Gauge invariance and hidden symmetries

K. Moriyasu

Citation: American Journal of Physics **49**, 819 (1981); doi: 10.1119/1.12400
View online: https://doi.org/10.1119/1.12400
View Table of Contents: https://aapt.scitation.org/toc/ajp/49/9
Published by the American Association of Physics Teachers

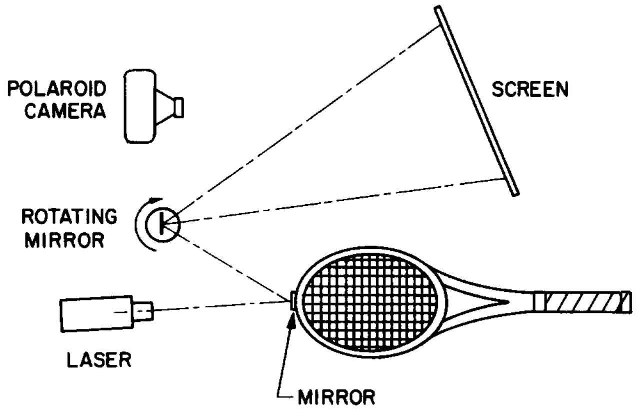

Fig. 6. Alternate method to determine position of racket head as a function of time.

this the resultant output from the liquid potentiometer should be frequency analyzed to get the ratio of the amplitude of the higher mode to the amplitude of the fundamental mode.

The shift in the position of the node was also determined for a single racket when a small mass was taped to the tip. 10 g moved the node approximately 2 cm.

The three definitions of sweet spot all have merit and, in general, the points corresponding to them are located in different places. It is assumed that if the ideal tennis racket could be designed, it would have all three of these points located at the center of the stringed area and have a power region and nodal region covering most of the face of the racket.

ACKNOWLEDGMENTS

I would like to thank Enoch Durbin for several helpful discussions and Jim Hobbs and Oleh Hnatiuk for their help in taking some of the data presented in this paper.

APPENDIX

A second method for observing both the fundamental and higher modes of oscillation with the handle clamped was used and it has the virtue of not requiring an oscilloscope, scope camera, or explanation of why the water beaker system works. (The liquid potentiometer is a pedagogical gem in this author's opinion.) This second method requires a light source (laser), mirror chip glued to the racket tip, rotating mirror, screen, and a camera (Fig. 6). It does require a certain amount of dexterity and coordination to have the ball hit the racket during the time interval when the beam of light is traversing the screen, but with some practice it was possible to get a good picture most of the time.

[1]H. Head, U.S. Patent 3999756, 28 December 1976.
[2]H. Brody, Am. J. Phys. **47**(6), 482 (1979).
[3]F. R. Lacoste, U.S. Patent 3941380, 2 March 1976.
[4]K. Hedrick, R. Ramnath, and B. Mikic, World Tennis **27**(4), 78 (1979).
[5]E. Durbin, patent pending and private communication.

Gauge invariance and hidden symmetries

K. Moriyasu
Visual Techniques Laboratory, Department of Physics, University of Washington, Seattle, Washington 98195
(Received 10 March 1980; accepted 10 December 1980)

We present a simple introduction to the basic physical ideas involved in the flux quantization number carried by systems such as magnetic monopoles and vortices. By using simple geometrical arguments, we see that the flux number is associated with a hidden symmetry in gauge theory that is topological in origin.

I. INTRODUCTION

Gauge invariance has emerged as one of the most important fundamental developments of modern physics. The successful unification of electromagnetism and the weak force by the Weinberg–Salam theory[1] has raised the principle of local gauge invariance to a level of significance equal to that of relativity. At the same time, gauge theory has provided new insight into some unexplained problems that predate gauge theory by many years. One of these problems is the existence of so-called superselection rules for special quantum numbers like the lepton number. Such quantum numbers appear to obey simple additive conservation rules that are not associated with any known symmetry. The origin of the lepton number is still a mystery but recently it has been discovered that a similar type of quantum number exists in a variety of gauge theory phenomena ranging from flux quantization in superconductors to the understanding of how quarks might be confined within hadrons. What is interesting about this new quantum number is that it not only obeys a simple additive superselection rule but it also originates from a "topological" source rather than from a conventional symmetry. Thus it appears that a "hidden" topological symmetry exists in gauge theory that might help us to eventually understand the origin of other superselection rules.

In this paper, we present a simple introduction to the basic ideas involved in the application of topological concepts in gauge theory. Our purpose is to show that it is possible to appreciate at least some of the physical implications of the new topological features of gauge theory without first having to become an expert in topology. We will use the tried and true physicist's approach of starting from a specific familiar physical example, namely, the

 0002-9505/81/090819-08$00.50

phenomenon of flux trapping in superconductors, and use it to teach us most of the topology we need. There is a good precedent for taking this approach since many physicists first learn about the theory of abstract Lie groups by studying the more familiar angular momentum operators in quantum mechanics.

The sections of this paper are organized as follows. In Sec. II, we discuss why the superconductor is an ideal pedagogical device for studying topological effects in gauge theory. We then reinterpret the superconductor in terms of gauge theory language in Sec. III and use an intuitive geometrical picture in Sec. IV to show how the phenomenon of flux trapping contains "hidden" topological properties. In Sec. V, we see how the same topological properties can be uncovered from the canonical Lagrangian description of superconductivity. In Sec. VI it is shown that the rules for adding flux numbers define a new type of symmetry group. In Sec. VII, the topological ideas learned from flux trapping are used to study the Dirac magnetic monopole.

II. WHY STUDY THE SUPERCONDUCTOR?

In many respects, the phenomenon of flux trapping in a superconductor is an ideal pedagogical device for learning about hidden topological concepts in gauge theory. First of all, superconductivity itself is one of the simplest real-life examples of a local gauge theory with spontaneous symmetry breaking. As we shall see later, the Lagrangian for a superconductor even resembles those of simple models in elementary particle physics. At the same time, one has the advantage that a phenomenon like flux trapping can be described and some of its properties can be calculated with only the use of electromagnetism and elementary quantum mechanics. Thus no matter how abstract the topology may become, it can always be related to basic physical concepts.

As described in many texts,[2] flux trapping occurs in type-II superconductors because the magnetic field induces a flow of Cooper-pair current. This current forms a vortex that surrounds the magnetic field lines and confines them to a small region where the conductivity is still normal. The flux is quantized according to the condition

$$\text{Flux} = \oint \mathbf{A}\cdot d\mathbf{x} = 2\pi N\hbar c/q, \qquad (1)$$

where q is the Cooper-pair charge and N is an integer. Equation (1) is derived from the requirement that the Cooper-pair wave function be continuous along any closed path in the superconductor that encircles the flux. The only advanced mathematics required is Stokes's theorem. Thus in this description, there is no overt evidence of any topological complexities.

How then does one even know that there are any physically interesting topological properties to be uncovered in the superconductor? The existence of such properties is suggested by the fact that the choice of the closed path around the trapped flux is completely arbitrary. The quantization condition strongly resembles a contour integral whose value depends only on the residue of a singularity and not on the contour of integration. This resemblance is not entirely accidental. The definition of an analytic function involves a strong connection between the local and global properties of the function over the entire complex plane. A similar kind of global constraint also exists in a system with broken gauge symmetry. We will see in the following discussion how this leads to a relatively simple reinterpretation of flux quantization as a topological condition.

III. GEOMETRICAL SUPERCONDUCTOR

Topology is an area of mathematics that has something to do with the global properties of spaces; it tells us why a doughnut is similar to a coffee cup. Gauge theory, on the other hand, is based on the invariance of physical laws under a local internal symmetry. Thus how can a space with global topological properties be uncovered in the superconductor vortex? To find the answer, we must first reformulate the flux quantization problem in gauge theory language. We will then be able to see that the topological properties we are seeking are those of the internal symmetry space associated with the local gauge group.

The beauty of gauge theory is that it is inherently geometrical. Thus we can use simple geometrical arguments and even draw some diagrams in place of excessively complicated mathematics. Let us first briefly review the geometrical structure of gauge theory[3] and then see how to reinterpret the superconductor as a geometrical system. Gauge theory involves the marriage of physical space with an internal symmetry space at each point. The internal symmetry space of the superconductor is the space of phase factors that transform under the $U(1)$ gauge group of electromagnetism. As shown in Fig. 1, physical space is represented by a horizontal plane and the internal symmetry space by the vertical line [for the one-dimensional $U(1)$ space]. The internal space is called a "fiber" by mathematicians.[4] In this picture, the location of a charged particle is given by a coordinate point in the horizontal plane and the phase of the particle's wave function is indicated by a coordinate point on the vertical "fiber." As the particle moves through physical space, the phase point traces out a path in the internal space above the particle's trajectory. If there is an external gauge potential $\mathbf{A}$ that interacts with the particle, it is interpreted as a geometrical "connection" in the internal symmetry space in much the same way as a connection in general relativity. Thus as the particle moves from $\mathbf{x}$ to $\mathbf{x} + d\mathbf{x}$, the potential rotates the phase of the particle by an amount

$$\delta\theta \simeq \hbar c/q\,\mathbf{A}\cdot d\mathbf{x}, \qquad (2)$$

where q is the charge of the particle. This simple picture can be directly applied to a superconductor in order to obtain a geometrical interpretation. The superconductor is a "self-coherent" system in the sense that it has a built-in relation between the phase values at different points in space. This is due to the large overlap of the Cooper-pair wave functions that produces a long-range correlation that

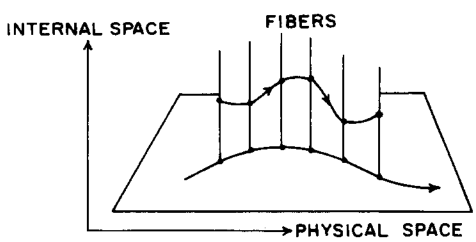

Fig. 1. Path of a charged test particle in the internal symmetry space as it moves through an external gauge field. The same picture also represents a graph of the intrinsic phase relation of the superconductor wave function.

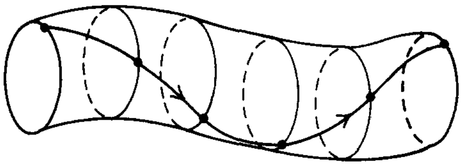

Fig. 2. Picture of the internal symmetry space after it has been wrapped around in a circle so that phase angles of zero and 2π coincide. The phase of a moving particle traces out a path on the surface of a cylinder.

locks the phases together coherently over macroscopic distances. The path in the internal symmetry space in Fig. 1 can thus be interpreted as a graph of the intrinsic phase relation of the superconductor.

When magnetic flux is trapped in the superconductor, a potential conflict arises between the intrinsic phase relation of the superconductor and the magnetic field. The magnetic field tries to rotate the local phase of the Cooper-pair wave function just as it would the phase of a free particle but it cannot because the phases are locked together. This conflict is resolved either by destroying the superconductivity if the magnetic field is too strong or by breaking the gauge symmetry of the magnetic field and forcing it to be consistent with the phase relation of the superconductor. This means that the trapped magnetic field, or more precisely the vector potential field **A**, when acting on a particle, will produce a phase change in the wave function that is the same as the phase relation of the surrounding superconductor.

We now see how to reinterpret the superconductor in terms of gauge theory language. The phase of the wave function is considered to be a geometrical coordinate in the internal symmetry space. At a certain phenomenological level, many of the interesting consequences of superconductivity, such as flux quantization, can be described in terms of the behavior of the phase factor. Since the phase is an internal space coordinate, this means that some of the physical properties of the superconductor can be related to the purely geometrical and topological properties of the internal space. This type of description is very similar to that of general relativity where the effect of the gravitational field is described by motion in a curved space.

IV. TOPOLOGY UNCOVERED

Let us first consider the topology of the internal space without flux trapping since it is easier to visualize. At any fixed position **x** in the superconductor, the phase of the wave function will have a value between 0 and 2π. Continuity of the wave function further requires that the phase values of 0 and 2π must coincide. Thus the internal space at each **x** is equivalent to a one-dimensional circular loop with the phase being a coordinate on the loop. As we move from **x** to an adjacent location $\mathbf{x} + d\mathbf{x}$, the phase must change continuously in accordance with the intrinsic phase relation of the superconductor. The phase coordinate will therefore trace out a path on the surface of a cylinder as shown in Fig. 2. Hence, from the point of view of an "internal observer," the superconductor looks like it has the global topology of a cylindrical space.

When flux is trapped in the superconductor, the topology of the internal space becomes more complicated because the phase relation does not hold in the region where the conductivity is still normal. We can map out the new topology by using the familiar device of a test charge and observing how its phase varies as we move it around the vortex region. How do we know that the test charge phase will really follow the phase of the superconductor and not deviate due to the effect of the trapped magnetic field? The answer is that the gauge symmetry breaking described in Sec. III guarantees that the magnetic field can only produce a phase change that is the same as the phase relation of the superconductor. Thus the test charge will faithfully trace out the convolutions of the superconductor phase.

After we move the test charge around a closed path encircling the vortex region, we see that it must have the same phase value as it did initially due to the continuity of the superconductor wave function. However, the phase of the test charge may have been rotated by the magnetic field through integer multiples of 2π while the charge moved along the path. Thus we conclude that the phase coordinate traces out a path on a surface shaped like a doughnut or torus as shown in Fig. 3. It might be argued that this conclusion is obvious because we have only taken the cylindrical space in Fig. 2 and wrapped it around a closed path and joined the ends together. However, if there is no trapped flux in the center of the torus, the closed path can be made smaller and smaller until there is no torus. The trapped flux makes a hole in the superconductor so that the path of the test charge cannot be arbitrarily shrunk down to a point. The presence of this hole is therefore essential to the topology of the internal symmetry space. This situation is very similar to the Aharonov–Bohm effect[5] where the particle also is not allowed to enter the magnetic field region.

Let us now see how the flux quantization condition, Eq. (1), can be reinterpreted in terms of the topological properties of the internal space. The vector potential **A** generates the phase rotation as given by Eq. (2). The loop integral in Eq. (1) is proportional to the total change in the phase during one complete trip around the vortex. Since the phase can only be rotated by integer multiples of 2π, we see that the flux quantum number N is equal to the number of times that the phase coordinate winds around the torus.

At this point, if we were to communicate our results to a real mathematician, he would tell us that the toroidal internal space we have uncovered is an example of a "multiply connected" space and that we have also just rediscovered a topological quantity known as the "winding number."[6] What is surprising about this result is not the complicated topology uncovered in the superconductor but rather the fact that a physically measurable quantum number N can be reinterpreted as a topological property of a geometrical symmetry space. It is interesting to compare this with general relativity that shows us that the classical gravitational force can be geometrized. Gauge

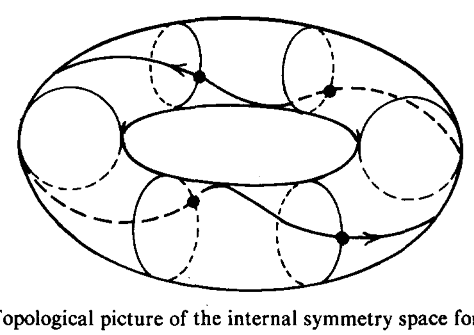

Fig. 3. Topological picture of the internal symmetry space for a closed path around the vortex. The cylinder in Fig. 2 has been wrapped around the trapped flux and the ends joined together, forming a torus.

theory seems to be telling us that certain types of quantum effects can also be interpreted as geometrical, albeit in an unusual type of space.

V. CANONICAL VORTEX

A. Lagrangian

Up to this point, our discussion has relied almost entirely on very intuitive geometrical arguments. In this section, we will compare our geometrical approach with the more formal canonical Lagrangian description of superconductivity. We will not duplicate the treatment already presented in existing publications. Instead, we will use our geometrical ideas as a guide in order to see what sort of arguments are needed to uncover the topological properties within a canonical Lagrangian formalism.

For our purposes, it is sufficient to use the simplest possible model Lagrangian for a charged scalar field Φ (of Cooper pairs) interacting with an external electromagnetic field. A particularly convenient choice is the phenomenological model of Higgs,[7] which also happens to be the relativistic version of the Ginzburg–Landau[8] model of superconductivity. It also is one of the simplest gauge models with broken symmetry encountered in elementary particle physics. The model Lagrangian is given by

$$\mathcal{L} = (\tfrac{1}{2})|D_\mu \Phi|^2 - V(\Phi) - (\tfrac{1}{4})|F_{\mu\nu}|^2, \qquad (3)$$

which is invariant under local gauge transformations of the electromagnetic gauge group $U(1)$. The first term is the kinetic energy, where

$$D_\mu \Phi = (\partial_\mu - iqA_\mu)\Phi \qquad (4)$$

is the "canonical momentum" or, more accurately, the gauge covariant derivative operator. Outside of the vortex region, the covariant derivative vanishes so that the kinetic energy and the Cooper-pair current are both zero. The potential

$$V(\Phi) = \mu^2|\Phi|^2 + \lambda(|\Phi|^2)^2 \qquad (5)$$

describes the self-interaction of the field Φ and can be identified with the free-energy density of the superconductor. The last term in Eq. (3), involving the Maxwell field tensor $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$, gives the usual energy in the electromagnetic field.

The polynomial form of $V(\Phi)$ results in a nonzero ground state for the superconductor given by

$$\Phi_0 = (-\mu^2/\lambda)^{1/2}\exp[iq\theta(x)], \qquad (6)$$

which is obtained by finding the minima of $V(\Phi)$ for negative values of the parameter μ^2. In principle, the ground state is degenerate because of the arbitrariness of the phase. Since $V(\Phi)$ itself is invariant under the group $G = U(1)$, a gauge transformation can rotate the superconductor ground state into an infinity of equivalent ground states:

$$G: \Phi_0 \rightarrow \Phi_0' = \exp[iq\theta']\Phi_0. \qquad (7)$$

However, since the value of the phase is determined by the intrinsic phase relation of the superconductor, such arbitrary local gauge transformations are not allowed. The ground state therefore breaks the symmetry even though the Lagrangian is manifestly gauge invariant.

B. Internal space and topology

How can we now relate this canonical Lagrangian description to our more intuitive geometrical and topological picture of the internal space? Clearly the first step is to locate whatever it is that corresponds to an internal symmetry space in the Lagrangian approach. This is not a trivial task. Since we cannot assume the *a priori* existence of a geometrical internal space within a strictly canonical Lagrangian formalism, we have only the gauge symmetry group itself to work with. Does this mean that we are now confronted with the problem that we avoided earlier, namely, having to learn abstract topology and more group theory as well? Fortunately, the answer is no because we already have all the mathematical tools that we need from our prior knowledge of the familiar three-dimensional rotation group $O(3)$ used in quantum mechanics to study spin states.

We recall that all orientations of a spin state can be generated by starting from a fixed initial spin direction, for example along the z axis, and rotating to the desired direction. The values of the three angles, which specify the rotation, can be considered as the coordinates of a point inside an abstract three-dimensional space.[9] Each point defines a rotation so that the spin states themselves can then be identified with the points in this angular space. Furthermore, since any rotation can be implemented as a continuous sequence of infinitesimal rotations, one can define "paths" between the points and use these paths to determine the topological structure of the space. If one considers a closed path, which starts from one point and returns to the same point, it can be seen that there are two distinct classes of such paths, namely, those that can be shrunk continuously down to the starting point and those that cannot. Examples of the two classes of paths are shown in Fig. 4. The existence of these two classes is related to the familiar double-valued representations of half-integer spin. Thus the rotation group has an angular space that is said to be "doubly connected" from a topological point of view.

For the $U(1)$ gauge group, the angular internal space is just the space defined by the local phase values of the superconductor wave function as we discussed earlier. We can draw this space as the ring of phases as shown in Fig. 5. The distinct closed paths in the space are those that wind around the ring a different number of times. Clearly, a closed path that winds twice around the ring cannot be continuously

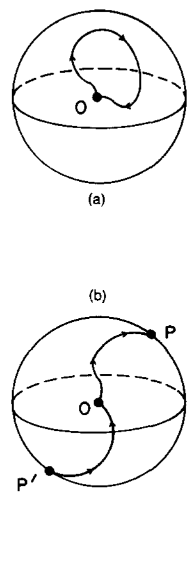

Fig. 4. Examples of the two distinct classes of closed paths in the angular space of the rotation group $O(3)$. (a) A path that can be shrunk back down to the point O. (b) The points P and P' on opposite ends of a diameter represent the same rotation. Thus the path cannot be shrunk to a point. For further details, see Ref. 9.

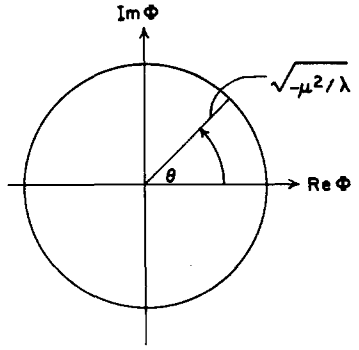

Fig. 5. Ring of phase values in the internal angular space of the group $U(1)$.

deformed or shrunk so that it only winds around once. Thus we see that there are an infinite number of distinct classes of paths so that the $U(1)$ group internal space must have an "infinitely connected" topology.

It is possible to associate the distinct classes of paths with the degenerate ground states of $V(\Phi)$. Each class of paths represents a phase rotation of $2\pi N$. Although we stated above that arbitrary gauge transformations cannot be performed on a ground state, rotations of $2\pi N$ are allowed. The reason for this is that a $2\pi N$ rotation has the same effect everywhere, i.e., it is global, and therefore it preserves the intrinsic phase relation of the superconductor. Hence, a $2\pi N$ rotation will transform a ground state into another perfectly legitimate ground state of $V(\Phi)$. However, the transformed ground state is not equivalent to the original ground state because their phases differ by $2\pi N$, which would violate the continuity requirement of the wave function. Thus we can associate each distinct class of paths with a unique ground state, and by doing so, we also associate the connectedness of the internal angular space with the degeneracy of the potential $V(\Phi)$.

C. Where has the torus gone?

The role of the winding number N is evident in the above discussion but there appears to be no sign of the torus in the canonical formalism. In fact, the torus is hidden because the spatial configuration of the vortex has not yet been taken into account. In the geometrical presentation in Sec. IV, we saw that the surface of the torus was traced out by the phase of the test charge moving along a closed path around the vortex. No explicit test charge is used in the canonical approach, but the same purpose is served by the axial symmetry of the vortex about the direction of the magnetic field lines. Because of this spatial symmetry, we can perform the following "gedanken" operation: imagine cutting through the torus and collapsing the phase windings together so that they look like a compressed spring. The angular space represented by this collapsed torus is precisely the ring of phases in Fig. 5. We note that this gedanken operation is just the reverse of taking the cylinder in Fig. 2 and wrapping it around the vortex as we did before. The important point is that there must be two types of closed paths for the torus; one path winds around N times in the internal space, the other goes once around the vortex in physical space. In the canonical Lagrangian formalism, the physical space path is implicit and only appears in the line integral of Eq. (1). Thus the torus appears to be hidden.

D. Degenerate vortex

The association of each ground state with a distinct class of path, and thus a unique winding number N, leads to an unusual topological picture of the vortex. The vortex can be interpreted as a "transition region" between pairs of different inequivalent ground states.

To see how this interpretation arises, we again use a test charge to trace out the ground-state phase around the vortex. We start with the test charge very far away from the vortex and align its phase with the ground-state wave function. We then transport the test charge up to the vortex, move it through the vortex region and see how the phase has changed. From our preceding discussion, we know that the phase of the test charge will be rotated by the magnetic field in the vortex region. Thus if we move the test charge completely around the vortex back to its initial position, the phase will have changed by $2\pi N$ at the same location. Since the test charge is just tracing out the phase of the superconductor ground state, this means that the test charge must have emerged into a different ground state. This leads to a very unusual picture of the vortex as shown in Fig. 6. If we imagine that the degenerate ground states are all superimposed on top of each other, then the vortex can be interpreted as a "transition region" between the different layers of ground states. It is interesting that this interpretation is analogous to a Riemann surface with multiple sheets.

An interesting consequence of the degeneracy is that it provides a simple topological interpretation for the dynamical stability of the vortex. We know that circulating currents actually confine the magnetic field in the vortex. The degenerate ground states prevent the vortex from spreading out and dissipating into the surrounding superconductor. The reason for this is that the vortex connects different ground states. If the vortex were to "decay," different ground states would then coincide and violate the continuity requirement for the ground-state wave functions. Thus the stability of the vortex can also be interpreted as a topological property.

We see from the preceding discussion that one has to resort to a relatively complicated group-theoretical argument within the canonical formalism in order to uncover geometrical or topological properties associated with the superconductor. It is clear, however, that the value of studying the model Lagrangian is that it provides some insights into the relationship between the details of the superconductor and the topology that are less obvious in the more general geometrical description.

VI. HOW TO ADD THE FLUX NUMBER

Let us now use the results of Secs. II–V to see how the

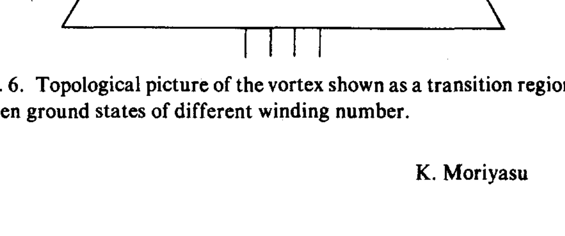

Fig. 6. Topological picture of the vortex shown as a transition region between ground states of different winding number.

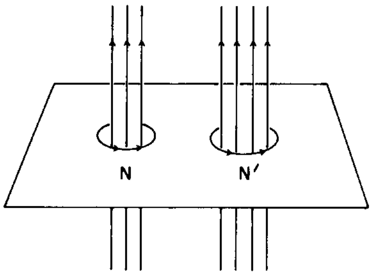

Fig. 7. Two separate vortices with different flux numbers N and N'.

flux number N can be interpreted as a good additive quantum number. The question we want to address is the following: given two separate vortices as in Fig. 7, with individual flux numbers N and N', can we consider the composite system of two vortices to be equivalent to a single vortex with flux number $N + N'$? To answer in the affirmative, we need to show that the sum $N + N'$ is a valid winding number.

We will again use our test charge to determine the properties of the two-vortex system. We want to show that if we move the test charge around each vortex separately, then the net change in phase is the same as that obtained from moving the test charge along a single loop around both vortices. We perform the first part of the operation as shown in Fig. 8(a). Starting from the point x, the test charge is moved around one vortex along the closed loop C and then around the second vortex along C'. The phase changes along C and C' are $2\pi N$ and $2\pi N'$, respectively. In order to add the phase changes, we must be sure that there is no additional phase change at x when the test charge is switched from C to C'. This question arises because each vortex is surrounded by its own ground states and these ground states must be matched up at x. It is clear that there cannot be a phase difference greater than 2π at x because this would allow a transition between ground states. Phase differences of less than 2π could occur because the ground states are not unique; they actually belong to distinct classes of gauge equivalent ground states. However, equivalent ground states will give the same contribution to the net winding number. We therefore conclude that the flux numbers of the two vortices can be added together to yield a net flux number $N + N'$ for the system.

To see that $N + N'$ is also the flux number for a single path around both vortices, we will show that the loops C and C' can be changed into a single loop. We are allowed to distort the loops as long as the flux is completely encircled by the distorted loops and the winding number is unchanged. The procedure for altering the loops is illustrated in Fig. 8(b). The endpoint of C and the starting point of C' are moved from x to a new point y. We are effectively cutting the torus and "patching in" a new piece between x and y. The net flux number remains unchanged because any extra phase change that we have introduced from x to y along C is cancelled by a change from y to x along C'. The path segment between x and y can then be ignored and we are left with a single continuous loop around both vortices. This argument is also reversible since the segment between x and y can be shrunk back down to a point. Thus we finally conclude that the flux number of the single loop is indeed the same number as the sum of the two loops.

The geometrical path approach for adding flux numbers has the virtue that it leads to a general symmetry structure that is valid for systems other than the vortex. The manipulations we used to merge the loops C and C' into a single loop actually define a new "hidden" symmetry group for the flux number. This unusual group consists of the closed loops themselves with their associated winding numbers plus a definition of the "product" of two loops. A closed loop with winding number N is taken to be the Nth element of the group. The identity or null element is a loop with $N = 0$; it does not encircle any net flux so that it can be shrunk down to a point. The inverse of a loop is another loop that winds in the opposite direction and thus has negative winding number. As we saw above, two loops with winding numbers N and N' can be combined to form a new single loop with winding number $N + N'$. This defines the "product" of two loops. Since the order of the loops is not physically relevant, the product is clearly commutative and the group is Abelian. A product can be formed of loops around separate vortices or around one vortex. For example, the two cases shown in Fig. 9 both give a product loop with net flux number equal to zero but with very different physical interpretations. The first case is just a loop that does not circle the vortex while the second case shows a system consisting of a vortex and an antivortex with net flux number equal to zero.

The group of closed paths is called the "fundamental group"[10] and is symbolized by

$$\Pi_1[U(1)], \qquad (8)$$

where the subscript refers to the one-dimensional closed path in physical space around the vortex. The $U(1)$ gauge group is shown explicitly in the brackets to indicate that the closed loops, which are the group elements, are defined in

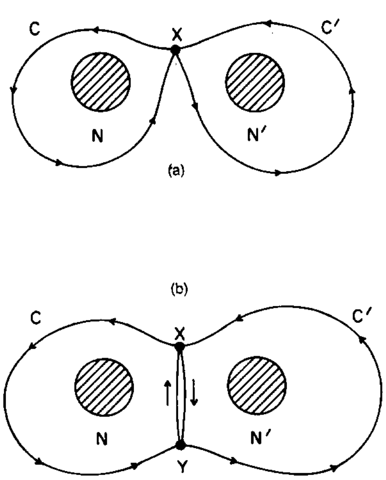

Fig. 8. (a) Closed loops C and C' used to determine the phase change around separate vortices with flux numbers N and N', respectively. (b) Loops C and C' are converted into a single loop around both vortices by moving endpoints from x to y. Phase changes along pieces of loop patched in between x and y cancel so that they can be ignored.

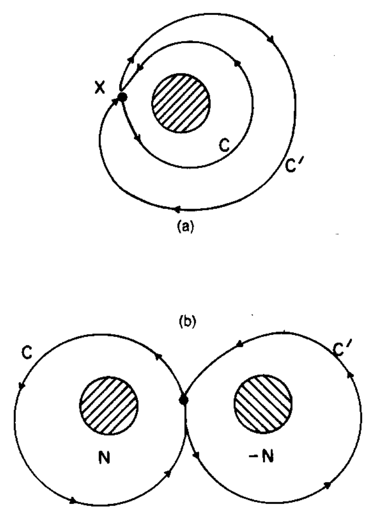

Fig. 9. Examples of product of two loops giving net flux number zero. (a) Product that does not circle the vortex. (b) Product that circles vortex–antivortex system with total flux zero.

the $U(1)$ angular internal space. Since each element of the group is labelled by a unique integer winding number N, the group is also said to be equivalent to the integers themselves, which form an additive Abelian group (denoted by Z); hence, one can write symbolically

$$\Pi_1[U(1)] \simeq Z. \qquad (9)$$

This equivalence relation neatly summarizes the "hidden" symmetry that underlies the flux quantum number. It also graphically demonstrates just how well the symmetry is hidden from view. A group like Π_1 is rarely encountered in other areas of physics because it is not directly related to any coordinate transformation like most familiar symmetry groups. In fact, Π_1 could be regarded as uniquely characteristic of gauge theory because it arises out of the relationship between sets of paths in the two entirely different types of spaces that are married together by gauge theory.

VII. DIRAC MAGNETIC MONOPOLE

One of the most interesting applications of the topological ideas we have discussed is the recent resurrection of the classic Dirac magnetic monopole. The magnetic monopole has been sought both experimentally and theoretically as the missing link needed to symmetrize Maxwell's equations. Dirac[11] showed 50 years ago that if the magnetic monopole existed, then the flux must be quantized according to the relation

$$\text{Flux} = 2\pi N\hbar c/e, \qquad (10)$$

where N is the same flux number as in the case of the vortex. Recently, Wu and Yang[12] re-examined Dirac's results within the context of gauge theory and found that the quantization of the monopole flux also could be interpreted as a topological condition.

We can apply the concepts uncovered in the study of the vortex to gain some insight into the nature of the monopole. Dirac's quantization condition is essentially identical to the vortex condition given by Eq. (1) and originally was derived by using very similar physical arguments. In analogy with the electron, the monopole is supposed to be a pointlike source of a magnetic "Coulomb field"

$$\mathbf{B} = -g\nabla(1/r), \qquad (11)$$

where g is the strength or "magnetic charge." By using an electron as a test charge, the flux can be related to the change in phase of the electron wave function. Assuming that the field around the monopole is spherically symmetric, one can calculate the phase change along a closed path C on a spherical surface surrounding the monopole as shown in Fig. 10.

The net phase change is given by

$$\delta\theta = (e/\hbar c) \oint_C \mathbf{A}\cdot d\mathbf{x} = 2\pi N. \qquad (12)$$

The quantization condition results from the continuity requirement imposed on the electron wave function just as in the case of the vortex.

An essential complication now arises when one tries to use Stokes's theorem to relate the phase change to the monopole flux. The closed loop C can be taken as the boundary of either one of the two hemispheres S and S' in Fig. 10. Thus we see that

$$\oint_C \mathbf{A}\cdot d\mathbf{x} = \text{Flux}(S) = \text{Flux}(S'). \qquad (13)$$

The net flux out of the sphere is the difference of Flux(S) and Flux(S'), which gives zero and clearly makes no sense. Dirac resolved this contradiction by concluding that there must be some kind of singularity in the magnetic field that passes through the spherical surface. In order to calculate the net flux, a hole has to be cut in the surface to avoid the singularity. The net flux out of the remainder of the sphere would then be nonzero, thus resolving the contradiction and giving the Dirac quantization condition.

Wu and Yang recently showed that the quantization condition could be derived without explicitly invoking singularities if the magnetic field was reinterpreted as a multiply connected field. The vector potential $\mathbf{A}$ around the monopole is considered to be multivalued so that different potentials $\mathbf{A}$ and $\mathbf{A}'$ are used to calculate the flux through S and S', respectively. Along the closed path C, the potentials are related by a gauge transformation

$$\mathbf{A} = \mathbf{A} - i(\hbar c/e)U^{-1}\nabla U, \qquad (14)$$

where the transformation U has the form

$$U = \exp[i\lambda(x)]. \qquad (15)$$

In order for U to be single valued along C, the phase factor λ is required to satisfy the condition

$$\oint_C \nabla\lambda\cdot d\mathbf{x} = 2\pi N. \qquad (16)$$

This equation takes the place of the continuity condition on the electron's wave function. Using Eq. (14) in Stokes's theorem then gives the phase change

$$\begin{aligned}\delta\theta &= (e/\hbar c) \oint_C [\mathbf{A}' - \mathbf{A}]\cdot d\mathbf{x} \\ &= \oint_C \nabla\lambda\cdot d\mathbf{x} = 2\pi N \\ &= (e/\hbar c)[\text{Flux}(S') - \text{Flux}(S)], \end{aligned} \qquad (17)$$

which leads once again to Dirac's quantization condition.

The Wu–Yang derivation appears to give the impression that there are no singularities in the field. However, it has been argued by Barut,[13] using different techniques, that the singularities must be real. If this is so, how can a singularity be equivalent to a multiply connected field? We can answer this question with a simple analogy from the theory of complex variables. Let us consider the case of a function such as log$z$, which has a branch cut singularity. It is well known[14] that if we replace the z plane with a Riemann surface with multiple sheets, then log$z$ can be treated like a single-valued function without a branch cut. A similar situation occurs in the Wu–Yang derivation. The Riemann

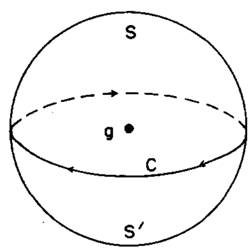

Fig. 10. Spherical surface of integration around Dirac magnetic monopole of magnetic charge g.

surface is the analog of the multiply connected internal phase space and the gauge transformation, Eq. (14), tells us how the potential changes from one "sheet" to the next. Thus we see that the interpretations of Dirac and Wu–Yang can be perfectly compatible.

The properties of the monopole and the vortex provide an interesting contrast. The flux quantization conditions are essentially identical, yet the Dirac monopole is supposed to be a fundamental particle like the electron while the vortex is a complicated dynamical system. In addition, the gauge symmetry of the monopole is not broken. There also are no degenerate ground states associated with the monopole; thus the multivaluedness of the field cannot be blamed on the degeneracy of the ground states and the monopole cannot be interpreted as a "transition region" like the vortex. In fact, one might say that the monopole resembles a vortexlike system in which all of the underlying physical mechanisms have been hidden. Thus the topological properties of the monopole appear more "abstract" than those of the vortex.

VIII. DISCUSSION

We have seen that an amazingly rich topological structure can be uncovered from a systematic study of the well-known phenomenon of flux trapping. By using the familiar pedagogical device of a test charge and simple geometrical arguments, the flux quantum number N is shown to be associated with a new type of "hidden" symmetry. The new symmetry is not based on the usual type of transformation laws but rather is topological in origin.

The two examples of the vortex and the Dirac magnetic monopole demonstrate that topological quantum numbers can arise in very different types of physical systems. We saw that the flux quantization equations were essentially identical, yet the underlying physics could hardly be more different. The vortex is an intricate dynamical system while the monopole is supposed to be an elementary particle like the electron. This contrast between the vortex and the monopole clearly shows that the topological quantum number is independent of the details of the particular system and is based on very general principles.

The discovery of physically relevant topological properties in gauge theories has stimulated many new investigations. By examining gauge theory models with more complex non-Abelian gauge groups G. t'Hooft and others[15] have found new monopolelike solutions with interesting properties in a variety of gauge models. It has also been suggested[16] that the kind of topological stability seen in the vortex may provide some insight into how quarks are confined inside hadrons.

ACKNOWLEDGMENT

This work was supported by the National Science Foundation.

[1]S. Weinberg, Phys. Rev. Lett. **19**, 1264 (1967); A. Salam, in *Elementary Particle Theory Nobel Symposium 1968*, edited by N. Svartholm (Wiley, New York, 1968).

[2]R. P. Feynman, R. B. Leighton, and M. Sands, *Feynman Lectures in Physics* (Addison-Wesley, Reading, MA, 1965), Vol. 3.

[3]K. Moriyasu, Am. J. Phys. **46**, 274 (1978).

[4]For further details of the fiber bundle formalism, see R. Hermann, *Lie Groups for Physicists* (Benjamin, New York, 1966); M. E. Mayer, *Fiber Bundle Techniques in Gauge Theories* (Springer-Verlag, Heidelberg, 1977).

[5]Y. Aharonov and D. Bohm, Phys. Rev. **115**, 485 (1959).

[6]For an introduction to elementary topology, see P. Roman, *Some Modern Mathematics for Physicists and Other Outsiders* (Pergamon, New York, 1975), Vol. 1.

[7]P. W. Higgs, Phys. Lett. **12**, 132 (1964); **13**, 508 (1964); Phys. Rev. **145**, 1156 (1966).

[8]V. L. Ginzburg and L. D. Landau, Zh. Eksp. Theor. Fiz. **20**, 1064 (1950).

[9]K. Gottfried, *Quantum Mechanics* (Benjamin, Reading, MA, 1966), Vol. I.

[10]This group is also called a "homotopy" group since the elements are sets of paths that are equivalent or "homotopic" to each other. For further details, see G. H. Thomas, "Introductory Lectures on Fiber Bundles and Topology for Physicists," Argonne National Laboratory, ANL-HEP-PR-78-23, May 1978 (unpublished).

[11]P. A. M. Dirac, Proc. R. Soc. London A **133**, 60 (1931). See also P. A. M. Dirac, *Directions in Physics* (Wiley, New York, 1978).

[12]T. T. Wu and C. N. Yang, Phys. Rev. D **12**, 3845 (1975); Nucl. Phys. B **107**, 365 (1976). See also L. H. Ryder, J. Phys. A **13**, 437 (1980).

[13]A. Barut, Sov. J. Part. Nucl. **10**, 209 (1979).

[14]R. V. Churchill, *Complex Variables and Applications* (McGraw-Hill, New York, 1960), p. 273.

[15]G. t'Hooft, Nucl. Phys. B **79**, 276 (1974); A. M. Polyakov, JETP Lett. **20**, 194 (1974).

[16]S. Mandelstam, in Proceedings of Joseph H. Weis Memorial Symposium on Strong Interactions, University of Washington, 1978 (unpublished).

PROBLEM: QUANTIZED RELATIVISTIC MOTION OF A HARMONIC OSCILLATOR

Use the Bohr quantization condition to find the energy levels of a relativistic one-dimensional harmonic oscillator. (The solution is on page 849.)

I. Richard Lapidus

Department of Physics and Engineering Physics
Stevens Institute of Technology
Hoboken, New Jersey 07030