International Journal of Theoretical Physics, Vol. 24, No. 8, 1985

Topological Geometrodynamics. I. Basic Theoretical Framework

Matti Pitkänen[1]

Received March 11, 1985

Topological geometrodynamics (TGD) is an attempt to a unified description of fundamental interactions based on the assumption that physically allowed space-times are representable as submanifolds of the space, which is a Cartesian product of Minkowski space (or possibly of its light cone) and of some compact space S. This paper is the first one in the series intended for the presentation of TGD. The basic ideas TGD are represented and it is shown that CP_2, the complex projective space of two complex dimensions, is the simplest choice of the space S providing explanation for the known elementary particle quantum numbers provided a topological explanation for the family replication phenomenon, emerging naturally in TGD framework, is accepted.

1. INTRODUCTION

During the past years considerable progress has been made in the understanding of the fundamental interactions. The concept of the broken gauge invariance has shown its power in the description of the electroweak interactions (Weinberg, 1967; Salam, 1968; Glashow, 1961; Abers and Lee, 1973; Bailin, 1977). In the description of strong interactions the QCD approach based on the idea of unbroken gauge invariance has had remarkable successes (Politzer, 1974; Close, 1979; Reya, 1981). It must be admitted, however, that this approach has not yet led to a satisfactory understanding of the confinement problem (Susskind and Kogut, 1976; Brander, 1981): thus the status of QCD as the correct model of the strong interactions is not yet established.

The above-mentioned successes have motivated even more ambitious attempts aiming at the unification of electroweak and strong interaction theories using only the concept of the broken gauge invariance (GUTS;

[1]Neste Oy Research Center, Group of Advanced Information Technology, 06850 Kulloo, Finland.

0020-7748/85/0800-0775$04.50/0

Georgi and Glashow, 1974; Fritsch and Minkoski, 1975; Georgi, 1975). The invention of supersymmetry (Gelfand and Likhtman, 1971; Volkov and Akulov, 1973; Wess and Zumino, 1974) has had a decisive influence on the model building since it makes possible to generalize the concept of the local gauge invariance to that of local supersymmetry and as a consequence bosons and fermions can be handled on the same footing.

Despite the considerable work done in the model building one can fairly say that there exists no compelling unification scenario yet and that the concept of gauge invariance, even in its supersymmetric form, is perhaps not all what is needed to carry out the unification program.

What could this "missing something" be? It is well known that the Yang-Mills gauge potentials can be regarded as components of a connection in an appropriate bundle over space-time (Husemoller, 1960; Madore, 1981) and therefore they correspond to a geometric object. This suggests that it is not the gauge invariance alone but rather the geometric nature of the gauge fields, which underlies the successes of, say the GWS model of the electroweak interactions. This is turn suggests that the program of geometrization of physics, initiated by Einstein, is the correct path to follow, when one tries to solve the unification problem.

Indeed, one can regard the Kaluza-Klein approach (Kaluza, 1921; Klein, 1926; Witten, 1980) as an attempt in this direction. The basic idea of this approach is to replace the four-dimensional space-time with a higher-dimensional space, which is typically a bundlelike structure having "ordinary space-time" as base space; the fiber is some compact "internal space" having size of order Planck length. What makes the Kaluza-Klein approach so attractive is that it relates the Yang-Mills fields to the geometry of the higher-dimensional space and also raises the hope of explaining the so-called internal quantum numbers geometrically.

The idea that the geometrization of physics, by suitably generalizing and modifying the basic notions of particle and space-time, might solve the unification problem, is also central in the TGD approach (Topological GeometroDynamics) program, which has developed into its present form during the past five years (Pitkänen, 1981, 1983).

The starting point of the TGD approach is what might be called the energy problem of general relativity. In GRT it turns out difficult to find any completely satisfactory definition of the energy concept, and more generally, that of four momentum and angular momentum (Faddeev, 1982; Denisov and Mestrishvili, 1981). In fact TGD can be regarded as an attempt to construct a Poincaré invariant theory of gravitation and of other fundamental interactions based on the following assumptions.

(1) The "physically allowed" space-times X^4 are representable as submanifolds of some space $H = M^4 \times S$, where M^4 denotes Minkowski space and S denotes some compact space with spacelike metric.

(2) The distances in X^4 are measured using the length units of the space H so that the metric of X^4 is induced from the metric of the space H (Eisenhart, 1964).

(3) The theory is invariant under the isometries of the space H so that the fundamental conservation laws follow from the isometries of the M^4 factor of the space H.

In this and subsequent papers our aim is to give an understandable account of the present state of TGD. The plan of the series is roughly the following:

In the first paper we do the following: (1) discuss the basic ideas of TGD (partially familiar from the previous publications); (2) propose a topological description of particles and particle reactions; and (3) consider the constraints on the choice of the space H and show that the space $H = V \times CP_2$, where V denotes either Minkowski space or its light cone and CP_2 is the complex projective space of real dimension 4 (Eguchi et al., 1980; Gibbons and Pope, 1978; Hawking and Pope, 1978), is a respectable candidate.

It should be emphasized that several ideas and results represented in these papers have appeared already in the previous publications but not in their final form.

In the second paper of the series we formulate and study the dynamics at semiclassical level. The third paper of the series is devoted to the construction of quantum TGD.

Notations

Symbol	*Meaning*
$H = V \times S$	imbedding space, which is a Cartesian product of V and S
M^4/M^4_+	Minkowski space/light cone of Minkowski space
CP_2	complex projective space of complex dimension 2
$h^k/m^k/s^k$	coordinates for space $H/M^4/S$
$\xi^k, \bar{\xi}^k, k = 1, 2$	complex coordinates for CP_2
x^α, ξ^α	coordinates for the interior/boundary component of a submanifold X^n
$h^k_{1\alpha}/m^k_{1\alpha}/s^l_{1\alpha}$	partial derivatives of the coordinate variables of $H/M^4/S$ with respect to the coordinate variables of X^n
$h_{kl}/m_{kl}/s_{kl}$	components of the metric tensor for $H/M^4/S$
e^A_k	components of the vielbein in H
$V_k/B_k/A_k$	components of vielbein connection/Kahler potential/spinor connection in H

Symbol	*Meaning*
$V_{kl}/J_{kl}/F_{kl}$	curvature form of vielbein connection/Kahler form/curvature form of spinor connection
$g_{\alpha\beta}=h_{kl}h^k_{1\alpha}h^l_{1\beta}$	induced metric in X^n
$V_\alpha/B_\alpha/A_\alpha$	induced vielbein connection/Kahler potential/spinor connection in X^n
$H^k_{\alpha\beta}=D_\beta h^k_{1\alpha}$	second fundamental form for X^n
$H^k=g^{\alpha\beta}H^k_{\alpha\beta}$	trace of the second fundamental form
Γ_k	gamma matrices for the space H
$\Gamma_\alpha=\Gamma_k h^k_{1\alpha}$	gamma matrices for X^n
$\gamma^A/\Sigma_{AB}=[\gamma_A,\gamma_B]/4$	flat space gamma/sigma matrices for space H
$\tilde{\Gamma}_k/\tilde{\Gamma}_\alpha$	modified gamma matrices of the space H/X^n
D_k/D_α	covariant derivative in H/X^n
X^n	n-dimensional submanifold of H
Int X^n	interior of X^n
$\delta_i X^n$	ith boundary component of X^n

2. BASIC IDEAS

2.1. A Poincaré Invariant Theory of Gravitation

In general relativity it has turned out difficult to find any completely satisfactory definition of the conserved four momentum and angular momentum (Faddeev, 1982; Denisov and Mestrishvili, 1981).

The reason for this circumstance is easy to understand. The basic assumption of GTR is that the presence of the matter makes the surrounding space-time curved so that the symmetries of the empty space-time are lost; space-time is no more homogenous and isotropic. Since the basic conservation laws follow as consequences of these symmetries it is not surprising that one encounters difficulties when trying to find a general coordinate invariant definition of the four momentum and angular momentum as conserved quantities. In practice these difficulties are passed over by arguing that the gravitational interaction is so weak that one can safely do, say particle physics, in Minkowski space.

Unfortunately the extreme weakness of the gravitational interaction makes it difficult to experimentally test whether these conservation laws are exact or not. Therefore it is of considerable interest to try to construct a physical theory based on the following assumptions:

(1) The gravitational interactions are describable in terms of the space-time geometry.

(2) Poincaré transformations act as exact symmetries of the theory.

As far as we know, the only way to construct this kind of theory is based on the following assumptions:

(1) The physically allowed space-times X^4 are representable as submanifolds of some space $H = M^4 \times S$, where M^4 denotes Minkowski space and S denotes some "internal space," which is compact and has spacelike metric.

(2) The distances in X^4 are measured using the length scale of the space H so that the metric $g_{\alpha\beta}$ of X^4 is induced from the metric of the space H or in component form:

$$g_{\alpha\beta} = h_{kl} h^k_{1\alpha} h^l_{1\beta} \tag{1}$$

(here the quantities h^k refer to the coordinate variables of H; see Notations).

(3) The isometries of H act as symmetries of the theory to be constructed so that the basic conservation laws follow from the invariance of the theory under the isometries of the M^4 factor of the space H.

Our attempt to construct a description of fundamental interactions based on these assumptions is based on the following general strategy:

(1) Only the mathematical structures closely related to the natural geometric structures of the space H will be used to build the theory and the mere hypothesis "$X^4 \subset H$" should fix the quantum theory uniquely.

(2) We try to derive the consequences of the basic assumptions in their full generality without imposing any ad hoc restrictions, say, on the topology of the "allowed" space-times.

2.2. The Analogy Between Color and Gravitational Interactions

The assumption that the isometries of H are symmetries of the theory implies a number of new, probably exact, conservation laws associated with the isometries of the space S. Perhaps the most natural identification of these symmetries is as color symmetries.

To our best knowledge, the space CP_2, the complex projective space of four real dimensions (Eguchi et al., 1980; Gibbons and Pope, 1978; Hawking and Pope, 1978), is the lowest-dimensional space having $SU(3)$ as its isometry group [strictly speaking; the isometry group of CP_2 is $SU(3)/Z_3$, not $SU(3)$]. We shall later show that this choice of S indeed seems to be realistic.

The proposed physical identification of the S isometries suggests a deep connection between color and gravitational interactions deriving from the assumption that both interactions couple to the isometry charges. Taking the idea of the color gravitational analogy seriously one can (1) identify the gluonic field variables as quantities closely related to the metric of X^4 and (2) derive a relationship between color and gravitational couplings.

Consider first the identification of the gluonic field variables. There are quite general requirements, which should be taken into account in the construction of these quantities. Gluonic fields should (1) be invariant under the general coordinate transformations of the spaces X^4 and H and (2) define a linear representation for the isometries of the space H.

Field quantities with the required properties are however obtained by contracting the metric tensor $g^{\alpha\beta}$ with the projections

$$j_\alpha^A = j^{Ak} h_{1\alpha}^l h_{kl} \tag{2}$$

of the infinitesimal generators j^A of the isometries:

$$\delta h^k = \varepsilon j^{Ak} \tag{3}$$

to obtain

$$g^{AB} = g^{\alpha\beta} j_\alpha^A j_\beta^B \tag{4}$$

To see that the quantities g^{AB} have the required properties consider first the case that both j^A and j^B generate translations in M^4. In standard Minkowski coordinates m^A for M^4 one has

$$g^{AB} = g^{\alpha\beta} m_{1\alpha}^A m_{1\beta}^B \tag{5}$$

Using Minkowski coordinates also for X^4 (this is possible if X^4 in regions of X^4 representable as a graph for a map $M^4 \to S$) so that the condition $m_{1\alpha}^A = \delta_\alpha^A$ holds, the components of g^{AB} reduce to the components of the metric tensor numerically and deserve an interpretation as the components of the gravitational field (in perturbative approach to GRT the field strengths are closely related to the components of the metric in coordinates, which reduce to Minkowski coordinates in the limit $G \to 0$).

When, say, j^A generates translations in M^4 and j^B generates an $SU(3)$ isometry one obtains a quantity transforming as M^4 vector and color octet. The identification as gluon field is suggestive.

When both j^A and j^B generate isometries of S one obtains a Minkowski scalar transforming as a symmetrized product of two color octets:

$$(8 \otimes 8)_s = 27 \oplus 8 \oplus 1 \tag{6}$$

The gluonic field variables are analogous to the gauge fields of the Kaluza-Klein theories, which are also closely related to the componets of an appropriate metric. In fact, under the general coordinate transformations of H having the following infinitesimal form,

$$\delta s^k = T_A(h) j^{Ak}, \qquad \delta m^k = 0 \tag{7a}$$

and thus resembling local color rotations, the projections of the color currents transform according to the formula

$$\delta j_{\alpha}^{A} = C_{BC}^{A} j_{H}^{B} T^{C} + j^{Ak} j_{Bk} T_{1\alpha}^{B} \tag{7b}$$

Here the quantities C_{BC}^{A} are the structure constants of $SU(3)$ Lie algebra. Thus the transformation law of gluonic fields under local color rotations resembles that associated with ordinary gauge fields.

To find the relationship between color and gravitational couplings we start by writing the Coulombic interaction energy between two colored objects

$$V(r) = -G\bar{Q}_1 \cdot \bar{Q}_2/r = \alpha_s \bar{q}_1 \cdot \bar{q}_2/r \tag{8}$$

Here $\bar{Q}_i$ are dimensional color charges analogous to four momenta and related to the ordinary dimensionless color charges $\bar{q}_i$ via the equation

$$\bar{Q} = (a/R)\bar{q} \tag{9}$$

Here R is the length scale of S: in the case of CP_2 R is fixed by the requirement that CP_2 geodesics have length

$$L = \pi R \tag{10}$$

and a is some numerical constant to be determined.

The determination of the constant a is based on a very simple idea. We merely generalize the representation of the mass squared operator as the d'Alambertian of M^4 ($m^2 = -\Box$) so that one obtains

$$\bar{Q} \cdot \bar{Q} = -\nabla_s^2 \tag{11}$$

This is indeed reasonable in the case of CP_2 since the eigenvalues of the CP_2 Laplacian are proportional to those of the Casimir operator $\bar{q} \cdot \bar{q}$ of $SU(3)$ (Hawking and Pope, 1978)

$$\bar{Q} \cdot \bar{Q} = -a^2 \bar{q} \cdot \bar{q} \tag{12}$$

so that one has

$$a = 2/R \tag{13}$$

Thus the relationship between color and gravitational couplings strengths is given by

$$\alpha_s = 4G/R^2 \tag{14}$$

A crucial step in the above argument is the proportionality between the CP_2 Laplacian and Casimir operator of $SU(3)$. This relationship does not hold for all spaces having $SU(3)$ as isometry group (example $\mathcal{C}^3$) and it would be interesting to know whether there are any other spaces than CP_2 with this property.

Of course, one might argue that the dimensional coupling associated with gravitational interactions associated with M^4 and CP_2 factors need not be the same at long length scales so that the result obtained should hold only at short enough length scales in accordance with the generally accepted picture that the renormalized color coupling becomes infinite at some hadronic length scale. It turns out however that the divergence of the color coupling does not seem to be necessary for the understanding of the color confinement in TGD approach.

2.3. Induction Procedure

In the case of the metric the induction procedure means the restriction of the line element of the space H to the surface X. From the formula defining the induced metric it is clear that the induction procedure for the various tensor quantities in practice means that they are projected to the surface in question.

For a Yang-Mills connection A defined in H, the induction procedure means that the parallel translation in X^4 is performed using the connection A or equivalently:

$$A_\alpha = A_k h^k_{1\alpha} \tag{15}$$

The most promising candidate for the connection to be induced is the spinor connection of the space H, which is defined uniquely for simply connected spaces from the requirement that gamma matrices are covariantly constant quantities (Shanahan, 1978) and is given in terms of vielbein e^A_k (Eguchi et al., 1980) by

$$V_k = D_k e^A_l e^l_B \Sigma^B_A \tag{16}$$

Here D_k denotes the ordinary covariant derivative defined by the standard metric connection of H. The spin matrices are proportional to the commutators of the flat space gamma matrices (see Notations).

The curvature form of the induced Yang-Mills connection is given by

$$V_{\alpha\beta} = V_{kl} h^k_{1\alpha} h^l_{1\beta} \tag{17a}$$

$$V_{kl} = (1/2) R_{klmn} \Sigma^{mn} \tag{17b}$$

Here the tensor R_{klmn} is the curvature tensor of the space H; Yang-Mills interactions thus reflect the nonflatness of the space H.

The gauge group of the vielbein connection is for the generic space H with one timelike direction the noncompact group $SO(\dim H-1, 1)$, but reduces to the group $SO(\dim S)$ for the product decomposition $H = M^4 \times S$. Hence the pathologies associated with noncompact gauge groups afford the "reason why" for the otherwise rather ad hoc choice $H = M^4 \times S$.

Accepting the proposed identification of the color interactions the identification of the components of the induced spinor connection as electroweak gauge potentials becomes natural. In fact, it turns out that the generalized spinor connection of the space CP_2 provides gauge potentials with correct electroweak properties.

In order to induce spinor structure assume that the space H itself allows spinor structure (not true, when H is for instance nonorientable) so that the space H allows globally defined gamma matrices satisfying the anticommutation relations

$$\{\Gamma_k, \Gamma_l\} = 2h_{kl} \tag{18}$$

The gamma matrices in a submanifold X^4 of H are defined as the projections of the gamma matrices of H

$$\Gamma_\alpha = h^k_{1\alpha}\Gamma_k \tag{19}$$

The obvious requirement

$$\{\Gamma_\alpha, \Gamma_\beta\} = 2g_{\alpha\beta} \tag{20}$$

stating that the induced gamma matrices generate a Clifford algebra, is satisfied.

For the spinors themselves the induction procedure means simply restriction to the submanifold X. The chirality (handedness) concept generalizes; the spinors of H carry H-chirality +1 provided space H is even dimensional (Shanahan, 1978). For the spaces containing M^4 as a Cartesian factor also the definition of the ordinary M^4 handedness (and also of S-handedness, if S is even dimensional) is possible.

It turns out that the notion of the chiral invariance generalizes; in its generalized form the chiral invariance implies the existence of two separately conserved fermion numbers corresponding to two H chiralities. One could identify these as baryon and lepton numbers respectively but this is not necessary as it turns out.

The ordinary chiral invariance characteristic to gauge theories (somewhat troublesome since it is not a symmetry of Nature) is lost. The loss is due to the compactness of the space S, which introduces a natural length scale into the theory.

Two important features distinguishing between the induced spinor structure and conventional spinor structure should be mentioned. First, the induced spinor structure is defined for all submanifold topologies unlike the ordinary one, which, for example, fails to be well defined for nonorientable manifolds.

Second, the induced gamma matrices are not covariantly constant:

$$D_\alpha\Gamma_\beta = H^k_{\alpha\beta}\Gamma_k \tag{21}$$

Here the quantities

$$H^k_{\alpha\beta} = D_\beta h^k_{1\alpha} \tag{22}$$

are the components of the so called second fundamental form associated with the submanifold X^4 (see Notations).

This property implies the presence of a "mass term" in the generalized Dirac equation obtained by extremizing the chirally invariant (in generalized sense) Dirac action as will be found later. The mass term resembles the Higgs term and the Higgs field is closely related to the CP_2 part of the second fundamental form.

2.4. Stringlike Objects and the Generalization of Space-Time and Particle Concepts

Stringlike objects appear as extremals of practically any action constructable from local invariants of space X^4. Stringlike objects are typically surfaces $X^4 = X^2 \times S^2$, where X^2 is a minimal surface in M^4 and S^2 a so-called geodesic submanifold of S (Helgason, 1978) having vanishing second fundamental form (S is expected to have size of the order of Planck length). Since the orbit of the string in various string models (Nambu, 1970; Jacob, 1974; Anderson et al., 1983) is classically a minimal surface the identification of the stringlike objects as TGD counter parts of the hadronic strings seems natural.

Accepting this identification one is led to a generalization of the ordinary string model (Nambu, 1970; Jacob, 1974; Anderson et al., 1983). The counterpart of the assumption "Quarks reside at the ends of the string" is the following assumption:

(1) The boundary components of the 3-manifold are carriers of quark like, and more generally, fermionic and possibly also bosonic quantum numbers.

This assumption has served as our guideline for a long time. It has however turned out that the localization of the quantum numbers on the boundaries of the 3-manifold is not necessary and the assumption can be replaced with the following milder one:

(1′) Quarks and also other elementary particles correspond to 3-manifolds with a single boundary component, which are very small in the length scale determined by the hadronic string.

The boundary components are closed two-dimensional manifolds and thus can have several nonequivalent topologies. The orientable, closed 2-manifolds are classified by the so-called genus, which tells how many handles must be attached to a 2-sphere in order to obtain the manifold in question (torus is obtained by adding one handle, etc.). Thus drilling g wormholes starting from and ending at the spherical boundary component, one obtains a boundary component carrying a topological quantum number g.

The idea that g is responsible for the family replication phenomenon (Fritch and Minkowski, 1981) of quarks and leptons seems natural. Above all so because the different fermion families seem to behave identically with respect to known interactions; this behavior is to be expected if the family degeneracy has the proposed origin.

The simplest model for quark/lepton would be a small piece of Minkowski 3-space having naturally an exterior boundary. In the ground state (particles of the lowest generation) boundary has the topology of sphere; topological excitations are obtained by drilling wormholes on the boundary. The size of this manifold is expected to be very small as compared with the size of, say, hadron.

Accepting the genus generation correspondence and either the assumption (1) and (1′) one is led to two alternative descriptions of a hadron:

(2) Hadron is a 3-surface obtained by drilling an arbitrary number of holes with various boundary topologies to the string of type $D^1 \times Y^2$ and by adding appropriate quantum numbers to the resulting boundary components in order to obtain valence quarks and the sea particles.

(2′) Hadron is obtained by "gluing" an arbitrary number of quarklike and gluonic 3-manifolds to the stringlike object having length large as compared with the size of elementary particle 3-manifolds. The gluing operation corresponds to the formation of the so-called connected sum (drill holes D^n to the n-dimensional summands and connect the boundaries S^{n-1} of the holes by a cylinder $D^1 \times S^{n-1}$).

In either case we have a natural descripton for the valence quarks and the sea.

It should be noticed that in the scenario (2′) the family replication phenomenon might be related to the topological degrees of freedom associated with the interior of elementary particle 3-manifold. The roughest classification for the interior topology is in terms of handles ["wormholes" of Wheeler (Adler et al., 1975)] and the number of interior handles might equally well label the various particle families. Whether the family replication phenomenon is related to the number of interior or boundary handles, is at this stage an open question; personally we prefer the explanation based on boundary topology.

Summarizing, if a topological explanation of the generation degeneracy is accepted, an amazingly simple scenario for understanding the basic interactions and particle spectroscopy suggests itself.

(1) Unification takes place at the level of a single particle generation.

(2) Electroweak interactions are the only gauge interactions and the corresponding gauge potentials are induced from the spinor connection of the space H.

(3) Color gravitational interactions are the interactions coupling to the isometry charges of the space H.

3. TOPOLOGICAL CONSIDERATIONS

3.1. Topological Particle Classification

The coarsest topological particle classification is based on the topology of the boundary of the 3-manifold identified as a free particle (only the simplest topologies are expected to correspond to particles in the usual sense of the word). The boundary is specified topologically by the number of the boundary components N and by the topology of the individual boundary components.

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

$$X^2 = P^2 \sharp P^2 \cdots \sharp P^2 \equiv nP^2 \tag{23}$$

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

$$X = B \cup \left(\bigcup_k n_k P^2 \right) \tag{24}$$

where B is a disjoint union of orientable 2-manifolds and in the latter union the constraint

$$\Sigma n_k = 0 \bmod 2 \tag{25}$$

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

A finer classification takes into account also the interior topology. The number of handles ("wormholes") serves as a rough classificational tool for the interior topologies of X^3 (Wallace, 1968).

The topology of the space H can also contribute to the particle classification. When the second homology group $H_2(H)$ (Hilton and Wiley, 1966) is nontrivial one can characterize each boundary component of X^3 by an element of $H_2(H)$, which we shall call homology charge. By the very definition of the concept the total homology charge vanishes for all 3-manifolds.

The concept of the homology charge is different in the orientable and nonorientable categories; the reason is that in the nonorientable category the sign of the homology charge is not defined. For the space $M^4 \times CP_2$, the group in question is isomorphic to the group Z of integers and to two element group Z_2 in orientable and nonorientable categories, respectively. The identification of the homology charge as magnetic charge turns out to be possible and the condition

$$\Sigma n_k = 0 \bmod 2 \tag{26}$$

states that no free magnetic monopoles exist.

3.2. Topological Description of Particle Reactions

The identification of free particles as compact 3-manifolds not only implies a topological particle classification but also a topological description of particle reactions. In this section we therefore address ourselves to the question "How does the dual diagrammatics of the string model generalize?"

We are interested in the "basic vertices" for the particle reactions described as topology changes and in possible topological selection rules. We base our approach on an intuitive idea that any particle state corresponds to a set of spacelike 3-manifolds imbedded in some spacelike ($\dim H - 1$) manifold of H. The problem can be stated in more mathematical terms as follows:

Given two three-dimensional spacelike submanifolds X_i^3 and X_f^3 in an ($\dim H - 1$)-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 spacelike boundaries so that the 4-manifold in question mediates a transition between the initial and final states?

Can one decompose the transition into more elementary ones? Which are the "basic vertices" and are there any selection rules of a purely topological origin?

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

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

(1) The changes in the purely internal topology of the 3-manifold; the number of the components of X^3 and the boundary topology remain unchanged.

(2) The reactions changing the particle number defined as the number of components of X^3 (only the simplest 3-topologies are expected to correspond to particles in the ordinary sense).

(3) The transitions involving a change in the boundary topology; either the topology of a single boundary component is changed or the number of the boundary components is changed.

3.2.1. Changes in the Internal Topology

Since the topology of the boundary is unchanged in these reactions it is reasonable to restrict the consideration to the cobordism of the closed (boundaryless) 3-manifolds. One obtains a rough idea about what is involved by observing that the problem reduces to a homology problem (Hilton and Wiley, 1966) if one gives up the requirement that the surfaces are manifolds, i.e., they can for instance intersect themselves. The selection rules for the homology result from the nontriviality of the third homology group of H, which is trivial for example for the space $H = M^4 \times CP_2$.

Therefore the possible selection rules result from the requirement that the intermediate surface is a manifold: both the internal topology of the manifold and the finite dimension of the imbedding space can in principle 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 the imbedding space (the requirement of causality might also set restrictions on the cobordism).

The problem of constructing basic vertices for these changes is solved (Wallace, 1968; Milnor, 1965). The characteristic property of the vertices is the localizability of the topology change; the change takes place via an intermediate 3-space, which fails to be a manifold at a single point (in the generic case). The transition from the topology of torus to that of a sphere serves as a simple two-dimensional illustration of this property.

3.2.2. Reactions Changing Particle Number

There are three basic vertices corresponding to change in particle number. We shall refer to these vertices as (1) connected sum ($\sharp$), (2) boundary connected sum ($\sharp_B$), and (3) fusion ($\natural$) vertex, respectively.

Connected sum vertex ($\sharp$) is a generalization of the "trouser" vertex of the string model. The "reactants" merge together at a point common to their interiors. This vertex is the only vertex leading to a change of the component number of n-manifold in the cobordism of closed n-manifolds (Wallace, 1968).

Boundary connected sum ($\#_B$) vertex corresponds to a process, where two 3-manifolds join along their boundary components. Since this vertex is not completely local one might ask whether this vertex is "generic" and thus whether it has any significant dynamical role in three-dimensional cases.

The diagrams obtained using $\#$ and $\#_B$ vertices correspond to diagrams encountered already in the ordinary string model. The homological triviality of the leptonic boundary components explains why they do not participate in those strong interactions, which correspond to the joining of the strings along their ends.

The fusion ($\natural$) vertex of two 3-manifolds at a point common to their boundaries (common to interiors in $\#$ vertex) is a vertex having no counterpart in the ordinary string model. The fusion of the liquid droplets described as manifolds with a spherical boundary proceeds via this vertex. Clearly, the handle number and the homology charge are conserved in this vertex. It is attractive to identify this vertex as responsible for the emission and absorption of elementary particles (gauge boson, leptons,...).

3.2.3. Reactions Changing the Boundary Topology

The reactions changing the boundary topology can be described by suitably generalizing the vertices already found.

For boundary components belonging to same 3-manifold $\#_B$ vertex leads to the quark annihilation diagram forbidden by the so-called OZI rule (Chew and Rosenzweig, 1978); the nonlocal nature of $\#_B$ vertex might explain this rule.

The $\natural$ vertex generalized so that the boundary components belong to same 3-manifold leads to a change in the number of boundary components. The local equivalence of ordinary and "internal" $\natural$ vertices suggests that the decay of, say, a hadronic boundary component via the internal $\natural$ vertex corresponds physically to the emission of a "sea particle."

We can generalize the $\natural$ vertex even further by assuming that the two regions belonging to same boundary component merge together at a common point. Since the genus of the boundary component changes in this transition by one unit, a natural expectation is that it is this transition, which is responsible for the various mixing phenomena [Cabibbo and neutrino mixings (Kobayashi and Maskawa, 1973; Bilenki and Pontocorvo, 1978)]. An interesting question is, whether the $\Delta g = 1$ nature of this transition could lead to any experimentally verifiable predictions concerning the nature of, say, Cabibbo mixing.

Concluding, we have found that three basic vertices for the cobordism of the 3-manifolds (not taking into account the vertices changing the interior topology, but not the component number of 3-manifold). The vertices $\natural$

and possibly also $\sharp_B$ might be the vertices responsible for the various reactions between elementary particles; this is certainly the case if elementary particle quantum numbers reside on the boundaries.

3.3. Classical 3-Space as a Topological Many-Particle Phenomenon

The proposed generalization of the space-time and particle concepts is rather radical, and, one could well argue, in contradiction with everyday experience. We however think that this need not be the case.

Consider first what one understands with the notion of the classical space-time in different contexts. The particle physicist in practice identifies the classical space-time as the empty Minkowski space, which via its symmetries is responsible for the basic conservation laws. It should be obvious that the space H corresponds to the generalization of the particle physicist's conception of the space-time, making possible, as we hope, to understand all elementary particle characteristics in terms of geometry and topology.

In various applications of GRT, such as cosmology, the space-time (or rather 3-space) is understood as a dynamical object and it is natural to define it as an extremal of some suitably defined effective action in accordance with the basic ideas of the quantum theory. It is however not at all obvious that surfaces with those properties, which we believe to be characteristic to the space-time of general relativity should have so marked a dynamical role as the everyday experience suggests.

Thus it is natural to seek a mechanism, which would in some sense generate the classical space-time. In this respect the following observations are crucial.

(1) Any classical action is expected to allow extremals representable as a graph for some map $M^4 \to S$ (simplest example: imbedding of M^4 as the surface $M^4 \times \{s\}$, where s is some point of S). The properties of these surfaces are in accordance with the conventional notion of space-time.

(2) The spaces H having dimension $d < 9$ are exceptional in the sense that the intersection of two 4-surfaces is stable under small perturbations of the surfaces. This property can be understood by using for the surfaces in question the representation

$$f^k = 0, \qquad k = 1, \ldots, d-4$$

where f^k are some functions defined in H.

The first observation raises the question "What happens for a particle like 3-manifold in the presence of the 'vacuon' that is 3-surface representable as graph and having macroscopic size?" It is the second observation, which suggests a possible answer to this question.

What we believe to happen is that the particle "collides" with the vacuon with a high probability and that under certain conditions the connected sum of the particle and of the vacuon is formed. A convenient visualization of the situation is obtained by considering the motion of a particlelike 2-surface in the presence of a 2-plane in a finite volume of Euclidian 3-space.

The observations just made suggest that the classical space-time with matter is in a certain sense a topological many-particle phenomenon, which we call hereafter "# condensation": in presence in the vacuon the transition

$$\text{particle} \cup \text{vacuon} \rightarrow \text{particle} \# \text{vacuon}$$

occurs with a high probability provided that the condition "dim $H < 9$" is satisfied. As a consequence we obtain something, which might be called "classical space-time with matter": matter being presented as topological inhomogeneities glued to the topologically trivial background.

What we have obtained is of course not yet the topologically trivial space-time of GRT. It is however intuitively clear that the classical space-time of GRT is obtained, when one defines the classical space-time as a length-scale-dependent concept. The classical space-time in length scale L is by definition topologically trivial in length scales smaller than L and is an extremal of some length-scale-dependent effective action.

The #-condensed matter corresponds to topological inhomogeneities of size smaller (!) than L. Because these inhomogeneities are pointlike in the length scale considered we expect them to be representable using continuous mass, charge . . . distributions defined in the "classical space-time in length scale L."

We shall return to a more precise formulation of the concept of the classical space-time as a length-scale-dependent concept in the second paper of the series.

Some further remarks concerning the #-condensation phenomenon are in order. First, the identification of the # condensate as a many-particle bound state is attractive since it would provide a general topological description of bound states applying to macroscopic gravitationally bound states as well as to molecules, atoms, and nuclei. Even hadrons might be understood as quark–gluon # condensates around the hadronic strings (or possibly bags depending on the length scale used in the description).

Second, if this interpretation is correct one expects that the reverse phenomenon ("# evaporation") occurs if the binding energy of the particle in condensate is too small. Thus in temperatures of the order of the binding energy of the particle "# evaporation" should occur. Finally, one must take in the description of matter into account both the #-condensed phase and the "vapor" phase. We shall return to this subject in the second paper of the series.

4. CHOICE OF THE SPACE H

In this section we shall consider the problem of finding a realistic candidate for the space H. We study the various constraints, which the space H should satisfy and show that the space $H = V \times CP_2$, where V denotes Minkowski space or its light cone is a respectable candidate.

4.1. Constraints on the Choice of the Space H

In order to attack the problem of finding a realistic candidate for the space H let us first recall some basic ideas of TGD approach.

(1) The gauge group of the spinor connection is compact.

This requirement favors strongly the product composition $H = V \times S$, where V is flat.

(2) Poincaré invariance is exact.

The choice $H = M^4 \times S$ guarantees the exact Poincaré invariance. The alternative choice $H = M^4_+ \times S$, where M^4_+ is in the light cone of Minkowski space, realizes Poincaré invariance in noncosmological scales. This alternative must be considered seriously since big bang cosmologies reduce to M^4_+ in the limit of vanishing mass density as can be seen from the expression for the line element of M^4_+:

$$ds^2 = da^2 - a^2(dr^2/(1+r^2) + r^2\, d\Omega^2) \tag{27}$$

Here the coordinates (a, r, θ, ϕ) are related to the standard spherical coordinates (m^0, r_M, θ, ϕ) of M^4 via the equations

$$\begin{aligned} a^2 &= (m^0)^2 - r_M^2 \\ ra &= r_M \end{aligned} \tag{28}$$

Clearly, this choice makes the big bang cosmology a necessity.

(3) Free particle corresponds to a compact 3-manifold.

This generalization of the particle concept leads to a topological explanation for the family replication phenomenon. The hypothesis that various boundary component topologies correspond to different elementary particle generations, "generation genus correspondence," nicely explains the identical properties of the various fermion generations with respect to the fundamental interactions.

Furthermore, this hypothesis simplifies decisively the task of finding realistic candidates for the space H; the remaining quantum numbers are (besides mass and spin) baryon and lepton numbers (or perhaps a suitable combination of them), color and electroweak quantum numbers.

(4) Color symmetries correspond to the isometries of the space S.

This identification is natural since color and Poincaré symmetries are believed to be exact. Moreover, it implies a deep relationship between color and gravitational interactions. The space CP_2, the complex projective space is to our best knowledge the lowest-dimensional space having $SU(3)/Z_3$ as its isometry group.

(5) The spinor connection of the space S (defined almost uniquely by the metric of S) defines Yang–Mills gauge potentials on the surface X^4 via the induction procedure. It is natural to identify these field quantities as gauge potentials.

Since the spinors of $2N$-dimensional space have $2^{N-1}+2^{N-1}$ components corresponding to two different chiralities (Anderson et al., 1983) it is tempting to identify the spinors with different chiralities as quark- and leptonlike spinors, respectively. The chiral invariance, which has a somewhat problematic nature in the conventional gauge theories would in its generalized form account for the conservation of baryon and lepton numbers. The separate conservation of baryon and lepton numbers is however not a necessary outcome of the TGD approach as will be found in the sequel.

In the case of CP_2 the spinors with definite H chirality have eight components each. Assuming that the right-handed neutrinos do exist, this number equals to the number of electroweak degrees of freedom inside a single fermionic generation.

Clearly, the color quantum numbers cannot be spinlike as in QCD but rather resemble angular momentum. That is, the colored fermions correspond to nontrivial partial waves in the partial wave expansion of the spinor field in H.

Thus CP_2 appears to be a promising candidate for the space S. There are however some serious problems.

(1) The presently known leptons/quarks are color singlets/triplets and thus satisfy the triality rule:

$$Q_{\text{em}} = t/3 \text{ mod integer} \qquad [t \text{ is the } SU(3) \text{ triality}]$$

How can one understand this rule? How are the $t \neq 0$ states even possible [recall that the isometry group of CP_2 is $SU(3)/Z_3$]?

(2) The right-handed neutrinos should effectively decouple from the electroweak interactions.

(3) The vielbein group of CP_2 is $SO(4) = SU(2)_l \times SU(2)_r$, where the labels refer to CP_2 chiralities. For a fixed H chirality these labels correspond to M^4 chiralities. It is clear that the vielbein group cannot contain the electroweak gauge group $SU(2)_L \times U(1)$ as its subgroup; the $U(1)$ factor is missing. And last but not least:

(4) CP_2 does not allow any spinor structure!

This list of problems is rather impressive. It turns out, however, that the Kahler structure of CP_2 (Eguchi et al., 1980; Gibbons and Pope, 1978; Hawking and Pope, 1978), which implies the existence of a covariantly constant 2-form and the associated $U(1)$ gauge potential, provides a solution to the purely mathematical problem (4) and also to the other problems.

4.2. Basic Properties of CP_2

4.2.1. CP_2 as a Manifold

CP_2, the complex projective 2-space is defined by identifying the points of the complex 3-space $\mathbb{C}^3$ under the equivalence

$$(z^1, z^2, z^3) \equiv \lambda (z^1, z^2, z^3) \tag{29}$$

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 each other (e.g., CP_2 is a complex manifold). The points $z^3 \neq 0$ 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 the 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; Gibbons and Pope, 1978) will be used and their relation to the complex coordinates is given by

$$\begin{aligned} \xi^1 &= z + it \\ \xi^2 &= x + iy \end{aligned} \tag{30}$$

These are related to "spherical" coordinates via the equations

$$\begin{aligned} \xi^1 &= r \exp(i(\psi + \phi)/2] \cos(\theta/2) \\ \xi^2 &= r \exp[i(\psi - \phi)/2] \sin(\theta/2) \end{aligned} \tag{31}$$

The ranges of the variables r, θ, ϕ, and ψ 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 $b = 1$. The last property stems from the fact that the second homology group $H_2(CP_2)$ is isomorphic to integers.

4.2.2. Metric and Kahler Structures of CP_2

In order to obtain a natural metric for CP_2 observe that CP_2 can be thought as a set of the orbits of the isometries $z^i \to \exp(i\alpha) z^i$ on the sphere

S^5: $\Sigma|z^i|^2 = R^2$. The metric of CP_2 is obtained by projecting the metric of S^5 orthogonally to the orbits. Therefore the distance between the points of CP_2 is that between the representative orbits of S^5. The line element has the following form in the complex coordinates

$$ds^2 = g_{a\bar{b}}\, d\xi^a\, d\xi^{\bar{b}} \tag{32}$$

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

$$g_{a\bar{b}} = R^2_{\partial a \partial \bar{b}} \ln F \tag{33}$$

The quantity F is defined as

$$F = 1 + r^2 \tag{34}$$

An explicit representation of the metric is given by

$$ds^2/R^2 = (dr^2 + r^2\sigma_3^2)/F + r^2(\sigma_1^2 + \sigma_2^2)/F^2 \tag{35}$$

where the quantities σ_i are defined as

$$\begin{aligned} &r^2\sigma_1 = \mathrm{Im}(\xi^1\, d\xi^2 - \xi^2\, d\xi^1), \qquad r^2\sigma_2 = -\mathrm{Re}(\xi^1\, d\xi^2 - \xi^2\, d\xi^1) \\ &r^2\sigma_3 = -\mathrm{Im}(\Sigma \xi^k\, d\bar{\xi}^k) \end{aligned} \tag{36}$$

The vierbein forms, which satisfy the defining relation

$$s_{kl} = R^2 e_k^A e_{Al} \tag{37}$$

are given by

$$\begin{aligned} &e^0 = dr/F, \qquad && e^1 = r\sigma_1/\sqrt{F} \\ &e^2 = r\sigma_2/\sqrt{F}, \qquad && e^3 = r\sigma_3/F \end{aligned} \tag{38}$$

The vierbein connection satisfying the defining relation

$$de^A = -V_B^A \wedge e^B \tag{39}$$

is given by

$$\begin{aligned} &V_{01} = -e^1/r, \qquad && V_{23} = e^1/r \\ &V_{02} = -e^2/r, \qquad && V_{31} = e^2/r \\ &V_{03} = (r - 1/r)e^3, \qquad && V_{12} = (2r + 1/r)e^3 \end{aligned} \tag{40}$$

The representations of the curvature (components of the curvature tensor in the vielbein basis) are constant reflecting the fact CP_2 is a constant curvature space:

$$\begin{aligned} &R_{01} = e^0 \wedge e^1 - e^2 \wedge e^3, \qquad && R_{23} = -e^0 \wedge e^1 + e^2 \wedge e^3 \\ &R_{02} = e^0 \wedge e^2 - e^3 \wedge e^1, \qquad && R_{31} = -e^0 \wedge e^2 + e^3 \wedge e^1 \\ &R_{03} = 4e^0 \wedge e^3 + 2e^1 \wedge e^2, \qquad && R_{12} = 2e^0 \wedge e^3 + 4e^1 \wedge e^2 \end{aligned} \tag{41}$$

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

$$J = -ig_{a\bar{b}}\, d\xi^a \wedge d\bar{\xi}^b \tag{42}$$

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

$$J^k_{\ l} J^l_{\ m} = -\delta^k_m \tag{43}$$

The form $2J$ is integer valued and by its covariant constancy satisfies free Maxwell equations. Hence it can be regarded as a curvature form of a $U(1)$ gauge potential B carrying a magnetic charge of unit $1/2g$ (g denotes the gauge coupling). Locally one has therefore

$$J = dB \tag{44}$$

where B is the so-called Kahler potential.

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 representation of J and B are given by

$$\begin{aligned} B &= 2re^3 \\ J &= 2(e^0 \wedge e^3 + e^1 \wedge e^2) \end{aligned} \tag{45}$$

The vielbein curvature form and Kahler form are that in complex coordinates are covariantly constant and in complex coordinates they have only components of type $V_{a\bar{b}} = -V_{\bar{b}a}$ and $J_{a\bar{b}} = -J_{\bar{b}a}$, respectively ($V_{ab} = V_{\bar{a}\bar{b}} = 0$ and $J_{ab} = J_{\bar{a}\bar{b}} = 0$). These properties turn out to be of central importance in the sequel.

4.2.3. Spinors in CP_2

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

To see how 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 with a base point x leads to a rotated vierbein at x: $\mathbf{e}^A = R^A_B e^B$ 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 S^2 in M and the elements $R^A_B(v)$ define a closed path in $SO(4)$. When the sphere 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 elements of the homotopy group $[SO(4)] = Z_2$. However, for homologically nontrivial 2-surface S^2 the associated path in $SO(4)$ can be homotopically nontrivial and therefore corresponds to nonclosed path in the covering group Spin (4) (leading from the matrix 1 to -1 in the matrix representation). Assume this is the case.

Assume now that the space allows a spinor structure. Then we can parallely 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 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 Φ is the magnetic flux through the surface. This factor has the value -1 provided the $U(1)$ potential carries half odd multiple of Dirac charge $1/2g$. In case of CP_2 the required gauge potential is half odd multiple of the Kahler connection B defined previously. In the case of $M^4 \times CP_2$ we can in addition couple the spinor components with different chiralities independently to an odd multiple of $B/2$.

4.3. ELECTROWEAK COUPLINGS

The delicacies of the spinor structure of CP_2 make it a unique candidate for the space S. First, the coupling of the spinors to the $U(1)$ gauge potential defined by the Kahler structure provides us with the missing $U(1)$ factor. Second, it is possible to couple the different H chiralities independently to a half odd multiple of the Kahler potential. Thus the hopes of obtaining a correct spectrum for the electromagnetic charge are considerable. In this section we demonstrate that the couplings of the induced spinor connection are indeed those of the GWS model (Weinberg, 1967; Salam, 1968; Glashow, 1961) and in particular, that the right-handed neutrinos decouple completely from electroweak interactions.

We begin by recalling that the space H allows us to define three different chiralities for spinors. Spinors with a fixed H, CP_2 chirality (l, r) and M^4 chirality (L, R) are defined by the condition

$$\Gamma\psi = e\psi, \qquad e = \pm 1 \tag{46}$$

where Γ denotes the matrix $\gamma_5 \times \gamma_5$, $1 \times \gamma_5$, and $\gamma_5 \times 1$, respectively. Clearly, for a fixed H chirality the M^4 and CP_2 chiralities are correlated.

The spinors with H chirality $e = \pm 1$ ($\Gamma_1 \psi = e\psi$) could be tentatively identified as quark- and leptonlike spinors respectively but is not necessary. For spinors of definite H chirality one can perform the identification $SO(4) = SU(2)_L \times SU(2)_R$ for the vielbein part of the gauge group.

The covariant derivatives are defined by the spinorial connection

$$A = V + (B/2)(n_+ 1_+ + n_- 1_-) \tag{47}$$

Here V and B denote the projections of the vielbein and Kahler gauge potentials, respectively and 1_e, $e = +1/-1$ projects to subspace with H chirality $+1/-1$. The integers n_e are odd from the requirement of a respectable spinor structure.

The explicit representation of the vielbein connection V and of B is given by the equations

$$\begin{aligned} V_{01} &= -e^1/r, & V_{23} &= e^1/r \\ V_{02} &= -e^2/r, & V_{31} &= e^2/r \\ V_{03} &= -(r - 1/r)e^3, & V_{12} &= (2r + 1/r)e^3 \\ B &= 2re^3, \end{aligned} \tag{48}$$

The explicit representation of the vielbein will not be needed here.

Let us first show that the charged part of the spinor connection couples purely left handedly. Identifying Σ_3^0 and Σ_2^1 as the diagonal (neutral) Lie-algebra generators of $SO(4)$ one finds that the charged part of the connection is given by

$$A_{\mathrm{ch}} = 2V_{23} I_L^1 + 2V_{13} I_L^2 \tag{49}$$

where we have defined

$$\begin{aligned} I_L^1 &= (\Sigma_{01} - \Sigma_{23})/2 \\ I_L^2 &= (\Sigma_{02} - \Sigma_{13})/2 \end{aligned} \tag{50}$$

A_{lh} is clearly purely left-handed so that we can perform the identifications

$$W^{+(-)} = 2(e^1 \underset{(-)}{\overset{+}{}} ie^2)/r \tag{51}$$

where $W^{+(-)}$ denotes the intermediate vector boson.

Consider next the identification of the neutral gauge bosons γ and Z^0 as appropriate linear combinations of the two functionally independent quantities

$$X = re^3$$
$$Y = e^3/r \tag{52}$$

appearing in the neutral part of the spinor connection. We show first that the mere requirement "photon couples vectorially" implies the basic coupling structure of the GWS model leaving only the value of the Weinberg angle undetermined.

To begin with we define

$$\begin{pmatrix} \bar{\gamma} \\ \bar{Z}^0 \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} X \\ Y \end{pmatrix} \tag{53}$$

imposing the normalization condition

$$ad - bc = 1 \tag{54}$$

The physical fields γ and Z^0 are related to $\bar{\gamma}$ and $\bar{Z}^0$ via simple normalization factors.

Expressing the neutral part of the spinor connection using these fields one obtains

$$A_{\rm nc} = [(c+d)2\Sigma_{03} + (2d-c)2\Sigma_{12} + d(n_+1_+ + n_-1_-)]\bar{\gamma}$$
$$+ [(a-b)2\Sigma_{03} + (a-2b)2\Sigma_{12} - b(n_+1_+ + n_-1_-)]\bar{Z}^0 \tag{55}$$

Identifying Σ_{12} and $\Sigma_{03} = 1 \times \gamma_5\Sigma_{12}$ as vectorial and axial Lie algebra generators, respectively, the requirement that γ couples vectorially leads to the condition

$$c + d = 0 \tag{56}$$

Substituting the result to equation (3.29) one obtains for the neutral part of the connection the expression

$$A_{\rm nc} = \gamma Q_{\rm em} + Z^0(I_L^3 - \sin^2\theta_w Q_{\rm em}) \tag{57}$$

Here we have defined the electromagnetic charge $Q_{\rm em}$ and the weak isospin by

$$Q_{\rm em} = \Sigma_{12} + (n_+1_+ + n_-1_-)/6$$
$$I_L^3 = (\Sigma_{12} - \Sigma_{03})/2 \tag{58}$$

The fields γ and Z^0 are defined via the relations

$$\gamma = 6d\bar{\gamma} = [6/(a+b)](aX + bY)$$
$$Z^0 = 4(a+b)Z = 4(X - Y) \tag{59}$$

The value of the parameter θ_W (Weinberg angle) is given by

$$\sin^2 \theta_W = 3b/[2(a+b)] \tag{60}$$

Observe that right-handed neutrinos decouple completely from the electroweak interactions.

The determination of the actual value of the Weinberg angle is a dynamical problem. A possible way to determine it is to express the Yang–Mills part of the effective (length-scale-dependent) bosonic action in terms of the physical fields γ and Z^0 and to require that no nondiagonal terms of form γZ^0 appear in it. Obviously the pure Yang–Mills action leads to a definite value of the Weinberg angle. However, the addition of the term proportional to

$$I = J^{\alpha\beta} J_{\alpha\beta} \tag{61}$$

where $J_{\alpha\beta}$ is the projection of the Kahler form, to the effective action clearly changes the value of the Weinberg angle. This can be interpreted as a term describing the different evolution of $SU(2)_L$ and $U(1)$ coupling constants. We will discuss this question in more detail in the forthcoming paper of the series devoted to the study of the dynamics on a semiclassical level.

We have identified the physical field quantities in a definite gauge. The identification procedure is however invariant under the electroweak gauge group as we will now demonstrate. The invariance follows from the following properties of identification procedure.

(1) The charged gauge bosons are identified as the charged left-handed part of the gauge field. Thus the definition of the charged part is invariant under $SU(2)_L$ and right-handed rotations generated by the third component of the right-handed isospin.

(2) The neutral gauge bosons can be expressed in the form

$$\begin{pmatrix} \gamma \\ Z^0 \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} A_L \\ A_R \end{pmatrix} \tag{62}$$

Here A_L/A_R is the neutral left/right-handed part of the spinor connection

$$\begin{aligned} A_L &= (r + 2/r) e^3/2 \\ A_R &= 3 r e^3/2 \end{aligned} \tag{63}$$

The invariance of the definition under left-handed vielbein rotations is thus obvious. Since the right-handed part is proportional to the Kahler potential (the geometry of CP_2 again!) the definition of the field quantities is also invariant under gauge transformations, which preserve this proportionality. These transformations are just the electroweak $U(1)$ transformations generated by a definite linear combination of the right-handed neutral isospin generator and of unit matrix.

4.4. Higgs Mechanism and $SU(2)_L$ Confinement

The Higgs mechanism plays a crucial role in GWS model as a mechanism giving masses to the intermediate gauge bosons. Hence it is natural to seek for the counterpart of the Higgs mechanism in the TGD approach. In this section we shall show the following:

(1) The propagator associated with the most general chirally invariant spinorial action describing leptons and quarks as elementary particles contains terms having the properties of the Higgs type mass term.

(2) Our candidate for the Higgs field indeed has correct group theoretical properties in the sense that ordinary gauge theory describing the interaction of gauge bosons and a Higgs field with the same group theoretical properties leads to the same predictions as GWS model for the mass ratios of the electroweak gauge bosons.

In addition we perform a transition to the so-called unitary gauge, where the Higgs field has no charged components, and find out rather surprisingly the following:

(3) The members inside the leptonic and quark doublets have equal masses unless the action contains a CP-odd term made possible by the Kahler structure of CP_2 and that the requirement "neutrinos are massless" fixes the CP-odd term uniquely.

(4) The field quantities obtained by making the transition to the unitary gauge can be regarded as $SU(2)_L$ singlets of type Higgs fermion and Higgs–Higgs gauge boson. This unexpected result means that in TGD approach the phenomena of $SU(2)_L$ confinement and symmetry breaking via the Higgs mechanism seem to be one and the same thing.

Consider first the construction of the spinorial action. The most general spinorial action based on the assumption that both leptons and quarks are elementary fermions (we shall represent an alternative scenario later) is given by

$$L = i\bar{\psi}[\tilde{\Gamma}^{\alpha}(l)\vec{\mathbf{D}}_{\alpha} - \overleftarrow{\mathbf{D}}\tilde{\Gamma}^{\alpha}(-l)]\psi\sqrt{(-|g|)}$$
$$\tilde{\Gamma}_{\alpha}(l) = \tilde{\Gamma}_{k}(l)h^{k}_{1\alpha} \tag{64}$$

Here the quantities $\tilde{\Gamma}_k(l)$ are related to the gamma matrices of H via the equations

$$\tilde{\Gamma}_{k(l)} = (h^{r}_{k} + ilJ^{r}_{k})\Gamma_{r} \tag{65}$$

l being a real number. Observe that hermiticity requires the appearence of the both values of l in the Lagrangian.

The covariant derivative D_{α} is defined by the induced spinor connection

$$D_{\alpha} = \partial_{\alpha} + V_{\alpha} + (B_{\alpha}/2)(n_{+}1 + n_{-}1_{-}) \tag{66}$$

The term V denotes the vielbein part of the spinor connection and the term B denotes the projection of the Kahler connection to the surface X^4.

The Kahler structure of CP_2 makes possible the term proportional to the parameter l. The CP oddness of this term will be shown in the context of the discrete symmetries to follow from the oddness of the Kahler form under the charge conjugation, which corresponds geometrically to complex conjugation in CP_2.

The anticommutation relations satisfied by the "modified gamma matrices" are given by

$$\{\tilde{\Gamma}_k, \tilde{\Gamma}_l\} = 2(m_{kl} + (1 - l^2)s_{kl}) \tag{67}$$

When the condition $l^2 = 1$ is satisfied the matrix defined by the anticommutators becomes noninvertible and thus the actions satisfying this condition are in a mathematically distinguished position.

By varying the generalized Dirac action with respect to the spinor variables one obtains something, which can be regarded as a generalization of the Dirac equation

$$\Gamma^\alpha D_\alpha \Psi = -(1/2)H^k \tilde{\Gamma}_k \Psi \tag{68}$$

The quantities H^k define the trace of the second fundamental form $H^k_{\alpha\beta}$ associated with the surface X^4. The components of the second fundamental form are defined as the covariant derivatives of the quantities $h^k_{1\alpha}$, which as vectors of H are parallel to the surface X^4:

$$H^k = g^{\alpha\beta} H^k_{\alpha\beta} = g^{\alpha\beta} D_\beta h^k_{1\alpha} \tag{69}$$

Observe that the addition of the CP-breaking term has only changed the term containing the second fundamental form (CP operation changes effectively the sign of the parameter l in this term).

The operator

$$\Gamma^\alpha D_\alpha + H^k \tilde{\Gamma}_k / 2 \tag{70}$$

can be identified as the inverse of the fermionic propagator in X^4. The part of $H^k \tilde{\Gamma}_k$ involving only the gamma matrices of CP_2, which we shall denote by the symbol M, is given by

$$M = S^A \tilde{\Gamma}_A = M^A \Gamma_A \tag{71}$$

and is a good candidate for mass matrix. The reason is that the vector S^A transforms under the vielbein group $SO(4)$ as a (1/2, 1/2) representation and thus under $SU(2)_L$ it has the same transformation properties as the complex Higgs doublet of the GWS model transforming as (1/2, 0) under $SO(4)$.

Consider first the group theoretical properties of our might be Higgs. The predictions of the GWS model for the ratio $r = m_w/m_z$ and for m_γ are both in agreement with experiment. Owing to the sensitivity of the ratio r to the details of the Higgs mechanism it is important to check whether the Higgs field transforming according to the real $(1/2, 1/2)$ representation of $SO(4)$ can in principle reproduce the predictions of the GWS model.

To resolve this question consider the ordinary field theory in Minkowski space obtained by coupling $(1/2, 1/2)$ Higgs field $\mathcal{M}$ to the gauge boson field and assign the following gauge invariant action:

$$L = L_1 + \mathrm{Tr}(D^\alpha \mathcal{M} D_\alpha \mathcal{M}) + V(\mathcal{M}) \tag{72}$$

Here L_1 denotes the pure YM action density. The convariant derivative appearing in the Higgs part of the action is defined as

$$D_\alpha \mathcal{M} = \partial_\alpha M + [A_\alpha, \mathcal{M}] \tag{73}$$

The term $V(M)$ is a self-interaction term and is assumed to have minimum for a nonvanishing vacuum expectation of the Higgs field.

A nonvanishing vacuum expectation of S

$$\mathcal{M} = M^0\Gamma^0 + M^3\Gamma^3 \tag{74}$$

leads to effective mass terms in the Yang-Mills part of the action. The photon remains massless because the mass matrix commutes with it and for the ratio r one obtains

$$r = m_w/m_z = g_w/g_z \tag{75}$$

just as in the GWS model (Weinberg, 1967; Salam, 1968; Glashow, 1961) so that our Higgs candidate passes the first test.

In order to obtain additional information about the properties of the Higgs candidate we shall perform transition to the "unitary gauge" (Weinberg, 1967; Salam, 1968; Glashow, 1961) defined as a gauge, where the Higgs field has no charged components.

We define the mass matrix via the following decomposition:

$$\begin{aligned} \bar{\psi}\mathcal{M}\psi &= \bar{\psi}^+_{R,r} M^+_{L,l} \psi^+_{L,l} + \bar{\psi}^+_{L,l} M^+_{R,r} \psi^+_{R,r} \\ &\quad + \bar{\psi}^-_{R,r} M^-_{L,r} \psi^-_{L,r} + \bar{\psi}^-_{L,r} M^-_{R,l} \psi^-_{R,l} \\ \mathcal{M} &= (\delta^A_B + ilJ^A_B) S^B \gamma_5 \times \gamma_A \end{aligned} \tag{76}$$

The mass operator indeed has the correct properties, i.e., it connects different Minkowski space chiralities of the spinor field ψ and its conjugate $\bar{\psi}$.

In order to perform the diagonalization we use the representation of the flat space gamma matrices of CP_2 defined in the spinor basis with definite CP_2 chiralities. This representation is given by

$$\gamma_A=\begin{pmatrix} 0 & i\sigma_A \\ i\sigma_A & 0 \end{pmatrix}, \quad A=1,2,3, \qquad \gamma_0=\begin{pmatrix} 0 & \sigma_0 \\ -\sigma_0 & 0 \end{pmatrix} \tag{77}$$

Here the matrices σ_A, $A=1,2,3$ are the ordinary 2×2 Pauli spin matrices and σ_0 is the 2×2 unit matrix.

In this representation the mass matrix is given by

$$\begin{aligned} M_R^+ &= -M_L^- = M^0\sigma_0 - iM^3\sigma_3 \\ M_L^+ &= -M_R^- = M^0\sigma_0 + iM^3\sigma_3 \end{aligned} \tag{78}$$

The action of the group $SU(2)_L$ on the spinors and on the Higgs field takes place via left/right multiplication by a unitary 2×2 matrix in this representation. The corresponding representation of the left-hand part of the vielbein connection is given by

$$V^L = i\varepsilon^{ABC}V_{BC}^L\sigma_A \equiv iV^A\sigma_A \tag{79}$$

and the action of the group $SU(2)_L$ is given by

$$V \to UVU^+ - dUU^+ \tag{80}$$

The conditions that the vielbein rotation accomplishing the transition to the unitary gauge must satisfy are as follows:

(1) The definitions of the physical gauge bosons must remain invariant. For the purely left-handed vierbein rotations this is certainly true.

(2) The transformation must diagonalize the mass matrix and make it real.

The right/left multiplication of M_R^+/M_L^+ by the unitary matrix

$$U = S^A\sigma_A/\surd(S^AS_A)^{1/2} \tag{81}$$

indeed accomplishes this task. This transformation casts the mass matrix in the form

$$\begin{aligned} M_R^+ &= -M_L^- = (S^AS_A)^{1/2}\begin{pmatrix} 1+l & 0 \\ 0 & 1-l \end{pmatrix}, \\ M_L^+ &= -M_R^- = (S^AS_A)^{1/2}\begin{pmatrix} 1-l & 0 \\ 0 & 1+l \end{pmatrix} \end{aligned} \tag{82}$$

It is clear that the presence of the CP-breaking term in the action is necessary if one wants nondegenerate masses for fermions inside $SU(2)_L$ multiplets. In order to obtain vanishing mass for neutrinos one must assume that the condition

$$l=1 \tag{83}$$

holds.

It is perhaps needless to emphasize the potential importance of this rather unexpected result; the proposed geometric Higgs mechanism promises to explain, not only the fermionic mass spectrum but also to give a first principle explanation for the phenomena of (1) the matter–antimatter asymmetry, made in principle possible by the *CP*-breaking term in the bare action (Yoshimura, 1978; Weinberg, 1979), (2) the *CP* breaking observed in the kaon–antikaon system (Cronin, 1981; Fitch, 1981).

A second important result is that one can regard the field quantities obtained by performing the transition to the unitary gauge as $SU(2)_L$ singlets obtained as local composites of the spinor or gauge field and of the Higgs field normalized to unity. This result follows directly from the property of the diagonalizing Higgs gauge transformation of being linear with respect to the normalized Higgs field.

This result suggests that the phenomena of $SU(2)_L$ confinement and symmetry breaking via the Higgs mechanism should be regarded as one and the same thing. Symmetry breaking occurs in the sense that when the Higgs field is nonvanishing the unitary gauge is uniquely defined and is thus dynamically distinguished from the other gauges. $SU(2)_L$ confinement occurs in the sense that the field quantities in the unitary gauge can be regarded as $SU(2)_L$ singlets. For example, the mass splittings inside the electroweak multiplets can be understood as a manifestation of the fact that the members of these multiplets are in fact gauge singlets not related by any symmetry transformation.

The idea of $SU(2)_L$ confinement, when taken seriously has rather deep consequences concerning the understanding of the dynamics of the theory. The experience with QCD suggests that length-scale-dependent $SU(2)_L$ gauge coupling must become infinite at some finite length scale so that the Weinberg angle, which measures the ratio of $U(1)$ and $SU(2)_L$ couplings, must vanish at the limit of long length scales. In the second paper of the series we will show that also the requirement "the bosonic effective action reproduces the Maxwell electrodynamics at long length scales" implies $SU(2)_L$ confinement in the sense that $SU(2)_L$ coupling becomes infinite at this limit and the bosonic effective action becomes a pure Maxwell action.

5. SYMMETRIES OF THE THEORY

In this section we shall consider the symmetries of the theory. In Section 5.1 we shall discuss the problem of realizing the action of the isometries on the spinors of *H*. In Section 5.2 we show that the isometries of CP_2 indeed deserve the interpretation as color symmetries. In Section 5.3 chiral invariance and discrete symmetries are considered. In Section 5.4 we discuss the basic particle phenomonology in TGD framework. In Section 5.5 we study the various dynamical scenarios possible in TGD framework.

5.1. The Representation of the Isometries of H

In this section we wish to show that the isometries of the space H 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; we regard the invariance of, say Yang-Mills action under isometries as obvious.

It is instructive to prove the infinitesimal representability of the isometries first in the general case, when H allows ordinary spinor structure. So, let

$$\delta h^k = \varepsilon j^k \tag{84}$$

be an infinitesimal isometry so that its infinitesimal generator j^k satisfies the well-known Killing identities (Bjorken and Dreil, 1965)

$$D_k j_l + D_l j_k = 0 \tag{85}$$

The following theorem holds true:

Theorem. The quantity

$$L = \bar{\psi}\Gamma^\alpha \mathbf{D}_\alpha \psi \tag{86a}$$

is invariant under under the infinitesimal isometries of H realized according to

$$\delta\psi = i\varepsilon(\partial_l j_k \Sigma^{kl}/2 + j^k V_k) \equiv i\varepsilon X\psi \tag{86b}$$

Proof. The metric of X^4 (or any submanifold of H) is invariant under the action of the isometries and therefore it suffices to consider the term

$$L_{\alpha\beta} = \bar{\psi}\Gamma_\alpha \mathbf{D}_\beta \psi \tag{87}$$

The change of this quantity can be written as

$$\delta L_{\alpha\beta} = \bar{\psi}(K_\alpha \mathbf{D}_\beta + \Gamma_\alpha L_\beta)\psi \tag{88}$$

where the quantities K and L are defined as

$$\begin{aligned} K_\alpha &= \delta\Gamma_\alpha + \varepsilon[\Gamma_\alpha, X] \\ L_\alpha &= \delta V_\alpha + \varepsilon\partial_\alpha X + \varepsilon[V_\alpha, X] \end{aligned} \tag{89}$$

Obviously the conditions $K = 0$ 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 induced spinor connection and of the induced gamma matrices

$$\begin{aligned} &\partial_r \Gamma_k j^r + [\Gamma_k, X] + \Gamma_r \partial_k j^r = 0 \\ &\partial_r V_k j^r + V_r \partial_k j^r + \partial_k X + [V_k, X] = 0 \end{aligned} \tag{90}$$

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

$$D_m D_n j_k - D_n D_m j_k = R^s_{kmn} j_s \tag{91}$$

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

$$\Sigma_c j^r R_{rkmn} = 0 \tag{92}$$

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 (Adler et al., 1975); thus also the equation (90) is true.

Next we turn to the case of $M^4 \times CP_2$. Clearly the presence of the Kahler potential introduces an additonal term to δL given by

$$\delta L^1 = i\varepsilon\bar{\psi}\Gamma^\alpha Y_k \partial_\alpha s^k (n_+ 1_+ + n_- 1_-) + g.c. \tag{93}$$

where the quantity Y_k is defined as

$$Y_k = \partial_r B_k j^r + B_r \partial_k j^r/2 \tag{94}$$

Provided the quantity Y_k is expressible as a gradient of some quantity, say Z, we can indeed compensate L^1 by an infinitesimal $U(1)$ gauge transformation performed for the spinor field

$$\delta_1 \psi = -i\varepsilon Z(n_+ 1_+ + n_- 1_-)\psi \tag{95}$$

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 Kahler form J; i.e., for an infinitesimal isometry j^k there exists a Hamiltonian H so that the equation

$$j^k = J^{kl} H_{1l} \tag{96}$$

holds true; as a consequence the quantity Y_k is expressible as a gradient of the quantity

$$Z = j^r B_r/2 + H \tag{97}$$

Proof. The existence of H satisfying equation (96) follows from the fact that J is invariant under the isometries and closed as a two-form.

To see this observe that (96) is equivalent to the integrability condition

$$\partial_r(J_{mn} j^n) - \partial_m(J_{rn} j^n) = 0 \tag{98}$$

Using the closedness property of J

$$\Sigma_c \partial_m J_{nr} = 0 \tag{99}$$

one can transform the equation

$$\partial_m J_{rs} j^m + J_{ms} \partial_r j^m + J_{rm} \partial_s j^m = 0 \tag{100}$$

expressing the infinitesimal invariance of J under the isometries to the equation (98) and thus the existence of the Hamiltonian is shown. The representation of the quantity Y_k as a gradient is obtained by a direct calculation expressing J^k in terms of H.

Some comments concerning the results just derived are in order. First, one might argue that the Hamiltonian associated with the infinitesimal isometry is determined only apart from an arbitrary constant. This is not however true; with respect to the Poisson bracket the Hamiltonians associated withe the various infinitesimal generators form a Lie algebra, which must be isomorphic to the Lie algebra of $SU(3)$; this condition fixes various Hamiltonians uniquely.

Second, we have earlier stated (Pitkänen, 1983) that the isometries are not represented in a conventional manner and that the commutator of two infinitesimal isometries is given by the formula

$$[J_1, J_2] = J_{[1,2]} + j_1^k j_2^l F_{kl} \tag{101}$$

containing an additional term proportional to the curvature form of the spinor connection, which leads to a deformation of the original Lie-algebra structure.

We have, however, recently found that this statement is erroneous: a careful consideration of the various terms appearing in the commutator has shown that the deformation term is absent so that the representation of the isometries is an ordinary group representation.

5.2. Color Symmetries

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. The action of $SU(3)$ on the coordinate variables can be deduced from its action on the complex coordinates z^k, $k = 1, 2, 3$ of $\mathbb{C}^3$ (ordinary matrix multiplication). The action of $SU(3)$ is clearly nonlinear; the maximal linearly realized subgroup is $SU(2) \times U(1)$ representable as matrix multiplication with matrices of the form

$$\begin{pmatrix} U_{2\times 2} & 0 \\ 0 & 1/\det(U) \end{pmatrix} \tag{102}$$

Observe that the center of $SU(3)$ is represented trivially so that the ordinary "spherical harmonics" of CP_2 must have triality zero.

The infinitesimal action of the subgroup $SU(2)\times U(1)$ on the components of the spinor connection is found directly from the representation of the vielbein in complex coordinates (Eguchi et al., 1980). The spinor connection is invariant under $SU(2)$ and under the group $U(1)$ the components of the connection transform as objects having an "anomalous" color hypercharge

$$Y_A = 2Q_{\text{em}} \tag{103}$$

The action of the $SU(3)$ isometries on the spinor fields is found by noticing that the tangent space part of the infinitesimal isometry j^A is proportional to the quantity

$$X = [j^A, e^B]^k e^C_k \Sigma_{BC} \tag{104a}$$

The bracket represents the commutator of the vielbein vector field and the infinitesimal generator of the isometry defined as

$$[j^A, e^B]^k = j^{Ar}\partial_r e^{Bk} - e^{Br}\partial_r j^{Ak} \tag{104b}$$

Consider first the action of the isotropy group H_0 of the point $(\xi^1 \xi^2) = (0, 0)$ isomorphic to $SU(2)\times U(1)$. Since vielbein vectors are invariant under transformations of $SU(2)$ the action of this group is trivial; in fact any $SU(2)$ subgroup being related by a conjugation to the group H_0 acts trivially on spinors.

The action of the $U(1)$ subgroup is nontrivial and induces a rotation in the subspace spanned by the vielbein vectors e^1 and e^2. Thus the vielbein rotation is proportional to the generator Σ_{12} and is diagonal and same for both CP_2 chiralities. By a direct calculation one finds that also the spinors carry an anomalous hypercharge $Y_A = 2Q_{\text{em}}$.

The action of the remaining isometries on the right-hand (in CP_2) spinors is also diagonal. The vanishing of the part mixing the components of the spinor field is easy to verify by using the complex coordinate representation of the vielbein vectors (e^1 and e^2 (e^0 and e^3) are real and imaginary parts of the same complex vector).

The hypercharge changing infinitesimal isometries mix the components of the left-handed (in CP_2) spinors in the gauge used and one might argue that gluons corresponding to these infinitesimal generators are electromagnetically charged. We don't believe that this is the case. The point is that one can find a gauge where the action of all $SU(3)$ isometries is diagonal as we wish to demonstrate now.

Since CP_2 is a coset space $(SU(3)/SU(2)\times U(1))$ one can associate with each point s of CP_2 an element of $SU(3)$ defined to have the property that it carries some reference point s_0 (say the origin in complex coordinates) to the point in question. This group element is not unique; given a choice $g(s)$ any other choice related to the original one by the equation

$$g_1(s) = h_s g(s) h_{s_0}, \qquad h_s \in H_s, \qquad h_{s_0} \in H_{s_0} \tag{105}$$

is equally good. Since $SU(3) \to CP_2$ is nontrivial as bundle, the map $s \to g(s)$ cannot be made continuous in whole CP_2.

The gauge transformation in question is defined by the action of the group element $g(s)$ to the spinor field. If the spinors are eigenstates of the $U(1)$ generator in the original gauge the different choices of the map $s \mapsto g(s)$ induce a mere phase transformation of the spinor field. In a similar manner, the gauge transformation associated with a given $SU(3)$ isometry reduces to a mere phase transformation.

Clearly, this representation of color isometries is analogous to the massless representations of Poincaré group in the sense that anomalous hypercharge and thus the electromagnetic charge corresponds to the helicity of the massless particle.

Concerning the physical interpretation of the results obtained there are two alternative scenarios depending on whether the equation $Y_A = 2Q_{\mathrm{em}}$ is assumed to be true

(1) completely generally so that physical particles cannot correspond to ordinary irreducible representations of the color group (which are electromagnetically neutral).

(2) only for the spinor fields and gauge potentials and physical particles are assumed to correspond to ordinary irreducible representations of the color group.

We have believed the scenario 2 to be the correct one hitherto but this belief need not be correct as we will find soon.

In the first scenario only the leptonic spinors need to be introduced as elementary fermion fields since quarks can be identified as leptons moving in pseudotriplet color partial wave carrying anomalous hypercharge $-2/3$. In the scenario 2 one must assume that quarks and possibly also leptons are elementary fermion fields in order to obtain fractionally charged states.

Common ingredients for the both scenarios are the pseudotriplet $\underset{\sim}{3} = (f^k, k = 1, 2, 3)$ defined as

$$f^k = \xi^k/(1 + r^2)^{1/2} (\xi^3 \equiv 1) \tag{106}$$

and its complex conjugate $\bar{3} = \{\bar{f}^k\}$.

Using the representation of the pseudotriplet in $\mathcal{C}^3$ coordinates

$$f^k = z^k(\bar{z}^3/|z^3|)/(\Sigma_k z_k \bar{z}_k)^{1/2} \tag{107a}$$

one finds that pseudotriplet transforms as an ordinary $SU(3)$ triplet apart from the phase factor U defined as

$$U = \bar{z}^3/|z^3| \tag{107b}$$

This phase factor carries an anomalous hypercharge $Y_A = -2/3$.

In the scenario 1 quarks correspond to leptons moving in pseudotriplet partial wave in the sense to be described more precisely later. The fractional charges $+(-)(2/3, -1/3)$ for quarks (antiquarks) result by assuming that antileptons (leptons) can move only in pseudotriplet (conjugate pseudotriplet) partial waves.

Of course, one can wonder why antileptons move only in the triality 1 pseudotriplet but not in its complex conjugate carrying charges $(4/3, 1/3)$ instead of the conventional quark charges. We believe that the explanation is related to the nontriviality of the spinor bundle structure. The requirement that spinor partial waves are sections in the spinor bundle excludes the conjugate triplet.

In fact, in the case of S^2 with $U(1)$ bundle structure corresponding to a magnetic charge $n = 1$ this is what occurs. The pseudodoublet $(z, 1)/(1 + z\bar{z})^{1/2}$ corresponds to a section in this bundle but not its complex conjugate. The reason is that the rigorous definition of the spinor section requires the introduction of two coordinate batches (upper and lower hemisphere). The pseudodoublet represents the spinor section only in the upper hemisphere and its representation in the lower hemisphere is obtained by multiplying with the phasefactor $\bar{z}/|\bar{z}|$. Complex conjugate doublet is not a section, since it corresponds to a spinor field, which is infinitely manyvalued on the south pole of S^2.

Consider next the probably mathematically inconsistent scenario II based on the assumption that the connection between anomalous hypercharge and electromagnetic charge holds true only for the spinor fields and gauge fields.

One of the basic assumptions of the quantum theory is that physical states must correspond to irreducible representations of the symmetry group of the theory. Since the induced spinor field is not $SU(3)$ singlet but carries anomalous hypercharge, this requirement implies that one must multiply the spinor field by suitable functions of the coordinate variables in order to obtain an irreducible representation of $SU(3)$.

In order to cancel the anomalous hypercharges one can use besides the pseudotriplets $\underline{3}$ and $\bar{\underline{3}}$ the phase factor V defined as

$$V = \xi^1\xi^2/|\xi^1\xi^2| \tag{108}$$

This phase factor carries an anomalous hypercharge $Y_A = -2$.

The spinorial spherical harmonics are constructed using the following recipe:

(1) Perform a gauge transformation g given by

$$g = (I_L^3 + I_R^3 + Id/2)V \tag{109}$$

This gauge transformation cancels the anomalous hypercharge associated with gauge bosons and charged leptons. Clearly, one can construct leptonic partial waves by multiplying leptonic spinors with definite electromagnetic charge with ordinary triality zero CP_2 partial waves.

(2) When the spinor is quark (antiquark) like the remaining anomalous charge ($= -2/3$ for U and D type spinors) is cancelled by multiplying the complex conjugate of the pseudotriplet so that the total anomalous hypercharge is cancelled.

We have only recently realized that the scenario 2 is probably not mathematically consistent. The weak point is the use of the phase factor V to gauge away the anomalous hypercharge associated with leptons and gauge bosons; this gauge transformation is not single valued at the origin of CP_2! As a consequence the leptonic spinors become infinitely many valued at the origin of CP_2 in the new gauge. This result suggests that the delicacies of the spinor bundle structure prevent the construction of the ordinary irreducible representations.

We conclude this section with an argument, that gives additional support to idea that the connection between electromagnetic charge and anomalous hypercharge is completely general. The argument is based on the representation of the spinor field in terms of the spherical harmonics of the space H. The study of the covariant derivative of the spinor field partial wave expansion shows that the ordinary derivatives of the CP_2 partial waves can be interpreted in terms of the ordinary gluonic couplings when the CP_2 partial waves in question correspond to ordinary irreducible color representations. In particular, the electromagnetic charge doesn't depend on the ordinary hypercharge of the particle.

In case of the pseudotriplet one can divide the gluonic coupling into two parts: the ordinary gluonic coupling plus an anomalous term. The anomalous term reduces to a coupling to Kahler potential with unit coupling strength. Thus one can say that the pseudotriplet differs from the ordinary triplet in that it has an additional electromagnetic charge $Q_{em} = Y_A/2$.

To begin with consider the definition of the gluonic gauge potentials. We have already found that the projections of the isometry generators j^A behave as gluon fields. Indeed these quantities are gauge potentials since they define a covariant derivative with respect to which the Hamiltonians of the CP_2 isometries are covariantly constant.

$$\partial_k H^B + ij_k^A T_{A)C}^B H^C = 0 \tag{110}$$

Here the quantities T_{AjC}^B are the representation matrices of the Lie-algebra generators in the adjoint representation.

The identities leading to the result are the following one

$$j^{A;l}_{k} j^{l}_{A} = s^{l}_{k} \tag{111a}$$

$$j^{k}_{A}\partial_k H^B = iT^{B}_{A)C} H^C \tag{111b}$$

Next we introduce the partial wave expansion for the spinor field using a complete spinor basis of the space H.

$$\psi(x) = \Sigma_n c_n(x)\psi^n(h(x)) \tag{112a}$$

The spinor field is understood as a general solution of the Dirac equation as a functional of the surface X^4. The partial waves can be written as tensor products of M^4 and CP_2 spinor partial waves

$$\psi^n(h) = f_k(m) \otimes Y_l(s) \tag{112b}$$

Consider now the covariant derivative of the spinor field represented in terms of the partial waves. We shall first show that when the partial waves correspond to an ordinary irreducible $SU(3)$ representation the ordinary derivatives of the CP_2 partial waves can be transformed to ordinary gluon coupling terms. Using the identities (111) and the definition of the gluonic field the derivative of the CP_2 partial wave can be cast to the form

$$\partial_k Y_l = ij^{A;r}_{k} j^{r}_{A}\partial_r Y_l \tag{113}$$

For the ordinary irreducible representations the term involving the action of the infinitesimal generator can be written in terms of the representation matrices of the Lie-algebra generators and one obtains just the ordinary gauge coupling term. Thus the color interactions are gauge interactions and gluons are neutral since the electromagnetic coupling doesn't depend on the ordinary hypercharge.

Next we treat the case of the pseudotriplet partial waves. Since this triplet differs from the ordinary color triplet only by a phase factor transforming as the quantity $U = \bar{z}^3/|z^3|$ it is clear that the action of the infinitesimal generator on the pseudotriplet can be divided into two parts. The first part is simply the linear action giving rise to an ordinary color coupling between gluons and color triplet and the second term can be written in the form

$$X_M Y_l = ig^A \delta_A U U^{-1} Y_l \tag{114}$$

The quantity $\delta_A U$ is simply the derivative of the phase factor U with respect to the infinitesimal isometry j^A interpreted as an isometry of $\mathbb{C}^3$.

What is important is that the quantity $\partial_A U U^{-1}$ and thus also X_M is function of CP_2 coordinates only.

By a direct calculation using the explicit expressions of the infinitesimal generators one finds that this term equals to the coupling of the Kahler potential to the spinor field and the coupling strength equals to unity. Recalling how the coupling to the Kahler potential controls the values of the electromagnetic coupling one can conclude that the equation $Q_{em} = Y_A/2$ applies completely generally.

The above reasoning leads to a concrete definition of the spinor partial wave concept and shows that gluons indeed exist and mediate gauge interactions. In addition, it strongly suggests that the connection $Y_A = 2Q_{em}$ holds completely generally so that the scenario 1 in which quarks are leptons in pseudotriplet partial waves is the correct one. Probably this scenario is also the simplest scenario one can imagine. In the second paper of the series we shall find that this scenario predicts for the bare Weinberg angle θ_w the value $\sin^2\theta_w = 1/4$, which is remarkably close to the experimental value ($\cong 0.23$). In this scenario one must however give up the idea that particles correspond to ordinary irreducible representations of the color group and it is not probably wise to draw any final conclusions at this stage.

5.3. Generalized Chiral Invariance and Discrete Symmetries

The generalized chiral invariance is necessary if one wants two conserved fermion numbers (or to pose chirality condition on spinors). The action of the chiral transformation on the spinor field is given by

$$\psi \to \exp(i\alpha X)\psi \tag{115}$$

Here the quantity X is a linear combination of the matrices 1 and Γ_9.

The most general spinorial action invariant under the chiral transformations is of the form given by equation (64). The generalized chiral invariance together with the assumption that both leptons and quarks are elementary fermions implies the absolute stability of the proton against spontaneous decay provided the quarks are massive enough.

We base our treatment of discrete symmetries C, P, and T on the following requirements. First, the symmetries must be realized as purely geometric transformations. Second, the transformation properties of the field variables should be essentially the same as in the conventional field theories (Bjorken and Drell, 1965). Finally, the assumption about the Grassmann valuedness of the spinors is made.

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

$$\begin{aligned} m^k &\to P(m^k) \\ \psi &\to \gamma^0 \times \gamma^0 \psi \end{aligned} \tag{116}$$

in the representation chosen for the gamma matrices. Indeed, the gauge bosons $W^{\pm}$ and Z^0 break the symmetry because they do not commute with the matrix $\gamma^0 \times \gamma^0$.

The guess that a complex conjugation in CP_2 is associated with the T transformation of the physicist turns out to be correct. One can verify by a direct calculation that the pure Dirac action ($l=0$) is invariant under T realized according to

$$\begin{aligned} m^k &\to T(m^k) \\ \xi^k &\to \bar{\xi}^k \\ \psi &\to \gamma^1 \gamma^3 \times 1 \psi \end{aligned} \tag{117}$$

When the parameter l is nonvanishing the action is not invariant under T; the reason is that the term involving the Kahler form is T odd. The T oddness follows from the fact that the Kahler form changes sign under the complex conjugation in CP_2 unlike the metric.

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

$$\begin{aligned} \xi^k &\to \bar{\xi}^k \\ \psi &\to \gamma^2 \times \gamma^0 \psi^+ \end{aligned} \tag{118}$$

From the results just derived we can conclude that the discrete symmetries CP and T are exact symmetries of the theory defined by pure Dirac action. When the parameter l is nonvanishing the part of the action involving the Kahler form is T and CP odd so that only CPT is an exact symmetry of the action. It is rather encouraging that the Kahler structure makes possible a CP breaking action. It is even more encouraging that the requirement "Neutrinos are massless" leads to a unique CP breaking term.

5.4. Particle Phenomenology and TGD

One of the main virtues of the TGD approach is that it provides a possibility to geometrize particle spectroscopy. The particle classification in CP_2 framework is based on the use of baryon and lepton numbers; electroweak quantum numbers, color quantum numbers and topological quantum numbers g (genus of the boundary component) and h (homology equivalence class of the boundary component).

Since there are no symmetries transforming the various particle generations to each other they are expected to behave identically with respect to the basic interactions; this prediction is in accordance with the experimental facts.

We think that the approximate flavor and even spin flavor symmetry is a combinatorial artefact, which can be understood by assuming that hadrons are describable as 3-quark states to a good approximation at low energies. The complete symmetry of the 3-particle statefunction with respect to one particle labels (g, I, spin) implies that the states can be formally regarded as a basis for an irreducible representation of an appropriate spin flavor group.

Of course, this group has a purely formal meaning; it is the transformation properties of the individual states under permutations of the one particle labels, which leads to the illusion about the presence of a real symmetry.

The proposed explanation for the family replication phenomenon besides being in concordance with the idea of the universality, also explains the basic phenomenological rules characterizing electroweak interactions of hadrons.

Consider first the well-known $\Delta I = 1/2$ rule observed in the decay of the strange particles. The rule is understood once it is realized that the notion of the strong isospin is artificial and instead of it one should label the hadronic states using the weak isospin having a genuine group theoretical meaning. In general the multiplet assignments obtained using I_s and I_w differ, but in such a way that the above-mentioned nonconservation law 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 $K \to \pi^+\pi^-$ the final state with $I_w = 2$ is suppressed relative to that with $I_s = I_w = 0$.

Cabibbo mixing is another peculiar phenomenon, where both the weak and the strong interactions are involved (Kobayashi and Maskawa, 1973). TGD approach predicts a topological transition changing the genus of the boundary component by one unit. Therefore we expect mixing between different particle generations to take place and thus to be observable in weak transitions mediated by the charged bosons $W^{\pm}$.

The mixings can be described phenomenologically using unitary matrices U and D defined as

$$\begin{aligned} U_i &\to U_i^j U_j = \tilde{U}_i \\ D_i &\to D_i^j D_j = \tilde{D}_i \end{aligned} \tag{119}$$

Since the emission of the weak boson does not change the boundary topology (we neglect the possible mixing between possible bosonic generations), the amplitude for the process $\tilde{U}_i \to \tilde{D}_j$ must be proportional to the quantity

$$U_i^k \bar{D}_{jk} \tag{120}$$

which is nonvanishing unless the mixings suffered by U and D differ only by a phase multiplication performed for the "initial" and "final" states.

The absence of generation changing neutral currents is also a peculiar phenomenon having a nice explanation as a manifestation of the $\Delta g = 0$ nature of the weak boson emission; the corresponding matrix element $\tilde{U}_i \to \tilde{U}_j(\tilde{D}_i \to \tilde{D}_j)$ is proportional to the quantity $U_i^k \bar{U}_{jk}(D_i^k \bar{D}_k)$ and thus vanishes.

5.5. Possible Dynamical Scenarios

When discussing the representation of the color isometries we found the emergence of two profoundly different physical scenarios.

(1) The equation $Y_A = 2Q_{\rm em}$ holds completely generally and quarks correspond to leptons moving in pseudotriplet partial waves of CP_2. Physical particles do not correspond to ordinary irreducible representations of the color group.

(2) The equation $Y_A = 2Q_{\rm em}$ is true only for spinor fields and gauge potentials. Physical states correspond to ordinary irreducible representations of the color group with no anomalous hypercharge.

We found that the first scenario is probably the correct one but that at this stage it is impossible to make any final conclusions.

The nice property of the first scenario is that it leads to an essentially unique theory in which leptonic spinors are the only elementary spinor fields.

The second scenario allows several variants depending on whether both leptons and quarks are considered to be elementary particles or not. In the most simple minded scenario leptons and quarks correspond to different chiralities of H-spinors and baryon and leptons numbers are separately conserved; in this respect the scenario differs from the typical unification scenarios, say GUT's (Georgi and Glashow, 1974; Fritch and Minkosk, 1975).

The prediction of two conserved fermion numbers is however not a necessary outcome of this scenario since one can imagine, in some respects simpler variants of the scenario 2, in which the quarks are the only elementary fermions.

The first scenario of this kind is roughly the following:

(1) Leptons are assumed to correspond to 3-manifolds with boundary and carrying quark number $N = -3$ so that leptons are 3-quark composites of type $L = \bar{q}\bar{q}\bar{q}$. This assumption has the highly desirable consequence that the topological explanation of the family replication phenomenon need not be given up.

It is not necessary to assume that quark number resides on the boundary; leptons (quarks) might well be analogous to many quark states constructed using quark field operators defined in the interior of the leptonic (quark) 3-manifold. If elementary particles are modelled as little pieces of Minkowski space the presence of the boundary component is a necessity.

(2) Quarks move in color triplet partial waves and leptonic states are color singlets and thus completely antisymmetric under the permutation of the color labels. In this respect leptons resemble antibaryons.

(3) The part of the leptonic state function, which corresponds to spin and electroweak spin transforms according to the direct product representation

$$(2\times 2)\otimes(2\times 2)\otimes(2\times 2)=(4\oplus 2\oplus 2)\times(4\oplus 2\oplus 2) \tag{121}$$

of $SU(2)\times SU(2)_L$, where the first factor refers to the rotation group. This scenario predicts a rich structure inside each leptonic family. There should exist four leptonic states with essentially identical electroweak properties inside each family besides the excitations having spin or (and) weak isospin equal to 3/2.

If the naive expectation that leptons correspond to 3-manifolds of size $L \ll 1/m_L$ is true, one expects that the description of the leptons as purely point like objects should be a good approximation and the corrections resulting from the compositeness to the leptonic magnetic moments should be very small. In fact, a naive dimensional argument based on the assumption that the vacuum expectation of Higgs field determines the mass of the lepton, leads to the order of magnitude estimate

$$m_L \cong R/L^2 \tag{122}$$

Thus the size of lepton is roughly given by the geometric mean of Planck length and leptonic Compton length and is roughly of the order of intermediate boson Compton length for electron.

In the second scenario leptons correspond to local composites of type $q(x)q(x)q(x)$ formed from quarklike spinors. Fermi statistics requires that the spin-weak isospin part of the leptonic state functional corresponds to the completely antisymmetric part of the product of three 2×2's (color quantum numbers are not spinlike!)

$$((2\times 2)\otimes(2\times 2)\otimes(2\times 2))_A=2\times 2 \tag{123}$$

This kind of scenario results naturally if both quarks and leptons correspond to components of a superfield (Gelfand and Likhtman, 1971; Volkov and Akulov, 1973; Wess and Zumino, 1974a, b)

$$\Omega(x)=\Sigma\Omega^{\beta_1\cdots\beta_m}_{\alpha_1\cdots\alpha_n}(x)\theta_{\beta 1}\cdots\theta_{\beta m}\bar{\theta}^{\alpha_1}\cdots\bar{\theta}^{\alpha_n} \tag{124}$$

where the theta parameters are anticommuting numbers. Quarks would correspond to the term $\sim\theta$ and leptons to the term $\bar{\theta}\bar{\theta}\bar{\theta}$ in this expansion. It should be noticed that the term $\bar{\theta}\theta\bar{\theta}\bar{\theta}\bar{\theta}$ also gives rise to a particle with the leptonic quantum numbers. No excited leptons with nonconventional quantum number assignments are predicted.

Baryons are expected to be unstable against spontaneous decay in all scenarios satisfying the chirality condition. An estimate for the decay rate baryon → lepton $+X$ is obtained by imagining hadron as a ball of radius of the order of a typical hadronic length and quark as a ball of radius R_q placed on the hadronic "bag." The rate of decay is obtained by the following line of reasoning.

The rate for two quark collisions is proportional to the quark density and to the area of the quark sphere

$$r_2 \sim R_q^2/R_h^3 \tag{125}$$

The rate for 3-quark collisions is in turn proportional to the rate for two quark collisions and the probability that a given quark happens to be inside a given ball of radius R_1

$$r_3 = (R_q^2/R_h^3) * (R_q/R_h)^3 \tag{126}$$

By requiring that the lifetime of proton is larger than 10^{30} years one obtains the lower bound

$$R_q/R_h < 10^{-15} \tag{127}$$

Thus quarks should have a rather small size.

Although the question "Which of the proposed scenarios is closer to the truth" cannot be settled at this stage, one could argue that the scenario with minimal number of elementary fermions (either leptons or quarks) is probably the correct one. Whether the Nature has followed this kind of argumentation is of course a question to be solved experimentally. In fact, some indications about the existence of excited leptons exists (Collab, 1983b). Our personal opinion is that the scenario assuming only leptonic spinors as fundamental field variables is probably the correct one.

ACKNOWLEDGMENTS

I wish to thank all those people who have helped me to carry out this work. I am indebted to Prof. D. Finkelstein for criticism and encouragement at the early stage of this work. Also I gratefully ackowledge the thorough criticism of Prof. J. A. Wheeler concerning the basic ideas of TGD approach. My sincere thanks are due to Prof. H. Römer for interesting and clarifying discussions. Also I thank Prof. C. Cronström, Doc. R. Keskinen, and Prof. J. Mickelson for encouragement and help as well as for enlightening discussions. Also I am very much obliged to Dr. T. Perko and Doc. A. Kupiainen for useful conversations.

REFERENCES

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

Adler, Basin, and Schiffer, *Introduction to General Relativity*. McGraw-Hill, New York.

Anderson, B., Gustafson, C., Ingelman, G., and Sjostrand, T. (1983). *Physics Reports*, **97**.

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

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

Bjorken and Drell, (1965). *Relativistic Quantum Fields*. McGraw-Hill, New York.

Brander, M. (1981). *Physics Reports*, **75**, 4.

Chew, G., and Rosenzweig, C. (1978). *Physics Reports*, **41C**.

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

Croñin, J. W. (1981). *Reviews of Modern Physics*, **53**.

Denisov, V. I., and Mestrishvili, M. A. (1981). *Soviet Journal Particles and Nuclei*, **12**, 1.

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

Eisenhart, (1964). *Riemannian Geometry*. Princeton University Press, Princeton, New Jersey.

Faddeev, L. D. (1982). *Soviet Physics Uspekhi*, **25**(3).

Fitch, V. L. (1981). *Reviews of Modern Physics*, **53**.

Fritch, H., and Minkowski, P. (1981). *Physics Reports*, **73**, 2.

Fritsch, H., and Minkoski, P. (1975). *Annals of Physics (New York)*, **193**.

Gelfand, Y. A., and Likhtman, E. P. (1971). *J.E.T.P. Letters*, **13**, 323.

Georgi, H. (1975). *Particles and Fields*-1974, C. E. Carlson, ed. A.I.P., New York.

Georgi, H., and Glashow, S. L. (1974). *Physical Review Letters*, **32**, 430.

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

Glashow, S., (1961). *Nuclear Physics*, **22**, 579.

Hawking, S. W., and Pope, C. N. (1978). *Physics Letters*, **73B**, 1, 42.

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

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

Husemoller, (1960). *Fibre Bundles*. McGraw-Hill, New York.

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

Kaluza, Th. (1921). *Sitzungsberichte der Preussischen Akademie der Wissenschaften Berlin, Mathematische-Physikalisch Klasse*.

Klein, O. (1926). *Zeitschrift für Physik*, **37**, 895.

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

Madore, J. (1981). *Physics Reports*, **75**(3).

Misner, Thorne, and Wheeler, *Gravitation*. W. H. Freeman and Company, San Francisco, 1975.

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

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

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

Pitkänen, M. (1983). *International Journal of Theoretical Physics*, **22**, 575.

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. Almqwist and Wicksells, Stockholm.

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

Susskind, L., and Kogut, J. (1976). *Physics Reports*, **23**, 348.

Thom, R. (1954). *Commentarii Math. Helvet.*, **28**, 17.

UA1 Collab., G. Arnison et al., *Physics Letters*, **126B** (1983a); *Physics Letters*, **135B** (1984).
UA2 Collab., P. Backala et al., *Physics Letters*, **124B** (1983b).
Wallace, (1968). *Differential Topology*, Chaps. 6 and 7. W. A. Benjamin, New York.
Weinberg, S. (1967). *Physical Review Letters*, **19**, 1264.
Weinberg, S. (1979). *Physics Review Letters*, **42**, 850.
Volkov, D. V., and Akulov, V. P. (1973). *Physics Letters*, **46B**, 109.
Wess, J., and Zumino, B. (1974a). *Nuclear Physics*, **B70**, 39; Wess, J., and Zumino, B. (1974b); *Nuclear Physics*, **B78**, 1.
Witten, E. (1980). Princeton University preprint.
Yoshimura, M. (1978). *Physics Review Letters* **41**, 281; *Phys. Rev. Lett.* **42**, 746(E) (1979).