Renormalizability aspects of massive Yang-Mills field models

C. N. Ktorides*
Nuclear Research Center "Demokritos," Athens, Greece
(Received 10 June 1975)

We confront the problem concerning the renormalizability of massive Yang-Mills theories in which the mass term for the vector fields has been inserted by hand. Explicitly, our starting Lagrangians are of the type studied in the past by Veltman and Boulware and found to be nonrenormalizable. We rely heavily on Boulware's analysis in which the basic point of view is to split the massive Yang-Mills fields into transverse and longitudinal components. The latter carry all the nonrenormalizability pathologies which manifest themselves in terms of certain nonpolynomial factors involving the longitudinal fields. The fact that these factors cannot be removed via a redefinition of the longitudinal fields leads to the conclusion of nonrenormalizability. We call the problem on hand, namely the removal of the bad nonpolynomial terms, Boulware's problem. We study this problem closely, within the context of the adjoint representation of the gauge group [we restrict ourselves to SU(2) for the most part] employing the language of differential geometry. We prove a theorem according to which a necessary condition for solving Boulware's problem is the introduction of extra fields. In the case of SU(2) we find an explicit solution which requires the introduction of a triplet of scalar fields belonging to the adjoint representation of SU(2). We interpret the additional fields as ghost, or superfluous fields—most probably corresponding to the ghost fields of spontaneously broken gauge theories in the R gauge. Here we note a basic difference between our program and that of Cornwall *et al.* First of all, our interpretation of the fields which combine with the longitudinal ones in order to remove the nonpolynomial factors as ghost fields is not evident in the treatment of Cornwall *et al.* Finally, unlike the case in Cornwall *et al.*, we do not just show the existence of the transformation which removes the undesirable terms but also give the explicit conditions which bring about this result in the case of SU(2). A proposition relating the models under consideration to spontaneously broken gauge ones is also presented. We argue, without explicit proof, that the combination of this proposition with our main theorem corresponds to building a spontaneously broken gauge theory in the R gauge, having started from a non-Abelian theory with mass inserted by hand.

I. INTRODUCTION

Explicit renormalization of Lagrangians containing massive Yang-Mills fields was first carried out by 't Hooft[1] and, subsequently, by Lee and Zinn-Justin.[2] The procedure followed in Refs. 1 and 2 to construct renormalizable models of the massive Yang-Mills type can be summarized by the following two observations: (a) A set of scalar fields belonging to a representation of the symmetry group is introduced into a manifestly gauge-invariant Yang-Mills Lagrangian.[3] (b) The mechanism of spontaneous symmetry breaking is employed in order to generate a mass term for the Yang-Mills fields.

In the present study we adopt a reverse point of view. Namely, we start with a massive Yang-Mills Lagrangian which is nonrenormalizable and systematically study those extra conditions which open the way to renormalizability. The models under study are of the following type: Every term in the Lagrangian is manifestly gauge-invariant except for the vector-field mass term. More explicitly, we shall concentrate on the kind of models studied in the past by Veltman[4] and by Boulware[5] and found by them to be nonrenormalizable. We shall show that, under a certain assumption, what is needed for opening the way to renormalizability is the introduction of additional, nonphysical scalar fields which, in the case of SU(2) at least, belong to the adjoint representation of the group.

To be specific, we prove a theorem to the following effect. Suppose we are given a massive non-Abelian theory of the type described above. A necessary condition for removing from such a theory certain nonpolynomial factors in the longitudinal fields, whose presence is responsible for the nonrenormalizability, is the introduction into the theory of additional fields. In the case of SU(2) we have an explicit solution in terms of a triplet of ghost scalar fields belonging to the adjoint representation.

Having dealt with renormalizability aspects we shall proceed to examine "gauge invariance" aspects of our massive Yang-Mills models. We shall show that certain conclusions arrived at in our recent study on the Abelian case[6] can readily be extended to the non-Abelian case. What we find can be summarized as follows. Suppose we add to a Yang-Mills model of the aforementioned type an appropriate number of physical scalar fields. Then, if we require a certain kind of "equivalence," to be explained in Sec. IV, with a manifestly gauge-invariant model this addition would force the latter to be of the spontaneously broken symmetry variety.

We shall interpret the simultaneous addition of the unphysical and the physical scalars mentioned above as building up a spontaneously broken gauge theory in the R gauge—having started with a theory in which the mass was inserted by hand. However, we shall not attempt to solidify this statement in the present study through explicit calculations.

At this point we must note several basic differences between our program and that of Cornwall *et al.*[7] These authors also confront the problem of removing nonpolynomial factors from their formalism. Tree unitarity, a requirement of gentle high-energy behavior of S-matrix elements in the tree approximation, guarantees the existence but does not give the explicit form of the point transformation which leads to the removal of the nonpolynomial factors. Our approach *does* lead to explicit conditions which bring about the desired result, in the case of SU(2) at least. Finally, our interpretation of the (extra) fields, which combine with the longitudinal fields to remove the nonpolynomial factors, as ghost fields is not evident in the treatment of Ref. 7. Our belief is that the present work and that of Ref. 7 are in many ways complementary. Each study emphasizes different aspects of converse routes to spontaneously broken gauge theories. In our case the preoccupation is with respect to those extra ingredients which pave the way to the construction of renormalizable models of the massive Yang-Mills type. The authors of Ref. 7, on the other hand, face head-on the problem of high-energy behavior of S-matrix elements (in the tree approximation), thereby responding more directly to the ultraviolet problems of massive Yang-Mills theories.

In Sec. II we review Boulware's work on massive Yang-Mills Lagrangians. Given the bad behavior of the massive vector field propagator $(g_{\mu\nu}-p_\mu p_\nu/m^2)(p^2-m^2)^{-1}$, it is important that we separate its transverse projection $(g_{\mu\nu}-p_\mu p_\nu/p^2)\times(p^2-m^2)^{-1}$ so that the renormalizability difficulties can be isolated. This splitting helps to isolate those factors which are responsible for the nonrenormalizability of the theory.

We concentrate on these factors in Sec. III. We characterize our difficulties as intimately connected with the fact that we are dealing with infinite-dimensional Lie groups and algebras. It becomes of central importance to study some aspects of the adjoint representation of such groups. This we do in a differential geometric context. We prove our main theorem which gives an explicit solution to Boulware's problem once we have introduced nonphysical scalar fields into our theory. These fields belong to the adjoint representation of SU(2).

In Sec. IV we deal with the symmetry aspects of our program. Our underlying point of view is the following. A prerequisite for a given massive non-Abelian vector-field model to be renormalizable in the true sense[8] is that it must possess a number of dynamical degrees of freedom equal to those of some manifestly gauge-invariant model. If this did not hold true, it would not be reasonable to expect that the symmetry, which the massless counterpart model manifestly exhibits, could be somehow buried inside the massive one. In this connection, we establish the previously mentioned relation between the massive Yang-Mills models we are studying and ones which exhibit spontaneous breaking of the local symmetry.

In Sec. V we give our interpretation of our findings in the preceding two sections. However, it must be noted that an explicit renormalization program which incorporates the features of Secs. III and IV must be carried out before our interpretations can be substantiated. This we intend to do elsewhere.

II. "NONRENORMALIZABLE" MASSIVE YANG-MILLS LAGRANGIAN MODELS

By definition, Yang-Mills fields are vector fields belonging to the adjoint representation of a local symmetry group. Accordingly, the massive Yang-Mills Lagrangians we want to study will have a large sector identical to that of a manifestly gauge-invariant theory. Following Veltman[4] and Boulware,[5] we consider Lagrangians which are gauge-invariant in all respects except for the vector-field mass term. The latter is inserted by hand. Prior to the works of Refs. 1 and 2, one hoped to show renormalizability for such models in analogy to familiar massive quantum-electrodynamical formulations in which the vector field is coupled to a conserved current. Unfortunately, the non-Abelian cases exhibited unmanageable difficulties which prevented the programs in Refs. 4 and 5 from succeeding.

The philosophy behind Boulware's general approach, which constitutes the main subject of this section, was the following. Given the propagator of a massive vector field with its bad power behavior, one should first try and split it into transverse and longitudinal parts. It is in the latter that all problems should reside because the propagator corresponding to the former has a well-behaved form. Once the renormalization difficulties have been isolated in this manner one concentrates one's attention on the longitudinal fields.

Before proceeding we need to establish our notation. We find it preferable to use a language which both saves us from a maze of indices and shows the intimate connection between Yang-Mills fields and the (local) algebra of the symmetry group. Thus, an SU(2) Yang-Mills field $\mathcal{B}_\mu(x)$ can be expressed as (we adopt the summation convention throughout)

$$\mathcal{B}_\mu(x)=B_\mu^a(x)t_a\,, \tag{2.1}$$

where the t_a, $a=1,2,3$, are thought of as the basis vectors of the SU(2) algebra and $B_\mu^a(x)$ are the components of $\mathcal{B}_\mu(x)$. It should be noted that it is sometimes advantageous, for the sake of clarity, to employ components instead of algebra elements.

The antisymmetric tensor $\mathcal{G}^{\mu\nu}$ as an algebra element has the form

$$\begin{aligned}\mathcal{G}_{\mu\nu}&=\partial_\mu\mathcal{B}_\nu-\partial_\nu\mathcal{B}_\mu-ig[\mathcal{B}_\mu,\mathcal{B}_\nu]\\&=G_{\mu\nu}^a t_a\,.\end{aligned} \tag{2.2}$$

We have the familiar component relation

$$G_{\mu\nu}^a=\partial_\mu B_\nu^a-\partial_\nu B_\mu^a-igt_{abc}B_\mu^b B_\nu^c\,, \tag{2.3}$$

where the t_{abc} are the structure constants of SU(2).

A gauge transformation on $\mathcal{B}_\mu(x)$ is given by

$$\begin{aligned}\mathcal{B}_\mu(x)\to\mathcal{B}'_\mu(x)=&S(x)\mathcal{B}_\mu(x)S^{-1}(x)\\&+\frac{1}{g}S(x)\partial_\mu S^{-1}(x)\,,\end{aligned} \tag{2.4}$$

where $S(x)$ belongs to the local SU(2) group (adjoint representation). As we shall discuss to some extent in the next section, $S(x)$ can be put in the form

$$S(x)=e^{\xi_a(x)t_a}\,, \tag{2.5}$$

where $\xi(x)\equiv\xi_a(x)t_a$ belongs to the local SU(2) algebra.

The gauge transformation (2.4) on $\mathcal{B}_\mu(x)$ implies that $\mathcal{G}^{\mu\nu}$ transforms under the adjoint action of the group, i.e.,

$$\mathcal{G}_{\mu\nu}(x)\to\mathcal{G}'_{\mu\nu}(x)=S(x)\mathcal{G}_{\mu\nu}(x)S^{-1}(x)\,. \tag{2.6}$$

In general, we shall admit into our models matter fields which belong to some (irreducible for simplicity) representation of the group. Suppose we form an array ψ from the matter fields. Then the gauge action on ψ will be given by

$$\psi(x)\to\psi'(x)=R(x)\psi(x)\,, \tag{2.7}$$

where $R(x)$ is the representation of $S(x)$ in the space of fields.

Let us assume that the basis $\{t_a\}$ chosen for the algebra is orthonormal. We write symbolically, i.e., avoiding to show explicitly that we are in essence calculating traces,

$$(t_a,t_b)=\delta_{ab}\,. \tag{2.8}$$

It follows that

$$(\mathcal{B}_\mu,\mathcal{B}_\mu)=B_\mu^a B_a^\mu\,. \tag{2.9}$$

For convenience, we shall write $\mathcal{B}_\mu\mathcal{B}^\mu$ in place of $(\mathcal{B}_\mu,\mathcal{B}_\mu)$.

With these conventions in mind, the simplest massive Yang-Mills field Lagrangian we can write and which subscribes to the specifications we have already given is

$$\mathcal{L}=-\tfrac{1}{4}\mathcal{G}_{\mu\nu}\mathcal{G}^{\mu\nu}-\tfrac{1}{2}m^2\mathcal{B}_\mu\mathcal{B}^\mu\,. \tag{2.10}$$

A more general version admits the presence of matter fields ψ belonging to an (irreducible) representation of SU(2). It has the form

$$\mathcal{L}=-\tfrac{1}{4}\mathcal{G}_{\mu\nu}\mathcal{G}^{\mu\nu}-\tfrac{1}{2}m^2\mathcal{B}_\mu\mathcal{B}^\mu+g\mathcal{B}^\mu j_\mu+\mathcal{L}_\psi\,, \tag{2.11}$$

where $\mathcal{L}_\psi$ contains purely matter field terms, consistent with gauge invariance, and j_μ is a conserved current made up of minimal-coupling interactions. The latter has an expansion in the SU(2) algebra

$$j_\mu=j_\mu^a t_a\,. \tag{2.12}$$

The Lagrangians given by (2.10) and (2.11) describe precisely the kind of theories we want to study. Henceforth, whenever we refer to "starting Lagrangians" or "models of interest" and so on, we shall mean theories of the type (2.10) and/or (2.11).

By their very nature, Lagrangians (2.10) and (2.11) remain invariant under gauge transformations (2.4) and (2.7) on the field variables in all but the mass term. For example, it can easily be shown[5] that, under a gauge transformation, (2.11) becomes

$$\begin{aligned}\mathcal{L}'=&-\tfrac{1}{4}\mathcal{G}_{\mu\nu}\mathcal{G}^{\mu\nu}+g\mathcal{B}^\mu j_\mu+\mathcal{L}_\psi\\&-\tfrac{1}{2}m^2\left[\mathcal{B}_\mu+\frac{1}{g}S^{-1}(x)\partial_\mu S(x)\right]^2.\end{aligned} \tag{2.13}$$

The following question is of immediate interest. Can (2.4) be utilized so that $\mathcal{B}_\mu(x)$ is replaced by a triplet of transverse vector fields $\mathcal{A}_\mu=A_\mu^a t_a$ plus a scalar field triplet $\xi=\xi^a t_a$ corresponding to the longitudinal degrees of freedom? This would be desirable because, as can be seen from (2.13), a large part of the resulting Lagrangian will involve transverse Yang-Mills fields whose propagators have the well-behaved (power-wise) form[9]

$$D_{\mu\nu,ab}^a(p;m)=-i\delta_{ab}\left(g_{\mu\nu}-\frac{p_\mu p_\nu}{p^2}\right)(p^2-m^2)^{-1}\,. \tag{2.14}$$

We then wish to know whether we can always perform a decomposition of the form

$$\mathcal{B}_\mu(x) = e^\xi \mathcal{A}_\mu e^{-\xi} + \frac{1}{g} e^\xi \partial_\mu e^{-\xi}, \qquad (2.15)$$

where the components A_μ^a and ξ_a, $a=1,2,3$, correspond, respectively, to transverse and longitudinal parts for our original Yang-Mills field B_μ^a.

This problem is solved in Ref. 5. It is shown that there indeed always exists an $S(x)=e^{\xi(x)}$ such that any given Yang-Mills field $\mathcal{B}_\mu(x)$ can be written as the gauge transform of a transverse Yang-Mills field $\mathcal{A}_\mu(x)$. The components of the latter can be expressed by

$$A_\mu^a = \sum_{i=1}^{3} \int P_\mu^i(x-x')B_i^a(x')dx', \qquad (2.16)$$

where the $P_\mu^i(x-x')$ have the properties

$$\partial^\mu P_\mu^i(x-x') = 0, \qquad (2.17a)$$

$$\int dx\, P_\mu^i(x-x')P_\mu^j(x''-x) = \delta_{ij}(x'-x''), \qquad (2.17b)$$

$$\sum_i \int dx\, P_\mu^i(x'-x)P_\nu^i(x''-x) = g_{\mu\nu}\delta(x'-x'') - \partial_\mu\partial_\nu D(x'-x''). \qquad (2.17c)$$

It now follows from (2.13) that Lagrangian (2.11) splits into two parts $\mathcal{L}_1 + \mathcal{L}_2$, where

$$\mathcal{L}_1 = -\tfrac{1}{4}\mathcal{G}_{\mu\nu}\mathcal{G}^{\mu\nu} - \tfrac{1}{2}m^2\mathcal{A}_\mu\mathcal{A}^\mu + g\mathcal{A}_\mu j^\mu + \mathcal{L}_\psi \qquad (2.18)$$

and

$$\mathcal{L}_2 = \frac{m^2}{2g}\left\{\mathcal{A}_\mu S^{-1}(x)\partial_\mu S(x) - \frac{1}{g}[S^{-1}(x)\partial_\mu S(x)]^2\right\}. \qquad (2.19)$$

We note in passing that (2.15) with the ensuing relations (2.16) and (2.17) has a welcome implication in conjunction with, at least, path-integral quantization of the theory. As shown in Ref. 5, when one passes from the measure $[d\mathcal{B}] = \prod_{a,\mu,x} dB_\mu^a(x)$ to the measure $[d\mathcal{A}][dS]$ a determinant factor enters which can be represented as a Faddeev-Popov (FP) ghost[10] particle integral. Thus, transformation (2.15) plays exactly the same role that gauge invariance transformations play in massless Yang-Mills theories. We conclude that Lagrangian (2.18) describes *in all respects* a non-Abelian *gauge* theory with an extra advantage: It has a built-in immunity against infrared divergences. Accordingly, $\mathcal{L}_1$ behaves nicely as far as renormalizability is concerned, and we can concentrate exclusively on $\mathcal{L}_2$ given by (2.19). We shall therefore turn our attention to this part. We find it advantageous to express $\mathcal{L}_2$ in terms of the longitudinal scalar fields ξ_a. We note that

$$\partial_\mu S(x) = \frac{\delta S(x)}{\delta\xi(x)}\partial_\mu \xi(x). \qquad (2.20)$$

Since $S(x)$ belongs to the SU(2) group, $\delta S(x)$ can be thought of as a tangent vector, i.e., as an element of the respective algebra. As readily calculated by Boulware[5] $\delta S(x)$ has the expansion

$$\delta S(x) = S(x)Q_{ab}(\xi)\delta\xi_a t_b, \qquad (2.21)$$

where[11]

$$Q_{ab}(\xi) = \left(\frac{e^\xi - 1}{\xi}\right)_{ab}, \quad \xi = \xi_a(x)t_a. \qquad (2.22)$$

Let us also introduce $G_{ab}(\xi)$ by

$$G_{ab} = (QQ^T)_{ab}. \qquad (2.23)$$

Formally, G can be thought of as the scalar product

$$G = (Q, Q). \qquad (2.24)$$

Inserting (2.20)–(2.23) in (2.19) we finally obtain

$$\mathcal{L}_2 \equiv \mathcal{L}_\xi = -\tfrac{1}{2}\partial_\mu\xi_a G_{ab}(\xi)\partial_\mu\xi_b + gm^2\partial_\mu\xi_a Q_{ab}(\xi)A_\mu^b. \qquad (2.25)$$

The nonrenormalizability of the theory is an aftermath of the nonpolynomial character of $Q_{ab}(\xi)$ and $G_{ab}(\xi)$. This can be seen best if we consider what happens in the corresponding (massive) Abelian case. We take the Abelian analog of (2.11)

$$\mathcal{L} = -\tfrac{1}{4}F_{\mu\nu}F^{\mu\nu} - \tfrac{1}{2}m^2 B_\mu B^\mu + gj^\mu B_\mu + \mathcal{L}_\psi. \qquad (2.26)$$

Suppose we quantize the theory by the path-integral method. The generating functional for the vector-field Green's functions is, of course, given by

$$Z[J] = Z_B^{-1}\int [dB]\exp\left[i\int(\mathcal{L} + J^\mu B_\mu)dx\right], \qquad (2.27)$$

where J_μ is a classical c-number source and Z_B^{-1} is the usual normalization factor.

After a Stückelberg decomposition[12] which separates the transverse from the longitudinal parts of B_μ, denoted by A_μ and ξ, respectively, one can readily show[5] that the integration over ξ can be done separately, and that it contributes a factor which can be absorbed in the wave-function renormalization for each charged field present in the theory. Actually, this factor is of unlimited degree corresponding to, formally, an infinite number of subtractions. Accordingly, one does not have, in this respect, renormalizability in the sense, e.g., of usual quantum electrodynamics. However, as Boulware argues, the S matrix shows no effects of nonrenormalizability on account of such a factor. Crudely speaking, one can think of having replaced a charged spinor ψ, e.g., by $\psi' = e^{i\xi(x)}\psi$.

The question facing us, then, is whether an analogous procedure works in the non-Abelian case. Thus, one wants to know what happens, within the context of path-integral quantization at least, when the integral over the longitudinal fields is evaluated. To be explicit, consider the generating functional for longitudinal field Green's functions

$$G_\xi[J]=Z_\xi^{-1}\int [d\xi]\,(\det{}^{1/2}G)\exp\left[-\frac{m^2}{g^2}\int dx\,(\tfrac{1}{2}\partial_\mu\xi_a G_{ab}\partial_\mu\xi_b+g\partial_\mu\xi_a Q_{ab}J^b_\mu)\right], \qquad (2.28)$$

where $\det^{1/2}G$ enters because we have changed integration variables from $S(x)=e^\xi$ to ξ. The source functions J^b_μ are related to the sources K^b_μ of the transverse vector fields[13] by

$$J^b_\mu=A^b_\mu-\frac{1}{m^2}K^b_\mu. \qquad (2.29)$$

Now, the essential difference between the Abelian and the non-Abelian cases is that in the latter the longitudinal modes interact with the transverse modes. Accordingly, there is no hope that the integration over the ξ_a's will yield a factor independent of all the other fields leading to the renormalization factor interpretation of the Abelian example. On the other hand, there is still some hope left. It can be expressed by the following question. Is it possible that $\mathcal{L}_\xi$ describes a "subtheory" in which the ξ_a's behave, for all practical purposes, just as charged fields behave in the presence of gauge vector fields? Put in a different way: Could $\mathcal{L}_\xi$ describe, apart from a term $\frac{1}{4}G^a_{\mu\nu}G^a_{\mu\nu}$ which is already included in $\mathcal{L}_1$ anyway, a theory which is equivalent to a non-Abelian *gauge* theory in which the ξ_a's play the role of the charged fields? If this answer were affirmative it would follow that the $\mathcal{L}_\xi$ part of the theory would be renormalizable by standard arguments. But this would have been the case if there were a transformation of variables $\xi\to\xi(\xi')$ such that

$$\partial_\mu\xi'_a=Q_{ab}(\xi)\partial_\mu\xi_b. \qquad (2.30)$$

In that case, the $\mathcal{L}_\xi$ part would assume the familiar form of a nicely behaving, non-Abelian gauge theory.[14] However, Eq. (2.30) is not integrable and the sought-for transformation is not available. As a result the A^a_μ will be coupled, through the $\partial_\mu\xi_a Q_{ab}(\xi)A^b_\mu$ term, to propagators of the form

$$\langle T(\Xi_{\mu_1}\cdots\Xi_{\mu_n})\rangle,$$
$$\Xi_{\mu_i}\equiv\partial_{\mu_i}\xi_{a_i}(x_i)Q_{a_ib_i}(\xi(x_i)),\quad i=1,\ldots,n \qquad (2.31)$$

which behave badly at high momenta. One concludes that the nonpolynomial character of $Q_{ab}(\xi)$ and $G_{ab}(\xi)$ in (2.25) is responsible for the nonrenormalizability of the theory. In passing we note that had we been able to diagonalize the two nonpolynomial terms in (2.25), i.e., if (2.30) were to hold true, then $\det^{1/2}G$ would be equal to unity.

We now wish to remark on the presence of matter fields in our model. The redefinition (2.15), which splits the massive vector fields into transverse and longitudinal parts, must be accompanied by a redefinition $\psi\to R(x)\psi$ of the charged matter fields. Now, $R(x)$ has the form

$$R(x)=e^{\xi_a(x)T_a}, \qquad (2.32)$$

where the T_a are representation matrices of the group generators in the space of the fields. In turn, (2.32) implies that a renormalization factor of unlimited degree which makes its appearance in the Abelian case will also appear in the non-Abelian case.

Finally, it is of interest to note the analogy with the following result obtained in Ref. 7. Tree unitarity, a concept closely related to renormalizability, implies that the scalar sector of a Lagrangian which originally has any general (even nonpolynomial) form, such as (2.25), should be brought into a diagonal form. Thus, transformation (2.30) *must* exist by tree unitarity. Of course, the authors of Ref. 7 do not face the same impasse as Boulware because they have additional scalar fields.

We shall proceed, in the next section, to lift the renormalizability barriers by investigating the conditions under which the diagonalizability of the nonpolynomial terms becomes possible.

III. NECESSARY CONDITIONS FOR DIAGONALIZABILITY

Our considerations in the preceding section have led us to the unwelcome terms in our Lagrangian

$$\mathcal{L}_\xi=-\tfrac{1}{2}\partial_\mu\xi_a G_{ab}(\xi)\partial_\mu\xi_b+g\partial_\mu\xi_a Q_{ab}(\xi)A_{\mu b}, \qquad (3.1)$$

where $G_{ab}(\xi)$ and $Q_{ab}(\xi)$ are nonpolynomial expressions in the longitudinal scalar fields $\xi_a(x)$, $a=1,2,3$. The nonpolynomial character of $G_{ab}(\xi)$ and $Q_{ab}(\xi)$ is responsible for the renormalizability problems surrounding the models we have been studying. If we were in a position, through a transformation of variables $\xi \to \xi(\xi')$, to bring $\mathcal{L}_\xi$ to the form

$$\mathcal{L}_{\xi'} = -\tfrac{1}{2}\partial_\mu \xi'_a \partial_\mu \xi'_a + g'\partial_\mu \xi'_a A_{\mu a}\,, \qquad (3.2)$$

then we would not have had renormalizability difficulties in the first place. Unfortunately, no transformation $\xi \to \xi(\xi')$ which takes (3.1) to (3.2) is available.[5]

To make progress we turn our attention to the manifold of the longitudinal fields. We shall employ, to this end, the language of differential geometry. For the convenience of the reader we gather a number of facts which will prove useful to subsequent developments.

Let us recall that the Lie algebra of a Lie group is the tangent space at the identity element of the group manifold. As such, it is a flat linear vector space. An element belonging to the adjoint representation of a Lie group G is, by definition, an element of the Lie algebra g.[15] The adjoint representation of G on g is a mapping $G \to \mathrm{GL}(g)$, where $\mathrm{GL}(g)$ stands for the general linear group of transformations on g as a vector space. Let $\sigma \in G$. The adjoint action of σ on $A \in g$ is specified by

$$\mathrm{ad}(\sigma)A = A', \quad e^{A'} = e^{\xi} e^{A} e^{-\xi}, \quad e^{\xi} = \sigma\,. \qquad (3.3)$$

The above prescription makes sense once we have introduced the exponential mapping exp, sometimes e for short.[16] It is a mapping from g to G, i.e., if $A \in g$ then $e^A \in G$, such that $e^0 = \epsilon$, where 0 stands for the zero vector in g and ϵ denotes the identity element of G. In addition, exp maps lines in g onto geodesics in G.

Note that since g can be thought of as a (flat) manifold in its own right exp is a mapping between manifolds. Consequently, its differential $d\exp$ is a mapping between their respective tangent spaces. But since g is flat, it can be identified with its tangent space at each point. Hence, $d\exp$ maps g onto itself. The differential of the exponential mapping will play a crucial role in our subsequent analysis.

Given now that g is a linear vector space, we can always choose for it an orthonormal basis $t_1, \ldots, t_n$. Each element A of g is then of the form

$$A = A_a(x) t_a\,. \qquad (3.4)$$

Through the exponential mapping, we can assign the coordinates $(A_1(x), \ldots, A_n(x))$ to the group element e^A. These coordinates form the so-called canonical or normal coordinate set for the group manifold.

Having gathered the above useful facts we shall proceed to study in a systematic manner various matters of importance connected with the adjoint representation of a gauge group. We shall concentrate, for the most part, on SU(2).

Naturally, the subtleties inherent to gauge transformations can be traced to their local character. Accordingly, one must deal with infinite-dimensional groups. Since, in particular, a coordinate set for the group manifold is an n-tuple of functions, questions related to the presence of derivatives immediately come up. Thus, consider the su(2) element $\xi = \xi_a(x) t_a$, $a = 1, 2, 3$. Just as ξ_a is a function, assigning the number $\xi_a(x)$ at x, $\partial_\mu \xi_a$ is a function as well, assigning the number $\partial_\mu \xi_a(x)$ at x. [We assume that the index μ is fixed so that one must not worry about the different Lorentz character of $\xi_a(x)$ and $\partial_\mu \xi_a(x)$.] Then, certainly, $\partial_\mu \xi = \partial_\mu \xi_a(x) t_a$ is an element of su(2). Can we use the $\partial_\mu \xi_a(x)$, μ fixed, to replace the $\xi_a(x)$ in providing a (normal) coordinate set serving to describe the group manifold? Alternatively, can we parametrize su(2) through the $\partial_\mu \xi_a(x)$, μ fixed?

The answer to the above question is, in fact, negative. Had it been affirmative we would have actually had, at the same time, the solution to our main problem posed at the beginning of this section. This statement should become evident once we have proved the following.

Lemma 3.1. For each element $\partial_\mu \xi$ of su(2) there exists another su(2) element $D_\mu \xi$ which cannot be put in the form $\partial_\mu \xi'$, $\xi' \in$ su(2). (The index μ is fixed throughout.)

Proof. Consider the exponential mapping. As pointed out earlier, it can be thought of as a mapping between manifolds, su(2) and SU(2) in the present case. However, let us be more general for the moment and talk about g and G. The flatness of g implies that $d\exp$ maps g onto itself. But since g is identified with its tangent space at each point, we must specify that point of g whose tangent space we are considering. In particular, $D_\mu \xi$ defined by

$$D_\mu \xi = d\exp_\xi \partial_\mu \xi\,, \qquad (3.5)$$

where ξ is that su(2) point whose tangent space we are mapping, is an element of su(2).

Recall now the well-known formula[17] for any Lie algebra g

$$d\exp_X Y = \frac{1 - e^{-\mathrm{ad}X}}{\mathrm{ad}X}\, Y, \quad X, Y \in g \qquad (3.6)$$

where $(1 - e^{-\mathrm{ad}X})/\mathrm{ad}X$ stands formally for $\sum_0^\infty (-\mathrm{ad}X)^n/(n+1)!$. Accordingly

$$(D_\mu \xi)_a = \partial_\mu \xi_b Q_{ba}(\xi)\,, \qquad (3.7)$$

where $Q_{ba}(\xi)$ is the same expression appearing in (3.1).

We also see, in comparison to (3.1), that

$$\partial_\mu \xi_a G_{ab} \partial_\mu \xi_b = (D_\mu \xi, D_\mu \xi), \quad (3.8)$$

where (,) denotes the scalar product in su(2).

But we have already pointed out that one cannot find a transformation of variables $\xi \to \xi(\xi')$ such that $\partial_\mu \xi_a(x) Q_{ab}(\xi)$ and $\partial_\mu \xi_a G_{ab}(\xi) \partial_\mu \xi_b$ become $\partial_\mu \xi'_a$ and $(\partial_\mu \xi'_a)^2$, respectively.[5] This simply means that there exists no $\xi' \in$ su(2) such that $D_\mu \xi = \partial_\mu \xi'$.

From the arguments in the above proof there are a number of realizations to be made. To begin with, it becomes quite obvious that the diagonalizability problems which arose in the first place are intimately connected with the "bad character" or insufficiency of the gradient functions to parametrize su(2).

Secondly, the nonpolynomial factor $G_{ab}(\xi)$ entering (3.1) cannot be interpreted literally as a metric tensor for the longitudinal scalar field manifold. In fact, the latter *is* flat since it is merely the vector space su(2). It is, perhaps, more appropriate to view $G_{ab}(\xi)$ as a factor which relates the inner su(2) product $(\partial_\mu \xi, \partial_\mu \xi)$ to the inner su(2) product $(D_\mu \xi, D_\mu \xi)$. Its nonpolynomial character is responsible for the fact that the latter product cannot be put in the form $(\partial_\mu \xi', \partial_\mu \xi')$, i.e., that the $\partial_\mu \xi_a(x)$, μ fixed, are insufficient to furnish by themselves a coordinate set suitable for the description of the group manifold. On the other hand, $G_{ab}(\xi)$ has some formal features of a metric tensor. For one thing, it is a symmetric matrix. Accordingly, its diagonalization is subject to specifications which, in general, apply for the metric tensor of a Riemannian manifold.

Finally, lemma 3.1 has introduced the object $D_\mu \xi$ which we shall call the *exponential gradient* of ξ. D_μ is some kind of a gradient operator. However, it lacks several basic properties of such operators, e.g., $D_\mu(\xi + \zeta) \neq D_\mu \xi + D_\mu \zeta$. It will become of central importance for the balance of this section because of its fundamental involvement in gauge transformations. This can immediately be seen by the following.

Lemma 3.2. $D_\mu \xi$ transforms, under the action of the adjoint representation of SU(2), in a manner similar to the gauge transformation of a vector field. The transformation properties of $D_\mu \xi$ under the adjoint action guarantee the multiplicative group property of gauge transformations for a vector field.

Proof. We start by recalling the identity

$$D_\mu \xi = e^{\xi} \partial_\mu e^{-\xi}. \quad (3.9)$$

Comparing with (2.4), one immediately sees that a non-Abelian vector field transforms under the gauge action by

$$\mathfrak{B}'_\mu(x) = \mathrm{ad}(e^{\zeta}) \mathfrak{B}_\mu(x) + D_\mu \zeta. \quad (3.10)$$

To obtain (2.4) explicitly observe, first, that by (3.3) we have

$$e^{\mathrm{ad}(e^{\zeta}) \mathfrak{B}_\mu(x)} = e^{\zeta} e^{\mathfrak{B}_\mu(x)} e^{-\zeta}. \quad (3.11)$$

Recalling the operator identity

$$A e^{B} A^{-1} = e^{ABA^{-1}}, \quad (3.12)$$

we obtain

$$\mathrm{ad}(e^{\zeta}) \mathfrak{B}_\mu(x) = e^{\zeta} \mathfrak{B}_\mu(x) e^{-\zeta} \quad (3.13)$$

and thus recover (2.4).

Applying (3.13) to $D_\mu \xi$ we have

$$\mathrm{ad}(e^{\zeta}) D_\mu \xi = e^{\zeta} e^{\xi} (\partial_\mu e^{-\xi}) e^{-\zeta}. \quad (3.14)$$

But

$$e^{\zeta} e^{\xi} (\partial_\mu e^{-\xi}) e^{-\zeta} = e^{\zeta} e^{\xi} \partial_\mu (e^{-\xi} e^{-\zeta}) - e^{\zeta} \partial_\mu e^{-\zeta}. \quad (3.15)$$

Now $e^{\zeta} e^{\xi}$ is just another group element e^{σ}, where

$$\sigma = \zeta + \xi + \frac{1}{2!}[\xi, \zeta] + \frac{1}{3!}[[\xi, \zeta], \zeta] + \cdots. \quad (3.16)$$

Accordingly,

$$\mathrm{ad}(e^{\zeta}) D_\mu \xi = D_\mu \sigma - D_\mu \zeta, \quad (3.17)$$

which shows, once we formally denote $D_\mu \sigma \equiv (D_\mu \xi)'$, that $D_\mu \xi$ transforms under the adjoint action in a manner analogous to that in which $\mathfrak{B}_\mu(x)$ transforms under the gauge action.

Consider now two successive gauge transformations on $\mathfrak{B}_\mu(x)$ parametrized by ξ and ξ', respectively,

$$\mathfrak{B}'_\mu = \mathrm{ad}(e^{\xi}) \mathfrak{B}_\mu + D_\mu \xi \quad (3.18)$$

and

$$\mathfrak{B}''_\mu = \mathrm{ad}(e^{\xi'}) \mathfrak{B}'_\mu + D_\mu \xi' = \mathrm{ad}(e^{\xi'}) \mathrm{ad}(e^{\xi}) \mathfrak{B}_\mu + \mathrm{ad}(e^{\xi'}) D_\mu \xi + D_\mu \xi'. \quad (3.19)$$

But

$$\mathrm{ad}(e^{\xi'}) \mathrm{ad}(e^{\xi}) \mathfrak{B}_\mu = e^{\xi'} e^{\xi} \mathfrak{B}_\mu e^{-\xi} e^{-\xi'} = e^{\rho} \mathfrak{B}_\mu e^{-\rho} = \mathrm{ad}(e^{\rho}) \mathfrak{B}_\mu, \quad (3.20)$$

where

$$\rho = \xi + \xi' + \frac{1}{2!}[\xi', \xi] + \frac{1}{3!}[[\xi', \xi], \xi] + \cdots.$$

On the other hand,

$$\mathrm{ad}(e^{\xi'}) D_\mu \xi = D_\mu \rho - D_\mu \xi. \quad (3.21)$$

Therefore,

$$\mathfrak{B}''_\mu = \mathrm{ad}(e^{\rho}) \mathfrak{B}_\mu + D_\mu \rho, \quad (3.22)$$

which establishes the (basic) multiplicative group property of gauge transformations.

From lemma 3.2 we immediately have the following.

Corollary 3.1. $D_\mu \xi$ is the covariant gradient of gauge transformations.

Proof. Since $D_\mu \xi$ has a Lorentz vector-field index, its gauge transform is given by (3.10), i.e.,

$$(D_\mu \xi)' = \mathrm{ad}(e^{\xi'}) D_\mu \xi + D_\mu \xi' . \tag{3.23}$$

Using (3.17) we deduce

$$(D_\mu \xi)' = D_\mu (\xi + \xi' + \tfrac{1}{2}[\xi, \xi'] + \cdots), \tag{3.24}$$

or, in other words, $(D_\mu \xi)'$ is of the form $D_\mu \rho$, $\rho \in \mathrm{su}(2)$.

It becomes evident through what we have shown so far that the presence of exponential gradients in the expressions in (3.1) is not accidental but an integral part of the gauge scheme. Furthermore, the inability to put $D_\mu \xi$ in the form $\partial_\mu \xi'$ suggests that no contemplation which restricts itself to the longitudinal field variables $\xi_a(x)$ only can succeed in replacing the exponential by regular gradients of fields. Explicitly, at this point we can list the following two conclusions:

(a) The presence of nonpolynomial factors in (3.1) must inevitably be accepted as an integral part of the covariant gradient of gauge transformations, i.e., $D_\mu \xi$.

(b) Any attempts to circumvent the present impasse must look beyond the longitudinal scalar field manifold; an enlargement of some sort is needed.

We shall now turn our attention to the nonpolynomial factors we have encountered and establish necessary conditions for their removal. This we do by proving the following.

Theorem 3.1. A necessary condition for removing the exponential gradients from (3.1) in favor of regular gradients, or, alternatively, to diagonalize $G_{ab}(\xi)$ [and $Q_{ab}(\xi)$] in (3.1), is the introduction of additional fields. Such an introduction should not increase the rank of $G_{ab}(\xi)$ as a matrix. In the case of SU(2) the only ambiguity-free solution is the complexification of ξ [$\in \mathrm{su}(2)$] in the exponential gradient, or, more generally, the replacement $\xi \to (\xi, \eta)$, $\eta \in \mathrm{su}(2)$, in $D_\mu \xi$. This means that, in the case of SU(2), the additional (scalar) fields belong to the adjoint representation.

Proof. We shall concentrate on $G_{ab}(\xi)$, or, symbolically, $G(\xi)$. It is a symmetric $n \times n$ matrix where n is the dimension of the Lie algebra g. It is constructed from functions $\xi_a(x)$ which are components of elements belonging to a space, i.e., g, whose discrete dimension is n. Consequently, its rank r at each point x must be $r \leqslant n$. We assume, as is the case for su(2), that all components of $\xi(x)$ participate in the construction of $G(\xi)$. Accordingly, $r = n$. From this follows that $G(\xi)$ has $\frac{1}{2}n(n+1)$ independent elements. Consequently, its diagonalization involves $\frac{1}{2}n(n+1)$ conditions. Once more we see that a mere replacement $\xi \to \xi(\xi')$ cannot, in general, achieve the diagonalization because ξ' introduces only n components to satisfy $\frac{1}{2}n(n+1)$ conditions. But

$$\tfrac{1}{2}n(n+1) > n \quad \text{for } n \geqslant 2 .$$

We conclude that for any non-Abelian gauge group ($n > 2$) the diagonalization of $G(\xi)$ must involve more than the n longitudinal scalar variables which are available. Hence, a necessary condition for the above diagonalization is the introduction of additional (field) variables—in general $\frac{1}{2}n(n-1)$, less if the rank of $G(\xi)$ is not maximal.

Consider the most general way of introducing the additional fields, namely the replacement $\xi \to \xi(\Xi)$ where $\Xi(x)$ is a $\frac{1}{2}n(n+1)$-component field. Let us recall that $G(\xi)$ enters as a factor between regular gradients in (3.1). Thus, the transformation $\xi \to \xi(\Xi)$ implies

$$G_{ab}(\xi) \to G'_{kl}(\Xi) = \frac{\delta \xi_a}{\delta \Xi_k} \frac{\delta \xi_b}{\delta \Xi_l} G_{ab}(\xi(\Xi)), \tag{3.25}$$
$$k, l = 1, 2, \ldots, \tfrac{1}{2}n(n+1) .$$

This means that $G(\xi)$ is actually replaced by the $[\frac{1}{2}n(n+1)] \times [\frac{1}{2}n(n+1)]$ matrix $G'(\Xi)$. Furthermore, since in general all components of Ξ participate in building up $G'(\Xi)$, the rank of the latter could be maximal, i.e., $\frac{1}{2}n(n+1)$. We see that, in general, to diagonalize $G'(\Xi)$ we must impose, once more, $\frac{1}{2}m(m+1)$ conditions on m variables, where $m = \frac{1}{2}n(n+1)$. In other words, an arbitrary substitution $\xi \to \xi(\Xi)$, where Ξ has more components than ξ, will not solve our problem but, in fact, worsen it. On the other hand, Ξ cannot have fewer components than ξ since that would correspond to the physically meaningless situation of reducing dynamical degrees of freedom.

There is only one way out of our present impasse. The addition of any field variables must not be accompanied by a proportional increase in the number of conditions for diagonalizing G'. Now, the size of G inevitably increases with the introduction of extra variables, as can be clearly seen from (3.25). The only way to minimize the rate of increase in the number of conditions is to impose that the rank of G remains constant.[18]

Let there be given a symmetric $m \times m$ matrix A of rank $r < m$. It can be shown[19] that A has only $\frac{1}{2}m(m+1) - \frac{1}{2}(m-r)(m-r+1) = \frac{1}{2}r(2m-r+1)$ independent elements. Now, if the size of G' is $m \times m$ it means, from (3.25), that we must have introduced m variables (i.e., fields) in the place of the original n. Therefore, for the diagonalization of G' it must be true that

$$\tfrac{1}{2}r(2m-r+1)=m. \tag{3.26}$$

This is actually the best we can do in balancing the number of conditions against the number of available variables. Now (3.26) has solutions

$$r=1, \quad r=2m. \tag{3.27}$$

The first solution is rejected outright, for a non-Abelian group, on physical grounds; it implies the reduction in dynamical variables. The second solution, although mathematically nonsensical (it suggests that the rank of a symmetric matrix is larger than its number of rows or columns), will be interpreted by us in a backward manner. Our interpretation will lead us to an unambiguous solution concerning the diagonalization of $G(\xi)$ in the case of SU(2). Furthermore, the uniqueness of this condition will guarantee the uniqueness of our solution.

We start by recalling that the sum of two symmetric $m\times m$ matrices A and B of rank r each is a matrix of rank $\leqslant 2r$, provided, of course, that $r\leqslant\frac{1}{2}m$. Suppose, on the other hand, that we insert A and B into larger $2m\times 2m$ matrices and add them as follows:

$$\begin{pmatrix}A&0\\0&0\end{pmatrix}+\begin{pmatrix}0&0\\0&B\end{pmatrix}=C. \tag{3.28}$$

Assuming that the ranks of both A and B were maximal, i.e., $r(A)=r(B)=m$, then, indeed, the rank of C could be as much as $2m$. Of course, we have not added properly the two $m\times m$ matrices. However, we may think of A as representing the first matrix in (3.28) and B as representing the second. Accordingly, C is also thought of as an $m\times m$ matrix of rank $2m$. This point of view constitutes the key to our present interpretation.[20]

Before proceeding further we should point out that the way A and B were placed into the larger matrices is crucial. Had we added them, e.g., as

$$\begin{pmatrix}A&0\\0&0\end{pmatrix}+\begin{pmatrix}0&B\\0&0\end{pmatrix},$$

we would never have been in a position to obtain a matrix of rank greater than m.

Identifying the matrix A above with $G(\xi)$ we see that, in the case of SU(2) ($n=m=r=3$), we need a second 3×3 matrix to fit the scheme just presented. With this observation as a guideline we shall now go ahead to exhibit an explicit solution to our original problem for SU(2). Once we have given our solution, we shall examine it in the light of the above interpretation of the unique condition $r=2m$.

Consider the replacement of the exponential gradient $D_\mu\xi$ by

$$D_\mu\xi=d\exp_\xi\partial_\mu\xi\to d\exp_\xi\partial_\mu(\xi+i\eta)\equiv\mathfrak{D}_\mu\zeta, \tag{3.29}$$

where $\eta\in\mathrm{su}(2)$. Note that $\mathfrak{D}_\mu\zeta$ is not an exponential gradient, i.e., it is not of the form $d\exp_{\xi+i\eta}\partial_\mu(\xi+i\eta)$.

Next we introduce the following scalar product to replace $(D_\mu\xi,D_\mu\xi)$:

$$(\mathfrak{D}_\mu\zeta,\mathfrak{D}_\mu\zeta)\equiv\mathfrak{D}_\mu\zeta_a^*\mathfrak{D}_\mu\zeta^a. \tag{3.30}$$

Now

$$\mathfrak{D}_\mu\zeta_a=\partial_\mu\xi_a+i\partial_\mu\eta_a+\tfrac{1}{2}C_a^{bc}\xi_c(\partial_\mu\xi_b+i\partial_\mu\eta_b)+\frac{1}{3!}C_a^{bc}C_c^{de}\xi_b\xi_e(\partial_\mu\xi_d+i\partial_\mu\eta_d)+\cdots \tag{3.31}$$

and similarly for $\mathfrak{D}_\mu\zeta_a^*$.

For (3.30) one straightforwardly obtains

$$(\mathfrak{D}_\mu\zeta,\mathfrak{D}_\mu\zeta)=\partial_\mu\xi_aG_{ab}(\xi)\partial_\mu\xi_b+\partial_\mu\eta_aG_{ab}(\xi)\partial_\mu\eta_b\,. \tag{3.32}$$

Note that G_{ab} enters *both* right-hand side terms as an expression of the ξ_a's only.

Consider the redefinition $\xi\to\xi(\xi')$, $\eta\to\eta(\eta')$, where ξ', $\eta'\in\mathrm{su}(2)$. We have

$$(\mathfrak{D}_\mu\zeta,\mathfrak{D}_\mu\zeta)=\partial_\mu\xi_a'\frac{\delta\xi_b}{\delta\xi_a'}\frac{\delta\xi_c}{\delta\xi_d'}G_{bc}(\xi')\partial_\mu\xi_d'+\partial_\mu\eta_a'\frac{\delta\eta_b}{\delta\eta_a'}\frac{\delta\eta_c}{\delta\eta_d'}G_{bc}(\xi')\partial_\mu\eta_d'\,. \tag{3.33}$$

For simplicity, we denote $f_a^b=\delta\xi_b/\delta\xi_a'$ and $h_a^b=\delta\eta_b/\delta\eta_a'$. We then immediately obtain the diagonalization conditions

$$f_a^b(\xi')f_d^c(\xi')G_{bc}(\xi')=\delta_{ad} \tag{3.34}$$

and

$$h_a^b(\eta')h_d^c(\eta')G_{bc}(\xi')=\delta_{ad} \tag{3.35}$$

or

$$\tfrac{1}{2}[f_a^b(\xi')f_d^c(\xi')+h_a^b(\eta')h_d^c(\eta')]G_{bc}(\xi')=\delta_{ad}\,. \tag{3.36}$$

But (3.36) describes six conditions on the six variables ξ_a', η_a', $a=1,2,3$, so that it is, in general, solvable. This is enough to establish the necessity of introducing the η_a's. Viewing conditions (3.36) as an algebraic system of six equations for six unknowns, for each point x, one can resort to numerical methods for a solution.

In passing we note that the same procedure which diagonalizes $G(\xi)$ also diagonalizes $Q(\xi)$ in (3.1). Indeed, (3.34) and (3.35) are basically the squares of the conditions necessary to diagonalize $Q(\xi)$.

Looking at (3.32) it becomes straightforward to recognize the parallelism with the way we interpreted condition $r = 2m$ in (3.27). The two matrices on the right-hand side correspond to A and B and the way they were introduced, i.e., through the complexification of ξ in $D_\mu \xi$, corresponds to putting them together as in (3.28). Thus, C is viewed by us as a 6×6 matrix which, after diagonalization, has the form

$$C = \begin{bmatrix} (\partial_\mu \xi_1')^2 & & & \\ & \cdot & & \\ & & \cdot & \\ & & & \cdot & \\ & & & & (\partial_\mu \eta_3')^2 \end{bmatrix}. \qquad (3.37)$$

Note that the replacement $\partial_\mu \xi_a G_{ab}(\xi)\partial_\mu \xi_b \to \partial_\mu \xi_a G_{ab}(\xi)\partial_\mu \xi_b + \partial_\mu \eta_a G_{ab}(\eta)\partial_\mu \eta_b$ would not have led to a solution. It would have corresponded to having a matrix C of the form $\left(\begin{smallmatrix} A & B \\ 0 & 0 \end{smallmatrix}\right)$ so that its rank could not possibly have been $2m$.

It follows from the above that the role of the exponential gradient $D_\mu \xi$ has once more proved crucial. In fact, one could trace the success of our program for SU(2) to the property $D_\mu(\xi + i\eta) \neq D_\mu \xi + iD_\mu \eta$. This property is responsible for the nontrivial mixing of the ξ and η terms in (3.32).

There are, now, only two other possibilities along the same lines which could have led to six conditions for six variables:

(a) $D_\mu \xi \to d\exp_{\xi+i\eta}\partial_\mu(\xi + i\eta)$,

(b) $D_\mu \xi \to d\exp_{\xi}\partial_\mu(\xi + \eta)$.

It can easily be seen that either case would correspond to C built up of three $m\times m$ matrices, i.e., of the form

$$C = \begin{pmatrix} A & D \\ D & B \end{pmatrix}. \qquad (3.38)$$

There is nothing formally wrong with such an approach since the requirement $r = 2m$ can be generally fulfilled by C in (3.38). However, at least the first of the above substitutions would have ambiguous physical interpretations. In fact, (a) actually implies the enlargement of the gauge group. The reason is that $d\exp_{\xi+i\eta}(\xi + i\eta) = D_\mu \zeta \neq \mathfrak{D}_\mu \zeta$ is *properly* an exponential gradient of an element belonging to the complexified su(2) vector space. It would then appear that we have enlarged our gauge group to $SU(2)\times SU(2)$ and, therefore, increased the number of gauge fields.

Replacement (b) is not as problematical as (a). However, since $\xi + \eta = \theta \in su(2)$, the introduction of $d\exp_{\xi}\partial_\mu \theta$ into our formulas casts an ambiguity with respect to gauge transformations. This is because $D_\mu \xi$ is the covariant derivative of gauge transformations (recall lemma 3.2) so that the addition of η without the simultaneous addition of a new algebraic direction (namely, the imaginary axis) may lead to inconsistencies. Whether this criticism really applies or not is an open question.

Even assuming now that our criticism of replacement (b) is not valid, one may safely conclude the following. The procedure which works *and* is consistent with our interpretation of the unique condition $r = 2m$ in (3.27) regarding the diagonalization of $G(\xi)$ is, in the case of SU(2), the introduction of additional variables via the extension $\xi \to (\xi, \eta), \eta \in su(2)$. The replacement $\xi \to (\xi, \eta)$ occurs in the exponential gradient $D_\mu \xi$ appearing in (3.1). Finally, such a procedure implies that the additional scalar fields must belong to the adjoint representation of SU(2), i.e., they form an SU(2) triplet. This completes our proof.

The introduction according to theorem 3.1 of extra fields via the complexification procedure, or, more generally, $\xi \to (\xi, \eta)$, $\eta \in su(2)$, shows that we have not really gone to a space of fields with a higher discrete dimension. In fact, we have drawn our extra fields from the space of the ξ_a's rejecting, at the same time, the replacement $D_\mu \xi \to D_\mu \zeta$ ($\zeta = \xi + i\eta$) which would have corresponded to a true complexification of su(2) and the doubling of (real) dimensions. What all this signifies is that it would be incorrect to view (ξ, η) as a vector belonging, for each given x, to a six-dimensional space. Consequently, we are led to interpret the η_a's as variables without dynamical significance; *the η_a's are not new dynamical variables*. Instead, we interpret them as ghosts, or superfluous fields.[21]

As a final remark we should point out that our explicit solution for SU(2) is not contingent on our interpretation of condition $r = 2m$ in (3.27). The latter was simply invoked in order to guarantee the uniqueness of our solution.

IV. CONNECTION WITH SPONTANEOUSLY BROKEN GAUGE THEORIES

The quantization of theories which involve gauge vector fields presents the well-known problem of how to handle redundant components which do not correspond to dynamical degrees of freedom. The choice of gauge is, in fact, a necessary step for eliminating nondynamical field components. Well-known questions immediately arise surrounding the choice of gauge, e.g., covariant gauge vs indefinite-metric Hilbert space, manifest unitarity in the S matrix, etc. In the case of spontaneously broken non-Abelian gauge theories, it has been recognized that the choice of gauge is also intimately connected with renormalizability. In the so-called unitary (U) gauge[22] in which all redundant field variables are eliminated, renormalization is not obvious. In the renormalizable (R) gauge, on the other hand, where extra fields, i.e., fields not corresponding to dynamical variables, are retained, the renormalization program goes through.[1,2] The role of the extra fields is vital for unitarity.

What we would like to examine now is whether our findings in the preceding section have any implications connected with spontaneously broken gauge theories. From now on we follow Ref. 7 in abbreviating the latter by SBGT. Before proceeding with our analysis it is important at this point that we make our position clear on the following interpretation: We adopt Utiyama's[23] point of view according to which gauge fields enter a theory not *a priori* but through the requirement that the gauge transformation, which is responsible for charge conservation, be local in character. This (physical) requirement leads to the condition that the gauge fields be "manifestly" massless.[24] But masslessness together with irreducibility under Poincaré transformations implies that a (physical) gauge field has two degrees of freedom. Accordingly, the reduction of the dynamical components of a gauge vector field to two, from the available four, is not a (mathematical) consequence of fixing the gauge. Rather, fixing the gauge follows from the (physical) requirement that each gauge field has only two dynamical components. Furthermore, we maintain that the original, manifestly symmetric Lagrangian of an SBGT involves vector fields of two degrees of freedom (each) even before a choice of gauge has been made. This is important in view of the fact that the R gauge, at least, is fixed not in the original, visibly symmetric configuration but after the spontaneous breaking has been accounted for via the redefinition of the (scalar) fields.

As is well known, the symmetric configuration in an SBGT involves a number of scalar fields, some of which have nonzero vacuum expectation values. A redefinition in terms of shifted fields, which have vanishing vev's, produces mass terms for all or some of the vector fields. Suppose, then, that there were m gauge vector fields in the original (symmetric) configuration and n scalar fields ($n > m$). The dynamical degrees of freedom are $2m+n$. The redefinition produces m massive vector fields (we assume, for simplicity, that the whole symmetry breaks) which account for $3m$ degrees of freedom. The m would-be Goldstone bosons have given the gauge fields a longitudinal part. This means that the scalar sector of the Lagrangian must now account for only n-m degrees of freedom. The difference between the R and U gauges, now, is the following. The first keeps n scalar fields anyway, calling the m-redundant fields ghosts.[21] We prefer the term superfluous. The U gauge does away with the superfluous fields and keeps as many field components as there are dynamical degrees of freedom. We view our extra fields η_a, introduced in the preceding section, as exactly those m-redundant fields which are an integral part of the R gauge. In fact, according to what we have argued at the end of that section, the very manner by which the η_a were introduced shows that they should not assume any dynamical significance. Their role is simply to cure the renormalizability pathologies which our starting Lagrangians presented. However, such an assertion cannot be truly established before we show a connection between the models we have been studying and models of the SBGT variety. The reason is that unless we know that our Lagrangians of interest are in some sense equivalent to manifestly gauge-invariant Lagrangians, the bridge between the two being perhaps spontaneous symmetry breaking, there would not have been a gauge to fix in the first place. Therefore, we must now turn our attention to the following question: Under what circumstances is a Lagrangian such as (2.11), or one that has been augmented by the addition of the η_a's, equivalent to a (manifestly) gauge-invariant Lagrangian? Furthermore, we would also like to know what is the nature of such a link, if any.

We have dealt with exactly the same question for the Abelian case in Ref. 6, where we argued as follows. Suppose we are given a theory which involves a massive vector field minimally coupled to all matter fields. Then, the addition of a single physical, real scalar field χ together with the requirement that the above theory (including χ) be related, in a way to become evident shortly, to a "strictly" gauge-invariant one, forces the latter to be an SBGT. By a strictly gauge-invariant Lagrangian, we mean one in which all matter fields are charged and are minimally coupled to the gauge fields. Any extra fields are inserted in a way consistent with gauge invariance (e.g., a real scalar field belonging to the trivial representation of the gauge group) spoil strict but not manifest gauge invariance.

With this distinction in mind let us begin our present considerations by looking at the following Lagrangian:

$$\mathcal{L}_m = -\tfrac{1}{4}G^{\,a}_{\mu\nu}G^{\,\mu\nu}_{a} - \tfrac{1}{2}m^2 B^a_\mu B^\mu_a + \hat{g} B^a_\mu j^{\,\mu}_{a} + \mathcal{L}_\psi + \mathcal{L}_\chi + \mathcal{L}_{\chi\psi}\,. \qquad (4.1)$$

This expression is actually (2.11) enlarged by the addition of the scalar field χ. The B^a_μ are coupled minimally to all matter fields, generically denoted by ψ, except χ. Thus, the j_a^μ in (4.1) is not necessarily the same as the j_a^μ appearing in (2.11). The latter is conserved whereas the former is not, if χ-B_μ^a couplings are present. $\mathcal{L}_\chi$ includes kinetic and self-interacting terms of the real field χ, and $\mathcal{L}_{\chi\psi}$ stands for all, if any, interaction terms between χ and the other matter fields.

Consider next the strictly gauge-invariant theory described by the following Lagrangian

$$\mathcal{L}_{\rm sym} = -\tfrac{1}{4} G'^a_{\mu\nu} G'^{\mu\nu}_a + \mathcal{L}_\psi - \tfrac{1}{2}\nabla_\mu \phi_k \nabla^\mu \phi_k^* + \hat{g} A'^a_\mu j'^\mu_a + V(\phi_k \phi_k^*), \tag{4.2}$$

where ∇_μ denotes the usual covariant derivative and the ϕ_k, $k=1,2$, are complex components of a scalar isospin field ϕ. For convenience, we introduce the notation

$$\phi = \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix}. \tag{4.3}$$

Also, $j'^a_\mu(x)$ in (4.2) must be a current which results from minimal-coupling terms.

Suppose now we make a decomposition of the massive vector fields B_μ^a into transverse and longitudinal parts. We adopt the infinitesimal version of this decomposition, strictly for convenience. We have

$$B_\mu^a = A_\mu^a + i\frac{g}{m} t_{abc}\, \xi_b A_\mu^c + \frac{1}{m}\partial_\mu \xi_a + O(\xi^2)\,, \tag{4.4}$$

where A_μ^a is a two-dynamical-component transverse vector field.

We shall now imitate the second of the two approaches in Ref. 6 based on identifying, as $m\to 0$, the transverse components of B_μ^a with gauge fields, i.e., we identify the A'^a_μ's in (4.2) with the A_μ^a's in (4.4). Such an identification is based on an analogy with the Abelian case where we know that for $m \neq 0$, but small, the transverse component of a massive vector field obeys the same equation as a massless gauge field. The existing difference in the zero-mass limit of non-Abelian fields do not affect the above conclusion. As there is now a one-to-one correspondence between dynamical degrees of freedom in (4.1) and (4.2), we make the pairings $\phi_a \sim \xi_a$, $a=1,2,3$, and $\phi_4 \sim \chi$. Since scalar fields are one-dimensional objects, we can write $\phi_a = \xi_a + \Delta\xi_a$, $\phi_4 = \chi + \Delta\chi$. These relations actually *define* $\Delta\xi$, $\Delta\chi$ (which must not be thought of as infinitesimal). For simplicity, we shall identify the ϕ_a with ξ_a, $a=1,2,3$, even though this is incorrect, strictly speaking. Our present conclusions are insensitive to the above simplification while our task becomes significantly lighter. In a forthcoming study we shall investigate in detail the zero-mass limit implications of non-Abelian fields in connection with the present discussion. For now we rely on our analysis in Ref. 6 for all assumptions and simplifications we are making.

From (4.1) we have the following equation of motion for B_μ^a:

$$\partial_\nu G^a_{\nu\mu} - m^2 B_\mu^a + \hat{g} j_\mu^a + ig t_{abc} B_\nu^c G^b_{\mu\nu} = 0\,. \tag{4.5}$$

Similarly, for A'^a_μ we obtain

$$\partial_\nu G'^a_{\nu\mu} + ig' t_{abc} A_\nu^c G'^b_{\mu\nu} + \hat{g}' j'^a_\mu + i\tilde{g}' \vartheta_{aij}\phi_j(\partial_\mu\phi_i + i\tilde{g}'\vartheta_{bik}\phi_k A'^b_\mu) = 0\,. \tag{4.6}$$

In writing (4.6) we have followed Weinberg's notation.[22] Note that in our specific case i, j, k run from 1 to 4.

Inserting decomposition (4.4) into (4.5) we obtain

$$\partial_\nu G^a_{\mu\nu} - m^2 A_\mu^a - igm t_{abc}\xi_b A_\mu^c - m\partial_\mu\xi_a - m^2 O(\xi^2) + \hat{g} j_\mu^a + ig t_{abc} A_\nu^c G^b_{\mu\nu} - \hat{g} g t_{abc} t_{cde} \xi_d A_\nu^e G^b_{\mu\nu} + ig t_{abc}\partial_\nu \xi^c G^b_{\mu\nu} + ig t_{abc} G^b_{\mu\nu} O(\xi^2) = 0. \tag{4.7}$$

We now identify the A_μ^a with the A'^a_μ, the ξ_a with the ϕ_r, $r=1,2,3$, and g with g'. It follows that $G^a_{\mu\nu} = G'^a_{\mu\nu}$ which leads to the identity

$$\hat{g}' j'_{\mu a} + \tilde{g}' \vartheta_{aij}\phi_j(\partial_\mu\phi_i + i\tilde{g}'\vartheta_{bik}\phi_k A_\mu^b) \equiv -m^2 A_\mu^a - igm t_{abc}\phi_b A_\mu^c - m\partial_\mu\phi_a - m^2 O(\phi^2) + \hat{g} j_\mu^a - \hat{g} g t_{abc} t_{cde}\phi_d A_\nu^e G^b_{\mu\nu} + ig t_{abc}\partial_\nu\phi^c G^b_{\mu\nu} + ig t_{abc} G^b_{\mu\nu} O(\phi^2). \tag{4.8}$$

We assume that the original theory (4.1) is stable about its vacuum so that $\langle A_\mu^a\rangle = \langle \phi^a\rangle = \langle\chi\rangle = 0$, $a, =1,2,3$. What remains is to examine $\langle\phi_4\rangle$ $(=\langle\Delta\chi\rangle)$. For this purpose we follow Higgs[25] and expand the various fields about their equilibrium points, i.e., $A_\mu^a \to \langle A_\mu^a\rangle + \delta A_\mu^a$, $\phi^k \to \langle\phi_k\rangle + \delta\phi_k$. Substituting in (4.8), retaining terms to only first order in the infinitesimal variations, and keeping only the vacuum expectation value (vev) $\langle\phi_4\rangle$ (whenever it appears) we obtain

$$\tilde{g}\vartheta_{ab4}\langle\phi_4\rangle\delta(\partial_\mu\phi_b) + \hat{g}'\langle\delta j'^a_\mu\rangle - \tilde{g}'^2(\vartheta_{bi4})^2\langle\phi_4\rangle^2\delta A_\mu^a \equiv -m^2\delta A_\mu^2 - m\delta(\partial_\mu\phi_a) + \hat{g}\langle\delta j_\mu^a\rangle\,, \tag{4.9}$$

where the notation $\langle \delta j^a_\mu \rangle$, denotes that δj^a_μ is to be evaluated at equilibrium values of the fields.

It is not hard to convince ourselves that

$$\langle \delta j^a_\mu \rangle = \langle \delta j'^a_\mu \rangle = 0 \ . \tag{4.10}$$

Indeed j'^a_μ involves only matter fields from the set ψ which have not been varied in the first place. On the other hand, j^a_μ would present problems only if it had a term linear in χ which would introduce a $\delta\phi_4$ factor. But, clearly, such a coupling would be of the form $\partial^\mu \chi A^a_\mu$, a fixed, which can easily be transformed away. Any other coupling involving χ could not possibly be linear in χ, in which case vanishing vev's, i.e., $\langle\chi\rangle$ (=0), would be present in front of the variations.

Taking into account the anti-Hermiticity properties of the ϑ_{aij} matrices[22] our identity finally assumes the form

$$(\tilde{g}\langle \phi_4\rangle - m)\delta(\partial_\mu \phi_a) + (\tilde{g}^2\langle\phi_4\rangle^2 - m^2)\delta A^a_\mu \equiv 0 \ . \tag{4.11}$$

Since the variations are arbitrary, their coefficients should vanish individually. We conclude

$$0 \neq \langle \phi_4 \rangle = m/\tilde{g} \ . \tag{4.12}$$

The following two points should now be noted.

(1) The infinitesimal nature of decomposition (4.4) is inconsequential to our proof. Indeed, terms in $O(\phi^2)$ would not have contributed to identity (4.9) even if they were not negligible. The reason is that they would have involved vanishing vev coefficients in front of each infinitesimal variation.

(2) The strictly gauge-invariant model (4.2) does not necessarily have to be restricted so as to admit only an isospin scalar doublet of the form (4.3). Any representation of SU(2) which involves four or more scalar fields would do. That it is necessary to have at least four becomes evident from (4.9). Indeed, if we identify the first three scalar fields in (4.2) with the longitudinal modes ξ_a, $a = 1, 2, 3$, we would need a fourth scalar with nonvanishing vev if identity (4.8) or (4.9) is to hold. Otherwise we would be led to the conclusion $m = 0$. If, on the other hand, we had more than four scalar fields, identity (4.9) would imply that at least one of the ϕ_k, $k \geq 4$, must have a nonvanishing vev for the (bare) mass m to be different from zero. Note that, in general, if (4.2) involves n scalar fields ϕ_k then we must introduce into (4.1) n-3 extra physical scalar fields.

In conclusion we see that a model involving massive non-Abelian vector fields cannot possibly be equivalent to a strictly gauge-invariant model unless the latter describes an SBGT. We have thus extended our previous findings for the Abelian case to the non-Abelian case. We can summarize our conclusions as follows.

Proposition. Consider a Lagrangian model involving massive non-Abelian vector fields B^a which belong to the adjoint representation of SU(n). In particular, there will be n^2-1 longitudinal modes [n^2-1 is the dimensionality of su(n)]. Suppose we add to this model k additional scalar fields ($k>0$) such that n^2+k-1 is equal to the real dimension of some complex representation of SU(n). Further, suppose the following:

(a) The B^a_μ are minimally coupled to all matter fields except to the k extra scalar fields.

(b) The transverse components of B^a_μ can be identified, in the zero-mass limit with non-Abelian gauge fields A^a_μ belonging to a strictly gauge-invariant field which includes the n^2-1 longitudinal (scalar) modes as part of an $\frac{1}{2}(n^2+k-1)$-dimensional complex representation of SU(n).

Then, the aforementioned strictly gauge-invariant model must correspond to an SBGT. Furthermore, assuming that all vector fields in the original model had the same (bare) mass m, we have that $m = g\langle\phi\rangle$, where g is the strength of the coupling of the n^2+k-1 charged scalar fields to the gauge fields and $\langle\phi\rangle$ is a factor built up from nonvanishing vev's of some or all of these scalar fields other than the n^2-1 which correspond to the longitudinal modes in the original model.

V. CONCLUDING REMARKS

In Sec. III we found that a necessary condition to lift the nonpolynomial factors which presented obstacles to the renormalizability of massive Yang-Mills Lagrangians is the introduction of superfluous scalar fields belonging to the adjoint representation of SU(2). In the preceding section we showed that the addition of extra physical scalar fields to a massive non-Abelian model such that there is a one-to-one correspondence in dynamical variables with a strictly gauge-invariant model forces the latter to correspond to an SBGT. Our feeling is that the simultaneous addition of the superfluous and the physical scalar fields corresponds to building up an SBGT in the R gauge. If this is correct, it means that we have been led to an R-gauge non-Abelian massive vector-field formulation of an SBGT starting from a model in which the mass was inserted by hand.

As already pointed out in the Introduction, our study as well as that of Ref. 7 investigate different aspects of converse, with respect to renormalizability, routes to SBGT. The ultimate objective is to formulate a uniqueness theorem regarding such theories. As far as our approach is concerned, the next step is to explicitly carry out a renormalization program for the models we have studied augmented by the addition of the extra scalar fields (both the superfluous and the physical ones). We shall not undertake any such program here. We only wish to comment briefly on what we think the role of the extra physical scalar fields of Sec. IV could be. After all, is it not enough to use the η_a's of Sec. III for the renormalizability program to go through?

Let us recall, at this point, Boulware's program[5] which motivated the present study. As he clearly shows in the case of the massive Abelian vector field and as he argues for the non-Abelian case, the diagonalization of the longitudinal kinetic terms leads to renormalizability in the sense of a finite number of necessary renormalizations. One is still left with an infinite, i.e., nonpolynomially bounded, wave-function renormalization. It is our feeling that the role of the additional physical scalar fields, which are necessary for the equivalence of a massive model to a strictly gauge-invariant model, is in fact to solve this problem.[26] It is mostly for this reason that it becomes important to carry out an explicit renormalization program which shows precisely the roles of the η_a's as well as those of the additional physical scalars of Sec. IV. We also remark that a transverse massive vector field—such as we obtain after the diagonalization—with its nonlocal propagator corresponds, in isolation by itself, to a nonunitary theory. The hope for rectifying this pathology lies in the presence of physical matter fields which couple to the transverse (massive) vector field. The physical scalar(s) of Sec. IV provide such an alternative.

A second point of curiosity, which also merits further investigation, is that both the work of Sec. III and that of Sec. IV seems to favor two unitary groups, namely U(1) and SU(2). We noted that the diagonalization involves $\frac{1}{2}m(m+1)$ conditions. The complexification method we employed involves $2m$ variables. Irrespective now of our interpretation of the condition $r=2m$ in (3.27), we see that the straightforward necessity requirement of having as many conditions as variables gives

$$2m=\tfrac{1}{2}m(m+1). \tag{5.1}$$

But (5.1) is satisfied only for $m=3$. Thus, ignoring for the moment gauge groups which are direct products (i.e. nonsemisimple), we see that the complexification method is especially tailored for SU(2). Furthermore, we notice that for $m=1$, $\frac{1}{2}m(m+1)=1$. This means that for the Abelian case, U(1), the longitudinal mode of the vector field does not need a partner in order to bring its kinetic term into a nonpolynomial well-behaved form. This is indeed what happens. On the other hand, we notice that for groups with $m\geq 4$ the complexification method is insufficient unless $G_{ab}(\xi)$ possesses other symmetries which reduce the number $\frac{1}{2}m(m+1)$.

In a similar way, our approach in Sec. IV also seems to favor U(1) and SU(2), however, not in as prohibitive a way as theorem 3.1. To be explicit, consider the unitary group SU(n). Its definition as a group of transformations is the following. An element of SU(n) is a unimodular matrix which leaves invariant the quadratic form

$$|z_1|^2+|z_2|^2+\cdots+|z_n|^2\,, \tag{5.2}$$

where the z_i are complex numbers, i.e., coordinates of an n-dimensional complex vector. Let us now note that the most basic terms of Lagrangians (kinetic and mass terms) are bilinear forms. Furthermore, if a given Lagrangian possesses an internal symmetry, then such bilinear forms must be left invariant by the action of the underlying symmetry group. Suppose we restrict ourselves to global unitary transformation on the scalar sector of a field Lagrangian. Substituting the complex numbers by complex fields it follows that the natural action of SU(n) is given on an n-tuplet of complex scalar fields. Now, a complex m-dimensional representation of SU(n) would still involve Lagrangian kinetic terms which make up the bilinear form

$$-\tfrac{1}{2}(\partial_\mu\phi_1\partial_\mu\phi_1^*+\cdots+\partial_\mu\phi_m\partial_\mu\phi_m^*). \tag{5.3}$$

We note that such a bilinear form is, in fact, also left invariant by the natural action of the group SU(n) which is meant to be the symmetry group rather than SU(m). Suppose now that the group action is localized, i.e., we pass to the gauge group. One would only have to look at the number of gauge fields in order to read the internal group of symmetries. This number should equal the dimensionality of su(n). We know that

$$\dim[\mathrm{u}(n)]=n^2,\quad \dim[\mathrm{su}(n)]=n^2-1. \tag{5.4}$$

We now recall that the premises of our proposition requires the introduction of k scalar fields, $k\geq 1$, in addition to the n^2-1 longitudinal fields. Furthermore, k must be such that $\frac{1}{2}(n^2+k-1)$ is the complex dimension of some representation of the gauge group. Since the natural representation of a unitary group can be so conveniently reflected in the kinetic terms, we may ask whether the equations, for unitary and unimodular unitary, respectively,

$$\frac{n^2+k}{2}=n,\quad \frac{n^2+k-1}{2}=n\quad (k\geq 1), \tag{5.5}$$

always have solution. In fact, they do not except for U(1) and SU(2). In both cases we get $k=1$. For higher unitary groups $\frac{1}{2}(n^2-1)$ is by itself already larger than the dimension of the natural representation of SU(n). Thus, adding the extra k scalar fields forces the strictly gauge-invariant model to incorporate a scalar multiplet belonging to a representation of a higher dimensionality than the natural one.

As a closing comment we wish to say that a perhaps more significant aspect to our program, provided it solidifies our present conclusions once it is completed, could be the following. The Higgs mechanism,[25] in conjunction with spontaneous symmetry breaking, has played a key role in gauge theoretical attempts toward unified field theories. Whether a mechanism can be divorced from a physical process is not totally clear. There have, in fact, been questions voiced regarding the fundamental meaning, if any, of the Higgs mechanism.[27] However, the route we are adopting does not have to employ the Higgs mechanism as such; rather, we show "equivalence" to it and SBGT, in general, once renormalizability and gauge content have been required of theories which contain massive vector fields.

ACKNOWLEDGMENTS

The author would like to thank Professor Abdus Salam, the International Atomic Energy Agency, and UNESCO for hospitality at the International Centre for Theoretical Physics, Trieste.

---

*Present address: University of Athens, Department of Theoretical Physics, Athens, Greece.

[1]G. 't Hooft, Nucl. Phys. B35, 167 (1971).

[2]B. W. Lee and J. Zinn-Justin, Phys. Rev. D 5, 3121 (1972); 5, 3137 (1972); 8, 4654(E) (1973); 5, 3155 (1972).

[3]We are ignoring, by this statement, gauge-fixing terms. These are, in fact, introduced after the redefinition of the scalar fields, i.e., after effecting spontaneous symmetry breaking.

[4]M. Veltman, Nucl. Phys. B7, 637 (1968); B21, 288 (1970).

[5]D. G. Boulware, Ann. Phys. (N.Y.) 56, 140 (1970).

[6]C. N. Ktorides, Nucl. Phys. B97, 256 (1975).

[7]J. M. Cornwall, D. N. Levin, and G. Tiktopoulos, Phys. Rev. D 10, 1145 (1974).

[8]By renormalization in the true sense we mean a finite number of renormalizations *and* a finite number of subtractions for each renormalization. Thus, the appearance of divergences of unlimited degree as, e.g., in the case of massive electrodynamics with a conserved current (see Ref. 5) corresponds, from our point of view, to "quasi-renormalizability."

[9]The following should be noted in connection with the form of the propagator for a massive vector field. We restrict our comments, for the most part, to the case when no matter fields are present, i.e. (2.10). Propagator (2.14) can be thought of as that of a gauge field in the Landau gauge with $(p^2-m^2)^{-1}$ replacing $1/p^2$. Now the presence of $(p^2-m^2)^{-1}$ supposes that the $B_\mu^a(x)$ belong to an irreducible representation of the Poincaré group, or, equivalently, that they describe particles of spin 1. This means that the B_μ^a must satisfy the Klein-Gordon equation $(\Box-m^2)B_\mu^a=0$. Otherwise each B_μ^a would describe a spin-1 and a spin-0 particle at the same time. Recall that in the Abelian case (Proca field) the operator condition $\partial_\mu B^\mu=0$ is imposed on a massive vector field. This must be so because the equation of motion for B_μ is $\partial_\nu F_{\nu\mu}-m^2B_\mu=0$, so that the aforementioned (Proca) condition is necessary to make the last relation equivalent to $(\Box-m^2)B_\mu=0$. In the presence of interactions, $\partial_\mu B^\mu=0$ should be viewed as an expectation-value condition rather than as an operator identity; see, e.g., F. Coester, Phys. Rev. 83, 798 (1951). In the non-Abelian case the equation of motion is $\partial_\nu G_{\nu\mu}^a+igt_{abc}B_\nu^cG_{\mu\nu}^b-m^2B_\mu^a=0$. For this equation to be equivalent to $(\Box-m^2)B_\mu^a=0$, an operator condition (expectation-value condition in the presence of interactions) should be imposed on B_μ^a which can easily be obtained. Henceforth, we shall assume that the presence of $(p^2-m^2)^{-1}$ in the propagator of massive non-Abelian vector fields is justified via the imposition of an appropriate condition.

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

[11]We shall be more explicit on these matters in Sec. III, where we shall employ differential geometric language to describe the manifold of the longitudinal fields $\xi_a(x)$.

[12]E. G. C. Stückelberg, Helv. Phys. Acta 11, 299 (1938).

[13]In other words, the whole theory is quantized before the splitting takes place. Thus, if $\mathcal{K}_\mu$ denotes the source function for $\mathcal{B}_\mu$, there will be a term of the form $K_\mu^a(A_\mu^a+Q^{ab}\partial_\mu\xi^b)$ after the splitting. To obtain (2.29) we combine the longitudinal part of this term along with the term $m^2\partial_\mu\xi_bQ_{ba}A_\mu^a$ appearing in (2.25).

[14]Admittedly, the term $\frac{1}{4}\mathcal{G}_{\mu\nu}\mathcal{G}^{\mu\nu}$ is missing from $\mathcal{L}_\xi$. However, we are only integrating over the ξ_a's. The vector-field kinetic term is already present in the other piece of our theory, i.e., in $\mathcal{L}_1$.

[15]Given a Lie group G, say SU(n), we shall reserve the capital letters to refer to the group itself whereas the respective algebra g will be denoted by small letters, e.g. su(n) stands for the algebra of SU(n). We shall refrain from repeating the word local each time we are referring to a local group or algebra.

[16]For unitary groups (more generally, general linear groups) the exponential mapping coincides with the familiar exponential expansion $e^A=1+A+\frac{1}{2}A^2\ldots$; see, e.g. S. Helgason, *Differential Geometry and Symmetric Spaces* (Academic, New York, 1962), p. 101.

[17]See, e.g., p. 96 of the reference quoted in footnote 16.

[18]Reducing, on the other hand, the rank of G would mean a reduction in the number of dynamical variables.

[19]R. M. Santilli and C. N. Ktorides, Phys. Rev. D 7, 2447 (1973).

[20]We would like to emphasize that the explicit solution which will be given for the SU(2) case is independent of this interpretation. Only its uniqueness depends on this way of looking at the unique condition $r = 2m$ in (3.27).

[21]We do not mean ghosts in the sense of Faddeev and Popov in any case.

[22]S. Weinberg, Phys. Rev. D 7, 1068 (1973).

[23]R. Utiyama, Phys. Rev. 101, 1597 (1956).

[24]The characterization "manifestly" means that we are excluding from consideration dynamical mechanisms for generation of mass; see, e.g., J. M. Cornwall and R. E. Norton, Phys. Rev. D 8, 3338 (1973); J. M. Cornwall, *ibid.* 10, 500 (1974); R. Jackiw and K. Johnson, *ibid.* 8, 2386 (1973).

[25]P. W. Higgs, Phys. Rev. Lett. 13, 508 (1964).

[26]Compare, e.g., with the model in Sec. VI of Ref. 1. 't Hooft's Z may very well be the counterpart of our scalar field χ.

[27]See, e.g., J. Sucher and C. H. Woo, Phys. Rev. D 8, 2721 (1973).