PHYSICAL REVIEW D VOLUME 21, NUMBER 4 15 FEBRUARY 1980

Restricted gauge theory

Y. M. Cho
Max-Planck-Institut für Physik und Astrophysik, München, Federal Republic of Germany
(Received 27 April 1979)

A restricted gauge theory is obtained as a self-consistent subset of a non-Abelian gauge theory by imposing an extra magnetic symmetry to the gauge symmetry. The theory describes the dual dynamics between the color isocharges (i.e., the electric charges) and the topological charges (i.e., the magnetic charges) of the non-Abelian symmetry, and contains two potentials, the electric and the magnetic potentials, in a dual-symmetric way. The topological charge is identified as the dual of the Noether charge of the magnetic symmetry of the theory. A possible role of the restricted chromodynamics for quark confinement in quantum chromodynamics is speculated.

I. INTRODUCTION

Since Dirac[1] introduced his magnetic monopole within the context of Abelian gauge theory, attempts[2] have been made to construct a local field theory of the monopole which could exhibit an explicit duality between the electric and the magnetic charges. In the meantime it has become clear that monopoles can better be described in non-Abelian gauge theories. In non-Abelian gauge theories the monopoles appear as classical solutions of the system[3] which have a definite topological meaning.[4] This observation has led people to speculate that perhaps there may exist a built-in duality[5,6] in non-Abelian gauge theories. Furthermore, it has been emphasized that the monopoles and the supposedly existing dual structure might play an important role in non-Abelian gauge theories, in particular in connection with the issue of quark confinement in quantum chromodynamics (QCD).[6,7] If this point of view has a trace of truth, one should be able to construct, out of QCD, a local gauge theory of the non-Abelian monopole which exhibits a built-in duality and at the same time could give us a better understanding of the dynamics of quark confinement. So far, however, there seems to exist no such self-consistent theory of the non-Abelian monopole. The purpose of this paper is to present such a theory.

The theory that we propose here consists of a well-defined and self-consistent subset of a non-Abelian gauge theory of a given symmetry group G which has an *additional* symmetry which we call the magnetic symmetry for a reason that will become clear soon. The virtue of the additional magnetic symmetry is that while keeping the full gauge degrees of freedom intact it restricts and reduces the dynamical degrees of freedom, providing us with a self-consistent but nontrivial subset of the original gauge theory. For this reason we will call the theory *the restricted gauge theory* of the group G when compared to the canonical (i.e., the unrestricted) gauge theory that does not have any additional symmetry. As we will see in detail in the following, owing to the magnetic symmetry one can choose a particular gauge (the magnetic gauge) in which the nonvanishing components of the restricted gauge field become only those of a smaller subgroup H that is uniquely determined by the magnetic symmetry. Thus one can formally reduce the theory to the gauge theory of the subgroup H. However, in this gauge *the gauge field of the subgroup H is now made of two parts, the "electric" part which is by definition not restricted by the magnetic symmetry and the "magnetic" part which is completely determined by the magnetic symmetry*. Furthermore, the electric part describes the electric flux of the color isocharges while the magnetic part describes the magnetic flux of the topological charges of the symmetry group G. This observation guarantees us the dual structure of our theory, which can be made explicit in the magnetic gauge.

In the magnetic gauge one can describe the gauge field in terms of two potentials, the electric and the magnetic potentials, and make the duality explicit at the level of the potential. Thus, the theory can be written explicitly in terms of the two potentials, together with whatever source one has in the theory. Furthermore, the magnetic potential couples to the source in a well-defined gauge-invariant way. However, although the two potentials appear in a symmetric way in the Lagrangian, there still exists a significant disparity between them. While the electric potential is regular and plays the role of the ordinary gauge potential the magnetic one is singular and contains the string singularity. In addition, the magnetic potential describes the monopoles with a "spacelike" potential while the electric one describes the isocharges with a "timelike" potential. To remove these apparent asymmetries we introduce the concept of *the dual magnetic potential*. The beauty of the dual magnetic potential is that it does not contain the string singularity anymore and describes the magnetic charges with a timelike potential. This enables us to make the duality more transparent. More significantly the existence of the dual magnetic potential allows us to view the magnetic symmetry as a genuine Noether symmetry of the theory. This means that a topological charge could, indeed, be viewed as the dual counterpart of a Noether charge.

One of the striking aspects of the restricted theory is the fact that within the framework of the restricted chromodynamics (RCD) one obtains the confinement of the color by breaking the magnetic symmetry dynamically. Logically, RCD can have *two* phases, the normal phase and the confined phase. In the normal phase the magnetic symmetry is left unbroken so that not only the quarks but also the monopoles should appear as physical particles. In the confined phase, however, both the quarks and the monopoles should disappear from the physical spectrum.

The paper is organized as follows: In Sec. II we start from the classical gauge theory and give the exact mathematical definition of what we mean by the magnetic symmetry in its most general form. The corresponding restricted gauge theory is then outlined. Although the theory can be formulated with an arbitrary gauge group G, in this paper we will mainly concentrate on the simplest group, i.e., the SU(2) group, to avoid unnecessary complications. The more realistic color SU(3) gauge group will be treated in a subsequent paper. In Sec. III we introduce the concept of the magnetic potential as a local field. As we will see, both the electric and the magnetic potential enjoy independently the explicit gauge degrees of freedom of the subgroup H as a part of the full gauge degrees of freedom of the group G. In Sec. IV we introduce the quarks to the system and show how the two potentials couple to the source in a gauge-invariant way. Again the gauge invariance determines the dynamics of the theory uniquely. In Sec. V we present *all* the possible homotopically inequivalent classes of the magnetic symmetry and the corresponding classical monopole configurations, which can describe arbitrary integral magnetic charges of the group G. Also we show that by choosing a proper magnetic symmetry one can describe any classical dynamical system of the monopoles. In Sec. VI we present a quantum-field-theoretical description of the restricted theory. This is done by introducing the dual magnetic potential and at the same time a charged scalar field to represent the monopole. The dual magnetic potential removes the string singularity and enables us to treat the monopole as a local field. Finally, in the last section we briefly outline a possible way that the restricted chromodynamics could explain quark confinement in QCD.

II. MAGNETIC SYMMETRY

In this section we will start from the classical gauge theory and give a precise mathematical definition of what we mean by magnetic symmetry. To do this, however, a better understanding of the geometrical structure of the gauge theory is necessary. For this reason we will start from a brief review of the relevant geometrical aspects of the gauge theory.

The gauge theory can be viewed as Einstein's theory of gravitation in a higher-dimensional unified space[8,9] which consists of the four-dimensional external space-time and the n-dimensional internal group space. If the metric g_{AB} ($A,B = 1,2,\ldots,4+n$) in this $(4+n)$-dimensional unified space has an n-dimensional isometry group G whose Killing vector fields span the internal space, the corresponding Einstein theory becomes essentially the gauge theory of the group G in curved space-time.[9,10] To be precise let us choose the n Killing vector fields ξ_i ($i=1,2,\ldots,n$) to satisfy the canonical commutation relations of the isometry group G,

$$[\xi_i,\xi_j]=f_{ij}{}^k\xi_k. \quad (1)$$

By definition these Killing vector fields must also satisfy

$$\mathcal{L}_{\xi_i} g_{AB}=0 \quad (i=1.2.\ldots,n), \quad (2)$$

where $\mathcal{L}_{\xi_i}$ is the Lie derivative along the direction of ξ_i. Now if one further assumes that n Killing vector fields ξ_i are orthonormal to each other with respect to the metric g_{AB} so that the internal metric ϕ_{ik} ($i,k=1,2,\ldots,n$) of the n-dimensional internal space

$$\phi_{ik}=\xi_i^A\xi_k^B g_{AB} \quad (3)$$

becomes of the Cartan-Killing form (here we are assuming that the group G is semisimple), then the corresponding Einstein theory in this $(4+n)$-dimensional space becomes the canonical Yang-Mills gauge theory (coupled with gravitation if the external space-time is assumed to be curved). In general, however, if the Killing vector fields are not kept to be orthonormal so that the internal metric (3) is left arbitrary, then one has, in addition to the gauge fields (and the external gravitons), the internal gravitons ϕ_{ik} as nontrivial dynamical fields in the theory. These scalar fields couple to the other fields in a gauge and generally invariant way.[10,11] We will call this theory the *generalized* gauge theory when it is necessary to distinguish it from the canonical one.

Now we are ready to discuss the gauge theory of the monopoles, or rather the restricted gauge theory. The theory is defined as the generalized gauge theory which has an extra internal symmetry made of some additional Killing vector fields which are internal and which commute with the already existing fields ξ_i. We will call the additional symmetry magnetic. To be precise, let us assume that there exists only one such vector field which we denote by m. Then, by assumption, one has

$$[m,\xi_i]=0 \quad (i=1,2,\ldots,n) \tag{4}$$

and

$$\mathcal{L}_m g_{AB}=0 . \tag{5}$$

Since m is assumed to be internal one can write

$$m=m^i\xi_i . \tag{6}$$

Then the condition (4) tells us that the multiplet $\hat{m}$, made of the components m^i,

$$\hat{m}=\begin{pmatrix} m^1 \\ m^2 \\ \cdot \\ \cdot \\ \cdot \\ m^n \end{pmatrix}$$

must form an adjoint representation of the group. Now the second condition (5) will undoubtedly restrict the internal metric (3) as well as the gauge potential. Indeed it is not difficult to show that the condition (5) can be written as

$$m^i(f_{ij}{}^l\phi_{kl}+f_{ik}{}^l\phi_{jl})=0 \tag{7}$$

and

$$D_\mu\hat{m}=\partial_\mu\hat{m}+g\vec{B}_\mu\times\hat{m}=0 , \tag{8}$$

where $\vec{B}_\mu$ is the gauge potential of the group G. Now let us assume for simplicity that the internal metric ϕ_{ik} is of the Cartan-Killing form. The general case in which the internal metric is left arbitrary will be discussed separately.[12] With this simplification the condition (7) is automatically satisfied. Now the condition (8) implies, among others, that the multiplet $\hat{m}$ must have a constant length,

$$\hat{m}^2=\text{const} , \tag{9}$$

which one can choose to be the unit without loss of the generality, and we will do so in the following. To see how the condition (8) restricts the gauge potential $\vec{B}_\mu$ let us consider the case when the isometry group is SU(2). In this case the condition can be solved exactly for $\vec{B}_\mu$,

$$\vec{B}_\mu=A_\mu\hat{m}-\frac{1}{g}\hat{m}\times\partial_\mu\hat{m} , \tag{10}$$

where A_μ is the (Abelian) component of $\vec{B}_\mu$ which is *not* restricted by the condition (8). Notice that the potential (10) is made of two parts, the unrestricted part A_μ and the other part which is completely determined by the magnetic symmetry. This decomposition property will have a far-reaching consequence in the following. For an obvious reason we will call the unrestricted part A_μ electric and the other restricted part magnetic.

The field strength $\vec{G}_{\mu\nu}$ corresponding to the potential (10) can easily be figured out. One finds

$$\begin{aligned}\vec{G}_{\mu\nu}&=\partial_\mu\vec{B}_\nu-\partial_\nu\vec{B}_\mu+g\vec{B}_\mu\times\vec{B}_\nu\\ &=(F_{\mu\nu}+H_{\mu\nu})\hat{m} ,\end{aligned} \tag{11}$$

where

$$\begin{aligned}F_{\mu\nu}&=\partial_\mu A_\nu-\partial_\nu A_\mu ,\\ H_{\mu\nu}&=-\frac{1}{g}\hat{m}\cdot(\partial_\mu\hat{m}\times\partial_\nu\hat{m}) .\end{aligned} \tag{12}$$

There are two important general aspects in the above results. First notice that $\vec{G}_{\mu\nu}$ is parallel to $\hat{m}$. This is not an accident. In fact, from the simple identity

$$[D_\mu,D_\nu]\hat{m}=g\vec{G}_{\mu\nu}\times\hat{m} , \tag{13}$$

which holds for an arbitrary group G, one can immediately see that the only nonvanishing components of $\vec{G}_{\mu\nu}$ that satisfy the magnetic symmetry (8) must necessarily be those of the little group H of $\hat{m}$. Thus for SU(2), for example, $\vec{G}_{\mu\nu}$ has to be parallel to $\hat{m}$. The other general aspect of the above results is that $\vec{G}_{\mu\nu}$ is made of two parts, $F_{\mu\nu}$ which comes from the unrestricted potential A_μ and $H_{\mu\nu}$ which comes from the part restricted by the magnetic symmetry $\hat{m}$. This dual structure which has already appeared at the level of the potential (10) is again a general aspect independent of what symmetry G one considers. Actually, from (8) and (13) it becomes intuitively clear that with an arbitrary gauge group G the same dual structure will persist in $\vec{B}_\mu$ and $\vec{G}_{\mu\nu}$, so that, in general, one can always set

$$\vec{G}_{\mu\nu}=\vec{F}_{\mu\nu}+\vec{H}_{\mu\nu} , \tag{14}$$

where both $\vec{F}_{\mu\nu}$ and $\vec{H}_{\mu\nu}$ should have the components of the little group H which can in general be non-Abelian. Naturally we will call $\vec{F}_{\mu\nu}$ electric and $\vec{H}_{\mu\nu}$ magnetic. This intriguing dual structure[13] of the theory can be made more dramatic when one realizes that one can actually go further and introduce the magnetic potential for $\vec{H}_{\mu\nu}$. This is our next subject.

III. MAGNETIC POTENTIAL

To introduce the magnetic potential corresponding to the magnetic field $\vec{H}_{\mu\nu}$ let us first remember that our theory still has the full gauge degrees of freedom of the group G. So for SU(2) one can rotate the magnetic vector $\hat{m}$ to a prefixed space-time independent direction, say, the third direction of the group space, by a gauge transformation U,

$$\hat{m} \underset{U}{\rightarrow} \hat{\xi}_3 = U\hat{m} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}. \tag{15}$$

To be specific, let us parametrize $\hat{m}$ by

$$\hat{m} = \begin{pmatrix} \sin\alpha \cos\beta \\ \sin\alpha \sin\beta \\ \cos\alpha \end{pmatrix}. \tag{16}$$

Then clearly one can choose

$$U = e^{-\alpha t_2} e^{-\beta t_3} \tag{17}$$

to fulfill the condition (15), where $t_i (i=1,2,3)$ is the adjoint representation of the generators. Now consider the following $\vec{C}_\mu^{*(m)}$ defined by

$$\vec{C}_\mu^{*(m)} = -\frac{1}{g}\hat{m} \times \partial_\mu \hat{m} - \frac{1}{g}\mathrm{tr}(-\tfrac{1}{2}\vec{t}\, U^{-1}\partial_\mu U). \tag{18}$$

From (16) and (17) one finds

$$\vec{C}_\mu^{*(m)} = C_\mu^* \hat{m}, \tag{19}$$

where

$$C_\mu^* = \frac{1}{g}\cos\alpha\, \partial_\mu \beta. \tag{20}$$

Then it is straightforward to show that under the gauge transformation (15) one has

$$\begin{aligned} \vec{B}_\mu &= A_\mu \hat{m} - \frac{1}{g}\hat{m} \times \partial_\mu \hat{m} \\ &= A_\mu \hat{m} + \vec{C}_\mu^{*(m)} + \frac{1}{g}\, tr(-\tfrac{1}{2}\vec{t}\, U^{-1}\partial_\mu U) \\ &= (A_\mu + C_\mu^*)\hat{m} + \frac{1}{g}\, tr(-\tfrac{1}{2}\vec{t}\, U^{-1}\partial_\mu U) \\ &\underset{U}{\rightarrow} \vec{B}'_\mu = (A_\mu + C_\mu^*)\hat{\xi}_3 \end{aligned} \tag{21}$$

and

$$\begin{aligned} \vec{G}_{\mu\nu} &= (F_{\mu\nu} + H_{\mu\nu})\hat{m} \\ &\underset{U}{\rightarrow} \vec{G}'_{\mu\nu} = (F_{\mu\nu} + H_{\mu\nu})\hat{\xi}_3, \end{aligned} \tag{22}$$

where now $H_{\mu\nu}$ is expressed in terms of C_μ^*,

$$\begin{aligned} H_{\mu\nu} &= -\frac{1}{g}\hat{m} \cdot (\partial_\mu \hat{m} \times \partial_\nu \hat{m}) \\ &= -\frac{1}{g}\sin\alpha(\partial_\mu \alpha\, \partial_\nu \beta - \partial_\nu \alpha\, \partial_\mu \beta) \\ &= \partial_\mu C_\nu^* - \partial_\nu C_\mu^*. \end{aligned} \tag{23}$$

Thus in this gauge, which we will call the magnetic gauge, the potential $\vec{B}_\mu$ has a remarkably simple form. It consists of two Abelian ones A_μ and C_μ^* in a dual-symmetric way. Furthermore, we have identified C_μ^* as the potential for the magnetic field $H_{\mu\nu}$. Naturally, we will call C_μ^* the magnetic potential. So in our theory we have, in addition to the electric potential, the magnetic one which is completely fixed by the magnetic symmetry.

The fact that one can introduce the magnetic potential for SU(2) is not accidental. In general, even when the little group H becomes non-Abelian as may be the case for a larger symmetry G, one can introduce a similar non-Abelian magnetic potential $\vec{C}_\mu^*$ for $\vec{H}_{\mu\nu}$, as we will see in the accompanying paper. Furthermore, one can show that the potential $\vec{C}_\mu^*$ even enjoys the gauge degrees of freedom of the subgroup H. To see this let us go back to the SU(2) case and notice that the gauge transformation U that fulfills the condition (15) is not uniquely determined. Consequently our definitions (18) and (19) have the corresponding degrees of freedom. To see exactly what degrees of freedom they have, notice that U has the following (and no other) degrees of freedom

$$U \Rightarrow U' = Ue^{-\gamma \vec{t}\cdot\hat{m}} = e^{-\gamma t_3} U, \tag{24}$$

which comes from the little group degrees of freedom of the magnetic symmetry $\hat{m}$. Under the replacement (24) one can easily show that $\vec{C}_\mu^{*(m)}$ transforms as

$$\begin{aligned} \vec{C}_\mu^{*(m)} &= C_\mu^* \hat{m} \\ &\Rightarrow \vec{C}_\mu^{*(m)\prime} = C_\mu^{*\prime} \hat{m}, \end{aligned} \tag{25}$$

where

$$C_\mu^{*\prime} = C_\mu^* + \frac{1}{g}\partial_\mu \gamma. \tag{26}$$

There are two points to be noticed here. First, no matter what U one chooses to define $\vec{C}_\mu^{*(m)}$, it always has the same form (19). This is what is expected. Indeed, from the general considerations discussed in the last section, the magnetic potential (if it exists at all) must behave like that of the little group H. The other point is that the little group degrees of freedom (24) indeed guarantees the magnetic potential the corresponding gauge degrees of freedom (26). This is again a general aspect independent of what group G one considers.

Here we have the Abelian degrees of freedom (26) simply because for SU(2) the little group H has to be Abelian.

In the magnetic gauge the potential C^*_μ is completely fixed by whichever U one has chosen. However, notice that even after one has fixed U (and accordingly the magnetic gauge degrees of freedom) one still has the *electric* degrees of freedom

$$H = e^{-\theta t_3} \tag{27}$$

that leaves $\hat{\xi}_3$ invariant, under which A_μ enjoys the corresponding gauge degrees of freedom

$$A_\mu \to A'_\mu = A_\mu + \frac{1}{g}\partial_\mu \theta \,. \tag{28}$$

Clearly the two gauge degrees of freedom (26) and (28) have one and the same origin. The stability subgroup of $\hat{m}$ which led us to the magnetic gauge degrees of freedom (26) now provides us with the electric gauge degrees of freedom (28) in the magnetic gauge. In other words, the two gauge invariances come from the same little group degrees of freedom except that they are extracted from two different points of view.

The beauty of the magnetic gauge is that here all the nonessential gauge degrees of freedom of the original symmetry G have been removed. Furthermore, only those components of the potential which have explicitly the structure of the little group H appear in a dual-symmetric way, so that the structure of the theory becomes more transparent. Thus, the magnetic gauge becomes the natural one to consider when one wants to figure out the gauge-invariant coupling of the two potentials to the source and to discuss the dynamics of the theory. This is our next subject.

IV. DUAL DYNAMICS

Now that the magnetic potential (as well as the electric one) is introduced one can include an arbitrary source in the theory and discuss the dynamics of the theory. Clearly the magnetic symmetry (8) does not restrict the symmetry structure of the source of the theory so that it must remain a multiplet of the full group G. Now to figure out its gauge-invariant coupling to the restricted potential (10) the magnetic gauge is the best one to choose, and we will do so in the following.

Let us first consider an SU(2) isodoublet spinor source Ψ

$$\Psi = \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix} ,$$

which one can write down in the magnetic gauge fixed by (17) as Ψ_m:

$$\Psi_m = \begin{pmatrix} \psi_+ \\ \psi_- \end{pmatrix} = \begin{pmatrix} \cos\frac{1}{2}\alpha e^{i\beta/2} & \sin\frac{1}{2}\alpha e^{-i\beta/2} \\ -\sin\frac{1}{2}\alpha e^{i\beta/2} & \cos\frac{1}{2}\alpha e^{-i\beta/2} \end{pmatrix} \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix} . \tag{29}$$

Now suppose one starts from the following SU(2) chromodynamics Lagrangian

$$\mathcal{L} = -\tfrac{1}{4}\vec{G}_{\mu\nu}{}^2 + \bar{\Psi} i\gamma^\mu D_\mu \Psi - m\bar{\Psi}\Psi \,.$$

Then one can easily write down the corresponding Lagrangian $\mathcal{L}^{(R)}$ of the restricted theory. From (21) and (29) one finds

$$\begin{aligned}\mathcal{L}^{(R)} = & -\tfrac{1}{4}F_{\mu\nu}{}^2 - \tfrac{1}{2}F_{\mu\nu}H_{\mu\nu} - \tfrac{1}{4}H_{\mu\nu}{}^2 \\ & + \bar{\psi}_+ i\gamma^\mu[\partial_\mu + \tfrac{1}{2}g(A_\mu + C^*_\mu)]\psi_+ \\ & + \bar{\psi}_- i\gamma^\mu[\partial_\mu - \tfrac{1}{2}g(A_\mu + C^*_\mu)]\psi_- \\ & - m(\bar{\psi}_+\psi_+ + \bar{\psi}_-\psi_-) \,. \end{aligned} \tag{30}$$

Notice that the source couples not only to the electric potential but also to the magnetic potential explicitly in a symmetric way. With the Lagrangian, one formally obtains the following equations of motion

$$\begin{aligned} (i\gamma^\mu \partial_\mu - m)\psi_\pm &= \mp\tfrac{1}{2}g\gamma^\mu B_\mu \psi_\pm \,, \\ \partial^\mu G_{\mu\nu} = j_\nu &= -\frac{g}{2}(\bar{\psi}_+\gamma_\nu\psi_+ - \bar{\psi}_-\gamma_\nu\psi_-) \,, \end{aligned} \tag{31}$$

where

$$B_\mu = A_\mu + C^*_\mu \,,$$

$$G_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu = F_{\mu\nu} + H_{\mu\nu} \,.$$

Thus it looks as if the theory were Abelian, with one gauge potential B_μ. However, notice that the dual field strength $G^*_{\mu\nu}$

$$G^*_{\mu\nu} = \tfrac{1}{2}\epsilon_{\mu\nu\rho\sigma} G^{\rho\sigma}$$

must now satisfy

$$\partial^\mu G^*_{\mu\nu} = \partial^\mu H^*_{\mu\nu} = k_\nu \neq 0 \tag{32}$$

when the magnetic symmetry $\hat{m}$ contains topological singularities, where k_μ is the monopole current. This assures us that indeed the theory has a nontrivial dual structure. In this respect, notice that although the Lagrangian (30) appears to be simple and Abelian in terms of the potential B_μ, it becomes highly nonlinear if one uses the coset variables α and β of $\hat{m}$. In the absence of the source, e.g., one obtains

$$\begin{aligned}\mathcal{L}_G^{(R)} &= -\tfrac{1}{4}G_{\mu\nu}{}^2 \\ &= -\tfrac{1}{4}F_{\mu\nu}{}^2 - \frac{1}{2g}F_{\mu\nu}\hat{m}\cdot(\partial_\mu\hat{m}\times\partial_\nu\hat{m}) \\ &\quad -\frac{1}{4g^2}(\partial_\mu\hat{m}\times\partial_\nu\hat{m})^2 \\ &= -\tfrac{1}{4}F_{\mu\nu}{}^2 - \frac{1}{g}\sin\alpha F_{\mu\nu}\partial_\mu\alpha\,\partial_\nu\beta \\ &\quad -\frac{1}{g^2}\frac{\sin^2\alpha}{2}[(\partial_\mu\alpha)^2(\partial_\nu\beta)^2-(\partial_\mu\alpha\,\partial_\nu\beta)^2]\,. \end{aligned} \tag{33}$$

Although the gauge-invariant coupling appears simple in terms of the magnetic potential C_μ^* in the Lagrangian (30), one may encounter a difficulty in treating C_μ^* as an ordinary potential since it must contain the well-known string singularity when $\hat{m}$ has a topological singularity. Later, in Sec. VI, we will remove this difficulty by introducing the dual magnetic potential which is regular.

V. MONOPOLES–A CLASSICAL DESCRIPTION

The dual dynamics obtained in the above can best describe a classical system. This is so since the monopoles appear as classical pointlike objects in the theory as topological singularities of $\hat{m}$. In this section we briefly review the classical monopole configurations of the theory and show that by assigning a set of properly chosen isolated singularities to $\hat{m}$ one can describe any classical dynamical system of pointlike monopoles in terms of a single multiplet $\hat{m}$.

Let us start by defining the topological charges in our theory. As is well known, the non-Abelian magnetic charge is topological in its origin.[4] For example, for SU(2) it is described by the homotopy class of the mapping $\Pi_2(S^2)$ of the two-dimensional spatial sphere S_R^2 to the coset space $S^2=\mathrm{SU}(2)/\mathrm{U}(1)$ of the internal space. Now to define the mapping one needs a scalar triplet in one's theory, at least on S_R^2. For this reason it has often been claimed that a proper definition (and thus a proper theory) of the non-Abelian magnetic charge is possible only when one has a scalar triplet *explicitly* in one's theory as in the Higgs-type theory.[4,14] However, notice that in our theory one does not need to introduce a scalar triplet explicitly since the magnetic symmetry $\hat{m}$ can be used to define the mapping $\Pi_2(S^2)$. Indeed, we will define the magnetic charge of the restricted theory by the homotopy class of the mapping $\Pi_2(S^2)$ given by $\hat{m}$:

$$\hat{m};\; S_R^2 \to S^2 = \mathrm{SU}(2)/\mathrm{U}(1)\,. \tag{34}$$

It is precisely because of this role of $\hat{m}$ that we call the additional symmetry m magnetic.

With this preliminary remark we will now present one magnetic potential per each homotopically different mapping. In other words, we will present *all* the potentials which can represent any inequivalent class of the mapping. To do this first notice that the homotopy class $\Pi_2(S^2)$ defined by

$$\hat{m} = \begin{pmatrix}\sin\theta\cos\varphi \\ \sin\theta\sin\varphi \\ \cos\theta\end{pmatrix}, \tag{35}$$

where θ and φ are the angular spherical coordinates of S_R^2, must describe the unit magnetic charge. Indeed, one can easily show that the magnetic field $H_{\mu\nu}$ obtained from (35),

$$\begin{aligned} H_{\mu\nu} &= -\frac{1}{g}\sin\alpha(\partial_\mu\alpha\,\partial_\nu\beta - \partial_\nu\alpha\,\partial_\mu\beta) \\ &= \begin{cases} -\dfrac{1}{g}\sin\theta\,, & \mu=\theta\,,\ \nu=\varphi \\ 0\,, & \text{otherwise}\end{cases} \end{aligned} \tag{36}$$

is nothing more than the Wu-Yang unit monopole.[3] Now it is easy to find out all the homotopically inequivalent classes of the mapping (34) and the corresponding monopole configurations. To do this notice that $\hat{m}$ described by (16) with

$$\alpha=\theta\,,$$
$$\beta=n\varphi\,,$$

or

$$\hat{m} = \begin{pmatrix}\sin\theta\cos n\varphi \\ \sin\theta\sin n\varphi \\ \cos\theta\end{pmatrix} \quad (n \text{ integer}) \tag{37}$$

will describe all the homotopically inequivalent mapping of (34) with the homotopy class Z

$$Z=n\,. \tag{38}$$

On the other hand, the corresponding magnetic potentials

$$C_\mu^* = \begin{cases} \dfrac{n}{g}\cos\theta\,, & \mu=\varphi \\ 0\,, & \text{otherwise}\,,\end{cases} \tag{39}$$

and their magnetic fields

$$H_{\mu\nu} = \begin{cases} -\dfrac{n}{g}\sin\theta\,, & \mu=\theta,\ \nu=\varphi \\ 0\,, & \text{otherwise}\,,\end{cases} \tag{40}$$

clearly describe the magnetic flux of the pointlike charges g_m

$$g_m = \frac{4\pi n}{g}\,. \tag{41}$$

Obviously the fact that there are no other homotopically inequivalent classes of the mapping (34) other than those described by (37) guarantees that the above magnetic potentials describe all the possible homotopically inequivalent mapping of $\Pi_2(S^2)$ with integral magnetic charges n (in the unit $4\pi/g$). In particular, as far as a static magnetic system is concerned, any magnetic field (not only the total flux but also the flux itself) must be described by one of those given above in the asymptotic region where the magnetic flux must be spherically symmetric.

The above solutions have, in fact, been known.[3,14] However, their topological meaning, in particular the existence of the magnetic direction $\hat{m}$ and its geometrical role, has so far not been fully understood. More importantly, we find the physical significance of $\hat{m}$ deep and far reaching. In fact, with a proper choice of $\hat{m}$ one could describe any classical dynamical system of the monopoles. To see this, suppose $\hat{m}$ is time dependent and has several isolated singularities at $z_\mu^{(i)} (i=1,2,\ldots,k)$, each of which carry a definite homotopy class n_i. Each singularity could be regarded as representing a monopole with the magnetic charge $4\pi n_i/g$. These will interact according to the dual dynamics of Sec. V with the magnetic current k_μ given by

$$k_\mu(x) = \sum_i \frac{4\pi n_i}{g} \int_{l_i} \delta^4(x-z)\, dz_\mu^{(i)}, \tag{42}$$

where l_i is the world line of the ith monopole. The conservation of the total magnetic charge (fixed by the homotopy of $\hat{m}$ at the infinity) is then guaranteed by the topological reason. It is amusing to notice that a single multiplet $\hat{m}$ can actually describe any classical magnetic system in such a simple manner.

VI. FIELD-THEORETIC DESCRIPTION

So far our description of the monopole has been classical; it appears as a pointlike singular object. Thus the theory obtained above can best be suited for a classical description of the dual dynamics. For a field-theoretic formulation of the theory, however, the Lagrangian (30) does contain a few undesirable features. First, in the magnetic gauge the potential C_μ^* becomes singular and carries the well-known string singularity when the monopole is present. Second, the monopole is described by a spacelike C_μ^* whereas the electric charge is described by a timelike A_μ. This second point (which clearly is related to the first one) implies that in the static limit the quarks should decouple to C_μ^* although the magnetic coupling cannot be neglected in the relativistic limit. Finally, and perhaps most importantly, in the Lagrangian (30) the monopole is simply *not* represented by a field as the source of the magnetic potential; it appears only as a classical pointlike object. In this section we will remove these undesirable features and present a field-theoretic description of the theory.

Let us start by removing the string singularity first. To do this it is crucial to observe that $H_{\mu\nu}^*$ can be described by a regular potential with *no* string singularity. This is so since $H_{\mu\nu}^*$ describes the electromagnetic field $(\vec{H}_m, -\vec{E}_m)$ created by the monopole. So we define *the dual magnetic potential* C_μ by

$$H_{\mu\nu}^* = (\partial_\mu C_\nu - \partial_\nu C_\mu). \tag{43}$$

The beauty of C_μ is that now it can describe the monopole by a timelike potential, and does not contain the string singularity anymore. Now one can replace C_μ^* in favor of C_μ in the Lagrangian (30). Then, in terms of A_μ and C_μ, the dual equations of motion (31) and (32) can be written as

$$\begin{aligned} \partial^\mu G_{\mu\nu} &= \partial^\mu F_{\mu\nu} = \partial^\mu(\partial_\mu A_\nu - \partial_\nu A_\mu) = j_\nu, \\ \partial^\mu G_{\mu\nu}^* &= \partial^\mu H_{\mu\nu}^* = \partial^\mu(\partial_\mu C_\nu - \partial_\nu C_\mu) = k_\nu. \end{aligned} \tag{44}$$

Furthermore, in terms of the dual potential the magnetic symmetry (26) could naturally be regarded as an ordinary Abelian gauge symmetry of C_μ. This means that we have really succeeded in making the magnetic symmetry a genuine Noether symmetry of the Lagrangian. In short, we have demonstrated that a topological charge can, indeed, be viewed as the dual object of a Noether charge. This allows us to treat the monopole as an ordinary source of the theory.

Now, for a field-theoretic description of the monopole there seems to be no other way than to introduce a new field for the monopole. Since the monopole as a pointlike object does not have any obvious spin structure one may describe it by a complex scalar field ϕ, and we will do so in the following. A similar field operator has been proposed by Mandelstam[6] in (3 + 1)-dimensional QCD, and earlier by 't Hooft[7] in (2 + 1)-dimensional QCD, to describe the topological objects of the corresponding theories. Naturally one would expect that ϕ should couple to C_μ minimally. However, to preserve the duality of the theory it would be better for us to include the *electric* coupling to the monopole. This can be done by formally introducing the dual electric potential A_μ^* by

$$F_{\mu\nu}^* = (\partial_\mu A_\nu^* - \partial_\nu A_\mu^*). \tag{45}$$

Clearly A_μ^* describes the electric charge by a spacelike potential, and will contain a string singularity. Now ϕ can couple minimally to B_μ^*,

$$B_\mu^* = A_\mu^* + C_\mu \tag{46}$$

with the strength $4\pi/g$. This leads us to the following Lagrangian for the restricted chromodynamics (RCD) of the color SU(2) gauge group

$$\begin{aligned}\mathfrak{L}^{(R)} = &-\tfrac{1}{4}F_{\mu\nu}{}^2 - \tfrac{1}{2}F_{\mu\nu}H_{\mu\nu} - \tfrac{1}{4}H_{\mu\nu}^{*\,2}\\ &+\overline{\psi}_+ i\gamma^\mu(\partial_\mu + \tfrac{1}{2}gB_\mu)\psi_+ + \overline{\psi}_- i\gamma^\mu(\partial_\mu - \tfrac{1}{2}gB_\mu)\psi_-\\ &+m(\overline{\psi}_+\psi_+ + \overline{\psi}_-\psi_-) + \left|\left(\partial_\mu + i\frac{4\pi}{g}B_\mu^*\right)\phi\right|^2\\ &-V(\phi^*\phi)\,.\end{aligned} \tag{47}$$

There are two points to be clarified in the above Lagrangian. First, notice that the term $H_{\mu\nu}{}^2$ has been replaced by $H_{\mu\nu}^{*\,2}$ with *no* change of the signature. At first glance this replacement appears to be wrong, but in fact it is correct. The correct signature is obtained by requiring that the Hamiltonian of the theory should remain the same while one changes the potential C_μ^* to the regular one C_μ. One must be careful not to change the physics of the theory by changing the variables. The other point is that although the singular potentials A_μ^* and C_μ^* do appear in the Lagrangian, only the regular ones A_μ and C_μ must be regarded as dynamically independent variables. Now, of course, we must point out that one should not take the singular potentials and the couplings involving them too literally. In field-theoretic formulation the singular potentials are indeed ill-defined quantities. At best, A_μ^* and C_μ^* in (43) and (45) are to be interpreted as some functionals of the regular potentials. Consequently the couplings involving A_μ^* and C_μ^* in the above Lagrangian appear to be ill-defined. Nonetheless, from the physical point of view the interaction between the quarks (the monopoles) and the magnetic (the electric) field created by the monopoles (the quarks, respectively) surely exists and should not be excluded from the Lagrangian. These interactions are represented symbolically in terms of the singular potentials in the above Lagrangian as a mnemonic means, and the Lagrangian must be interpreted accordingly.

Admittedly our derivation of the regular Lagrangian (47) for RCD from the singular one (30) is not so rigorous in that we have not established the existence of the regular monopole field ϕ from the Lagrangian (30), but introduced the field by hand from physical grounds. So it remains to be seen whether the Lagrangian (47) can really be derived from the singular one (30) in a mathematically rigorous way. Nonetheless, one may take the Lagrangian (47) as an effective Lagrangian which can govern the dual dynamics at the phenomenological level, just as one could view the Ginsburg-Landau Lagrangian as an effective Lagrangian for the theory of superconductivity, and may pursue its consequences. An immediate consequence is that indeed the Lagrangian (47) can explain color confinement in QCD, which we discuss in the following section.

VII. DISCUSSIONS

In this paper we have proposed a gauge theory of non-Abelian monopoles which has a built-in dual structure. A crucial feature of the theory is the fact that it consists of a self-consistent subset of a full non-Abelian gauge theory. The theory is obtained by imposing an extra magnetic symmetry to the original theory. The virtue of the magnetic symmetry is that it preserves essential features of the original gauge theory. It keeps the full gauge degrees of freedom intact, and *a priori* is not worse than the symmetry that has already been imposed, namely the gauge symmetry itself. Besides, it preserves the natural topological structure of the full gauge symmetry, and thereby describes the dual dynamics of the non-Abelian monopoles.

Certainly the restricted theory is interesting in its own right. Furthermore, it could well be that for a certain kind of problem the restricted theory could give us results which would not be qualitatively different from what one would expect in the unrestricted theory, and could actually simplify the problems. Thus one may hope that by understanding the restricted theory one could obtain a better insight of the complicated non-Abelian structure of the fully unrestricted theory.

Now we will briefly outline how the restricted chromodynamics (RCD) could actually explain the confinement of the color in QCD. It has recently been emphasized by Mandelstam[6] and by 't Hooft[7] that the monopoles and a possible dual structure of QCD might play an important role for quark confinement. This kind of speculation can be made more precise within the framework of RCD. So let us for the time being take the attitude that the requirement of the magnetic symmetry somehow does not make a real difference for the discussion of quark confinement in QCD, and look for the possible confinement mechanism in RCD. Now the confinement in RCD could be obtained by the following observations. First, in the absence of the quarks *the effective potential* of the Lagrangian (47) (which could be estimated using the loop expansion approximation) will either break the magnetic symmetry or else will preserve it. So, logically, the theory has two phases, the normal phase where the magnetic symmetry is preserved and the abnormal phase where the symmetry is dynamically broken. In the strong-coupling limit the dynamical symmetry breaking indeed could occur, as has been argued by Coleman and Weinberg.[15] In this case the physical vacuum is made of the Bose condensation of the monopoles. The dual dynamics will then ensure us the perfect color dielectric effect (i.e., the dual Meissner effect) that will confine any colored flux in a finite region of space. Consequently both the quarks and the monopoles will disappear from the physical spectrum of the theory. In short, *the dynamical breaking of the magnetic symmetry will guarantee quark confinement in RCD.* In the weak-coupling limit, however, the magnetic symmetry is not likely to be broken since in this limit the magnetic coupling may become too strong to allow us the monopole condensation for the physical vacuum. In this case not only the quarks but also the monopoles become unavoidable as physical states. Now in view of the fact that the monopoles have not been observed one may say that in reality the magnetic symmetry is indeed broken, and the confinement is enforced. But one has to bear in mind that this argument holds only within the framework of RCD. A more detailed discussion on this issue is available in a separate paper.[16]

Although the group SU(2) is examined in detail in this paper as an example, we should like to emphasize that the theory can be defined for an arbitrary group G. Furthermore, the general structure of the theory, i.e., the existence of the dual structure and the magnetic potential with the corresponding magnetic gauge degrees of freedom, the existence of two phases and the magnetic confinement of the colored flux all remain unchanged, as we will see in a later paper. The only complication is that for a group of rank higher than one, there are different types, indeed as many as the number of the possible little groups, of the magnetic symmetry which one properly has to take into account. A more detailed treatment of the theory with an arbitrary group G, especially the color SU(3), will be presented in a later paper.

ACKNOWLEDGMENTS

It is a great pleasure to thank Dr. D. Maison and Dr. S. Schlieder for many critical discussions, the careful reading of the manuscript, and the constant encouragement throughout the period of the work. Also special thanks are due to S. O. Ahn for the never-failing encouragement.

[1]P. A. M. Dirac, Proc. R. Soc. London A133, 60 (1931); Phys. Rev. 14, 817 (1948).
[2]See, e.g., J. Schwinger, Phys. Rev. 144, 1087 (1966); Phys. Rev. D 12, 3105 (1975); D. Zwanziger, *ibid.* 3, 880 (1971); R. A. Brandt, F. Neri, and D. Zwanziger, Phys. Rev. Lett. 40, 147 (1978).
[3]T. T. Wu and C. N. Yang, in *Properties of Matter under Unusual Conditions*, edited by H. Mark and S. Fernbach (Interscience, New York, 1969); G. 't Hooft, Nucl. Phys. B79, 276 (1976); A. M. Polyakov, Pis'ma Zh. Eksp. Teor. Fiz. 20, 430 (1974) [JETP Lett. 20, 194 (1974)].
[4]J. Arafune, P. G. O. Freund, and G. G. Goebel, J. Math. Phys. 16, 433 (1975); S. Coleman, in *New Phenomena in Subnuclear Physics*, Proceedings of the 14th Course of the International School of Subnuclear Physics, Erice, 1975, edited by A. Zichichi (Plenum, New York, 1977).
[5]P. Goddard, J. Nuyts, and D. Olive, Nucl. Phys. B125, 1 (1977); C. Montonen and D. Olive, Phys. Lett. 76B, 117 (1977); P. G. O. Freund, Int. J. Theor. Phys. 17, 287 (1978).
[6]S. Mandelstam, Phys. Rep. 23C, 245 (1976); Phys. Rev. D 19, 249 (1979).
[7]G. 't Hooft, Nucl. Phys. B138, 1 (1978).
[8]This idea, which is the non-Abelian generalization of the Abelian Kaluza-Klein-Jordan-Thirry theory, has been suggested by B. de Wit and A. Trautman. An attempt to confirm the idea has subsequently been made by R. Kerner. Unfortunately his calculation misleads the metric theory to a (gauge-noninvariant) massive vector theory. See B. de Wit, in *Dynamical Theories of Groups and Fields*, edited by E. W. Montroll, G. H. Vineyard, and M. Levy (Gordon and Breach, New York, 1965); A. Trautman, Rep. Math. Phys. 1, 29 (1970); R. Kerner, Ann. Inst. Henri Poincaré 9, 143 (1968).
[9]Y. M. Cho, J. Math. Phys. 16, 2029 (1975). This paper contains a correct confirmation of the idea as well as all the earlier historical references.
[10]Y. M. Cho and P. S. Jang, Phys. Rev. D 12, 3189 (1975).
[11]Y. M. Cho and P. G. O. Freund, Phys. Rev. D 12, 1711 (1975).
[12]Y. M. Cho and D. Maison (unpublished).
[13]Our duality, however, appears to be different from the ones proposed in Refs. 5 and 6.
[14]E. Corrigan, D. Olive, D. Fairlie, and J. Nuyts, Nucl. Phys. B106, 475 (1976).
[15]S. Coleman and E. Weinberg, Phys. Rev. D 7, 1888 (1973).
[16]Y. M. Cho, MPI Report No. MPI-PAE/PTh 11/79, revised (unpublished).