An introduction to magnetic monopoles for novices

or, What I Did In My Graduate Studies

Jack M^c^Kenna

SMSAS PGR seminar

25^{th} January 2019

Table of Contents

Table of Contents

Classical electromagnetism: Maxwell's equations

The laws of classical electromagnetism were empirically determined and unified in the 19th century, and are expressed in Maxwell's equations:

$$\nabla \cdot \mathbf{E} = \rho \qquad\qquad \nabla \cdot \mathbf{B} = 0$$

$$\frac{\partial \mathbf{E}}{\partial t} - \nabla \times \mathbf{B} = -\mathbf{j} \qquad\qquad \frac{\partial \mathbf{B}}{\partial t} + \nabla \times \mathbf{E} = 0$$

together with the Lorentz force law:

$$\mathbf{F} = q\left(\mathbf{E} + \mathbf{v} \times \mathbf{B}\right).$$

Here:

- $\mathbf{E}(t, \mathbf{x})$ and $\mathbf{B}(t, \mathbf{x})$ are the electric and magnetic field strength vectors resp.,
- ρ and $\mathbf{j}$ are the electric charge and electric current distributions resp.,
- $\mathbf{F}$ is the force felt by a particle of electric charge q at position $\mathbf{x}$ with velocity $\mathbf{v}$.

Classical electromagnetism: Maxwell's equations

Maxwell originally wrote these equations without the notational benefit of modern vector calculus.

Differentiating (112) with respect to x, y, and z respectively, and substituting, we find

$$\frac{de}{dt} = \frac{1}{4\pi E^2}\frac{d}{dt}\left(\frac{dP}{dx}+\frac{dQ}{dy}+\frac{dR}{dz}\right) \ldots\ldots\ldots\ldots (114);$$

whence

$$e = \frac{1}{4\pi E^2}\left(\frac{dP}{dx}+\frac{dQ}{dy}+\frac{dR}{dz}\right) \ldots\ldots\ldots\ldots (115),$$

the constant being omitted, because $e=0$ when there are no electromotive forces.

We have seen that electromotive force and electric displacement are connected by equation (105). Differentiating this equation with respect to t, we find

$$\frac{dR}{dt} = -4\pi E^2 \frac{dh}{dt} \ldots\ldots\ldots\ldots (111),$$

shewing that when the electromotive force varies, the electric displacement also varies. But a variation of displacement is equivalent to a current, and this current must be taken into account in equations (9) and added to r. The three equations then become

$$\left.\begin{aligned} p &= \frac{1}{4\pi}\left(\frac{d\gamma}{dy}-\frac{d\beta}{dz}-\frac{1}{E^2}\frac{dP}{dt}\right)\\ q &= \frac{1}{4\pi}\left(\frac{d\alpha}{dy}-\frac{d\gamma}{dx}-\frac{1}{E^2}\frac{dQ}{dt}\right)\\ r &= \frac{1}{4\pi}\left(\frac{d\beta}{dx}-\frac{d\alpha}{dy}-\frac{1}{E^2}\frac{dR}{dt}\right)\end{aligned}\right\} \ldots\ldots\ldots\ldots (112),$$

where p, q, r are the electric currents in the directions of x, y, and z; α, β, γ are the components of magnetic intensity; and P, Q, R are the electromotive forces. Now if e be the quantity of free electricity in unit of volume, then the

Let us then find three quantities F, G, H from the equations

$$\left.\begin{aligned} \frac{dG}{dz}-\frac{dH}{dy} &= \mu\alpha\\ \frac{dH}{dx}-\frac{dF}{dz} &= \mu\beta\\ \frac{dF}{dy}-\frac{dG}{dx} &= \mu\gamma\end{aligned}\right\} \ldots\ldots\ldots\ldots (55),$$

with the conditions

$$\frac{1}{4\pi}\left(\frac{d}{dx}\mu\alpha+\frac{d}{dy}\mu\beta+\frac{d}{dz}\mu\gamma\right) = m = 0 \ldots\ldots\ldots\ldots (56),$$

This equation being true for all values of α, β, and γ, first let β and γ vanish, and divide by α. We find

$$\frac{dQ}{dz}-\frac{dR}{dy} = \mu\frac{d\alpha}{dt}$$

Similarly,

$$\frac{dR}{dx}-\frac{dP}{dz} = \mu\frac{d\beta}{dt} \ldots\ldots\ldots\ldots (54).$$

and

$$\frac{dP}{dy}-\frac{dQ}{dx} = \mu\frac{d\gamma}{dt}$$

Photographs of Maxwell's paper *On Physical Lines of Force* (1861), Philosophical Journal **90** 11-23, as reproduced in *The scientific papers of James Clerk Maxwell*, Dover Publications Inc. New York, 1965.

Classical electromagnetism: fields and potentials

The homogeneous equations for the magnetic field tell us we can obtain two potential functions: a vector potential **A** for the magnetic field, and an associated scalar potential ϕ for the electric field, satisfying

$$\mathbf{B} = \nabla \times \mathbf{A}, \quad \mathbf{E} = \nabla\phi - \frac{\partial \mathbf{A}}{\partial t}. \tag{1}$$

We can observe that **E** and **B** will then remain unchanged if we transform the potentials according to

$$\mathbf{A} \mapsto \mathbf{A}' = \mathbf{A} + \nabla\chi, \quad \phi \mapsto \phi' = \phi + \frac{\partial \chi}{\partial t} \tag{2}$$

Such transformations are called *gauge transformations*. Note that χ is a function of space and time, so the transformation can be different at different points: we say gauge symmetry is a *local* symmetry.

Galilean and Lorentzian invariance

Classical physical theories preserve Euclidean distances, and are said to have a "rigid" symmetry group of *Galilean* transformations: translations and rotations. Maxwell's equations do not have this Galilean invariance, but instead satisfy the *Lorentzian* invariance of special relativity, which allows the rigid transformations which preserve distances of the form

$$\mathrm{d}s^2 = -\mathrm{d}t^2 + \mathrm{d}x^2 + \mathrm{d}y^2 + \mathrm{d}z^2. \tag{3}$$

Equivalently, Lorentz transformations are transformations which preserve the pseudo-inner product given by the quadratic form

$$\eta_{\mu\nu} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{4}$$

4-vectors and the 4-potential

The Lorentz invariance of Maxwell's theory lets us write it in the language of 4-vectors, where the position vector in space and time is denoted

$$x^\mu = (x^0, x^1, x^2, x^3) = (t, \mathbf{x}) \tag{5}$$

and similarly all vectors have one time component and three space components. We combine the electric and magnetic potentials into the 4-potential $A_\mu = (\phi, \mathbf{A})$. Then a gauge transformation is a transformation of the form

$$A_\mu \mapsto A'_\mu = A_\mu + \partial_\mu \chi \tag{6}$$

The field strength tensor

Now we define the *field strength* tensor, which captures all components of the electric and magnetic fields:

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu = \begin{pmatrix} 0 & -E_x & -E_y & -E_z \\ E_x & 0 & B_z & -B_y \\ E_y & -B_z & 0 & B_x \\ E_z & B_y & -B_x & 0 \end{pmatrix} \tag{7}$$

In the language of differential geometry, the 2-form F is the exterior derivative of the 1-form A.

Equivalently, given $F_{\mu\nu}$, we can define the electric and magnetic fields to be

$$E_i = F_{i0}, \quad B_i = \tfrac{1}{2}\epsilon_{ijk}F_{jk} \quad (i, j, k \text{ the spatial indices}) \tag{8}$$

Maxwell's equations in Lorentz co-ordinates

By its antisymmetry, $F_{\mu\nu}$ is automatically gauge invariant. Maxwell's equations are reduced to two properties of the field strength. The homogeneous ("magnetic") equations are the components of the *Bianchi identity*:

$$\partial_\lambda F_{\mu\nu} + \partial_\mu F_{\nu\lambda} + \partial_\nu F_{\lambda\mu} = 0 \tag{9}$$

The inhomogeneous ("electric") equations may be expressed

$$\partial_\nu F^{\mu\nu} = j^\mu, \tag{10}$$

where $j^\mu = (\rho, \mathbf{j})$ is called the 4-current and captures the charge and current distributions.

Maxwell's equations, co-ordinate free

Equivalently, Maxwell's equations can be expressed via the field strength tensor in the co-ordinate free language of differential geometry.

- Gauge symmetry arises because the field strength is an *exterior derivative*: $F = \mathrm{d}A$ where

$$A = A_\mu \, \mathrm{d}x^\mu, \quad F = \frac{1}{2} F_{\mu\nu} \, \mathrm{d}x^\mu \wedge \mathrm{d}x^\nu. \tag{11}$$

- The homogeneous Maxwell equations are due to a broad geometric identity shared by all such exterior derivatives.
- The inhomogeneous, dynamical Maxwell equations also have a straightforward differential-geometric expression,

$$\mathrm{d}^\dagger F = j, \tag{12}$$

where $\mathrm{d}^\dagger$ is a related operation called the *adjoint* exterior derivative.

Table of Contents

Sources and fluxes: electric flux

The static Maxwell equations have the interpretation of describing *sources* of potential in the field. Consider the *flux* Φ_E or perpendicular flow of the electric field through the sphere S^2:

$$\Phi_E = \int_{S^2} \mathbf{E} \cdot \mathbf{n} \, \mathrm{d}S \tag{13}$$

where **n** is a unit vector normal to the surface of the sphere. By the divergence theorem (Stokes' theorem), this is equal to an integral over the ball B^3 enclosed by the sphere:

$$\Phi_E = \int_{S^2} \mathbf{E} \cdot \mathbf{n} \, \mathrm{d}S = \int_{B^3} \nabla \cdot \mathbf{E} \, \mathrm{d}x^3 = \int_{B^3} \rho \, \mathrm{d}x^3 \tag{14}$$

The flux is determined entirely by the total electric charge within the sphere. This holds for any closed surface.

Sources and fluxes: magnetic flux

Now comparing with the expression for magnetic flux,

$$\Phi_B = \int_{S^2} \mathbf{B} \cdot \mathbf{n} \, \mathrm{d}S = \int_{B^3} \nabla \cdot \mathbf{B} \, \mathrm{d}x^3 = 0. \tag{15}$$

There can be no magnetic flux through a closed surface, so there can be no isolated magnetic charge.

Specifically, Stokes' theorem tells us that if we have a well-defined global magnetic potential **A**, there is no magnetic flux, so no magnetic monopoles exist.

Why hunt for a monopole?

So why try make sense of what apparently does not exist? Several reasons:

1. Maxwell's electromagnetism shows a complete duality between electricity and magnetism, e.g. ferromagnetism/ferroelectricity, magnetic motors/capacitor motors – *except* for the lack of isolated magnetic charge.
2. Some physical systems in condensed matter theory (e.g. spin ice, spinor Bose-Einstein condensate) contain "effective" magnetic monopoles, or phenomena that are mathematically analogous.
3. Dirac observed that in quantum mechanics, the presence of a single magnetic monopole in the universe would explain why electric charge is discretely quantised.

The Dirac monopole

To model a non-zero magnetic flux which mimicked the workings of electric flux, Dirac (1931) introduced a point magnetic charge of strength g sitting at the origin, satisfying

$$\nabla \cdot \mathbf{B} = 4\pi g \delta(\mathbf{x}). \tag{16}$$

This can be solved to obtain the magnetic field:

$$\mathbf{B} = g\mathbf{x}/|\mathbf{x}|^3 \tag{17}$$

The Dirac monopole admits a non-zero flux,

$$\Phi_B = \int_{S^2} \mathbf{B} \cdot \mathrm{d}\mathbf{S} = \int_{B^3} 4\pi g \delta(\mathbf{x}) \, \mathrm{d}x^3 = 4\pi g. \tag{18}$$

However, it cannot admit a well-defined potential, i.e. it can't be incorporated into the rest of Maxwell's theory.

Local potentials

It is possible to define potentials that give the correct field strength and flux in neighbourhoods of $\mathbb{R}^3$ rather than the total space. Dirac gave the following example with two vector potentials, $\mathbf{A}^N$ and $\mathbf{A}^S$:

$$A_x^N = \frac{-gy}{r(r+z)} \quad A_y^N = \frac{gx}{r(r+z)} \quad A_z^N = 0 \tag{19}$$

$$A_x^S = \frac{gy}{r(r-z)} \quad A_y^S = \frac{-gx}{r(r-z)} \quad A_z^S = 0 \tag{20}$$

where $r = |\mathbf{x}|$. $\mathbf{A}^N$ is well-defined everywhere except the non-positive z-axis (where $r + z = 0$), so it's well-defined in the "north" half of the universe ($z \geq 0$, except the origin). Similarly, $\mathbf{A}^S$ is well-defined everywhere except the non-negative z-axis, so models the potential of a magnetic monopole in the "south" half of the universe.

The Wu-Yang monopole

Wu & Yang (1975) suggested that, in the absence of a global potential for a monopole, we can "glue together" Dirac's local potentials via a gauge transform. On the unit sphere S^2, the difference between $\mathbf{A}^N$ and $\mathbf{A}^S$ at the equator may be expressed in polar co-ordinates as

$$\mathbf{A}^N - \mathbf{A}^S = \frac{2g}{r \sin\theta}\hat{\mathbf{e}}_\phi = \nabla(2g\phi) \tag{21}$$

These potentials only differ by a change in gauge: $\mathbf{A}^N = \mathbf{A}^S + \nabla\chi$, with $\chi = 2g\phi$. Maxwell's equations are unaffected. Moreover, using Stokes' theorem we can show that this combination of *local* potentials has the desired total *global* flux:

The Wu-Yang monopole

Let S^2_+ and S^2_- be the north and south hemispheres respectively of the unit sphere S^2. The total flux through the sphere is given by the following.

$$\begin{aligned}
\Phi_B &= \int_{S^2} \mathbf{B} \cdot \mathrm{d}\mathbf{S} = \int_{S^2} \nabla \times \mathbf{A} \cdot \mathrm{d}\mathbf{S} \\
&= \int_{S^2_+} \nabla \times \mathbf{A}^N \cdot \mathrm{d}\mathbf{S} + \int_{S^2_-} \nabla \times \mathbf{A}^S \cdot \mathrm{d}\mathbf{S} \\
&= \int_{\text{equator}} \mathbf{A}^N \cdot \mathrm{d}\mathbf{s} - \int_{\text{equator}} \mathbf{A}^S \cdot \mathrm{d}\mathbf{s} \\
&= \int_{\text{equator}} (\mathbf{A}^N - \mathbf{A}^S) \cdot \mathrm{d}\mathbf{s} \\
&= \int_{\text{equator}} \nabla(2g\phi) \cdot \mathrm{d}\mathbf{s} = 4\pi g
\end{aligned}$$

The geometry of the Wu-Yang model

This description of the Wu-Yang monopole involves seperate neighbourhoods or *patches* of spacetime, to each of which we associate information about the choice of gauge representative of the potential.

The geometric structure is captured by the gauge transformation which glues the patches of configuration space together smoothly. This idea is made rigorous in the theory of *fibre bundles*.

Table of Contents

Fibre bundle: technical definition

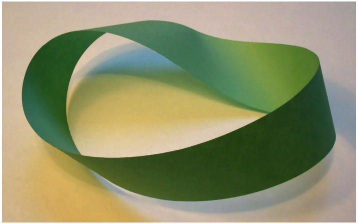
A *fibre bundle* (E, π, M, F, G) or $\pi : E \to M$ or $E \xrightarrow{\pi} M$ is

1. A differentiable manifold E, the *total space*
2. A differentiable manifold M, the *base space*
3. A differentiable manifold F, the *(typical) fibre*
4. A surjection $\pi : E \to M$, the *projection*, satisfying $\pi^{-1}(p) = F_p \cong F$. F_p is called the fibre at p.
5. A Lie group G, the *structure group*, acting on F from the left
6. A set of open coverings $\{U_i\}$ of M and diffeomorphisms $\phi_i : U_i \times F \to \pi^{-1}(U_i)$ such that $\pi\phi_i(p, f) = p$.
7. The *transition functions* on overlapping patches of the bundle $U_i \cap U_j \neq \emptyset$ given by $t_{ij}(p) = \phi_i^{-1}\phi_j(p, \cdot) : F \to F$ should be diffeomorphisms in G.

Fibre bundles: simple example

The cylinder and the Möbius strip are both examples of a fibre bundle of the interval over the circle. The cylinder is the trivial bundle $[0,1] \times S^1$, while the Möbius strip is non-trivial, and has non-trivial transition functions.

A Möbius strip[1]



[1] Image copyright David Benbennick, licenced under Creative Commons Attribution-ShareAlike 3.0 Unported.

$\mathrm{U}(1)$ bundles and electromagnetism

The gauge freedom in a model of electromagnetism is always parametrised by an angle or equivalently a phase $\mathrm{e}^{\mathrm{i}\alpha} \in \mathrm{U}(1)$, where

$$\mathrm{U}(1) = \{z \in \mathbb{C} \big| |z| = 1\}. \tag{22}$$

$\mathrm{U}(1)$ is topologically equivalent to the circle S^1. (Cf. the Wu-Yang monopole, where the gauge freedom was just proportional to the azimuthal angle ϕ.)

An electromagnetic system can be modelled by a $\mathrm{U}(1)$ bundle over a spacetime manifold.

The Wu-Yang monopole as a $\mathrm{U}(1)$ bundle

We can now construct the Wu-Yang monopole as a $\mathrm{U}(1)$ bundle over the sphere S^2. We take neighbourhoods S^2_+ and S^2_- the hemispheres of S^2 as before, and in each neighbourhood specify the gauge by choices of phase $\mathrm{e}^{\mathrm{i}\phi_+}$ and $\mathrm{e}^{\mathrm{i}\phi_-}$ resp.

The transition function on the equator S^1 must be an element of $\mathrm{U}(1)$ that depends only on the azimuthal coordinate ϕ. Thus it takes the form $\mathrm{e}^{\mathrm{i}\alpha(\phi)}$, and must satisfy

$$t_{NS}(\phi) : \mathrm{U}(1) \to \mathrm{U}(1) : \mathrm{e}^{\mathrm{i}\phi_-} \mapsto \mathrm{e}^{\mathrm{i}\phi_+} = \mathrm{e}^{\mathrm{i}\alpha(\phi)}\mathrm{e}^{\mathrm{i}\phi_-} \tag{23}$$

Solutions in the homotopy group

Suppose we seek a transition function of the form $t_{NS}(\phi) = e^{in\phi}$. Then for the gauge to be well-defined on the equator, we must have $n \in \mathbb{Z}$ This *winding number* is how many times we orbit the gauge space (topologically, the circle) on one orbit of the equator of S^2 (also topologically the circle) Thus the possible $\mathrm{U}(1)$ bundles over S^2 are classified by the homotopy group of the circle.

- The case $n = 0$ has trivial transition function. The fibre bundle is globally $S^1 \times S^2$, the potential is globally defined, and there can be no magnetic monopole.
- The case $n = 1$ corresponds to the Wu-Yang monopole with one unit of magnetic charge g.
- Other positive cases can be interpreted as monopoles with magnetic charge ng, as the total flux can be observed to be $4n\pi g$.

S^3 as a $\mathrm{U}(1)$ bundle

A $\mathrm{U}(1)$ over S^2 has the mathematical interpretation that it describes a smooth map of the 3-dimensional sphere S^3 onto the 2-dimensional sphere S^2. The case $n = 1$ is called the Hopf fibration and is relatively simple to describe. Embed S^3 in $\mathbb{R}^4$ according to

$$S^3 = \{(x_1, x_2, x_3, x_4) \in \mathbb{R}^4 \big| x_1^2 + x_2^2 + x_3^2 + x_4^2 = 1\}. \tag{24}$$

Define the projection $\pi : S^3 \to \mathbb{R}^3$ by

$$\begin{aligned}
\xi_1 &= 2(x_1 x_3 + x_2 x_4) \\
\xi_2 &= 2(x_2 x_3 - x_1 x_4) \\
\xi_3 &= x_1^2 + x_2^2 - x_3^2 - x_4^2.
\end{aligned}$$

Then $\xi_1^2 + \xi_2^2 + \xi_3^2 = 1$, so this is a projection to S^2. It can be shown that this is the $n = 1$ $\mathrm{U}(1)$ bundle.

Hopf co-ordinates

A co-ordinate change helps visualise the fibres. Set

$$
\begin{aligned}
x_1 &= \cos(\varphi_1 + \varphi_2)\cos\eta \\
x_2 &= \sin(\varphi_1 + \varphi_2)\sin\eta \\
x_3 &= \cos(\varphi_2 - \varphi_1)\cos\eta \\
x_4 &= \sin(\varphi_2 - \varphi_1)\sin\eta
\end{aligned}
$$

This implies

$$
\begin{aligned}
\xi_1 &= \sin(2\eta)\cos\varphi_1 \\
\xi_2 &= \sin(2\eta)\sin\varphi_1 \\
\xi_3 &= \cos(2\eta).
\end{aligned}
$$

φ_2 can be interpreted as the fibre co-ordinate which is eliminated by the projection π.

Visualising the Hopf bundle

The following video gives a beautiful visualisation of the Hopf bundle, illustrating aspects such as its non-trivial structure and links to topics such as knot theory.

Hopf fibration – fibers and base (Niles Johnson)