6

Magnetic monopoles in larger gauge groups

Having studied in some detail the monopole that arises when SU(2) is broken to U(1), let us now consider the possibilities that occur with larger, and possibly more realistic, gauge groups. Perhaps the most important of these are the groups that occur in grand unified theories. As we will see, magnetic monopoles are a generic prediction of all such theories.

6.1 Larger gauge groups—the external view

Although our primary interest is in monopoles with nonsingular cores, which arise when a gauge group G is spontaneously broken to a subgroup H with $\pi_2(G/H)$ nontrivial, I will begin by focusing on the long-range field, the part outside the monopole core, and determine the possible magnetic charges. This is essentially an extension of Dirac's analysis to the case where the unbroken gauge group H is larger than U(1) and, in general, non-Abelian. Like Dirac's analysis, this analysis applies equally well to singular point monopoles and to nonsingular solitons. Also as with the Dirac analysis, we will find that in some cases there are allowed magnetic charges for which there are no nonsingular monopoles.

Let us therefore start by assuming that we have an unbroken gauge group H, and that the magnetic components F_{ij} of the field strength have long-range tails that fall as $1/r^2$. Let us further assume that the vector potential can be expanded in inverse powers of r, with its Cartesian components falling as $1/r$ or faster, and that[1]

$$A_i = \frac{f_i(\theta, \phi)}{r} + O(1/r^2), \tag{6.1}$$

where A_i and f_i should be understood to be matrices representing elements of the Lie algebra of H. The $1/r^2$ terms do not contribute to the $1/r^2$ part of the

[1] In order to simplify the notation, the gauge fields in this section have been rescaled so as to absorb the gauge coupling constant.

 https://doi.org/10.1017/CBO9781139017787.007

magnetic field, and so will henceforth be omitted. If we now go over to spherical components, these assumptions imply that A_θ and A_ϕ are independent of r at large distance. Finally, we will restrict ourselves to static fields with no electric charge, and with $A_0 = 0$.

The first step is to choose a gauge where $A_r = 0$ outside a sphere of radius R. (Excluding the region $r < R$ avoids any issues of singularities at the origin, and has no effect on the fields at large distance.) This can be done by means of a gauge transformation with

$$U^{-1}(r, \theta, \phi) = P \exp\left[\frac{i}{g}\int_R^r dr'\, A_r(r', \theta, \phi)\right], \tag{6.2}$$

with the P denoting path ordering. With this gauge choice and our dropping of the $1/r^2$ terms in A_i, the only nonzero component of the field strength is $F_{\theta\phi}$, which is independent of r.

Next, we set $A_\theta = 0$ in this asymptotic region by means of an r-independent gauge transformation obtained in a similar fashion by integrating along arcs of fixed ϕ starting at the north pole and moving downward. The field equation

$$0 = D_a(\sqrt{g}\, F^{a\phi}) \tag{6.3}$$

then implies that A_ϕ, the only nonzero component of the gauge potential, is of the form

$$A_\phi = C(\phi) + \frac{Q_M(\phi)}{4\pi}\cos\theta. \tag{6.4}$$

To avoid a singularity at the north pole, $\theta = 0$, we must set $C(\phi) = -Q_M(\phi)/4\pi$. The field equation

$$0 = D_a(\sqrt{g}\, F^{a\theta}) \tag{6.5}$$

then tells us that Q_M is independent of ϕ, so that

$$A_\phi = \frac{1}{4\pi} Q_M(\cos\theta - 1) \tag{6.6}$$

and (returning to Cartesian components)

$$B_i = -\frac{1}{2}\epsilon^{ijk}F_{jk} = \frac{Q_M}{4\pi}\frac{\hat{r}_i}{r^2}. \tag{6.7}$$

This looks very much like the U(1) case, with a Coulomb magnetic field and a Dirac vector potential with its string along the negative z-axis. There are two crucial differences. First, the U(1) charge was simply a number, whereas now Q_M is a matrix in the appropriate representation of the Lie algebra. Second, Q_M, like B_i itself, is not gauge invariant, but only gauge covariant. Had we made other gauge choices, it could even have varied with angle.

Despite these differences, the derivation of the quantization condition goes through pretty much as before. Requiring that the potential of Eq. (6.6) be related by a single-valued gauge transformation to an equivalent potential with the Dirac string along the positive z-axis gives the condition [96–98]

$$e^{iQ_M} = I. \tag{6.8}$$

The corresponding condition in the U(1) case, Eq. (5.12), was required to hold for all electric charges q that are present in the theory. Similarly, Eq. (6.8) must hold with Q_M a matrix in any representation of the Lie algebra that actually occurs in the theory.

Let us start by considering the case where the unbroken gauge symmetry has the algebra of SU(2), with the generators taken to be the T_a ($a = 1, 2, 3$) with the standard normalization. Since any element in the Lie algebra can be rotated by a global gauge transformation to be proportional to T_3, there is no loss of generality in writing

$$Q_M = 4\pi k T_3. \tag{6.9}$$

The quantization condition then becomes

$$e^{4\pi k T_3} = I, \tag{6.10}$$

so that for every eigenvalue t_3 we must have

$$2kt_3 = n, \tag{6.11}$$

with n an integer.

If all of the fields in the theory transform under integer "spin" representations of SU(2), then the t_3 will all be integers and k can be either an integer or an integer plus 1/2. If there are also fields transforming under half-integer representations, then k must be an integer.

These conditions can be rephrased by being more careful about specifying the gauge group, keeping in mind that SO(3) and SU(2) share the same Lie algebra. The allowed eigenvalues of T_3 are known as weights. These are all integers for SO(3), and either integers or half-integers for SU(2). If the theory only has fields in integer representations, and thus only integer weights, then the gauge group can be taken to be either SO(3) or SU(2); let us choose the former. If there are also fields transforming under half-integer representations, then the gauge group is unambiguously SU(2). What we have found is that if the "electric" gauge group is SU(2), then the "magnetic weight" k must be a weight of the "magnetic group" SO(3). Conversely, if the electric group is SO(3), then the allowed magnetic weights are those of SU(2).

Configurations with different values of k are not all physically distinct. Because a global gauge transformation can reverse the sign of T_3, monopoles with magnetic weights k and $-k$ are gauge equivalent and can be continuously transformed one into the other. (It is possible to continuously connect these two discrete values because for the interpolating configurations Q_M is not simply a multiple of T_3.)

This leads to a curious phenomenon. Consider a configuration with two monopoles, each with magnetic charge $Q_M = 2\pi T_3$, held fixed at some separation much larger than their core size. Outside the monopole cores there is an obvious solution of the field equations in which the magnetic field is equal to $2\pi T_3$ times the ordinary electrostatic field corresponding to two like-sign unit charges. However, changing the sign of one of the charges by a gauge rotation leads us to another quasi-Abelian solution, of lower energy, formed from the electrostatic field of two charges with opposite sign. If the field and charges are initially in the former solution, even the slightest perturbation should be enough to cause them to transform into the latter, radiating off the excess energy in the process.

In this process, the total charge (i.e., that seen at spatial infinity) will have changed from $4\pi T_3$ to 0. By considering configurations with more monopoles, we can see that configurations with total charges differing by any integer multiple of $4\pi T_3$ can be continuously connected (although the intermediate configurations may not be solutions of the static field equations). Hence, two configurations can be continuously deformed into one another if they both have integer, or both have half-integer magnetic weights. What if one has an integer and the other a half-integer magnetic weight?

One can define a topological quantity that shows that these cannot be connected [99]. Consider a sphere at very large radius, where the gauge fields can be assumed to take on their asymptotic form. This sphere can be covered by a family of loops, each of which begins and ends at the north pole. One such family consists of loops $C(\tau)$ that go from the north pole ($\theta = 0$) to the south pole ($\theta = \pi$) along the arc $\phi = 0$ and then back along the arc $\phi = 2\pi\tau$. As τ ranges from 0 to 1, these cover the sphere.

Each such loop defines a group element

$$h(\tau) = P \exp\left[-i \int_{C(\tau)} d\boldsymbol{\ell} \cdot \mathbf{A}\right], \tag{6.12}$$

where A_i is understood to be a matrix and the P indicates path ordering. This Wilson loop is gauge invariant, so we can evaluate it in a gauge where $A_\theta = 0$ on the sphere. For nonzero Q_M there is a gauge singularity at the south pole, with $A_\phi \neq 0$. Including an infinitesimal section around this singularity, we can view the Wilson loop as being composed of three parts: (1) $0 \le \theta \le \pi$, with $\phi = 0$; (2) $0 \le \phi \le 2\pi\tau$, with $\theta = \pi$; and (3) $0 \le \theta \le \pi$, with $\phi = 2\pi\tau$. Since $A_\theta = 0$, the first and last segments do not contribute to the line integral. With A_ϕ given by Eq. (6.6), the integral is straightforward to evaluate, and we have

$$h(\tau) = \exp\left[i\tau Q_M\right]. \tag{6.13}$$

Clearly $h(0) = I$. By the quantization condition of Eq. (6.8), we also have $h(1) = I$. Thus, $h(\tau)$ for $0 \le \tau \le 1$ defines a loop in the (electric) gauge group H and so can be assigned to an element of $\pi_1(H)$. If H is the simply connected

SU(2), then π_1 has only a single element, and the topological invariant is trivial. If instead $H = \mathrm{SO}(3)$, then π_1 has two elements, one arising when k is an integer and the other when it is a half-integer. Hence, a configuration with integer magnetic weight cannot be continuously deformed into one with a half-integer magnetic weight.

Now note that the energy in the long-range Coulomb field of a monopole is proportional to $\mathrm{tr}\, Q_M^2$. Since configurations with the same topological charge can have different values for this quantity, one might wonder if solutions with higher values were unstable and could decay to solutions with lower values of $\mathrm{tr}\, Q_M^2$. This instability is not completely obvious, both because there might be an energy barrier along the path joining the two solutions, and because we have not considered the fields in the core region. These concerns are dispelled by detailed analysis of small fluctuations about a pure Coulomb solution, as was done by Brandt and Neri [100] and by Coleman [101].[2] They showed that, even if one imposes the restriction that the field be held fixed inside some sphere of radius R, there is always such an instability. Stable solutions must have the minimum value of $\mathrm{tr}\, Q_M^2$ consistent with their topological charge. In the present case, monopoles can only be stable if $k = \pm 1/2$; monopoles with integer magnetic weights are all unstable.

The language of roots and weights[3] is ideally suited for generalizing this discussion to the case of an arbitrary semisimple unbroken gauge group H [97, 98]. The Cartan subalgebra can be chosen to include any given element of the Lie algebra.[4] Hence, there is no loss of generality in taking Q_M to be a linear combination of the generators H_a of the Cartan subalgebra and writing

$$Q_M = 4\pi \mathbf{k} \cdot \mathbf{H}. \tag{6.14}$$

The components k_a are called the magnetic weights of the monopole. In a basis where the H_a, and hence Q_M, are simultaneously diagonalized, the diagonal elements of Q_M are of the form $4\pi i \mathbf{k} \cdot \mathbf{w}$, where $\mathbf{w}$ is a weight vector of the representation in which the generators are being expressed. The quantization condition of Eq. (6.8) becomes the requirement that for every weight $\mathbf{w}$ that appears in the theory

$$2\mathbf{k} \cdot \mathbf{w} = n \tag{6.15}$$

for some integer n. If the group is SU(2), Eqs. (6.14) and (6.15) reduce to our previous results, Eqs. (6.9) and (6.11), respectively.

[2] Their analysis does not apply in the BPS limit, in which the Higgs field also has a long-range tail.

[3] Root and weight vectors are reviewed in Appendix A.

[4] For the important case of SU(N), where the Cartan subalgebra can be taken to be the elements that are diagonal in the fundamental representation, this is just the statement that any Hermitian matrix can be diagonalized.

Any weight $\mathbf{w}$ and root $\boldsymbol{\alpha}$ must satisfy

$$\frac{2\mathbf{w}\cdot\boldsymbol{\alpha}}{\boldsymbol{\alpha}^2} = N, \tag{6.16}$$

with N an integer. Hence one solution of the quantization condition is given by

$$\mathbf{k} = \sum n_{\boldsymbol{\alpha}} \frac{\boldsymbol{\alpha}}{\boldsymbol{\alpha}^2} = \sum n_{\boldsymbol{\alpha}} \boldsymbol{\alpha}^*, \tag{6.17}$$

with the $n_{\boldsymbol{\alpha}}$ integers; i.e., by taking $\mathbf{k}$ to be an element of the root lattice of a dual group H^v whose roots are the dual roots $\boldsymbol{\alpha}^* = \boldsymbol{\alpha}/\boldsymbol{\alpha}^2$. If H is simple and the $\boldsymbol{\alpha}^*$ differ from the $\boldsymbol{\alpha}$ only by an overall rescaling, then H and H^v share the same Lie algebra, although they will not in general be the same group. For a semisimple H, equality of the two Lie algebras only requires that the rescaling be the same within each simple factor.

If H is the universal covering group of the Lie algebra, so that all representations appear, Eq. (6.17) gives the only solution for $\mathbf{k}$. If not, there are additional solutions. For example, if H is the adjoint group, whose representations all have weights lying on the root lattice, then any weight $\mathbf{w}$ can be written as an integral sum of the roots $\boldsymbol{\alpha}$. Applying Eq. (6.15) to the adjoint representation, whose weights are equal to the roots, yields the requirement that

$$2\mathbf{k}\cdot\boldsymbol{\alpha} = \frac{2\mathbf{k}\cdot\boldsymbol{\alpha}^*}{\boldsymbol{\alpha}^{*2}} = N', \tag{6.18}$$

be an integer. This is the same as requiring that $\mathbf{k}$ be a weight of the Lie algebra of H^v.

Thus, just as with the case of SU(2), we have an electric group whose weights correspond to the representations of the elementary fields and a magnetic group whose weights correspond to the allowed magnetic charges. The former has roots $\boldsymbol{\alpha}$, while their duals $\boldsymbol{\alpha}^*$ are the roots of the latter. The larger one group is, the smaller the other must be. If one is the universal covering group of its algebra, the other is the adjoint group.[5]

We saw for the case of SU(2) that solutions with magnetic weights k and $-k$ are gauge-equivalent. In the general case, magnetic weight vectors $\mathbf{k}$ and $\mathbf{k}'$ that are related by Weyl reflections lead to physically equivalent solutions.

For SU(2) we also found that any two configurations with integer k or any two with half-integer k could be continuously deformed into one another. On the other hand, a configuration with integer k and one with half-integer k could not be continuously connected, a fact that could be verified by noting that the loop defined by $h(\tau)$ associated them with different elements of $\pi_1(H)$.

The generalization of this result can be expressed in terms of sublattices of the weight lattice. The weight lattice of the simply connected covering group can be

[5] The possible intermediate cases are discussed in [98].

decomposed into sublattices such that any two weights in the same sublattice differ by an integer sum of root vectors, while the difference between weights in different sublattices is never of this form. Just as in the case of $H = \mathrm{SU}(2)$, with its two sublattices, these sublattices are in one-to-one correspondence with the elements of $\pi_1(H)$. Configurations with magnetic weights in the same sublattice can be continuously deformed into one another, while those with weights in different sublattices cannot.

Finally, the Brandt–Neri–Coleman analysis generalizes immediately to larger groups, and shows that stable solutions must have the minimum value of $\mathrm{tr}\, Q_M^2$ consistent with their topological charge. In particular, solutions with nonzero magnetic weights lying in the root lattice, which includes the origin, are all unstable.

The discussion thus far has assumed that H is semisimple. If this assumption is relaxed, then the U(1) electric and magnetic charges also contribute to the quantization condition. For simplicity, let us assume that H contains a single U(1) factor with generator $T_{\mathrm{U}(1)}$. The magnetic charge will then have a non-Abelian component, given by Eq. (6.14), and an Abelian component equal to $4\pi k_{\mathrm{U}(1)} T_{\mathrm{U}(1)}$. If there is an electrically charged particle with non-Abelian electric weight $\mathbf{w}$ and U(1) electric charge $q_{\mathrm{U}(1)}$, then Eq. (6.15) is replaced by

$$2\mathbf{k} \cdot \mathbf{w} + 2k_{\mathrm{U}(1)} q_{\mathrm{U}(1)} = n, \tag{6.19}$$

with an obvious generalization if there are multiple U(1) factors. The theory always includes the gauge bosons of the semisimple part of H, which carry no U(1) charge and have weights in the root lattice. Imposing the quantization condition using these shows that $\mathbf{k}$ must lie in the magnetic weight lattice.

Let us consider an explicit example. In the real world, there is an unbroken SU(3)×U(1) gauge group, with the first factor corresponding to QCD and the latter to electromagnetism. Experimentally, there is a correlation between the color and electromagnetic charges, in that particles invariant under SU(3) or transforming under zero triality representations have integer U(1) charges ne, while particles corresponding to triality ± 1 representations (e.g., quarks and antiquarks, respectively), have fractional electric charges of the form $(N \pm \frac{1}{3})e$.

If the monopole's magnetic charge has no SU(3) component, we recover the original Dirac quantization condition

$$2k_{\mathrm{U}(1)} q_{\mathrm{U}(1)} = n\,. \tag{6.20}$$

The same is true if the monopole has a magnetic weight in the dual root lattice, since then $2\mathbf{k} \cdot \mathbf{w}$ is an integer for any $\mathbf{w}$. For either case the existence of down quarks with U(1) electric charge $-e/3$ implies that the minimum U(1) magnetic charge is $4\pi k_{U(1)} = 6\pi/e$.

Now suppose that the magnetic weight $\mathbf{k}$ of the monopole lies in the triality 1 sublattice. Applying Eq. (6.19) with an electrically charged particle of SU(3) triality 0 leads again to Eq. (6.20). Because all triality 0 particles have integer electric charges, this quantization condition is satisfied if the magnetic charge is a multiple of $2\pi/e$. However, we must also consider the electrically charged particles with nonzero triality and fractional U(1) charge. For any triality ± 1 representation with weight $\mathbf{w}$ the product $2\mathbf{k} \cdot \mathbf{w}$ is of the form $n_1 \pm \frac{2}{3}$, and the U(1) electric charge is $(n_2 \pm \frac{1}{3})e$. Substituting these into Eq. (6.19) leads to the requirement that $2k_{\mathrm{U}(1)}e$ be an integer, so that the U(1) component of the magnetic charge is still of the form

$$4\pi k_{\mathrm{U}(1)} = \frac{2\pi n}{e}. \tag{6.21}$$

The net effect of all of this is that if a color singlet electron, with charge $-e$, goes around the Dirac string of this monopole, it acquires a phase of 2π and the string is unobservable, as required. If a quark is carried around the same path, the U(1) charges give a phase that is less than 2π, but the deficit is made up by the phase from the SU(3) magnetic and electric charges.

6.2 Larger gauge groups—topology

The analysis in the previous section focused on the long-range fields, extending Dirac's analysis to determine what magnetic charges are allowed by the requirement that the Dirac string be unobservable. However, not every allowed magnetic charge can be realized in a nonsingular soliton. For such a soliton to arise and be topologically stable, the manifold of vacuum solutions, $\mathcal{M} = G/H$, must have a nontrivial second homotopy group. In this section I will illustrate some of the possible behaviors. In these examples Eq. (4.51), which reduces the calculation of $\pi_2(G/H)$ to the calculation of $\pi_1(H)$, will be of considerable help.

6.2.1 SU(3) broken to SU(2)×U(1)

Consider a $G = \mathrm{SU}(3)$ gauge theory with gauge coupling g and an octet Higgs field that can be viewed as a traceless 3×3 Hermitian matrix. Such a matrix can always be diagonalized, and so can be characterized by its eigenvalues. If the scalar field potential is such that the Higgs vacuum expectation value has three unequal eigenvalues, the symmetry is broken to U(1)×U(1). Let us focus instead on the other possibility, with two equal eigenvalues and a vacuum expectation value

$$\phi_0 = \mathrm{diag}\,(2b, -b, -b). \tag{6.22}$$

The generators of the unbroken symmetry can then be taken to be

$$T_8 = \mathrm{diag}\left(\frac{1}{\sqrt{3}}, -\frac{1}{2\sqrt{3}}, -\frac{1}{2\sqrt{3}}\right) \tag{6.23}$$

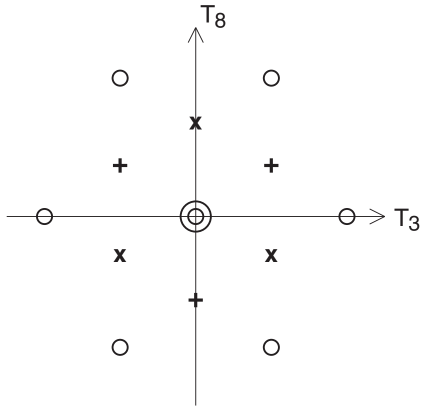

Fig. 6.1. The weights of the octet (open circles), triplet (+'s), and antitriplet (×'s) representations of SU(3). Note that the octet has a pair of weights at $T_3 = T_8 = 0$.

together with the three generators

$$T_a = \begin{pmatrix} 0 & 0 \\ 0 & \dfrac{\tau_a}{2} \end{pmatrix}, \quad a = 1, 2, 3, \tag{6.24}$$

obtained by embedding the Pauli matrices in the lower right 2×2 block. These generate an $\mathrm{SU}(2) \times \mathrm{U}(1)$ algebra, but because $e^{2\pi i T_a} e^{2\sqrt{3}\,\pi i T_8} = I$, the actual unbroken symmetry group is $H = [\mathrm{SU}(2) \times \mathrm{U}(1)]/Z_2$. The generators of the Cartan subalgebra can be chosen to be $H_1 = T_3$ and $H_2 = T_8$.

After the symmetry breaking the spectrum of states from the octet fields includes massive scalars and massless vectors in the singlet and triplet representations of SU(2), all with $T_8 = 0$, and two SU(2) doublets of massive vectors with $T_8 = \pm\frac{\sqrt{3}}{2}$. The corresponding weight vectors are shown in Fig. 6.1. Also shown are the weights that would appear if the theory also included fields transforming under the triplet and antitriplet representations of the original SU(3).

Nonsingular monopoles correspond to nontrivial elements of $\pi_2(G/H)$. Because SU(3) is simply connected we can make use of Eq. (4.51), which gives

$$\pi_2(G/H) = \pi_1(H) = \pi_1\{[\mathrm{SU}(2) \times \mathrm{U}(1)]/Z_2\} = Z, \tag{6.25}$$

with the Z arising from the U(1). Hence, there are topologically stable nonsingular monopoles carrying U(1) magnetic charge. We will see that they can also carry SU(2) magnetic charge.

A nonsingular spherically symmetric configuration with unit magnetic charge can be obtained by embedding the 't Hooft–Polyakov ansatz of Eq. (5.60) in the SU(2) subgroup lying in the upper left 2×2 block of the SU(3) matrices; i.e., the subgroup generated by $\frac{1}{2}\lambda_1$, $\frac{1}{2}\lambda_2$, and $\frac{1}{2}\lambda_3$, where the λ_a are the Gell-Mann matrices defined in Eq. (A.3). This by itself would not give the correct eigenvalues for ϕ at spatial infinity, so a term proportional to λ_8 must be added. Thus, we have [102, 103]

$$\begin{aligned}\phi &= \frac{1}{2}\sum_{a=1}^{3}\hat{r}^a\lambda_a h(r) + \frac{1}{2}\lambda_8 j(r),\\ A_i &= \frac{1}{2}\sum_{a=1}^{3}\epsilon^{iam}\hat{r}^m\lambda_a\left[\frac{1-u(r)}{gr}\right].\end{aligned} \tag{6.26}$$

Requiring that ϕ be nonsingular at the origin and be gauge equivalent to ϕ_0 at spatial infinity gives the boundary conditions

$$\begin{aligned} h(0) &= 0, \quad j'(0) = 0,\\ h(\infty) &= 3b, \quad j(\infty) = \sqrt{3}\,b\end{aligned} \tag{6.27}$$

for the scalar field, while for the gauge field we have $u(0) = 1$ and $u(\infty) = 0$, as in the SU(2) monopole. Solving the field equations with this ansatz and these boundary conditions gives a monopole with a mass

$$M_{\rm mon} \approx \frac{4\pi(3b)}{g}. \tag{6.28}$$

According to the analysis of the previous section, the asymptotic magnetic field should take the Coulomb form of Eq. (6.7), where Q_M can be gauge rotated to

$$Q_M = \frac{4\pi}{g}(k_1H_1 + k_2H_2) = \frac{4\pi}{g}(k_1T_3 + k_2T_8). \tag{6.29}$$

Indeed, the monopole solution that follows from the ansatz of Eq. (6.26) has a magnetic charge that along the positive z-axis is given by

$$Q_M = \frac{4\pi}{g}\,\mathrm{diag}\left(\frac{1}{2}, -\frac{1}{2}, 0\right), \qquad k_1 = -\frac{1}{2}, \quad k_2 = \frac{\sqrt{3}}{2}. \tag{6.30}$$

Our ansatz was obtained by making use of the SU(2) subgroup lying in the upper left 2×2 block of the SU(3) matrices. We could equally well have used the SU(2) subgroup defined by the four corner elements of the SU(3) matrices. In this case the magnetic field along the positive z-axis would have given

$$Q_M = \frac{4\pi}{g}\,\mathrm{diag}\left(\frac{1}{2}, 0, -\frac{1}{2}\right), \qquad k_1 = \frac{1}{2}, \quad k_2 = \frac{\sqrt{3}}{2}. \tag{6.31}$$

The two ansatzes are related by a global gauge rotation in the unbroken SU(2).

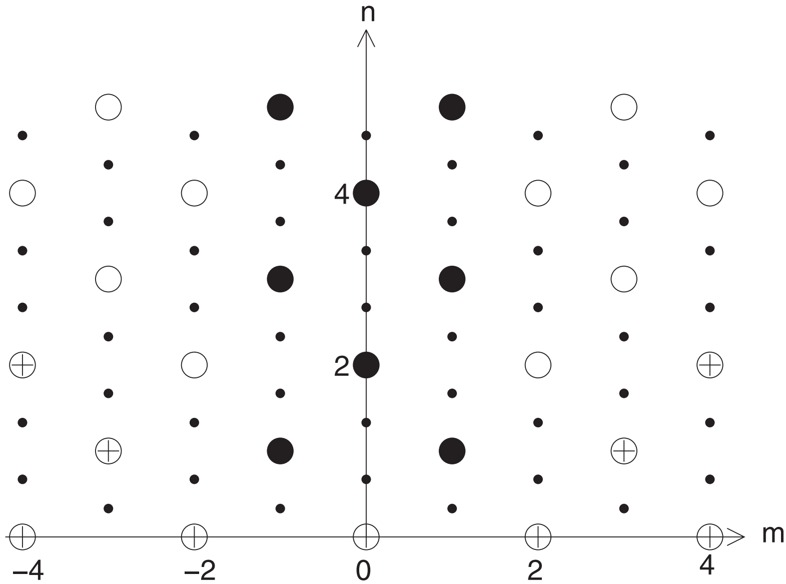

Fig. 6.2. Allowed magnetic weights for SU(3) broken to SU(2)×U(1). Here $m = 2k_1$ and $n = 2k_2/\sqrt{3}$, with the k_j defined as in Eq. (6.29). The large circles represent magnetic weights that are consistent with any representation for the electrically charged particles, while the weights denoted by small circles are only allowed if the electrically charged particles are all in triality zero representations. Only the former can be obtained from configurations containing collections of nonsingular monopoles ($n = 1$) and antimonopoles ($n = -1$). Of these, the large solid circles represent weights that are stable by the Brandt–Neri–Coleman analysis; these can all be obtained from configurations containing only monopoles. The large open circles can also be obtained using only monopoles, but are unstable by this analysis. The large circles with crosses require assemblies of monopoles and antimonopoles, and are also unstable. The pattern of weights for $n < 0$ is similar, with the roles of monopoles and antimonopoles interchanged.

By assembling a number of these monopoles, using various combinations of the two forms, we can construct configurations with $k_1 = m/2$ and $k_2 = n\sqrt{3}/2$, where m and n are either both even or both odd integers and $|m| \le n$. These points are indicated in Fig. 6.2. It is n, from the coefficient of the U(1) generator, that is the conserved topological charge. Configurations with different values of m can be deformed into one another and, by the results of Brandt, Neri, and Coleman, will reduce their long-range non-Abelian components until $k_1 = \pm\frac{1}{2}$ or 0. If the only fields in the theory are the SU(3) adjoint Higgs and gauge fields, then the generalized Dirac quantization condition allows a larger set of charges, which are also shown in Fig. 6.2. The absence of nonsingular solitons with these charges can be understood by noting that they would be forbidden if fields with nonzero triality were added to the theory.

6.2.2 A Z_2 monopole

Let us again consider an SU(3) gauge theory with gauge coupling g, but this time with a Higgs field S that transforms according to the **6** representation of SU(3) [104]. This can be viewed as a symmetric 3×3 matrix that transforms as

$$S \to USU^T, \tag{6.32}$$

where U is an SU(3) matrix and a superscript T denotes the transpose. If the Higgs potential is minimized by

$$S_0 = \sigma \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \tag{6.33}$$

the unbroken symmetry is the SO(3) subgroup generated by the antisymmetric matrices λ_2, λ_5, and λ_7. Because the triplet of the original SU(3) transforms as a vector under SO(3), only integer-spin representations of SO(3) appear, confirming that the unbroken group really is SO(3), and not SU(2).

Making use of Eq. (4.51), we find that the second homotopy group of the vacuum manifold is

$$\pi_2[\mathrm{SU}(3)/\mathrm{SO}(3)] = \pi_1[\mathrm{SO}(3)] = Z_2. \tag{6.34}$$

Thus, we should expect to find monopoles with Z_2 topological charges. A combination of two of these should be topologically trivial, and so the monopole should be its own antiparticle.

To see more explicitly how this can occur, let us start with a singular string gauge configuration in which $S = S_0$ is spatially uniform at large distance while the gauge potential has the Dirac form

$$A_r = A_\theta = 0, \quad A_\phi = \frac{n}{2g}(\cos\theta - 1)\lambda_2, \tag{6.35}$$

with the Dirac string also along the negative z-axis. This has a magnetic charge

$$Q_M = \frac{4\pi}{g}\frac{n}{2}\lambda_2. \tag{6.36}$$

If we take λ_2 to be the single generator of the Cartan subalgebra, this corresponds to magnetic weight $k = n/2$.

This string singularity can be removed by a gauge transformation generated by the gauge function

$$U_n(\theta, \varphi) = e^{in\lambda_2\varphi/2} e^{i\lambda_3\theta/2} e^{-in\lambda_2\varphi/2}, \tag{6.37}$$

which is singular along the negative z-axis. [See the analogous transformation given by Eq. (5.74).] In particular, the asymptotic Higgs field at $r = \infty$ becomes

$$S_n = U_n S_0 U_n^T$$
$$= \begin{pmatrix} \cos\theta + i\sin\theta\cos(n\varphi) & -i\sin\theta\sin(n\varphi) & 0 \\ -i\sin\theta\sin(n\varphi) & \cos\theta - i\sin\theta\cos(n\varphi) & 0 \\ 0 & 0 & 1 \end{pmatrix}. \quad (6.38)$$

The covariant derivatives of S were rapidly vanishing at large distance before the gauge transformation, and so must also be afterwards. Hence, we have a finite energy configuration, and the usual arguments show that for $n = \pm 1$ there are actual static solutions with this asymptotic behavior. These would be expected to have a mass $\sim 4\pi\sigma/g$.

A Z_2 monopole should be its own antiparticle. We can see that this is so by noting that

$$S_{-1} = e^{i\pi\lambda_5} S_1 \left(e^{i\pi\lambda_5}\right)^T, \quad (6.39)$$

so that the $n = 1$ monopole and $n = -1$ antimonopole Higgs fields (and in fact their vector potentials also) are gauge equivalent. Furthermore, let us define a unitary matrix

$$V_n(\theta,\varphi) = e^{i\lambda_5\theta} e^{in\lambda_2\varphi/2} e^{-i\lambda_5\theta} e^{-i\lambda_3\theta/2} e^{-in\lambda_2\varphi/2}. \quad (6.40)$$

This is multiple-valued for odd n, because $V_n(\theta,\varphi) \neq V_n(\theta,\varphi + 2\pi)$. For even n, on the other hand, it is nonsingular and single-valued, with $V_n(0,\varphi) = I$, and has the property that

$$V_n S_n V_n^T = S_0. \quad (6.41)$$

Hence, any configuration with even n is equivalent by a smooth gauge transformation to one with $n = 0$. Because $\pi_2(\mathrm{SU}(3)) = 0$, this gauge transformation at $r = \infty$ can be smoothly deformed to the identity, thus giving a homotopy connecting S_n and S_0. Note that Eqs. (6.39) and (6.41) both required gauge transformations involving matrices that went outside the 2×2 block that contains the twisting of S_n.

6.2.3 A light doubly charged monopole

Let us now add an SU(3) triplet Higgs field ψ to the model of the previous example [104]. Let us also assume that ψ has a nonzero vacuum expectation value, and that in the vacuum the two Higgs fields are (up to a gauge transformation)

$$S_0^{ab} = \sigma\delta^{ab},$$
$$\psi_0^a = v\delta^{a3}. \quad (6.42)$$

The unbroken gauge group is now U(1), so the relevant homotopy group is

$$\pi_2[\mathrm{SU}(3)/\mathrm{U}(1)] = \pi_1[\mathrm{U}(1)] = Z. \quad (6.43)$$

The topological charge on the monopoles is now an ordinary additive integer charge. Choosing the asymptotic Higgs fields so that $S = S_n(\theta, \varphi)$ as in Eq. (6.38) and $\psi = \psi_0$ gives a configuration with topological charge n.

Ordinarily, we would expect to have a static particle-like solution only for $n = \pm 1$, with all larger values of n corresponding to multimonopole configurations. A new feature appears if $v \ll \sigma$, so that the symmetry breaking can be viewed as a two-step process

$$\mathrm{SU}(3) \underset{S}{\longrightarrow} \mathrm{SO}(3) \underset{\psi}{\longrightarrow} \mathrm{U}(1). \tag{6.44}$$

At the first stage we obtain a Z_2 monopole with mass $M_1 \sim \sigma/g^2$. This remains a solution, with only slight modifications, at the second stage. However, the transformation of Eq. (6.40), which turned the monopole into an antimonopole, is no longer possible, because λ_5 is not a generator of the unbroken group.

Now consider an $n = 2$ configuration. With just the breaking to SO(3), this would be topologically trivial and could be unwound by applying the V_2 of Eq. (6.40). With the breaking to U(1), it has topological charge 2 and is topologically nontrivial. We can still use V_2 to unwind S, but this would have the effect of twisting ψ. However, because the mass scale associated with ψ is much less than that associated with S, shifting the winding from S to ψ reduces the energy considerably, thus allowing us to obtain a charge 2 monopole with $M_2 \sim v/g^2 \ll M_1$.

All configurations with $n \geq 3$ presumably relax to multimonopole solutions.

6.2.4 Electroweak monopoles?

The previous examples in this section were illustrative, but not of direct phenomenological significance. Let us now consider the standard electroweak model, with SU(2)×U(1) broken to U(1) by a complex doublet Higgs field. Because the full gauge group is not simply connected, we cannot use Eq. (4.51) to determine $\pi_2(\mathcal{M}) = \pi_2(G/H)$. This is no problem, because we showed in Sec. 4.5 that the space of vacua for this theory is a three-sphere [see Eq. (4.26)]. Since $\pi_2(S^3) = 0$, there are no topologically stable monopoles in the Weinberg–Salam model.

6.3 Monopoles in grand unified theories

The idea of a grand unified theory (GUT) whose spontaneous breakdown leads to the observed gauge symmetries of the Standard Model remains an attractive possibility. Various implementations of this idea, often with several stages of symmetry breaking, have been proposed. By definition, all begin with a simple gauge group G that is ultimately broken down to the SU(3)×U(1) of QCD and electromagnetism. If we take G to be the covering group of the Lie algebra, Eq. (4.51) tells us that

$$\pi_2(G/H) = \pi_1(H) = \pi_1[\mathrm{SU}(3) \times \mathrm{U}(1)] = Z, \tag{6.45}$$

with the Z arising from the unbroken electromagnetic U(1). Thus any grand unified theory must contain topologically stable magnetic monopoles. Their mass will be of the order of $4\pi v/e$, where v is the vacuum expectation value of the Higgs field responsible for the symmetry breaking that first gives rise to non-trivial topology. Since this is typically a GUT scale of roughly 10^{16} GeV, these monopoles will be supermassive.

Let us consider two important examples.

6.3.1 SU(5) monopoles

The prototypical grand unified theory is based on an SU(5) gauge group, with a gauge coupling g, that is broken in two stages,

$$\mathrm{SU}(5) \underset{\Phi}{\longrightarrow} \mathrm{SU}(3) \times \mathrm{SU}(2) \times \mathrm{U}(1) \underset{\chi}{\longrightarrow} \mathrm{SU}(3) \times \mathrm{U}(1). \tag{6.46}$$

The first breaking is due to an adjoint representation Higgs field Φ that acquires a GUT-scale vacuum expectation value, while the second is due to a fundamental representation Higgs field χ that includes the Weinberg–Salam doublet with an electroweak scale vacuum expectation value.

Because the fermions fall into the $\bar{\mathbf{5}}$ and $\mathbf{10}$ representations, the initial group is indeed the covering group, SU(5), and not a factor group. By arguments similar to those for the SU(3) example of Sec. 6.2.1, the final unbroken subgroup is actually $[\mathrm{SU}(3) \times \mathrm{U}(1)]/Z_3$, with the factoring by Z_3 explaining the observed correlation between SU(3) triality and fractional electric charge.

To start, let us focus on the first breaking and set $\chi = 0$. The scalar field potential can then be chosen so that Φ has a vacuum expectation value of the form

$$\Phi_0 = \begin{pmatrix} v & 0 & 0 & 0 & 0 \\ 0 & v & 0 & 0 & 0 \\ 0 & 0 & v & 0 & 0 \\ 0 & 0 & 0 & -\frac{3}{2}v & 0 \\ 0 & 0 & 0 & 0 & -\frac{3}{2}v \end{pmatrix}. \tag{6.47}$$

The generators of the unbroken symmetry then take on a block diagonal form, with SU(3) generators $\lambda_a/2$ lying in the upper left 3×3 block, SU(2) generators $\tau_a/2$ in the lower right 2×2 block, and the U(1) generator being

$$T_{\mathrm{U}(1)} = \frac{1}{\sqrt{15}}\,\mathrm{diag}\left(1, 1, 1, -\frac{3}{2}, -\frac{3}{2}\right). \tag{6.48}$$

Twelve of the SU(5) gauge bosons acquire a mass $M_X = \sqrt{25/8}\,gv$ at this stage of symmetry breaking.

Because $\pi_1[\mathrm{SU}(3) \times \mathrm{SU}(2) \times \mathrm{U}(1)] = Z$, monopoles already appear at this first stage of symmetry breaking. Classical solutions can be obtained by following a strategy similar to that used for the SU(3) example of Sec. 6.2.1 [105]. We choose an ansatz such that the twisting of the Higgs field lies entirely within a 2×2 SU(2) subgroup corresponding to the intersections of columns and rows 1, 2, or 3 with columns and rows 4 or 5, and then add diagonal components to Φ to ensure the correct eigenvalues at spatial infinity. Choosing, for example, the subgroup defined by rows and columns 3 and 4 gives a Higgs field ansatz of the form

$$\Phi = \begin{pmatrix} a(r) & 0 & 0 & 0 \\ 0 & a(r) & 0 & 0 \\ 0 & 0 & h(r)\hat{\mathbf{r}} \cdot \boldsymbol{\tau} + b(r) I_2 & 0 \\ 0 & 0 & 0 & -2[a(r) + b(r)] \end{pmatrix}. \tag{6.49}$$

The nonzero components of A_i all lie within the chosen SU(2), and lead to an asymptotic magnetic field with magnetic charge along the positive z-axis

$$\begin{aligned} Q_M &= \frac{4\pi}{g} \operatorname{diag}\left(0, 0, \frac{1}{2}, -\frac{1}{2}, 0\right) \\ &= \frac{4\pi}{g} \left[\operatorname{diag}\left(\frac{1}{6}, \frac{1}{6}, \frac{1}{6}, -\frac{1}{4}, -\frac{1}{4}\right) + \operatorname{diag}\left(-\frac{1}{6}, -\frac{1}{6}, \frac{1}{3}, 0, 0\right)\right. \\ &\qquad \left. + \operatorname{diag}\left(0, 0, 0, -\frac{1}{4}, \frac{1}{4}\right)\right], \end{aligned} \tag{6.50}$$

where the second equality shows the decomposition into U(1), SU(3), and SU(2) components, respectively. The classical energy of this monopole is approximately $4\pi M_X/g^2$, and its core radius is of the order of M_X^{-1}.

The electroweak symmetry breaking is driven by the vacuum expectation value of χ, which is at a mass scale 14 or so orders of magnitude lower than the GUT scale. The effects of this symmetry breaking only become significant at length scales of order M_W^{-1}, so the corrections to the monopole core structure and mass are negligible. However, at distances much larger than M_W^{-1} the Coulomb magnetic field must lie within the unbroken gauge group. Thus, whatever the orientation of the SU(2) magnetic field near the core, at large distances Q_M must be rotated so that it is a linear combination of an SU(3) charge and the electromagnetic charge generator

$$Q_{\text{em}} = \operatorname{diag}\left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}, -1, 0\right). \tag{6.51}$$

The normalization of Q_{em} is such that $e = \sqrt{3/8}\, g$ (evaluated at the GUT scale). Taking this into account, we find that the minimally charged monopole, with a core profile given by Eq. (6.49), has an electromagnetic magnetic charge $2\pi/e$.

In Eq. (6.49) the choice of the SU(2) subgroup that contained the twisting of Φ was somewhat arbitrary. We could, for example, have used the first and fifth rows and columns, leading to a solution with asymptotic magnetic charge

$$Q'_M = \frac{4\pi}{g} \operatorname{diag}\left(\frac{1}{2}, 0, 0, 0, -\frac{1}{2}\right). \tag{6.52}$$

Now consider a configuration composed of two monopoles, one with charge Q_M and one with Q'_M. At distances that are small compared to the electroweak scale, the Coulomb interaction between the two is proportional to $\mathrm{tr}\,(Q_M Q'_M) = 0$. What has happened is that their long-range U(1) repulsion has been exactly canceled by the attractive SU(2) and SU(3) forces. The interaction between the two is then determined by the Yukawa forces mediated by the massive Higgs and gauge bosons. By proper choice of the masses of the bosons, one can arrange for the net effect of these to be attractive, giving a stable monopole with two units of U(1) magnetic charge. In fact, by this mechanism one can also obtain solutions with three, four, and six units of U(1) magnetic charge [106].

6.3.2 SO(10) monopoles

A second widely studied model is based on SO(10). One possible symmetry breaking pattern is commonly written as

$$\mathrm{SO}(10) \underset{\phi_1}{\rightarrow} \mathrm{SU}(4)\times\mathrm{SU}(2)\times\mathrm{SU}(2) \underset{\phi_2}{\rightarrow} \mathrm{SU}(3)\times\mathrm{SU}(2)\times\mathrm{U}(1) \underset{\phi_3}{\rightarrow} \mathrm{SU}(3)\times\mathrm{U}(1). \quad (6.53)$$

This is correct as far as the Lie algebras go, but to get the homotopy right we need to be careful about specifying the groups.

To start, we note that the Standard Model quarks and leptons of one generation, together with a right-handed neutrino, fill out a 16-component SO(10) spinor. With a spinor representation present, the original gauge group G is unambiguously the covering group, Spin(10). In the first stage of symmetry breaking, a Higgs field ϕ_1 transforming under the 54-dimensional traceless symmetric tensor representation obtains a vacuum expectation value of order v_1 that breaks this symmetry down to a subgroup H_1 that is locally $\mathrm{SO}(6) \times \mathrm{SO}(4) = \mathrm{SU}(4) \times \mathrm{SU}(2) \times \mathrm{SU}(2)$. Under this breaking the fundamental spinor of SO(10) decomposes into $(\mathbf{4},\, \mathbf{1},\, \mathbf{2}) + (\bar{\mathbf{4}},\, \mathbf{2},\, \mathbf{1})$. Here the $\mathbf{4}$ and $\bar{\mathbf{4}}$ are conjugate SO(6) spinors [or, equivalently, the fundamental and antifundamental representations of SU(4)], while the $\mathbf{2}$'s are SU(2) spinors. A rotation by 2π multiplies a spinor by -1, so simultaneous rotations by 2π in the SO(6) and SO(4) subgroups give two factors of -1 and thus act as the identity on the fermions. Hence,

$$H_1 = [\mathrm{SU}(4) \times \mathrm{SU}(2) \times \mathrm{SU}(2)]/Z_2, \quad (6.54)$$

which is not simply connected. Thus, we have

$$\pi_2(G/H_1) = \pi_1(H_1) = Z_2, \quad (6.55)$$

which means that there is a monopole carrying a Z_2 charge with a mass of order v_1/g. Like the Z_2 monopole of Sec. 6.2.2, this monopole is its own antiparticle.

At the next stage of symmetry breaking an SO(10) spinor Higgs field ϕ_2 gets a vacuum expectation value v_2 that breaks the symmetry down to

$$H_2 = [\mathrm{SU}(3) \times \mathrm{SU}(2) \times \mathrm{U}(1)]/Z_6. \quad (6.56)$$

The crucial point here is the appearance of the U(1) factor, so that

$$\pi_2(G/H_2) = \pi_1(H_2) = Z \tag{6.57}$$

and we now have monopoles with ordinary additive charges. If $v_2 \ll v_1$, the situation is essentially the same as in the example of Sec. 6.2.3. The Z_2 monopole that appeared at the first stage remains, but now with a unit Z charge. In addition, there is a new monopole, with two units of magnetic charge, associated with a nontrivial winding of ϕ_2 but a topologically trivial ϕ_1. This monopole has a core size of order $(gv_2)^{-1}$ and a mass $M_2 \sim v_2/g$ that can be several orders of magnitude smaller than that of the unit monopole [107].

Both monopoles survive the final stage of symmetry breaking with negligible corrections to their masses.

6.4 Chromodyons

We have seen that when a soliton is not invariant under a symmetry of the theory, the spectrum of fluctuations about the soliton includes a zero mode that requires the introduction of a collective coordinate z. Exciting this mode in a time-dependent fashion gives a nonzero conjugate momentum $p = I\dot{z}$, where I can be thought of as a generalized moment of inertia, and leads to a tower of excited states with energies $p^2/2I$ above the ground state. Thus, any soliton breaks translation invariance, and solitons with time-dependent position collective coordinates have nonzero linear momentum; here I is simply the soliton mass. If there is an unbroken U(1) internal symmetry that acts nontrivially on a soliton, then a time-dependent phase rotation gives the soliton a U(1) charge Q; for the case of the 't Hooft–Polyakov monopole, this yields the dyons studied in Sec. 5.5.

The GUT monopole solutions described in the previous section have Coulomb magnetic fields with nonzero components in the unbroken color SU(3). These are acted upon by the SU(3) generators, and so one might expect to obtain monopoles with SU(3) electric-type charges—chromodyons—from solutions that rotate in the internal SU(3) space.

Matters are not so simple. The first indication that there might be a problem is the slow falloff of the zero modes. A magnetic field falling as $1/r^2$ corresponds to a vector potential falling as $1/r$. Acting on such a field, an SU(3) transformation that was nontrivial at spatial infinity and did not commute with the magnetic charge would produce an infinitesimal transformation δA_j that also fell as $1/r$, making the resulting zero mode non-normalizable. One might naïvely view this as corresponding to an infinite moment of inertia, and conclude that the tower of chromodyon states collapses to a set of degenerate states. However, in a gauge theory one must proceed more carefully, making sure that the Gauss's law constraints are satisfied, as was done for the U(1) dyon in Sec. 5.5. This entails finding an A_0 that satisfies Gauss's law and that has a $1/r$ behavior consistent with the chromodyonic charge. It turns out that this cannot be done[6] [108].

The underlying explanation for these difficulties is that the long-range non-Abelian fields of the monopole create a topological obstruction that makes it impossible to define a set of generators for the unbroken gauge group that is nonsingular everywhere on the sphere at spatial infinity. Without these generators, one cannot define the global gauge rotations that would give rise to the chromodyons [109–114].

To see this more explicitly [109], consider the example of Sec. 6.2.1, where the unbroken gauge group is SU(2)×U(1). At spatial infinity the adjoint Higgs field is

$$\phi(\theta,\varphi) = b\begin{pmatrix} \frac{1}{2}+\frac{3}{2}\hat{\mathbf{r}}\cdot\boldsymbol{\tau} & 0 \\ 0 & -1 \end{pmatrix} = b\,U^{-1}\begin{pmatrix} 2 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix} U, \tag{6.58}$$

where U is a 3×3 matrix with the block diagonal form

$$U(\theta,\varphi) = \begin{pmatrix} \mathcal{U} & 0 \\ 0 & 1 \end{pmatrix}, \tag{6.59}$$

with $\mathcal{U}$ being the 2×2 SU(2) matrix given in Eq. (5.74).

At the north pole, $\theta = 0$, the unbroken SU(2) corresponds to the lower right 2×2 block, and a standard choice for the U(1) generator T_0 and the SU(2) generators T_k is

$$T_0 = \begin{pmatrix} 2 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}, \qquad T_1 = \frac{1}{2}\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix},$$
$$T_2 = \frac{1}{2}\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix}, \qquad T_3 = \frac{1}{2}\begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}. \tag{6.60}$$

By acting on these with U, we can obtain a set that commutes with $\phi(\theta,\varphi)$ and has the correct commutation relations. Two of these,

$$T_0(\theta,\varphi) = \begin{pmatrix} \frac{1}{2}+\frac{3}{2}\cos\theta & \frac{3}{2}\sin\theta e^{-i\varphi} & 0 \\ \frac{3}{2}\sin\theta e^{i\varphi} & \frac{1}{2}-\frac{3}{2}\cos\theta & 0 \\ 0 & 0 & -1 \end{pmatrix} \tag{6.61}$$

and

$$T_3(\theta,\varphi) = \frac{1}{2}\begin{pmatrix} \frac{1}{2}-\frac{1}{2}\cos\theta & -\frac{1}{2}\sin\theta e^{-i\varphi} & 0 \\ -\frac{1}{2}\sin\theta e^{i\varphi} & \frac{1}{2}+\frac{1}{2}\cos\theta & 0 \\ 0 & 0 & -1 \end{pmatrix}, \tag{6.62}$$

[6] One can also work in $A_0 = 0$ gauge, in which case Gauss's law must be imposed as an additional constraint. In this approach, the construction of the chromodyon only goes through if there is a zero mode corresponding to a gauge transformation with a gauge function Λ that is nonzero at spatial infinity and satisfies $D_k D_k \Lambda + g^2[\phi,[\phi,\Lambda]] = 0$. This equation has no solutions if Λ does not commute with the magnetic charge [108].

are well defined everywhere, but the other two,

$$T_1(\theta,\varphi) = \frac{1}{2}\begin{pmatrix} 0 & 0 & -\sin(\theta/2)e^{-i\varphi} \\ 0 & 0 & \cos(\theta/2) \\ -\sin(\theta/2)e^{i\varphi} & \cos(\theta/2) & 0 \end{pmatrix} \tag{6.63}$$

and

$$T_2(\theta,\varphi) = \frac{1}{2}\begin{pmatrix} 0 & 0 & i\sin(\theta/2)e^{-i\varphi} \\ 0 & 0 & -i\cos(\theta/2) \\ -i\sin(\theta/2)e^{i\varphi} & i\cos(\theta/2) & 0 \end{pmatrix}, \tag{6.64}$$

are singular at $\theta = \pi$, the south pole.[7]

In this example, the generators that commuted with the magnetic charge (T_0 and T_3) could be defined globally. It was only the two that did not commute with Q_M that failed to be well defined. This can be understood as follows. One way to define a global gauge rotation is to choose a Lie algebra element Ω at one point P on a sphere at large r and then use parallel transport to obtain Ω at any other point P' on the sphere. This only works if the result of the parallel transport is independent of the path from P to P'. This in turn requires that the surface integral of $[\mathbf{B}, \Omega]$ over the area between any two such paths vanishes. In the limit of infinite radius only the $1/r^2$ part of $\mathbf{B}$, i.e., the magnetic charge, contributes to this integral. Hence, only the generators of the subgroup that commutes with the magnetic charge are well defined.

This suggests a loophole that might allow chromodyons to exist. Consider a monopole with a purely Abelian magnetic charge in a theory with an unbroken non-Abelian subgroup. The Abelian magnetic charge would not be an obstacle to defining global color transformations. Although these would have no effect on the asymptotic magnetic field, there might well be fields nearer the core that were not invariant under color transformation. The corresponding zero modes would be normalizable, and would provide the basis for constructing a chromodyonic solution.

A monopole of just this sort can be constructed in a gauge theory with SO(5) broken to SU(2)×U(1). However, numerical study of the classical evolution of the chromodyon solution reveals that the rate of rotation in color space decreases with time, corresponding to a loss of color charge [115]. This is apparently due to radiation of energy and color charge via the massless gauge boson field, with all indications being that the radiation continues until the charge has been completely lost. Thus, even when there are no topological obstacles to their existence, chromodyons appear to be dynamically unstable.

[7] An alternative approach is to use two patches, with sets of generators that are nonsingular in the upper and lower hemispheres, respectively. Consistency then requires that the two sets be related by a nonsingular gauge transformation in the overlap region. Again, this turns out to be impossible.

6.5 The Callan–Rubakov effect

We saw in Sec. 5.7 that the scattering of massless fermions off an 't Hooft–Polyakov monopole has some unusual aspects. In the $J = 0$ sector one finds only half of the expected incoming states and half of the expected outgoing states; this is ultimately a consequence of the extra charge-monopole contribution to the angular momentum. The matching of incoming to outgoing states requires that either the fermion chirality or the fermion electric charge must change. The analyses of Rubakov [91] and of Callan [92, 93] showed that the former is the case, and that no electric charge is deposited either on the monopole or on the surrounding fermion condensate.

An analogous effect occurs with monopoles in larger gauge groups, in particular those that arise in grand unified theories [91, 93, 116–118]. The new feature here is that the incoming and outgoing states that are paired have different baryon and lepton numbers. To be specific, let us consider a monopole in the SU(5) theory that has a magnetic charge given by Eq. (6.50). This is essentially an embedding of the SU(2) monopole in the subgroup corresponding to the third and fourth rows and columns.

Each family of fermion fields in the SU(5) model can be assembled into two multiplets of Weyl fields. The first family, which can be treated as approximately massless, contains an antifundamental $\bar{\mathbf{5}}$ representation,

$$\psi = (d_1^c, d_2^c, d_3^c, e^-, \nu)_L^t, \tag{6.65}$$

and a symmetric tensor **10** representation,

$$\chi = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & u_3^c & -u_2^c & -u_1 & -d_1 \\ -u_3^c & 0 & u_1^c & -u_2 & -d_2 \\ u_2^c & -u_1^c & 0 & -u_3 & -d_3 \\ u_1 & u_2 & u_3 & 0 & -e^+ \\ d_1 & d_2 & d_3 & e^+ & 0 \end{pmatrix}_L . \tag{6.66}$$

(Here subscripts are SU(3) color indices and a superscript c denotes charge conjugation and the d's should be understood as the CKM-rotated mixtures.) When these SU(5) multiplets are decomposed into representations of the SU(2) defined by the monopole embedding, we find four doublets,

$$\begin{pmatrix} e^+ \\ d_3 \end{pmatrix}_L , \quad \begin{pmatrix} d_3^c \\ e^- \end{pmatrix}_L , \quad \begin{pmatrix} u_1 \\ u_2^c \end{pmatrix}_L , \quad \begin{pmatrix} u_2 \\ -u_1^c \end{pmatrix}_L . \tag{6.67}$$

In the $J = 0$ sector, the upper components of the doublets only appear as incoming waves, and the lower ones as outgoing waves.

As with the SU(2) theory considered in Sec. 5.7, the analysis of the system is most easily done by bosonizing the theory. In the previous case there were two Dirac, or four Weyl, fermion fields, leading to two scalar fields. Now, with eight Weyl fermion fields, we have four scalar fields, but the analysis is otherwise similar.

The monopole is surrounded by a fermion condensate formed from the doublets listed above. The energy of the monopole-fermion system is minimized by requiring vanishing charge under all components of the unbroken gauge group. A set of particles from the relevant doublets that meets this criterion is the electrically neutral, color singlet combination $e^- u_1 u_2 d_3$ (or the corresponding set of antiparticles). The analysis of the bosonized theory shows that the ground state of the system is a superposition of states with arbitrary numbers of this set of fermions. This allows scattering processes that effectively add or subtract particles in this combination. An example is $u_1 + \text{Monopole} \rightarrow \bar{u}_2 + \bar{d}_3 + e^+ + \text{Monopole}$, a process that violates the conservation of both baryon number B and lepton number L (but not of $B - L$). With the initial u being a valence quark in a proton, this process could lead to the monopole-catalyzed decay of a proton to a positron plus a $\pi^+\pi^-$ pair or to a positron plus a photon.

The possibility of a baryon number violating process is not surprising, since it is well known that the SU(5) theory allows proton decay. The striking feature is that there is no suppression by factors of the masses of the superheavy gauge bosons, or of the monopole core size. Instead, the cross-section is essentially geometric, and so is expected to be of typical strong interaction size.

This analysis in the SU(5) theory depended in a detailed manner on the way in which the light fermions transformed under the SU(2) defined by the magnetic charge. With a different embedding of the magnetic charge in the GUT gauge group, this catalysis of baryon number violation might not occur. This is confirmed by a detailed examination of a number of theories. The monopoles in the SO(10) theory considered in Sec. 6.3.2 are important examples [119]. The heavier, singly charged monopoles that arise at the first stage of symmetry breaking catalyze baryon number violation, but the lighter, doubly charged ones do not.