about the origin. Obviously the breaking of translational and Lorentz invariance has no consequences observable in the laboratory scale. However, the breaking of the macroscopic time reversal invariance finds a nice explanation: the retarded boundary conditions used for the macroscopic fields follow from the requirement that nothing enters into M^4_+ from "outside." It should be noted, however, that the breaking of translational, time reversal and CPT symmetry becomes microscopic at the very early times and should be taken into account in cosmological modeling.

Next we consider in more detail the basic features of the resulting cosmology. It deserves to be noticed that M^4_+ corresponds to empty, hyperbolic cosmology with a Robertson–Walker-type metric (Misner et al., 1973; Adler et al., 1975)

$$ds^2 = dt^2 - t^2\left[dr^2/(1+r^2) + r^2\,d\Omega^2\right] \tag{64}$$

where the variables t and r are related to the M^4 coordinates via the formulas

$$\begin{aligned} t &= \left[(m^0)^2 - r_M^2\right]^{1/2} \\ r &= r_M/t \end{aligned} \tag{65}$$

where r_M denotes the radial coordinate of M^4. Clearly, t is the proper time for the world line connecting the point $\{m^k\}$ to the origin and is Lorentz invariant. The "nothing from outside" assumption obviously implies that matter is comoving in these coordinates (on the average, of course). Moreover, the assumption of isotropy and homogeneity (e.g., the so-called cosmological principle) corresponds to the invariance of the matter distribution with respect to the Lorentz transformations leaving the origin fixed. The Hubble constant is given by $H = t^{-1}$ and thus the age of the universe is predicted to be rather large: $t_U \sim 18\times 10^9$ yr. Of course, gravitational corrections could change both the definition of the cosmic time and the relation between H and cosmic time so that the age of the universe gets shortened.

The most natural approach to the history of the universe in the proposed scheme is to neglect the effects related to the $\sharp$ condensation and to do cosmology in fixed M^4_+ background. This approximation is expected to be reasonable for sufficiently early times because the high temperature is expected to make the $\sharp$ condensate an unstable phase. More concretely: the assumption means only that we drop the basic equation of the GRT cosmology. $T^{\alpha\beta} = \kappa G^{\alpha\beta}$, idealize particles with pointlike particles in M^4_+, and write the kinetic equations (Dolgov and Zeldovich, 1981) governing the time behavior of the Lorentz-invariant particle densities in the fixed M_+^4 background.

The principal differences between the proposed and the usual GRT cosmology are the following. First, M_+^4 has no particle horizons, which means that the isotropy of the 3-K radiation poses no problems (Dolgov and Zeldovich, 1981). Also, concerning the explanation of matter–antimatter asymmetry (Dolgov and Zeldovich, 1981) in the QGD context this feature has a welcome consequence. Since baryon number is conserved in QGD, the asymmetry must be a local phenomenon, perhaps related to the ♯ condensation and the CP breaking associated with the vacuum solutions (Sections D1), which act as condensation centers. In GRT cosmology the main argument against the local matter–antimatter asymmetry (Dolgov and Zeldovich, 1981) is that the regions ontaining only matter–antimatter must have been formed rather early ($t < 10^{-3}$ s) and so the galaxies should contain typically the number of nucleons inside the maximal causally connected region at the time of the matter–antimatter separation, typically about 10^9 nucleons. Now, of course, the radius of the minimal causally connected region is infinite.

A second difference is that now one does not have Einstein equations, which together with the equation of state fix the relation between cosmic time and temperature. One can however fix the dependence between temperature and the cosmic time at early times by assuming that the matter is radiation dominated, e.g., the matter density is proportional to the fourth power of the temperature T and that the time evolution is adiabatic. As a consequence one obtains the relation

$$T = C/t \tag{66}$$

where the constant C remains undetermined now and can be fixed from the observed helium abundance of the universe. So this simple approach does not predict helium abundance unlike GRT cosmology (Dolgov and Zeldovich, 1981). A welcome consequence is that in this approach no bound for the number of the light neutrino types is predicted either (the number is, of course, not finite if the generation genus correspondence is accepted) (Dolgov and Zeldovich, 1981).

It is rather natural to ask whether it could be possible to idealize away all details of the matter distribution and obtain models for the universe as Lorentz-invariant, graphlike solutions to the field equations. In Section D4 it is shown that this kind of solutions can be found but that their energy density is negative so that any simple-minded cosmological interpretation is impossible.

11. SUMMARY AND CONCLUSIONS

The basic features of the suggested elementary particle theory deserve a short discussion.

(a) Topological Level. Particles are identified as three-dimensional submanifolds of some metric space H. The hypothesis leads to a rough particle classification using the boundary topology and a topological explanation for the generation degeneracy is proposed. Also the topology of H brings an important element to the particle classification. The boundary components of a 3-manifold are classified by their homology equivalence classes in the group $H_2(H)$, which for $S = CP_n$ is isomorphic to integers. It turns out that the baryonic and mesonic valence quarks indeed carry homology charges $\{1, -1, 0\}$ and $\{1, -1\}$, respectively, but that the attempt to explain color homologically fails.

A topological classification of the interaction vertices is performed leading to a generalization of the dual diagrammatics so that also the 3-particle vertices of field theory find their topological analogs.

(b) Construction of the Dynamics. The construction of the dynamics of the model is based on three basic hypotheses. First, the boundary components of a 3-manifold representing particles are carriers of various dynamical charges besides the topological ones. Secondly, the theory should have the formal structure of an Einstein–Yang–Mills theory defined on X^4 and finally, the YM structure and the accompanying metric and spinor structure should result from the natural geometric structures of the space H. The mathematical device used to obtain these structures on X^4 is the so-called induction procedure. It is remarkable that the decomposition $H = M^4_{(+)} \times S$ provides an elegant way to avoid the pathologies associated with the noncompact gauge groups.

(c) Quantization of the Theory. A correspondence principle based on the geometric representation of an ordinary field theory is used as a guideline in a heuristic formulation of the quantized theory. The essence of the correspondence principle is that the Feynman propagator $G(x, y)$ is expressible as a path integral over the paths from x to y. The suggested theory can be thought to present a generalization of a field theory, obtained by thickening the paths contributing to the above-mentioned functional integral to 4-manifolds in H or equivalently, by generalizing the particle concept (three-dimensional manifold instead of a point particle). Using this correspondence principle we obtain a description for the generalized Green's functions, a rather unique definition of one-particle states, and finally, a formal definition of transition amplitudes and probabilities.

(d) Choice of H. Accepting the topological explanation for the generation degeneracy (generation genus correspondence) and the group-theoreti-

cal explanation for color (the isometries of S give rise to color symmetry), the only remaining quantum numbers needed for the classification of the observed particles are those associated with the standard model of the electroweak interactions. It is remarkable that the choice $S = CP_2$ probably uniquely leads to the coupling structure of the standard model predicting for the Weinberg angle the value $\sin^2\theta_W = 9/26$. Moreover, the isometry group of CP_2 is indeed $SU(3)$. The right-handed neutrinos are predicted to decouple totally from the gauge interactions. Furthermore, the baryon and lepton numbers are separately conserved as a consequence of the generalized chiral invariance so that proton is absolutely stable against the spontaneous decay provided the colored states are massive enough.

(e) Classification of the Strongly Interacting Particles. Concerning the classification of the strongly interacting particles a few remarks are in order. First, the classification uses only the topological quantum numbers g and h besides the quantum numbers associated with the gauge group of the standard model and with the color group and leads to the understanding of the broken flavor spin symmetry as an effective symmetry resulting from the properties of the hadronic states under the permutations of the labels I_3 and g. Second, the basic peculiarities associated with the weak interactions of hadrons—Cabibbo mixing, absence of the neutral strangeness changing currents, and the $\Delta I = 1/2$ rule—find their natural explanation in the suggested scheme. Finally, even if the valence quarks turn out to be carriers of homology charges $|h| = 1, 0$, the correct explanation of the color is group theoretical. The key observations in this respect are the following. First, various field quantities are not quite $SU(3)$ singlets but carry an anomalous hypercharge $Y = 2Q_{em}$. Second, one can generate from the coordinate variables of CP_2, besides the irreducible triality zero representations, also pseudorepresentations with nonvanishing triality but with an anomalous hypercharge. Associating with field quantities partial waves so that the total anomalous hypercharge vanishes one finds that quarklike and leptonic state functionals transform according to $t = 1, 2$ and $t = 0$ representations of $SU(3)$. Moreover, quarks and antiquarks correspond to triality 1 and -1 representations, respectively, so that fractionally charged states are always colored.

(f) Dynamics of the Strongly Interacting Particles. As regards the dynamics of the strongly interacting particles, some remarks should be made. First, with respect to the electroweak interactions quarks are expected to behave like fractionally charged pointlike particles so that the theory should be consistent with the parton picture used in the description of the various high-energy phenomena (the idea of spinors, in particular, quarklike spinors on boundaries, was partly motivated by the parton picture). Second, stringlike objects appear as solutions to the classical equations of motion and one

obtains also baryonic strings. As a consequence the theory leads to a generalization of dual diagrams as a graphical description of the strong interactions so that baryons find their natural place in this description. Moreover, homology charge plays the role of coupling constant in planar dual diagrams. Thirdly, the elimination of the singularities associated with the transverse excitations of the string leads to topologically nontrivial excitations (holes in a hadron), which represent natural candidates for gluons. The fact that these might be gluons interact with quarks via the ♮ vertex makes QCD a natural candidate for an approximate description of the quark gluon interactions. It should be emphasized, however, that in the proposed scheme QCD is expected to describe only a single aspect of strong interactions. Stated more precisely: dual diagrams are not only a phenomenological description of the nonperturbative QCD effects but rather it is the topological generalization of the dual diagrammatics which gives rise also to QCD-like diagrams in the description of the strong interactions. Finally, the fact that color degeneracy is closely related to the translational degrees of freedom of CP_2 suggests a purely kinematical explanation for the color confinement. Colored states perform a nontrivial center-of-mass motion in CP_2 degrees of freedom and since the radius of CP_2 is of the order of Planck length, the mass scale of the colored states is most naturally given by Planck mass (uncertainty principle).

(g) Gravitation. Concerning the description of the gravitational phenomena in the proposed framework, the basic problem to be understood is, why the classical, general relativistic space-time is so useful a concept. It is proposed that the classical space-time can be understood as a kind of a topological many-particle phenomenon. First, the theory allows both particle and vacuumlike solutions to the equations of motion so that the latter can have an arbitrarily large scale. Second, when $\dim H < 9$ the transition particle $\cup$ vacuon $\rightarrow$ particle $\sharp$ vacuon takes place with high probability. We call this phenomenon $\sharp$ condensation. Moreover, the theory allows graphlike solutions with a metric, which is asymptotically Schwartzchild metric. The metric in the asymptotic region could be interpreted as that created by the matter, which has suffered $\sharp$ condensation around vacuon.

The utmost importance of gravitation in understanding the gross features of the elementary particle mass spectrum comes as a surprise. Indeed, depending on whether one adds to the action the curvature scalar term or not, the prediction that the scale of CP_2 should be given by Planck length or by a typical elementary particle length, follows!

(h) Cosmology. The basic cosmological facts afford a crucial test for the theory. Indeed, the choice $V^4 = M^4_+$, e.g., the light cone of M^4, makes the big bang cosmology a kinematical necessity and at the same time allows Poincaré invariance at the laboratory scale. Moreover, the light cone cos-

mology avoids the problems associated with the finite horizons in GRT-based cosmologies.

(i) Problems. Some remarks concerning the open problems in the theory are in order. Besides the formidable technical problems related to the quantization of the theory there are various aspects, which might/should be understood before the quantization of the theory. First, the semiclassical description of leptons represents an open problem. The pointlike nature of the leptons suggests that at least the charged leptons should correspond to the classical solutions of the form $M^1 \times X^3$, where M^1 is a geodesic of M^4 and X^3 is a submanifold of CP_2. However, the possibility that leptons correspond to the closed strings of type I (Section C4) having Regge slope of order G is not completely excluded. Second, the qualitative features of the mass spectrum should be understood. Semiclassical considerations suggest that the mass scale of a generic particle is given by Planck mass so that the masses of ordinary elementary particles result from rather miraculous cancellations (topological simplicity, color neutrality). Concerning the more detailed features of the mass spectrum such as the dependence of the lepton and quark masses on generation index, the situation is completely open. Thirdly, since baryon and lepton numbers are separately conserved, the explanation of the matter–antimatter asymmetry represents a crucial test for the theory. Because vacuum solutions and also the Schwartzchild-type solutions break CP symmetry, it is tempting to conjecture that the asymmetry as well as the CP breaking observed in a $K\bar{K}$ system are closely related to the phenomenon of $\sharp$ condensation. Finally, there is a problem created by the theory itself. Poincaré energy turns out to be a nonpositive definite quantity [membranelike solutions and Lorentz-invariant solutions (Sections C3 and D4, respectively)].

Summarizing, we suggest that the unified description of the basic interactions might proceed, not via the naive extension of the gauge group, but via the generalization of the elementary particle concept itself, so that the topological concepts play a fundamental role both in classification of the elementary particles and in the description of their interactions.

APPENDIX A: BASIC PROPERTIES OF CP_2

A1. CP_2 as a Manifold. CP_2 or the complex projective 2-space is defined by identifying the points of the complex 3-space $\mathbb{C}^3$ under the equivalence

$$(z_1, z_2, z_3) \equiv \lambda(z_1, z_2, z_3) \tag{A1}$$

Here λ is any nonzero complex number. The pair z_i/z_j for a fixed j and

$z_j \neq 0$ defines a complex coordinate chart for CP_2. As j runs from 1 to 3 one obtains an atlas of three charts covering CP_2, the charts being holomorphically related to one another (e.g., CP_2 is a complex manifold). The points with $z_3 \neq 0$ form a subset of CP_2 homeomorphic to R^4 and the points with $z_3 = 0$ a set homeomorphic to $CP_1 = S^2$. Therefore CP_2 is obtained from R^4 by adding "a 2-sphere at infinity."

Besides the complex coordinates $\xi^i = z^i/z^3$, $i = 1,2$, the coordinates of Eguchi and Freund (Eguchi et al., 1980) will be used and their relation to the complex coordinates is given by

$$\xi^1 = z + it$$
$$\xi^2 = x + iy \tag{A2}$$

These are related to the "spherical coordinates" via the equations

$$\xi^1 = r\exp[i(\psi + \phi)/2]\cos(\theta/2)$$
$$\xi^2 = r\exp[i(\psi - \phi)/2]\sin(\theta/2) \tag{A3}$$

The ranges of the variables r, θ, ψ, and ϕ are $[0, \infty]$, $[0, \pi]$, $[0, 4\pi]$, and $[0, 2\pi]$, respectively.

Considered as a real four-dimensional manifold, CP_2 is compact and simply connected, with Euler number 3, Pontryagin number 3, and second Betti number $b_2 = 1$. The last property stems from the fact that the second homology group $H_2(CP_2)$ is isomorphic to integers.

A2. Metric and Kähler Structure of CP_2. In order to obtain a natural metric for CP_2 observe that CP_2 can be thought of as a set of the orbits of the isometries $z_j \to \exp(i\alpha)z_j$ on the sphere S^5: $\sum_{k=1}^{3} |z_k|^2 = R^2$. The metric of CP_2 is obtained by projecting the metric of S^5 orthogonally to the orbits. Therefore the distance between two points of CP_2 is that between the representative orbits in S^5. The line element has the following form in complex coordinates:

$$ds^2 = g_{a\bar{b}}\, d\xi^a\, d\bar{\xi}^b \tag{A4}$$

where the Hermitian metric $g_{a\bar{b}}$ is defined by

$$g_{a\bar{b}} = R^2 \partial_a \partial_{\bar{b}} \ln F \tag{A5}$$

Here the quantity F is defined as

$$F = 1 + \left(\sum_k \xi^k \bar{\xi}^k \right) = 1 + r^2 \tag{A6}$$

An explicit representation of the metric is given by

$$ds^2/R^2 = \left(dr^2 + r^2\sigma_3^2\right)/F + r^2\left(\sigma_1^2 + \sigma_2^2\right)/F^2 \tag{A7}$$

where the quantities σ_k are defined as

$$\sigma_1 = (1/r^2)\mathrm{Im}(\xi^1 d\xi^2 - \xi^2 d\xi^2), \qquad \sigma_2 = -(1/r^2)\mathrm{Re}(\xi^1 d\xi^2 - \xi^2 d\xi^1)$$
$$\sigma_3 = -(1/r^2)\mathrm{Im}\left(\sum \xi^k d\bar{\xi}^k\right) \tag{A8}$$

The vierbein forms e^A are given by

$$\begin{aligned} e^0 &= dr/F \\ e^1 &= r\sigma_1/F^{1/2} = r\,d\theta/2F^{1/2} \\ e^2 &= r\sigma_2/F^{1/2} = r\sin\theta\, d\phi/2F^{1/2} \\ e^3 &= r\sigma_3/F = r[d\psi + \cos\theta/2\, d\phi]/2F \end{aligned} \tag{A9}$$

The vierbein connection associated with the vierbein forms (A9) is given by

$$\begin{aligned} V_{01} &= -e^1/r, & V_{23} &= e^1/r \\ V_{02} &= -e^2/r, & V_{31} &= e^2/r \\ V_{03} &= (r - 1/r)e^3, & V_{12} &= (2r + 1/r)e^3 \end{aligned} \tag{A10}$$

The components of the curvature are constant

$$\begin{aligned} R_{01} &= e^0 \wedge e^1 - e^2 \wedge e^3, & R_{23} &= -e^0 \wedge e^1 + e^2 \wedge e^3 \\ R_{02} &= e^0 \wedge e^2 - e^3 \wedge e^1, & R_{31} &= -e^0 \wedge e^2 - e^3 \wedge e^1 \\ R_{03} &= 4e^0 \wedge e^3 + 2e^1 \wedge e^2, & R_{12} &= 2e^0 \wedge e^3 + 4e^1 \wedge e^2 \end{aligned} \tag{A11}$$

The metric defines a real, covariantly constant, and therefore closed 2-form J:

$$J = -ig_{a\bar{b}}\, d\xi^a \wedge d\bar{\xi}^b \tag{A12}$$

Because J is closed CP_2 is by definition a Kähler manifold. The form J defines in CP_2 a symplectic structure because it satisfies

$$J^{kl}J_{lm} = -\delta^k{}_m \tag{A13}$$

The form $2J$ is integer valued and by its covariant constancy satisfies free Maxwells equations. Hence it can be regarded as a curvature form of a $U(1)$

connection B carrying a magnetic charge of one Dirac unit $1/2g$. Locally one has therefore

$$J = dB/2 \tag{A14}$$

It should be noticed that the magnetic flux of J through a 2-surface in CP_2 is proportional to its homology equivalence class, which is integer valued. The explicit representations of B and J are

$$B = 2re^3$$
$$J = 2(e^0 \wedge e^3 + e^1 \wedge e^2) \tag{A15}$$

A3. Spinors in CP_2. As Hawking has shown (Hawking and Pope, 1978), CP_2 does not allow spinor structure in the conventional sense. However, the coupling of the spinors to a half-odd multiple of the Kähler connection B leads to a respectable spinor structure. Because the intricacies associated with the spinor structure of CP_2 play a fundamental role in the construction of the proposed theory we repeat the arguments of Hawking here.

To see how the space can fail to have an ordinary spinor structure consider the parallel transport of the vierbein in a simply connected space M. The parallel propagation around a closed curve γ with a base point x leads to a rotated vierbein at x: $\bar{e}^A = R^A{}_B e^B$ and one can associate to each closed path an element of $SO(4)$. Consider now a one-parameter family of closed curves $\gamma(v)$, $v \in [0, 1]$ with the same base point x and $\gamma(0)$ and $\gamma(1)$ trivial paths. Clearly these paths define a sphere in M and the elements $R^A{}_B(v)$ define a closed path in $SO(4)$. When S^2 is contractible to a point, e.g., homologically trivial the path in $SO(4)$ is also contractible to a point and therefore represents a trivial element of the homotopy group $\Pi_1(SO(4)) = Z_2$. However, for a homologically nontrivial 2-surface S^2 the associated path in $SO(4)$ can be homotopically nontrivial and therefore corresponds to a nonclosed path in the covering group Spin(4) (leading from the matrix 1 to the matrix -1 in the matrix representation). Assume that this is the case.

Assume now that the space allows spinor structure. Then we can parallelly propagate also spinors and by the above construction associate a closed path of Spin(4) to the surface S^2. Now, however, this path corresponds to a lift of the corresponding $SO(4)$ path and therefore cannot be closed. Thus we have a contradiction.

From the preceding argument it is clear that one could compensate the nonallowed -1 factor associated with the parallel transport of the spinor around the sphere S^2 by coupling it to a gauge potential in such a way that in the parallel transport the gauge potential introduces a compensating -1 factor. For a $U(1)$ gauge potential this factor is given by the exponential $\exp(i2\Phi)$, where the quantity Φ is the magnetic flux through the surface. This factor has the value -1 provided the $U(1)$ potential carries a half unit of Dirac charge $1/2g$ or, more generally, half-odd multiple of it. In the case of CP_2 the required gauge potential is an odd multiple of the quantity $B/2$ defined previously. In the case of $M^4_{(+)} \times CP_2$ we can in addition couple the spinor components with different H chiralities independently to the odd multiple of $B/2$.

A4. Geodesic Submanifolds of CP_2. By restricting the coordinate variables of CP_2 to a geodesic submanifold (defined as a submanifold for which the geodesic lines are also geodesic lines of imbedding space) S of CP_2, one obtains a "subtheory" with $H_{\text{eff}} = V^4 \times CP_2$. So one can ask, What kind of geodesic submanifolds CP_2 has besides the geodesic lines? One can answer this question by observing that CP_2 is a symmetric space (Helgason, 1978), e.g., representable as a coset space $SU(3)/SU(2)$.

In the book by Helgason a general characterization of the geodesic submanifolds of an arbitrary symmetric space G/H is given. The geodesic submanifolds are in 1:1 correspondence with the so-called Lie triple systems of the Lie algebra g of the group G. The Lie triple system t is defined as a subspace of g characterized by the closedness property with respect to double commutation:

$$[X,[Y,Z]] \in t \quad \text{for } X, Y, Z \in t \tag{A16}$$

It is rather straightforward to verify that $SU(3)$ allows, besides the geodesic lines, two nonequivalent (not isometry related) two-dimensional geodesic submanifolds, which have the topology of 2-sphere. This can be understood by observing that $SU(3)$ allows two nonequivalent $SU(2)$ subalgebras. The first subalgebra corresponds to the group $SO(3)$ of the 3×3 orthogonal matrices and the second algebra integrates to the usual isospin group $SU(2)$. The nonequivalence of the algebras is obvious from the fact that they integrate to nonisomorphic groups. By taking a subset of any two generators from these algebras one obtains the two nonequivalent Lie triple systems.

Convenient representatives for the geodesic spheres of CP_2 are given by the equations

$$\text{I:} \quad \xi^1 = \xi^2 \text{ or equivalently } (\theta = \pi/2, \phi = 0)$$

$$\text{II:} \quad \xi^1 = \bar{\xi}^2 \text{ or equivalently } (\theta = \pi/2, \psi = 0)$$

The nonequivalence of these submanifolds is clear from the fact that isometries act as holomorphic transformations in CP_2. The vanishing of the second fundamental form, which is equivalent to the geodesic submanifold property, is easy to verify by a direct calculation.

APPENDIX B: THE REPRESENTATION OF ISOMETRIES OF *H*

The aim of this Appendix is to show that the isometries of CP_2 are representable as symmetries of the action defining the theory. The problem is essentially that of finding the action of the isometries on the spinors because the bosonic part of the action is invariant under isometries (Kähler form is covariantly constant and thus invariant under isometries).

It is instructive to prove the infinitesimal representability of the isometries first in the general case, when H allows ordinary spinor structure. So, let $h^k \to h^k + \varepsilon j^k$ be the infinitesimal isometry so that the infinitesimal generator satisfies the well-known Killing identities (Misner et al., 1973; Adler et al., 1975)

$$D_l j_k + D_k j_l = 0 \tag{B1}$$

The following theorem holds true:

Theorem. The quantity $L = \overline{\Psi}\Gamma^\alpha D_\alpha \Psi$ is invariant under the infinitesimal isometries of H realized according to

$$\delta\Psi = i\varepsilon\left(\partial_l j_k \Sigma^{kl}/2 + j^k V_k\right)\Psi \equiv i\varepsilon X\Psi \tag{B2}$$

Before going to the proof observe that the transformation formula has a simple interpretation. The isometry is interpreted as a flow in H and the spinor field is translated along the flow lines by parallel translation: besides the usual rotation, the spinor field suffers a gauge transformation given by the so-called nonintegrable phase factor (Wu and Yang, 1976).

Proof. The metric of X^4 is invariant under the action of isometries and therefore it suffices to consider the term $L_{\alpha\beta} = \overline{\Psi}\Gamma_\alpha D_\beta \Psi$. The change of this quantity can be written as

$$\delta L_{\alpha\beta} = \overline{\Psi}\left(K_\alpha D_\beta + \Gamma_\alpha L_\beta\right)\Psi \tag{B3}$$

where the quantities K_α and L_α are defined as

$$K_\alpha = \delta\Gamma_\alpha + \varepsilon[\Gamma_\alpha, X] \tag{B4}$$

$$L_\alpha = \delta V_\alpha + \varepsilon\partial_\alpha X + \varepsilon[V, X] \tag{B5}$$

Obviously the requirements $K = 0$ and $L = 0$ guarantee the invariance of the action, the latter requirement implying that the isometries act as gauge transformations. These conditions can be transformed into a simpler form by using the definitions of the connection and induced gamma matrices:

$$\partial_r\Gamma_k j^r + [\Gamma_k, X] + \Gamma_r\partial_k j^r = 0 \tag{B6}$$

$$\partial_r V_k j^r + V_r\partial_k j^r + \partial_k X + [V_k, X] = 0 \tag{B7}$$

The first condition is found to be true by using the covariant constancy of the gamma matrices and Killing identities. The derivatives of j^k contained in the second equation can be eliminated by using the equation defining the curvature tensor

$$D_m D_n j_k - D_n D_m j_k = R^s{}_{kmn} j_s \tag{B8}$$

and Killing identities. Using the representation of the vielbein curvature in terms of the curvature tensor the second equation can be cast into the form

$$j^r \sum_c R_{rkmn} = 0 \tag{B9}$$

where the sum is over the cyclic permutations of the indices k, m, n. Because of the so called cyclic identities satisfied by the curvature tensor this sum however vanishes so that also the equation (B.7) is true. ■

Next we turn to the case of $M^4 \times CP_2$. Clearly the presence of the Kähler connection introduces an additional term to δL:

$$\delta L^{\mathrm{add}} = i\varepsilon\bar{\Psi}\Gamma^\alpha Y_k \partial_\alpha s^k (1_+ + 31_-)\Psi + \mathrm{g.c.} \tag{B10}$$

where the quantity Y_k is defined as

$$Y_k = \partial_r B_k j^r + (B_r/2)\partial_k j^r \tag{B11}$$

So, provided the quantity Y_k is expressible as a gradient of some quantity, say, Z, we can indeed compensate δL^{add} by an infinitesimal gauge transformation performed for the spinor field

$$\delta\Psi^{\mathrm{add}} = -i\varepsilon Z(1_+ + 31_-)\Psi \tag{B12}$$

The representation of Y^k as a gradient in turn follows from the fact that the isometries of CP_2 can be regarded as Hamiltonian flows:

> *Theorem.* The infinitesimal isometries of CP_2 can be regarded as infinitesimal Hamiltonian flows with respect to the symplectic structure defined by the Kähler form J, e.g., for an infinitesimal isometry j^k there exists a Hamiltonian H so that
>
> $$j^k = J^{kl}\partial_l H \tag{B13}$$
>
> and as a consequence the quantity is expressible as a gradient of the quantity
>
> $$Z = (j^r B_r/2 + H) \tag{B14}$$

Proof. The existence of H satisfying (B13) follows from the fact that J is invariant under isometries and closed as a 2-form. To see this observe that (B13) is equivalent to the integrability condition

$$\partial_r(J_{mn}j^n) - \partial_m(J_{rn}j^n) = 0 \tag{B15}$$

Using the closedness property of J,

$$\partial_m J_{rs} + \partial_r J_{sm} + \partial_s J_{mr} = 0 \tag{B16}$$

one can transform the equation

$$\partial_m J_{rs} j^m + J_{ms}\partial_r j^m + J_{rm}\partial_s j^m \tag{B17}$$

expressing the infinitesimal invariance of J_{mn} under isometries to the equation (B15) and thus the existence of H is shown. The representation of the quantity Y^k as gradient is obtained by a direct calculation expressing j^k in terms of H. ■

Summarizing, a spinor field transforms under the infinitesimal isometries according to the formula

$$\delta\Psi = i\varepsilon[X - Z(1_+ + 31_-)]\Psi \equiv J\Psi \tag{B18}$$

Since the isometries are represented as gauge transformations infinitesimally, the representation cannot be integrable in the ordinary sense. This can be seen by calculating the commutator of two infinitesimal isometries j_i^k, $i=1,2$, having the representation matrices J_i:

$$[J_1, J_2] = J_{[1,2]} + j_1^k j_2^l F_{kl} \tag{B19}$$

The additional term, where F_{kl} denotes the curvature form of the spinorial connection, implies that the representation is not integrable in the ordinary sense. One can speak of a group representation modulo gauge transformations, which can be thought of as a generalization of the ordinary projective representations (Varadarjan, 1970).

APPENDIX C: CLASSICAL ASPECTS OF THE THEORY I

This Appendix is devoted to the classical aspects of the theory. Classical equations of motion are given and some families of particlelike solutions to the equations of motion will be obtained. The solutions should provide semiclassical models for hadrons and various massless particles.

C1. Classical Equations of Motion. The classical equations of motion are obtained as extremum conditions for the action defining the theory. The equations for the coordinate variables h^k are

$$D_\alpha\{[T^{\alpha\beta} - (1/8\pi G)G^{\alpha\beta}]h^k{}_\beta\} - \mathrm{Tr}(j^\alpha F^k{}_l h^l{}_\alpha) = 0 \tag{C1}$$

where the covariant energy momentum tensor and the gauge current are defined as

$$T^{\alpha\beta} = -(1/g^2)\mathrm{Tr}\left[F^\alpha{}_\gamma F^{\gamma\beta} + (1/4)g^{\alpha\beta}F^{\mu\nu}F_{\mu\nu}\right] \equiv \bar{T}^{\alpha\beta}/g^2 \tag{C2}$$

$$j^\alpha = (1/g^2)D_\beta F^{\alpha\beta} \tag{C3}$$

It should be noticed that the gauge field $F_{\alpha\beta}$ is the projection of the curvature form of the connection of H and therefore the vierbein part of F is the projection of the curvature tensor of H.

These equations can be regarded as a generalization of the equation defining Lorentz force in ordinary electrodynamics. In fact, by taking projections to the directions of the tangent vectors of X^4 one obtains

$$D_\beta T^{\alpha\beta} = \mathrm{Tr}(j_\beta F^{\alpha\beta}) \tag{C4}$$

which is simply the definition of the Lorentz force and holds identically true reflecting the fact that the freedom to choose the coordinates of X^4 arbitrarily reduces the number of dynamical degrees of freedom from $n+4$ to n ($=\dim S$).

The equations of motion for the spinor field, which for Grassmann-valued spinors have sense only as functional expectation values [e.g., in the sense that the quantities $\int \exp(iS)\,\delta S/\delta\Phi^k\, D\Phi$ vanish since the integral is functional gradient] are given by

$$\Gamma^\alpha D_\alpha \Psi = -i\frac{H}{2}\Psi \tag{C5}$$

where we have defined

$$H = g_3^{\alpha\beta} H^k_{\alpha\beta}\Gamma_k = H^k\Gamma_k \tag{C6}$$

Here the quantities $H^k_{\alpha\beta}$ define the so-called second fundamental form (4). Equation (C5) can be written also in the form resembling massless Dirac equation

$$(\Gamma^\alpha D_\alpha + D^\alpha\Gamma_\alpha)\Psi = 0 \tag{C7}$$

The equations resulting from the variation of the coordinate variables h^k on the boundaries of X^4 are given by a rather awkward expression:

$$\begin{aligned}&\left[(T^{n\beta} - kG^{n\beta})h^k{}_\beta - \mathrm{Tr}\left(F^{n\beta}F^k{}_l h^1{}_\beta\right)\right](-g_4)^{1/2}\\ &\quad = \left[D_\alpha\left(T_1^{\alpha k}\right) + \bar\Psi\Gamma^\alpha F^k{}_l h^l{}_\alpha\Psi\right](-g_3)^{1/2} + \left(D_\alpha^\beta\right)^k{}_l\left[T_2^{\alpha k}(-g_4)^{1/2}\right]\end{aligned} \tag{C8}$$

The quantities $T_1^{\alpha k}$ and $T_2^{\alpha k} = T_2^{\alpha\beta}h^k_\beta$ can be thought as generalizations of energy momentum densities associated with the fermionic and bosonic degrees of freedom, respectively:

$$\begin{aligned}T_1^{\alpha k} &= \left[\bar\Psi(\Gamma^\alpha\vec{D}^\beta - \overleftarrow{D}^\alpha\Gamma^\beta)\Psi + (\alpha\leftarrow\ \rightarrow\beta)\right]h^k{}_\beta\\ &\quad + \bar\Psi(\Gamma^k\vec{D}^\alpha - \overleftarrow{D}^\alpha\Gamma^k)\Psi\end{aligned} \tag{C9}$$

$$\begin{aligned}T_2^{\alpha\beta} &= \left[g_4^{\alpha\beta}g_4^{\mu\nu} - g_4^{\alpha\mu}g_4^{\beta\nu} - \left(g_4^{\alpha\beta} - g_3^{\alpha\beta}\right)g_4^{\mu\nu} - \left(g_4^{\mu\nu} - g_3^{\mu\nu}\right)g_4^{\alpha\beta}\right]\left\{ {n \atop \mu\nu} \right\}\\ &\quad + \frac{1}{2}\left[g_3^{\alpha\mu}\left(g_4^{nn}g_{3\nu}^{\beta}\left\{ {\nu \atop \mu n} \right\} + g_4^{n\beta}g_{3\nu}^{\delta}\left\{ {\nu \atop \mu\delta} \right\}\right)\right] + (\alpha\leftarrow\ \rightarrow\beta)\end{aligned} \tag{C10}$$

The covariant derivative appearing in (C8) is defined as

$$\left(D_\alpha^B\right)^k{}_l=\delta^k{}_l\partial_\alpha+\left\{ {k \atop lm} \right\} h^m{}_\alpha \tag{C11}$$

The equations have clearly the nature of boundary conditions guaranteeing the conservation of various isometry charges.

C2. Explicit Form of the Action. In sequel the explicit form of the action density will be useful. The YM part of the action density can be written in the form

$$L_{\mathrm{YM}}=-\left(8/g^2\right)\left[18\left(e^0\wedge e^3\right)^2+18\left(e^1\wedge e^2\right)^2+32e^0\wedge e^3\cdot e^1\wedge e^2 + \left(e^0\wedge e^1-e^2\wedge e^3\right)^2+\left(e^0\wedge e^2-e^3\wedge e^1\right)^2\right] \tag{C12}$$

where we have used the same notation for the projections of the vierbein components as for vierbeins themselves.

The curvature scalar can be written in the form $R=R_1+R_2$, where the part R_1, which is present only when H is curved, can be written in the case of $H=V^4\times CP_2$ as

$$R_1=-4R^2\left[4\left(e^0\wedge e^3\right)^2+4\left(e^1\wedge e^2\right)^2+\left(e^0\wedge e^1\right)^2+\left(e^0\wedge e^2\right)^2 - \left(e^2\wedge e^3\right)^2-\left(e^1\wedge e^3\right)^2-2e^0\wedge e^1\cdot e^2\wedge e^3+2e^0\wedge e^2\cdot e^1\wedge e^3 + 4e^0\wedge e^3\cdot e^1\wedge e^2\right] \tag{C13}$$

The part R_2 can be expressed using the second fundamental form

$$R_2=h_{kl}\left(H^{k\alpha\beta}H^l_{\alpha\beta}-H^{k\alpha}_{\alpha}H^{l\beta}_{\beta}\right) \tag{C14}$$

Of particular interest will be the two nonequivalent Abelian subtheories obtained by restricting CP_2 variables either to the geodesic sphere S^{I}: $(\theta=\pi/2,\ \phi=0)$ or S^{II}: $(\theta=\pi/2,\ \Psi=0)$. It is advantageous to use ordinary spherical coordinates (Θ,Φ) instead of the coordinate variables (r,ψ) and (r,ϕ) in the cases I and II, respectively. The relation between the two coordinate sets is given by the equations

$$\mathrm{I:}\quad \cos\Theta=2/\left(1+g^2\right)-1,\quad \Phi=\psi/2,\qquad \mathrm{II:}\quad \cos\Theta=1/\left(1+g^2\right)^{1/2},\quad \Phi=\phi \tag{C15}$$

In these coodinates the line element has the form

$$ds^2 = \frac{R^2}{4}(d\Theta^2 + \sin^2\Theta\, d\Phi^2) \tag{C16}$$

Both geodesic spheres have same radius, as is clear from the fact that CP_2 allows geodesics of only one type.

The nonvanishing components of curvature have the expressions

$$R_{03} = 2R_{12} = 4e^0 \wedge e^3 = -du \wedge d\Phi \tag{C17}$$

and

$$R_{02} = -R_{31} = e^0 \wedge e^2 = du \wedge d\Phi/2 \tag{C18}$$

in the cases I and II, respectively. Here we have used the notation $u = \cos\Theta$. The YM part of the action can be written in the form

$$L^i_{\mathrm{YM}} = -(1/4)(F^i)^2 \tag{C19}$$

where we have defined the gauge field as

$$F^i = n_i\, du \wedge d\Phi \tag{C20}$$

Here one has $n_{\mathrm{I}}^2 = 36/g^2$ and $n_{\mathrm{II}}^2 = 8/g^2$ corresponding to the two cases $i = \mathrm{I}$ and II, respectively.

The part R_1 of the curvature scalar can be written as

$$R_1 = R^2(du \wedge d\Phi)^2 \tag{C21}$$

and has same form in cases I and II.

C3. Particlelike Solutions with Vanishing Gauge Fields. The action defining the theory is quadratic in spinor and gauge fields so that the surfaces having $F_{\alpha\beta} = 0$ are vacuum solutions for the theory defined by a pure YM action. Geometrically the condition means that the projection of the 4-surface to the space S is at most one dimensional, as is seen by choosing the coordinates so that X^4 has the representation $s^k = \mathrm{const}$ for $k \neq k_0$.

For the theory defined by the action $S = S_{\mathrm{YM}} + S_{\mathrm{gr}}$ the interior equations of motion reduce to

$$G^{\alpha\beta} H^k_{\alpha\beta} = 0 \tag{C22}$$

so that the vacuum degeneracy is removed to a great extent. It is easy to see that surfaces of the form $M^1 \times X^2 \times S^1$ ("membranes"), where M^1 is a timelike geodesic of M^4, X^2 an arbitrary surface in the orthogonal complement of M^1, and S^1 a geodesic in CP_2, satisfies the equations of motion. The equations of motion reduce to geodesic equations for M^1 and S^1 because the Einstein tensor for the product manifold $X^2 \times Y^2$ is given by the expression

$$G^{\alpha\beta} = -\left(g_1^{\alpha\beta} R_2 + g_2^{\alpha\beta} R_1\right)/2 \tag{C23}$$

where the indices 1 and 2 refer to the manifolds X^2 and Y^2, respectively (Einstein tensor of a 2-manifold vanishes identically).

The mass of the solution is quantized being proportional to the Euler characteristic of the manifold X^2 (Eguchi et al., 1980, p. 344)

$$M = (1/2G) L(g + n/2 - 1) \tag{C24}$$

Here g and n denote the genus and the number of holes in X^2 and L denotes the length of the CP_2 geodesic which is equal to πR. Since R is of the order of Planck length, the mass is enormous except for the solutions having either the topology of the torus of a 2-sphere with two holds. A rather peculiar feature of the solution is that the rest energy is negative for the standard choice of the time orientation when $g = 0$ and n is smaller than 2.

C4. Stringlike Solutions. Stringlike objects are obtained as solutions of type $X^4 = X^2 \times Y^2 \subset A \times B$, where A is a time linear submanifold of M^4 and B is the orthogonal complement of A in H. The separability requirement for the equations of motion implies that either (a) Y^2 is a geodesic sphere S^2 in CP_2 and X^2 is a minimal surface in M^4, e.g., the trace of the second fundamental form vanishes:

$$M^k = 0 \tag{C25}$$

[for the definition of the second fundamental form see (4)] or (b) that X^2 corresponds to a static string in M^4, e.g., is a timelike plane strip and that Y^2 minimizes its energy. Observe that the Einstein tensor has nonvanishing components only for X^2 indices.

The equations of motion for X^2 are exactly the same as those satisfied by the orbit of the string in the string model. The 4-momentum of the solution can be expressed in the same form as in string model. In case (a) one obtains

$$P^k = k_i \int g^{0\beta} m^k{}_\beta (-g_2)^{1/2} \, dx \tag{C26}$$

The mass is proportional to the length of the string, the proportionality factor k_i, $i = \mathrm{I}, \mathrm{II}$, is different for the two nonequivalent geodesic spheres and is given by the expression

$$k_i = 8\pi\left(l_i - g^2R^2/16\pi G\right)/g^2R^2 \tag{C27}$$

The parameter l_i obtains the values $l_{\mathrm{I}} = 9$ and $l_{\mathrm{II}} = 2$ in the cases I and II, respectively. The inverse of the quantity k_i can be identified as Regge slope $\alpha_R \sim 1\ \mathrm{GeV}^{-2}$ provided the spinorial contributions to the value of the Regge slope can be neglected. As a consequence the constants G and $g^2R^2/16\pi Gl_i \equiv G_i$ differ by an extremely small amount: $(G_i - G)/G \sim 10^{-38}$! Observe that without the curvature term in action one would obtain for the constant g^2R^2 a value, which is 10^{38} times larger than G.

Which of the two alternatives then corresponds to the hadrons? The requirement that the energy should be positive for both kinds of strings favors the alternative II. The results found studying graphlike solutions to the equations of motion (Appendix D) also favor this choice.

For the solutions of type (b) the inverse of the Regge slope is given by the expression

$$k = \rho_M + (1/2G)(g-1) \tag{C28}$$

where the quantity ρ_M is the magnetostatic energy per unit length of the string. Hence the Regge slope is expected to be of the order of G and the corresponding objects should be extremely pointlike and probably also very massive (the same is expected to hold also for the strings of type I).

Clearly, by adding spinors on the boundaries of the open string of type II, one obtains a model for pion (or ρ-meson). How then to construct other mesons? A little thought shows that one can build, not only mesons, but also baryons from the pionic string. The point is that the genus of the boundary component Y^2 can be increased by simply drilling "wormholds" starting from and ending at the boundary component in question [visualization: $\delta(\text{apple}) = S^2$ and δ (apple with wormhole $= S^1 \times S^1$]. Obviously one can build all mesons using this procedure. Now, since baryons behave like strings in many respects (linear Regge trajectories), one might argue that baryons in fact correspond classically to 3-manifolds obtained by drilling a hole in a pionic string and by adding spinors on the resulting three boundary components. Of course, one can add an arbitrary number of sea quarks to the hadron in this way and perhaps also gluons. So, one can say that the hadronic interior solution is in definite sense universal (and probably so also the leptonic one).

Clearly, one could also start from closed string solutions and apply to them a similar procedure. The resulting "hadrons" would differ from the open string hadrons in that they would have only homologically trivial boundary components. Since the homology charge plays the role of coupling constant in the interactions described by dual diagrams in the sense that only the ends of the open string, having $|h| = 1$, are involved in these diagrams, one expects that this second type of hadron and ordinary hadrons have rather different strong interaction properties. It is perhaps worth noticing that the magnetic flux through the homologically charged boundary components is nontrivial, indeed the magnetostatic energy of the string has an interpretation as the interaction energy of two magnetic monopoles (in case II the gauge field is nondiagonal).

C5. Small Perturbations around the Static Hadronic String. The equations governing the small perturbations for the stringlike solutions are easily obtained from the equations of motion (C1) using the fact that the second fundamental form vanishes for the static string solutions. One obtains for the purely transversal perturbation, e.g., for the perturbations with $\delta s^k = 0$, the equations

$$\Box_X 2\delta m_T^k = \kappa_i \Box_S 2\delta m_T^k \tag{C29}$$

Here the symbol $\Box$ denotes the d'Alambert operator for the space specified by the lower index and the index T expresses that the perturbation is perpendicular to the string. The parameter κ_i is given by the expression

$$\kappa_i = l_i / \left(l_i - g^2 R^2 / 16\pi G \right) \tag{C30}$$

($l_{\rm I} = 9$ and $l_{\rm II} = 2$).

Now, what do we mean with small perturbations in our topological context? In the conventional approach one would of course restrict the perturbations to be small and certainly nonsingular everywhere. Now the situation is however different since one can always cut away the possible singularities and the resulting boundary component has a natural interpretation as a particlelike excitation. In the following we shall apply both of these approaches.

The everywhere regular solutions to (C30) can be written as a superposition of the products of plane waves and spherical harmonics:

$$f^k_{plm} = \varepsilon^k e^{ip\cdot m} Y^L{}_m \tag{C31}$$

where ε^k is a polarization vector orthogonal to the string orbit. The

2-momentum p^k satisfies the dispersion relation

$$p^2 = -\kappa_i L(L+1)/R_i^2 \tag{C32}$$

Obviously, the wave vector p is either spacelike or has an imaginary time component. In the latter case the perturbation has an exponential time dependence $\exp(\omega m^0)$ ($\omega \sim R_h/G$, where R_h is a typical hadronic length!) and as a consequene, the string can be regarded unstable classically. The result is indeed puzzling. Of course, one could get rid of it by changing the overall sign in the definition of energy so that the quantity κ_i would change sign, making the modes with $L > 0$ oscillatory. One could, however, argue that the conventional stability criteria should not be taken too seriously because now also the topologically trivial perturbations are allowed, and therefore the original definition of energy should not be given up.

In the more general approach we give up the regularity requirement altogether, writing the perturbation as a superposition of the separable solutions:

$$f^k(p, n, k) = e^{ip\cdot m} z^n f_k^n(|z^2|) \tag{C33}$$

Here z denotes the complex coordinate for S^2 and the function f satisfies the differential equation

$$f'' + \frac{(n+1)}{u} f' - \frac{kf}{u} = 0 \tag{C34}$$

Moreover, the dispersion relation

$$p^2 = -\kappa_i k \tag{C35}$$

is satisfied.

An important special case of (C33) is obtained, when the parameter k vanishes: $k = 0$. These kinds of solutions give rise to perturbations, which can be written in the form

$$m^k = f_+^k(z, m_+) + f_-^k(z, m_-) \tag{C36}$$

where $f_{+(-)}^k$ is an analytic function of the variable z and the dependence on the light cone coordinate $m_{+(-)} = m^0{}_{(-)}^{+} m^3$ is arbitrary. Clearly, these solutions represent pulses propagating along the string with the velocity of light. The analyticity of $f_{+(-)}^k$ implies the presence of singularities, e.g., of cuts or poles unless the dependence on z is trivial. Polelike singularity can be eliminated by cutting a hole around the pole, and as a consequence one has a hole in a hadron. Of course, the mere presence of a cut introduces a hole also. Since the parameters describing the position of the discontinuity or pole in general depend on the variable $m_{+(-)}$, the boundary component associated with the eliminated singularity is expected to move with a velocity of light or to be stationary (provided that the dependence on $m_{+(-)}$ is trivial).

Concerning the interpretation of these singularities one should take into account that the "quanta" in question are associated with the transversal excitations of the string described by a vector quantity with two physical components and they satisfy a massless wave equation. Therefore, the interpretation as the classical counterparts of gluons is suggestive for the nonstationary boundary components. The stationary boundary components might correspond to sea quarks.

As regards the nature of the more general "massive modes" it should be noticed that one obtains both oscillatory and tachyonic modes depending on the sign of the parameter k. Also it is to be expected that the elimination of singularities also now leads to particlelike excitations, but now they should behave like massive particles. Perhaps the interpretation as off mass shell gluons might be appropriate for the associated bosonic quanta. It is amusing to notice that the solutions of (C34) correspond to radially symmetric, zero-energy solutions of an $(n+2)$-dimensional Schrödinger equation in Coulombic potential k/u.

C6. Massless Particlelike Solutions. A rather general set of solutions is obtained by assuming first that X^4 is a submanifold of $M^4 \times S^2$, where S^2 is a geodesic sphere, and second that X^4 is given either by the equations

$$F(k \cdot m, \Theta) = 0$$

$$G(e \cdot m, \Phi) = 0 \tag{C37}$$

or by the equations

$$F(k \cdot m, \Phi) = 0$$

$$G(\rho, \Phi) = 0 \tag{C38}$$

Here k is a lightlike vector, e a spacelike polarization vector orthogonal to k, the variable ρ is a radial variable in the plane orthogonal to the wave vector associated with k: $\rho = (m_1{}^2 + m_2{}^2)^{1/2}$, and (Θ, Φ) denotes the coordinates of S^2.

The equations of motion are satisfied because the quantities $G^{\alpha\beta}$, $T^{\alpha\beta}$, and j^{α} appearing in them are proportional to the quantities $k^{\alpha}k^{\beta}$ and k^{α}, respectively (k can be regarded as a vector of X^4 also when Minkowsky coordinates are used for X^4) and because the field equations involve only contractions of type $k \cdot e$, $k \cdot k$, and $k \cdot e_{\rho}$ (e_{ρ} unit vector in the direction of ρ), which all vanish.

Concerning the interpretation of the solutions of the first type, which clearly represent massless particles, some remarks should be made. First, the solution manifold has necessarily a boundary since then it must have a finite extension in direction orthogonal to the polarization and wave vectors. Because the vector e defines a natural polarization vector, the interpretation as a classical counterpart of gauge boson or neutrino is suggestive (provided the number of boundary components is equal to 1). Second, since CP_2 has two nonequivalent geodesic submanifolds one obtains two solution types: the standard choices for the geodesic sphere S^2 give rise to purely diagonal and purely nondiagonal gauge fields corresponding to the cases I and II, respectively. Thirdly, there is *a priori* no restriction on the genus of the boundary component. The bosons with $g > 0$ induce generation changing transitions and they should be massive in the quantized theory. This is indeed expected since the transitions changing the boundary topology introduce mixing between different boson generations and therefore the mass matrix for these states has nonvanishing nondiagonal components. Diagonalization should provide masses for all the gauge bosons except the photon. Finally, also solutions with more than one boundary component are possible. The question whether they correspond to anything stable enough to be observable, is interesting.

The second solution type has properties which make it a good candidate for classical graviton. First, the solution can be closed and, in particular, can have S^3 topology. Second, there is no vector polarization associated with the solution because of the rotational symmetry around the direction of motion. However, the deviation of the metric from flat metric defines a tensor polarization and so the solution has the nature of spin-2 particle.

APPENDIX D: GRAPHLIKE SOLUTIONS

This Appendix is devoted to the study of the solutions representable as graphs of some map f: $M^4 \to CP_2$. We shall first consider vacuum solutions and then shall concentrate on the Abelian subtheories.

D1. Vacuum Solutions. The theory allows a great number of vacuum solutions besides the trivial ones (e.g., regions of M^4). One interesting type of vacuum solution is obtained using the ansatz

$$s^k = s^k(m^{k_0}) \tag{D1}$$

where m^{k_0} is the preferred Cartesian coordinate of M^4. The requirement

$$s_{kl}\partial_0 s^k \partial_0 s^l = \text{const} \tag{D2}$$

guarantees the flatness of the induced metric. A rather general solution satisfying (D2) is obtained, when s^k satisfies the equations

$$\partial_0{}^2 s^k + \left\{ \begin{matrix} k \\ lm \end{matrix} \right\} \partial_0 s^m \partial_0 s^n = F^k{}_l \partial_0 s^1 \tag{D3}$$

Here $F^k{}_1$ is an antisymmetric tensor in CP_2. Clearly (D3) corresponds mathematically to a motion of a charged particle in the electromagnetic field.

These vacuum solutions are as such rather uninteresting. The small perturbations around these solutions have some rather amusing properties. One finds easily that small perturbations satisfy the equation

$$\Box_3 \delta s^k = 0 \tag{D4}$$

where $\Box_3$ is d'Alembert operator in the orthogonal complement of the linear subspace associated with the coordinate variable m^{k_0}. When m^{k_0} is spacelike these perturbations clearly represent a massless gauge field polarized in the direction of m^{k_0}. When m^{k_0} is timelike the perturbation represents scalar potential ("timelike polarization"). A rather interesting feature of the perturbation is that the dependence on the coordinate variable is arbitrary. As a consequence the propagator associated with these perturbations is effectively that in a three-dimensional space. A first consequence is that timelike perturbations do not propagate at all in the ordinary sense. A second consequence is that the perturbation theoretic approach to the calculation of the transition amplitudes in the background defined by this kind of a vacuum solution, should be free of divergence difficulties. This because the loop integrals in three dimensions need no infinite subtractions, when dimensional regularization is used ('t Hooft and Veltman, 1973). Of course, the same result should hold also for the spinorial functional integrals, because spinors are restricted to three-dimensional boundaries of the background X^4. A second interesting feature of these solutions is that they in general are CP noninvariant (C performs a complex conjugation in CP_2).

D2. Approximate Solutions to the Field Equations. In this section we study the approximate field equations obtained by expanding the exact field equations to the lowest order in the parameter R^2. In this approximation the field equations read as

$$\mathrm{Tr}\left(j^{\alpha}F^{k}{}_{l}s^{1}{}_{\alpha}\right)=0 \tag{D5}$$

As expected these equations follow from the M^4 gauge theory action with the constraints expressing the gauge field in terms of the coordinate variables of CP_2.

An interesting property of (D5) is that, when the CP_2 coordinates are restricted to a geodesic sphere, the field equations have an infinite parameter group of symmetries: The gauge field F_i (see Section C2) can be thought as a 2-form induced from the area form of $S^2{}_i$ and thus is invariant under the area-preserving transformation of a 2-sphere, e.g., the canonical transformations.

It is rather easy to find solutions to the equations (D5). One class of solutions is obtained via the ansatz

$$u=\cos\Theta=h(m^i)$$

$$\Phi=\omega m^0+f(u) \tag{D6}$$

The field equations reduce to the condition

$$\Box_3 h=0 \tag{D7}$$

e.g., the quantity h is proportional to a potential associated with a sourceless $U(1)$ gauge field. The function f is arbitrary in this approximation (canonical invariance).

The following features of these solutions appear to be rather general:

(a) The compactness of CP_2 introduces regions where the solution is undefined. For instance, for $h=c/r$ (Coulombic potential) the cutoff radius is $r=C$.

(b) The massive solutions are necessarily gauge charged (electrically neutral solutions are of course obtained in the S^2_{II} case). To see this, observe that the g_{00} component of the metric is equal to

$$g_{00}=1-R^2\sin^2\Theta\omega^2 \tag{D8}$$

Therefore one must have

$$u=a+b/r+O\left(1/r^2\right) \tag{D9}$$

for massive solutions. This in turn implies that the gauge field has $1/r^2$ behavior so that the charge is proportional to the parameter M/ω. The charge can be made arbitrarily small by a suitable choice of the parameters a and b but solutions with exactly vanishing charge are impossible. For the solution type II the nonvanishing gauge charge should lead to no long-range forces because the associated field quanta are expected to be massive.

(c) The gravitational potential $\Phi_{gr} = g_{00} - 1$ satisfies Poisson equation

$$\Box_3 \Phi_{gr} = -4\pi G_i T^{00}_{YM} \tag{D10}$$

Here the gravitational constant G_i is given by the expression

$$G_i = g^2 R^2/16\pi l_i \tag{D11}$$

For the solution type II the constant G_{II} is practically equal to Newton's constant: $(G_{II} - G)/G \sim 10^{-38}$. The identification of the vacuum II as the physical vacuum is favored because the cancellation between the gauge field and gravitational energies is almost complete.

D3. Stationary, Spherically Symmetric Solution Ansatz. A solution ansatz leading to a stationary, spherically symmetric metric and Abelian gauge field is given by

$$\begin{aligned} u &= g(r), \qquad m^0 = \lambda x^0 + h(r) \\ \Phi &= \omega x^0 + f(r), \qquad r_M = r \end{aligned} \tag{D12}$$

[r_M denotes the radial coordinate of M^4: $r_M^2 = \Sigma_{i=1}^3 (m^i)^2$]. The metric and gauge field associated with the solution ansatz are

$$\begin{aligned} g_{00} &= \lambda^2 - R^2 \sin^2\Theta\, \omega^2 \\ g_{0r} &= h' - R^2 \sin^2\Theta\, \omega f' \\ g_{rr} &= -1 - R^2 \sin^2\Theta (f')^2 + (h')^2 - (g')^2 \end{aligned} \tag{D13}$$

and

$$F^i_{0r} = n_i \omega u'/2 \tag{D14}$$

respectively. The condition $g_{0r} = 0$ fixes the function h apart from an additive constant

$$h' = \sin^2\Theta\, \omega f'/4\lambda \tag{D15}$$

The parameter λ is fixed by the boundary condition $g_{00}(\infty)=1$.

$$\lambda^2 = 1 + R^2 \sin^2\Theta_\infty \omega^2 \tag{D16}$$

Next we proceed to study the field equations associated with this ansatz. The isometry $\Phi \rightarrow \Phi + \varepsilon$ gives rise to a conservation law

$$(T^{rr} - \kappa G^{rr}) f' \sin^2\Theta(-g)^{1/2} = C \tag{D17}$$

Only the value $C = 0$ is physically acceptable in (D17) by the requirement of asymptotic flatness and therefore the equation

$$T^{rr} = \kappa G^{rr} \tag{D18}$$

results. As a consequence the metric behaves asymptotically like the Schwartzchild metric (Misner et al., 1973; Adler et al., 1975) since the requirement $G^{rr} = O(1/r^4)$ implies: $g_{00} - 1 = g_{rr} + 1 + O(1/r^2)$ asymptotically. The equation (D18) makes it possible to express the quantity f' in terms of the variable g and its derivatives. By straightforward manipulation one finds for the function f the asymptotic behavior $f \sim r^{1/2}$ and $f \sim r$ in the cases $u(\infty) \neq 1$ and $u(\infty) = 1$, respectively.

Instead of devoting ourselves to the study of the rather complicated equation associated with the coordinate variable u, we shall proceed to look at whether it is possible to imbed the Reissner–Nordström solution (Misner et al., 1973; Adler et al., 1975)

$$g_{00} = -g_{rr}^{-1} = 1 - 2GM/r + \kappa q^2/r$$

$$F_{0r} = q/r^2 \tag{D19}$$

satisfying the Einstein–Maxwell equations

$$T^{\alpha\beta} = \kappa G^{\alpha\beta}$$

$$j^\alpha = 0 \tag{D11}$$

to H for some preferred values of the parameters gR and G. Clearly, this kind of imbedding affords a solution also to the field equations of our theory. The ansatz

$$u = a + b/r \tag{D12}$$

indeed solves (D19) provided one has

$$GM = -abR^2\omega^2$$
$$\kappa q^2/2 = b^2R^2\omega^2$$
$$n_i g\omega/2 = q \tag{D13}$$

The last two equations lead to the condition

$$G = g^2R^2/16\pi Gl_i = G_i \tag{D14}$$

The condition (D14) clearly fails to be satisfied but the failure is extremely small for the solution type II: $(G_{II} - G)/G \sim 10^{-38}$. Therefore we expect that the Reissner–Nordström solution is a good approximate solution to the field equations in this case and that the corrections to Reissner–Nordström solution can be expanded in the powers of the extremely small parameter $(G_{II} - G)/G$.

D4. Lorentz-Invariant Solutions. Perhaps the simplest possible solution ansatz one can imagine is the Lorentz-invariant graphlike solution ansatz having the form

$$m^0 = (1 + r^2)^{1/2} t$$
$$r_M = rt$$
$$s^k = s^k(t) \tag{D15}$$

The induced metric has the Robertson–Walker form (Misner et al., 1973; Adler et al., 1975)

$$ds^2 = \left(1 - s_{kl}\partial_t s^k \partial_t s^l\right) dt^2 - t^2\left(dr^2/(1 + r^2) + r^2 d\Omega^2\right) \tag{D16}$$

The components of the Einstein tensor are given by the expressions

$$G^0{}_0 \equiv \rho = 3\left[(a'/a)^2 - 1/a^2\right]$$
$$G^i{}_j = \delta^i{}_j(-2a''/a - \rho/3) \tag{D17}$$

Here the prime denotes derivative with respect to the proper time $\tau = \int (g_{00})^{1/2} dt$.

The basic property of the solution is that it is necessarily restricted to $M^4 \times S^1$, where S^1 is a geodesic line in S. This can be seen by choosing the coordinates of S so that the solution can be represented in the form

$$s^k = \text{const}, \qquad k \neq k_0$$

The equations of motion for the coordinate variables s^k, $k \neq k_0$, are: $G^{00}s^k{}_{00} = 0$ and imply $s^k{}_{00} = 0$. Therefore a geodesic line is in question. This properly makes it possible to concentrate on a specific solution, say, the one given by equations

$$r = \infty, \qquad \psi = 0$$
$$\theta = \pi/2, \qquad \phi = f(t) \tag{D18}$$

The equation of motion is most easily expressed as a conservation law associated with the infinitesimal isometry $\phi \to \phi + \varepsilon$.

$$G^{00}\partial_0 f(-g)^{1/2} = \text{const} \tag{D19}$$

This equation has the integral

$$\partial_0 f = (1/r)y^{1/2}/(1+y)^{1/2} \tag{D20}$$

Here the quantity y is defined as

$$y = (C/t)^{1/3} \tag{D21}$$

This integral, of course, fixes the metric completely.

The solution satisfies the "equation of state" $\rho = -9p$, where the pressure p is defined as the quantity $p = G^i{}_i/3$. The quantity ρ behaves in the limits $t \to 0$ and $t \to \infty$ as $\rho \sim t^{-5/3}$ and $\rho \sim t^{-7/3}$, respectively. The fact that the Poincaré energy density associated with the solution is negative makes the interpretation as some kind of an idealization for cosmology impossible. The interpretation as a negative energy vacuum solution might be more appropriate.

ACKNOWLEDGMENTS

I thank D. Finkelstein for criticism in the early phase of the work. I gratefully acknowledge J. A. Wheeler for a thorough criticism of the basic ideas in the work. I am deeply grateful to R. Keskinen for encouragement and help. Also I would like to thank M. Lehto for enlightening discussions.

REFERENCES

Abers, E. S. and Lee, B. W. (1973). *Physics Reports*, **9**, 20.

Adler, Basin, and Schiffer, (1975). *Introduction to General Relativity*, McGraw–Hill, New York.

Bailin, D. (1977). *Weak Interactions*, Sussex University Press, Sussex, England.

Berezin, F. A. (1966). *Method of Second Quantization*, Academic Press, New York, Chap. 1.

Bilenki, S. M., and Pontecorvo, P. (1978). *Physics Reports*, **41C**, 225.

Björken and Drell, (1965). *Relativistic Quantum Fields*, McGraw–Hill, New York, p. 108.

Chew, G., and Rosenzweig, C., (1976). *Nuclear Physics B* **104**, 1220.

Close, F. E. (1979). *An Introduction to Quarks and Partons*, Academic Press, New York.

Dolgov, A. D., and Zeldovich, Ya. B., (1981). *Reviews of Modern Physics*, **53**, 1.

Eguchi, T., Gilkey, B., and Hanson, J. (1980). *Physics Reports*, **66**, 6.

Fritzsch, H. and Minkowski, P., (1981). *Physics Reports*, **73**, 1.

Georgi, H., and Glashow, S. (1974). *Physical Review Letters*, **32**, 438.

Gibbons, G. W., and Pope, C. N. (1978). *Communications in Mathematical Physics*, **61**, 239–248.

Hawking, S. W., and Pope, C. N., (1978). *Physics Letters*, **73B1**, 42–44.

Helgason, S. (1978). *Differential Geometry, Lie Groups, and Symmetric Spaces*, Academic Press, New York.

Hilton and Wylie, (1966). *Homology Theory*, Cambridge University Press, Cambridge, England.

Jackson, J. D. (1970). *Classical Electrodynamics*, 2nd ed. J. Wiley & Sons, New York, p. 364.

Jacob, M. (1974). *Dual Theory*, North-Holland Publishing Company, Amsterdam.

Johnson, K. (1975). *Acta Physica Polonica*, **B6**, 865.

Kobayashi, M. and Maskawa, K. (1973). *Progress in Theoretical Physics*, **49**, 652.

Lichnerowicz, A. (1968). In *Battelle Rencontres 1967 Lectures in Mathematics and Physics*, ed. C. M. Dewitt and J. A. Wheeler, W. A. Benjamin, New York.

Lee, T. D. (1979). *Physics Reports*, **96**, 143.

Mahantappa, K. T., and Randa, J. (1980). *Quantum Flavourdynamics, Quantum Chromodynamics and Unified Field Theories*, Plenum Press, New York.

Milnor, J. (1965). *Topology from Differential Point of View*, The University Press of Virginia, Charlottesville, Virginia, Chap. 7.

Misner, Thorne, and Wheeler, (1973). *Gravitation I, II*, W. H. Freeman and Company, San Francisco.

Nambu, Y. (1970). Lecture Notes prepared for the Summer Institute of the Niels Bohr Institute (SINBI).

Papini, C., and Valluri, S. R. (1977). *Physics Reports*, **33C**(2), 60.

Pitkänen, M. (1981). *International Journal of Theoretical Physics*, **20**, 843.

Politzer, H. D. (1974). *Physics Reports*, **14C**.

Reya, E. (1981). *Physics Reports*, **69**, 197.

Salam, A. (1968). In *Proceedings of the 8th Nobel Symposium*, *Stockholm*, N. Swartholm, ed., Almquwist and Wicksells, Stockholm.

Schwinger, J. S. (1957). *Physical Review*, **82**, 664.

Shanahan, P. (1978). *Atyiah-Singer Index Theorem*, Lecture Notes in Mathematics 638, Springer Verlag, New York, Chap. 2.

Symanzik, K. (1969). Varenna Lectures.

Thom, R. (1954). *Commentarii Math. Helvetica*, **28**, 17–86.

’t Hooft, G., and Veltman, M. (1973). Diagrammar, CERN 73-9.

Varadarjan, V. S. (1970). *Geometry of Quantum Theory II*, Van Nostrand Reinhold Company, New York, p. 95.

Wallace, (1968). *Differential Topology*, W. A. Benjamin, New York, Chaps. 6 and 7.
Weinberg, S. (1967). *Physical Review Letters*, **19**, 1264.
Weinberg, S. (1976). *Physical Review Letters*, **31**, 657.
Weinberg, S. (1977). *The First Three Minutes*, Basic, New York.
Veneziano, G. (1974). *Nuclear Physics*, **B74**, 365; (1976). **B117**, 519.
Wu, T. T., and Yang, C. N. (1976). *Physical Review D*, **14**(2), 437.