REVIEW

Magnetic monopoles in field theory and cosmology

BY ARTTU RAJANTIE*

Department of Physics, Imperial College London, Prince Consort Road, London SW7 2AZ, UK

The existence of magnetic monopoles is predicted by many theories of particle physics beyond the standard model. However, in spite of extensive searches, there is no experimental or observational sign of them. I review the role of magnetic monopoles in quantum field theory and discuss their implications for particle physics and cosmology. I also highlight their differences and similarities with monopoles found in frustrated magnetic systems.

Keywords: magnetic monopoles; cosmology; particle physics; quantum field theory; grand unified theory

1. Introduction

Experience tells us that magnetic north and south poles cannot be separated into magnetic monopoles, i.e. isolated magnetic charges. This was discussed in the thirteenth century by Peregrinus [1]. On the other hand, there are strong theoretical reasons to believe that they should exist.

Curie [2] speculated on the possibility of free magnetic charges in 1894, and magnetic monopoles were occasionally used as mathematical examples in electrodynamics, but physicists only started to consider them seriously after Dirac had shown in 1931 that their existence would explain the observed quantization of electric charge [3]. In 1974, ’t Hooft and Polyakov showed that magnetic monopoles exist as nonlinear objects in many theories of particle physics [4,5]. In fact, they are an inevitable prediction of grand unification of elementary particle interactions [6], and the same applies generally also to more modern ‘theories of everything’ such as superstring theory [7].

Finding a magnetic monopole particle would be an incredible breakthrough in high-energy physics. They derive their properties from processes that take place at extremely high energies, and yet, because they are stable particles and they interact through the electromagnetic field, they would be relatively easy to study experimentally. Therefore, they could open up a new window to particle physics at

*a.rajantie@imperial.ac.uk

One contribution of 8 to a Theo Murphy Meeting Issue ‘Emergent magnetic monopoles in frustrated magnetic systems’.
A. Rajantie

energies that are out of reach of any foreseeable accelerator experiments. There have been numerous attempts [8,9] to detect them in different ways, but with no success.

Even though magnetic monopoles have not been found, they have played an important role in theoretical high-energy physics. They have given powerful theoretical tools to investigate properties of strongly coupled non-Abelian gauge field theories, such as quantum chromodynamics (QCD) and especially its supersymmetric variants [10]. Also, an attempt to explain the apparent absence of monopoles led to the formulation of the theory of cosmological inflation [11], which has later been confirmed by astronomical observations [12].

Recently, it has been discovered that spin ices, frustrated magnetic systems, have effective quasi-particle excitations with magnetic charges [13]. These effective monopoles appear to have very similar properties to actual fundamental magnetic monopoles [14,15], and this may also provide a way to learn more about their physics.

The aim of this paper is to give an brief overview of magnetic monopoles in particle physics. For a more pedagogical introduction, see Rejantie [16], and for a more detailed review, see Milton [17]. A comprehensive and up-to-date magnetic monopole bibliography is also available [18]. This paper is organized as follows: §2 shows how magnetic monopoles arise from different field theories. Section 3 reviews experimental attempts to search for them. Section 4 discusses their role in cosmology, and §5 their significance for understanding gauge field theories in theoretical high-energy physics.

2. Monopole solutions

(a) Dirac monopole

Classical electrodynamics can describe magnetic charges without any difficulty, and in fact, their existence would make the theoretical description more symmetric. Denoting the magnetic charge density by \( \rho_M \) and magnetic current density by \( J_M \), and their electric counterparts by \( \rho_E \) and \( J_E \), respectively, the Maxwell equations for the electric field \( E \) and magnetic field \( B \) can be written in a complex form,

\[
\begin{align*}
\nabla \cdot (E + iB) &= \rho_E + i\rho_M, \\
\nabla \times (E + iB) &= i \frac{\partial}{\partial t} (E + iB) + i(J_E + iJ_M).
\end{align*}
\]

This is invariant under the electric–magnetic duality transformation, which is a rotation of the complex phase,

\[
\begin{align*}
E + iB &\to e^{i\theta} (E + iB), \\
\rho_E + i\rho_M &\to e^{i\theta} (\rho_E + i\rho_M) \\
J_E + iJ_M &\to e^{i\theta} (J_E + iJ_M).
\end{align*}
\]

Magnetic charges cause a problem when one tries to describe the electromagnetic field using the vector potential \( A \), because then the magnetic...
Magnetic monopoles

The magnetic field \( B = \nabla \times A \) is sourceless and therefore the magnetic charge is necessarily zero. However, one can write down a singular vector potential

\[
A(r) = \frac{g}{4\pi |r|} \frac{r \times n}{|r| - r \cdot n},
\]

where \( r \) is the position vector and \( n \) is a constant unit vector. Apart from the direction of the vector \( n \), this is smooth and describes a radial Coulomb-like magnetic field

\[
B(r) = \frac{g}{4\pi} \frac{r}{|r|^3},
\]

which is identical to the field that a magnetic monopole with charge \( g \) would have. Along vector \( n \), one finds a singular ingoing magnetic field that carries a total flux of \( \Phi = g \) and therefore makes the whole configuration magnetically neutral. This configuration corresponds to an infinitesimally thin solenoid, which ends at the origin and carries electric current \( j_E = g/2\pi R \), where \( R \) is the radius of the solenoid cross section. In classical electrodynamics, such an infinitesimally thin solenoid would be unobservable, and therefore one would appear to have only the monopole field (2.4).

Dirac [3] pointed out that in quantum theory, this solenoid, or ‘Dirac string’, is observable by the effect it has on the complex phase of a charged particle’s wave function. When moving along a path, the complex phase changes by the amount

\[
\Delta \theta = e \int dr \cdot A,
\]

where \( e \) is the electric charge of the particle. For a closed loop, this is equal to \( e\Phi \), where \( \Phi \) is the magnetic flux through the loop. Therefore, the magnetic flux of the solenoid can be detected without touching, for example, by carrying out an Aharonov–Bohm interference experiment. However, if the monopole charge is \( g = 2\pi/e \), then the phase change is a multiple of \( 2\pi \), and it will not have any effect on the interference pattern. Therefore, the Dirac string is unobservable, and one would only see the magnetic monopole.

This means that every monopole has to be connected to a Dirac string, but the location of the string is not observable. Because of this, the string is merely a theoretical construction required for the mathematical description of the monopole, and does not have any physical relevance. The only physical object in the solution is the point-like magnetic monopole located at the origin. From this, Dirac concluded that quantum electrodynamics (QED) allows the existence of point-like magnetic monopoles with charge \( g = 2\pi/e \) or an integer multiple of it.

This has the remarkable implication that if such monopoles exist, then in order for their Dirac strings to be completely unobservable, the electric charge \( e \) of every particle species has to satisfy the Dirac quantization condition

\[
\frac{eg}{2\pi} \in \mathbb{Z}.
\]

Electric charges do, in fact, seem to obey such a condition, which led Dirac to conclude that it would be surprising if magnetic monopoles did not exist [3,19]. More recently, Polchinski [20] has argued along the same lines that any theory which explains charge quantization would predict the existence of magnetic monopoles.

It is important to note that although, as Dirac showed, magnetic monopoles can be consistently described in quantum theory, they do not appear in the particle
spectrum of QED. One can see this by calculating the energy of the monopole configuration (2.4), which turns out to be infinite. Of course, divergences such as this are very common in quantum field theory. For example, the quantum correction to the electron mass due to the electromagnetic field is also divergent. However, when the theory is renormalized, the electron mass and any other physical observables become finite, but the divergence in the monopole energy remains. This means that in order to have monopoles with finite energy, one needs to modify the theory in some way. One simple way to do that is to define the theory on a discrete space–time lattice.

(b) Monopoles on lattices

In lattice field theory, space–time is assumed to be discrete. For simplicity, let us consider a three-dimensional theory, which can be thought of as a classical statistical system. The lattice consists of points \( \mathbf{x} = n \mathbf{a} \), where \( \mathbf{n} \) is an integer-valued component vector and \( a \) is the lattice spacing. In gauge field theories such as QED, the gauge field is represented by link variables \( U_i \) defined on links between the lattice points.

There are essentially two different ways one can formulate QED on a lattice: a compact formulation, which has magnetic monopoles, and a non-compact one which does not. In the compact formulation, the link variable is related to the continuum vector potential \( A_i \) as

\[
U_i = \exp(i a e A_i),
\]

which means that it is a complex number with unit norm. The dynamics of the theory is determined by the action

\[
S = \beta \sum_{\mathbf{x}} \sum_{i<j} P_{ij}(\mathbf{x}),
\]

where the plaquette \( P_{ij} \) is the path-ordered product of four link variables around an elementary closed loop,

\[
P_{ij}(\mathbf{x}) = U_i(\mathbf{x}) U_j(\mathbf{x} + \hat{i}) U^*_i(\mathbf{x} + \hat{j}) U^*_j(\mathbf{x}).
\]

The complex phase of the plaquette gives the magnetic field strength,

\[
P_{ij}(\mathbf{x}) = \exp[i a e (A_i(\mathbf{x}) + A_j(\mathbf{x} + \hat{i}) - A_i(\mathbf{x} + \hat{j}) - A_j(\mathbf{x}))] \to \exp(i a^2 \epsilon_{ijk} B_k).
\]

The magnetic charge inside a lattice cube is equal to the total magnetic flux out of the cube, which is given by the six plaquettes on the sides of the cube. In this formulation, the magnetic charge can be a non-zero integer multiple of \( g = 2\pi/e \), in accordance with the Dirac quantization condition. Because the lattice spacing provides an ultraviolet cut-off, the energy of a monopole is finite. Therefore, the theory actually has dynamical magnetic charges. Their energy is proportional to the inverse lattice spacing, and therefore they disappear in the continuum limit \( a \to 0 \).

The monopoles discovered in spin ices \([13]\) are in many ways similar to those in lattice QED. One can think of the spin ice as a lattice of magnetic dipoles \( s_i = \pm 1 \) on links between lattice sites. In the ground state of the system, the spins
are arranged in such a way that they satisfy the ‘ice rule’ that the dipoles cancel at each lattice site. There is a very large number of different spin configurations that satisfy this, and if the temperature of the system is higher than the energy difference between them, then they are practically degenerate, and the typical state of the system consists of highly disordered spins satisfying the ice rule.

Violation of the ice rule at a lattice site means that the site has effectively a non-zero magnetic charge. Of course, these effective monopoles do not actually violate the magnetic Gauss law $\nabla \cdot B = 0$; so they are not real magnetic monopoles. Instead, they are end-points of a line of ‘flipped’ dipoles, which can be thought of as the analogue of a Dirac string. Because the spins are disordered, the whole system is actually filled with tangled Dirac strings, and therefore it is impossible to locate and identify the string that is connecting a given pair of monopoles. To a good approximation, the monopoles behave like free magnetic charges interacting through a magnetic Coulomb force. As in lattice QED, this picture is valid at distances much longer than the lattice spacing.

There is, however, an important difference between monopoles in lattice QED and in spin ices. In the former case, the Dirac strings are completely unphysical, but in the latter, they can be ignored only because they are so disordered. This can be seen by considering the spin ice system in an external magnetic field [14]. Initially, the monopoles start to move in the direction of the magnetic field and a magnetic current is generated. However, this motion makes the spin configuration more ordered, and eventually the current stops when the spins are completely aligned and the monopoles can no longer move.

(c) ’t Hooft–Polyakov monopoles

The lattice monopoles are a simple example of how QED can be modified to allow magnetic charges. They also illustrate the general result that the mass of the monopoles is at the scale at which the new, modified physics appear. In the case of lattice QED, this is the inverse lattice spacing. Therefore, the monopoles are really lattice objects.

However, it is possible to modify QED at short distances to allow magnetic charges without having to discretize the theory. In 1974, ’t Hooft [4] and Polyakov [5] found that this happens whenever the U(1) gauge group of electrodynamics is embedded in a larger unified gauge group. In particular, this is the case in all grand unified theories (GUTs) [21], which unify electrodynamics with weak and strong interactions into one unified force. In these theories, magnetic monopoles appear as topological solitons: smooth, nonlinear field configurations that are stable for topological reasons.

They considered a simplified model of a GUT consisting of a three-component real scalar field $\Phi^a$, ($a = 1, 2, 3$), coupled to a non-Abelian SO(3) gauge field. The potential of the scalar field has the form

$$V(\Phi) = \frac{\lambda}{4} (\Phi^a \Phi^a - v^2)^2. \quad (2.11)$$

The classical vacuum state corresponds to the minimum of the potential, but in this case, there is a sphere of degenerate minima given by the condition $\Phi^a \Phi^a = v^2$. Although the theory itself is symmetric under three-dimensional rotations of $\Phi$, which form the SO(3) Lie group, each of these vacuum states is only invariant...
under rotation around the $\Phi$ vector. Therefore, the SO(3) symmetry is said to be broken spontaneously to U(1). As in QED, this U(1) gauge symmetry gives rise to the electromagnetic field. For example, if the vacuum state is $\Phi = (0, 0, v)$, then the component $A_3^\mu$ behaves as the electromagnetic vector potential.

In addition to the classical vacuum solution, one can find another simple time-independent solution by assuming a ‘hedgehog’ shape in which the direction of the scalar field in the internal space is coupled to the direction in space,

$$\Phi^\alpha(\mathbf{x}) = v f(|\mathbf{x}|) \frac{x_\alpha}{|\mathbf{x}|}. \quad (2.12)$$

By continuity, the field $\Phi$ has to vanish at the origin, so $f(0) = 0$, and in order for it to approach the vacuum at infinite distance, one must have $\lim_{r \to \infty} f(r) = 1$. Then, one has a smooth field configuration that cannot be continuously transformed into a vacuum solution. Because of this topological obstruction, it is classically stable.

Without the gauge field, the gradient energy of the scalar field would make the total energy of the solution divergent. However, the gauge field $A_\mu^a$ couples to the scalar field through the covariant derivative

$$(D_\mu \Phi)^a = \partial_\mu \Phi^a + e e^{abc} A_\mu^b \Phi^c. \quad (2.13)$$

Far from the origin, the gauge field cancels the gradient energy if

$$A^a_\mu(r) \to -\frac{1}{e^2} \epsilon_{\alpha\beta\gamma} \frac{x_\beta}{|\mathbf{x}|^2}, \quad (2.14)$$

and one finds a solution with a finite energy. The precise forms of the scalar and gauge fields can be only obtained numerically, but generally the mass of the monopole is determined by the energy scale at which the symmetry scale is restored.

Because the $\Phi$ field is now position dependent, one cannot simply identify $A^3_\mu$ with the electromagnetic vector potential. However, ’t Hooft [4] found a general expression that gives the electromagnetic field strength tensor in an arbitrary field configuration,

$$\mathcal{F}_{\mu \nu} = \hat{\Phi}^a F_{\mu \nu}^a + \frac{1}{e} \epsilon_{abc} \hat{\Phi}^a (D_\mu \hat{\Phi})^b (D_\nu \hat{\Phi})^c, \quad (2.15)$$

where $\hat{\Phi}^a = \Phi^a / \sqrt{\hat{F}^a \hat{F}^a}$ and $F_{\mu \nu}$ is the non-Abelian field strength tensor,

$$F_{\mu \nu}^a = \partial_\mu A^a_\nu - \partial_\nu A^a_\mu - e \epsilon_{abc} A^b_\mu A^c_\nu. \quad (2.16)$$

The ’t Hooft tensor $\mathcal{F}_{\mu \nu}$ is singular at the origin, and around it, the corresponding magnetic field $B_\mu = \epsilon_{ijk} F_{jk}^\mu$ has the magnetic monopole form with magnetic charge $g = 4 \pi / e$. Therefore, the ’t Hooft–Polyakov monopole is a magnetic monopole.

More generally, one can see that whenever the field values on a sphere in space cannot be continuously transformed to a constant, without leaving the set of vacua, there must be a monopole inside the sphere. In topology, homotopy theory states that such topologically non-trivial configurations are possible if and only if the second homotopy group $\pi_2(\mathcal{M})$ of the set $\mathcal{M}$ of possible vacua is non-trivial [22]. One can further show that this is always the case when the original symmetry group of the theory is a simple Lie group and the residual group in the broken phase contains a U(1) factor [23]. What is meant by a GUT
in particle physics is precisely a theory that unifies electrodynamics with weak and strong nuclear interactions into one elementary interaction in this way [21], and therefore one has to conclude that the existence of magnetic monopoles is an unavoidable prediction of GUTs. The mass of these GUT monopoles would be around $10^{16}$ GeV.

More recently, the focus in high-energy physics has shifted from GUTs based on quantum fields to ‘theories of everything’ such as superstring theory, which aim to include also gravity. The same conclusion applies also to them, although the topological argument given above cannot be applied in the same form.

### 3. Monopole searches

If monopoles are indeed such a generic prediction of particle-physics models, why have we not found them? There have been many attempts to detect them, both directly in experiments or indirectly by astronomical observations. The following is a brief summary, and more details can be found in recent studies [8,9].

**(a) Accelerator searches**

In principle, magnetic monopoles should be produced in particle accelerator experiments if the collision energy is sufficiently high, higher than $2\sqrt{s}$. For GUT monopoles, the required energy is at least 12 orders of magnitude higher than the energies available at the Large Hadron Collider (LHC). Therefore, it is unrealistic to expect that they could be produced in any foreseeable particle accelerators.

However, intermediate-mass monopoles, by which we mean monopoles with mass well below the GUT scale, could potentially be produced at the LHC, and there is a dedicated experiment, the Monopole and Exotics Detector at the Large Hadron Collider [24], which attempts to detect them by looking for damage that the monopoles would cause to plastic sheets placed around the collision point. Monopoles have also been searched for in other accelerators in the past, such as Tevatron, the Large Electron–Positron collider and the Hadron Electron Ring Accelerator. There are also indirect constraints arising from the way virtual monopole–antimonopole pairs would modify other scattering processes. There is some uncertainty in these calculations because of the strong magnetic charge of the monopoles, which makes perturbation theory unreliable, and non-perturbative numerical methods are not yet sufficiently developed for them [25]. Nevertheless, the accelerator experiments can currently exclude monopoles of mass less than roughly 1 TeV [26].

**(b) Direct searches**

Instead of trying to produce monopoles in an experiment, one can also try to look for monopoles that already exist in the universe. Monopoles are stable particles, and therefore even monopoles created in the early universe should still be around. Because of the Dirac quantization condition, their magnetic field is strong, and their behaviour is very different from other, electrically charged particles. Therefore, they should be fairly easy to detect and identify. For example, if a monopole passes through a conductor loop, it generates an electric current
by induction. In a constant electric or magnetic field, the monopole would follow a characteristic trajectory. Because of their strong electromagnetic interaction, they would lose energy much more rapidly than electrically charged particles when travelling through matter. In many GUTs, monopoles would also induce nucleon decay [27,28]. On the other hand, the likelihood of finding a monopole is, obviously, proportional to their flux because it determines how often a monopole would hit the experiment.

There have been many attempts to look for magnetic monopoles in cosmic rays. Some early experiments appeared to show evidence for them [29,30], but these turned out to be false. However, they motivated several large-scale experiments that have provided significant limits on magnetic monopoles and also paved the way for dark matter searches. The Monopole, Astrophysics and Cosmic Ray Observatory operated underground in Gran Sasso, Italy, from 1989 to 2000, and was specially designed to detect magnetic monopoles. Its failure to find any puts an upper limit of roughly $F \lesssim 10^{-16} \text{cm}^{-2} \text{s}^{-1} \text{sr}^{-1}$ on the flux of magnetic monopoles [31]. Several other experiments such as the Antarctic Muon and Neutrino Detector Array [32] and Baikal [33] have produced similar bounds. For intermediate-mass monopoles, i.e. lighter than GUT monopoles, the radio ice Cherenkov experiment (RICE experiment) at the South Pole gives an even stronger bound of roughly $F \lesssim 10^{-18} \text{cm}^{-2} \text{s}^{-1} \text{sr}^{-1}$ [34]. There have also been attempts to find magnetic monopoles trapped in materials such as moon rock, meteorites and sea water, but with no success [35].

(c) Astrophysical bounds

Magnetic monopoles would also have astrophysical effects, which can be used to look for them and constrain their flux. One of these, the Parker bound [36], follows from the existence of a magnetic field of roughly $3 \mu G$ in our galaxy. This field would make any magnetic monopoles accelerate, which drains energy from the field. The rate of this dissipation depends on the flux of the monopoles, and therefore one obtains an upper bound on the flux. If the monopoles are very heavy, $M \gtrsim 10^{17} \text{GeV}$, this effect is weaker because their motion is dominated by gravitational forces, but for lighter monopoles, one finds the bound

$$F \lesssim 10^{-15} \text{cm}^{-2} \text{s}^{-1} \text{sr}^{-1}. \quad (3.1)$$

An even simpler bound is due to the total mass of the monopoles in the universe. From astronomical observations [12], we know that matter particles make up approximately 23 per cent of the total energy in the universe, and roughly one-third of this can be accounted for by known particles. In principle, some or all of the remaining dark matter could be magnetic monopoles. If they are light enough, $M \lesssim 10^{-17} \text{GeV}$, they are not gravitationally bound to galaxies, and one can approximate that they are roughly uniformly distributed in the universe. From the observed dark matter density, one obtains an upper bound

$$F \lesssim 10^{-16} \left(10^{16} \text{GeV}/M\right) \left(\frac{v}{10^{-3}c}\right) \text{cm}^{-2} \text{s}^{-1} \text{sr}^{-1}, \quad (3.2)$$

where $v$ is the average speed of the monopoles. Because of the mass dependence, this bound is only relevant for heavy monopoles, such as GUT ones.
4. Monopoles in cosmology

In principle, if one had a good enough understanding of the early universe, one should be able to calculate a prediction for the flux of monopole and compare it with the experimental and observational bounds. However, for monopoles, this prediction is many orders of magnitude too high [6]. This is known as the monopole problem.

The precise process by which monopoles would have been produced in the early universe depends on their details. If the grand unified symmetry was initially unbroken, GUT monopoles would have been formed in the phase transition in which this symmetry became spontaneously broken. One obtains a simple lower bound for their number density by considering the Kibble mechanism [22], which is a consequence of causality.

Before the transition, the universe is in the symmetric phase, and there is a unique vacuum state, \( \Phi = 0 \). When the transition takes place, the system needs to choose one possible vacuum state, characterized by different values of \( \Phi \) with \(|\Phi| = v\). All of these vacua are identical because they are related to each other by symmetry. Therefore, the system has no reason to favour any vacuum state, and the choice is random. Furthermore, the transition takes place in a finite time, and information can travel only a finite distance \( \hat{\xi} \) during this time; therefore; the choice will be uncorrelated at distances longer than this. In Big Bang cosmology, this correlation length \( \hat{\xi} \) has to be shorter than the particle horizon \( \sim 1/H \), where \( H \) is the Hubble rate.

After the transition, the universe will therefore consist of domains of size \( \hat{\xi} < 1/H \), each of which is in a different vacuum, and these vacua are uncorrelated between different domains. Consider, now, a point where four of these domains meet. On a sphere around that point, the system is in different random vacuum state in different directions. With a probability of order one, the field values on the sphere are topologically non-trivial so that they cannot be continuously deformed to a constant without leaving the set of vacuum states. As discussed in §2c, there then has to be a monopole inside the sphere. This means that one would expect roughly one monopole per domain, and therefore one has a lower bound for the number density \( n \gtrsim H^3 \). Even if one takes into account annihilation reactions between monopoles and antimonopoles, this density is massively higher than the total matter density in the universe [6].

This monopole problem prompted Guth [11] to propose the theory of inflation, according to which the universe expanded at an accelerating rate, thereby diluting the monopole density to an acceptable level. This also solves other puzzling properties of the universe, such as the horizon and flatness problems. Inflation also generates primordial density perturbations, which acted as the initial seeds for formation of the large-scale structure in the universe and can be observed as temperature fluctuations in the cosmic microwave background radiation. Measurements of these fluctuations, by the Wilkinson Microwave Anisotropy Probe satellite [12] and other experiments, have confirmed the predictions of inflation to high accuracy.

In many inflationary models, inflation takes place at energies well below the GUT scale, and therefore if monopoles were formed at all, that would have happened before inflation. Inflation would then have diluted their density down to an unobservably small level. However, in some otherwise viable models, GUT
or intermediate mass monopoles are formed in a phase transition at the end of inflation or shortly afterwards. In that case, the dynamics of the monopole formation can be very different [37], and it is important to understand the constraints one obtains in that case. It is also possible that thermal effects and the dynamics of the gauge field play an important role and modify the predictions [38,39]. It is therefore important to understand the formation of monopoles and their subsequent dynamics.

5. Monopoles and gauge field theory

From a purely theoretical point of view, magnetic monopoles have been extremely useful for understanding the dynamics of strongly coupled gauge field theories. The most important example of such a theory is QCD, which describes strong interactions. The theory is defined in terms of gluon and quark fields, but these are confined into hadrons such as protons and neutrons at low energies. There is a large amount of strong numerical evidence from lattice field theory simulations that the theory can describe this confinement correctly, but there is no actual proof because it is a strongly coupled theory and therefore standard perturbative quantum field theory techniques are not applicable.

’t Hooft [40] and Mandelstam [41] suggested that confinement could be explained as a ‘dual Meissner effect’. The chromoelectric field between quarks would be confined into flux tubes in a similar way as magnetic flux is confined to Abrikosov flux tubes in superconductors, and as a result, no free quarks could exist. In superconductors, this can be understood as a consequence of condensation of electrically charged Cooper pairs, and the dual of this would be condensation of magnetic monopoles in QCD.

It is difficult to make this idea precise in QCD because the theory has no classical monopole solution. However, in some simpler theories, such as 2+1-dimensional lattice QED [42] and some supersymmetric theories [43], it has been shown rigorously that confinement is caused by magnetic monopoles.

More generally, this can be seen as an example of a generalization of the electric–magnetic duality (2.2) to non-Abelian gauge field theories, as first proposed by Montonen & Olive [44]. This duality is strictly valid only in a certain supersymmetric version of QCD, but it is very useful for many calculations because it turns a strongly coupled theory into a weakly coupled one, which can be treated using standard methods. Similar dualities have also turned out to be very important in string theory [45].

In non-supersymmetric theories such as QCD, the dualities are not believed to be exact. The idea of confinement as a dual Meissner effect has been studied extensively using numerical lattice Monte Carlo simulations [46], but in the case of QCD, these suffer from a lack of a clear formulation of the problem. In the Georgi–Glashow model (see §2c), one can study monopoles in more detail [25,47,48], and there are hints that the Montonen–Olive duality may emerge asymptotically near the phase transition.

6. Conclusions

The study of magnetic monopoles in quantum field theories has been both a great success and a disappointment. Disappointingly, no fundamental monopoles...
Magnetic monopoles have been found in nature, but at the same time, theoretical monopole solutions and the use of electric–magnetic duality have given theorists new ways to understand the physics of gauge field theories. The absence of magnetic monopoles in the universe also led to the theory of inflation, which is the cornerstone of modern cosmology.

The recent discovery of effective magnetic monopole quasi-particles in spin ices raises the question whether one could make use of them to take these theoretical advances further. Although the quasi-particles do behave in many ways as fundamental magnetic monopole particles, it is also clear that there are important differences. Perhaps most importantly, the Dirac strings connecting the monopoles are not completely unphysical.

This does not mean that one cannot use spin ice experiments to draw conclusions for fundamental monopoles, but one has to be aware of the differences between the systems and the limitations they impose. Situations in which random thermal fluctuations play an important role would therefore appear most promising, such as studying the formation of monopoles in phase transitions.

The author thanks David Weir for collaboration and extensive discussions, and acknowledges funding by the STFC grant no. ST/G000743/1 and Royal Society International Joint Project JP100273.

References


