Chapter 20
Monopoles

December 4, 2019

The theoretical possibility of magnetic monopoles is explored, ending in the modern understanding that monopoles are allowed, but only in quantized units. The analysis also leads to a deeper appreciation of the extent to which the 4-vector potential is needed and unique.

Contents

1 Introduction

The question being posed

Even beginning students must have been struck by an apparent symmetry between electricity and magnetism in Maxwell's equations, especially in vacuum. Perfect analogy would require magnetic charges, also called monopoles. This Chapter deals with the question:

$\boxed{\text{Can monopoles exist?}}$

We shall assume that magnetic charge, if it exists, is conserved. (Since the possible existence is motivated by an analogy with electric charge, it would seem very odd if magnetic charge were not conserved.)

Three different questions

Before launching into the analysis, it has to be made clear that there are three different questions:

A Can monopoles exist?

B Do monopoles exist?

C Is a bar magnet made up of two monopoles?

The answer to C is definitely "No": the magnetic field inside a bar magnet goes from the S pole to the N pole, so there is no net flux from one end of the pole. This question can be put aside.

The difference between A and B can be illustrated if "monopoles" is replaced with "cats with five legs and three tails". In the latter case, the answers would be as follows.

A Yes, such cats *can* exist: no law of biology would forbid such a cat.

B No, such cats *do not* exist, as an observational fact.

This Chapter deals mainly with question A for monopoles, and only briefly with question B in Section 7. For the moment, it suffices to say that no magnetic monopole has been definitively found, so the presumed answer to B is "No" at least for now.

Being more specific

Suppose magnetic monopoles exist, with a density

$\rho_{\rm m}$. Then to be specific, our question is whether it is possible to have

$$\nabla \cdot \mathbf{B} \quad = \quad C\,\rho_{\rm m} \neq 0$$

in analogy to Gauss' law for electricity. There are at least three subsidiary questions:

- What is the constant C? Since the unit of magnetic charge is not yet defined, this is a matter of convention only.
- How does the corresponding current $\mathbf{J}_{\rm m}$ enter Maxwell's equations? In short, what would the four Maxwell's equations look like?
- What would be the Lorentz force law for a magnetic monopole?

The answer to these subsidiary questions turn out to be not difficult, if $\rho_{\rm m}$ can exist.

Three levels to the analysis

The attempt to answer question A has gone through a number of twists and turns historically.

- At the classical level, monopoles are possible, as shown in Section 2.
- At the quantum level, it seems monopole are not possible. The argument is presented in Section 3. This would seem to provide a good reason why monopoles have not been found.
- But a more careful analysis, starting with Dirac in the 1930s [1], showed that monopoles are in fact possible in quantum mechanics, provided they are quantized. A charge e and a monopole g can coexist only if[1]

  $$\boxed{eg = nK} \tag{1}$$

  where n is an integer and K is a fundamental constant.[2]

In fact, the most important part of this Chapter is this quantization condition, to be presented in Sections 4 to 6. There are at least three ways to arrive at this result.

- The original method of Dirac [1] is technically difficult and will only be sketched.
- A more illuminating derivation was given by Wu and Yang [2, 3] in the 1970s, leading to a new perspective on the potential A_μ.
- The quantization condition can also be finessed in a much simpler derivation that does not even mention potentials [4].[3]

The most interesting development is the new perspective on the potential, in an extremely elegant formalism [2, 3] that stands at the interface between physics and differential geometry. The method due to Wu and Yang, as well as the theoretical concepts that it throws up, will be the main focus of this Chapter.

One cannot but help recall the very first sentence of the Dirac paper [1]:

> The steady progress of physics requires for its theoretical formulation a mathematics that gets continually more advanced. This is only natural and to be expected.

2 Monopoles in classical EM

This Section analyzes how monopoles could be accommodated within classical EM.

2.1 Duality

The apparent symmetry between electricity and magnetism in vacuum is made precise by the idea of *duality*, as manifested by the mapping

$$F^{\mu\nu} \quad \mapsto \quad \tilde{F}^{\mu\nu}$$

or equivalently

$$\begin{aligned} \mathbf{E}/c \quad &\mapsto \quad \mathbf{B} \\ \mathbf{B} \quad &\mapsto \quad -\mathbf{E}/c \end{aligned}$$

which can be thought of as a 90 degree rotation in the *duality plane* (**Figure 1**), with one direction being electric and the other being magnetic. For compact notation, define the vector[4]

$$\tilde{\mathbf{E}} \quad = \quad \mathbf{E}/c + i\mathbf{B}$$

so the duality map is

$$\tilde{\mathbf{E}} \quad \mapsto \quad -i\,\tilde{\mathbf{E}}$$

Indeed one can rotate in the duality plane by any angle θ, or equivalently multiply these complex quantities by any phase $e^{i\theta}$. (This vector $\tilde{\mathbf{E}}$ was called $\mathbf{Z}$ in a previous Chapter.)

A trivial word about units

Since the key idea is symmetry between $\mathbf{E}$ and $\mathbf{B}$, it is somewhat inconvenient that they have different units. Therefore we shall consider $\mathbf{B}$ together with

$$\mathbf{E}' \quad = \quad \mathbf{E}/c$$

[1] The symbol e rather than q is used since this should refer to the fundamental charge rather than an arbitrary charge.

[2] The value of K is deferred since the unit for magnetic charge is not yet defined.

[3] Not many years after the papers by Wu and Yang [2, 3], Yang was giving a graduate course at CUHK. In that course, he worked out the scattering between a charge and a monopole, but took care to avoid the singular (but in fact simpler) case of a head-on collision. We asked ourselves what would happen in that case, and discovered an alternate derivation of the quantization condition.

[4] The meaning of $\tilde{}$ here is different from that in $\tilde{F}$.

which is the quantities that appear in $\tilde{\mathbf{E}}$ and in $F^{\mu\nu}$. Moreover, for the charge density we shall define

$$\rho' = \rho c$$

which has the same units as $\mathbf{J}$.

Problem 1
In terms of these variables, show that Maxwell's equations can be written as

$$\begin{aligned} \nabla \cdot \mathbf{E}' &= \mu_0 \rho' \\ \nabla \cdot \mathbf{B} &= 0 \\ \nabla \times \mathbf{E}' + \frac{\partial}{\partial(ct)}\mathbf{B} &= 0 \\ \nabla \times \mathbf{B} - \frac{\partial}{\partial(ct)}\mathbf{E}' &= \mu_0 \mathbf{J} \end{aligned} \tag{2}$$

and check that in vacuum this set of equations satisfy duality as defined above. §

2.2 Introducing monopoles

Maxwell's equations
If there is a monopole density ρ_m with corresponding current density $\mathbf{J}_\mathrm{m}$, then the obvious generalization of the Maxwell equations must be

$$\begin{aligned} \nabla \cdot \mathbf{E}' &= \mu_0 \rho' \\ \nabla \cdot \mathbf{B} &= \mu_0 \rho'_\mathrm{m} \\ \nabla \times \mathbf{E}' + \frac{\partial}{\partial(ct)}\mathbf{B} &= -\mu_0 \mathbf{J}_\mathrm{m} \\ \nabla \times \mathbf{B} - \frac{\partial}{\partial(ct)}\mathbf{E}' &= \mu_0 \mathbf{J} \end{aligned} \tag{3}$$

The two novel terms are explained in turn.

Choice of units
In the $\nabla \cdot \mathbf{B}$ equation, the coefficient of the ρ'_m term simply represents the most natural choice, such that ρ and ρ_m have the same units. Thus magnetic charges or monopole strengths g are measured in the same units as charge — Coulomb or esu as the case may be.

Coulomb's law
Solving the two divergence equations for a δ-function at the origin gives Coulomb's law for a point charge q and its analog for a point monopole g:

$$\begin{aligned} E &= \frac{\mu_0 c^2}{4\pi}\frac{q}{r^2} \\ B &= \frac{\mu_0 c}{4\pi}\frac{g}{r^2} \end{aligned} \tag{4}$$

In the E-equation, one power of c comes from changing ρ' to ρ, and another power comes from changing $\mathbf{E}'$ to $\mathbf{E}$. In the B-equation, there is only one power of c, coming from changing ρ'_m to ρ_m.

The current term
There may be some question about the sign with which $\mathbf{J}_\mathrm{m}$ appears. There are two ways to approach this. First, define

$$\begin{aligned} \tilde{\rho} &= \rho + i\rho_\mathrm{m} \\ \tilde{\mathbf{J}} &= \mathbf{J} + i\mathbf{J}_\mathrm{m} \end{aligned}$$

and it is obvious that the full duality transformation should be

$$(\tilde{\mathbf{E}}, \tilde{\rho}, \tilde{\mathbf{J}}) \mapsto -i\,(\tilde{\mathbf{E}}, \tilde{\rho}, \tilde{\mathbf{J}}) \tag{5}$$

or indeed multiplication by any $e^{i\theta}$. Application of this law determines the sign of the $\mathbf{J}_\mathrm{m}$ term.

The second approach relies on conservation.

Problem 2
Show that the modified equations, as shown, are consistent with the conservation of magnetic charge, but that there would be an inconsistency if the sign of $\mathbf{J}_\mathrm{m}$ is reversed. §

Consolidating Maxwell's equations
In fact, Maxwell's equations can be summarized into two equations:

$$\begin{aligned} \nabla \cdot \tilde{\mathbf{E}} &= \mu_0 \tilde{\rho}' \\ \nabla \times (i\tilde{\mathbf{E}}) + \frac{\partial}{\partial(ct)}\tilde{\mathbf{E}} &= -\mu_0 \tilde{\mathbf{J}} \end{aligned} \tag{6}$$

Problem 3
Check that (6) reproduces the four Maxwell's equations. Pay attention to signs. §

Lorentz force law
The force on a point charge q is

$$\mathbf{F} = q\,(\mathbf{E}'c + \mathbf{v} \times \mathbf{B}) \tag{7}$$

The duality transformation then gives the force on a point monopole g:

$$\mathbf{F} = g\,(\mathbf{B}c - \mathbf{v} \times \mathbf{E}')$$

or, expressing back in terms of $\mathbf{E}$:

$$\mathbf{F} = g\,(\mathbf{B}c - \mathbf{v} \times \mathbf{E}/c) \tag{8}$$

In particular, the force on a stationary g' due to another stationary monopole g at a distance r is given by

$$\begin{aligned} F &= g'Bc \qquad \text{from (8)} \\ &= g' \cdot \frac{\mu_0 c}{4\pi}\frac{g}{r^2} \cdot c \qquad \text{from (4)} \\ &= \frac{1}{4\pi\epsilon_0}\frac{gg'}{r^2} \end{aligned} \tag{9}$$

which is *exactly* the same as the corresponding formula for two charges. This should not be surprisingly, since monopoles strengths are measured in the same unit as electric charges, say in coulombs.

One might object that the analogy leading to (8) does not constitute a derivation. Let us sketch a better justification, as follows.

- The energy and momentum density and flux of the fields involve terms such as $E_i E_j/c^2 + B_i B_j$, which are invariant under a duality transformation. (For the Poynting vector, see Problem 4 below.)
- It is known that the force on electric charges and currents (LHS) is equal to the rate of change of EM momentum (RHS). Perform a duality rotation. The LHS becomes the force on magnetic monopoles and their currents as stated above; the RHS remains unchanged. This shows that the appropriate force expression for monopoles is the dual transform of the force on charges. An explicit demonstration can be found in [5], which can be summarized by saying that every formula is the dual transform of the standard one.

Problem 4
In discussing the energy and momentum of the fields, there are also quantities such as $\mathbf{E} \times \mathbf{B}$. Show that these are also invariant under the duality transformation. §

2.3 Freedom to rotate in duality space

Duality is the same as rotations between the electric and magnetic directions (**Figure 1**). Suppose there are only electric charges (the "usual" situation that we deal with). There is no reason — other than convention — to prevent us from making a 90 degree rotation and calling them monopoles; in other words, to say that there are only monopoles and no charges!

Therefore we need to be slightly more careful and rephrase our question A as:

In the convention in which an electron is called a pure charge, can there be monopoles?

Henceforth this qualification (which also applies to questions B and C) will be understood without being stated explicitly.

To put this in another way, suppose there are particles j with electric and magnetic charges given by (q_j, g_j). The only question is whether these, regarded as vectors in the duality plane, all lie along the same line through the origin (**Figure 2a**) or have non-zero *relative* angles (**Figure 2b**). The absolute angle is a matter of convention only. In the former case, these vectors can be rotated to the electric axis, and we say — only as a matter of convention — that every particle has an electric charge but no magnetic charge.

How about a situation such as that shown in **Figure 3**: Most particles (e.g., electrons e, protons p) are said to have only electric charges, while there is possibly some particle X with both electric and magnetic charge, say (q, g). Then we can think of X as the sum of two particles, one with pure q and one with pure g. As far as the analysis in the rest of this Chapter is concerned, this "breakup" does not matter. So henceforth we shall study the possibility of a pure monopole g in the presence of other particles (e.g., electrons) with pure charges q.

2.4 Summary

In classical EM, monopoles are allowed, and both Maxwell's equations and the Lorentz force law exhibit symmetry under duality transformations. Which direction in the duality plane is called electric and which is called magnetic is a matter of convention and not physics. When we ask whether monopoles can (or does) exist, we do so under the convention that the usual particles are purely electric.

3 Simplistic quantum treatment

The situation is quite different in quantum mechanics (QM). To be specific, the rest of this Chapter investigates only the case of a point monopole g situated at the origin, and a point charge, say an electron with $q = -e$, moving around it. For simplicity, consider only time-independent situations. In that case, one has to solve the equation

$$H\,\Psi \quad = \quad E\,\Psi$$

The Hamiltonian H involves the vector potential $\mathbf{A}$, and the question is whether $\mathbf{A}$ can be constructed.

3.1 Differential treatment

The simplest objection is that since

$$\boldsymbol{\nabla} \cdot \mathbf{B} \quad \neq \quad 0 \tag{10}$$

then it is impossible to define a vector function $\mathbf{A}$ such that

$$\mathbf{B} = \nabla \times \mathbf{A}$$

If **A** cannot be constructed, then QM falls apart. The conclusion seems to be (if we believe in QM) that monopoles are not allowed.

But one might object: the difficulty (10) happens at only one point. If we can somehow exclude that point, can the difficulty be circumvented?

3.2 Integral treatment

The integral treatment will show that the trouble is not limited to one point.

Evaluating the potential
A point monopole g is situated at the origin (**Figure 4a**); the magnetic field **B** is radial with magnitude given by (4). Take **A** to be along the azimuthal direction:

$$\mathbf{A} = A_\phi(r,\theta)\,\hat{\mathbf{e}}_\phi \tag{11}$$

Consider a spherical surface of radius r and a loop γ at polar angle θ (i.e., a line of latitude if the surface is regarded as the earth); see **Figure 4b**. Integrate **A** around the loop:

$$\oint_\gamma \mathbf{A}\cdot d\boldsymbol{\ell} = \int \mathbf{B}\cdot\hat{\mathbf{n}}\,dS \tag{12}$$

By axial symmetry, A_ϕ is independent of ϕ, so

$$\text{LHS} = A_\phi \cdot 2\pi(r\sin\theta)$$

where $r\sin\theta$ is the radius of the loop. On the RHS of (12), the integral is to be evaluated over a surface with the loop γ as the boundary, which we *choose* to be the "northern cap" with polar angle θ' in the range $0 \le \theta' \le \theta$, i.e., the surface shaded in **Figure 4b**. The usual right-hand rule on γ defines the normal for this surface to be outwards. Thus

$$\begin{aligned}\text{RHS} = B\int_0^\theta r^2\,2\pi\sin\theta' d\theta' \\ = \quad 2\pi B r^2(1-\cos\theta) = \frac{\mu_0 c g}{2}(1-\cos\theta)\end{aligned}$$

where in the last step we have used (4) for B. Thus

$$A_\phi = \frac{\mu_0 c g}{4\pi}\cdot\frac{1}{r}\cdot\frac{1-\cos\theta}{\sin\theta} \tag{13}$$

The singularity
The denominator $\sin\theta$ vanishes at $\theta = 0$ and at $\theta = \pi$. The former does not lead to any problems, because the numerator also vanishes (in fact, has a second-order zero). But the latter is indeed a singularity everywhere along the negative z axis. In cylindrical coordinates (ρ,ϕ,z), near the negative z axis

$$\begin{aligned} r\sin\theta &= \rho \\ 1-\cos\theta &\approx 2 \\ A_\phi &\approx \frac{\mu_0 c g}{2\pi\rho} \end{aligned} \tag{14}$$

obviously singular everywhere on the negative z axis (i.e., as $\rho \to 0$). The result is obvious: the numerator is the total flux from the monopole, nearly all captured when the loop γ is near $\theta = \pi$ (**Figure 4c**); the denominator is the length of the loop γ.

Therefore, the attempt to construct **A** runs into trouble not just at the monopole, but everywhere along *a string of singularity* which starts from the monopole and goes to infinity (along the negative z axis). Thus, staying away from one point is not good enough. The QM formalism seems to collapse, and the apparent conclusion is that monopoles are not allowed — they would lead to singularities in the potential. That as the belief until Dirac's 1931 paper [1], and seemed to offer an explanation why monopoles had never been found.

Problem 5
Suppose we had chosen the "southern cap" in the above evaluation, with polar angle θ' in the range $\theta \le \theta' \le \pi$, i.e., the surface shaded in **Figure 4c**. What would **A** be and where would it be singular?
§

4 Dirac treatment

Dirac's insight can be discussed in three parts:
- The line of singularity can be placed elsewhere. Moving it corresponds to a gauge transformation.
- Therefore there cannot be a genuine problem along any line.
- Analysis of the gauge transformation and the associated singular behavior then leads to the quantization condition.

4.1 Moving the singularity

In contrast to **Figure 4**, we could have analyzed the problem with a z' axis that is inclined at some angle, as in **Figure 5**. The line of singularity would then appear along the z' axis. In this manner, the line of singularity can be *moved*. Problem 5 is a case in which the line of singularity is moved 180 degrees.

In fact the line of singularity need not be straight. Think of performing the analysis in the last Section separately for each spherical surface of radius

r. There is one *point* on the surface where $\mathbf{A}$ is singular. This point can be placed anywhere on the surface by a choice of the z' axis. Of course these points must change continuously as r is changed. So the collection of all such points would lead to the following (**Figure 6**):

The line of singularity can be any curve from the origin to infinity. This line can be arbitrarily moved.

For each choice of the line of singularity, $\mathbf{A}$ is different. Yet all these different $\mathbf{A}$'s describe the same $\mathbf{B}$. Thus

Different choices of the line of singularity are related by gauge transformations.

We shall not write out these transformations; they are quite messy in the general case.

4.2 No real problem?

Suppose we choose $\mathbf{A}$ as in (13), and there is an apparent singularity along the negative z axis. The electron wavefunction Ψ would be singular there as well, and it would seem that this would affect measurable quantities. But one can argue: since this line of singularity can be moved to another place, perhaps there is no "genuine" problem. That is the insight of Dirac [1]. But then the wavefunction has to be solved not just locally, but must be patched together for a global solution. Dirac proved there is no problem provided the quantization condition is satisfied. The proof is technically difficult and subtle, and will not be presented here. Instead, the next Section outlines the Wu–Yang approach, which develops the Dirac insight more elegantly, avoiding all singularities.

5 Wu–Yang treatment

5.1 Key idea

The Wu–Yang method has one additional insight that goes beyond Dirac. The idea can be explained by the analogy of maps. A map (in the geographical sense) is a mapping f (in the mathematical sense) from the surface of the earth (say S) to 2D Euclidean plane P (i.e., a flat piece of paper, say a page on the atlas):

$$f : \qquad S \to P$$

But it is impossible to have a single non-singular mapping for the whole surface of the earth. The solution is the following.

- Define subsets $S_1, \ldots, S_N$ of S such that

$$\bigcup_j S_j \quad = \quad S$$

  In other words, they cover the whole surface of the earth.
- For each of these there is a mapping in the conventional sense:

$$f_j : \qquad S_j \to P$$

  This is what we see on each page of an atlas.
- But in addition, in the overlap regions $S_j \bigcap S_k$, the two mappings f_j and f_k are consistent in a suitable sense. (For example, if the distance between two points is determined by reference to either of these mappings, the answers must be the same.)

The idea of a function is now replaced by such a "patchwork" which is everywhere non-singular.

It turns out that this idea applied to the the vector potential $\mathbf{A}$ solves the problem in a very nice way.

5.2 Dividing into two patches

Again consider a point monopole g at the origin, and one spherical surface of radius r (which should later be regarded as merged together into all space). Divide the surface into two patches or regions (**Figure 7**):

$$\begin{aligned} \text{Region } a &= \text{northern hemisphere plus a bit} \\ \text{Region } b &= \text{southern hemisphere plus a bit} \end{aligned}$$

They overlap in a circular band around the equator.

Now apply the integral treatment as in Section 3.2 to each region.

Northern cap
Repeat the argument in Section 3.2. There is now no choice: the surface has to be chosen to be the "northern cap", and the result is, from (13)

$$\mathbf{A}^a \quad = \quad \frac{\mu_0 cg}{4\pi} \cdot \frac{1}{r} \cdot \frac{1 - \cos\theta}{\sin\theta}\, \hat{\mathbf{e}}_\phi \qquad (15)$$

Since this is to be used only in region a, it never encounters any singularity (which occurs at $\theta = \pi$).

Southern cap
Repeat the same argument for region b. Now the surface has to be chosen to be the "southern cap", and the result is, from Problem 5:

$$\mathbf{A}^b \quad = \quad \frac{\mu_0 cg}{4\pi} \cdot \frac{1}{r} \cdot \frac{-1 - \cos\theta}{\sin\theta}\, \hat{\mathbf{e}}_\phi \qquad (16)$$

The overall sign is negative, because the positive direction along γ translates into the inward normal on the surface. Since (16) is to be used only in the region b, it never encounters any singularity (which occurs at $\theta = 0$).

5.3 Consistency in overlapping region

Difference is gauge transformation
In the overlapping region, both formulas apply. Thus we must have

$$\mathbf{A}^a - \mathbf{A}^b \quad = \quad \boldsymbol{\nabla}\Lambda$$

for some gauge function Λ.

Loop integrals
Usually Λ has to be well defined as a function of position, which would require

$$\oint \boldsymbol{\nabla}\Lambda \cdot d\mathbf{x} \quad = \quad 0$$

or the corresponding integrals of $\mathbf{A}^a$ and $\mathbf{A}^b$ to be the same. In other words, and generalizing to 4D spacetime, the usual requirement is that the closed loop integrals

$$\boxed{A[\gamma] \equiv \oint_\gamma A_\mu \, dx^\mu} \tag{17}$$

are unchanged under gauge transformations; these are the physical quantities. This would be the usual statement as to how much of the potential is needed — that we only need $A[\gamma]$ for all γ. If these are all zero, then there is no physical effect — it is a pure gauge.

Region that is simply connected
If the relevant domain is simply connected (in the sense that every closed loop γ can be continuously shrunk to a point), then a closed loop γ can be regarded as the sum of infinitesimal loops γ_j. For an infinitesimal loop, $A[\gamma_j]$ can be expressed in terms of the value of $\partial^\mu A^\nu - \partial^\nu A^\mu$ inside the loop. So the physical quantity is reduced back to $\mathbf{B}$ and $\mathbf{E}$. We have simply come full circle in the argument.

Region that is not simply connected
The situation however is different if the region is not simply connected. The present overlapping region, a band, is an example. For simplicity, let us specialize to the equator $\theta = \pi/2$; taking the difference between (15) and (16), we get

$$\boldsymbol{\nabla}\Lambda \quad = \quad \frac{\mu_0 c g}{4\pi} \cdot \frac{2}{r}\,\hat{\mathbf{e}}_\phi \tag{18}$$

Now we come to a subtlety. As far as the EM fields are concerned, we only need $\boldsymbol{\nabla}\Lambda$ (which is something with zero curl); we do not really need Λ itself. But the full gauge transformation includes changing the phase of the wavefunction by Θ, where

$$\Theta \quad = \quad -\frac{e}{\hbar}\Lambda$$

where it is assumed that $q = -e$ to be precise. Thus

$$\boldsymbol{\nabla}\Theta \quad = \quad -\frac{\mu_0 c g e}{4\pi\hbar} \cdot \frac{2}{r}\,\hat{\mathbf{e}}_\phi \tag{19}$$

We need Θ, not just its gradient. For Θ to be well defined, the integral of (19) around a close loop must be zero; the integral around the equator is (up to an irrelevant sign)

$$I \quad = \quad \mu_0 c \, \frac{eg}{\hbar}$$

It would seem that QM requires Θ to be uniquely defined, and therefore $I = 0$ — and this would forbid any nonzero g.

There is an escape!
But there is an escape. We do not really need Θ; we only need $e^{i\Theta}$. This means Θ can be ambiguous by a multiple of 2π. Thus a consistent quantum theory can be constructed provided $I = 2n\pi$ for any integer n. A little arithmetic then leads to (1), with

$$\boxed{K = \frac{2\pi\hbar}{\mu_0 c}} \tag{20}$$

In terms of fine structure constant
The condition can be expressed more intuitively via the fine structure constant

$$\alpha \quad \equiv \quad \frac{e^2}{4\pi\epsilon_0 \,\hbar c} \approx \frac{1}{137}$$

Since this is dimensionless, its value is the same in all systems of units.

Problem 6
(a) Check that α as defined is dimensionless and check its numerical value.
(b) For a hydrogen atom, express the binding energy E_B in terms of mc^2 and factors of α, where m is the (reduced) mass of the electron. §

Problem 7
Show that the quantization condition can be written as

$$\boxed{\frac{g}{e} = \frac{1}{2\alpha}\, n} \tag{21}$$

For our choice of units, in which monopoles are measured in the same units as charges, the fundamental monopole strength is ~ 68 times the fundamental electric charge. §

Quantization of charge
If a single monopole g exists, then charges would be quantized — thereby explaining what would otherwise be a mystery. That is certainly an attractive reason for supposing that monopoles would actually exist.

5.4 To what extent is vector potential needed?

We can phrase the different levels of understanding in terms of the answer to the question:

How much of the potential A_μ is really needed?

- In classical physics A_μ is not really needed; it is merely a calculational convenience.
- The naive treatment of QM would say that A_μ is needed, say in the Hamiltonian.
- A more careful analysis would provide a weaker condition, that only $A[\gamma]$ is needed for every closed loop γ.
- Finally, in the above analysis, we see that we only need $A[\gamma]$ (mod h/e). (Check that this follows from I being $2n\pi$.)

This constitutes a precise statement of exactly what aspects of the potential is needed.

6 Head-on collisions

This Section introduces the analysis by Lai and Young [4] of the head-on collision of a point charge and a point monopole, leading to another derivation of the quantization condition — without having to mention the potential.

6.1 Defining the situation

A point charge e is situated at the origin (say constrained so that it cannot undergo translational motion) and a point monopole g approaches it head-on,[5] along the z axis (**Figure 8**). To be definite, suppose $e > 0$, $g > 0$. This scattering event has two interesting features, which are indeed superficially conflicting, i.e., a paradox.

- The charge produces only an electric field $\mathbf{E}$, which is radially outwards. The monopole moves parallel to $\mathbf{E}$, and therefore is never subject to any force. By Newton's third law, the monopole also does not exert any force on the charge. So nothing happens — the monopole travels in a straight line with constant velocity, and passes through the charge.
- Examine the fields at an arbitrary point P before the monopole hits the charge (**Figure 8a**). The field directions are shown. There is a momentum density in the direction $\mathbf{E} \times \mathbf{B}$. It is easy to see that there is an angular momentum in the $-z$ direction. On the other hand, the situation after the monopole has passed through the charge is shown in **Figure 8b**; there is an angular momentum along $+z$. So the scattering causes a change of angular momentum.

One way to analyze the situation is to write down the momentum density in the two situations, and hence obtain the angular momentum density. Integrating this over all space gives the total angular momentum before and after the scattering, which then allows us to proceed to the condition (22) below. This calculation would be quite complicated.

6.2 Regularization

The paradox occurs because we have made the idealization of two *point* particles, a singular limit. The usual method is to solve the problem first in a non-singular case, and then take the limit. For this purpose, first replace the point charge e by a distribution over a small sphere, say of radius a; we shall later let $a \to 0$. The spherical distribution of charge can be regarded as the sum of many rings perpendicular to the z direction. **Figure 9** shows one such ring with charge q (a part of the total charge e), and radius $R \le a$, through which the monopole passes. Now the monopole and the charge always stay a finite distance apart, and there should be no singularity.

As the monopole approaches, the magnetic flux φ through the ring changes, leading to an EMF $\mathcal{E}$ of magnitude[6]

$$\mathcal{E} \quad = \quad \frac{d\varphi}{dt}$$

in which Faraday's law has been used. But the EMF is the line integral of $\mathbf{E} = E\,\hat{\mathbf{e}}_\phi$, so

$$E \quad = \quad \frac{1}{2\pi R}\frac{d\varphi}{dt}$$

This causes a force $F = qE$ on the ring, and a torque $\tau = RF = qRE$, causing an increase in the angular momentum j of the ring (in the z direction)

$$\frac{dj_z}{dt} = \tau \quad = \quad qRE = \frac{q}{2\pi}\frac{d\varphi}{dt}$$

The total change of angular momentum (comparing the initial state with the monopole at $z = -\infty$

[5] It is also possible to consider the reverse situation, with a stationary monopole and a moving charge. We have chosen the present scenario so that only the more familiar Faraday's law need to be invoked, rather than its dual analog.

[6] In this Section, we can afford to be sloppy about signs.

and the final state with the monopole at $z = +\infty$) is then

$$\Delta j_z = \frac{q}{2\pi}\varphi$$

where φ is the total flux that has passed through, namely the flux emanating from the monopole:

$$\varphi = \mu_0 gc$$

Thus

$$\Delta j_z = \frac{\mu_0 gc}{2\pi} \cdot q$$

This is for one ring. Notice it depends on the particular ring only through its charge q. Add this up for all the rings to obtain the analogous expression for the angular momentum J of the whole charge distribution;[7] this simply involves changing $q \mapsto e$ on the RHS:

$$\Delta J_z = \frac{\mu_0 gc}{2\pi} \cdot e \qquad (22)$$

Up to this point, the analysis is purely classical.

Now imagine that the charge e is a quantum mechanical particle. Then its angular momentum component must be quantized in steps of $\hbar$, so $\Delta J_z = n\hbar$ for an integer n. Therefore

$$\frac{\mu_0 gc}{2\pi} \cdot e = n\hbar \qquad (23)$$

which gives the same quantization condition as before.

Being careful about the flux

Care is needed in applying Faraday's law. There are two related issues.

- The flux is evaluated for any surface that has the ring of radius R as its boundary. For example, we can choose either the flat surface bounded by the ring (**Figure 10a**) or a long "sock" extending to $z \to +\infty$ (**Figure 10b**).
- When there are magnetic monopoles, Faraday's law is modified: there is an extra term due to $\mathbf{J}_m$. For example, if we choose the flat surface in **Figure 10a**, there is a sudden and large magnetic current (in fact a δ-function in time) when the monopole passes the surface.

The easiest way to deal with these subtleties is to choose the long "sock" as the surface. Then when when $z \to -\infty$ (**Figure 11a**), the flux is $\varphi = 0$; when $z \to +\infty$ but still to the left of this surface (**Figure 11b**), $\varphi = \mu_0 gc$, i.e., capturing all the flux from the monopole. Over this entire time interval, there is no magnetic current through the surface.

[7] This J should not be confused with the current.

Our earlier derivation should be understood in this manner.

Problem 8

(a) Give a general argument as to why the derivation using any surface with the same boundary, i.e., a ring of radius R, would give the same result.

(b) Specifically, choose the flat surface S bounded by the ring. As the monopole moves along the z axis, sketch as functions of z (i) the magnetic flux φ through S and (ii) the total magnetic charge that has passed through S. Discuss these in the light of your answer to (a). §

7 Experimental evidence

At the beginning of this Chapter, the argument was advanced that questions A and B are logically different. Although it has been shown that monopoles (of definite strengths) *can* exist, it does not imply that they *must* exist. Many searches over the years have not found a monopole, and a summary of the current status (as of 2015) can be found in Ref. [6]. Note that with the quantization condition, the searches can target particular values of g. There was one intriguing event of the right magnitude detected in the 1980s [7], but the author was careful to call it a "candidate event" and not a discovery; that event is now generally discounted. As of now, very stringent limits have been set on the density of monopoles in matter, or their flux in cosmic rays, or their production in high energy experiments.

The search for monopoles has taken on more importance for two reasons.

- There is a general belief that in high energy collisions, anything that can be produced (i.e., not forbidden by any principle) will be produced, provided there is enough energy. To that extent, our earlier distinction between questions A and B becomes somewhat blurred.
- In certain theories such as Grand Unified Theories (GUT) or string theories, monopoles are actually predicted to exist. But these theories also predict very large masses.

Therefore the experimental search is likely to go on for many years to come, and the general belief is that monopoles will be found one day.

References

[1] PAM Dirac, "Quantised singularities in the electromagnetic field", Proc. Roy. Soc. **A 133**, 60 (1931).
DOI: 10.1098/rspa.1931.0130

[2] TT Wu and CN Yang, "Concept of nonintegrable phase factors and global formulation of gauge fields", Phys. Rev. **D 12**, 3845 (1975). DOI: 10.1103/PhysRevD.12.3845

[3] TT Wu and CN Yang, "Dirac's monopole without strings: classical Lagrangian theory", Phys. Rev. **D 14**, 437 (1976). DOI: 10.1103/PhysRevD.14.437

[4] HM Lai and K Young, "Classical head-on collisions between charges and monopoles", Phys. Rev. **D 28**, 1552 (1983). DOI: 10.1103/PhysRevD.28.1552

[5] PT Leung, "Magnetic monopole and Poynting's theorem", Euro. J. Phys., **16**, 43 (1995).

[6] L Patrizii1 and M Spurio, "Status of searches for magnetic monopoles", Ann. Rev. Nucl. and Part. Science, **65**, 279 (2015). DOI: 10.1146/annurev-nucl-102014-022137 The preprint version is more accessible: `arXiv:1510.07125 [hep-ex]`

[7] B Cabrera, "First results from a superconductive detector for moving magnetic monopoles", Phys. Rev. Lett. **48**, 1378 (1982). DOI: 10.1103/PhysRevLett.48.1378

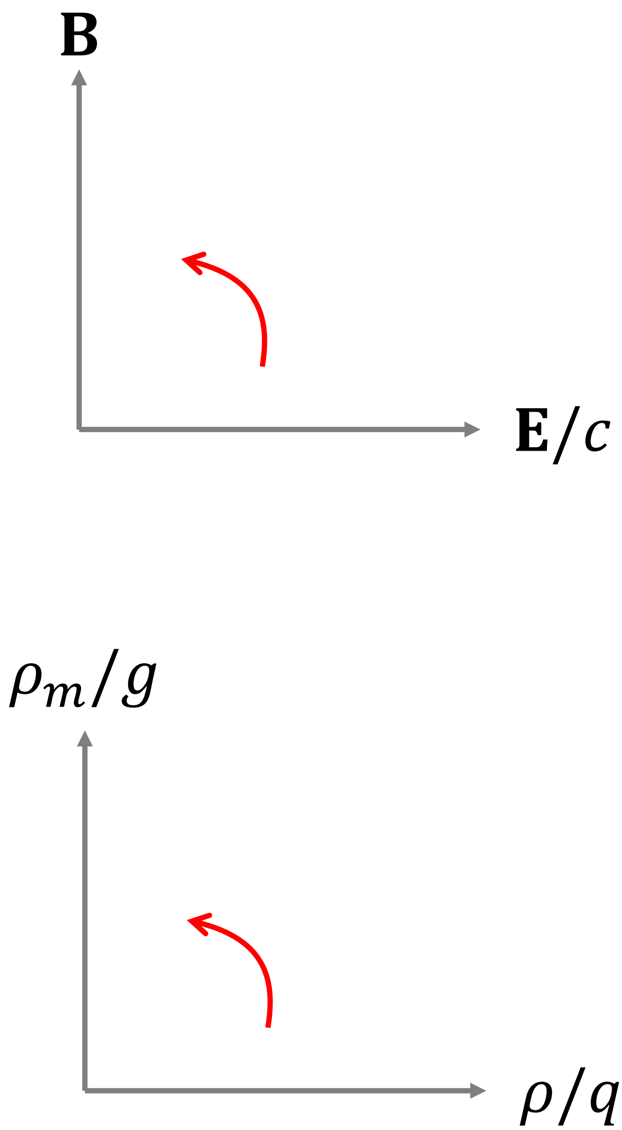
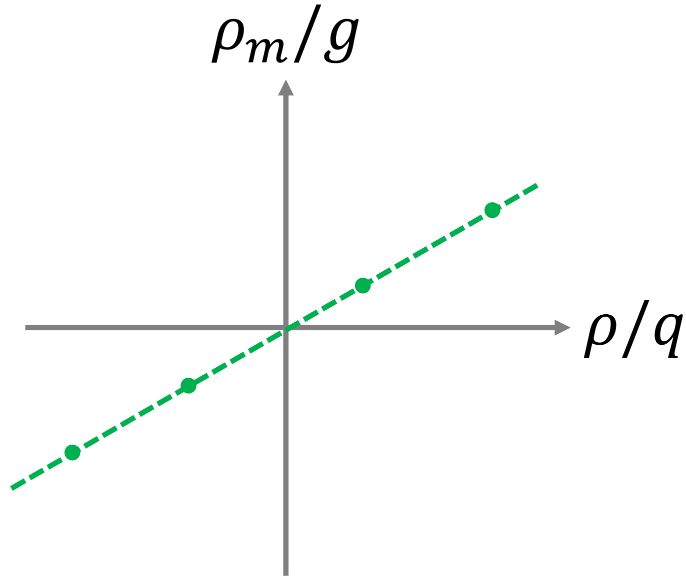


Figure 1 Duality is a rotation in the electric-magnetic plane.

(a)



(b)

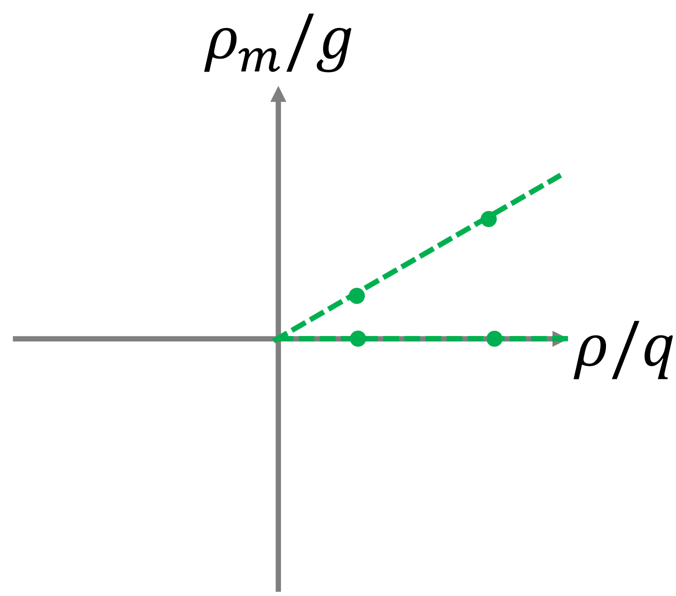


Figure 2 (a) A system of particles with electric charges q_j and magnetic charges g_j having a constant ratio. Such a system is the same as one with pure charges only.

(b) If different particles have g_j/q_j with different values, then there is a nontrivial monopole effect.

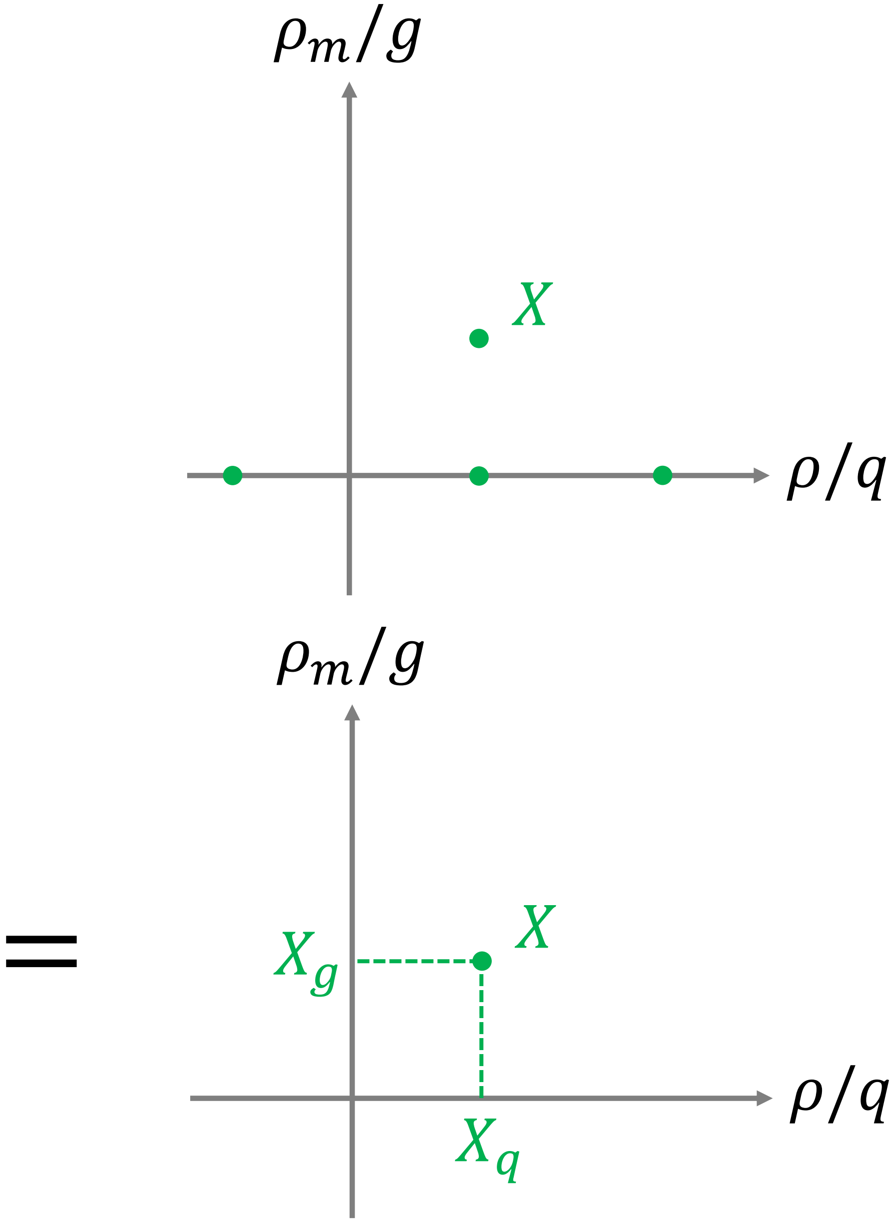


Figure 3 A particle X with both electric charge and magnetic charge (in the convention where normal particles have only electric charges) can be broken up in a pure electric charge X_q (which is trivial) and a pure magnetic charge X_g. The question is then whether the latter can exist.

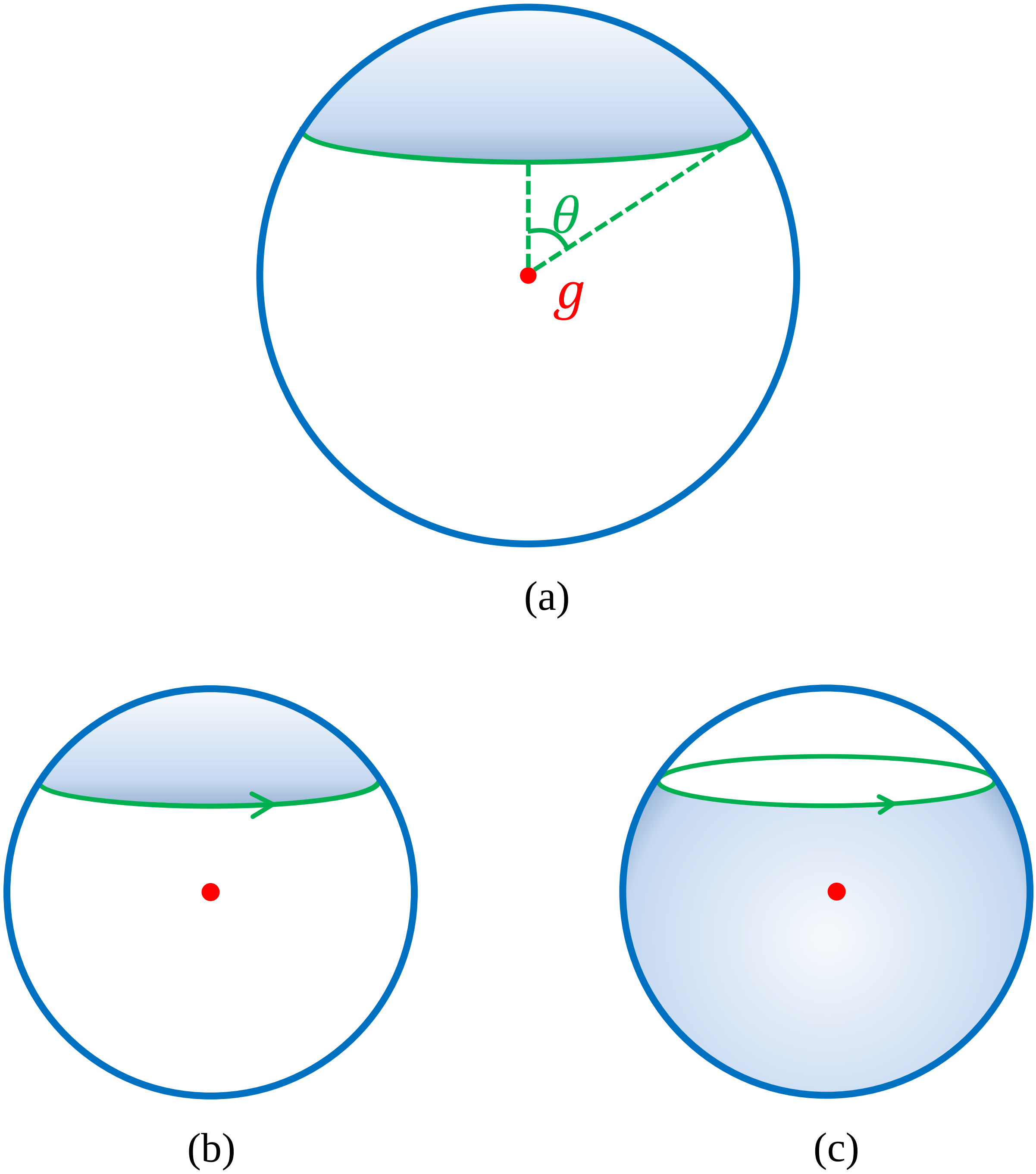


Figure 4 (a) A surface of radius r around a monopole g, and a loop at polar angle θ.
(b) One choice of surface for the loop.
(c) Another choice of surface for the loop

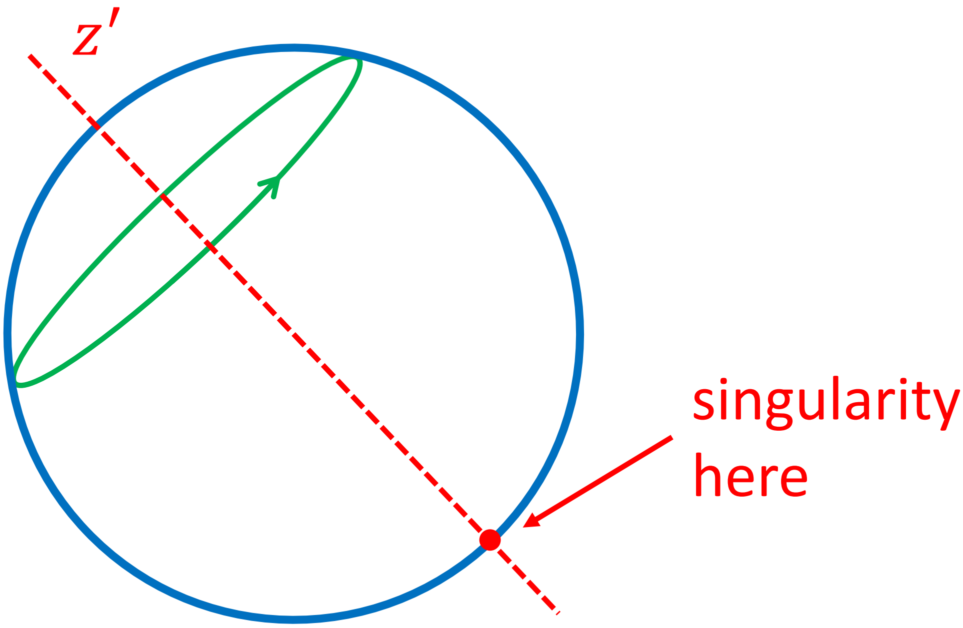


Figure 5 Choose a z' axis in an arbitrary direction.
Singularity occurs on the negative z' axis.

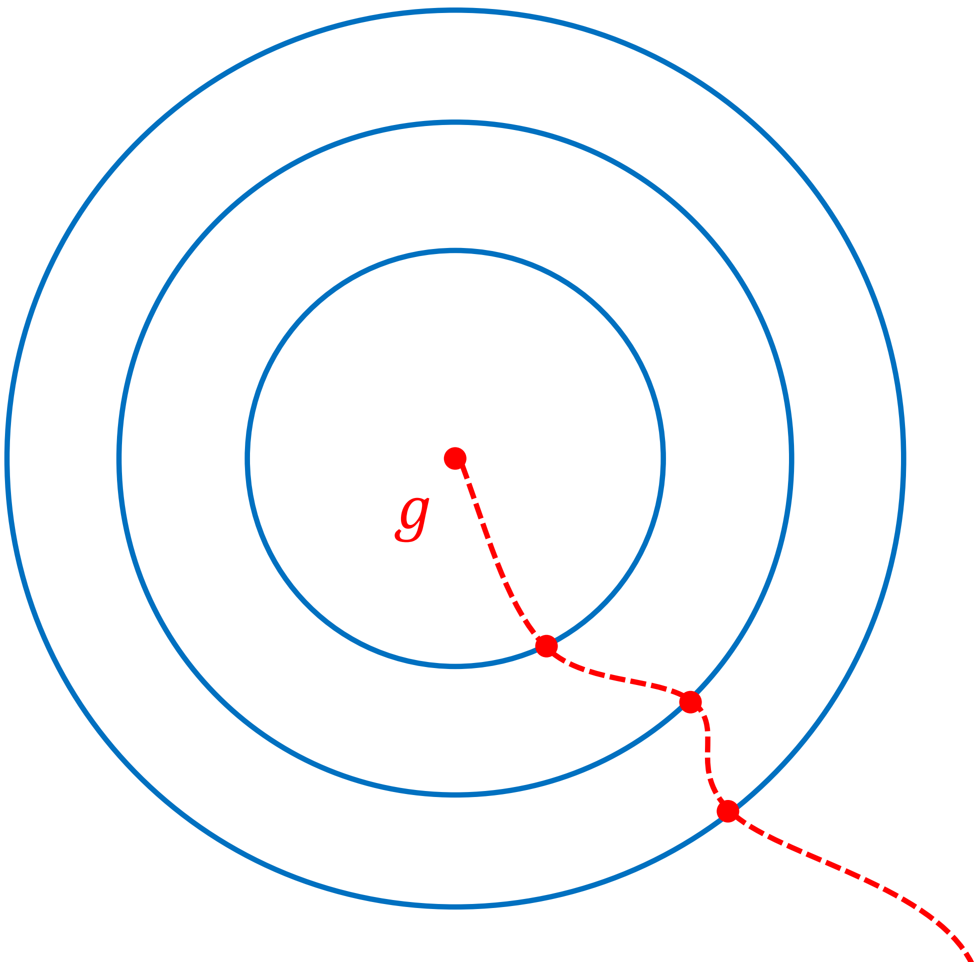


Figure 6 There is a point of singularity on each spherical surface, together forming a line of singularity from the monopole to infinity.

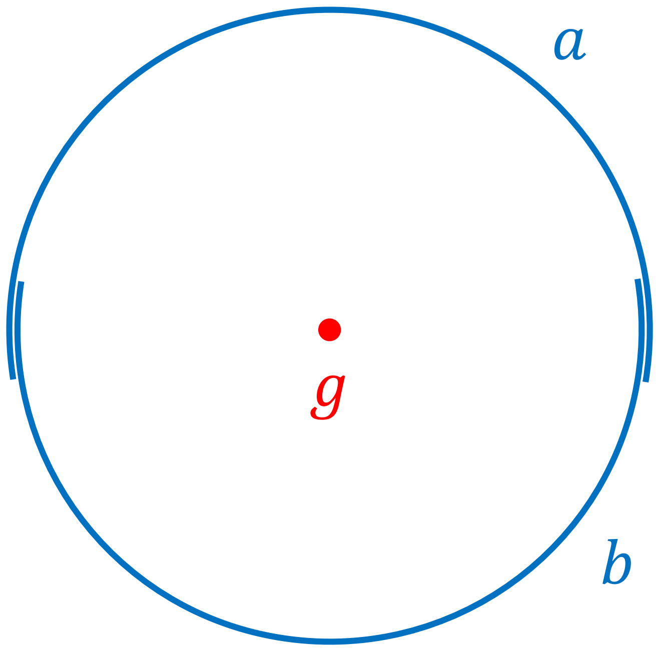


Figure 7 Define two regions on a spherical surface:
a is the northern hemisphere plus a bit;
b is the southern hemisphere plus a bit;
the intersection is a band around the equator.

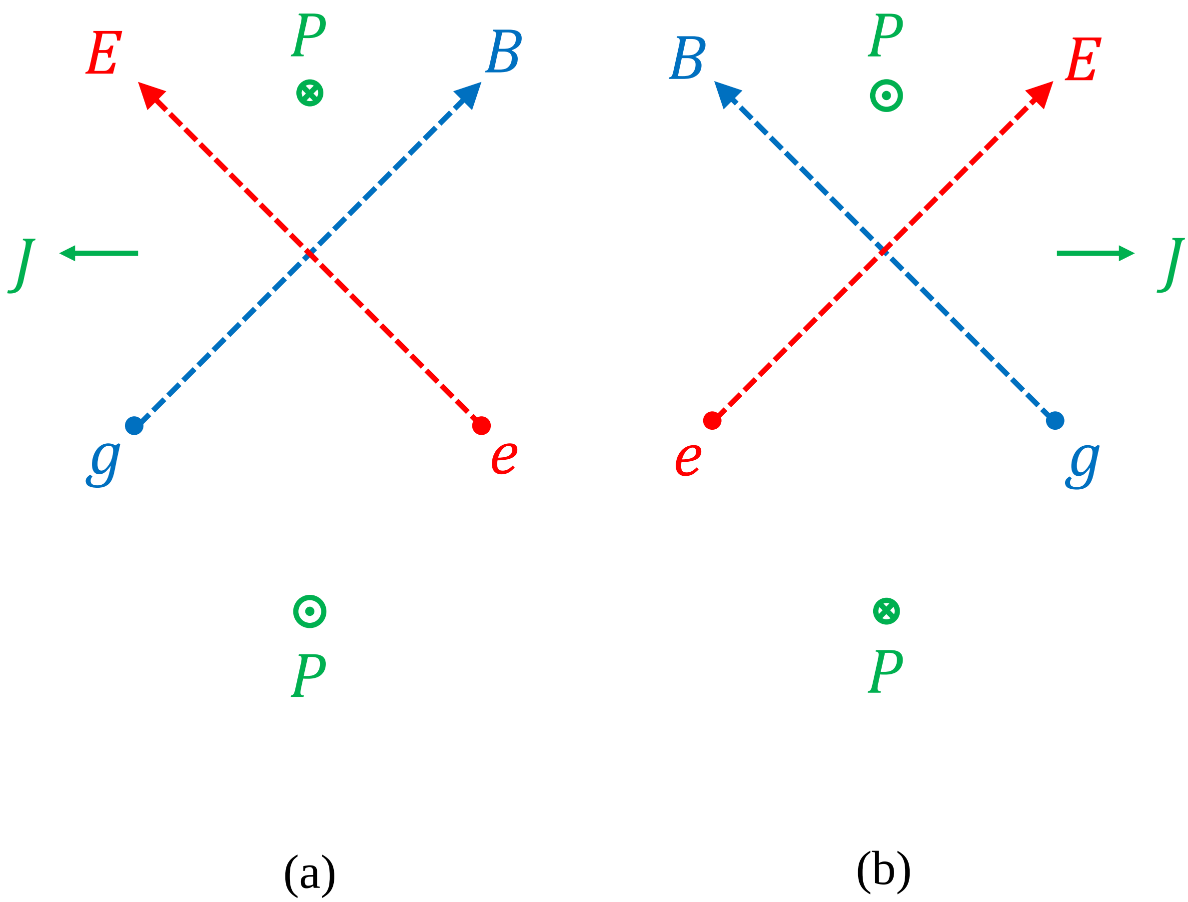


Figure 8 (a) The monopole g to the left of e. There is a momentum density P as shown, leading to an angular momentum density J to the left.

(b) The monopole g to the right of e. There is a momentum density P as shown, leading to an angular momentum density J to the right.

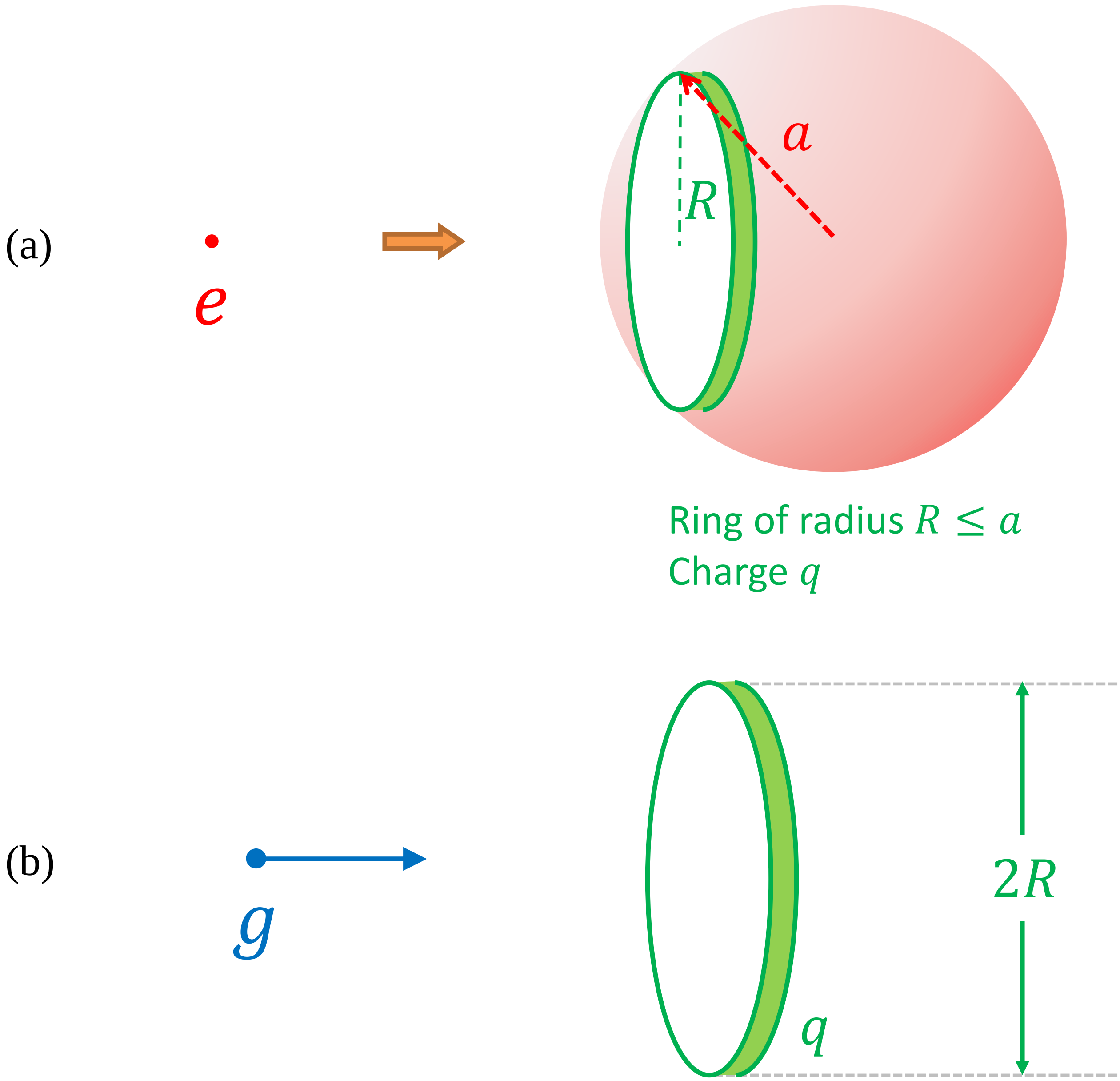


Figure 9 (a) A point charge e replaced by a spherical distribution of radius a, which is cut into rings of radius R.

(b) A monopole g approaches the ring of radius R. The flux through the ring changes, leading to a tangential electric field and an EMF.

(a)

(b)

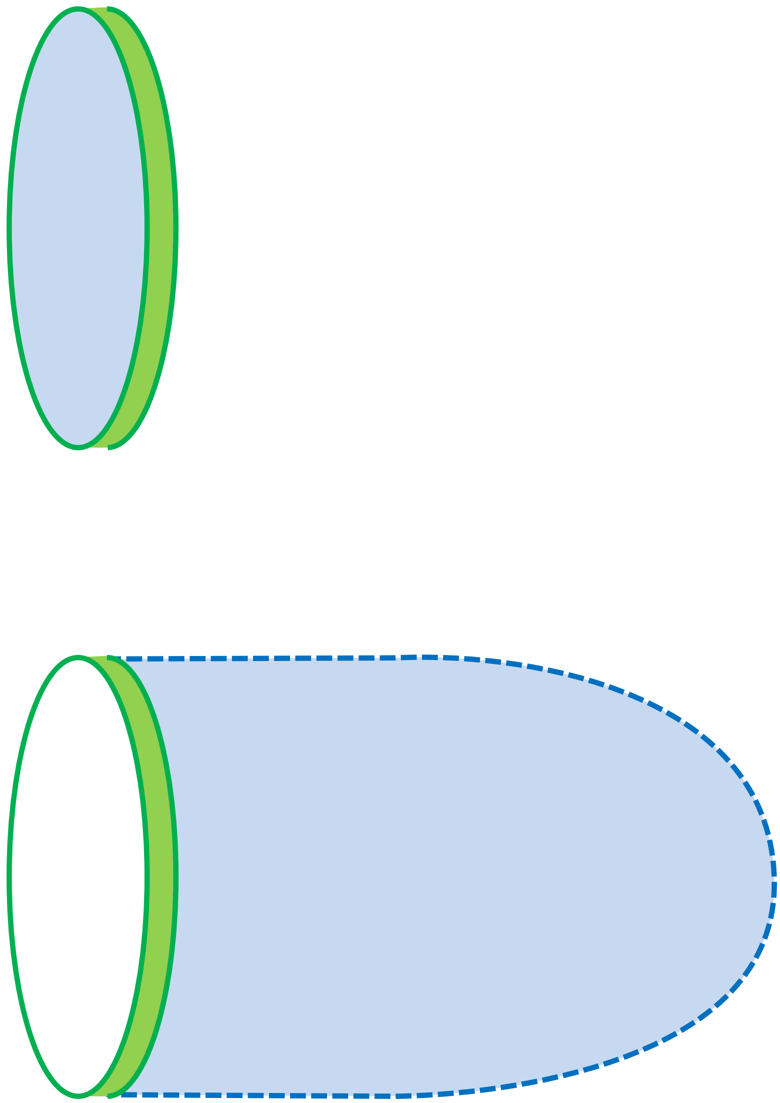


Figure 10 For the loop defined by the ring, there are many ways of associating a surface.

(a) The planar surface

(b) A long "sock" extended to the right

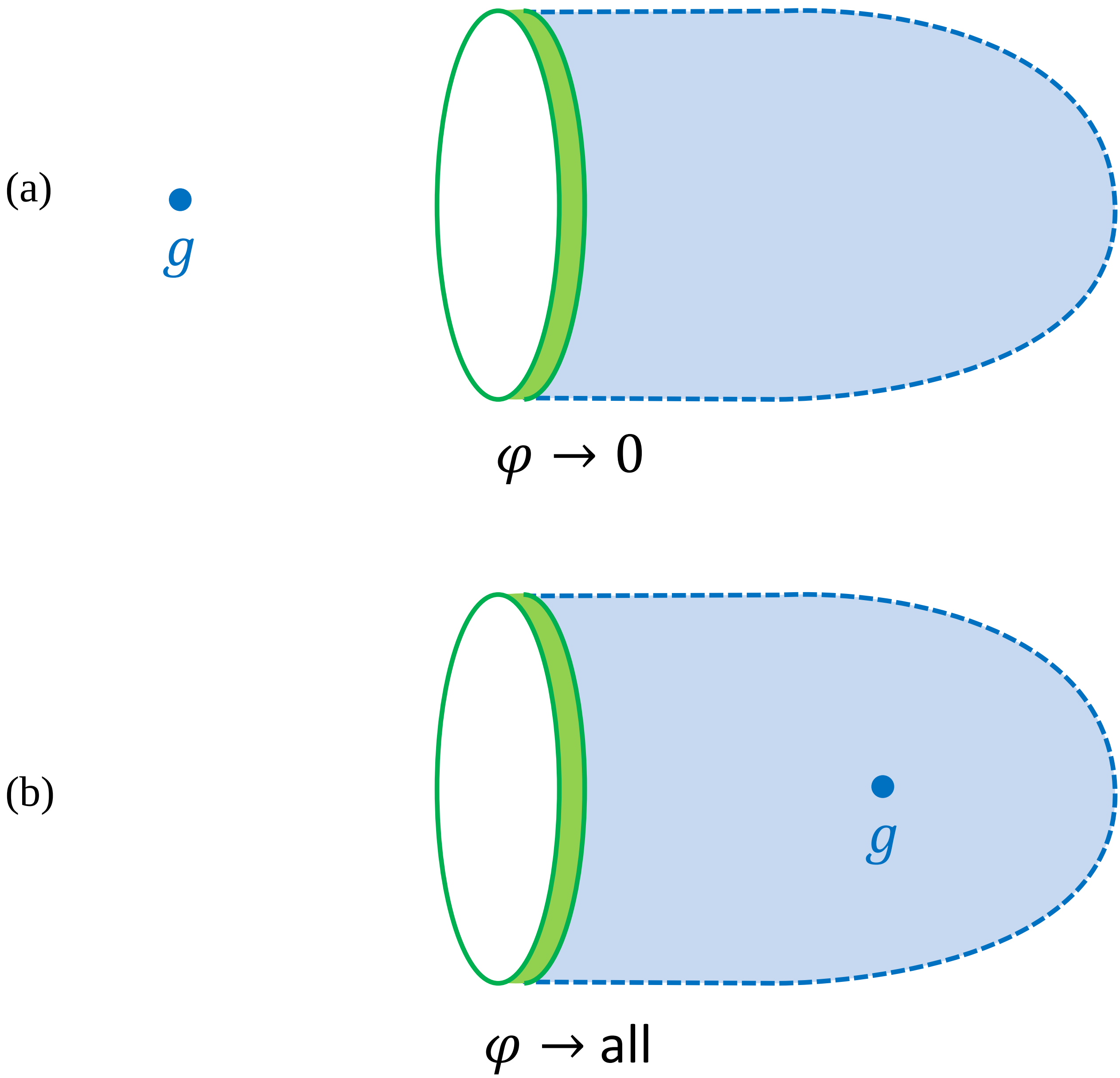


Figure 11 Using the surface defined by a long “sock”.

(a) When g is far to the left of the ring, the flux through the “sock” is nearly zero.

(b) When g is far to the right of the ring, the flux is nearly the total flux emanating from the monopole.