CALIFORNIA INSTITUTE OF TECHNOLOGY

Synchrotron Laboratory

Pasadena, California

GAUGE THEORIES OF VECTOR PARTICLES[*]

Sheldon L. Glashow

and

Murray Gell-Mann

April 24, 1961

$AT(11-1)-68$

DISCLAIMER

DISCLAIMER

Portions of this document may be illegible
in electronic image products. Images are
produced from the best available original
document.

Contents

ABSTRACT

The possibility of generalizing the Yang-Mills trick is examined. Thus we seek theories of vector bosons invariant under continuous groups of coordinate-dependent linear transformations. All such theories may be expressed as superpositions of certain "simple" theories; we show that each "simple" theory is associated with a simple Lie algebra. We may introduce mass terms for the vector bosons at the price of destroying the gauge-invariance for coordinate-dependent gauge functions.

The theories corresponding to three particular simple Lie algebras -- those which admit precisely two commuting quantum numbers -- are examined in some detail as examples. One of them might play a role in the physics of the strong interactions if there is an underlying super-symmetry, transcending charge independence, that is badly broken.

The intermediate vector boson theory of weak interactions is discussed also. The so-called "schizon" model cannot be made to conform to the requirements of partial gauge-invariance. It is possible, however, to find a formal theory of four intermediate bosons that are partially gauge-invariant and gives an approximate $\left| \Delta I \right| = 1/2$ rule.

I. Introduction

The electromagnetic interaction of elementary particles is remarkably simple. It is of universal strength and form and is associated with a principle of gauge invariance. In fact, starting with the idea of invariance under gauge transformations with coordinate - dependent gauge functions, one can deduce the existence of a massless vector field coupled to a conserved current. If all charged fields are subjected to the same gauge transformation, then the electric charges of all particles are the same.

The fact that the weak interactions are vectorial in character (apart from nonconservation of parity) and nearly universal in strength has suggested to many physicists that they may be mediated by vector fields (1,2) and that there may be a useful parallel between them and electromagnetism, perhaps even extending to the notion of gauge invariance (3,4,5,6).

The strong interactions, too, seem to exhibit some degree of universality. Moreover, the approximate conservation laws of isotopic spin and of strangeness, as well as the exact law of conservation of baryons, present an analogy with the conservation of charge and suggest that some principles of gauge invariance may be at work. Until recently, it seemed that the strong couplings were not vectorial, but there is mounting evidence that there are objects (like the $I = 1$, $J = 1$, $\pi\pi$ resonance) that can be interpreted as vector mesons and that may play a very significant role in the strong interactions (7,8).

There are two great difficulties in the way of constructing theories of weak and strong interactions by analogy with electrodynamics. One is that some of the relevant currents are not conserved. The isotopic spin and strangeness currents that may enter into a vectorial theory of the strong couplings fail to be conserved on account of electromagnetic and weak interactions, while the conservation of the weak current is broken not only by electromagnetism but, in the case of the axial vector and strangeness-changing parts, by masses and perhaps by strong interactions as well.

The other difficulty is that whereas photons are massless (as the quanta must be in a theory that is fully gauge invariant with a coordinate-dependent gauge function) the vector particles that mediate the strong and weak interactions must be massive if they exist at all.

Thus the notion has arisen (3,4,5,6,7,8) of a theory that is partially gauge-invariant. In each case we have a Lagrangian like the electromagnetic one, fully invariant under coordinate-dependent gauge transformations, plus other terms. The remaining terms are of two kinds:

a) those which break the full gauge invariance, while leaving intact the conservation law and the invariance under constant gauge transformations;

b) those which destroy the gauge invariance altogether, along with the conservation law.

In the case where the conservation law is exact (conservation of baryons) the terms of type b) are, of course, absent.

Now the idea of partial gauge-invariance poses a number of questions, to which we shall return briefly in Section VII. For the moment, let us concentrate on the straightforward part of the problem, the construction of the fully gauge-invariant part of the theory.

The coupling of a vector meson field to a single quantity like baryon number follows exactly the pattern of electromagnetic coupling to the charge, as long as the complete gauge-invariance is maintained. But, when we go over to the case of three non-commuting quantities like the components of the isotopic spin current, the situation becomes different and a more sophisticated theory becomes necessary. The intermediary vector meson field now carries isotopic spin 1 and its own isotopic spin current contributes a source term. Thus the theory of the vector meson field becomes non-linear. The problem of constructing the theory in question has been solved by Yang and Mills (9) and by Shaw (10).

In the next two sections, we review the simple case of charge or baryon number and the more complicated case of isotopic spin. Then, in Section IV, we go on to the main point of this article -- the description of all possible straightforward generalizations of the Yang-Mills trick. We are interested in such generalizations because we do not know, for either the strong or the weak interactions, exactly how many intermediate vector fields may be involved (if any). To give just one example, it has been suggested (11,12,13) that there may be four such (hermitian) fields for the weak interactions -- the

so-called schizon model, set up to give $|\Delta \underline{I}| = 1/2$ and $\Delta S = 0, \pm 1$
for the nonleptonic weak interactions of baryons and mesons. We shall
show in Section VII that the ideas of partial gauge-invariance lead to
severe restrictions on four-field models; in fact, the restrictions are
so strong as to make it impossible to construct the schizon model
according to the gauge principles of this article.

The classification of generalized Yang-Mills theories discussed
in Section IV is described further in Section V; some examples are given
in Section VI; and some possible physical applications are touched on
briefly in Section VII.

II. The One-Parameter Gauge Theory

The gauge formalism of electromagnetism is, of course, well-
known. The generalization from charge to baryon number was discussed
by Yang and Lee (14); it is clear from their work that the generalization
contradicts experiment unless either the coupling constant is ridi-
culously small or the gauge invariance is broken, say by a mass term
for the vector field. Let us review the method.

We start with an additive quantity like charge or baryon number;
call it Q. Let the fields $\psi_a(x)$ destroy particles of charge Q_a and
create their antiparticles. We then discuss invariance under the
infinitesimal gauge transformations

$$\psi_a(x) \longrightarrow \psi_a(x) - i Q_a \Lambda(x) \psi_a(x) \qquad . \qquad (2.1)$$

Whenever the coordinate derivative ∂_α acts on ψ_a, it undergoes the transformation

$$\partial_\alpha \longrightarrow \partial_\alpha - i\, Q_a\, \partial_\alpha \Lambda \qquad \text{(on } \psi_a) \quad . \qquad (2.2)$$

In order to cancel this change, we introduce a vector field $A_\alpha(x)$ that suffers the gauge transformation

$$A_\alpha(x) \longrightarrow A_\alpha(x) - \partial_\alpha \Lambda(x) \qquad\qquad (2.3)$$

and a field Lagrangian density L_A invariant under this transformation, say

$$L_A = - 1/4 \, (\partial_\alpha A_\beta - \partial_\beta A_\alpha)^2 \qquad . \qquad (2.4)$$

In the absence of the field A_α and its couplings, let the Lagrangian be $L_o(\psi_a)$ and let it conserve Q. Then the "minimal" gauge-invariant Lagrangian including A_α is

$$L = \widetilde{L}_o \, (\psi_a) + L_A \qquad , \qquad\qquad (2.5)$$

where $\widetilde{L}_o$ is obtained from L_o by the replacement

$$\partial_\alpha \longrightarrow \partial_\alpha - i\, Q_a\, A_\alpha(x) \qquad \text{(on } \psi_a) \quad . \qquad (2.6)$$

It is evident that (2.5) gives us a gauge-invariant Lagrangian and certainly the procedure described by (2.5) is the usual one. But what do we mean by "minimal"? The point is that we could add to the Lagrangian (2.5) further gauge-invariant terms involving the field

strength $\partial_\alpha A_\beta - \partial_\beta A_\alpha$. However, nature, in the case of electromagnetism, does not seem to make use of such terms.

Consider, for example, a Dirac particle of charge e, for which ψ is a spinor and the free Lagrangian density is

$$L_o = - \bar{\psi} (\gamma_\alpha \partial_\alpha + m_o)\psi \qquad . \qquad (2.7)$$

The substitution (2.6) gives the usual coupling ,

$$i e \bar{\psi} \gamma_\alpha A_\alpha \psi \qquad , \qquad (2.8)$$

but no Pauli moment. We generally suppose that the effective Pauli moments of nucleons arise from the ordinary electrical interaction of the meson cloud around the nucleon and not from a basic Pauli moment term in the Lagrangian:

$$i \mu \bar{\psi} \sigma_{\alpha\beta} \psi (\partial_\alpha A_\beta - \partial_\beta A_\alpha) \qquad . \qquad (2.9)$$

Hence the attempt (15) to state a principle of minimal electromagnetic interaction, that the electromagnetic field interacts only with electric charges in the normal way (as in (2.6)) and not through special field-dependent terms like (2.9) in the basic Lagrangian.

The difficulty (16,17) with any attempt to put the idea of minimal electromagnetic interaction in definite mathematical form is the following. Various Lagrangian densities (differing by divergences of four-vectors) can lead to the same equations of motion. But if we choose in this way a new L_o, the resulting electromagnetic coupling (and the equations of motion including electromagnetism) may become radically

different. Thus we can obtain the Pauli moment term (2.9) by the "minimal" procedure (2.6) if we just add to the usual L_o in (2.7) the term

$$\frac{2\mu}{e} \quad \partial_\alpha \, (\bar{\Psi} \, \sigma_{\alpha\beta} \partial_\beta \, \psi) \qquad . \tag{2.10}$$

We see that the procedure (2.6) defines the "minimal" interaction only if the original Lagrangian density L_o is chosen in a "minimal" way. We must assign a physical meaning to L_o and say that (2.7) describes a Dirac particle properly, while if the term (2.10) is added we obtain the wrong Lagrangian density for a Dirac particle, even though the equation of motion without electromagnetism is just the Dirac equation in both cases.

Of course we have still not specified in a clear-cut way how to find the "minimal" L_o in all cases. But that difficulty is not restricted to the problem of electromagnetic couplings. Even without electromagnetic interactions and without strong and weak interactions, we must still assign a physical significance to L_o because it determines the gravitational coupling. If we add a term like (2.10) to L_o and follow the usual procedure for constructing the stress-energy-momentum tensor, we will get a different answer. In fact, the gravitational interactions are constructed from L_o in a way that is closely analogous to the method given in (2.6) for electromagnetism.

Now let us return to the theory described by the Lagrangian density (2.5). The equation of motion for the field A_α is

$$\Box^2 A_\alpha - \partial_\alpha \partial_\beta A_\beta = - D \widetilde{L}_o(\psi_a)/D A_\alpha = - j_\alpha \quad , \qquad (2.11)$$

where D/DA_α is the Lagrangian derivative $\partial/\partial A_\alpha - \partial_\beta \partial/\partial(\partial_\beta A_\alpha)$.[1]
The formula for the current can be re-expressed as follows. Consider
a gauge transformation in which the ψ fields are affected as in (2.1)
but A_α is not transformed. Denote partial derivatives with respect to
Λ and $\partial_\alpha \Lambda$ under this condition by $\left[\partial/\partial\Lambda\right]_{\delta A=0}$ and $\left[\partial/\partial(\partial_\alpha \Lambda)\right]_{\delta A=0}$.
Then we remark that since $\widetilde{L}_o$ is totally gauge-invariant, the derivative
$\left[\partial/\partial(\partial_\alpha \Lambda)\right]_{\delta A=0}$ has the effect of the negative of a derivative with
$\delta\psi_a = 0$ and only A_α affected by the gauge transformation. But such a
negative derivative is exactly $-D/DA_\alpha$. Thus we have the result

$$j_\alpha = \left[\partial \widetilde{L}_o/\partial(\partial_\alpha \Lambda)\right]_{\delta A=0} \quad . \qquad (2.12)$$

The current is calculated from the Lagrangian (either $\widetilde{L}_o$ or L) by a
gauge transformation involving only the ψ_a fields and not A_α.

Next we note (6) that in any local gauge transformation, the
Euler-Lagrange equation applies to the gauge function, even though it
is not a field variable, as a consequence of the Euler-Lagrange equa-
tions for the field variables themselves. Thus we have

$$\partial_\alpha j_\alpha = \partial_\alpha \left[\partial \widetilde{L}_o/\partial(\partial_\alpha \Lambda)\right]_{\delta A=0} = \left[\partial \widetilde{L}_o/\partial\Lambda\right]_{\delta A=0}$$

$$= \partial \widetilde{L}_o/\partial\Lambda \quad . \qquad (2.13)$$

But the Lagrangian is invariant under gauge transformations with con-
stant gauge function. Therefore the current is conserved:

$$\partial_\alpha j_\alpha = \partial \widetilde{L}_o / \partial \Lambda = 0 \qquad . \tag{2.14}$$

Looking back at the equation of motion (2.11), we see that the supplementary condition $\partial_\alpha A_\alpha = 0$ may be imposed.

Finally, we may identify the constant of the motion $-i \int j_4 \, d^3x$ with the charge Q. So far we have looked at the equations classically; but in quantum mechanics, of course, Q is an operator and has the commutation relations

$$\left[\psi_a, Q \right] = Q_a \psi_a \qquad . \tag{2.15}$$

Now that we have sketched the fully gauge-invariant theory, we may discuss what happens when a term is added to L that breaks the full gauge invariance but leaves the invariance under gauge transformations of the first kind, that is, with constant Λ. We shall take the simple case of a mass term for the vector meson

$$- \mu_o^2 \, A_\alpha \, A_\alpha / 2 \qquad .$$

Evidently all that happens is that the equation of motion (2.11) becomes

$$(\Box^2 - \mu_o^2) \, A_\alpha - \partial_\alpha \partial_\beta A_\beta = - j_\alpha \qquad , \tag{2.16}$$

while the expression (2.12) for the current and the conservation law (2.14) remain unchanged. We have a vector meson coupled to a conserved current in a "partially gauge-invariant" theory.

III. The Three-Parameter Gauge Theory of Yang and Mills

We now turn from the simple case of charge or baryon number to the case of the isotopic spin $\underline{I}$, obeying the commutation relations

$$\left[I_i, I_j \right] = i\, e_{ijk}\, I_k \quad . \tag{3.1}$$

This time our fields ψ carry isotopic spin; let us consider for simplicity a field N of isotopic spin 1/2 (the nucleon) and a field $\underline{\pi}$ of isotopic spin 1 (the pion)(9). The relations analogous to (2.15) are

$$\left[N, I_i \right] = \tau_i N/2 \quad ,$$

$$\left[\pi_j, I_i \right] = -i\, e_{ijk}\, \pi_k \quad . \tag{3.2}$$

The infinitesimal gauge transformations analogous to (2.1) are then

$$N(x) \longrightarrow N(x) - i\, \gamma_o\, \underline{\tau} \cdot \underline{\Lambda}(x)\, N(x) \quad ,$$

$$\underline{\pi}(x) \longrightarrow \underline{\pi}(x) + 2\gamma_o\, \underline{\Lambda}(x) \times \underline{\pi}(x) \quad . \tag{3.3}$$

(We have denoted by γ_o the bare coupling parameter.) Thus the coordinate derivative ∂_α acting on the fields N and $\underline{\pi}$ suffers the change

$$\partial_\alpha \longrightarrow \partial_\alpha - i\, \gamma_o\, \underline{\tau} \cdot \partial_\alpha \underline{\Lambda} \qquad (\text{on } N) \ ,$$

$$\partial_\alpha \longrightarrow \partial_\alpha + 2\gamma_o\, \partial_\alpha \underline{\Lambda} \times \qquad (\text{on } \underline{\pi}) \ . \tag{3.4}$$

corresponding to (2.2).

To form a gauge-invariant theory, we must introduce a vector field $\underline{A}_\alpha(x)$ with isotopic spin one; its gauge transformation is

essentially different from (2.3) because A_α carries isotopic spin, whereas the photon carries no charge. Thus the field is not only displaced by the gradient of Λ but also undergoes isotopic rotation as $\pi(x)$ does in (3.3). The gauge transformation is thus

$$A_\alpha(x) \longrightarrow A_\alpha(x) - \partial_\alpha \Lambda(x) + 2\gamma_0 \, \Lambda(x) \times A_\alpha(x) \qquad . \qquad (3.5)$$

For the field Lagrangian density, we must choose an expression invariant under this gauge transformation. We note that

$$G_{\alpha\beta} \equiv \partial_\alpha A_\beta - \partial_\beta A_\alpha + 2\gamma_0 \, A_\alpha \times A_\beta \qquad\qquad (3.6)$$

transforms according to the rule

$$G_{\alpha\beta} \longrightarrow G_{\alpha\beta} + 2\gamma_0 \, \Lambda \times G_{\alpha\beta} \qquad . \qquad (3.7)$$

The simplest gauge-invariant Lagrangian is thus

$$L_A = - 1/4 \, G_{\alpha\beta} \cdot G_{\alpha\beta} \qquad , \qquad\qquad (3.8)$$

which is, of course, nonlinear, unlike (2.4). In the equation of motion deducible from (3.8), the source of the A field is its own isotopic spin current.

Now, given a Lagrangian $L_0(N,\pi)$ not involving the A field but conserving isotopic spin, we can introduce the "minimal" gauge-invariant Lagrangian density including A:

$$L = \widetilde{L}_0 \, (N,\pi) + L_A \qquad , \qquad\qquad (3.9)$$

with $\widetilde{L}_0$ obtained from L_0 by the substitutions

$$\partial_\alpha \rightarrow \partial_\alpha - i\,\gamma_0\,\underset{\sim}{T}\cdot\underset{\sim}{A}_\alpha \qquad\qquad \text{(on N)}$$

$$\partial_\alpha \rightarrow \partial_\alpha + 2\gamma_0\,\underset{\sim}{A}_\alpha\,x \qquad\qquad \text{(on } \underset{\sim}{\pi}) \qquad (3.10)$$

analogous to (2.6).

The current, source of the field $\underset{\sim}{A}_\alpha$, is given, exactly as in (2.12), by the formula

$$2\gamma_0\,\underset{\sim}{\mathcal{L}}_\alpha = \left[\partial\widetilde{L}_0 \big/ \partial(\partial_\alpha \underset{\sim}{\Lambda})\right]_{\delta A=0} \qquad . \qquad (3.11)$$

and is conserved, while the analog of the charge is just

$$-i\,2\gamma_0\int\underset{\sim}{\mathcal{L}}_4\,d^3x = 2\gamma_0\,\underset{\sim}{T} \qquad . \qquad (3.12)$$

If we now add a common mass term for the three kinds of vector meson,

$$-\frac{\mu_0^2}{2}\,\underset{\sim}{A}_\alpha\cdot\underset{\sim}{A}_\alpha \qquad\qquad ,$$

the gauge invariance is broken (except for constant $\underset{\sim}{\Lambda}$) but the isotopic spin current is still conserved. Unfortunately the renormalizability of the theory, at least in the conventional sense, is lost when a mass is added (18,19).

IV. Generalizations

We now come to grips with out problem, that of classifying the straightforward generalizations of the Yang-Mills trick. We imagine sets of N fields, like the two kinds of nucleon or the three kinds of pion in Section III, on which a gauge operation performs a linear

transformation as in (3.3). We may write

$$\psi_i(x) \qquad \psi_i(x) - 2i\gamma_o \sum_{j=1}^{n} \sum_{k=1}^{N} M_j^{ik} \Lambda_j(x) \psi_k(x) \qquad (4.1)$$

for our generalization. The n independent gauge functions $\Lambda_j(x)$ may be taken real, while the M_j are, for the moment, arbitrary complex N x N matrices.

The Lagrangian density $L_o(\psi_i)$ is presumed invariant under (4.1) for constant gauge functions Λ_j. Then (4.1) must be an infinitesimal unitary operation; the matrices M_j must be hermitian. The coordinate derivative acting on ψ changes according to the rule

$$\partial_\alpha \to \partial_\alpha - 2i\gamma_o \sum_{j=1}^{n} M_j (\partial_\alpha \Lambda_j) \qquad . \qquad (4.2)$$

To cancel this change, we introduce n hermitian fields $A_{\alpha i}$ to take up the gauges Λ_i. In place of (3.5) we have

$$A_{\alpha i}(x) \to A_{\alpha i}(x) - \partial_\alpha \Lambda_i(x) + 2\gamma_o \sum_{j,k=1}^{n} c_{ijk} \Lambda_j(x) A_{\alpha k}(x) \qquad , \qquad (4.3)$$

where all the indices in c_{ijk} run from 1 to n. The c's must be real to preserve the hermiticity of the A fields. We must determine the properties of c_{ijk} that will permit the Yang-Mills trick to go through.

First of all, we must be able to find a gauge-invariant field Lagrangian for the $A_{\alpha i}$. We seek a field strength that transforms simply, like $G_{\alpha\beta}$ in (3.7):

$$G_{\alpha\beta i} \equiv \partial_\alpha A_{\beta i} - \partial_\beta A_{\alpha i} + 2\gamma_o \sum_{j,k=1}^{n} b_{ijk} A_{\alpha j} A_{\beta k} \qquad . \qquad (4.4)$$

The transformation is, in general, very complicated. Let us use the summation convention. We obtain

$$G_{\alpha\beta i} \rightarrow G_{\alpha\beta i} + 2\gamma_o (c_{ijk} - b_{ijk}) \partial_\alpha \Lambda_j A_{\beta k} - 2\gamma_o (c_{ijk} + b_{ikj}) \partial_\beta \Lambda_j A_{\alpha k}$$

$$+ 2\gamma_o c_{ijk} \Lambda_j (\partial_\alpha A_{\beta k} - \partial_\beta A_{\alpha k})$$

$$+ 4\gamma_o^2 (b_{ijm} c_{jk\ell} + b_{i\ell j} c_{jkm}) \Lambda_k A_{\alpha\ell} A_{\beta m} \qquad . \qquad (4.5)$$

In order to obtain a law analogous to (3.7), we must put

$$b_{ijk} = c_{ijk}, \ b_{ikj} = -c_{ijk}, \ b_{ijm} c_{jk\ell} + b_{i\ell j} c_{jkm} = c_{ikj} b_{j\ell m} \qquad .$$

Then we have

$$G_{\alpha\beta i} \rightarrow G_{\alpha\beta i} + 2\gamma_o c_{ijk} \Lambda_j G_{\alpha\beta k} \qquad , \qquad (4.6)$$

with

$$G_{\alpha\beta i} = \partial_\alpha A_{\beta i} - \partial_\beta A_{\alpha i} + 2\gamma_o c_{ijk} A_{\alpha j} A_{\beta k} \qquad , \qquad (4.7)$$

$$c_{ijk} = - c_{ikj} \qquad , \qquad (4.8)$$

$$c_{ijm} c_{jk\ell} + c_{i\ell j} c_{jkm} + c_{ikj} c_{jm\ell} = 0 \qquad . \qquad (4.9)$$

Now the Lagrangian density

$$L_A = -\frac{1}{4} G_{\alpha\beta i} G_{\alpha\beta i} \qquad (4.10)$$

will indeed be gauge-invariant provided we have

$$c_{ijk} = - c_{kji} \qquad . \qquad (4.11)$$

The necessary and sufficient conditions for the construction of the generalized Yang-Mills field are thus:

$$c_{ijk} \text{ totally antisymmetric and real,} \tag{4.12a}$$

$$c_{ijm} c_{k\ell j} + c_{kjm} c_{\ell ij} + c_{\ell jm} c_{ikj} = 0 \quad . \tag{4.12b}$$

The Yang-Mills theory itself is the special case in which n = 3 and $c_{ijk} = e_{ijk}$, which obviously satisfies (4.12).

Now we must couple the field $A_{\alpha i}$ to the current generated by the gauge transformation (4.1) of the ψ_i. We have to construct from $L_0(\psi_i)$ a completely gauge-invariant quantity $\widetilde{L}_0(\psi_i)$. In order to make the gauge transformations (4.2) and (4.3) compensate each other, we use the prescription analogous to (3.10),

$$\partial_\alpha \longrightarrow \partial_\alpha - 2i\,\gamma_0 \sum_{i=1}^{n} M_i A_{\alpha i} \quad (\text{on } \psi) \tag{4.13}$$

to construct $\widetilde{L}_0$ from L. Under the unitary transformation (4.1), which we may rewrite in the form

$$\psi \rightarrow \psi - 2i\,\gamma_0 \,(\sum_j M_j \Lambda_j)\,\psi \quad ,$$

the M's transform according to the rule

$$M_i \longrightarrow M_i - 2i\,\gamma_0 \sum_j \left[M_i, M_j\right] \Lambda_j \quad , \tag{4.14}$$

while ∂_α and $A_{\alpha i}$ transform as in (4.2) and (4.3) respectively. Thus the prescription (4.13) yields a gauge-invariant Lagrangian density $\widetilde{L}_0$ if and only if we have

$$\left[M_i, M_j \right] = i \, c_{ijk} M_k \qquad , \qquad (4.15)$$

with the summation convention understood. Evidently this is the generalization of the commutation rule (3.1) for the isotopic spin.

In order that the c_{ijk} define the commutation relation (4.15), it must obey just two conditions. First, the rule

$$\left[M_i, M_j \right] = - \left[M_j, M_i \right] \qquad (4.16)$$

tells us that c_{ijk} must be antisymmetric in i and j; but we already know that from (4.12a). Second, the Jacobi identity

$$\left[M_i, \left[M_j, M_k \right] \right] + \left[M_j, \left[M_k, M_i \right] \right] + \left[M_k, \left[M_i, M_j \right] \right] = 0 \qquad (4.17)$$

gives us just (4.12b).

There remains the condition that c_{ijk} be antisymmetric not only in i and j but in the other pairs of indices as well. We shall return to the consequences of this further condition shortly.

Suppose we can divide the indices k into two sets such that $c_{ijk} = 0$ whenever i belongs to one set and j to the other. Then the fields $A_{i\alpha}$ of one set and those of the other set are completely unconnected to each other by any of the gauge transformations we have discussed. Likewise, the operators M_i belonging to one set of indices commute with those belonging to the other set. We are then dealing with a linear superposition of two completely independent Yang-Mills theories, which may have vastly different coupling strengths and no direct physical connection. We might as well restrict our attention

to one of these.

We can go further. Suppose our theory is not simplifiable as just discussed. We may apply any real rotation in the n-dimensional space of the $A_{i\alpha}$, rotating at the same time the M_i and the gauges Λ_i. The properties (4.12) of c_{ijk} are unaffected by such a rotation.) It may turn out then that our theory is simplifiable. In that case, let us restrict our attention to one of the parts. We continue this process until we reach an irreducible Yang-Mills theory, one for which we cannot, no matter how we rotate in the n-dimensional space, find two sets of indices that are unconnected by the c_{ijk}. From now on, we shall deal with these unsimplifiable or "simple" theories, from which the most general theory can be built up by ordinary superposition and rotation.

Simple theories with more than one vector meson have an important property -- they are characterized by a single _universal_ coupling constant. To see this, suppose there are two distinct multiplets of fermions $\psi^{(1)}$ and $\psi^{(2)}$, both coupled to the $A_{\alpha i}$ by means of the prescription (4.13) but possibly with different coupling strengths, $\gamma_{(1)}$ and $\gamma_{(2)}$. The scale of the matrices, $M_{(1)}$ and $M_{(2)}$, acting upon $\psi^{(1)}$ and $\psi^{(2)}$ is fixed by (4.15). For the interaction to be invariant under simultaneous rotations of the $\psi^{(\lambda)}$, (4.1), and the $A_{\alpha i}$, (4.3), it is clear that either: (a) the $A_{\alpha i}$ may be separated into two sets, one of which interacts only with $\psi^{(1)}$, the other with $\psi^{(2)}$ -- such a theory cannot be simple; (b) $c_{ijk} \equiv 0$ --- this is possible only if the

theory is a superposition of one or more trivial one-parameter theories; (c) $\gamma_{(1)} \equiv \gamma_{(2)}$ --- the non-trivial simple theory must also be a universal theory[2].

Now the condition that c_{ijk} be totally antisymmetric is easily shown to be equivalent to the condition[3],

$$\text{Tr } M_i M_j = (\text{const.}) \, \delta_{ij} \quad , \tag{4.18}$$

for a "simple" theory. For the original Yang-Mills theory, in which the M_i ($i = 1,2,3$) are isotopic spin matrices, Eq. (4.18) is evidently fulfilled.

We may now summarize the necessary and sufficient conditions for a simple generalized gauge theory. We must find an algebraic system, say of quantities S_i ($i = 1,\ldots n$), defined by a commutator $\left[S_i, S_j\right]$ obeying the antisymmetry and Jacobi laws (4.16 and 4.17) as well as the relation

$$\left[S_i, S_j\right] = i \, c_{ijk} \, S_k$$

with real, totally antisymmetric c_{ijk}. Furthermore, no real rotation of the S_i may result in a system that can be split into two commuting parts.

Such an algebraic system can always be represented by various sets of hermitian matrices M_i obeying the same rules as well as the condition (4.18). The construction of a Yang-Mills theory then follows the pattern we have outlined.

Now, the algebraic systems under discussion are well known to the mathematicians. One is the trivial one with n = 1, c_{ijk} = 0 that was discussed in Section II. All the others, including the Yang-Mills case with n = 3, c_{ijk} = e_{ijk}, are called simple Lie algebras (strictly speaking, simple Lie algebras in a special kind of real form). As such, they have been completely classified. All possible ones are known, and their representations by hermitian matrices M_i have been studied. In the next Section, we shall discuss the classification and some of the simpler cases.

Utiyama (20) has treated the connection of the Yang-Mills trick with Lie algebras, but he did not mention the severe restrictions of the Lie algebra that are necessary to obtain a vector meson theory with positive probabilities[4].

V. On Simple Lie Algebras

Let us mention first the listing of all the simple Lie algebras by Cartan (21). Each one, of course, may be regarded as the algebra of the infinitesimal generators of a continuous group, which is called a simple Lie group.

a) First of all, there is the infinite sequence of unitary unimodular groups $SU(\nu)$(ν = 2,3,4,....). The group $SU(\nu)$ is made up of all unitary transformations with unit determinant in an ν-dimensional complex space. The infinitesimal generators are then isomorphic to the traceless hermitian ν x ν matrices; evidently there are $\nu^2 - 1$ independent matrices of that kind and therefore the algebra has $\nu^2 - 1$ elements

S_i. We have $n = \nu^2 - 1$.

We have, incidentally, constructed the smallest representation of the S_i by matrices M_i; we simply use the $\nu^2 - 1$ traceless hermitian $\nu \times \nu$ matrices. They can, of course, be chosen to obey (4.18). Moreover, this representation is irreducible. (To avoid confusion, let us remark that the <u>algebra</u> of the S_i is already <u>simple</u>; no real rotation in the n-dimensional space of the $A_{i\alpha}$ can divide it into two parts that are unconnected by the c_{ijk}. However, the <u>representation</u> of the algebra by the $N \times N$ matrices M_i may be <u>reducible</u>. In other words, there may be a unitary transformation in the N-dimensional space of the ψ's that reduces all the M_i simultaneously to block form and allows us to pick out a smaller representation of the algebra. If such a reduction is impossible, the representation is irreducible.)

The isotopic spin algebra is that of SU(2); we have $n = 2^2 - 1 = 3$ and $S_i = I_i$ (i = 1,2,3). The irreducible representation by traceless hermitian 2×2 matrices $\tau_i/2$ satisfying (4.18) is just the familiar spin 1/2 representation. We know, too, all the other irreducible representations, classified according to the value $I(I + 1)$ of the matrix $\sum\limits_{i=1}^{3} M_i^2$, which commutes with all the I_i. The Ith representation (I = 0,1,2,....) is said to correspond to isotopic spin I and has dimension 2I + 1.

b) Next, we have the infinite sequence of rotation groups $O(\nu)$ in real ν-dimensional spaces (ν = 7,8,9,.....). We have omitted O(2) because it is just the one-parameter group of electromagnetism and that

degenerate case is not included among the simple Lie groups by the mathematicians. $O(3)$ is just the 3-dimensional rotation group and we know that is essentially the same as the isotopic spin group SU_2. The four-dimensional rotation group $O(4)$ is not simple; it is equivalent to the direct product $O(3) \times O(3)$. The groups $O(5)$ and $O(6)$ are omitted because they are essentially the same as $Sp(2)$ (see below) and $SU(4)$ respectively. Thus we begin with $O(7)$. The dimension n of the algebra of $O(\nu)$ is just the number of infinitesimal rotations $\frac{\nu(\nu-1)}{2}$. In fact, the infinitesimal $\nu \times \nu$ rotation matrices (imaginary and antisymmetric) form an irreducible matrix representation of the algebra of the group.

c) The third infinite sequence of simple Lie groups is that of the symplectic groups $Sp(\nu/2)$ with $\nu = 4,6,8,10,12,\ldots$ The algebra of the infinitesimal elements of $Sp(\nu/2)$ is just the algebra of the $\nu \times \nu$ skew-symplectic matrices[5]. Again we have a natural $\nu \times \nu$ irreducible matrix representation of the algebra. We note that $Sp(1)$ is omitted because it is the same as $SU(2)$.

d) Finally, there are five more simple Lie groups and the corresponding Lie algebras. These are called exceptional Lie algebras and their names and dimensions are as follows: G_2 -- 14, F_4 -- 52, E_6 -- 78, E_7 -- 133, E_8 -- 240.

In our listing, we have really defined each of the simple Lie algebras (except the exceptional ones, for which the same can be done) by exhibiting one of its matrix representations. In each case, we

understand that the n $\nu \times \nu$ matrices of the defining representation are to be taken hermitian and satisfying (4.18). We then have the simple Lie algebra in "real form" with real, totally antisymmetric c_{ijk}. In each case, we fix the value of the constant in (4.18) for the defining representation; that fixes the scale of the M's and of the c's.

For any given Lie algebra the matrix $\sum_i M_i^2$, which commutes with all the M_i (as we can see from (4.15)), equals some number for each irreducible representation. (This situation is familiar for the isotopic spin algebra, as mentioned above.) Let the value of $\sum_i M_i^2$ for representation R be V_R. Then for that representation the constant in (4.18) is $V_R d_R/n$, where d_R is the dimension of the representation.

In our generalized Yang-Mills theory, the various fields ψ that are coupled to the $A_{i\alpha}$ fall into multiplets, with each multiplet corresponding to an irreducible representation of the algebra. As long as the symmetry is maintained under gauge transformations <u>with constant gauge function</u>, the members of a multiplet are degenerate. The number of particles in the multiplet is, of course, the dimension of the representation.

Now the n vector fields $A_{i\alpha}$ represent n vector particles that also form a degenerate multiplet. They too correspond to an irreducible representation of the algebra, called the adjoint representation, with the same dimension as the algebra itself. (For example, in the case of isotopic spin, for which n = 3, the adjoint representation is that with

isotopic spin one, and the vector mesons form an isotopic triplet.) The matrices of the adjoint representation are easy to construct. They are simply

$$M_i^{jk} = -i \, c_{ijk} \qquad \text{for adjoint representation.} \qquad (5.1)$$

That this is so is obvious from the transformation properties of the $A_{i\alpha}$ in (4.3), with the gauge function taken constant, compared to those of the ψ_i in (4.1). In the adjoint representation, let

$$\text{Tr} \, M_i \, M_j = A \, \delta_{ij} \qquad \text{(adjoint representation)}; \qquad (5.2)$$

Then A defines the scale of the algebra. It is an arbitrary positive constant; from the above discussion it is clearly equal to the value of $\sum_i M_i^2$ for the adjoint representation. For any linear combinations S and T of the S_i, we can define a scalar product

$$(S,T) = \text{Tr} \, M(S) \, M(T) \qquad \text{(adjoint representation).} \qquad (5.3)$$

Then we have $(S_i, S_j) = A \, \delta_{ij}$, $iAc_{ijk} = (S_k, \, [S_i, S_j])$.

We might now characterize each simple Lie algebra by the constants c_{ijk}, but they are subject to arbitrary orthogonal transformations on the n-dimensional space of the $A_{i\alpha}$. An invariant and physically useful characterization is constructed as follows. (We quote without proof the usual mathematical results (22).)

Each simple algebra has a certain maximum number _f_ of elements that all commute with one another; let us call _f_ the rank of the algebra.

We may then enumerate the elements S_i of the algebra in this way: C_1, C_2, ... C_f; $\dfrac{E_1 + E_{-1}}{\sqrt{2}}$; $\dfrac{E_2 + E_{-2}}{\sqrt{2}}$, ... $\dfrac{E_g + E_{-g}}{\sqrt{2}}$; $\dfrac{E_1 - E_{-1}}{\sqrt{2}\,i}$, ... $\dfrac{E_g - E_{-g}}{\sqrt{2}\,i}$.

Here $g = (n-f)/2$. The C's are a maximal set of commuting elements. The E's are not real and are represented by non-hermitian matrices, but E_α and $E_{-\alpha}$ are represented by hermitian conjugate matrices. The corresponding vector fields are complex. The E's may be chosen to have these properties[6]:

$$(E_\alpha, E_\beta) = A\, \delta_{\alpha,-\beta} \qquad , \qquad (5.4)$$

$$\left[C_i, E_\alpha \right] = \lambda_i^\alpha E_\alpha \qquad , \qquad (5.5)$$

$$\left[E_\alpha, E_{-\alpha} \right] = \sum_i \lambda_i^\alpha C_i \qquad , \qquad (5.6)$$

$$\lambda_i^\alpha = -\lambda_i^{-\alpha} \qquad . \qquad (5.7)$$

The C_i are analogous to I_z in the isotopic spin algebra, while the $E_{\pm\alpha}$ are analogous to the raising and lowering operators $I_\pm$. The λ_i^α are the possible eigenvalue differences of the operators C_i in any representation. They are real and can be regarded as n-f distinct non-zero vectors in a real f-dimensional space.

The $(n-f)/2$ complex vector fields corresponding to the E_α give n-f vector particles carrying specific values of the quantities C_i, namely λ_i^α. Thus in the Yang-Mills theory the two charged mesons carry

$I_z = \pm 1.$

The vectors λ_i^α in the real f-dimensional space of the C_i are called roots. Their lengths and relative angles are invariant properties of the algebra (except for the overall scale of length, proportional to $\sqrt{A}$). We may define a scalar product for the roots:

$$\langle \alpha, \beta \rangle = \sum_i \lambda_i^\alpha \lambda_i^\beta \tag{5.8}$$

When we add to one root λ_i^β integral multiples $k \lambda_i^\alpha$ of another, we may find further roots. When this occurs, it always happens only for a sequence of successive integers $k = p_{\beta\alpha}, \ldots\ldots q_{\beta\alpha}$. Evidently $p \leq 0$ and $q \geq 0$. When $q \geq 1$, then $\lambda_i^\beta + \lambda_i^\alpha$ is a root. This situation is important for the commutation properties of the E_α:

$$\left[E_\alpha, E_\beta\right] = 0 \quad \text{unless} \quad \beta = -\alpha \text{ or } \lambda_i^\alpha + \lambda_i^\beta \text{ is a root;} \tag{5.9}$$

$$\left[E_\alpha, E_\beta\right] = i N_{\alpha\beta} E_\gamma \quad \text{when} \quad \lambda_i^\alpha + \lambda_i^\beta = \lambda_i^\gamma \quad ; \tag{5.10}$$

$$N_{\alpha\beta}^2 = \frac{1}{2} \langle \alpha, \alpha \rangle q_{\beta\alpha} (1 - p_{\beta\alpha}) \quad ; \tag{5.11}$$

$$N_{\alpha\beta} = N_{-\alpha,-\beta} \quad . \tag{5.12}$$

Even with the sign condition (5.12), the various relative signs of the E's must still be adjusted and a sign convention established for the $N_{\alpha\beta}$. But apart from that the algebra is now completely and invariantly described by its rank and the scalar products of its roots with one another. The commutation rules of all the C's and E's can then be constructed.

In the generalized Yang-Mills theory associated with the simple Lie algebra, we have gone over to a new particle representation. Instead of the n real fields $A_{i\alpha}$, we have f real fields coupled to the currents of commuting quantities and then $(n-f)/2$ complex fields coupled to the currents of raising and lowering operators for these commuting quantities. Instead of the c_{ijk}, we have the quantities λ_i^α and $N_{\alpha\beta}$ to describe the commutation rules and the amplitudes of the trilinear couplings among the vector mesons. (By going back to the real and imaginary parts of the complex fields, we can immediately recover the c_{ijk} in a particular form.) The particles of the complex fields carry the values λ_i^α of the quantities C_i and, since the C_i are conserved, the emission of the vector particle changes the value of C_i for the rest of the system by λ_i^α; the λ_i^α are indeed the possible eigenvalue differences of the C_i, whatever the representation.

In the next section, we shall give some examples of simple Lie algebras analyzed by the method of roots.

VI. Examples of Simple Lie Algebras

The simple Lie algebras of smallest dimension are those of the groups SU(2), with n = 3; SU(3), with n = 8; Sp(2), with n = 10; G_2, with n = 14; SU(4), with n = 15; Sp(3), with n = 21; O(7), with n = 21; SU(5), with n = 24; O(8), with n = 28; SU(6), with n = 35; Sp(4), with n = 36; and O(9), with n = 36. It is hard to imagine that any higher Lie algebras will be of physical interest.

The algebra of SU(2) or O(3) or Sp(1) is just the isotopic spin algebra and has, of course, rank one. The next three algebras are the only ones of rank two, and we shall use them as examples. (We might mention, however, that the next algebra after these, that of SU(4) or O(6), with rank three, is familiar to physicists. It is the algebra of the traceless Dirac matrices and is also the algebra of Wigner's old theory of nuclear supermultiplets (23).) For the three algebras of rank two, the roots are two-dimensional vectors, which are plotted in Figs. 1 - 3. The orientation and overall scale of length are arbitrary, as has been mentioned.

For any algebra, it is convenient to take one of the roots lying along the first axis and normalize its length to unity by proper choice of the constant A. For SU(3), it doesn't matter which root is chosen. For each of the other two cases, there are two inequivalent choices; we can take either a long or a short vector.

With the "first" root taken along the first axis with length one, the elements E_1, E_{-1}, and C_1 form the components $J_+/\sqrt{2}$, $J_-/\sqrt{2}$, and J_z of an angular momentum, as we can see from the commutation rules (5.5 and 5.6). Moreover, the second commuting quantity C_2 commutes with all three components of $\underset{\sim}{J}$.

Let us consider the algebra of SU(3). We may, using our convention, read off the values of the six roots in Fig. 1:

$$(1,0), \ (-1,0), \ (1/2,\sqrt{3}/2), \ (-1/2, \ \sqrt{3}/2), \ (1/2, \ -\sqrt{3}/2), \ (-1/2, \ -\sqrt{3}/2).$$

The commuting elements C_1, C_2, can be thought of as belonging to a

"root" $(0,0)$. With respect to $\underset{\sim}{J}$ spin and C_2, then, we have for the eight vector mesons the following: a triplet with $C_2 = 0$, a singlet with $C_2 = 0$, a doublet with $C_2 = \sqrt{3}/2$, and a doublet with $C_2 = -\sqrt{3}/2$. The triplet is coupled to the $\underset{\sim}{J}$-spin current, the singlet to the C_2 current, and the two doublets to the currents of raising and lowering operators that change $\underset{\sim}{J}$ by $1/2$ and C_2 by $\pm \sqrt{3}/2$.

Any representation of the algebra may be analyzed in terms of $\underset{\sim}{J}$ and C_2. For example, consider the defining representation, of dimension 3. In order to accommodate all the operators enumerated above, it must contain a singlet and a doublet, with values of C_2 differing by $\sqrt{3}/2$.

To obtain the quantities $N_{\alpha\beta}^2$ characteristic of the commutators in (5.10), we must ask what roots, in Fig. 1, can be added to make other roots. Evidently, the only case is that of two roots at 120° to each other; when added, they give the root in between. We see by inspection that the numbers p and q of Eq. (5.11) are zero and one respectively in this case. Thus, $N_{\alpha\beta}^2 = 1/2 \langle \alpha, \alpha \rangle$.

Next, let us look at the 10-dimensional algebra, with roots as in Fig. 2. If we take one of the short vectors to be $(1,0)$, then the root system is

$$(\pm 1, 1), \ (0, 1), \ (\pm 1, -1), \ (0, -1), \ (\pm 1, 0) \ .$$

Including the two vector mesons coupled to C_1 and C_2, both treated as $(0,0)$, we have a triplet with $C_2 = 0$, a triplet with $C_2 = +1$, a triplet with $C_2 = -1$, and a singlet with $C_2 = 0$.

If we treat the algebra as belonging to the group Sp(2), we get a four-dimensional defining representation which corresponds, using the above analysis, to a doublet with $C_2 = + 1/2$ and another with $C_2 = - 1/2$. If we consider the algebra in connection with O(5), then the defining representation is the 5-dimensional one, which consists, in our present language, of two singlets with $C_2 = \pm 1$ and a triplet with $C_2 = 0$.

Now we may consider the other possibility, taking one of the long vectors to be (1,0). The ten-dimensional adjoint representation then corresponds to two doublets with $C_2 = \pm 1/2$, three singlets with $C_2 = \pm 1, 0$, and a triplet with $C_2 = 0$. The four-dimensional representation yields a doublet with $C_2 = 0$ and two singlets with $C_2 = \pm 1/2$, while the five-dimensional one gives two doublets with $C_2 = \pm 1/2$ and a singlet with $C_2 = 0$. A conceivable physical application of this situation is **mentioned** in the next section.

The evaluation of $N_{\alpha\beta}^2$ for the ten-dimensional algebra involves two different situations in which adding two roots gives a third. As we see from Fig. 2, we can add a long vector to a short one at 135° from it, obtaining the short one at 45°. Or we can add two short vectors at right angles, obtaining the long one in between. In each case, N^2 comes out to equal the norm of the short vector.

Finally, we look at Fig. 3, showing the root system of G_2. [7] Here there are four different cases in which adding two roots gives a third. For three of these cases, N^2 is 3/2 times the norm of the short vector. The fourth case is that of adding two long vectors at 120° to

each other, obtaining the long one in the middle; N^2 is twice the norm of the short vector.
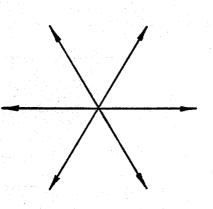
Again, the $\underset{\sim}{J}$-spin may be chosen in two ways. If a short vector is used, the adjoint representation corresponds to a singlet and a triplet with $C_2 = 0$, two quartets with $C_2 = \pm \sqrt{3}/2$, and two singlets with $C_2 = \pm \sqrt{3}$. If a long vector is taken to be $(1,0)$, then we get four doublets with $C_2 = \pm \frac{1}{2\sqrt{3}}$, $\pm \frac{\sqrt{3}}{2}$, a singlet and a triplet with $C_2 = 0$, and two singlets with $C_2 = \pm \sqrt{3}$.

For each of the three algebras we have taken as examples, one may work out all the representations of low dimension, analyze them according to $\underset{\sim}{J}$ and C_2, and calculate the matrix elements of the various operators. The whole procedure is a fairly straightforward generalization of what we do in the case of isotopic spin.

VII. Possible Applications to Physics

Sakurai (7) has discussed a vector meson picture of the strong interactions in which three simple gauge theories are superposed. We have a one-parameter theory of a meson ω^0 coupled to the hypercharge current, a three-parameter theory of a meson $\underset{\sim}{\rho}$ coupled to the isotopic spin current, and another one-parameter theory of a meson B^0 coupled to the baryon current. In all three cases, gauge invariance with variable gauge function is broken by some kind of meson mass term. In the first two cases, the conservation of the current itself, corresponding to gauge invariance with constant gauge function, is broken by weak and electromagnetic interactions respectively.
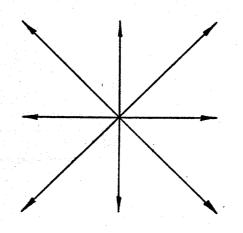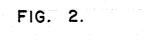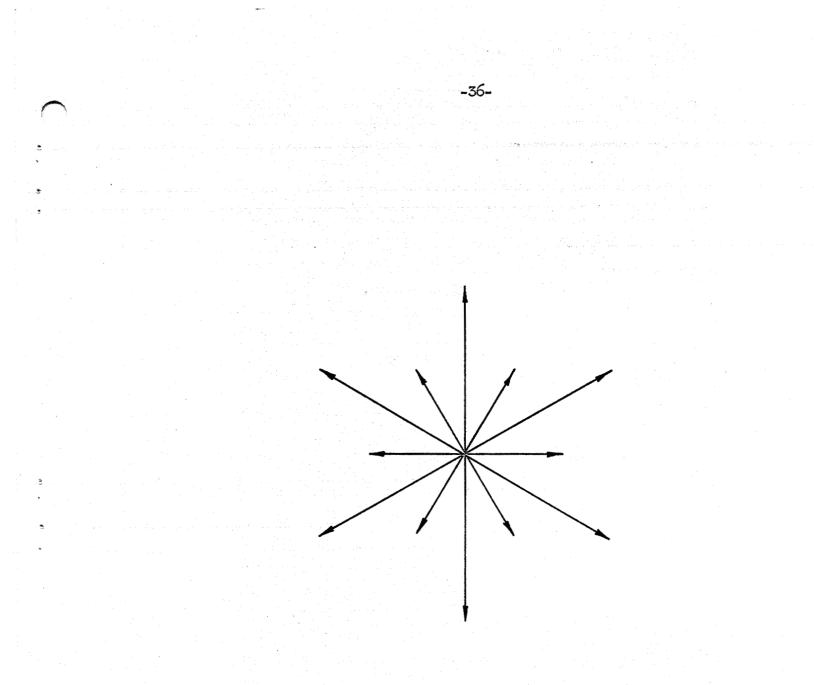
FIG. I.

Root vectors of SU(3).

FIG. 2.

Root vectors of $Sp(2) = O(5)$.

FIG. 3.

Root vectors of G_2.

If we are willing to let some large effect, such as the N-Λ mass difference or whatever causes it, break the gauge invariance with constant gauge function, then we may consider theories in which higher symmetries than isotopic spin play a role, and strangeness-changing currents are conserved to begin with. There are then strange vector mesons in the gauge theory, and we may be dealing with generalized Yang-Mills theories such as we have discussed.

It has been suggested (8) that the eight-dimensional algebra of SU(3) may be used for such a theory. The $\underset{\sim}{J}$ spin of the last section is taken to be the isotopic spin $\underset{\sim}{I}$ and C_2 is taken to be $\sqrt{3}/2$ times the hypercharge. Then if the baryons N, Ξ, Λ, and Σ all have the same spin and parity they can form an irreducible representation of the algebra. So can the pseudoscalar mesons K, $\overline{K}$, π, and χ, where χ is a hypothetical isotopic singlet of zero strangeness. The vector mesons of the gauge theory then follow the same pattern, and consist of Sakurai's $\underset{\sim}{\rho}$ and ω^o and a pair of strange doublets M and $\overline{M}$. The sources of the strange mesons are then strangeness-changing currents, the conservation of which is broken by such things as the baryon mass differences.

Alternatively, we may imagine that the baryon supermultiplet does not consist of N, Ξ, Λ, and Σ. Using the algebra of SU(3), we could take the three-dimensional irreducible representation and have it correspond to N and Λ (25).

We might even use the ten-dimensional algebra of Sp(2), taking one of the long vectors to correspond to the root (1,0), as discussed

in the previous section. If we interpret J_μ as I_μ and C_2 as half the hypercharge, then the baryons N, Λ, and Ξ could correspond to the five-dimensional irreducible representation.

Besides the strong interactions, we may consider the possible application of vector gauge theories to the weak couplings.

Since there is no sign of charge-retention weak couplings among the leptons, one might try to describe all weak couplings by a $J_\alpha^+ J_\alpha$ model in which just two intermediate vector particles $X^\pm$ are involved. Since the operators to which X^+ and X^- are coupled cannot commute, the algebraic system involved is not closed and the theory cannot be of our type.

The possibility has been discussed (4,5) of correcting the situation by introducing the electromagnetic field as the third member of a Yang-Mills triplet including $X^\pm$. The introduction of a huge mass for the $X^\pm$ would totally ruin the symmetry and account for the short range and feeble strength of the weak interactions. A major difficulty in this approach is that the generating operators of X^+ and X^- violate parity conservation and it is hard to make their commutator equal the electric charge operator which does conserve parity. For the leptons, the problem can be solved (26) only with the introduction of a fourth neutral gauge field. In the resulting "non-simple" $3 \oplus 1$ theory, the photon must be identified as a linear combination of the singlet gauge field and one member of the triplet.

A further difficulty is that with just $X^{\pm}$ for the weak interactions, one cannot justify the $\left| \Delta \underset{\sim}{I} \right| = 1/2$ rule for non-leptonic strangeness-violating weak couplings of baryons and mesons. One may, however, try forgetting the leptons and introducing, for baryons and mesons, charge-retention weak interactions mediated by neutral X's.

If just one X^o is used, then the strangeness-changing part of the operator to which it is coupled must carry both $\Delta S = +1$ and $\Delta S = -1$ in order to be hermitian. In the resulting weak interaction, one cannot then avoid having $\left| \Delta S \right| = 2$, which brings trouble with the $K_1^o - K_2^o$ mass difference.

The "schizon" model ($\underline{11},\underline{12},\underline{13}$) avoids $\left| \Delta S \right| = 2$ by using two neutral X's, X^o and $\overline{X^o}$, along with $X^{\pm}$. Say X^+ is coupled to an operator B, X^- to B^+, X^o to A, and $\overline{X^o}$ to A^+. Then we take $B = B_o + B_1$, where B_o conserves strangeness and changes isotopic spin by one unit, while B_1 lowers strangeness by one and changes isotopic spin by one-half unit. Similarly, $A = A_o + A_1$, where A_o conserves S and gives $\left| \Delta \underset{\sim}{I} \right| = 0,1$, while A_1 gives $\Delta S = -1$, $\left| \Delta \underset{\sim}{I} \right| = 1/2$. The operators B_1 and A_1 are chosen to be isotopic spin "partners"; the same is true of B_o and the $\left| \Delta \underset{\sim}{I} \right| = 1$ part of A_o. It is then easy to adjust the relative values of B_1 and A_1, B_o and A_o to give $\left| \Delta \underset{\sim}{I} \right| = 1/2$, $\left| \Delta S \right| = 1$ for the non-leptonic strangeness-violating weak interaction.

Unfortunately, it is impossible to apply the gauge principles we have discussed in this article to the "schizon" model, as we shall now see.

The only four-dimensional theories of our type are those made up of four one-dimensional theories superposed or else of one three-dimensional and one one-dimensional theory superposed. In the first case, the four operators A, B, A^+, and B^+ must all commute, which is clearly impossible. In the second case, if we take account of charge conservation, we must have (with suitable normalization and suitable choice of the arbitrary phase of A) the commutation rules:

$$\left[B, \frac{A + A^+}{2} \right] = B$$

$$\left[B, A - A^+ \right] = 0$$

$$\left[B^+, B \right] = \frac{A + A^+}{2}$$

$$\left[A + A^+, A - A^+ \right] = 0 \tag{7.1}$$

characteristic of a "3 $\oplus$ 1" theory. Writing the last equation as $\left[A, A^+ \right] = 0$ and taking the $\triangle S = 0$ part of the equation, we have

$$\left[A_o^+, A_o \right] + \left[A_1^+, A_1 \right] = 0 \qquad . \tag{7.2}$$

Now, A_1 is a strangeness-lowering operator. Let us suppose that a finite number of particles participate in the interaction. Consider those particles of lowest strangeness that are coupled to any of higher strangeness. For all these particles (treated as states $\bar{\Psi}_i$), we have

$$A_1 \bar{\Psi}_i = 0 \qquad . \tag{7.3}$$

On the other hand, $A_1^+ \bar{\Psi}_i$ cannot be zero for all the $\bar{\Psi}_i$, since some of these particles are coupled to states of higher strangeness. Thus

$$\sum_i (\bar{\Psi}_i, \left[A_1, A_1^+ \right] \bar{\Psi}_i) > 0 \qquad . \qquad (7.4)$$

But by (7.2) we then have

$$\sum_i (\bar{\Psi}_i, \left[A_0^+, A_0 \right] \bar{\Psi}_i) > 0 \qquad , \qquad (7.5)$$

which is impossible, since the matrices A_0 and A_0^+ conserve strangeness and connect states $\bar{\Psi}_i$ only to states of the same set; the expression in (7.5) is thus the trace of the commutator of two finite matrices and vanishes.

The algebraic system of the "schizon" model, then, has no interesting representations.

It is interesting that if we insist on trying to reconcile our gauge notions with the idea of a weak interaction with four X's that explains the $\left| \Delta I \right| = 1/2$ rule and the $K_1^0 - K_2^0$ mass difference, we can construct a model that does all of those things. As usual, we must gloss over the leptons (with their apparent lack of neutral currents) and treat just baryons and mesons.

To start with, we forget the $K_1^0 - K_2^0$ mass difference and allow $\left| \Delta S \right| = 2$. Then we can use just three vector fields: X^+, coupled to B, X^-, coupled to B^+, and X^0, coupled to A. Later, we will add in a field Y^0 coupled to an operator C. The X's are described by the original

3-parameter Yang-Mills theory. As before, we have $B = B_0 + B_1$, but this time $A = A_0 + A_1 + A_1^+$.

The commutation rules of the Yang-Mills theory

$$\left[B, A \right] = B$$

$$\left[B^+, B \right] = A \qquad\qquad (7.6)$$

yield, for the whole system of A_0, A_1, B_0, B_1, strangeness S, and charge Q, an algebra which (if we introduce a simplification by making A_0 a linear function of S and Q) is equivalent to that of SU(3). (Note that for the moment we do not introduce any isotopic spin rules; that would make the algebra still more complicated.) We may describe the theory, then, by using the smallest representation of SU(3), involving three particles, which we will take to be n, p, and Λ.

Let us introduce rotated "particles" (6,27)

$$n' = n \cos\Theta + \Lambda \sin\Theta$$

$$\Lambda' = \Lambda \cos\Theta - n \sin\Theta \qquad\qquad (7.7)$$

and work only with the left-handed parts $(1 + \gamma_5)/2 \; n' = n'_L$, etc., of these fields. Then we couple p_L and n'_L only, with $B = \tau'_-$, $B^+ = \tau'_+$, $A = \tau'_z$, where the τ' matrices are just like ordinary τ's, but with p_L, n'_L as a basis. We get the currents

$$\bar{p}_L \, \gamma_\alpha \, n_L^! = \bar{p}_L \, \gamma_\alpha \, n_L \cos\theta + \bar{p}_L \, \gamma_\alpha \, \Lambda_L \sin\theta \quad ,$$

$$\bar{n}_L^! \, \gamma_\alpha \, p_L = \bar{n}_L \, \gamma_\alpha \, p_L \cos\theta + \bar{\Lambda}_L \, \gamma_\alpha \, p_L \sin\theta \quad ,$$

$$\frac{1}{\sqrt{2}}\left(\bar{p} \, \gamma_\alpha \, p_L - \bar{n}_L^! \, \gamma_\alpha \, n_L\right) = \frac{1}{\sqrt{2}}\left\{ \bar{p}_L \, \gamma_\alpha \, p_L - \bar{n}_L \, \gamma_\alpha \, n_L \cos^2\theta \right.$$

$$\left. - \bar{\Lambda}_L \, \gamma_\alpha \, \Lambda_L \sin^2\theta - \cos\theta \sin\theta \, (\bar{\Lambda}_L \, \gamma_\alpha \, n_L + \bar{n}_L \, \gamma_\alpha \, \Lambda_L) \right\} \quad , \qquad (7.8)$$

which lead to $\Delta S = 0$, $|\Delta S| = 1$, and $|\Delta S| = 2$ interactions.

The $|\Delta S| = 1$ interaction contains both $|\Delta I| = 1/2$ and $|\Delta I| = 3/2$. However, if θ is small, then the $|\Delta I| = 1/2$ interaction is of order θ while the $|\Delta I| = 3/2$ interaction is of order θ^3. Then the strangeness-changing interaction is weaker than the strangeness-conserving one and the $|\Delta I| = 3/2$ interaction is weaker than $|\Delta I| = 1/2$, but not zero.[8]

We must still cancel out the $|\Delta S| = 2$ contribution to the $K_1^0 - K_2^0$ mass difference. That can be done (28) by cancelling just the scalar part of the $|\Delta S| = 2$ interaction. If we couple the fourth boson Y^0 with _appropriate strength_ to the current

$$i \, \bar{n}_R \, \gamma_\alpha \, \Lambda_R - i \, \bar{\Lambda}_R \, \gamma_\alpha \, n_R \quad , \qquad (7.9)$$

where we now work with the right-handed fields only, then the total $|\Delta S| = 2$ interaction can be made purely pseudoscalar and of order θ^2. It will give no $K_1^0 - K_2^0$ mass difference, but it will give a very small probability for $\Xi \rightarrow N + \pi$.

The model we have discussed is not seriously put forward as a physical theory, but it is a good illustration of the ideas involved in the gauge method.

We have discussed several ways in which the strong interactions may constitute a partially gauge-invariant theory, and have sketched a gauge-invariant model of the weak interactions. In general, the "weak" and "strong" gauge symmetries will not be mutually compatible. There will also be conflicts with the electromagnetic gauge symmetry, conflicts that must be resolved in favor of electromagnetism, since its gauge invariance is exact. We have not attempted here to describe the three kinds of interaction together, but only to speculate about what the symmetry of each one might look like in an ideal limit where symmetry-breaking effects disappear.

FOOTNOTES

p. 11: [1] We are using a classical, not a quantum, action principle.
The variational derivatives are thus with respect to
c-number quantities. Although we neglect the difficulties
encountered due to lack of commutativity of the quantized
fields, our results are presumably independent of this
omission in a properly formulated quantum theory.

p. 21: [2] The remarkable universality of the electric charge would be
better understood were the photon not merely a singlet, but
a member of a family of vector mesons comprising a simple
partially gauge-invariant theory. One of the authors (S.L.G.)
acknowledges a conversation with G. Feinberg in this connec-
tion.

p. 21: [3] Define $d_{ijk} = -i \, \mathrm{Tr} \, M_i \left[M_j, M_k \right]$; define $g_{k\ell} = \mathrm{Tr} \, M_k \, M_\ell$.
Clearly d_{ijk} is real and totally antisymmetric, $g_{k\ell}$ is real
and symmetric, and $d_{ijk} = \sum_\ell g_{i\ell} \, c_{\ell jk}$. Now if $g_{i\ell} = A \, \delta_{i\ell}$,
evidently $d_{ijk} = A \, c_{ijk}$ and c is totally antisymmetric.

To prove the converse, diagonalize $g_{i\ell}$ by an orthogonal trans-
formation of the fields $A_{i\alpha}$. Then $g_{i\ell} = F_i \, \delta_{i\ell}$ and
$d_{ijk} = F_i \, c_{ijk}$. But both $\underline{d}$ and $\underline{c}$ are totally antisymmetric
and therefore $d_{jik} = F_j \, c_{jik} = F_i \, c_{ijk}$ and $F_j = F_i$ whenever
i and j are connected by a non-zero coefficient c_{ijk}. For

FOOTNOTES (cont.)

a <u>simple</u> theory, however, we can ultimately convert all elements to one another in this way and prove all the F_i are equal. Thus $g_{i\ell} = A \, \delta_{i\ell}$ (unless, of course, it vanishes; that gives the one-parameter theory).

p. 22: [4] Note that if we set up the Einstein theory of gravity by gauge methods then the conclusions are slightly different. Instead of an isotopic rotation, we perform a Lorentz transformation at each point of space. Thus we have, in place of the isotopic index i, another Lorentz index β, giving us a tensor field $A_{\alpha\beta}$ or $h_{\alpha\beta}$. But whereas the metric in isotopic space must be positive to give positive probabilities, the metric in Minkowski space is both positive and negative, and this causes no trouble with positive probabilities. Thus the infinitesimal generator of a Lorentz transformation involving the time is not represented by a hermitian matrix, but such a situation, although it occurs in the theory of gravity, cannot be permitted in a gauge theory of vector fields.

p. 24: [5] A $\nu \times \nu$ matrix M is skew-symplectic if it is unitary and $M^T A M = A$, where $\left\| A_{ij} \right\| = \pm \, \delta_{\nu, i+j}$ depending upon whether or not $i > j$. $Sp(\nu/2)$ may alternatively be defined as the group of unitary transformations on $\nu/2$ quaternions. The generators of $Sp(\nu/2)$ are then the skew-hermitian $\nu/2 \times \nu/2$ matrices over quaternions; there are evidently $\nu(\nu + 1)/2$ such matrices.

FOOTNOTES (cont.)

p. 27: [6] The C_i and E_α are known as a Weyl basis to the Lie algebra.

p. 32: [7] A theory of the strong interactions of baryons and mesons whose invariance group is G_2 has been suggested by Behrends and Sirlin (24). They do not discuss the possibility of introducing vector gauge fields coupled to each of the fourteen conserved currents in order to secure invariance under coordinate-dependent transformations.

p. 43: [8] Observe that the sum of the squares of the coupling strengths to strangeness-saving charged currents and to strangeness-changing charged currents is just the square of the universal coupling strength. Should the gauge principle be extended to leptons -- at least, for the charged currents -- the equality between G_V and G_μ is no longer the proper statement of universality, for in this theory $G_V^2 + G_\Lambda^2 = G_\mu^2$ (G_Λ is the unrenormalized coupling strength for β-decay of Λ).

REFERENCES

1. R. P. Feynman and M. Gell-Mann, Phys. Rev. 109, 193 (1958).

2. J. Schwinger, Ann. Phys. 2, 407 (1957).

3. S. Bludman, Nuovo Cimento 9, 433 (1958).

4. S. L. Glashow, Nuclear Physics 10, 107 (1959).

5. A. Salam and J. C. Ward, Nuovo Cimento 11, 568 (1959).

6. M. Gell-Mann and M. Levy, Nuovo Cimento 16, 705 (1960).

7. J. Sakurai, Ann. Phys. 11, 1 (1960).

8. M. Gell-Mann, "The Eightfold Way: A Theory of Strong Interaction Symmetry", California Institute of Technology Synchrotron Laboratory Report No. CTSL-20 (1961). See also Y. Ne'eman, Nuclear Physics (to be published).

9. C. N. Yang and R. Mills, Phys. Rev. 96, 191 (1954).

10. R. Shaw, unpublished.

11. T. D. Lee and C. N. Yang, Phys. Rev. 119, 1410 (1960).

12. S. B. Treiman, Nuovo Cimento 15, 916 (1959).

13. M. Gell-Mann, Bull. Am. Phys. Soc. 4, 256(T) (1959).

14. C. N. Yang and T. D. Lee, Phys. Rev. 98, 1501 (1955).

15. M. Gell-Mann, Nuovo Cimento Suppl. 4, 848 (1956).

16. G. Wentzel, as reported by M. Goldberger, Proceedings of the 1960 Conference on High Energy Physics at Rochester (Interscience 1960).

17. C. G. Bollini, Nuovo Cimento 14, 560 (1959).

18. H. Umezawa and S. Kamefuchi, Nuclear Physics 23, 399 (1961).

REFERENCES (cont.)

19. A. Komar and A. Salam, Nuclear Physics 21, 624 (1960).

20. R. Utiyama, Phys. Rev. 101, 1597 (1956).

21. E. Cartan, "Sur la Structure des Groupes de Transformations Finis
    et Continus", Thèse (Paris, 1894, 2nd ed. 1933).

22. "The Sophus Lie Seminars, 1954-1955", (École Normale Supérieure,
    Paris, 1955).

23. E. Wigner, Phys. Rev. 51, 106 (1937).

24. R. E. Behrends and A. Sirlin, Phys. Rev. 121, 324 (1961).

25. A. Salam, to be published.

26. S. L. Glashow, Nuclear Physics 22, 579 (1961).

27. R. P. Feynman and M. Gell-Mann, Proceedings of the Geneva Conference
    on Peaceful Uses of Atomic Energy (1958).

28. S. L. Glashow, Phys. Rev. Letters 6, 531 (1961).