

Anyone who has inspected Maxwell's equations even briefly has probably speculated about the existence of magnetic monopoles. There is no experimental evidence for magnetic monopoles, but the equations would be far more symmetric if they existed. It was Dirac who first considered carefully the implications of monopoles, and he came to a striking conclusion: the existence of monopoles would require that *electric charge* be quantized in terms of a fundamental unit. The problem of describing a monopole lies in writing $\vec{B} = \vec{\nabla} \times \vec{A}$. We could simply give up this identification, but Dirac recognized that \vec{A} is essential in formulating quantum mechanics. To resolve the problem we can follow Wu and Yang and maintain that $\vec{B} = \vec{\nabla} \times \vec{A}$ but not require that the vector potential be single valued. Suppose that we have a monopole located at the origin. In the northern hemisphere we can take

$$\vec{A}_N = \frac{g}{4\pi r} \frac{1 - \cos\theta}{\sin\theta} \hat{e}_\phi, \quad (7.1)$$

while in the southern hemisphere we take

$$\vec{A}_S = -\frac{g}{4\pi r} \frac{1 + \cos\theta}{\sin\theta} \hat{e}_\phi. \quad (7.2)$$

By looking up the formulae for the curl operator in spherical coordinates, you can check that, in both hemispheres,

$$\vec{B} = \frac{g}{4\pi r^2} \hat{r}, \quad (7.3)$$

so indeed this does describe a magnetic monopole.

Each of expressions (7.1) and (7.2) is singular along a half-line: A_N is singular along $\theta = \pi$; A_S is singular along $\theta = 0$. These string-like singularities are known as *Dirac strings*. They are suitable vector potentials to describe long thin solenoids which start at the origin and go to infinity along the negative or positive z axis. With discontinuous \vec{A} , though, we need to ask whether quantum mechanics is consistent. Consider the equator ($\theta = \pi/2$). We have

$$\vec{A}_N - \vec{A}_S = \frac{g}{2\pi r} \hat{e}_\phi = -\vec{\nabla}\chi, \quad \chi = -\frac{g}{2\pi}\phi, \quad (7.4)$$

where ϕ is the azimuthal angle and χ is a general function. So, the difference has the form of a gauge transformation. But to be a gauge transformation it must act sensibly on particles of definite charge. In particular, it must be single valued. As such a particle circumnavigates the sphere, its wave function acquires a phase

$$\exp\left(ie \int d\vec{x} \cdot \vec{A}\right). \quad (7.5)$$

Potentially, this phase is different for \vec{A}_N and \vec{A}_S , in which case the string would be a detectable, real, object. But the phases are the same if

$$\exp\left(i\frac{eg}{2\pi}\oint d\vec{x}\cdot\vec{\nabla}\phi\right)=1 \quad \text{or} \quad eg=2\pi n. \quad (7.6)$$

This is known as the *Dirac quantization condition*. Dirac argued that, since e can be the charge of any charged particle, if there is even one monopole somewhere in the universe, this result shows that charge must be quantized.

In pure electrodynamics the status of magnetic monopoles is obscure; the \vec{B} field is singular and the energy is infinite. In non-Abelian gauge theories with scalar fields (Higgs fields), however, monopoles often arise as finite-energy non-dissipative solutions of the classical equations. Such solutions cannot arise in linear theories like electrodynamics; all configurations in such a theory spread with time. Non-dissipative solutions can only arise in non-linear theories, and even then, such solutions – known as solitons – can only arise in special circumstances.

The simplest theory which exhibits monopole solutions is $SU(2)$ (more precisely $O(3)$) Yang–Mills theory with a single Higgs particle in the adjoint representation. But, before considering this case, which is somewhat complicated, it is helpful to consider solitons in lower-dimensional situations.

7.1 Solitons in 1 + 1 dimensions

Consider a quantum field theory in 1 + 1 dimensions, with

$$\mathcal{L}=\frac{1}{2}(\partial_\mu\phi)^2-V(\phi). \quad (7.7)$$

Here

$$V(\phi)=-\frac{1}{2}m^2\phi^2+\frac{\lambda}{4}\phi^4. \quad (7.8)$$

This potential, which is symmetric under $\phi\rightarrow-\phi$, has two degenerate minima, $\pm\phi_0$. Normally, we would choose as our vacuum a state localized about one or the other minimum. These correspond to trivial solutions of the equations of motion. We can consider a more interesting configuration, a localized finite-energy solution known as a *soliton*, for which

$$\phi(x\rightarrow\pm\infty)\rightarrow\pm\phi_0. \quad (7.9)$$

Such a solution interpolates between the two different vacua. We can construct this solution much as one solves analogous problems in classical mechanics, by quadrature. Finding the solution for this particular model, known as a *kink*, is left for the exercises; the result is

$$\phi_{\text{kink}}=\phi_0\tanh[(x-x_0)m]. \quad (7.10)$$

This solution is shown in Fig. 7.1. The kink has finite energy. As we have indicated, there is a continuous infinity of solutions, corresponding to the fact that this kink can be

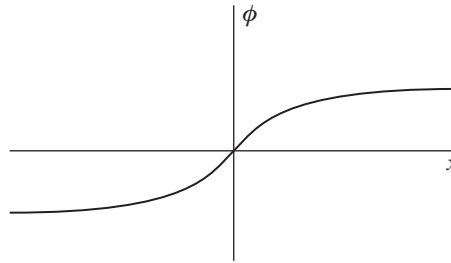


Fig. 7.1 Kink solution of the two-dimensional field theory.

located anywhere; this is a consequence of the underlying translational invariance. We can use this to understand in what sense the kink is a particle. Consider configurations which are not quite solutions of the equations of motion, in which x_0 is allowed to be a slowly varying function $x_0(t)$ of t . We can write down the action for these configurations:

$$S_{\text{kink}} = \int dt \int dx \left[\frac{1}{2} (\partial_\mu \phi_{\text{kink}})^2 - V(\phi_{\text{kink}}) \right]. \quad (7.11)$$

Only the $\dot{\phi}$ term contributes. The result is

$$S_{\text{kink}} = \int dt \frac{M}{2} \dot{x}_0^2. \quad (7.12)$$

Here M is precisely the energy of the kink. So, the kink truly acts as a particle. The quantity x_0 is called a *collective coordinate*. We will see that such collective coordinates arise for each symmetry broken by the soliton. These are similar to the collective coordinates we encountered in the Euclidean problem of the instanton.

7.2 Solitons in 2 + 1 dimensions: strings or vortices

As we go up in dimension, the possible solitons become more interesting. Consider a $U(1)$ gauge theory in 2 + 1 dimensions, with a single charged scalar field ϕ . This model is often called the Abelian Higgs model. The Lagrangian is

$$\mathcal{L} = |D_\mu \phi|^2 - V(|\phi|). \quad (7.13)$$

We assume that the potential is such that

$$\langle \phi \rangle = v. \quad (7.14)$$

Now we have a possibility that we have not considered before. Working in plane polar coordinates r, θ , if we consider only the potential then we can imagine obtaining finite-energy configurations for which, at large r ,

$$\phi \rightarrow e^{in\theta} v. \quad (7.15)$$

Because the potential tends to its minimum at infinity, such a configuration has finite potential energy. However, the kinetic energy diverges, since $\partial_\mu\phi$ includes $(1/r)\partial_\theta\phi$. We can try to cancel this with a non-vanishing gauge field. At ∞ , the scalar field is a gauge transformation of the constant configuration, so to achieve finite energy we want to gauge-transform the gauge field as well,

$$A_\theta \rightarrow n; \quad (7.16)$$

consequently, at ∞ , $D_\mu\phi \rightarrow 1/r^2$ or a higher negative power of r . It is not hard to construct such solutions numerically. As for the kinks, these configurations again have collective coordinates, corresponding to the two translational degrees of freedom and a rotational (or charge) degree of freedom.

We can take these configurations as configurations in a $(3+1)$ -dimensional theory, which are constant with respect to z . Viewed in this way, these are vortices, or strings. One has collective coordinates corresponding to transverse motions of the string, $x_0(z, t), y_0(z, t)$. These string configurations could be quite important in cosmology. Such a broken $U(1)$ theory could lead to the appearance of long strings, which could carry enormous amounts of energy. For a time, these were considered a possible origin of inhomogeneities leading to the formation of galaxies, but the data now disfavors this possibility.

7.3 Magnetic monopoles

Dirac's argument shows that, in the presence of a monopole, electric charges are all multiples of a basic charge. This means that the $U(1)$ symmetry is effectively compact. So, a natural place to look for monopoles is in gauge theories where $U(1)$ is a subgroup of a simple group. The $SU(5)$ grand unified theory is an example, in which electric charge is quantized.

We start, though, with the simplest example of this sort, an $SU(2)$ gauge theory with Higgs fields in the adjoint representation, ϕ^a . Such a theory was first considered by Georgi and Glashow as a model for weak interactions without neutral currents and is known as the Georgi–Glashow model. An expectation value for ϕ , $\phi^3 = v$ or

$$\phi = \frac{v}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad (7.17)$$

leaves an unbroken $U(1)$. The spectrum includes massive charged gauge bosons W^\pm and a massless gauge boson, which we will call the photon, γ . By analogy to the string or vortex solutions, we require finite energy at ∞ :

$$\phi \rightarrow gv. \quad (7.18)$$

In the $(2+1)$ -dimensional case we could think of the gauge transformation as a mapping from the space at infinity (topologically a circle) onto the gauge group (also a circle).

In three dimensions we want gauge transformations which map the two-sphere S_2 into the gauge group $SU(2)$. For example, we can take

$$g(\vec{x}) = i \frac{\hat{x}^i \sigma^i}{2}. \quad (7.19)$$

This suggests the following ansatz (guess) for a solution:

$$\phi^a = \hat{r}^a h(r), \quad A_i^a = -\epsilon_{ij}^a \frac{\hat{r}^j}{r} j(r). \quad (7.20)$$

This solution is very symmetric: it is invariant under a combined rotation in spin and isospin (rather similar to the sorts of symmetry of the instanton solution). Note that h and j satisfy coupled non-linear equations, which in general must be solved numerically. We can see from the form of the action that the mass is of order $1/g^2$. In the next section we show that an analytic solution can be obtained in a particular limit.

We can write down an elegant expression for the number of times $g(x)$ maps the sphere into the gauge group:

$$N = \frac{1}{4\pi} \int dS^i \epsilon_{ijk} \text{Tr}(g \partial_j g \partial_k g). \quad (7.21)$$

In terms of the field, ϕ ,

$$N = \frac{1}{8\pi v^3} \int dS^i \epsilon^{ijk} \epsilon^{abc} \partial_i \phi^a \partial_j \phi^b \partial_k \phi^c. \quad (7.22)$$

Finally, we need a definition of the magnetic charge. A natural choice is

$$\int d^3x \frac{1}{v} \partial_i (\phi^a B_i^a) = \frac{4\pi N}{e}.$$

Putting these statements together, we see that this solution, the 't Hooft–Polyakov *monopole*, has one Dirac unit of magnetic charge.

7.4 The BPS limit

Prasad and Sommerfield wrote down an exact monopole solution in the limit $V = 0$. This limit seemed originally rather artificial, but we will see later that some supersymmetric field theories automatically have a vanishing potential for a subset of fields. What simplifies the analysis in this limit is that the equations for the monopole, which are ordinarily second-order non-linear differential equations, become first-order equations. We will shortly see how to understand this in terms of supersymmetry. First, though, we will derive this result by looking directly at the potentials for the gauge and scalar fields. We start by deriving a bound, the Bogomol'nyi–Prasad–Sommerfield (BPS) bound, on the mass of a static field configuration. Again we call the gauge coupling e , to avoid confusion with the magnetic charge g :

$$M_m = \int d^3x \frac{1}{2} \left[\frac{1}{e^2} \vec{B}^a \cdot \vec{B}^a + (\vec{D}\Phi)^a \cdot (\vec{D}\Phi)^a \right]. \quad (7.23)$$

We can compare this with

$$\begin{aligned} A_{\pm} &= \int d^3x \left[\frac{1}{e} \vec{B}^a \pm (\vec{D}\Phi)^a \right]^2 \\ &= \frac{1}{2} \int d^3x \left[\frac{1}{e^2} \vec{B}^{a2} + (\vec{D}\Phi)^{a2} \right] \pm \int d^3x \frac{1}{e} \vec{B}^a (\vec{D}\Phi)^a. \end{aligned} \quad (7.24)$$

We can integrate the last term by parts. You can check that this works for both parts of the covariant derivative, i.e. this term becomes:

$$\frac{1}{e} \int d^3x (\vec{D} \cdot \vec{B})^a \Phi^a - \frac{1}{e} \int d^2a \Phi^a \hat{n} \cdot \vec{B}^a. \quad (7.25)$$

The first term vanishes by the Bianchi identity (the Yang–Mills generalization of the equation $\vec{\nabla} \cdot \vec{B} = 0$). The second term is v times what we have defined to be the monopole charge, g . So we have

$$A_{\pm} = \int d^3x \left[\frac{1}{e^2} \vec{B}^a \pm (\vec{D}\Phi)^a \right]^2 = M_m \pm \frac{vg}{e}. \quad (7.26)$$

The left-hand side of this equation is clearly greater than zero, so we have shown that

$$M_m \geq \left| \frac{vg}{e} \right|. \quad (7.27)$$

This bound, known as the Bogomol'nyi or BPS bound, is saturated when

$$\vec{B}^a = \pm \frac{1}{e} (\vec{D}\Phi)^a. \quad (7.28)$$

Note that while so far in this chapter we have worked in terms of $SU(2)$, this result generalizes to any gauge group with Higgs in the adjoint representation. But let us still focus on $SU(2)$ and try to find a solution which satisfies the Bogomol'nyi bound. As in the case of the 't Hooft–Polyakov monopole, it is convenient to write:

$$\Phi^a = \frac{\hat{r}^a}{er} H(evr), \quad A_i^a = -\epsilon_{ij}^a \frac{\hat{r}^j}{er} [1 - K(evr)]. \quad (7.29)$$

Here we are using a dimensionless variable, $u = evr$, in terms of which the Hamiltonian scales simply. We are looking for solutions for which $H \rightarrow 0$ and $K \rightarrow 1$ as $r \rightarrow 0$. Otherwise, the solutions would be singular at the origin. At ∞ , we want the configuration to look like a gauge transformation of the vacuum solution, so we require

$$K \rightarrow 0, \quad H \rightarrow evr \quad \text{as } r \rightarrow \infty. \quad (7.30)$$

We will leave the details to the exercises, but it is straightforward to show that these equations are solved by

$$H(y) = y \coth y - 1, \quad K(y) = \frac{y}{\sinh y}. \quad (7.31)$$

The monopole mass is

$$M_m = \frac{vg}{e} = \frac{2\pi v}{e^2}, \quad (7.32)$$

as predicted by the BPS formula.

7.5 Collective coordinates for the monopole solution

In lower-dimensional examples we witnessed the emergence of collective coordinates, which described the translations and other collective motions of the solitons. In the case of the monopole we have similar collective coordinates. Again, the solutions violate translational invariance. As a result we can generate new solutions on replacing \vec{x} by $\vec{x} - \vec{x}_0$. Now viewing x_0 as a slowly varying function of t , we obtain as before the action of a non-relativistic particle of mass M_m . The particle is non-relativistic in the weak coupling limit because its mass scales as $1/g^2$ and it becomes infinitely heavy as $g \rightarrow 0$.

There is another collective coordinate of the monopole solution, which has quite remarkable properties. In the monopole solution, charged fields are excited. So the monopole solution is not invariant under the $U(1)$ gauge transformations of electrodynamics. One might think that this is not important; after all, we have stressed that gauge transformations are not real symmetries but instead represent a redundancy of the description of a system. But we need to be more precise. In interpreting Yang–Mills instantons, we worked in the $A_0 = 0$ gauge. In this gauge the important gauge transformations are time-independent gauge transformations, and these fall into two classes: large gauge transformations and small gauge transformations. The small gauge transformations are those which fall rapidly to zero at infinity, and physical states must be invariant under these. For large gauge transformations this is not the case, and they can correspond to physically distinct configurations.

For the monopole configurations, the interesting gauge transformations are those which tend, at infinity, to a transformation in the unbroken $U(1)$ group. For large r , this direction is determined by the direction of the Higgs fields. We must be careful how we fix the gauge; again we will work in the $A_0 = 0$ gauge. For our collective motion, we want to study gauge transformations in this direction which vary slowly in time. It is important, however, that we remain in the $A_0 = 0$ gauge, so the transformations that we will study are not quite gauge transformations. Specifically, we consider

$$\delta A_i = \frac{D_i[\chi(t)\Phi]}{v}, \quad (7.33)$$

where $\chi(t)$ is a general time-dependent function, but we transform A_0 by

$$\delta A_0 = \frac{D_0(\chi\Phi)}{v} - \frac{\dot{\chi}\Phi}{v} \quad (7.34)$$

and, in order that the Gauss law constraint be satisfied, we require that $\delta\Phi = 0$. The action for χ has the form:

$$S = \frac{C}{2e^2} \dot{\chi}^2. \quad (7.35)$$

Note that χ is bounded between 0 and 2π , i.e. it is an angular variable. Its conjugate variable is like an angular momentum; calling this Q we have

$$Q = p_\chi = \frac{C}{e^2} \dot{\chi}, \quad H = \frac{1}{2C} e^2 Q^2. \quad (7.36)$$

In the case of a BPS monopole, the constant C is $e^2 M_m / (2v^2)$. So, each monopole has a tower of charged excitations, with energies of order e^2 above the ground state. These excitations of the monopole about the ground state are known as *dyons*. The mass formula for these states has the form, in the case of a BPS monopole:

$$M = vg + \frac{vQ^2}{g}. \quad (7.37)$$

We will understand this better when we embed this structure in a supersymmetric field theory.

7.6 The Witten effect: the electric charge in the presence of θ

We have argued that in a $U(1)$ gauge theory it is difficult to see the effects of θ . But, in the presence of a monopole, a θ term (see Section 5.3) has a dramatic effect, pointed out by Witten: the monopole acquires an electric charge that is proportional to θ .

We can see this first in an heuristic way. We will work in a gauge with non-zero A_0 and take all fields as static. Then

$$\vec{E} = -\vec{\nabla}A_0, \quad \vec{B} = \frac{g}{4\pi} \frac{\vec{r}}{r^2} + \vec{\nabla} \times \vec{A}. \quad (7.38)$$

For such a configuration, the θ term,

$$\mathcal{L}_\theta = \frac{\theta e^2}{8\pi^2} \vec{E} \cdot \vec{B}, \quad (7.39)$$

takes the form

$$\mathcal{L}_\theta = -\frac{\theta e^2 g}{32\pi^2} \int d^3r A_0 \vec{\nabla} \cdot \frac{\hat{r}}{r^2} = -\frac{\theta e^2 g}{8\pi^2} \int d^3r A_0 \delta(\vec{r}). \quad (7.40)$$

We started with a magnetic monopole at the origin, but we now also have an electric charge at the origin, $\theta e^2 g / (8\pi^2)$.

One might worry that in this analysis one is dealing with a singular field configuration, but in the non-Abelian case the configuration is non-singular. We can give a more precise argument. Let us go back to the $A_0 = 0$ gauge. In this gauge we can sensibly write down the canonical Hamiltonian. In the absence of θ , the conjugate momentum to \vec{A} is \vec{E} . But, in the presence of θ , there is an additional contribution,

$$\vec{\Pi} = -\frac{d\vec{A}}{dt} + \frac{\theta e^2}{8\pi^2} \vec{B}. \quad (7.41)$$

Now we will think about the invariance of states under small gauge transformations. For $\theta = 0$ we saw that the small gauge transformations, with gauge parameter ω , are generated by

$$Q_\omega = \int d^3x \vec{\nabla} \omega \cdot \vec{E}. \quad (7.42)$$

An interesting set of large gauge transformations is those with $\omega^a = \lambda \Phi^a / v$. For these, if we integrate by parts then we obtain a term which vanishes by Gauss's law (Gauss's law is enforced by the invariance under small gauge transformations), and a surface term. This surface term gives the total $U(1)$ charge times λ . We can think of this another way. For the low-lying excitations, multiplication by $e^{iQ\omega}$ corresponds to shifting the dynamical variable χ by a constant, λ . In general the wave functions for χ have the form $e^{iq\chi}$, where q is quantized. So the states pick up a phase $e^{iq\lambda}$. This is just the transformation of a state of charge q under a global gauge transformation with phase λ .

In the presence of θ , however, the operator which implements time-independent gauge transformations is modified. The field \vec{E} is replaced by the canonical momentum above. Now acting on states, the extra term gives a factor $g\theta/(2\pi)$ in the exponent. Even states with $q = 0$ pick up a phase, so there is an additional contribution to the charge,

$$Q = n_e e - \frac{\theta n_m}{2\pi}. \quad (7.43)$$

7.7 Electric–magnetic duality

As mentioned earlier, Maxwell's equations suggest a possible duality between electricity and magnetism. If there were magnetic charges, these equations would take the form

$$\vec{\nabla} \cdot \vec{E} = \rho_e, \quad \vec{\nabla} \cdot \vec{B} = \rho_m, \quad (7.44)$$

$$\vec{\nabla} \times \vec{E} = -\frac{\partial \vec{B}}{\partial t} + \vec{j}_m, \quad \vec{\nabla} \times \vec{B} = \frac{\partial \vec{E}}{\partial t} + \vec{j}_e. \quad (7.45)$$

These equations retain their form if we replace \vec{E} by $-\vec{B}$ and \vec{B} by \vec{E} and also let $\rho_e \rightarrow \rho_m$ and $\rho_m \rightarrow -\rho_e$ (and similarly for the electric and magnetic currents).

Now that we have a framework for discussing magnetic charges, it is natural to ask whether some theories of electrodynamics really obey such a symmetry. In general, however, this is a difficult problem. We have just learned that electric and magnetic charges, when they both exist, obey a reciprocal relation, $g \propto 1/e$. From the point of view of quantum field theory, this means that exchanging electric and magnetic charges also means replacing the fundamental coupling by its inverse. In other words, if there is such a duality symmetry, *it relates a strongly coupled theory to a weakly coupled theory*. We do not know a great deal about strongly coupled gauge theories, so investigating the possibility of such a duality is a difficult problem. That such a symmetry might exist in theories of the type we have been discussing is not entirely crazy. For example the monopole masses behave, at weak coupling, like $1/g^2$. So as the coupling becomes strong, these particles become light, even as the charged states become heavy. They have complicated quantum numbers (some monopole states are fermionic, for example).

Remarkably, there is a circumstance where such dualities can be studied, namely theories with more than one supersymmetry (in four dimensions): $N=4$ supersymmetric Yang–Mills theory turns out to exhibit an electric–magnetic duality. These theories will be

discussed in Chapter 15. Crucial to verifying this duality will be a deeper understanding of the Bogomol'nyi–Prasad–Sommerfield (BPS) condition, which will allow us to establish exact formulas for the masses of certain particles that are valid for all values of the coupling. These formulas will exhibit precisely the expected duality between electricity and magnetism.

Suggested reading

There are many excellent reviews and texts on monopoles. These include Coleman (1981) and Harvey (1996), and this chapter borrows ideas from both. You can find an introduction to the subject in Chapter 6 of Jackson's electrodynamics text (1999).

Exercises

- (1) Verify that Eqs. (7.1) and (7.2) are those for infinitely long, thin, solenoids ending at the origin.
- (2) Find the kink solution of the (1 + 1)-dimensional model. Show that the collective coordinate action is

$$S = \int dt \frac{1}{2} M_{\text{kink}} \dot{x}_0^2.$$

- (3) Verify that Eqs. (7.31) solve the BPS equations.