Ref.TH.2891-CERN

MONOPOLE CHARGES IN UNIFIED GAUGE THEORIES

CHAN Hong-Mo and TSOU Sheung Tsun
CERN - Geneva

A B S T R A C T

Monopole charges, being global quantities, depend on the gauge group of a theory, which in turn is determined by the representations of all its fields. For example, chromodynamics in its present form when combined with electrodynamics has as its gauge group not $SU(3) \times U(1)$ but a "smaller" group $U(3)$. The specification of monopole charges for a theory can thus be quite intricate. We report here the result of an investigation in several current gauge theories. Of particular interest is the possible existence in some theories of monopoles carrying multiplicative charges. As a by-product, we clarify some earlier assertions in the literature which seem to us incorrect.

Ref.TH.2891-CERN

25 June 1980

1. INTRODUCTION

From very general considerations of electrodynamics, Dirac predicted the existence of magnetic monopoles with a conserved quantized charge¹⁾. Although no magnetic monopole has yet been discovered, his prediction has long been recognized as being of fundamental theoretical significance. Given our present preoccupation with non-Abelian gauge theories, it is natural to ask whether Dirac's arguments can be extended to these theories. The general method for doing so is contained already in a rich and profound paper by Wu and Yang in 1975²⁾. It turns out, however, that for theories with gauge groups $SU(n)$ ^{*)}, which seem to be our most immediate concern, the answer is not very interesting. The concept which generalizes the magnetic charge of electrodynamics to non-Abelian theories is a certain topological quantum number which has to do with the degree of connectedness of the gauge group manifold; and the manifolds of $SU(n)$ being all simply-connected, theories with $SU(n)$ as gauge groups have no proper equivalent to the magnetic charge.

Hence it may appear that an extension of Dirac's considerations to include $SU(n)$ theories will lead to no new predictions, but this is not entirely true for the following reason. Combining an $SU(n)$ theory with electrodynamics, one obtains a theory with a larger symmetry corresponding locally^{**)} to the gauge Lie algebra $\mathfrak{su}(n) \oplus \mathfrak{u}(1)$ ^{*)}. However, topological charges being global^{**)} quantities, they depend not merely on the Lie algebra but rather on the gauge group, and several distinct groups can in general have the same Lie algebra. Now the gauge group of a theory is determined not by the Lagrangian but by the representations to which the physical fields of the theory belong. Indeed, an examination of chromodynamics, for example, in combination with electrodynamics reveals that the gauge group in

*) We adopt the convention of denoting a group and its Lie algebra by capital and small letters respectively; thus for example, $SU(n)$ and $\mathfrak{su}(n)$.

***) The terms 'local' and 'global' are used in the literature with respect both to the gauge group and to space-time. To avoid confusion, we shall use them here exclusively with respect to the gauge group. 'Local' and 'global' (in space-time) gauge transformations we shall refer to as 'x-dependent' and 'x-independent'.

the present theory is not $SU(3) \times U(1)$ but a "smaller" group $U(3)$. Depending thus on the gauge group under consideration, the classification of monopoles according to their charges can become quite intricate and lead to different predictions.

The same question arises also in the so-called unified theories embedding these symmetries into an even larger one, which is then broken to $su(n) \oplus u(1)$ by a Higgs mechanism. Both the classification of monopoles and the physical interpretation of their charges in these theories will depend on whether the subgroup corresponding to the exact symmetry is indeed $SU(n) \times U(1)$ or some other group having the same Lie algebra. Again, for the popular $SU(5)$ theory of Georgi and Glashow³⁾ for example, it turns out that the gauge group corresponding to electrodynamics and chromodynamics is $U(3)$ and not $SU(3) \times U(1)$.

It is our purpose in this paper to examine monopole charges in some current gauge theories. We believe that even in case that these theories were in the end to prove unrealistic, our considerations will serve as useful prototypes for theories of the future. Now, in common with Dirac's quantized magnetic charge, its generalizations have their origin in the geometrical structure of the theory which is independent of such details of the dynamics as the exact form of the Lagrangian or of the equations of motion. Indeed, one can well claim that their great beauty lies just in the fact that they are so primitive and therefore presumably fundamental. Hence, in order to exhibit this important feature, we have adopted a language appropriate for the purpose, which is somewhat out of the ordinary and avoids redundant reference to dynamical details. The language will be introduced in each section by a brief discussion of a familiar example.

As a by-product, we also clarify some earlier assertions in the literature on the subject which seem to us to be incorrect.

In this paper, we concentrate on physical concepts. For a more detailed and rigorous treatment of the mathematics involved, we refer the reader to a companion paper by one of us⁴⁾.

2. SPECIFICATION OF GAUGE GROUPS

The gauge group of a theory is determined by the representations to which the physical fields of the theory belong. Take the simplest example of invariance under a phase transformation:

$$\psi_j \rightarrow \psi'_j = \psi_j \exp(i e_j \alpha). \quad (1)$$

If all the fields ψ_j of a theory have charges e_j which are integral multiples of a certain minimal charge e , then for two values of α differing by $2\pi/e$ the transformations are indistinguishable. The gauge group is then compact and isomorphic to $U(1)$. Otherwise, if charge is not quantized, α must range over all values, in which case the gauge group is non-compact, its manifold being the real line $\mathbb{R}$ and not the circle S^1 as for $U(1)$ ⁵⁾.

Consider now the Weinberg-Salam theory for electromagnetic and weak interactions ⁶⁾, the gauge group for which is often referred to in the literature as $SU(2) \times U(1)$, but this is strictly speaking incorrect. By definition, the direct product $A \times B$ of two groups $A = SU(2)$ and $B = U(1)$ consists of the couples (a,b) , where $a \in A$ and $b \in B$, and different couples are to be regarded as distinct elements of $A \times B$. Now as the theory stands at present, the physical fields are either:

- i) $SU(2)$ doublets with half-integral hypercharges, e.g. $(\nu, e)_L$ with $Y = 1/2$;
 (ϕ^+, ϕ^0) with $Y = -1/2$, etc.

or

- ii) $SU(2)$ singlets and triplets with integral hypercharges, e.g. $(e)_R$ with $Y = 1$;
 B_μ, A_μ^i with $Y = 0$, etc.

Hence following the discussion in the last paragraph the period for the phase α_3 in $\exp(i g \alpha_3 T_3) \in SU(2)$ is $4\pi/g$, and for α' in $\exp(i g' \alpha' Y) \in U(1)$ is $4\pi/g'$.

Let, now

$$\hat{a} = \exp(i 2\pi T_3), \quad (\alpha_3 = 2\pi/g), \quad (2)$$

$$\hat{b} = \exp(i 2\pi Y) , \quad (\alpha' = 2\pi/g'). \quad (3)$$

Then $\hat{a}$ ($\hat{b}$) being different from the identity element of A (B), the couple $(\hat{a}, \hat{b})$ should be distinct from the identity element $(1,1)$ in $A \times B$. However, operating on the fields ψ_j in (i) and (ii), all of which have $T^3 + Y$ integer,

$$(\hat{a}, \hat{b}) \psi_j = \exp(i 2\pi (T_j^3 + Y_j)) \psi_j = \psi_j. \quad (4)$$

This means that $(\hat{a}, \hat{b})$ should be identified with $(1,1)$ and that the couple $(a \hat{a}, b \hat{b})$ for any a, b has always the same effect on our fields ψ_j as (a,b) . Our gauge group is therefore not $SU(2) \times U(1)$ but a "smaller" group obtained by identifying $(a \hat{a}, b \hat{b})$ with (a,b) in $SU(2) \times U(1)$, namely the group $SU(2) \times U(1)/\mathbb{Z}_2$. This group is isomorphic to $U(2)$, the group of 2×2 unitary matrices via the map: $(a,b) \rightarrow \underline{a} \cdot \underline{b}$, for $a = \exp(i g \alpha_k T_k)$, $b = \exp(i g' \alpha' Y)$, and $\underline{a} = \exp[(i/2) g \alpha_k \tau_k]$, $\underline{b} = \exp[(i/2) g' \alpha' I]$, τ_k and I being 2×2 matrices and $\cdot$ representing ordinary matrix multiplication⁴⁾.

Suppose next that we were to combine instead colour with electromagnetism, very similar considerations would apply. The minimal charge of our fields (by which we include the standard quarks and gluons) being now $e/3$, the period for the phase α in $\exp(i e \alpha Q)$ is $6\pi/e$, so that

$$d_r = \exp(i 2\pi r Q) , \quad r = 0, 1, 2, \quad (5)$$

are distinct elements of the electromagnetic gauge group $U(1)_{em}$. On the other hand, the 3×3 matrices,

$$c_r = \exp(i 2\pi r/3) I_3 , \quad r = 0, 1, 2, \quad (6)$$

being unitary and unimodular, represent three (also distinct) elements c_r of the group $C = SU(3)$. Hence the couples (c_r, d_r) for $r = 0, 1, 2$ correspond to three distinct elements of $SU(3) \times U(1)$. However, when operated on the quark fields q_j

belonging to the triplet representation of colour SU(3) and having electric charges $Q_j = -1/3, 2/3$, they have the same effect:

$$(c_r, d_r) q_j = \exp(i 2\pi r(\frac{1}{3} + Q_j)) q_j = q_j. \quad (7)$$

Furthermore, it is easily seen that all the other fields such as gluons and leptons which have integral charges and belong to the adjoint and the singlet colour representations, are left invariant under c_r and d_r separately. Hence one can conclude analogously that the gauge group is the one obtained by identifying in $SU(3) \times U(1)$ the triplets of elements $(c c_r, d d_r)$, for $r = 0, 1, 2$ and any $c \in SU(3)$ and $d \in U(1)$, namely the group $SU(3) \times U(1)/\mathbb{Z}_3$, which is isomorphic to the group U(3) of 3×3 unitary matrices⁴⁾.

Finally, combining colour, weak interactions, and electromagnetic interactions, the gauge group implied by the particles known at present and assigned to the usual representations, is the group obtained by identifying the following sextets of elements in $SU(3) \times SU(2) \times U(1)_Y$:

$$\begin{aligned} (c, a, y) &= (c c_1, a, y y_1) = (c c_2, a, y y_2) \\ &= (c, a \hat{a}, y \hat{y}) = (c c_1, a \hat{a}, y \hat{y} y_1) = (c c_2, a \hat{a}, y \hat{y} y_2), \end{aligned} \quad (8)$$

for any $c \in SU(3)$, $a \in SU(2)$, and $y \in U(1)_Y$, where $U(1)_Y$ consists of elements $\exp(i g' \alpha' Y)$ (Y being $T_3 - Q$ for all fields) with α' having period $12\pi/g'$, and

$$\hat{y} = \exp(i 6\pi Y). \quad (9)$$

$$y_r = \exp(i 4\pi r Y), \quad r = 1, 2. \quad (10)$$

We shall denote this group by⁴⁾

$$U_{2,3} = SU(3) \times SU(2) \times U(1)/\mathbb{Z}_6. \quad (11)$$

Of course, it is not excluded that future experiments may force us to change our theories to include particles belonging to other representations, and the gauge groups of our theories may change accordingly. For example, if particles with $T + Y$ half-integral were discovered, the gauge group for the Weinberg-Salam theory will have to be $SU(2) \times U(1)$ and not $U(2)$ as at present.

3. EXACT SYMMETRIES

Supposing we have a magnetic monopole somewhere in space, we can determine its magnetic charge as follows. We first enclose it with a closed surface which we may take without loss of generality to be topologically equivalent to a sphere. Next we measure the magnetic flux

$$H_i = -\frac{1}{2} \epsilon_{ijk} F_{jk} \quad (12)$$

at every point on this surface. The magnetic charge of the monopole is then given as

$$g_M = \frac{1}{4\pi} \oint\!\!\!\oint_{\text{sphere}} F_{ij} d\sigma_{ij}. \quad (13)$$

Equivalently, we may proceed as follows. Since $\text{div } \underline{H} = 0$ on the surface, we may write:

$$F_{ij} = \partial_i A_j - \partial_j A_i. \quad (14)$$

Because of the existence of the monopole, however, the potential A_i cannot be non-singular everywhere (hence the Dirac string), but one can cover the surface by two coordinate patches which we may take to be the northern and southern hemispheres overlapping along the equator. In each patch, the potentials denoted by A_μ^N and A_μ^S , respectively, are non-singular, and in the overlap region they are related by a gauge transformation,

$$A_\mu^S = A_\mu^N - \frac{i}{e} S \frac{\partial}{\partial x_\mu} S^{-1}. \quad (15)$$

Clearly, the magnetic charge in (13) can also be written as

$$g_M = \frac{1}{4\pi} \oint_{\text{equator}} \{A_j^N - A_j^S\} dx_j = \frac{1}{4\pi} \oint_{\text{equator}} \left\{ -\frac{\partial}{\partial x_j} \log S \right\} dx_j. \quad (16)$$

The form (16) for the magnetic charge, though less familiar, is more convenient for certain purposes. For example, from the fact that the patching function S must be single-valued, we have immediately

$$g_M = \frac{1}{4\pi} \oint \left\{ -\log S(\phi=2\pi) + \log S(\phi=0) \right\} = n/2e, \quad (17)$$

or that the magnetic charge is quantized. Next, from the fact that a potential with a magnetic charge $g_M = n/2e$ cannot change continuously to another with a different magnetic charge $g_M = n'/2e$, it can be deduced that the magnetic charge is conserved. These are just Dirac's results derived in a somewhat different way.

Suppose now that we were to attempt to generalize the concept of monopoles to non-Abelian theories; it can be seen that there will be difficulty in adopting directly the usual definition (13) of the magnetic charge. Since in general $\text{div } \underline{H} \neq 0$ for $\underline{H}$ given in (12), the integral (13) will depend on the surface over which it is performed. However, one may still give a meaning to the alternative (16) as follows. The patching function S which characterizes the gauge may be regarded as a map of the equator into the gauge group G , or equivalently as a closed curve C_S in the group manifold. Under an x -dependent gauge transformation in each hemisphere, C_S is deformed continuously into another such closed curve $C_{S'}$. It is clear then that gauges in the theory are divided into equivalence classes (gauge types) where members in each class are related by gauge transformations, while the classes themselves are in one-one correspondence with the homotopy classes of closed curves in the group manifold, namely $\pi_1(G)$. For electrodynamics,

$$\pi_1(U(1)) = \mathbb{Z}. \quad (18)$$

In other words, gauge types in electrodynamics are characterized by integers $[n]$, where n is equal to the number of times a curve in that homotopy class winds

around the group manifold S^1 . Now this is exactly $2e$ times the magnetic charge g_M of the monopole as given in (17).

It can therefore be seen that a natural generalization of the quantized magnetic charge of electrodynamics to non-Abelian theories is the gauge type defined over a sphere. To each potential $A_\mu(\underline{x})$ we may assign a quantity, the gauge type, which is quantized (i.e. discrete) and conserved, since, as for the Abelian theory, a potential here belonging to one gauge type also cannot change continuously into a potential belonging to another. Thus any potential belonging to some gauge type which does not contain the vacuum may be regarded as carrying a monopole charge of that type inside the sphere over which the gauge types are defined.

Since gauge types are somewhat abstract quantities, it may be convenient if possible to characterize them by a quantum number analogous to the magnetic charge. For $su(n)$ theories, we may use for this purpose the multiplicative quantum number^{7,*}:

$$\omega = \exp \oint_{\text{equator}} (S A_j^N S^{-1} - A_j^S) dx_j = \exp \oint_{\text{equator}} S \frac{\partial}{\partial x_j} S^{-1} dx_j, \quad (19)$$

where A_μ , S are the $n \times n$ matrices representing the potential $A_\mu(\underline{x})$ and the patching function S in the fundamental representation of the group $SU(n)$. Theories with gauge group $SU(n)$ have only one gauge type with $\omega = 1$ which contains the vacuum (i.e. no monopole), while in theories with gauge group $SU(n)/\mathbb{Z}_n$, ω can take values among the n^{th} roots of unity $\{\zeta_r\}$, $r = 0, \dots, n-1$, $(\zeta_r)^n = 1$.

Consider now the Weinberg-Salam theory for electromagnetic and weak interactions⁶). In this theory, the symmetry corresponding to the algebra $su(2) \oplus u(1)$ is broken by a Higgs mechanism to $u(1)$ of electrodynamics. However, we shall first discuss monopole charges in this theory as if the $su(2) \oplus u(1)$ symmetry were exact. Indeed, according to some modern extensions of the theory, an unbroken $su(2) \oplus u(1)$ symmetry is believed to correspond to the actual effective situation at some primordial stage of the universe. Here, however, we are studying in detail the unbroken $su(2) \oplus u(1)$ theory mainly because it is the simplest

*) In Ref. 7, a factor of i/e has inadvertently been omitted in the second term on the r.h.s. of (27) and in the last term of (28).

example of the type of theories encountered in Section 2, obtained by combining other gauge theories with electrodynamics. The question of monopole charges in the broken theory will be considered in the next section.

Although the Weinberg-Salam theory in its present form has the gauge group $U(2)$ as shown above, we shall consider here also the theory with gauge group $SU(2) \times U(1)$ so as both to illustrate the differences between the two cases and to cover the eventuality that the 'larger' gauge group be forced upon us by future experimental discoveries or theoretical considerations. The manifold for the group $SU(2) \times U(1)$ can be pictured as a doughnut whose cross-sections are $SU(2)$ manifolds, namely S^3 . Notice that we deliberately suggested the analogy of a doughnut (i.e. a solid) as distinct from that of a toroidal surface to bring out the fact that closed curves in $SU(2)$ can all be shrunk continuously to a point so that the number of times a curve 'wraps' around the doughnut has no topological significance. Thus, for example, the curve C_0 in Fig. 1 belongs to the same homotopy class in $\pi_1[SU(2) \times U(1)]$ as C_1 , which 'wraps' once round the doughnut. It is clear then that these homotopy classes, and hence also the gauge types in a theory with gauge group $SU(2) \times U(1)$, are characterized by integers $[n]$, n being the number of times the curves in that class 'wind' around the doughnut (e.g. $n = 1$ for both C_0 and C_1 in Fig. 1). Any potential belonging to the gauge type $[n]$ may then be regarded as carrying a monopole charge of type $[n]$ inside the sphere on which the gauge types are defined.

For certain purposes we may like to ask whether these potentials may be regarded also as carrying monopole charges corresponding to the subgroups $U(1)$ and $SU(2)$. To answer this, let us calculate according to (16) the hypermagnetic charge g_{MY} in the $U(1)$ subgroup for a potential belonging to type $[n]$. If the hypermagnetic charge is a well-defined quantity then it is gauge-independent and we need examine only one patching function in each class, for example:

$$S^{[n]} = \exp(i 2\pi \phi \gamma) \quad (20)$$

belonging to the homotopy class $[n]$. A potential

$$\Gamma_\mu = g A_\mu^i T_i + g' B_\mu Y \quad (21)$$

in a gauge corresponding to $S^{[n]}$, takes different forms in the northern and southern hemispheres which are related by the following gauge transformation along the equator:

$$\Gamma_\mu^S = S^{[n]} \Gamma_\mu^N (S^{[n]})^{-1} - i S^{[n]} \frac{\partial}{\partial x_\mu} (S^{[n]})^{-1}. \quad (22)$$

Substituting (20) and (22) into (16), we have immediately

$$g_{MY} = \frac{1}{4\pi} \oint (\beta_j^N - \beta_j^S) dx_j = n/g'. \quad (23)$$

[The apparent factor of 2 between (23) and (17) is due simply to the fact that here the minimal hypercharge is by convention $g'/2$.] Similarly, we can evaluate the quantum number ω_F defined as the ω in (19) for the subgroup $SU(2)$. Indeed, since $S^{[n]}$ in (20) does not depend on T_i , we have by (22) $A_\mu^{iN} = A_\mu^{iS}$, so that for all $[n]$, $\omega_F = +1$, i.e. the same as for the vacuum. Moreover, it can readily be checked that both g_{MY} and ω_F are gauge-independent quantities. It can therefore be seen that in a theory with gauge group $SU(2) \times U(1)$, monopoles have exactly the same charges as in a $U(1)$ theory. Except for the trivial replacement of the electric charge e by the hypercharge $g'/2$, we have obtained no new predictions besides those of Dirac, as indeed we have anticipated already in the Introduction.

However, the gauge group of present interest for the Weinberg-Salam theory is not $SU(2) \times U(1)$ but $U(2)$, which has a different group structure. Now $U(2)$ can be obtained by identifying diagonally opposite points in the doughnut for $SU(2) \times U(1)$, which can be proved to give just another doughnut shape⁴⁾, so that again

$$\pi_1(U(2)) = \mathbb{Z}, \quad (24)$$

and homotopy classes are again labelled by integers $[m]$. However, there are more classes in $\pi_1[U(2)]$ than in $\pi_1[SU(2) \times U(1)]$ in the sense that only for even values of m is a closed curve in the manifold of $U(2)$ also a closed curve in the manifold of $SU(2) \times U(1)$. Take, for example, a curve joining the identity $(1,1)$ to the element $(\hat{a}, \hat{b})$ in $SU(2) \times U(1)$. This is a closed curve in the $U(2)$ manifold since in the group $U(2)$ these elements are identified, but it is clearly not a closed curve in $SU(2) \times U(1)$. Potentials belonging to gauge types $[m = \text{odd}]$ in the $U(2)$ theory represent therefore a new class of monopoles which did not exist in the $SU(2) \times U(1)$ theory.

As before, these $U(2)$ monopoles may be regarded also as carrying hypermagnetic charges. It is sufficient again to consider only one particular patching function in each class, say

$$S^{[m]} = \exp(-i m \phi (T_3 - Y)). \quad (25)$$

The expressions for (21) and (22) are still valid, and on substitution into (16) we obtain

$$g_{MY} = \frac{1}{4\pi} \oint_{\text{equator}} (\beta_j^N - \beta_j^S) dx_j = m/2g', \quad (26)$$

which is quantized in only half Dirac units. For odd m , however, these are not pure hypermagnetic monopoles, since according to (19) they have

$$\omega_F = \exp \oint_{\text{equator}} \{ S^{[m]} \frac{\partial}{\partial x_j} (S^{[m]})^{-1} \} dx_j, \quad (27)$$

where $S^{[m]}$ is a 2×2 matrix representing $e^{-im\phi T_3}$ in the fundamental representation of $SU(2)$, namely

$$S^{[m]} = \begin{pmatrix} \exp -\frac{im}{2}\phi & 0 \\ 0 & \exp \frac{im}{2}\phi \end{pmatrix}. \quad (28)$$

In other words,

$$\omega_F = (-1)^m, \quad (29)$$

which differs from the vacuum for m odd. Here then, for the first time, we seem to have found something new which is not already covered by Dirac.

Essentially the same analysis as that given above for $SU(2) \times U(1)$ and $U(2)$ can be carried out for all the other gauge groups considered in Section 2. We shall therefore give only the results.

For theories with gauge groups $SU(3) \times U(1)$ and $SU(3) \times SU(2) \times U(1)$, there are only pure $U(1)$ monopoles satisfying Dirac's quantization condition for its $U(1)$ magnetic charge.

For theories combining colour with electromagnetism with gauge group $U(3)$, monopoles are characterized by an integer $[m]$. They have magnetic charges

$$g_M = m/2e, \quad (30)$$

and a quantum number ω for $su(3)$ defined as in (19):

$$\omega_c = \exp \frac{-2\pi i m}{3}. \quad (31)$$

For $m = 3n$ (n integral) they are pure magnetic monopoles satisfying Dirac's quantization condition for magnetic charges (i.e. $g_M = 3n/2e$, the minimal charge being $e/3$), but for other values of m , they are new objects not previously encountered.

For theories combining together colour, weak interactions, and electromagnetic interactions with a gauge group $U_{2,3} = SU(3) \times SU(2) \times U(1)/Z_6$, monopoles are again characterized by an integer $[m]$. They have hypermagnetic charges:

$$g_{MY} = m/2g', \quad (32)$$

a quantum number ω_F for SU(2)-flavour:

$$\omega_F = (-1)^m, \quad (33)$$

and a quantum number ω_C for SU(3)-colour:

$$\omega_C = \exp \frac{-2\pi i m}{3}. \quad (34)$$

Among these monopoles, those with $m = 6n$ (n integral) are pure hypermagnetic, satisfying the Dirac quantization condition $g_{MY} = 3n/g'$, the minimal hypercharge now being $g'/6$. All others are new objects not previously encountered.

It has been suggested in the literature that monopoles in non-Abelian theories form 'magnetic charge multiplets' belonging to representations of a 'magnetic group'⁸⁾. We believe this classification to be incorrect for reasons we have already explained in a recent note⁷⁾. The idea of 'magnetic multiplets' was recently adopted to the problem of monopoles in a unified theory with gauge group SU(5)⁹⁾. As we shall see in the next section, this problem reduces exactly to the questions we are considering here. Therefore our classification disagrees with the conclusions reached in that paper. To exhibit the difference, it suffices to consider the simplest example in a theory with gauge group U(2). A classification of monopoles according to the 'magnetic group' scheme gives a doublet of monopoles corresponding to two potentials $\Gamma_\mu^\pm$ pointing in fixed isospin directions for all x and μ in gauges with patching functions $S^\pm = \exp i\phi(\pm T_3 + Y)$, respectively. It is convenient to depict these patching functions as follows. The generators T_3 and Y generate a subgroup $U(1)_{T_3} \times U(1)_Y$ of $SU(2) \times U(1)$, which is a torus representable by the diagram of Fig. 2 whose boundary is meant to be cyclic. The length of the sides gives the period of the phase α in each U(1). The patching functions $S^\pm$ defined above are represented by $C^\pm$ in the figure where, because the elements (a, b) and $(a \hat{a}, b \hat{b})$ of $SU(2) \times U(1)$ are identified in U(2), one needs to consider only one half of the diagram and to regard both $C^\pm$ as closed curves on the U(2) manifold. Now C^+ and C^- differ only by a closed circuit in the manifold of SU(2) and must

therefore belong to the same homotopy class in $\pi_1[U(2)]$. This implies that the two potentials $\Gamma^\pm$ are of the same gauge type and can change continuously from one to the other. In other words, there is in general finite transition between the states represented by the two potentials⁷⁾. They cannot therefore, in our view, be interpreted as forming a charge doublet of an exact magnetic symmetry^{*)}.

4. BROKEN SYMMETRIES

In this section we are interested in broken symmetries for two reasons. First, in the Weinberg-Salam theory of electroweak interactions, the symmetry $su(2) \oplus u(1)$ is broken down by a Higgs mechanism to the $u(1)$ symmetry of electrodynamics; whereas up to now we have only analysed the charges of monopoles as if the $su(2) \oplus u(1)$ symmetry in this theory were exact. Secondly, one may wish to embed the gauge symmetries already considered into an even larger symmetry as is often done in the recent literature in the so-called grand unification schemes, which symmetry is then broken by an appropriate Higgs mechanism down to the symmetries we know.

In order to introduce the language appropriate for our purpose, consider first the original example of 't Hooft¹⁰⁾ and Polyakov^{11,12)} in broken $su(2)$ symmetry. At any time t , the system is described by the gauge fields $A_\mu^i(\underline{x})$ and the Higgs fields $\phi^i(\underline{x})$, $i = 1, 2, 3$, both belonging to the adjoint representation, which implies that the theory has gauge group $SO(3) = SU(2)/Z_2$. The requirement that the symmetry be broken is specified by a term in the Lagrangian $V(\phi^2)$ with minima at $\phi^2 = \rho_0^2 \neq 0$.

The manifold for $SO(3)$ being doubly connected, there are two distinct gauge types characterized as in (19) by the numbers $\omega = \pm 1$. Suppose now that we are given, on a large sphere Σ in space, the values of $A_\mu^i(\underline{x})$ and $\phi^i(\underline{x})$ in some gauge belonging to the gauge type $\omega = +$. One can always transform them into the gauge

*) We were advised by the authors of Ref. 9 in a recent discussion that their classification given in their paper was not meant to be exact as conjectured by Goddard et al.⁸⁾, but to be another speculation based on the assumed existence of a 'spontaneously broken magnetic group' which cannot at present be either proved or disproved. We thank these authors for this clarification, which was otherwise not obvious from our reading of the paper.

$\mathcal{G}$ where the patching function is trivial, i.e. $S = 1$. In this gauge, then, both A_μ^i and ϕ^i are single-valued and continuous over the whole sphere Σ . Now ϕ^2 cannot differ from its minimum value ρ_0^2 except over a finite area on the infinite sphere Σ ; otherwise the energy would become infinite, given the form of the Higgs potential. We may then write

$$\phi^i(\underline{x}) = \hat{\phi}^i(\underline{x}) \rho_0, \quad \hat{\phi}^2 = 1, \quad (35)$$

where $\hat{\phi}^i$ can be regarded as a map of the unit sphere into the unit isosphere except at isolated points, which we can neglect. These maps are divided into homotopy classes labelled by integers $[n]$, i.e.

$$\pi_2(S^2) = \mathbb{Z}. \quad (36)$$

Any ϕ_1^i whose $\hat{\phi}_1^i$ belongs to one homotopy class $[n_1]$ cannot be deformed continuously into another ϕ_2^i , whose $\hat{\phi}_2^i$ belongs to a different homotopy class $[n_2]$. Hence there can be no continuous field path with finite action, and therefore also no transition, which takes a state at time t_1 with $\phi^i(\underline{x}) = \phi_1^i(\underline{x})$ to a state at a later time t_2 with $\phi^i(\underline{x}) = \phi_2^i(\underline{x})$ if $n_1 \neq n_2$. In other words, the number n is conserved.

Next the electromagnetic potential A_μ is defined as the component of the potential A_μ^i in the direction of the Higgs field ϕ^i when the latter is pointing in a fixed direction in isospace. Let us therefore rotate $\phi^i(\underline{x})$ to a fixed direction, say along the 3-axis, by a gauge transformation. For a ϕ^i whose $\hat{\phi}^i$ belongs to the homotopy class $[n]$ with $n \neq 0$, the transformation cannot be continuous everywhere on Σ , but can be so in each of two patches, which we may take as usual to be the northern and southern hemispheres overlapping along the equator. In the new gauge $\mathcal{G}'$, the patching function S' is no longer trivial. However, since by definition $\phi'^N = \phi'^{Ni}_{Ti}$ and $\phi'^S = \phi'^{Si}_{Ti}$ are related by

$$\phi'^S = S' \phi'^N S'^{-1}, \quad (37)$$

and both point in the 3-direction, S' must be of the form

$$S' = \exp (ie\alpha(\underline{x})T_3), \quad (38)$$

which is represented by a closed curve in the group manifold lying entirely in the electromagnetic $U(1)$ subgroup of rotations about the 3-axis. Moreover, it can be seen that if $\phi^i(\underline{x})$ in the original gauge $\mathcal{G}$ belongs to the homotopy class $[n]$ in $\pi_2(S^2)$, then S' belongs to the gauge type $[2n]$ corresponding to the homotopy class $[2n]$ in $\pi_1[U(1)]$. In other words, for whatever $A_\mu^i(\underline{x})$, so long as it is single-valued and continuous in the original gauge $\mathcal{G}$, we have for the magnetic charge g_M , defined in (16), the value $g_M = n/e$. This is just a rederivation in our language of 't Hooft and Polyakov's original result.

Adopting this language, we turn now to the Weinberg-Salam theory of electroweak interactions. At any time t the system is described by gauge fields $A_\mu^i(\underline{x})$, $B_\mu(\underline{x})$, the isospinor Higgs field $\phi(\underline{x})$, and some matter fields which we need not specify here. The symmetry is broken by a Higgs potential $V(\phi^\dagger\phi)$ which has minima at $\phi^\dagger\phi = \rho_0^2 \neq 0$.

Consider first the simpler situation when the gauge group is $SU(2) \times U(1)$. There are infinitely many gauge types which may be labelled by integers $[n]$, as explained in Section 3. Let now $A_\mu^i(\underline{x})$, $B_\mu(\underline{x})$, and $\phi(\underline{x})$ be given over a large sphere Σ in space in some gauge $\mathcal{G}_{[n]}$ belonging to the gauge type $[n]$. To avoid infinite energies, ϕ cannot deviate from the minima of V except locally on Σ . It may thus be regarded as having effectively a fixed length ρ_0 . One can then always find an $SU(2)$ gauge transformation which rotates $\phi(\underline{x})$ to point in the same direction in isospace for every $\underline{x}$, thus

$$\phi(\underline{x}) \rightarrow \phi'(\underline{x}) = \begin{pmatrix} 0 \\ \rho_0 \end{pmatrix}. \quad (39)$$

As in the preceding example, the electromagnetic group is defined in this new gauge $\mathcal{G}'_{[n]}$ as that subgroup of the original gauge group which leaves $\phi'(\underline{x})$ invariant. Now, under the action of a general element of $SU(2) \times U(1)$, namely a couple of the form $g = (\exp ig\alpha^i T^i, \exp ig'\alpha' Y)$, the Higgs field ϕ' transforms as

$$\phi'(\underline{x}) \longrightarrow \tilde{\phi}'(\underline{x}) = \exp\left(\frac{ig}{2}\alpha^i \tau_i - \frac{ig'}{2}\alpha'\right) \phi'(\underline{x}). \quad (40)$$

The requirement that $\phi'(\underline{x})$ be left invariant then defines the electromagnetic subgroup as consisting of the following set of elements:

$$h = (\exp -i\alpha e \tau^3, \exp i\alpha e Y), \quad 0 \leq \alpha \leq 4\pi/e, \quad (41)$$

where

$$e = gg'/\sqrt{g^2 + g'^2}. \quad (42)$$

In this same gauge, the electromagnetic potential, defined as the component of Γ_μ^i lying in the electromagnetic subgroup, takes the particular simple form:

$$A_\mu = -\frac{e}{g} A_\mu'^3 + \frac{e}{g'} B_\mu'. \quad (43)$$

Moreover, since the Higgs field is the same in both the northern and southern hemispheres, the patching function S relating them,

$$\phi'^S = S \phi'^N \quad (44)$$

is represented by a closed curve on the gauge group manifold which lies entirely in the electromagnetic subgroup and winds around the group exactly n times. Hence the magnetic charge g_M as defined in (16) can immediately be calculated and yields the value $g_M = n/e$ for any potentials $A_\mu^i(\underline{x})$ and $B_\mu(\underline{x})$ defined in the original gauge $\mathcal{G}[n]$. One therefore sees that, in contrast to the example of 't Hooft and Polyakov, all the potentials belonging to one gauge type $[n]$ of the original gauge group $SU(2) \times U(1)$ carry the same magnetic charge n/e when the symmetry is broken down to $U(1)$ by the Higgs mechanism.

Of more immediate interest in electroweak interactions, however, is a theory with gauge group $U(2)$. In the present context, this differs from a theory with gauge group $SU(2) \times U(1)$ only in having more gauge types, in the sense that although

gauge types are also labelled by integers $[m]$ they correspond to gauge types of $SU(2) \times U(1)$ only when m is even. Applying the same considerations as before, one obtains that any potentials $A_\mu^i(\underline{x})$ and $B_\mu(\underline{x})$ belonging to the gauge type $[m]$ must carry a magnetic charge $m/2e$. For m even, these charges may be interpreted as pure magnetic monopoles. For m odd, however, they carry in addition an $SU(2)$ quantum number $\omega_F = -$, as already explained in Section 3, (29).

Essentially the same analysis applies when these gauge theories are combined with chromodynamics. If the gauge group for the combined theory is $SU(3) \times SU(2) \times U(1)$, we have only pure magnetic monopoles when the symmetry is broken to $SU(3) \times U(1)$ by the Higgs mechanism. On the other hand, if the gauge group is $U_{2,3}$ as defined in (11), potentials belonging to the gauge type $[m]$ carry magnetic charges $g_M = m/2e$, but are not pure magnetic monopoles in general since they carry also $SU(2)$ quantum numbers ω_F and $SU(3)$ quantum numbers ω_C as listed in (33) and (34).

Notice that the expression (43) for the electromagnetic potential is by construction supposed to be valid only in the gauge where the Higgs field ϕ is pointing in a fixed direction in isospace. Now it has been asserted in the literature¹³⁾ that monopoles belonging to the same gauge type in theories with gauge groups $SU(2) \times U(1)$ or $U(2)$ can acquire different magnetic charges when the symmetry is broken down to $U(1)$ of electromagnetism. This is in contradiction to the conclusions reached above, and is, we believe, an error due to using the formula (43) for A_μ in a wrong gauge. It is sufficient to illustrate our point by considering one particular example in a theory with gauge group $U(2)$. Take any potentials $A_\mu^i(\underline{x})$, $B_\mu(\underline{x})$ in a gauge with patching function $S = \exp i\phi(T_3 + Y)$ as represented by the closed curve C^+ along OE in Fig. 2. If the "magnetic charge" for these potentials is calculated according to (13) or (16) simply by substituting for A_μ from (43) in this gauge, we obtain $g_M = e/2 (-1/g^2 + 1/g'^2)$ [which for $\sin^2 \theta_W = g'^2/(g^2 + g'^2) = 3/8$ appropriate to $SU(5)$, is the value $1/8e$ given in Ref. 13]. However, if the magnetic charge is calculated by transforming first to the right

gauge with a patching function represented by a curve in the electromagnetic group, i.e. along DC in Fig. 2, we obtain the correct Dirac quantized value $1/2e^*$).

We next examine what happens when the symmetries considered above are embedded in an even larger symmetry. It is sufficient to illustrate the problems involved by investigating the popular Georgi-Glashow model with $SU(5)$ symmetry³⁾. Since some quarks and leptons in this theory are assigned to the fundamental 5-dimensional representation, the gauge group is necessarily $SU(5)$ and not the "smaller" group $SU(5)/Z_5$. The symmetry is broken by two sets of Higgs scalars belonging to the 24-dimensional adjoint and 5-dimensional fundamental representations, respectively, which we may consider as occurring in two stages.

At the first stage, the symmetry is broken by a term in the Lagrangian of the form

$$V_1 = -\frac{1}{2}\mu^2 (\phi_i^j \phi_j^i) + \frac{1}{4}\lambda_1 (\phi_i^j \phi_j^i)^2 + \frac{1}{4}\lambda_2 (\phi_i^j \phi_j^k \phi_k^l \phi_l^i), \quad (45)$$

where ϕ_i^j ($i, j = 1, \dots, 5$) with $\phi_i^i = 0$ represent the Higgs fields in the adjoint representation. For $\lambda_2 > 0$ it was shown by Li¹⁴⁾ that V_1 has minima, and that these minima can always be rotated by an $SU(5)$ transformation into the form

$$\tilde{\phi}_i^{'j} = \rho_0 \begin{pmatrix} 1 & & & & \\ & 1 & & & \\ & & 1 & & \\ & & & -\frac{3}{2} & \\ & & & & -\frac{3}{2} \end{pmatrix}, \quad (46)$$

where $\rho_0 > 0$ is some function of μ and λ_r , the explicit form of which does not concern us here. Now it can immediately be seen that $\tilde{\phi}_i^{'j}$ in (46) is invariant

*) If we were to define naively also a "Z-magnetic charge", simply replacing A_μ by

$$Z_\mu = \frac{e}{g'} A_\mu^3 + \frac{e}{g} B_\mu$$

in the gauge with patching function $S = \exp i \phi Y$, we would obtain a non-zero value also. But this has clearly no significance, as the formula for Z_μ has again been used in a wrong gauge. It would therefore appear that the argument in Ref. 13 based just on this non-vanishing Z-monopole charge, that primordial monopoles in theories with gauge group $SU(2) \times U(1)$ will automatically annihilate, is incorrect.

under a subgroup of $SU(5)$ isomorphic to the group $U_{2,3}$ defined in (11). In other words, the minima of V_1 together form a manifold $\mathcal{M}$ which is topologically equivalent to $SU(5)/U_{2,3}$.

Now the group manifold of $SU(5)$ being simply connected implies that there is only one gauge type. Given then the gauge fields A_μ^α and the Higgs fields ϕ_i^j over some large sphere Σ in space, one can always transform them into the gauge $\mathcal{G}$ where the patching function is trivial, i.e. $S = 1$, and both A_μ^α and ϕ_i^j are single-valued and continuous over Σ . Again, to avoid infinite energies, ϕ_i^j cannot deviate except locally on Σ from the minima of V_1 , and may therefore be regarded effectively as maps $\hat{\phi}_i^j$ of the unit sphere into the manifold $\mathcal{M}$, namely $SU(5)/U_{2,3}$. Such maps are divided into homotopy classes where the members of different classes cannot be deformed continuously into one another. Repeating now the above arguments for the 't Hooft-Polyakov example, we see that the class to which a certain ϕ_i^j belongs is a conserved quantity. Again, $SU(5)$ being simply connected, it can be shown⁴⁾ that these homotopy classes are in one-one correspondence with homotopy classes of closed curves on the manifold of $U_{2,3}$, namely

$$\pi_2(SU(5)/U_{2,3}) = \pi_1(U_{2,3}), \quad (47)$$

or, in other words, with gauge types in a theory with gauge group $U_{2,3}$. Moreover, by rotating ϕ_i^j via an $SU(5)$ gauge transformation from the gauge $\mathcal{G}$ to another gauge $\mathcal{G}'$ in which the Higgs fields $\phi_i'^j$ point in the fixed direction (46), the subgroup $U_{2,3}$ of $SU(5)$ which leaves $\tilde{\phi}_i'^j$ invariant can then be assigned the physical significance ascribed to it in Sections 2 and 3. The classification of monopoles in the $SU(5)$ theory at this first stage of symmetry breaking is therefore identical to the unbroken theory with gauge group $U_{2,3}$, as we have described in Section 3.

At the second stage, the residual $U_{2,3}$ symmetry is broken down to $U(3)$. The situation here is essentially the same as that in the breaking of a theory with gauge group $U_{2,3}$ down to $U(3)$, considered earlier in this section, and leads to the same classification.

5. REMARKS

The stated aim of this paper was to specify the charges of monopoles in some current gauge theories. This modest aim is now achieved. Now the charge, whether magnetic or otherwise, inside any region of space is a quantity definable at any time and should therefore be independent of the details of dynamics which govern the time evolution of the system. We have taken special pains to exhibit this fact by insisting on a language appropriate for the purpose. Although it may seem pedantic to do so, we have found a fair amount of confusing and incorrect assertions in the literature, which can be traced directly to a lack of clarity on this point.

The physical implications of our results are at present quite limited. The most interesting is probably the possible existence of monopoles carrying multiplicative charges. Clearly, if monopoles exist, as seems probable in some current theories such as $SU(5)$, the first questions to ask would concern such dynamical quantities as masses, production cross-sections in the primordial universe, and survival rate at our present time, none of which can be answered by our purely geometrical considerations. Nonetheless, unless one has a correct specification of the monopole charges -- as we claim to have done -- it is easy to come to wrong conclusions when investigating the dynamics, as in the example of Section 4.

Of course, one of the most interesting questions to ask about monopoles is whether they may not correspond to something we already know. Present dynamical theories have one unsatisfying feature in that the number and nature of fundamental matter fields are always assumed. It would be much more satisfactory if, in a certain theory, the matter fields occur automatically as monopoles. It is hoped that our analysis here may be useful for future attempts in this direction.

REFERENCES

- 1) P.A.M. Dirac, Proc. Roy. Soc. London A133 (1931) 60.
- 2) Tai Tsun Wu and Chen Ning Yang, Phys. Rev. D 12 (1975) 3845.
- 3) H. Georgi and S.L. Glashow, Phys. Rev. Letters 32 (1974) 438.
- 4) Tsou Sheung Tsun, CERN preprint TH.2892-CERN (1980).
- 5) Chen Ning Yang, Phys. Rev. D 1 (1970) 2360.
- 6) S. Weinberg, Phys. Rev. Letters 19 (1967) 1264.
A. Salam, *in* Proc. 8th Nobel Symposium, Elementary particle theory (ed. N. Svartholm) (J. Wiley and Sons, New York, 1968), p. 367.
- 7) Chan Hong-Mo and Tsou Sheung Tsun, CERN preprint TH.2883-CERN (1980), to appear in Phys. Lett.
- 8) P. Goddard, J. Nuyts and D. Olive, Nucl. Phys. B125 (1977) 1.
- 9) M. Daniel, G. Lazarides and Q. Shafi, CERN preprint TH.2800-CERN (1979).
- 10) G. 't Hooft, Nucl. Phys. B79 (1974) 276.
- 11) A. Polyakov, Zh. Eksp. Teor. Fiz. Pis'ma 20 (1974) 430 [JETP Lett. 20 (1974) 194].
- 12) A. Polyakov, Sov. Phys. JETP 41 (1976) 988.
- 13) G. Lazarides and Q. Shafi, CERN preprint TH.2821-CERN (1980).
- 14) Ling-Fong Li, Phys. Rev. D 9 (1974) 1723.

FIG 1

FIG 2