On the Theory of Non-Abelian Vortices and Cosmic Strings

Thesis by
Martin Bucher

In Partial Fulfillment of the Requirements
for the Degree of
Doctor of Philosophy

California Institute of Technology
Pasadena, California

1991
(Submitted September 13, 1990)

Table of Contents

Acknowledgements

First and foremost. I would like to thank my advisor Professor John Preskill for valuable insight. countless discussions. many ideas and suggestions. and much encouragement. John taught me how to think about physics in a clearer and more intuitive way. John made a lot of things seem quite simple that prior to my arrival at Caltech seemed hopelessly complicated and confusing.

I also want to acknowledge some of my other teachers who have helped me a great deal and who have had a great influence on me. I would like to thank Professor Charles Schwartz and Professor Robert Ely. I also want to thank two of my math professors. Professor Charles Rickart and Professor G. Milton Wing, from whom I learned most of the math that I know.

Finally, I would like to thank Professor Charles Schwartz and Professor Harry Morrison for their hospitality at UC Berkeley this summer.

Abstract

The thesis deals with the theory of non-Abelian vortices in two spatial dimensions and cosmic strings in three spatial dimensions that arise when a non-Abelian gauge symmetry G is broken to a non-Abelian unbroken symmetry group H by the condensation of a Higgs field. The first part of the thesis discusses the case in which H is discrete. In this case all of the gauge bosons acquire a large mass; however, at low energies discretely charged particles experience non-Abelian Aharonov-Bohm scattering off vortices, which can be used to measure the flux of the vortices. Vortices also experience non-Abelian Aharonov-Bohm scattering with each other. When there are more than three vortices in a system, the Aharonov-Bohm interaction, which is described by a path integral involving sums over elements of the braid group, becomes extremely complicated. The vortices are subject to a new kind of exotic statistics. The second part of the thesis discusses the physics that arises when the requirement that H be discrete is relaxed to allow H to have one continuous generator. The vortices or strings that result change the sign of the charge for charged particles. Loops of Alice string or pairs of vortices can carry charge without any apparent source. A quantization condition for the charge carried by such pairs and loops is derived. It is also found that loops of Alice string can carry magnetic charge and that topologically stable monopoles exist in any theory with Alice symmetry breaking.

Chapter 1
Introduction

In the last few decades the idea of broken symmetry has played a key role in theoretical physics—both of high energies and of condensed matter. Often it is the case that at high temperatures the full symmetry of a system is manifest; however, upon cooling, a system experiences a phase transition in which the symmetry is broken through the choice of orientation of an order parameter, which no longer vanishes in the lower temperature phase. Examples of such systems include the Heisenberg ferromagnet and the Standard Model, which must have been in a symmetric phase in the early universe when the temperature was much hotter. In general, for a sufficiently large system, during the process of cooling there is little or no communication between the various parts, and domains of the phase with broken symmetry with uncorrelated orientations of the order parameter start to form in various places. As these domains grow bigger and start to coalesce, the various vacuum orientations of the domains try to align themselves, for there is an energetic cost, generically to lowest order equal to the square of the gradient, associated with a non-uniform choice of vacuum orientation.

In one scenario the final state of the system, after as much energy as possible has been dissipated, approaches a uniform vacuum. However, for some systems, because of the topological properties of the manifold of degenerate vacua, it is impossible for the system to evolve continuously toward a state with a uniform choice of vacuum. Instead, the vacuum aligns itself as well as it can, and the discrepancies in the choice of vacuum become concentrated near submanifolds of lower dimension known as topological defects.

These topological defects can be classified using the following technique. Let Ψ be a spatially dependent order parameter subject to a potential $V(\Psi)$ and let G be the symmetry group of the theory so that $V(g\Psi) = V(\Psi)$. Suppose that V takes a minimum value at Ψ_0 and that G acts on Ψ_0 non-trivially. Then the set of values of the order parameter $M = \{g\Psi_0 \mid g \in G\}$ also minimizes V. [1.] For $\Psi \in M$, we define the unbroken symmetry group as $H(\Psi) \subset G$, whose embedding in G depends on the choice of $\Psi \in M$. Let $H_0 = H(\Psi_0)$. Since

$$M \cong G/H, \tag{1.1.1}$$

the coset space has the same topology as the manifold of degenerate vacua. Topological defects can be classified by the non-trivial homotopy classes of functions from an n-sphere S^n [2] enclosing the defect into (G/H). [3,4.] In three spatial dimensions defects described by $\pi_0(G/H)$ are domain walls, those described by $\pi_1(G/H)$ are cosmic strings, and those described by $\pi_2(G/H)$ are monopoles.

In this thesis we shall only consider defects described by $\pi_1(G/H)$—cosmic strings in three spatial dimensions and pointlike vortices in two dimensions. The discussion shall be further restricted to theories in which the symmetry is gauged. The fact that a symmetry is gauged has several important consequences. First, it assures that the energy of a vortex is finite and concentrated at the core. Suppose that there were no gauge field to compensate for the variation in Ψ as one traverses around the vortex. Since the magnitude of the derivative is of order $(1/r)$, the energy of the vortex (taking the energy of the vacuum to be zero) would be proportional to

$$\int_{r_<}^{r_>} r dr \left(\frac{1}{r}\right)^2 = \ln\left[\frac{r_>}{r_<}\right]. \tag{1.1.2}$$

The divergence as $r_< \to 0$ is naturally regulated by the vortex core size, but the divergence as $r_> \to \infty$ cannot be removed. When the symmetry is gauged, however, a vector field $\mathbf{A}$ can be chosen such that at large distances from the core the covariant derivative, which is the derivative that appears in the Hamiltonian, vanishes.

The organization of the thesis is the following. Chapter 2 discusses Abelian vortices, which occur both in the Ginzburg-Landau phenomenological model of superconductivity and in the relativistically invariant Abelian Higgs model. The phase structure and vortex solutions are discussed, and in the last section it is described how gauged Z_N discrete symmetry and anyons can result in a $U(1)$ Higgs model. In Chapter 3 some of the complications that arise when the unbroken symmetry group is non-Abelian are described. The non-Abelian magnetic flux of a vortex and the superselection rule for the total flux of all of the vortices are discussed. Problems associated with the global realization of the unbroken symmetry group H are also discussed. In Chapter 4 the physics of non-Abelian vortices for which the unbroken symmetry group H is discrete is discussed in detail. [5.] It is found that when H is non-Abelian, the vortices in the theory experience a long-distance interaction due to the Aharonov-Bohm effect. A path integral formulation for the general N vortex problem involving the braid group is presented. For two vortices the scattering is exactly analogous to scattering of an electromagnetically charged particle from an infinitely thin tube of magnetic flux, first calculated by Aharonov and Bohm. However, when more than two vortices are present, it is no longer possible to treat one of the vortices quantum mechanically and the remaining vortices classically. When a third vortex passes between two vortices, the relation between the two vortices changes. Consequently, a vortex that moves through a background containing more than one vortex changes the state of the background in a path dependent manner. For the N vortex problem, the path integral is the sum over e^{iS} where S is the free particle action with the summation restricted to the homotopy equivalence classes of vortex paths that create the same gauge field topology for the final state. Exotic statistics result for non-Abelian vortices because what types of vortices may be considered identical depends on what other kinds of vortices are present in the system. Finally, the path integral formulation for non-Abelian vortices is compared to the path integral formulation of anyon dynamics. In Chapter 5 the discussion is extended to the case in which H has exactly one linearly independent infinitesimal generator. If this generator Q does not commute with the discrete elements of H, interesting new physics results. The resulting strings, called Alice strings, prevent the global implementation of the $U(1)_Q$ symmetry in the presence of vortices, and the vortices flip the sign of the charge of a charged particle that travels around the Alice vortices

an odd number of times. Pairs of Alice vortices and closed loops of Alice string can carry electric charge that does not have any apparent source. The charge carried by a pair (or loop) is defined by surrounding the pair (or loop) with a surface and integrating the electric field over this surface. This definition of electric charge is extended to the quantized system, and a quantization condition for the charge carried by such pairs (or loops) is derived. It is shown that loops of Alice string can carry magnetic charge and that there must exist topologically stable magnetic monopoles in any theory that supports Alice strings.

Notes

1. We shall assume that all of the minima of V are related by the action of the symmetry group G—in other words, that there is no accidental degeneracy that is not a consequence of the symmetries of the theory.

2. S^n is the n-dimensional sphere. The zero-dimensional sphere is defined as the two points $S^0 = \{-1, 1\}$.

3. J. Preskill, "Vortices and Monopoles," in P. Ramond and R. Stora, Eds., *Architecture of the Fundamental Interactions at Short Distances*, Amsterdam: North Holland (1987).

4. D. Mermin, "The Topological Theory of Defects in Ordered Media," Rev. Mod. Phys. **51**, 591 (1979).

5. M. Bucher, "The Aharonov-Bohm Effect and Exotic Statistics for Non-Abelian Vortices," Caltech Preprint 68-1655 (1990). [To appear in Nucl. Phys. **B**.]

Chapter 2
Abelian Strings and Vortices

2.1. Introduction

In this chapter the physics of vortices that arise when a $U(1)$ symmetry is broken by the formation of a Higgs condensate is described. Most of the material in this section is old and well-known. In fact, most of what is known today about Abelian vortices was worked out in the fifties in the context of the Ginzburg-Landau model of superconductivity. [1.] It was only some twenty-three years later, in an attempt to construct from a gauge field theory strings that obey the Nambu action, that vortices arising in relativistically invariant gauge field theories were first discussed. [2.] The discovery by t'Hooft that non-Abelian gauge theories are renormalizable, even in a phase that spontaneously breaks all or part of the gauge symmetry, and the experimental success and widespread acceptance of the Weinberg-Salam model naturally led to a search for similar models with a higher gauge symmetry that would unify the strong and electroweak interaction in a more complete manner, generally using a single simple, compact, non-Abelian gauge group G, in order to reduce the number of adjustable parameters in the theory. This development in turn led to the study of the string-like defects that arise in many such models.

This chapter summarizes the most important results for Abelian strings. The next chapter extends the discussion to non-Abelian strings. Abelian strings arising from the breaking of $U(1)$ gauge symmetry are discussed, both in the context of the Ginzburg-Landau model and in the context of the relativistically invariant Abelian Higgs model. The theory of the superconducting state of condensed matter is of interest because the superconducting phase is essentially a Higgs phase characterized by the property that magnetic flux is confined to vortices. A somewhat extensive summary of the phase structure of Type I and Type II superconductors at various values of the surrounding magnetic field H_0 is included here because the properties of these phases offer a simple way to understand qualitatively the structure, interactions, and relative masses of the vortices that arise in the Abelian Higgs model. The Ginzburg-Landau model and the Abelian Higgs model (considered as a classical, unquantized field theory) are essentially identical. Except for the absence of a time-derivative term, the expression for the Ginzburg-Landau free energy is the same as the expression for the Abelian Higgs action. Solutions of the Ginzburg-Landau differential equations correspond to time-independent solutions of the Abelian Higgs model.

2.2. Ginzburg-Landau Model of Superconductivity

In 1950 Ginzburg and Landau published their seminal paper introducing a phenomenological theory of superconductivity now known as the Ginzburg-Landau model. [1.] Although the BCS theory of superconductivity [3] is more fundamental in the sense that it explains superconductivity in terms of microscopic interactions, in practice the Ginzburg-Landau model is often more useful because of the ease with which it treats spatial inhomogeneities. As Gor'kov showed in 1959, the Ginzburg-Landau model describes a limiting case of the BCS theory. [4.] In the Ginzburg-Landau model, the state of the medium capable of being in either a normal or superconducting phase is represented by a complex scalar order parameter field $\Psi(\mathbf{x})$ whose amplitude squared in some crude sense represents a density of superconducting electrons n_s. The overall phase of the complex order parameter field is physically unobservable and therefore meaningless. The $U(1)$ global symmetry is generalized to a local symmetry by the introduction of a vector potential $\mathbf{A}(\mathbf{x})$ and replacement of the ordinary derivative with the covariant derivative

$$\mathbf{D} = \left(\nabla + \frac{ie^*}{\hbar c}\mathbf{A}\right). \tag{2.2.1}$$

The gauge invariant current

$$\begin{aligned} \mathbf{J} &= \frac{e^*}{2m^*}\Big[\Psi^\dagger[(-i\hbar\nabla)\Psi] + [(-i\hbar\nabla)\Psi]^\dagger\Psi\Big] + \frac{e^{*2}}{m^*c^2}\cdot\mathbf{A}\cdot(\Psi^\dagger\Psi) \\ &= \frac{e^*}{2m^*}\cdot(-i\hbar)\cdot\left[\Psi^\dagger(\mathbf{D}\Psi) - (\mathbf{D}\Psi)^\dagger\Psi\right] \end{aligned} \tag{2.2.2}$$

represents the current that is due to the flow of the charged superfluid, which in the language of the BCS theory is a condensate of Cooper pairs with charge $e^* = -2|e|$. The free energy in terms of $\mathbf{A}(\mathbf{x})$ and $\Psi(\mathbf{x})$ is

$$F[\Psi(\mathbf{x}),\mathbf{A}(\mathbf{x})] = \int d^3x \left[f_{n0}(T) + \frac{\hbar^2}{2m^*}\left|\left\{\nabla - \frac{ie^*}{\hbar c}\mathbf{A}\right\}\Psi(x)\right|^2 + \frac{a(T)}{2}(\Psi^\dagger\Psi)^2 - b(T)(\Psi^\dagger\Psi) + \frac{\mathbf{B}^2}{8\pi}\right] \tag{2.2.3}$$

where $f_{n0}(T)$ is the free energy density of the normal phase at temperature T. The coefficients $a(T)$ and $b(T)$ that appear in the potential may either be derived from a microscopic theory or chosen to agree with the experimentally observed values of the critical magnetic field H_c and the magnetic penetration depth λ.

Physically, the Ginzburg-Landau variational principle, which requires that the free energy be at a local minimum, is a method of finding static (meaning time-reversal invariant as well as time-independent) configurations of the system at thermal equilibrium. By setting $\delta F = 0$ with respect to variations in $\Psi(\mathbf{x})$ and $\mathbf{A}(\mathbf{x})$, one obtains the Ginzburg-Landau equations

$$\begin{aligned} &\frac{-\hbar^2}{2m^*}\left(\nabla - \frac{ie^*}{\hbar c}\mathbf{A}\right)^2\Psi + a(T)(\Psi^\dagger\Psi)\Psi - b(T)\Psi = 0 \\ &\nabla\times\mathbf{B} = \nabla(\nabla\cdot\mathbf{A}) - \nabla^2\mathbf{A} = \frac{4\pi}{c}\mathbf{J}_s \\ &= \left(\frac{2\pi e}{m^* c}\right)\left[\Psi^\dagger[(-i\hbar\nabla)\Psi] - [(-i\hbar\nabla)\Psi]^\dagger\Psi\right] + \frac{4\pi e^{*2}}{m^* c^2}\mathbf{A}\,(\Psi^\dagger\Psi). \end{aligned} \tag{2.2.4}$$

Let us first consider solutions for which $\mathbf{A} = 0$ and Ψ is constant. Stability requires that $a(T) > 0$, for otherwise the free energy would have no lower bound. When $b(T) < 0$, $\Psi(x) = 0$ minimizes $F[\Psi, \mathbf{A}]$, indicating that the material is in its normal phase at $T > T_c$. When $b(T) > 0$, the free energy is minimized for $|\Psi(x)| = b(T)/a(T)$, implying that

$$f_{s0} = f_{n0} - \frac{b^2(T)}{2a(T)} \tag{2.2.5}$$

where f_{s0} is the free energy density of the uniform superconducting phase. As a consequence, the thermodynamic critical field [4] is

$$H_c = \sqrt{4\pi \frac{b(T)}{a(T)}}. \tag{2.2.6}$$

Next consider the situation in which a semi-infinite slab of superconducting material at ($z > 0$) is bounded by an empty region for ($z < 0$) in which a constant magnetic field of intensity H_0 is maintained parallel to the superconducting surface. At the boundary of a superconducting sample it is necessary to impose the boundary condition that the current normal to the boundary vanish. A surface current on the face of the superconductor, which is spread over a depth of order λ, will cause B to vanish inside the superconductor. Choosing a gauge in which $\mathbf{A} = \hat{x}A(z)$, one obtains the coupled differential equations

$$-\frac{\hbar^2}{2m^*}\left\{\frac{d^2}{dz^2} + \frac{e^{*2}}{\hbar^2 c^2}A^2(z)\right\}\Psi(z) + a(T)[\Psi^+(z)\Psi(z)])\Psi(z) - b(T)\Psi(z) = 0 \tag{2.2.7}$$

and

$$\frac{d^2}{dz^2}A(z) + \left\{\frac{4\pi e^{*2}}{m^* c^2}\right\}A(z) = 0. \tag{2.2.8}$$

For weak fields with $H_0 << H_c$, $\Psi = (b/a) + O(H_0^2)$. Therefore, for small H_0 the solution inside the superconductor is

$$H(z) = H_0 \exp\left[-(z/\lambda)\right] \tag{2.2.9}$$

where

$$\lambda = \sqrt{\frac{m^* c^2}{4\pi e^{*2}} \cdot \frac{a(T)}{b(T)}}. \tag{2.2.10}$$

As H_0 increases, $|\Psi(z)|$ drops near the boundary, allowing the magnetic field to penetrate deeper into the superconductor. In the limit of weak magnetic fields, the description of F. London's theory for electrodynamics in a superconductor, which assumes n_s to be constant, becomes a valid approximation. However, as the magnitude of the applied external field increases, effects non-linear in H_0 cause the density of superconducting electrons to decrease.

In the Ginzburg-Landau theory there is another length scale ξ, somewhat misleadingly called the correlation length, which is the length scale of the fluctuations of $|\Psi|$. For $A = 0$ the Ginzburg-Landau equation is

$$-\frac{\hbar^2}{2m^*}\nabla^2\Psi + a(T)(\Psi^\dagger\Psi)\Psi - b(T)\Psi = 0. \tag{2.2.11}$$

Perturbations about one of the solutions $\Psi = (b/a)e^{i\phi}$ with $0 \leq \phi < 2\pi$ that do not perturb the phase decay (or blow up) exponentially in the linear approximation with a characteristic length

$$\xi = \frac{\hbar}{\sqrt{4b(T)m^*}}, \tag{2.2.12}$$

which is distinct from λ.

Let us again consider the case of the thick superconducting slab in a strong magnetic field. For $H_0 = H_c$, $\Psi = 0$ at $z = 0$. Physically, the vanishing of Ψ at the boundary implies that the region $z < 0$ could be replaced with a slab of superconducting material in the normal phase with $H = H_0$ and $\Psi = 0$. The value of H_0 at which such a boundary is stable can be determined without solving the differential equation by considering the variation in F as the boundary is moved. Such variation must vanish, for otherwise the superconducting and normal phases could not coexist. One finds that [7]

$$f_{s0} = f_{n0} - \frac{H_c^2}{8\pi} \tag{2.2.13}$$

because the energetic cost of expelling magnetic flux must equal the drop in free energy associated with enlarging the volume of the superconducting region.

The ratio between these two length scales $\kappa = (\lambda/\xi)$ determines the qualitative behavior of the vortex solutions and the normal phase/superconducting phase interface because the sign of the surface energy depends only on κ. The surface tension is

$$\sigma = \int_{-\infty}^{\infty} dz \Bigg[\frac{H^2}{8\pi} - \frac{H \cdot H_c}{4\pi} + \frac{H_c^2}{8\pi} + \frac{\hbar^2}{2m^*} \left| (\nabla + \frac{ie^*}{\hbar c} A)\Psi \right|^2 + \frac{a(T)}{2}(\Psi^\dagger\Psi)^2 - b(T)(\Psi^\dagger\Psi)^2 \Bigg], \tag{2.2.14}$$

which by subtracting the integrated Ginzburg-Landau equation can be simplified to

$$\sigma = \int_{-\infty}^{\infty} dz \left[\frac{(H - H_c)^2}{8\pi} - \frac{a(T)}{2}(\Psi^\dagger\Psi)^2 \right]. \tag{2.2.15}$$

Qualitative arguments allow one to determine the sign of σ in the limit of very small and very large κ without evaluating the integral. When $\kappa << 1$, $\xi >> \lambda$. Consequently, it is energetically very unfavorable to have a large area because flux must be expelled without a concomitant increase in $|\Psi|$. Likewise, for $\kappa >> 1$, $\lambda >> \xi$, and it is energetically very favorable to have a large surface area because inside the boundary layer $\Psi(\mathbf{x})$ can take a value near the minimum of the potential without expelling much flux. Numerical integration is necessary to determine the exact value of κ at which $\sigma = 0$. This value turns out to be $\kappa = 1/\sqrt{2}$.

Physically, $\sigma > 0$ means that the surface is stable because the free energy is minimized for a surface of minimal area. On the other hand, $\sigma < 0$ implies that the surface is a mathematical artefact and does not correspond to a stable physical solution. The surface is unstable against any perturbation that increases its area, so the free energy is lowered by creating a phase that is essentially all surface. Such a mixed state is exactly what occurs in practice. At first the case with $\kappa > 1/\sqrt{2}$ received little attention because it was thought that for all real superconductors $\kappa << 1/\sqrt{2}$. However, there remained a class of superconductors, the hard or Type

II superconductors, whose properties evaded explanation within the framework of the then existing theories. In 1959 A. B. Abrikosov published a paper discussing the case $\kappa > 1/\sqrt{2}$. [6.]

For Type II superconductors, described by $\kappa > 1/\sqrt{2}$, there are two critical fields. For weak fields with $H < H_{c1}$, the superconducting phase is stable. However, at $H = H_{c1}$ the energy required to create an isolated vortex inside the superconductor becomes precisely equal to the drop in energy due to removing the required flux from the region outside of the surface. As H increases above H_{c1}, vortices continue to enter the material until a vortex density is reached at which the repulsion between vortices has raised the energy required to add another vortex sufficiently to match the decrease in energy that results from removing flux from the exterior. For $H_{c1} < H < H_{c2}$ the exclusion of magnetic flux is only partial, implying a partial Meissner effect.

2.3. Nucleation of the Mixed State

It is possible to calculate the critical field H_{c2} using the linearized Ginzburg-Landau equation because as H is varied about H_{c2}, the fields A and Ψ vary in a continuous manner, unlike the transition between a Type I superconductor and the normal phase. In the Type I superconductor, B inside the superconductor varies discontinuously from $B = 0$ for $H < H_c$ to $B = H$ when $H > H_c$. This fact means that the Type I superconducting/normal phase transition is a first order phase transition. In contrast, for Type II superconductors B varies continuously as H_0 is varied near H_{c2}. As H_0 increases from H_{c1}, vortices continue to permeate the superconducting material until at H_{c2} the magnetization of the superconductor vanishes as $|\Psi| \to 0$. The phase transition is second-order. For H_0 only slightly below H_{c2}, B differs only slightly from H_0. Therefore, in this regime it is justified to introduce the ansatz

$$\begin{aligned} A(\mathbf{x}) &= A_0(\mathbf{x}) + \delta A(\mathbf{x}) \\ \Psi(\mathbf{x}) &= \delta\Psi(\mathbf{x}) \end{aligned} \tag{2.3.1}$$

and to consider the linearized equations for $\delta\Psi(\mathbf{x})$ and $\delta\mathbf{A}(\mathbf{x})$.

The Ginzburg-Landau free energy can be divided into two parts—a part quadratic and a part quartic in Ψ. For F to have a non-trivial minimum, the quadratic part must have a negative eigenvalue because the quartic part always contributes positively to the free energy. Whenever the quadratic part has a negative eigenvalue, the global minimum of F is nontrivial. When $H > H_{c2}$, all eigenvalues are positive. When $H = H_{c2}$, zero becomes an eigenvalue. To find H_{c2} one solves the Ginzburg-Landau equation linearized in Ψ in a background with a constant magnetic field H, finding the highest value of H for which there exists a solution. The linearized Ginzburg-Landau equation obtained is identical to the Schroedinger equation for a particle in a constant magnetic field with $b(T)$ acting as the energy eigenvalue. For a particle in a constant magnetic field, the lowest energy state has a ground state energy $E_0 = \hbar\omega_c/2$ where

$$\omega_c = \frac{e^* H_{c2}}{m^* c} \tag{2.3.2}$$

is the cyclotron frequency. Setting $b(T) = \hbar\omega_c/2$ implies that

$$H_{c2} = \frac{2m^* c}{\hbar e^*} \cdot b(T). \tag{2.3.3}$$

The solutions of the linearized Ginzburg-Landau equation are degenerate, forming a vector space of infinite dimension; therefore, the quartic part of the free energy functional will determine which subspace of this space of solutions can be realized as H drops below H_{c2}. It has been calculated numerically that as H approaches H_{c2} from below, the vortices arrange themselves in a triangular lattice, whose spacing can be determined from H_{c2}.

2.4. Abelian Higgs Model

In the Abelian Higgs model a complex scalar field is minimally coupled to a $U(1)$ gauge field. The theory is described by the Lagrangian

$$\mathcal{L} = [\mathcal{D}_\mu\phi]^\dagger[\mathcal{D}_\mu\phi] - \lambda[(\phi^\dagger\phi) - v^2]^2 - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} \tag{2.4.1}$$

where $\mathcal{D}_\mu = [\partial_\mu + igA_\mu]$. The states that minimize the energy of the classical field theory consist of solutions of the form $A = 0$ and $\phi(x) = v\ \mathrm{e}^{i\phi}$, where $0 \leq \phi < 2\pi$.

We consider vertex solutions of the form

$$\phi(r) = v \cdot \mathrm{e}^{-im\theta} \cdot F(r)$$
$$A_\theta(r) = \frac{m}{g} \cdot \frac{1}{r} \cdot G(r) \tag{2.4.2}$$

where the winding number m is an integer. These solutions are further subject to the boundary conditions that $F(0) = G(0) = 0$ and that $F(r), G(r) \to 1$ as $r \to \infty$. The differential equations for $F(r)$ and $G(r)$ are

$$-\frac{1}{r}\frac{d}{dr}\left[r\frac{dF}{dr}\right] + g^2G^2F = F - F^3$$
$$\frac{d}{dr}\left[\frac{1}{r}\frac{d}{dr}(rG)\right] = gGF^2. \tag{2.4.3}$$

Exact numerical solutions have been obtained; however, it is possible to calculate the approximate character of the vortex solutions by minimizing the energy of a vortex solution expressed as a function of the core size r_C. The core size will lie somewhere between the inverse scalar mass and the inverse vector boson mass.

One might wonder whether vortices continue to exist in the quantum theory in which there is a Higgs condensate or whether, alternatively, quantum fluctuations and radiative effects might not render the vortex solutions unstable. The answer is, almost without doubt, that vortices are stable in the Higgs phase of the quantized theory. The stability of the vortex solutions is quite robust because to destroy a vortex it is necessary to alter the fields at infinity.

2.5. Type I and Type II Abelian Vortices

In a certain crude sense, a vortex can be thought of as an isolated island of material in the normal phase carrying a magnetic flux surrounded on all sides by material in the superconducting phase from which all magnetic flux has been expelled. The interface between the two regions contributes a surface energy roughly equal to its area multiplied by the surface tension σ calculated in the previous section. This analogy suggests that there is a profound, qualitative difference between the properties of Type I and Type II vortices.

For Type I vortices there is an attractive interaction at close distances when the cores start to touch, and for vortices that carry a large amount of flux, the energy per unit flux is smaller. Since $\xi >> \lambda$, the magnetic flux resides in the inner part of the core or core surface. So when vortices come close to each other so that the cores begin to overlap, the energy can be lowered if the vortices coalesce, smoothing the variation in $|\Psi|$. By trying to form a single vortex, the pair reduces its surface area and thus also its total energy.

On the other hand, in Type II vortices since there is a negative surface energy, only the vortices carrying a flux of unit magnitude are stable. Vortices repel because when they are brought close together, the magnetic fields interfere in a way that increases the total energy.

The qualitative discussion here is confirmed by numerical experiment. [8.] In fact, at $\kappa = 1/\sqrt{2}$ the vortices are found not to interact at all. It has been shown rigorously using a special case of the Atiyah-Singer index theorem that in this case n-vortex solutions of minimal energy have a $2n$-dimensional degeneracy. [9.]

2.6. Vortices, Discrete Symmetry, and the Aharonov-Bohm Effect

In the Abelian Higgs model, if the charge of the Higgs field is equal to the quantum of electric charge e, then there is no long-range interaction between the vortices and the charged particles of the theory. The quantum of magnetic flux for a particle of charge $q = me$ is

$$\Psi(q) = \frac{2\pi\hbar}{qc} = \frac{1}{m} \cdot \left(\frac{2\pi\hbar}{ec}\right) = \frac{1}{m} \cdot \Phi_0 \tag{2.6.1}$$

where Φ_0 is the quantum of flux with respect to a particle that carries a charge equal to the quantum of charge e. Therefore, a vortex with winding number M will carry an integral number of flux quanta defined with respect to a particle with charge q, and there will be no Aharonov-Bohm scattering.

However, if the charge of the field that condenses is a non-unit integral multiple k of the charge quantum e, then vortices will carry flux in multiples of Φ_0/k, and charged particles whose charge is not divisible by k will experience Aharonov-Bohm scattering. This Aharonov-Bohm scattering can be used to measure the flux of the vortex mod k, or alternatively the charge of the particle mod k. The unbroken symmetry group is Z_k and its charge is measurable. The existence of such discrete charge and operators for measuring it have been discussed. [10,11.] Another interesting property of such vortices is the possibility to construct particles that obey fractional statistics. [12, 13.]

Notes.

1. V. L. Ginzburg and L. D. Landau, "On the Theory of Superconductivity," Zh. Eksperim. i Teor. Fiz. **20,** 1064 (1950). [English translation: D. ter Haar, *Collected Papers of L. D. Landau,* New York: Gordon and Breach (1965).]

2. H. B. Nielsen and P. Olesen, "Vortex-Line Models for Dual Systems," Nucl. Phys. **B61,** 45 (1973).

3. J. Bardeen, L. N. Cooper and J. R. Schrieffer, "Theory of Superconductivity," Phys. Rev. **96,** 1175 (1957).

4. L. P. Gor'kov, "Microscopic Derivation of the Ginzburg-Landau Equations in the Theory of Superconductivity," Zh. Eksperim. i Teor. **36,** 1918 (1959). [English translation: Soviet Physics–JETP **9**, 1364 (1959).]

5. We call H_c, defined by the equation (2.2.6), the "thermodynamic" critical field to distinguish it from the actual experimentally observed critical field, which is the same for Type I superconductors but differs for Type II superconductors.

6. A. A. Abrikosov, "On the Magnetic Properties of Superconductors of the Second Group," Zh. Eksperim. i Teor. **32,** 1442 (1957). [English translation: Soviet Physics–JETP **5**, 1174 (1957).]

7. This relation is derived by setting the force on the surface separating the normal and superconducting regions equal to zero. For the total free energy to be stationary with respect to translations of the surface that increase the volume of the superconducting phase, it is necessary that the decrease in free energy due to the increase in volume of the superconducting phase be exactly balanced by the increase in magnetic free energy resulting from the expulsion of magnetic flux required to increase the volume of the superconducting region. Consequently,

$$\delta V \cdot [f_{s0} - f_{n0}] = \frac{1}{4\pi} \int d^3x \mathbf{H}_0 \cdot (\delta \mathbf{B}) = \frac{1}{4\pi} H_0 \ \Phi \ \delta L = \frac{1}{4\pi} H_0^2 \ \delta V.$$

Therefore,

$$\left[\frac{H^2}{4\pi} + f_{s0} - f_{n0}\right] \cdot \delta V = 0.$$

8. L. Jacobs and C. Rebbi, "Interaction Energy of Superconducting Vortices," Phys. Rev. **B19,** 4486 (1979).

9. E. Weinberg, "Multivortex Solutions of the Ginzburg-Landau Equations," Phys. Rev. **D19,** 3008 (1979).

10. L. Krauss and F. Wilczek, "Discrete Gauge Symmetry in Continuum Theories," Phys. Rev. Lett. **62,** 1221 (1989).

11. J. Preskill and L. Krauss, "Local Discrete Symmetry and Discrete Quantum Mechanical Hair," Caltech Preprint 68-1601 (1989).

12. F. Wilczek, "Quantum Mechanics of Fractional Spin Particles," Phys. Rev. Lett. **49,** 957 (1982).

13. F. Wilczek, "Magnetic Flux, Angular Momentum and Statistics," Phys. Rev. Lett. **48,** 1144 (1982).

Chapter 3
Some Remarks on Non-Abelian Vortices

3.1. Introduction

In this section the classification of non-Abelian vortices and some of the differences between Abelian and non-Abelian vortices are discussed. Like Abelian vortices, non-Abelian vortices have cores of finite size in which their mass and magnetic flux are concentrated. Outside the core the difference of T_{00} from its vacuum expectation value and F^a_{ij} decay exponentially, and a local observer who is far away and unable to travel around the vortex can see nothing but a uniform vacuum. Whereas the last section discussed in quite a bit of detail the energetics and core structure of Abelian vortices, this section and the following sections shall concentrate almost exclusively on the asymptotic structure of the non-Abelian vortices, idealizing vortices as pointlike objects without any internal structure. Most of the discussion about vortices can be extended in a straightforward manner to cosmic strings in three dimensions. When such an extension is obvious, we shall just discuss the case for vortices. When vortices are idealized as pointlike objects without internal structure, one can equivalently remove the pointlike singularity from the spatial manifold, creating a multiply-connected spatial manifold. The approximation that vortices are pointlike is valid at energies much lower than the symmetry breaking scale because the vortex size is approximately equal to the inverse of that scale.

Except for corrections that decay exponentially with distance from the vortex, far from the vortex, $|\Psi| = \Psi_0$, $(\mathcal{D}\Psi) = 0$, and $F_{\mu\nu} = 0$. Therefore, in any compact region that does not enclose any vortices, it is possible to choose a so-called "unitary gauge" in which $\Psi(x) = \Psi_0$ and $A_\mu = 0$. However, if we consider an annular region ($r_a < r < r_b$) that encloses a vortex at $r = 0$, no such gauge choice is possible. If one places a cut across the annulus making it simply connected, then it is possible to make a gauge transformation to a unitary gauge; however, at the cut the gauge transformation must be singular and A_μ would be infinite there.

3.2. Magnetic Flux and the Classification of Vortices

In the introduction it was described how topological defects may be classified by enclosing the defect with a sphere S^n and determining the homotopy class of the mapping from that sphere into the manifold of classical vacua $M \cong G/H$. For vortices in two spatial dimensions and the strings in three dimensions, the relevant homotopy group is $\pi_1(M, x_0)$, also called the fundamental group or the Poincaré group. [1.] A gauge is chosen so that at some arbitrary point x_0 far from the core $\Psi(x_0) = \Psi_0$. A path C starting and ending at x_0 and enclosing the vortex exactly once in the counterclockwise direction induces a mapping $f : S^1 \to G/H$, whose homotopy class is invariant under continuous deformations of C that do not make C cross through the core of the vortex. When the homotopy class of f is non-trivial, the stability of the vortex is insured by a topological conservation law. It is not possible for a local process to make the vortex disappear because the vortex has "hair" at spatial infinity. Viewed from another perspective, a process that lowers the energy of a state by destroying a vortex would have to tunnel through an infinite potential barrier of finite thickness and, hence, cannot take place, at least not at any order of the semi-classical approximation because in the Euclidean field theory there is no trajectory with finite action that connects the two states.

Consider a field theory consisting of a Higgs field Ψ minimally coupled to a gauge field whose symmetry group is G and that is described by the Lagrangian

$$\mathcal{L} = \frac{1}{2}(\mathcal{D}_\mu \Psi)(\mathcal{D}^\mu \Psi) - V[\Psi] - \frac{1}{4} F^a_{\mu\nu} F^{a\mu\nu}. \tag{3.2.1}$$

There may be other additional fields, but because they do not condense, they are not relevant to the discussion here. Here the potential V is symmetric under G in the sense that $V[g\Psi] = V[\Psi]$. Since there is spontaneous symmetry breaking, V takes a minimum at $\Psi \neq 0$. We shall assume that there is no accidental degeneracy, meaning that if $V(\Psi) = V(\Psi_0)$, then for some $g \in G$, $\Psi = g\Psi_0$. Define the manifold of classical vacua M as the set of Ψ for which V is minimum. Let $H_0 = \{h \in G | h\Psi_0 = \Psi_0\}$. Then the subgroups $H(\Psi_0)$ and $H(\Psi)$ are related by conjugation so that $H(\Psi) = gH(\Psi_0)g^{-1}$ because $\Psi = g\Psi_0$.

For this model, determining whether there exist stable vortices is equivalent to determining whether the fundamental group $\pi_1(G/H)$ is non-trivial. At this point there arises an ambiguity as to what is the correct Lie group G of the theory. A gauge field theory is usually described in terms of its Lie algebra $\mathcal{G}$, the fields of the theory transforming according to various representations of $\mathcal{G}$. Corresponding to a particular Lie algebra $\mathcal{G}$, there are in general several distinct Lie groups that are locally isomorphic (because they share the same Lie algebra) but not globally isomorphic. Among these locally isomorphic groups there exists a largest group $\tilde{G}$ that is simply connected and is the universal covering group of each of the other locally isomorphic groups. Each of the locally isomorphic G may be formed from $\tilde{G}$ by dividing out a subgroup of the center of $\tilde{G}$, which we shall denote $Z(\tilde{G})$.

We are only interested in compact Lie groups because since $\mathrm{tr}[F_{\mu\nu}F^{\mu\nu}]$ must be positive definite in Euclidean space to get an energy that is bounded from below), we need a positive definite, group invariant quadratic form. It suffices to consider the Lie groups isomorphic to the simple compact Lie algebras. Any compact Lie group G can be expressed as a direct product of $U(1)$ factors and simple Lie groups so that [2.]

$$G \cong \left[U(1)\right]^j \times G_1 \times \ldots \times G_k. \tag{3.2.2}$$

The compact simple Lie algebras consist of the special unitary groups $SU(n)$ for $n > 2$, the two series of orthogonal groups $SO(2n)$ and $SO(2n+1)$ for $n \geq 1$, and the five exceptional groups. The special unitary groups $SU(n)$ are simply connected in their fundamental representations. By Schur's lemma, if $M \in Z[SU(n)]$ then $M = cI$, implying that $c^n = 1$. Therefore, the elements of the center are the unit matrix I multiplied by the n roots of unity, so that $Z[SU(n)] \cong Z_N$. $SO(2n+1)$ is doubly connected, its covering group being $\mathrm{Spin}(2n+1)$. By Schur's lemma, the center of $SO(2n+1)$ is trivial; therefore, $Z[\mathrm{Spin}(2n+1)] \cong Z_2$. For $SO(2n)$ where $n > 2$ the group $\mathrm{Spin}(2n)$ in the universal covering group, and the projection mapping $\pi : \mathrm{Spin}(2n) \rightarrow SO(2n)$ is 2 to 1. However, since $2n$ is even, $-I \in Z[SO(2n)]$.

Therefore, $Z[\text{Spin}(2n)]$ is isomorphic to Z_4. The only groups that remain are the exceptional groups: G_2, F_4, E_6, and E_8.

We shall choose $G = \tilde{G}$ so that G is simply connected. Now we are prepared to determine $\pi_1(M, x_0)$. Since $M \cong G/H$, we shall consider paths in G that start and end in H_0. Because G is simply connected, paths that start and end in the same component of H_0 can be deformed into the trivial (constant) path. Therefore, $\pi_1(M, x_0) \cong \pi_0(M, x_0)$, whose elements correspond to the connected components of H_0.

The classification of vortices according to homotopy classes of

$$f : S^1 \to G/H \tag{3.2.3}$$

where $f(\theta) = \Psi(Re^{i\theta})$ for $(R > R_{core})$ determines which vortices cannot be deformed into each other or into the vacuum in a continuous manner; however, this classification is incomplete because there is also the gauge field A_μ chosen so that $(\mathcal{D}_\mu \Psi) = 0$. When H has a continuous part, there are several ways to choose this gauge field, and a finer classification of vortices is needed.

Consider the operator

$$h(C, x_0) = P[\exp(-ig \int_{(C,x_0)} dx^\mu A_\mu)] \tag{3.2.4}$$

where P indicates path ordering. The operator measures non-Abelian magnetic flux and transforms under gauge transformations as

$$h(C, x_0) \to \Omega(x_0) h(C, x_0) \Omega(x_0)^{-1}. \tag{3.2.5}$$

If we select a component of H not connected to the identity, then we can fix $\Psi(x)$ around the vortex. To minimize the energy $A_\mu(x)$ must be chosen so that $\mathcal{D}_\mu \Psi = 0$. However, because there is residual continuous symmetry, there are a variety of ways to choose A_μ that will result in different values of $h(C, x_0)$ within this connected component.

These different values of $h(C, x_0)$ within the connected component correspond to different types of vortices. However, there does not necessarily exist a stable vortex for every value of h within a connected component of H because no topological conservation law prevents h from moving within a component of H. But if H_i acting on an element h by conjugation creates a non-trivial orbit, then at the classical level all of these vortices will be either stable or unstable and will have the same masses and excitation spectra. However, as is discussed in Chapter 5, quantization lifts this degeneracy.

3.3. Global Non-Realizability of the Unbroken Symmetry Group

A property peculiar to vortices with non-Abelian flux [non-Abelian meaning here that $h(C, x_0)$ does not lie in the center of $H(x_0)$] is that in the presence of such vortices it is not possible to realize the unbroken symmetry H globally in the spatial manifold with the cores of the vortices excluded. Instead of speaking of the unbroken symmetry group H as an abstract group without taking into account its embedding in G, we have been careful to specify the point at which H is the unbroken symmetry group. Thus $H(x)$ is the group that stabilizes $\Psi(x)$, and is well-defined as long as x is not one of the core singularities. The vector field A_μ provides a natural connection between $H(x)$ and $H(y)$ dependent only on the homotopy class of the path that connects x to y.

In a background with no vortices it is always possible to define a mapping that maps $\mathcal{M} \times H$ into the various $H(x)$, where $\mathcal{M}$ is the spatial manifold with the vortex core singularities excluded. We shall call this mapping a global realization of the unbroken symmetry H. [3,4.] However, in the presence of vortices carrying flux that does not lie in the center of H, no such continuous mapping can be defined because of a topological obstruction. Physically, the difficulty lies in the fact that a global gauge transformation would change the flux of the vortex.

Consider a background with a single vortex carrying flux $h(C, x_0)$. If we take $H(x_0)$ and parallel transport its elements around the vortex along C (or any other path homotopic to C), then the induced automorphism

$$h(C,x_0) : H(x_0) \to H(x_0) \tag{3.3.1}$$

maps $\hat{h} \mapsto h(C,x_0)\ \hat{h}\ h^{-1}(C,x_0)$.

Let us consider the action of the automorphism on the continuous part of the symmetry H. Suppose that H_c has k continuous generators: $T_1,\ldots,T_k$. Then we can normalize the generators so that $< T_a, T_b >= \delta_{ab}$ where $<\ ,\ >$ indicates the Cartan or Killing form. Define the matrix M_{ab} by the equation

$$T_a M_{ab} = h(C,x_0) T_b h(C,x_0)^{-1}. \tag{3.3.2}$$

$M \in O(k)$ since the Killing form is invariant; therefore, M can be diagonalized with a unitary transformation. The generators with unit eigenvalues correspond to the part of the unbroken symmetry group that can be globally realized.

Notes

1. J. Preskill, "Vortices and Monopoles," Les Houches Lecture given in 1984 in P. Ramond and R. Stora, Eds., *Architecture of the Fundamental Interactions at Short Distances*. Amsterdam: North Holland (1987).

2. L. S. Pontryagan, *Topological Groups*, New York: Gordon and Breach (1966).

3. M. G. Alford, K. Benson, S. Coleman, J. March-Russell and F. Wilczek, "The Interactions and Excitations of Non-Abelian Cosmic Strings," Harvard Preprint HUTP 89-A051 (1990).

4. J. Preskill and L. Krauss, "Local Discrete Symmetry and Quantum Mechanical Hair," Caltech Preprint 68-1601 (1989).

Chapter 4—Discrete Non-Abelian Vortices

4.1. Introduction

Recently there has been much interest in "local discrete symmetry," which arises when a gauge theory based on the symmetry group G has a vacuum containing a Higgs condensate $< \phi > \neq 0$ that breaks G to a subgroup H in such a manner that H is either a discrete group or a continuous group with more than one connected component. [1,2,3,4.] When H is discrete, all of the gauge bosons acquire a mass of the same order as the symmetry breaking scale and become irrelevant at low energies. In the effective field theory that describes physics in the absence of vortices at energy scales much below the symmetry breaking scale, H becomes a discrete symmetry of the effective Lagrangian. One of the few remnants of the fact that the discrete symmetry H resulted from the breaking of a continuous, local symmetry is the existence of stable vortex excitations, which in 2+1 dimensions are pointlike objects and cosmic strings in 3+1 dimensions. In the Higgs phase whenever H (considered as a subgroup of the simply-connected gauge group G) has a discrete part, there exist stable vortices.

This paper discusses some of the peculiarities that result if H is non-Abelian. When H is Abelian, vortices do not interact except at short distances when their cores overlap. When H is non-Abelian, however, vortices experience a long-range, Aharonov-Bohm interaction that is a consequence of the topology of the vector potential A_μ that surrounds the vortices and the manner in which it is continuously deformed as vortices move. Aharonov-Bohm vortex-vortex scattering has already been noted by Wilczek and Wu, who calculated a scattering cross section for two vortices. [5.] In this paper we discuss the $N = 2$ case where N is the number of vortices using a different formalism and discuss the complications that result when $N \geq 3$. The presence of a third vortex introduces new features, even when the third vortex is distant from the two other vortices. When a vortex passes between two vortices, the relation between the two vortices is altered. Consequently, when a vortex is moved through a background with many vortices, the state of the background is changed in a way that depends on the homotopy class of the path, and it is not possible to consider the background as a fixed, classical object. One must consider the whole system quantum mechanically.

The organization of the paper is as follows. In Section 4.2 vortices are regarded as a classical background upon which the light fields propagate. The Aharonov-Bohm scattering of charged particles by vortices is discussed as a means of performing measurements on vortices to determine the part of the global structure of the $< A_\mu >$ background that cannot be gauged away. To make the discussion concrete a specific model in which SU(2) breaks to the eight element quaternion group Q is introduced. In Section 4.3 vortex-vortex interactions are studied, first by examining how a classical vortex background evolves as the singularities are moved, the motion continuously deforming the surrounding classical gauge field. The fact that final classical field configuration depends not only on the final positions of the singularities, but also on the homotopy class of the path used to move the singularities to their final positions necessitates considering quantum superpositions of gauge inequivalent A_μ backgrounds that surround the singularities. A quantum theory is thus derived in an ad-hoc manner. In Section 4.4 a quantum theory of the dynamics of non-Abelian vortices is developed in a more formal and systematic way based on a path integral. The classical vortex backgrounds from Section 4.2 serve as a basis for the quantum states of a vortex system. In the concluding remarks the exotic statistics of non-Abelian vortices and the similarities of the many non-Abelian vortex system to the many-anyon system are discussed.

4.2. Physics in a Classical Vortex Background

In the Higgs phase of a gauge field theory in the absence of vortices or in a simply connected domain containing no vortices, it is always possible to choose a unitary gauge in which the expectation value of the Higgs scalar field $< \phi >$ is constant everywhere. In this section we shall treat the expectation values of ϕ and A_μ as classical fields because the quantum fluctuations of these fields are not essential to the physics considered here. We consider the case in which H is finite and non-Abelian.

As a specific example, consider the SU(2) gauge theory with a real Higgs field that transforms under the five-dimensional real representation. [2] We may represent the Higgs field as a real, traceless, symmetric 3×3 matrix transforming as

$$M \to OMO^{-1} \tag{4.2.1}$$

where O is the SO(3) element corresponding to an SU(2) element. If M acquires a non-zero vacuum expectation value

$$< M >= \mathrm{diag}[\lambda_1, \lambda_2, -(\lambda_1 + \lambda_2)] \tag{4.2.2}$$

where none of the eigenvalues are equal, then the group H considered as a subgroup of SO(3) has the structure of the Abelian group $Z_2 \times Z_2$ of 180° rotations about the three axes, but when lifted to SU(2) becomes the eight element quaternion group **Q** generated by the elements $\hat{\imath}, \hat{\jmath}$, and $\hat{k}$, which satisfy the relations $\hat{\imath}^2 = \hat{\jmath}^2 = \hat{k}^2 = -1$, $\hat{\imath}\hat{\jmath} = -\hat{\jmath}\hat{\imath} = \hat{k}$, and the two additional relations that result from cyclically permuting $\hat{\imath}, \hat{\jmath}$, and $\hat{k}$ in the previous relation.

Because the continuous part of the gauge group is completely broken, all of the gauge bosons acquire a large mass and become irrelevant at energy scales of interest here. Whenever a group G breaks to a discrete group or a group with several connected components (which can be written as a semi-direct product of a finite group and a connected Lie group), there exist stable vortices. There is an ambiguity in defining the Lie group G. Usually a field theory is defined in terms of its Lie algebra $\mathcal{G}$, which acts upon the fields of the theory through various representations of $\mathcal{G}$. Corresponding to a given Lie algebra $\mathcal{G}$, there exist in general several distinct Lie groups that are locally isomorphic, but have different global structures and hence are not globally isomorphic. Among these different locally isomorphic Lie groups there exists a unique group that is simply connected. It is the largest of these groups, and all of the other locally isomorphic Lie groups can be formed from this group by dividing out a discrete subgroup of the center. In the sequel G shall always denote the simply connected group. In the previous example SU(2) is simply connected. Its center is Z_2 and $SO(3) = SU(2)/Z_2$. When G is simply connected, the elements of $H \subset G$ not equal to the identity identify the distinct vortices of the theory.

Before discussing the general case of N vortices, we consider a single vortex in two spatial dimensions. In the presence of a vortex it is no longer possible to choose a unitary gauge because the Higgs field ϕ cannot be continuously deformed to a value constant everywhere. This topological fact is what provides stability to the vortex. We shall ignore the core of the vortex and idealize it as a point singularity. The size of the vortex core is of the order of the inverse symmetry breaking scale, making the vortex almost pointlike at low energies. Outside of the core A_μ is pure gauge because $F_{\mu\nu} = 0$ to minimize the magnetic energy. Locally, but not globally, A_μ can be set to zero by an appropriate choice of gauge. For an oriented, closed curve C starting and ending at x_0 and enclosing the vortex, we define the operator

$$h(C, x_0) = P[exp(i \int_{C,x_0} dx^i A_i)]. \tag{4.2.3}$$

This integral is invariant under continuous deformations of C that keep the basepoint x_0 fixed and avoid the singularities at the vortex cores. $h(C, x_0) \in H(x_0)$ because $h(C, x_0)$ acting on the Higgs field parallel transports $\phi(x_0)$ around the curve back to its original value.

Now suppose that there are N vortices fixed at the positions $x_1, x_2, ..., x_N$ in the spatial plane $\mathcal{R}^2$ and choose a basepoint x_0 from which to start and terminate paths. $h(C, x_0)$ is defined as before. If we define $M = M(x_1, x_2, ..., x_N)$ as the punctured plane consisting of $\mathcal{R}^2$ with the points $x_1, x_2, ..., x_N$ excluded, then the elements of the fundamental group $\pi_1(M, x_0)$ are homotopy equivalence classes of closed curves based at x_0, and the integral is well-defined on these equivalence classes. The mapping

$$h : \pi_1(M, x_0) \to H(x_0) \tag{4.2.4}$$

is a homomorphism between the groups. As will be shown, the homomorphism h completely describes the vortex background except for a physically irrelevant choice of gauge, and for every mapping h between the groups that is a homomorphism there exists a vortex configuration described by the mapping h. For N punctures $\pi_1(M, x_0)$ has the structure of the free group generated by N elements, which we shall denote F_N. To choose a set of generators of π_1, we select N paths that enclose each of the N vortices. When $N > 1$ there is no natural way to choose the path from x_0 that encircles a particular vortex because the different ways to thread a path between the other vortices produce homotopically inequivalent paths. For example, in Figure 4.1 there are two vortices A and B. A basepoint x_0 is chosen. The paths a and a' indicated in Figure 4.1(a) and 4.1(b), respectively, are homotopically inequivalent. If the basis for the fundamental group $\pi_1(M, x_0)$ indicated in Figure 4.1(c) consisting of α and β is chosen, then $a = \alpha$ and $a' = \beta\alpha\beta^{-1}$. The different threadings a and a' around the vortex A are related by a conjugacy transformation. More generally, if $\gamma \in \pi_1(M, x_0)$, then α and $\gamma\alpha\gamma^{-1}$ merely represent different threadings. The class of equally valid threadings around a vortex is a conjugacy class in F_N, which contains an infinite number of elements. If H were Abelian, then this inequivalence would be irrelevant because the inequivalent paths would map into the same element of $H(x_0)$; however, such is not the case when H is non-Abelian. The structure of F_N is described in the following manner. Free group elements consist of strings of the generators raised to integer powers. The only relation available to simplify strings is the trivial relation: $g^m g^n = g^{m+n}$. With this relation strings can be simplified to a minimal form. Two strings with different minimal forms describe distinct group elements. F_N is of infinite order and is the largest group generated by N elements. All other groups generated by N elements are isomorphic to F_N/S where $S \subset F_N$ is a normal subgroup. To describe a homomorphism of F_N into some group it suffices to specify into which elements a set of generators of F_N are mapped.

Although the assignment of a group element to a vortex depends on the threading of the path surrounding the vortex from the basepoint to the vortex, the conjugacy class of group elements to which the path is assigned is invariant both with respect to the threading and the choice of basepoint. Different threadings of paths surrounding just one vortex in $\pi_1(M, x_0)$ lie in the same conjugacy class; therefore, their images in $H(x_0)$ lie in the same conjugacy class. Next we consider invariance with respect to basepoint. Suppose that z_1 and z_2 are two basepoints. The association between $H(z_1)$ and $H(z_2)$ is generally path dependent. Let C_a and C_b be homotopically inequivalent paths from z_1 to z_2, and let h^* be the magnetic flux contained in the closed path $C = C_b{}^{-1} C_a$. The isomorphisms $\rho_a : H(z_1) \rightarrow H(z_2)$ and $\rho_b : H(z_1) \rightarrow H(z_2)$ defined by the integral

(4.2.5) $$\rho = P[exp(i \int_{z_1}^{z_2} dx^i A_i)]$$

are obtained by integrating along C_a and C_b, respectively. The vector potential A_i that appears in the integral is in the adjoint representation because it is the adjoint representation that acts on the group elements. The isomorphisms are equivalent iff $h^* \in Z(H(z_1))$, which is the center of the group $H(z_1)$. However, even when the homomorphisms are inequivalent, the mapping of conjugacy classes into conjugacy classes is path independent. When there exist vortices containing flux not in the center of H, the fact that the isomorphisms are path dependent means that the symmetry H cannot be realized globally. In this case there exists no smooth mapping from the abstract group H into the various $H(x_0)$. [3.]

The Wilson loop operator can be used to identify the conjugacy class of a vortex. Unlike $h(C, x_0)$, the Wilson loop operator is gauge invariant, but representation dependent. $W^\alpha(C)$ and $h(C, x_0)$ are related in the following manner. Let the index α label the inequivalent, irreducible representations of H and $\Gamma^\alpha(h)$ be the matrix into which h is mapped by the representation α. Then the Wilson loop is the character

(4.2.6) $$W^\alpha(C) = Tr[\Gamma^\alpha(h(C, x_0))].$$

For a closed loop C, measuring $W^\alpha(C)$ for all representations uniquely determines the conjugacy class of the flux enclosed by C but does not determine $h(C, x_0)$ more exactly. This is because the number of conjugacy classes is equal to the number of inequivalent, irreducible representations, and the orthogonality theorem for characters insures that knowledge of all the characters of a group element determines its conjugacy class uniquely.

To identify the physical states of the pure vortex system, one must consider the effect of gauge transformations and associate configurations differing by a gauge transformation with the same physical state. Let $\Omega : \mathcal{R}^2 \rightarrow G$ be a smooth mapping. Then the mapping h transforms as

$$h(C, x_0) \rightarrow \Omega(x_0) h(C, x_0) \Omega(x_0)^{-1}. \tag{4.2.7}$$

To fix the gauge we fix $\phi(x_0)$ at the basepoint. (We can choose the corresponding unitary gauge in a open, simply-connected region containing x_0 and imagine an observer living in that region who ventures out to discover the global structure of the space.) This fixes $H(x_0)$, which is the embedding of the abstract group H in G. It might appear that there still remains a residual freedom to perform discrete gauge transformations lying in $H(x_0)$, but this is the case because in reality there are other fields that transform non-trivially under $H(x_0)$ and we want to be able to distinguish between different relations between these fields and the vortex background.

Among the paths in $\pi_1(M, x_0)$, the path surrounding all of the vortices plays a special role. This path measures the combined flux of all of the vortices at spatial infinity. Since no local operator can change the field at infinity without violating causality, the conservation of total flux is a superselection rule.

We now discuss Aharonov-Bohm scattering of charged particles by fixed vortices. [6.] The light fields of the theory are grouped into multiplets that transform irreducibly under the large group G. Under the remaining unbroken symmetry H these multiplets are, in general, reducible, and the fields can be further reduced into multiplets that transform irreducibly under H. Since H is finite, there exist a finite number of inequivalent, irreducible representations of H equal to the number of conjugacy classes of H. These representations can be divided into two classes:

the one-dimensional representations and the multidimensional representations. An Abelian group has only one-dimensional representations. A non-Abelian group has at least one representation of dimension greater than one.

The representations of the quaternion group are the following. There exist four one dimensional representations—a trivial representation $\Gamma^{(1)}$ mapping all group elments into 1, a representation, which we shall call $\Gamma^{(i)}$, such that $\hat{\imath} \mapsto 1, \hat{\jmath} \mapsto -1, \hat{k} \mapsto -1$, and $\Gamma^{(j)}$ and $\Gamma^{(k)}$, which are defined analogously. There also exists a two-dimensional representation, which we shall call $\Gamma^{(2)}$, that maps the generators of $\mathbf{Q}$ as follows

$$\hat{\imath} \mapsto -i\sigma_1, \ \hat{\jmath} \mapsto -i\sigma_2, \ \hat{k} \mapsto -i\sigma_3 \tag{4.2.8}$$

where σ_i are the Pauli matrices.

To describe the propagation of particles, one must sum over paths. Each path is given a kinetic factor, the complex phase obtained by integrating the Lagrangian in the absence of the gauge field A_μ along the path. Suppose that the particle propagates from P to P'. The respective multiplets Ψ at P and Ψ' at P' reside in different vectors spaces, which we shall call $\mathcal{V}(P)$ and $\mathcal{V}(P')$, respectively. The background gauge vector potential A_μ defines a connection between the vector spaces. For each homotopy class of curves from P to P' there exists a connection from $\mathcal{V}(P)$ to $\mathcal{V}(P')$ obtained by integrating A_μ. To calculate the propagator $K(P,t;P',t') : \mathcal{V}(P) \to \mathcal{V}(P')$, one must form a sum over homotopy classes of curves from P to P'. The terms in this sum consist of the transformation from $\mathcal{V}(P)$ to $\mathcal{V}(P')$ corresponding to parallel transport along a path in the homotopy class multiplied by the ordinary path integral restricted to paths lying in the homotopy class.

For one-dimensional representations, the particle–vortex scattering obtained is completely analogous to the scattering of electromagnetically charged particles by infinitely thin tubes of magnetic flux. For a single flux tube the only parameter of importance is the relative phase $e^{i\xi}$ between the two paths. The scattering cross section is the same as that calculated by Aharonov and Bohm [7]

(4.2.9) $$\frac{d\sigma}{d\theta} = \frac{1}{k} \cdot \frac{\sin^2 \xi}{2\pi} \cdot \frac{1}{\sin^2(\theta/2)}$$

where k is the momentum of the incident beam in units for which $\hbar = 1$.

For particles in representations of dimension greater than one the situation becomes more complicated. [8,9,10.] For a given vortex described by the group element h it is possible to choose a basis for the Ψ multiplet in which $\Gamma^\alpha(h)$ is diagonal. In this basis $\Gamma^\alpha(h)$ has the form $\Gamma^\alpha(h) = \text{diag}[e^{i\xi_1}, ..., e^{i\xi_D}]$ because the matrix is unitary. Traversing around the vortex (or, alternatively, moving between two points along paths in different homotopy classes) merely produces a complex phase. However, one cannot diagonalize all of the matrices of the representation simultaneously, for otherwise the representation would be reducible.

Physically, it is possible to decompose a beam of particles traveling at the same momentum in a fixed direction into its diagonal charges with respect to a given group element h. A diffraction grating can be constructed out of vortices by dividing the spatial plane in half by a line on which vortices of type h are placed at equally spaced intervals. A unitary gauge is chosen on each half plane. On the incident side the wave function is the plane wave

(4.2.10) $$\Psi(x) = \Psi_0 e^{ikx}$$

where a basis is chosen such that $\Gamma^\alpha(h)$ is diagonal. After the beam emerges on the other side of the grating, for the wave function of each component there is a different relative phase between adjacent gaps in the grating. This phase is equal to the corresponding diagonal entry in $\Gamma^\alpha(h)$. Consequently, each component is scattered in a distinct direction. Thus it is possible to separate the beam into its various charges with respect to the group element h.

For the quaternion group we consider a multiplet that transforms according to $\Gamma^{(2)}$. With a $\hat{k}$ grating the beam can be split into its $\hat{k}$ charges. Suppose that only one component is kept and then passed through a $\hat{\jmath}$ grating. Beams of $\hat{\jmath}$ and $-\hat{\jmath}$ charge emerge with equal intensity. Out of each of these beams both types of $\hat{k}$ charge can be regenerated by passing the beam of definite $\hat{\jmath}$ charge through another $\hat{k}$ grating. This situation is very analogous to the $K_0 - \bar{K}_0$ system.

The Aharonov-Bohm effect provides a means by which an observer situated at a basepoint x_0 can measure the mapping

$$h : \pi_1(M, x_0) \to H(x_0) \tag{4.2.11}$$

by Aharonov-Bohm scattering discretely charged particles off of the vortices in the background. We considered two mappings h' and h'' related for some $h^* \in H(x_0)$ by $h'' = h^* h' h^{*-1}$ to be distinct notwithstanding that they are connected by a gauge transformation because when charged light fields are added to the pure vortex system, we want to be able to distinguish different relations between the vortices and the charges.

In this paper we have used the fundamental group $\pi_1(M, x_0)$ to describe the gauge field background of a vortex configuration. Another method is to make the punctured plane simply-connected by introducing cuts. [2,5.] In this simply connected region a unitary gauge can be chosen, setting ϕ = constant and $A_\mu = 0$. Across the cuts A_μ becomes singular. Consequently, for covariant derivative $D_\mu = (\partial_\mu + igA_\mu)$ of the matter fields to be continuous across the cut, the matter fields must obey a discontinuous matching condition across the cut. The positions of the cuts are arbitrary because the cuts can be moved by singular gauge transformations. Whereas the method with cuts involves making a choice of gauge at every point, the method presnted here based on the fundamental group involves only making a choice of gauge at a single point x_0.

4.3. Vortex-Vortex interactions

In the previous section we described the physics of a vortex background in which vortices were placed at fixed locations in the spatial manifold. We saw that a classical vortex background with N vortices is described by specifying both the positions $\mathbf{x} = (x_1, x_2, ..., x_N)$ of the vortices and the mapping of the fundamental group of the resulting punctured space $M(\mathbf{x})$ to the unbroken symmetry group

$$h_x : \pi_1(M(\mathbf{x}), x_0) \to H(x_0) \tag{4.3.1}$$

with respect to some basepoint x_0. These classical vortex backgrounds will serve as a basis of quantum states for the Hilbert space describing the vortex system in the same way as the states $|x>$ serve as a basis for the Hilbert space of states of a non-relativistic particle moving in a single dimension. In this section it will be shown that when the singularities $\mathbf{x} = (x_1, x_2, ..., x_N)$ are moved along some continuous path to new positions $\mathbf{y} = (y_1, y_2, ..., y_N)$, the induced mapping describing the new configuration

$$h_y : \pi_1(M(\mathbf{y}), y_0) \to H(y_0) \tag{4.3.2}$$

that results from continuously deforming the potential A_i while keeping $F_{ij} = 0$ depends on the homotopy class of the path used to move from $\mathbf{x}$ to $\mathbf{y}$. The easiest way to determine $h_y : \pi_1(M(\mathbf{y}), y_0) \to H(y_0)$ is to drag a set of paths generating $\pi_1(M(\mathbf{y}))$ back to $\pi_1(M(\mathbf{x}))$ and to map corresponding generators into the same elements of $H(x_0)$. The dependence of the final configuration on the homotopy class of the path from $\mathbf{x}$ to $\mathbf{y}$ is what causes the Aharonov-Bohm effect for non-Abelian vortices.

Consider first the simplest case with only two vortices. In Figure 4.2(a) there are two vortices A and B and an observer at basepoint x_0. As indicated in Figure 4.2(b), the position of vortex B is fixed, and we move vortex A to a new position A' along two inequivalent paths labeled 1 and 2. Let a and b in Figure 4.2(c) be a set of generators of $\pi_1(M, x_0)$ for the initial configuration and a' and b' in Figure 4.2(d) be a set of generators of $\pi_1(M', x_0)$ for the final configuration. Suppose that the gauge field of the initial configuration maps a into h_a and b into h_b and that in the final configuration a' is mapped into $h_{a'}$ and b' into $h_{b'}$. If A is moved to A' along path 1, then

$$h_{a'} = h_a, \qquad h_{b'} = h_b. \tag{4.3.3}$$

However, if A is moved to A' along path 2, then

$$h_{a'} = (h_b h_a) h_a (h_b h_a)^{-1}, \qquad h_{b'} = (h_b h_a) h_b (h_b h_a)^{-1}. \tag{4.3.4}$$

Unless h_a and h_b commute, different final configurations result. When the system propagates quantum mechanically from the initial state shown in Figure 4.2(a), after some finite time if vortex A has moved to A', the flux at A' measured relative to the path a' will have non-zero probabilities of being either h_a or $h_b h_a {h_b}^{-1}$.

Next we allow A to wind around B an arbitrary number of times. With B fixed we allow A to move its new position A' via all possible homotopy classes of paths. The homotopy classes can be classified by a winding number around B. It suffices to consider what happens when A' moves around B back to its original position winding around n times in the clockwise direction. The group element a becomes conjugated by h_{ba}. We may write

$$h_a(n) = {h_{ba}}^n h_a(0) {h_{ba}}^{-n}. \tag{4.3.5}$$

Since H is finite, there must be a smallest n such that $h_a(n) = h_a(0)$. For the Aharonov-Bohm effect with two vortices it is convenient to choose a new basis for the $h_a(j)$. If one defines

$$q(j) = \frac{1}{\sqrt{n}} \sum_{k=0}^{n-1} \exp(i2\pi j \frac{k}{n})\ h_a(k), \tag{4.3.6}$$

then after winding M times, the state $q(j)$ acquires a phase $(e^{i2\pi\frac{j}{n}})^M$. The fact that this is a simple phase makes the situation completely analogous to that of an electromagnetically charged particle winding around a tube of magnetic flux. It is worth noting that one cannot state in a manner independent of the choice of position of the observer that A winding around B changes the state of the flux tube B. Suppose that we had chosen x_0 situated between B and A'. Then B would not become conjugated. From the point of view of an observer near a vortex, the motion of distant vortices has no local effect.

Before discussing more complicated situations, we first consider how the relation between two vortices is altered when a third vortex passes between them. To each homotopy class of paths connecting two vortices there is assigned a combined flux of the two vortices. Let A and B be two vortices and $\mathcal{C}$ be a representative path from a homotopy class of paths from A to B. Suppose that as indicated in Figure 4.3(a) $C(\mathcal{C})$ is a closed, counter-clockwise oriented path containing the point x_0 that encloses only $\mathcal{C}$. The combined flux of the vortices takes a value in the group $H(x_0)$. This flux could be identified independently of a basepoint up to a conjugacy class, which is always globally well-defined. Vortices A and B are compared along a path $\mathcal{C}$ before and after a vortex C passes between them cutting through $\mathcal{C}$ as indicated in Figure 4.3(b). Figure 4.3(c) shows a set of generators of the fundamental group of the punctured space. Initially the combined flux is $h(\alpha\beta)$. However, after C passes through $\mathcal{C}$, the combined flux becomes $h(\gamma\alpha\gamma^{-1}\beta)$. In the case where H is the quaternion group, suppose that $h(\alpha) = h(\beta) = \hat{\imath}$ and $h(\gamma) = \hat{\jmath}$. If A were moved to B along $\mathcal{C}$, the two vortices could, in principle, fuse to form a single vortex of flux -1. However, if C first passes through $\mathcal{C}$ before the vortices are brought together, then the two vortices could annihilate and radiate their mass away, or create any combination of vortices with total flux 1.

This phenomenon makes the situation in which $N \geq 3$ more complicated. When there are many vortices, it is not possible to quantize a single vortex and allow it to propagate through a classical vortex background experiencing Aharonov-Bohm scattering because the path taken by the vortex changes the background. For two vortices it is possible to regard one of the vortices as a classical object by fixing the basepoint an infinitesimal distance from the classical vortex, but when the number of vortices is increased, no such simplification is possible.

Let us consider again the two vortex scattering, but now in the presence of a third vortex. As indicated in Figure 4.4, A is allowed to move past B to A' along the same two paths as before, but now in the presence of vortex C. We draw a path $\mathcal{C}$ from B to C that is cut by path 2. Now let A propagate to its new position at A'. In general there will be non-zero amplitudes for propagation along both paths, and the flux of the vortex at A', which depends on the path taken, will be indeterminate. However, if we measure the relation between B and C along $\mathcal{C}$, which is the combined flux of the two vortices when the two vortices are brought together along path $\mathcal{C}$, then the flux of the vortex at A' is no longer indeterminate. Measuring the relation between B and C determines whether the path taken by A to A' cut the path $\mathcal{C}$ and, hence, whether path 1 or path 2 was taken. As a consequence of measuring this relation, the wave function of the vortex background is collapsed, giving the vortex A' a definite flux. For a system in which B is fixed, C is fixed far away, and A may wind around B an arbitrary number of times, the quantum states of the system are specified by the coordinates of the A vortex plus the flux of the A vortex and flux of the C vortex. For a given flux of one vortex, the flux of the other vortex is completely determined. If we want to consider just various windings of A to A' around B as before, we would have to sum over products of states of the A and the C vortices. We can measure the final state by measuring either A or C.

The quaternion group has two special properties that allow us to choose a basis for which homotopically inequivalent paths merely produce relative complex phases, even when many vortices are present. For the quaternion group, when the group acts on itself by conjugation, the resulting adjoint representation can be reduced to a sum of one-dimensional representations. Furthermore, since elements of the same conjugacy class are related by elements of the center of the group, elements of the new basis conjugate other elements in exactly the same manner. Set

$$|I^+> = \frac{1}{\sqrt{2}}[|+\hat{i}> + |-\hat{i}>],$$
$$|I^-> = \frac{1}{\sqrt{2}}[|+\hat{i}> - |-\hat{i}>]. \qquad (4.3.7)$$

and define $|J^+>$, $|J^->$, $|K^+>$, and $|K^->$ in an analogous manner. These superpositions correspond to various one-dimensional representations of H. All of the + representations are trivial. $|I^->$ transforms according to $\Gamma^{(i)}(h)$, $|J^->$ according to $\Gamma^{(j)}(h)$, and $|K^->$ according to $\Gamma^{(k)}(h)$.

4.4. Path Integral Formulation and the Braid Group

In this section we formulate a quantum theory for a pure vortex system using the Feynman path integral formalism. The path integral presented here provides a natural framework for discussing vortex-vortex scattering due to the Aharonov-Bohm interaction and describes the long-range physics of vortices exactly. However, because vortices are considered pointlike objects without internal structure or short-range interactions, short-range effects are ignored in this formulation. It is also assumed that the total number of vortices is conserved and all topologically allowed types of vortices are stable. In the conclusion we shall sketch how these defects can be remedied by introducing vertex interactions.

For a single vortex the path integral is trivial. The only parameter needed to describe the vortex dynamics is the vortex mass, and the path integral is exactly the same as the path integral for a free, non-relativistic particle of the same mass. There is no need to include a description of the surrounding gauge field in the wave function because the flux carried by the vortex, which is the flux measured at spatial infinity, is a constant of motion.

The Hilbert space of physical states for a system of N vortices $\mathcal{H}$ is constructed from classical vortex configurations described by specifying the position of the N vortices $\mathbf{x} = (x_1, ..., x_N)$ and the mapping $h : \pi_1 M(\mathbf{x}, x_0) \to H(x_0)$. (For reasons that shall soon become apparent, x_0 is chosen to be at spatial infinity in a certain direction.) Paths between configurations are specified by a classical trajectory $\mathbf{x}(t)$.

We first consider the propagation of two vortices. Suppose that in the initial configuration at $t = t_x$ two vortices are at x_1 and x_2 and that we want to calculate the amplitude for the vortices to propagate to positions y_1 and y_2 at $t = t_y$. Suppose that the group elements corresponding to x_1 and x_2 belong to different conjugacy classes. Hence we only need to consider paths that connect x_1 to y_1 and x_2 to y_2. Such paths can be classified into homotopy equivalence classes known as braidings, which can be labeled by an integer winding number. The winding number zero is assigned to an arbitrary braiding. By winding y_2 around y_1 N times in the counter-clockwise direction, the braiding N is created. In the descriptions of the initial and final states we specify $h_x : \pi_1(M(\mathbf{x}), x_0) \rightarrow H(x_0)$ and $h_y : \pi_1(M(\mathbf{y}), y_0) \rightarrow H(y_0)$, respectively. Let $\gamma_1, \gamma_2 \in \pi_1(M(\mathbf{y}), y_0)$ be paths enclosing only x_1 and x_2, respectively, in the counter-clockwise direction. If we define n as the order of the element $h_x(\gamma_1\gamma_2\gamma_1{}^{-1}\gamma_2{}^{-1})$, then every nth winding will result in the same homomorphism h_y; therefore, only every nth braiding should be included in the path integral sum for the propagator.

As the number of vortices increases, the complexity of the problem rapidly grows because the paths become hopelessly entangled. Braids can be classified by the transformations required to unwind them. For three paths with their endpoints fixed, one can first unwind one pair of paths, obtaining as before a winding number. The transformation required to unwind the third path is an element of the free group with two generators F_2. The braid group is a semi-direct product $Z \times F_2$ because conjugation of F_2 by elements of Z produces non-trivial automorphisms of F_2.

In summing the path integral, one must also consider statistics. In general, the path integral sum between given initial and final states contains not only different braiding of paths, but also braidings plus permutations that connect the points $x_1, ..., x_N$ to $y_{\pi(1)}, ..., y_{\pi(N)}$ where $\pi \in S_N$. Which vortices can be identical depends on what types of vortices are present in the system. For example, when H is the quaternion group, in a system with just $\hat{\imath}$ and $-\hat{\imath}$ vortices, these two kinds of vortices are distinguishable. However, after even just one $\hat{\jmath}$, $-\hat{\jmath}$, $\hat{k}$, or $-\hat{k}$ vortex is added, the $\hat{\imath}$ and $-\hat{\imath}$ vortices become identical upon winding around one of the other vortices an odd number of times. More generally, the vortices act on each other through conjugation. If vortices corresponding to all elements of H are present, then all vortices in a conjugacy class would be considered identical. If only a subset of the vortices in H are present, the orbits obtained by allowing this subset to act on H through conjugation are the classes of identical vortices of the system.

We now discuss the choice of basepoint. To describe the vortex configurations at t_i and t_f, it is necessary to choose basepoints z_i and z_f, respectively. The particular choices of these points are arbitrary and have no physical significance as far as the description of the pure vortex background is concerned. It was previously noted, however, that when the basepoint of the fundamental group is moved around a closed path, the induced automorphism of $H(x_0)$ is generally non-trivial. For similar reasons there is no natural correspondence between $H(z_i)$ and $H(z_f)$. To establish a correspondence between the two groups, one must specify a path from z_i to z_f. Physically, the path from z_i to z_f represents the trajectory of an observer carrying a set of test charges used to measure the vortex background. Since the charges carried by the observer are affected by the choice of braiding from z_i to z_f, one expects different paths to result in different correspondences between the groups.

If trajectories $\mathbf{x}(t)$ of interest are allowed to wind around the path between the basepoints in different ways, then additional complications are introduced. The complexity of the braid structures increase as if another vortex had been added. Fortunately, if one is interested only in the pure vortex background, the additional complication can be avoided by placing the basepoint near spatial infinity so that the vortices never braid around the path connecting the basepoints.

4.5. Concluding Remarks

In this paper we have calculated two-vortex scattering, discussed the complications that arise in the presence of more vortices, and developed a path integral formalism that expresses exactly the long-range dynamics of non-Abelian vortices. Although solving the two vortex problem is quite easy, the problem with three vortices is virtually intractable. Furthermore, there does not seem to exist for the case of many vortices any obvious approximation that can be used to extract even the qualitative behavior of the system.

In the previous section it was mentioned that the path integral formulation discussed here ignores short-range vortex-vortex interactions. The physics considered in this paper is independent of the details of the internal structure of the vortices, allowing one to idealize the vortices as pointlike in structure and without short-range interactions. In real models vortices have finite size and interact when brought close together. The neglect of the short-distance interactions, which depend on the details of the internal structure, can be remedied adding vertex interactions consistent with flux conservation. The detailed internal structure of the vortex core, which is specific to the particular model, determines the coupling constants.

The difficulties encountered in the many-vortex system are quite similar to those encountered in systems with many anyons. As for the N non-Abelian vortex system, the path integral for the N non-interacting anyon system distinguishes between different braidings, assigning a distinct complex phase to each braiding. There are, however, significant differences that make the anyon problem more tractable. The assignment to a braiding of a mapping from the punctured plane to the unbroken symmetry group is more complicated than just a complex phase. For non-Abelian vortices, it is winding a vortex around another vortex of a different kind that makes it necessary to distinguish between different braidings. If all the vortices were of the same type, then the many-vortex system would behave like a non-interacting Bose gas. Complications arise only when vortices of a different type are introduced because what vortices may be considered identical depends on what other kinds of vortices are present in the system.

For anyons the exotic fractional statics can be removed by introducing an additional gauge field with a Chern-Simons kinetic term in the action. The anyons are minimally coupled to the gauge field. At the cost of introducing an extra gauge field, the exotic statistics have been removed. Now it is possible to calculate the average Chern-Simmons magnetic field due to the anyon gas and use mean field theory to obtain an approximate solution. Using such an approximation, Laughlin and Chen, Wilczek, Witten, and Halperin have predicted a superconducting phase for the non-interacting anyon gas. [11,12,13.]

It is not possible to apply this mean field approach to an ideal gas of non-Abelian vortices because it is the braiding of vortices around vortices of a different type that forces one to distinguish between different classes of braidings. Perhaps it is possible to find some mean field approach for certain specific problems. The behavior of the many non-Abelian vortex system is a problem that remains to be solved.

Notes

1. L. Krauss and F. Wilczek. "Discrete Gauge Symmetry in Continuum Theories," Phys. Rev. **62**, 1221 (1989).

2. J. Preskill and L. Krauss. "Local Discrete Symmetry and Quantum Mechanical Hair," Caltech Preprint 68-1601, Dec. 1989.

3. M. G. Alford, K. Benson, S. Coleman, J. March-Russell and F. Wilczek, "The Interactions and Excitations of Non-Abelian Cosmic Strings," HUTP 89-A051 (1990).

4. M. G. Alford, J. March-Russell, and Frank Wilczek, "Discrete Quantum Hair on Black Holes and the Non-Abelian Aharonov-Bohm Effect," HUTP 89-A040 (1989).

5. F. Wilczek and Y. Wu, "Space-Time Approach to Holonomy Scattering," IASSNS HEP 90-2 (Dec. 1989).

6. M.G. Alford and F. Wilczek, "Aharonov-Bohm Interaction of Cosmic Strings with Matter," Phys. Rev. Lett. **62,** 1071 (1989).

7. Y. Aharonov and D. Bohm, "The Significance of Electromagnetic Potentials in the Quantum Theory," Phys. Rev. **115**, 485 (1959).

8. T. Wu and C.N.Yang, "Concept of Non-Integrable Phase Factors and Global Formulation of Gauge Fields," Phys. Rev. **D12**, 3845 (1975).

9. P. Horváthy, "Non-Abelian Aharonov-Bohm Effect," Phys. Rev. **D33**, 407 (1986).

10. R. Sundrum and L. Tassie, "Non-Abelian Aharonov-Bohm Effects, Feynmen Paths, and Topology," J. Math. Phys. **27**,1566 (1986).

11. R. B. Laughlin, "Superconducting Ground State of Non-Interacting Particles Obeying Fractional Statistics," Phys. Rev. Lett. **60,** 2677 (1988).

12. R. B. Laughlin, "The Relationship Between High-Temperature Superconductivity and the Fractional Quantum Hall Effect," Science **242,** 525 (1988).

13. Y. Chen, F. Wilczek, E. Witten, and B. Halperin, "On Anyon Superconductivity," Int. J. Mod. Phys. **B3**, 1001 (1989).

Chapter 5
Alice Strings and Vortices

5.1. Introduction

In the previous chapter the physics of discrete non-Abelian strings and vortices whose flux must take values in a discrete, non-Abelian unbroken symmetry group H was discussed. This chapter generalizes the discussion to cases in which H consists of several connected components and has exactly one infinitesimal generator, which we shall call Q. Of particular interest is when Q does not commute with at least some of the elements of H that lie in the connected components that are not connected to the identity. For such a situation, the $U(1)$ symmetry generated by Q cannot be globally defined in a background with vortices. Pairs of vortices in two spatial dimensions (or closed loops of string in three spatial dimensions) can carry an unlocalizable Q-charge that in the quantized theory must take values that are integral multiples of $2e$, where e is the quantum of charge of the theory. In three dimensions loops of string can also carry magnetic charge in integral multiples of $2g$, where g is the magnetic charge quantum as determined from e by the Dirac quantization condition.

The simplest model in which such symmetry breaking occurs was first discussed by A.S. Schwarz and Y. S. Tyupkin in 1982. [1,2.] This model is a gauge field theory with a $G = SO(3)$ symmetry and has a scalar field that transforms according to the 5-representation. The Higgs field may be represented by a 3×3 traceless, symmetric, real matrix M that under $O \in SO(3)$ transforms as $M \to OMO^{-1}$. Suppose that a potential is chosen in such a way that the Higgs field condenses acquiring a vacuum expectation value of the form

$$<M>= v \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix}. \tag{5.1.1}$$

In the presence of such a condensate the unbroken symmetry group H consists of two disconnected components: a component connected to the identity H_c, and another component H_d that is not connected to the identity. H_c is generated by the Lie algebra element

$$Q = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{5.1.2}$$

If $SO(3)$ is thought of as acting on ordinary 3-space, then Q generates the rotations about the z-axis. The elements of H_d consist of the 180° rotations about axes that lie in the xy-plane. If one arbitrarily chooses an $X \in H_d$, for definiteness the rotation about the x-axis so that

$$X = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}, \tag{5.1.3}$$

then every $Y \in H_d$ can be expressed as $Y = X\mathrm{e}^{i\phi Q}$. If one allows $\mathrm{e}^{i\phi Q}$ to act on X by conjugation, then

$$X \to \mathrm{e}^{i\phi Q} X \mathrm{e}^{-i\phi Q} = X\mathrm{e}^{-2i\phi Q} = \mathrm{e}^{+2i\phi Q} X, \tag{5.1.4}$$

which is a rotation about an axis rotated from the x-axis by an angle ϕ. Consequently, H_d is a single conjugacy class.

The stable vortices in this theory correspond to the components of H considered as a subgroup of $SU(2)$, which is the universal covering group of $SO(3)$. After $H \subset SO(3)$ is lifted to $SU(2)$, H_c and H_d have exactly the same structure as before. The only difference is that the angle through which one must rotate to return to the identity becomes 4π instead of 2π. Vortices carry flux that lies in the disconnected component $H_d = \{X\mathrm{e}^{i\varphi Q} | 0 \leq \varphi < 4\pi\}$. Pairs of vortices can carry a combined flux $h = \mathrm{e}^{i\varphi Q}$ where $0 \leq \varphi < 4\pi$. The complete phase, where $\mathrm{e}^{i4\pi Q} = 1$ but $\mathrm{e}^{i2\pi Q} \neq 1$, is at least in principle observable because in the field theory there can exist particles that transform according to half-integral representations of $SU(2)$. The relative phase of Alice vortices can be measured using ordinary Aharonov-Bohm scattering.

5.2. Classical Alice Electrodynamics

Since Alice strings carry a flux $Y = Xe^{i\varphi Q}$, which has the property

$$YQY^{-1} = -Q, \tag{5.2.1}$$

a charge flips its sign after traversing around an Alice string. The change in sign is not just a matter of semantics; rather it is a physically observable phenomenon. In Figure 5.1(a) are two positive charges and an Alice vortex. One of the two charges is taken around the vortex along the path C indicated in Figure 5.1(b) back to its original position while the other charge remains in its original position. In the initial configuration, shown in Figure 5.1(a), the charges have the same sign and repel; in the final configuration, shown in Figure 5.1(c), after the one charge moves around the vortex, the charges have opposite signs and attract. At first sight this property may seem paradoxical because locally, or in any simply connected region of space, and at energy scales much lower than the symmetry breaking scale, ordinary classical electrodynamics describes the physics of the unbroken $U(1)_Q$ symmetry exactly. In particular, since charge seems to be conserved, at least locally, one is led to ask: What happens to the missing charge $+2q$ that seems to disappear after a charged particle with charge $+q$ travels around the Alice vortex and changes its charge to $-q$?

The classical electrodynamics in the presence of Alice strings has been investigated first by Schwarz and Tyupkin [1,2], and more recently by Preskill and Krauss [3] and Alford, Benson, Coleman, March-Russell, and Wilczek [4]. A discussion of the quantization of the Alice string loop appears in the following section.

Since it is much easier to think about configurations with pairs (or even numbers) of Alice vortices, we shall first consider a configuration in which there are two Alice vortices whose positions are fixed. (It is possible to regard an isolated Alice vortex as a pair of vortices in the limit as one of the vortices is removed to spatial infinity.) The paradox of the missing charge is resolved by the peculiar fact that a pair of Alice vortices can carry electric charge without any apparent source. Since $(\nabla \cdot \mathbf{E}) = 0$

everywhere, it is not possible to attribute the source of the charge to any particular region. Nor, as Preskill points out, is it possible that the missing charge is transferred to the vortex cores (or the core of the string) because it is possible to carry out the process adiabatically, moving the charge through the vortices or loop infinitely slowly. If the core of the vortex or of the loop has charged states, then because of electrostatic self-energy these states must be separated by a finite energy gap. But as a consequence of the adiabatic theorem, in the limit as the process is carried out infinitely slowly, no transition will take place.

Using Gauss's law, we define the total charge by the surface integral

$$Q_{total} = \int_{\partial V} d\mathbf{S} \cdot \mathbf{E} \tag{5.2.2}$$

where the region V encloses the vortex pair. This integral indicates the charge seen at large distances by an observer unconcerned with how the long range field is produced. The field of the vortex pair behaves just as if it were created by an ordinary charge.

This total charge must be conserved, for otherwise causality would be violated. It is not possible for any local process to change the total charge, for the electric field extends to spatial infinity, and an operator that changes Q_{total} would have to destroy this field at infinity and therefore could not be local.

It is important to realize that the property that a vortex pair can carry charge is a global phenomenon. Locally, in a simply connected region, it is always possible to make a unitary choice of gauge. The heavy gauge boson degrees of freedom are not relevant, and we simply have ordinary electrodynamics with the usual effective Lagrangian $\mathcal{L} = (-1/4) \cdot F_{\mu\nu} F^{\mu\nu}$. Suppose that we want to consider not just a simply connected region but the entire region around the two vortices indicated in Figure 5.2. A pair of Alice vortices is placed at fixed locations A and B. The region can be made simply connected by introducing a cut that connects the two vortices. The gauge transformation required to implement a unitary choice of gauge in this region is singular along the cut if the initial choice of gauge is smooth. The vector potential

A_μ, which vanishes in the interior of this region, will have a δ-function singularity on the cut, and this singularity will require that the fields of the theory satisfy a discontinuous matching condition across the cut so that the covariant derivative $\mathcal{D}_\mu = (\partial_\mu + igA_\mu)$ is continuous across the cut. In particular, $F_{\mu\nu}$ must flip its sign upon crossing the cut.

For a pair of Alice vortices it is most convenient to choose as the cut the straight line segment connecting the two cores. To find the charged vortex pair solutions, we solve for electric field configurations that satisfy $(\nabla \cdot \mathbf{E}) = 0$ and the matching condition that $\mathbf{E}(\mathbf{x}) \to -\mathbf{E}(\mathbf{x})$ as $\mathbf{x}$ crosses the cut. If we imagine the cut to be a perfect conductor with a charge distributed over its surface, then the solution obtained satisfies the boundary conditions for the vortex pair problem because at the surface of a conductor $\mathbf{E}$ is normal, and because of the symmetry of the problem, $\mathbf{E}$ points in opposite directions on opposite sides of the cut. The solution is indicated in Figure 5.3. The similarity between the problems, however, is deceptive. For the conductor real charge produces the field as is indicated by evaluating $(\nabla \cdot \mathbf{E})$. The divergence in the Alice pair problem, which must be taken using the covariant derivative, vanishes because of the singular vector potential on the cut. Seen another way, the cut may be moved by making a gauge transformation, and the apparent source of charge will move. Therefore, the appearance of charge on the cut is nothing but a gauge artefact.

We next consider a system consisting of an Alice vortex pair and a charged particle. Suppose that initially the pair is uncharged and that the charge is brought toward the vortex pair from spatial infinity. In Figure 5.4 a charge $+q$ is brought toward the Alice vortex pair from the side. Because the vortex pair was initially uncharged, the lines of force that end on the cut must be compensated for by an equal number of lines of force that emanate from the cut. It is apparent that the energy of the electric field is lowered by the presence of the vortex pair; therefore, from the side the interaction between the charge and the pair is attractive. When the charge lies on the line connecting the vortices, the field on the cut again because of the reflection symmetry of the problem satisfies the conductor boundary conditions.

It is as if charge from a neutral conductor of the opposite sign is attracted toward the charged particle and charge of the same sign is repelled toward the part of the cut farther away from the charge.

In Figure 5.5 the charge approaches the pair from below rather than from the side. In this case the electric field energy is increased as the charge is brought toward the cut, so the charge and the vortex pair repel each other. As before, all the flux lines that emanate from the charge and end on the cut must re-emanate on the same side of the cut as is indicated in Figure 5.5(a.) To show what happens when the charge passes through the cut, we first bring the charge close to the cut, as indicated in Figure 5.5(b), and then instead of moving the charge further, we move the cut over the charge as indicated in Figure 5.5(c). With the cut moved over the charge, the cut appears to have acquired a charge $+2q$ and the charge of the charged particle appears to have changed from $+q$ to $-q$. Since the operation of moving the cut merely consists of choosing a different gauge, there is no gauge-invariant way to determine the distribution of charge between the pair and the charged particle. Only the total charge is well-defined. In Figure 5.5(d) the field that results after the charged particle has been pulled through the vortices and has been moved farther upward is shown.

For those who find the idea of introducing cuts distasteful, there is another more elegant, but perhaps less useful way to formulate electrodynamics in an Alice vortex background. The gauge field potential can be split into two parts

$$A_\mu = A_\mu^{(bac)} + \bar{A}_\mu \tag{5.2.3}$$

where $A_\mu^{(bac)}$ is the background vector potential due to the flux of the vortices and $\bar{A}_\mu$ is an additional part. Since locally $A_\mu^{(bac)}$ is pure gauge and can be set to zero by a gauge transformation, we should be able to write an action just in terms of $\bar{A}_\mu$ without any interference terms between the two parts of A_μ because one of the parts is pure gauge. However, we run into difficulties in trying to choose a basis for the generators for $\bar{A}_\mu$ because of the matching condition. If the vortex carries a flux

$h = \exp(i\pi T_1)$, then after parallel transport through the region between the vortices by $A_\mu^{(bac)}$, $\bar{T}_1$, $\bar{T}_2$, $\bar{T}_3$ transform into $\bar{T}_1$, $-\bar{T}_2$, $-\bar{T}_3$, respectively. Consequently, it is not possible to choose single valued generators. and a cut must be introduced. However, if we expand $\bar{A}_\mu = \bar{A}_\mu^{(1)}\bar{T}^{(1)} + \bar{A}_\mu^{(2)}\bar{T}^{(2)} + \bar{A}_\mu^{(3)}\bar{T}^{(3)}$, we see that $\bar{A}_\mu^{(2)}$ and $\bar{A}_\mu^{(3)}$ must be double valued since $\bar{A}_\mu$ is single valued. The effective Lagrangian then can be written as

$$\mathcal{L} = \frac{-1}{4}\bar{F}_{\mu\nu}^{(a)}\bar{F}^{\mu\nu(a)} + g^2v^2\left[\bar{A}_\mu^{(1)}\bar{A}^{(1)\mu} + \bar{A}_\mu^{(2)}\bar{A}^{(2)\mu}\right]. \tag{5.2.4}$$

If there exist monopoles in the theory, then by the same line of argument as with electric charge it can be shown that Alice string loops can carry magnetic charge. In a later section we have shown by another method that Alice string loops generically can carry magnetic charge. This fact has important cosmological consequences because magnetic monopoles are troublesome in Grand Unified theories. [7,8.]

5.3. Gauge Symmetry and Electric Charge

In the classical theory of electrodynamics in a background with an even number of Alice vortices, it is not possible to fix the sign of the charge in a continuous or gauge-invariant manner: however, the total charge of a collection of vortices and charges is well-defined and conserved. One defines

$$Q_{total} = \int_{\partial V} d\mathbf{S}\cdot\mathbf{E} \tag{5.3.1}$$

where V is a large volume that encloses the vortices and charged particles. In this section this definition of total charge is extended to situations in which both the vortices and the charges are quantum fields. We shall find that quantum mechanical operator that corresponds to

$$\hat{Q} = \lim_{R\to\infty} \int_{\partial V(R)} d\mathbf{S}\cdot\mathbf{E}, \tag{5.3.2}$$

where $V(R)$ is a volume of radius R, is the generator of the global gauge transformations. In a gauge theory physical states must be invariant under gauge transformations that are equal to the identity except in regions of compact support. However, there is no such requirement for global gauge transformations, and the transformation properties of a state under global $U(1)_Q$ transformations determine its total charge. [3.]

In the classical field theory in the Hamiltonian formulation in Hamiltonian gauge, in addition to the action principle there is an equation of constraint, which is equivalent to Gauss's law, that must be imposed in order to obtain the complete equations of motion for electrodynamics. An equation of constraint is required because the gauge fixing condition $A_0 = 0$ has forced variations with respect to A_0 to escape consideration. In the Hamiltonian formulation the dynamical variables are the vector potentials $A_1(\mathbf{x}), A_2(\mathbf{x}), A_3(\mathbf{x})$ and their conjugate fields $\Pi_1(\mathbf{x}), \Pi_2(\mathbf{x}), \Pi_3(\mathbf{x})$, which are equal to the electric field components $E_1(\mathbf{x}), E_2(\mathbf{x}), E_3(\mathbf{x})$. The equation obtained by varying the Lagrangian action with respect to A_0 is

$$[\partial^\mu \partial_\mu A^0 - \partial^0(\partial^\mu A_\mu)] - J_0 = 0, \tag{5.3.3}$$

which with $A_0 = 0$ becomes

$$\partial_i(\partial_0 A_i) - J_0 = \nabla \cdot \mathbf{E} - J_0 = \nabla \cdot \Pi - J_0 = 0, \tag{5.3.4}$$

which is Gauss's law. In the classical field theory the functional

$$\begin{aligned} F[\mathbf{A}_i, \mathbf{E}_i] &= \int d^2x \; \omega(\mathbf{x})[\nabla \cdot \mathbf{E}(\mathbf{x}) - J_0(\mathbf{x})] \\ &= \int d^2x \; [-[\nabla\omega(\mathbf{x})] \cdot \mathbf{E}(\mathbf{x}) - \omega(\mathbf{x}) J_0(\mathbf{x})] \end{aligned} \tag{5.3.5}$$

is the generator of infinitesimal gauge transformations. The pair of coupled ordinary differential equations

$$\begin{aligned} \frac{\partial A_i(\mathbf{z})}{\partial \xi} &= \{F[\mathbf{A}(\mathbf{x}), \mathbf{E}(\mathbf{x})], A_i(\mathbf{z})\} \\ \frac{\partial E_i(\mathbf{z})}{\partial \xi} &= \{F[\mathbf{A}(\mathbf{x}), \mathbf{E}(\mathbf{x})], E_i(\mathbf{z})\} \end{aligned} \tag{5.3.6}$$

with the Poisson bracket defined as

$$\{F,G\} = \int d^3x \left\{ \frac{\delta F}{\delta A_i(\mathbf{x})} \cdot \frac{\delta G}{\delta E_i(\mathbf{x})} - \frac{\delta F}{\delta E_i(\mathbf{x})} \cdot \frac{\delta G}{\delta A_i(\mathbf{x})} \right\}, \tag{5.3.7}$$

when integrated from $\xi = 0$ to $\xi = 1$, generates the finite time-independent gauge transformation

$$\mathbf{A}(\mathbf{x}) \to \mathbf{A}(\mathbf{x}) - \nabla\omega(\mathbf{x}). \tag{5.3.8}$$

In the quantum theory the shape states, which are the functions $\mathbf{A}(\mathbf{x})$, where $\mathbf{x}$ are just the spatial coordinates, form a basis of quantum states from which the Hilbert space of the system is constructed. At fixed time in Hamiltonian gauge the quantum state of the electromagnetic field is described by a state vector $|\Psi>$ in the Hilbert space $\mathcal{H}$, which is a complex valued linear functional defined on the space of shape states. This Hilbert space is too large, and the Gauss's law constraint must be imposed to obtain the subspace $\mathcal{H}_{ph} \subset \mathcal{H}$ of physical states. In the quantized theory (with $\hbar = 1$), the electric field becomes the functional derivative

$$E_i = \Pi_i = (-i) \cdot \frac{\delta}{\delta A_i(\mathbf{x})}, \tag{5.3.9}$$

and if $\omega(\mathbf{x})$ is a function of compact support, then one must require that the operator

$$\begin{aligned} G[\omega] &= \int d^3x \left[-[\nabla\omega(\mathbf{x})] \cdot \mathbf{E}(\mathbf{x}) - \omega(\mathbf{x}) J_0(\mathbf{x}) \right] \\ &= \int d^3x \left[-[\nabla\omega(\mathbf{x})] \cdot (-i) \cdot \frac{\delta}{\delta \mathbf{A}(\mathbf{x})} - \omega(\mathbf{x}) J_0(\mathbf{x}) \right] \end{aligned} \tag{5.3.10}$$

annihilate physical states. This condition can be expressed as

$$G[\omega(\mathbf{x})]|\Psi>_{phy} = 0 \tag{5.3.11}$$

where the ket $|\Psi>$ represents a complex linear functional $\Psi[\mathbf{A}(\mathbf{x})]$. Physically, this condition means that if two shape states differ by a time-independent gauge transformation of compact support, then a functional that represents a physical state must map them into the same amplitude.

In the quantum theory $G[\omega(\mathbf{x})]$ represents an infinitesimal gauge transformation, and a finite gauge transformation is formed using the exponential so that

$$U[\omega(\mathbf{x})] = \exp\left[iG[\omega(\mathbf{x})]\right]. \tag{5.3.12}$$

$U[\omega(\mathbf{x})]$ is unitary since $G[\omega(\mathbf{x})]$ is Hermitian.

This kind of gauge transformation probes whether Gauss's law is satisfied in the quantum theory. If we want to test Gauss's law in its usual integral form

$$\int_{\partial V} d\mathbf{S} \cdot \mathbf{E} = \int_{V} d^3x \; J_0 = Q_V, \tag{5.3.13}$$

we would choose

$$\omega(\mathbf{x}) = \begin{cases} 1, & \text{for } \mathbf{x} \in V, \\ 0, & \text{for } \mathbf{x} \notin V. \end{cases} \tag{5.3.14}$$

To avoid a singularity at ∂V the function $\omega(\mathbf{x})$ could be smeared, giving the boundary a finite thickness.

Now suppose that $\omega(\mathbf{x})$ does not have compact support but rather approaches a constant value at spatial infinity. Since differences in $\omega(\mathbf{x})$ in regions of compact support do not have any effect on physical states, we can assume that $\omega(\mathbf{x})$ is constant. Then in the expression $G[\omega(\mathbf{x})]$, defined as the integral in Equation (5.3.10), the J_0 term makes a contribution proportional to the total electric charge, but there is no contribution from the $(\nabla\omega) \cdot \mathbf{E}$ term to cancel it because ω is constant. Therefore, $G[\omega]$ measures the total charge. For states with definite charge $q = me$ where e is the quantum of charge

$$U_Q(\omega)|\Phi> = e^{i\omega\hat{Q}}|\Phi> = e^{im\omega}|\Phi>, \tag{5.3.15}$$

meaning that the state transforms under an irreducible representation of the global gauge transformations.

For non-Abelian gauge fields the quantization in Hamiltonian gauge is completely analogous. Restricting the gauge to $A_0 = 0$ necessitates restricting the Hilbert space $\mathcal{H}$ to a subspace of physical states $\mathcal{H}_{phys}$ by imposing the requirement that the wave functionals are invariant under time-independent gauge transformations of compact support. Problems arise only when one tries to identify a subset consisting of global gauge transformations to measure total charge. In non-Abelian gauge theories non-trivial global charge and global gauge transformations cannot be defined in a gauge invariant manner and, therefore, are not physically valid concepts. The difficulty really has nothing to do with quantization. In fact the problem is less tractable in the classical theory because electric confinement, which is a purely quantum phenomenon, makes it unnecessary to attempt to define the total color charge of a state in the quantum theory. In the presence of a Higgs condensate where the unbroken symmetry is $U(1)$, however, these difficulties disappear because the orientation of the Higgs condensate indicates a natural direction for Q, at least asymptotically when there are no vortices or an even number of Alice vortices. Therefore, the total charge can be defined using the transformation properties under $U(1)_Q$ in a manner exactly analogous to the simple QED case just discussed.

Now let us return to the system with a pair of Alice vortices. In Figure 5.6 are two vortices A and B, a basepoint x_0, and two paths α and β that generate $\pi_1(M, x_0)$. Using the methods developed in Chapter 4, we assign the following fluxes to the vortices

$$h(\alpha) = X\mathrm{e}^{i\varphi_a Q}$$
$$h(\beta) = X\mathrm{e}^{i\varphi_b Q} \tag{5.3.16}$$

These two fluxes determine completely all properties of the two-vortex background that are gauge invariant and, hence, those which have physically observable consequences. Since the total flux is conserved as a consequence of the total flux superselection rule,

$$h_{total} = h(\alpha\beta) = X\mathrm{e}^{i\varphi_a Q} X\mathrm{e}^{i\varphi_b Q} = \mathrm{e}^{i(\varphi_b - \varphi_a)Q} \tag{5.3.17}$$

is constant with time. Therefore, the difference $\varphi = (\varphi_b - \varphi_a)$ must be independent of time; however, φ_a and φ_b may rotate in opposite directions. We shall show later that such rotating solutions correspond to the charged classical solutions in Hamiltonian gauge.

Suppose that we allow a $U(1)_Q$ global gauge transformation to act on the vortex pair described above. Since

$$h(C, x_0) \rightarrow \Omega(x_0) h(C, x_0) \Omega^{-1}(x_0), \tag{5.3.18}$$

it follows that under $U_Q(\theta) = e^{i\theta Q}$ the vortex flux $h(\alpha)$ transforms as

$$X e^{i\varphi_a Q} \rightarrow e^{i\theta Q} X e^{i\varphi_a Q} e^{-i\theta Q} = X e^{i(\varphi_a - 2\theta) Q}. \tag{5.3.19}$$

Note that the total flux remains invariant because it lies in the center of H and, therefore, is unaffected by conjugation. Global gauge transformations rotate φ_a and φ_b in opposite directions.

The fact that the flux does not transform trivially under global gauge transformations is intimately related to the global unrealizability of global symmetry in the presence of the vortex pair. If the pair were enclosed by a simply-connected region R, the rotations $U(1)_Q$ could be globally defined outside of R. Trouble arises only when one tries to extend the global symmetry into the region between the vortices. It is impossible to do so without choosing a transformation that rotates the nonvanishing background gauge field A_μ. In a singular gauge with a cut, the global gauge transformations act on the matching conditions in a non-trivial manner. It turns out that all of the many ways of extending $U(1)_Q$ into the region between the vortices act equivalently on the space of physical states because the different extensions differ only by a gauge transformation of compact support.

Since states of definite charge transform irreducibly under $U(1)$, one must form superpositions of classical field configurations with different fluxes in order to produce charge eigenstates. Starting with an arbitrary classical field configuration $|\theta>$, one defines a one-parameter family of rotated states

$$|\theta>= e^{i\theta\hat{Q}}|\theta=0> . \tag{5.3.20}$$

Although $\theta = 4\pi$ is the smallest angle such that $e^{i\theta\hat{Q}} = 1$, $|\theta = 2\pi>=|\theta=0>$. Under $U(1)_Q$ global gauge transformations these states transform according to

$$e^{i\omega\hat{Q}}|\theta>=|\theta+2\omega> . \tag{5.3.21}$$

Since eigenstates of the charge operator merely acquire a phase under global gauge transformations, we must superimpose the θ-states to obtain quantum states with definite charge. Consequently, the state

$$|n>=\int_0^{2\pi} d\theta\ e^{in\theta}|\theta> \tag{5.3.22}$$

represents a pair of vortices carrying a charge nq where $q = 2e$ is twice the charge quantum of the theory.

We see that while in the classical theory the charge carried by the pair (or loop) can take any value, in the quantized theory the charge is constrained to take values that are integral multiples of $2e$, where e is the magnitude of the charge of a particle from a field that transforms under the 2-dimensional representation of $SU(2)$.

Let us reconsider the classical solution for a charged circular loop of Alice strings, discussed in Section 5.2. If the cut is chosen as the circular disk bounded by the loop of string, the problem can be readily solved because it is identical to the problem of a conducting or equipotential disk. Suppose that the disk has a radius R. The potential problem is solved in oblate spheroidal coordinates, in which the disk becomes a surface of constant ξ. [6.] It turns out that $\Phi(\xi,\eta,\phi) = \Phi(\xi)$ and that the disk has a capacitance $C = (2R/\pi)$ so that

$$\Phi_0 = \frac{\pi Q}{2R} \tag{5.3.23}$$

where Φ_0 is the potential on the surface of the disk. Consequently, there is an energy

$$U(Q) = \frac{\pi Q^2}{4R} \tag{5.3.24}$$

in the electric field that creates a potential that tries to expand the loop. This increase in potential means that loops can be stable if they carry charge because the decrease in potential energy due to shortening the string as R is decreased can be matched by the increase in electric field energy.

For our purposes, it is convenient to transform this solution to Hamiltonian ($A_0 = 0$) gauge. We divide the vector potential into two parts

$$A_i = A_i^{(cut)} + A_i^{(E)}, \tag{5.3.25}$$

which are

$$\mathbf{A}^{(cut)} = \hat{\mathbf{z}}\ \pi\ T_1\ \delta(z)\ \theta(R - \rho) \tag{5.3.26}$$

and a part that varies linearly with time

$$A_i^{(E)} = t\ T_3\ \nabla_i \Phi. \tag{5.3.27}$$

An observer at infinity at time t measures a flux,

$$\begin{aligned} h(C, x_0; t) &= \exp\Big[i \int\limits_{C_1} dx^i A_i(x)\Big] \exp\Big[i\pi T_1\Big] \exp\Big[\ i \int\limits_{C_2} dx^i A_i(x)\Big] \\ &= \exp\big[it\Phi_0 T_3\big] \exp\big[i\pi T_1\big] \exp\big[-it\Phi_0 T_3\big] \\ &= \exp\Big[i\pi[\cos(\Phi_0 t)T_1 + \sin(\Phi_0 t)T_2]\Big] \end{aligned} \tag{5.3.28}$$

where the contours are shown in Figure 5.7. Consequently, the flux $h(C, x_0, t)$ rotates in H_d at an angular velocity $\omega = \Phi_0$. Classically any angular velocity is allowed, but in the quantum theory the angular velocity must take discrete values.

4. Alice Strings and Magnetic Charge

Before considering magnetic charge carried by loops of Alice string in three spatial dimensions, we shall consider magnetic flux carried by pairs of Alice vortices in two spatial dimensions. In Figure 5.8 are shown two Alice vortices labeled A and B, with a basepoint x_0 and two paths C_a and C_b encircling the vortices A and B, respectively. These two paths generate the group $\pi_1(M, x_0)$. Since Alice vortices have a continuous degeneracy, two angles θ_a and θ_b are required to specify the flux. We set

$$\begin{aligned} h(C_a) &= X e^{i\theta_a Q} \\ h(C_b) &= e^{i\theta_b Q} X. \end{aligned} \tag{5.4.1}$$

The path $C = C_a C_b$ is used to measure the total flux, measurable at spatial infinity carried by the pair of vortices, which is

$$h(C, x_0) = e^{i(\theta_a + \theta_b)Q}. \tag{5.4.2}$$

This flux, which is ordinary $U(1)$ magnetic flux, can be measured by Aharonov-Bohm scattering.

Next consider the situation in three spatial dimensions. For parallel Alice strings of infinite length the case would be the same as that in the two-dimensional problem. The magnetic flux everywhere along the strings would have to be the same because otherwise there would be a radial magnetic field that falls off as $1/R$, corresponding to an energy per unit length that diverges logarithmically with R.

However, suppose we take a loop of string. There no longer is a log divergence, and we can put an integer number N kinks in the string when we close the two ends so that as a loop traces out a torus enclosing the loop, θ winds N times around $U(1)$. The kinks mean that if we take a path starting at x_0 that winds through the loop and back to x_0 and sweep this path around the loop back to its original position so as to sweep out a torus enclosing the loop, or equivalently a sphere that surrounds the loop, the winding number of $h(C, x_0)$ is the magnetic charge carried by the loop of Alice string.

Since this construction of magnetic charge used only the relation $XQX^{-1} = -Q$, magnetic charge generically exists in any theory that supports Alice strings. Therefore, such theories are plagued with all of the problems that arise in theories with monopoles. [7,8]

5. Generalizations of Alice Strings

In this section we consider more general Alice symmetry breaking scenarios. Consider $H \subset G$, where G is the simply connected universal covering group. Suppose that H is a one-dimensional Lie group with several connected components. Then H can be expressed as

$$H = U(1) \times_{s.d.} D \tag{5.5.1}$$

where D is a discrete group and $\times_{s.d.}$ indicates that the product is semi-direct. D must be finite because G is compact. The choice of $D \subset H$ is generally not unique because in the examples of interest to us $H_c = U(1)$ will act on D through conjugation non-trivially.

On the other hand, D acting on H_c by conjugation must leave H_c invariant. The action of an element $d \in D$ on H_c is determined by its action on the generator Q. There are two possibilities. Either

$$dQd^{-1} = +Q$$

or

$$dQd^{-1} = -Q$$

because D acting on H_c induces an automorphism of H_c, and $\text{Aut}[U(1)] \cong Z_2$.

Therefore, the elements of D, or equivalently the connected components of H, can be divided into two classes: those that commute with Q, and those that do not commute with Q and therefore display Alice behavior. Note that for there to be Alice behavior, H must be non-Abelian. However, D can be Abelian, as in our example where $D = Z_2$.

Notes

1. A. S. Schwarz, "Field Theories With No Local Conservation of Electric Charge," Nuc. Phys. **B208,** 141 (1982).

2. A. S. Schwarz and Y. S. Tyupkin, "Grand Unification and Mirror Particles," Nucl. Phys. **B209,** 427 (1982).

3. J. Preskill and L. Krauss, "Local Discrete Symmetry and Quantum Mechanical Hair," Caltech Preprint 68-1601, Dec. 1989.

4. M. G. Alford, K. Benson, S. Coleman, J. March-Russel and F. Wilczek, HUTP 89-A051(1990).

5. M. Bucher, "The Aharonov-Bohm Effect and Exotic Statistics for Non-Abelian Vortices," Caltech Preprint 68-1655 (1990). [To appear in Nucl. Phys. B.]

6. L. D. Landau and E. M. Lifshitz, *Electrodynamics of Continuous Media*. New York: Pergamon (1960).

7. A. Vilenkin, "Cosmic Strings and Domain Walls," Phys. Reports **121,** 263 (1985).

8. J. Preskill, "Cosmological Production of Superheavy Magnetic Monopoles," Phys. Rev. Lett. **43,** 1365 (1979).

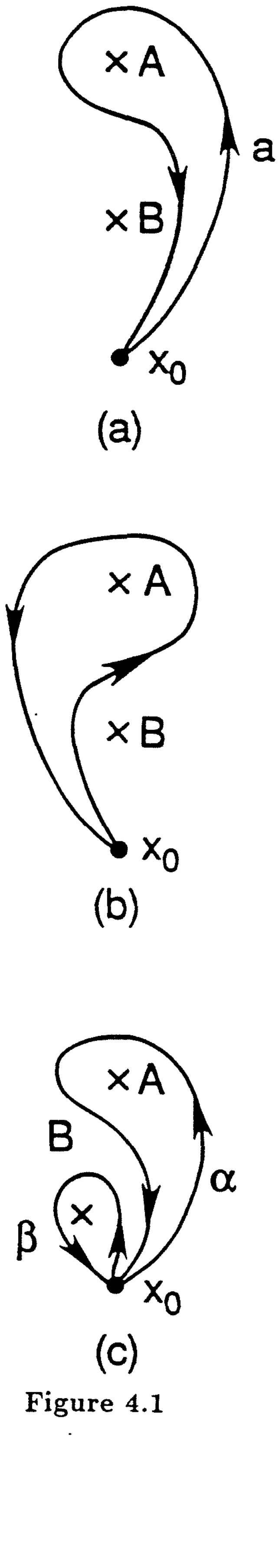


Figure 4.1

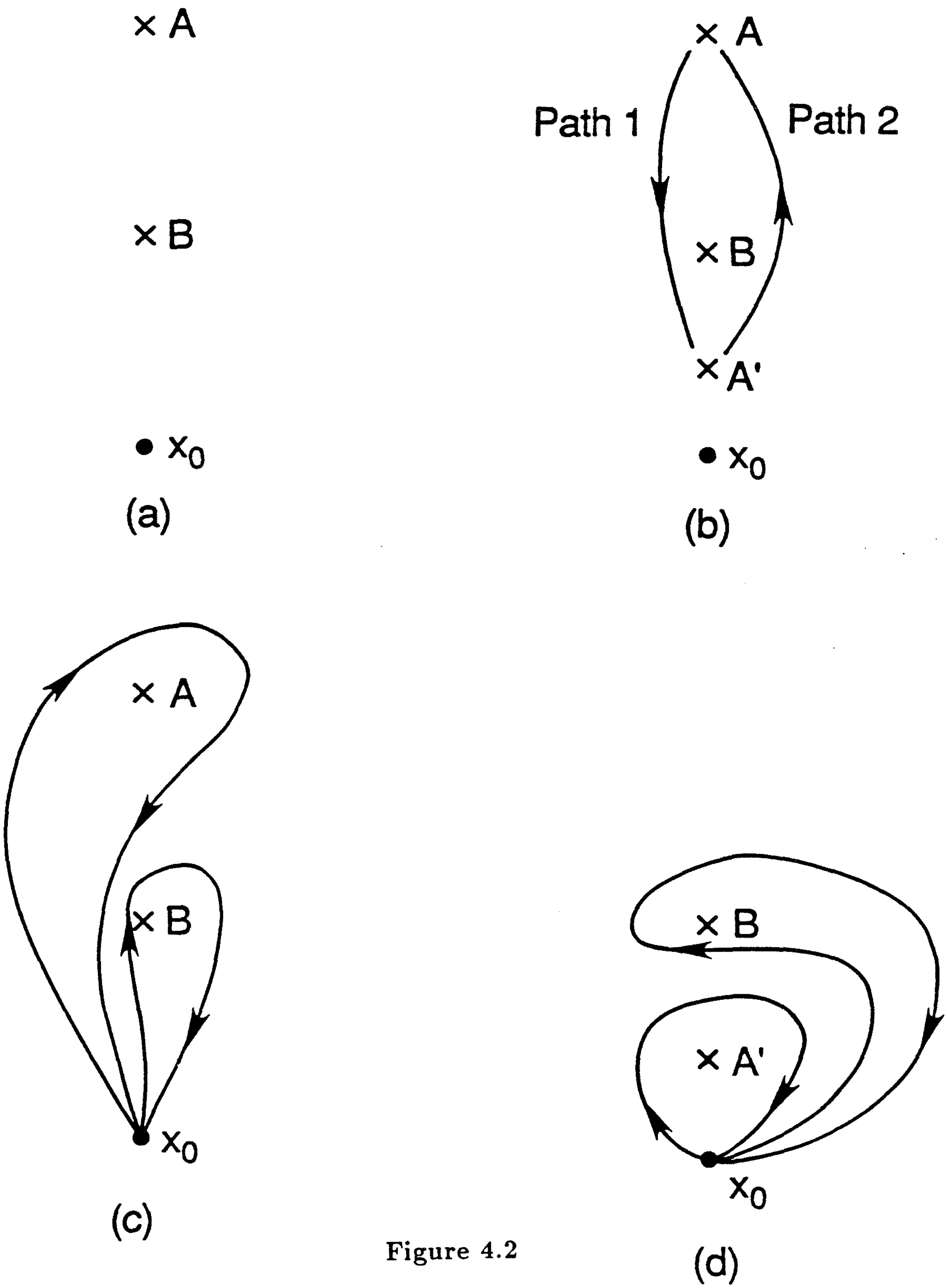


Figure 4.2

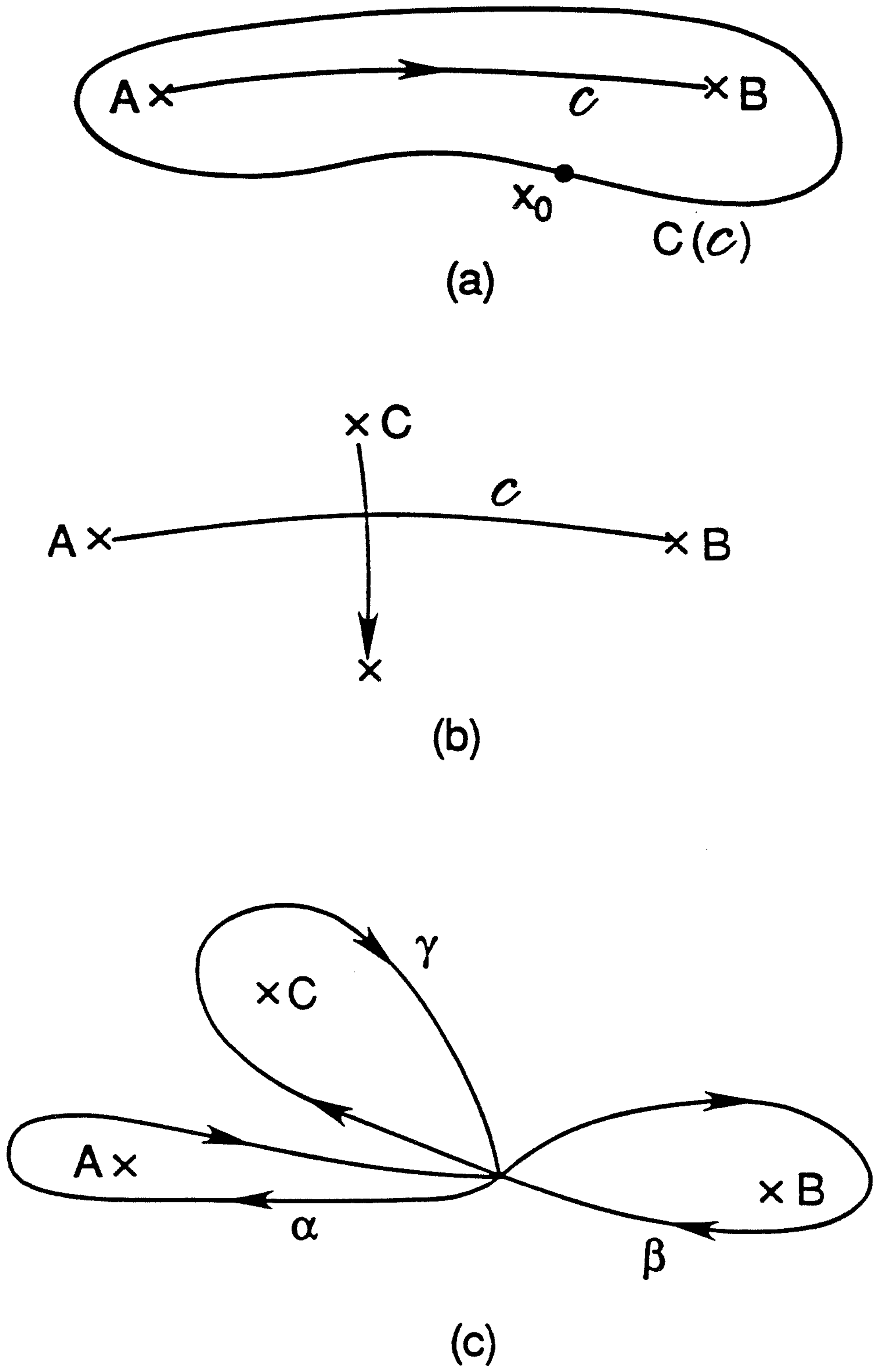


Figure 4.3

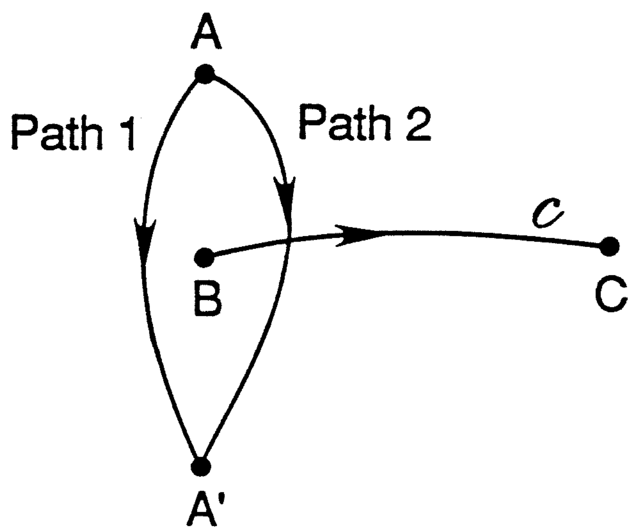
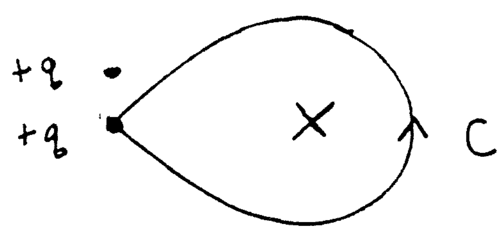
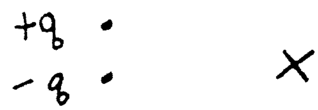


Figure 4.4

(a)



(b)



(c)

Figure 5.1

Figure 5.2

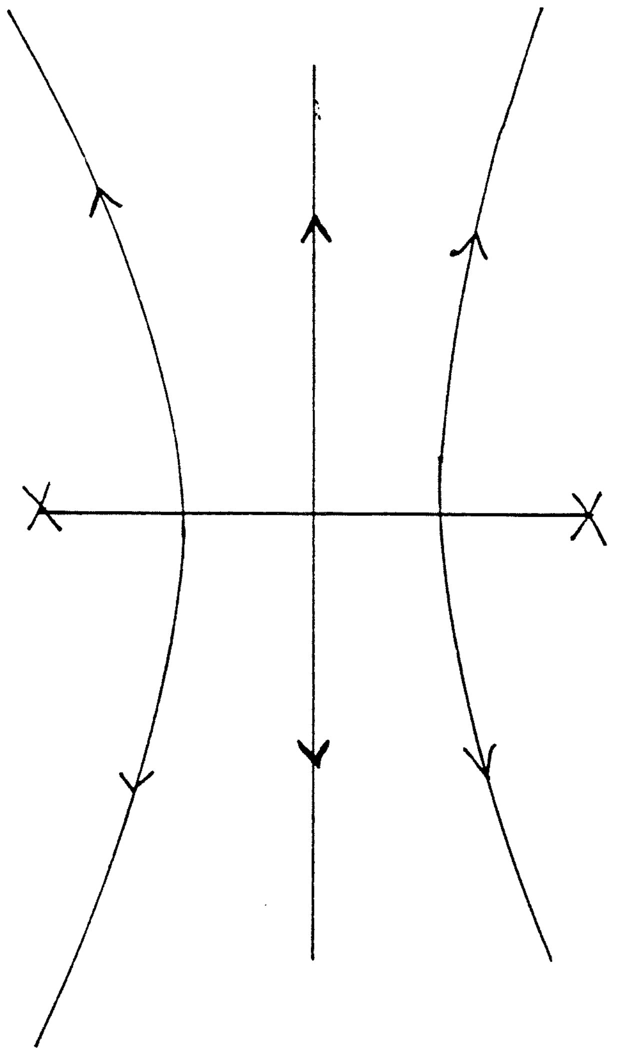


Figure 5.3

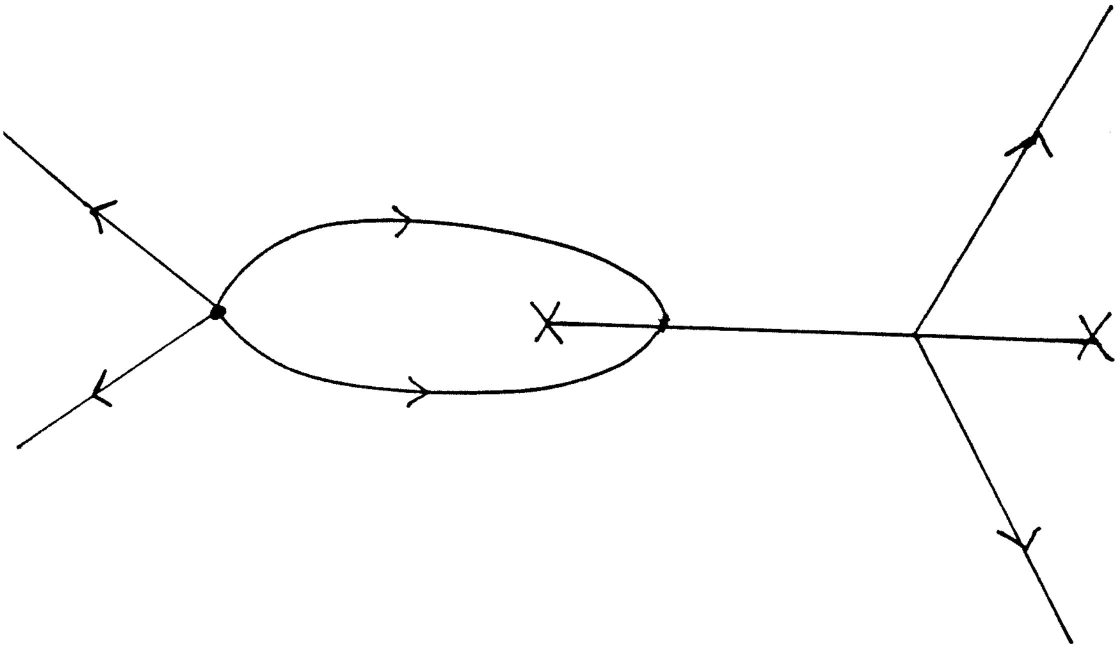


Figure 5.4

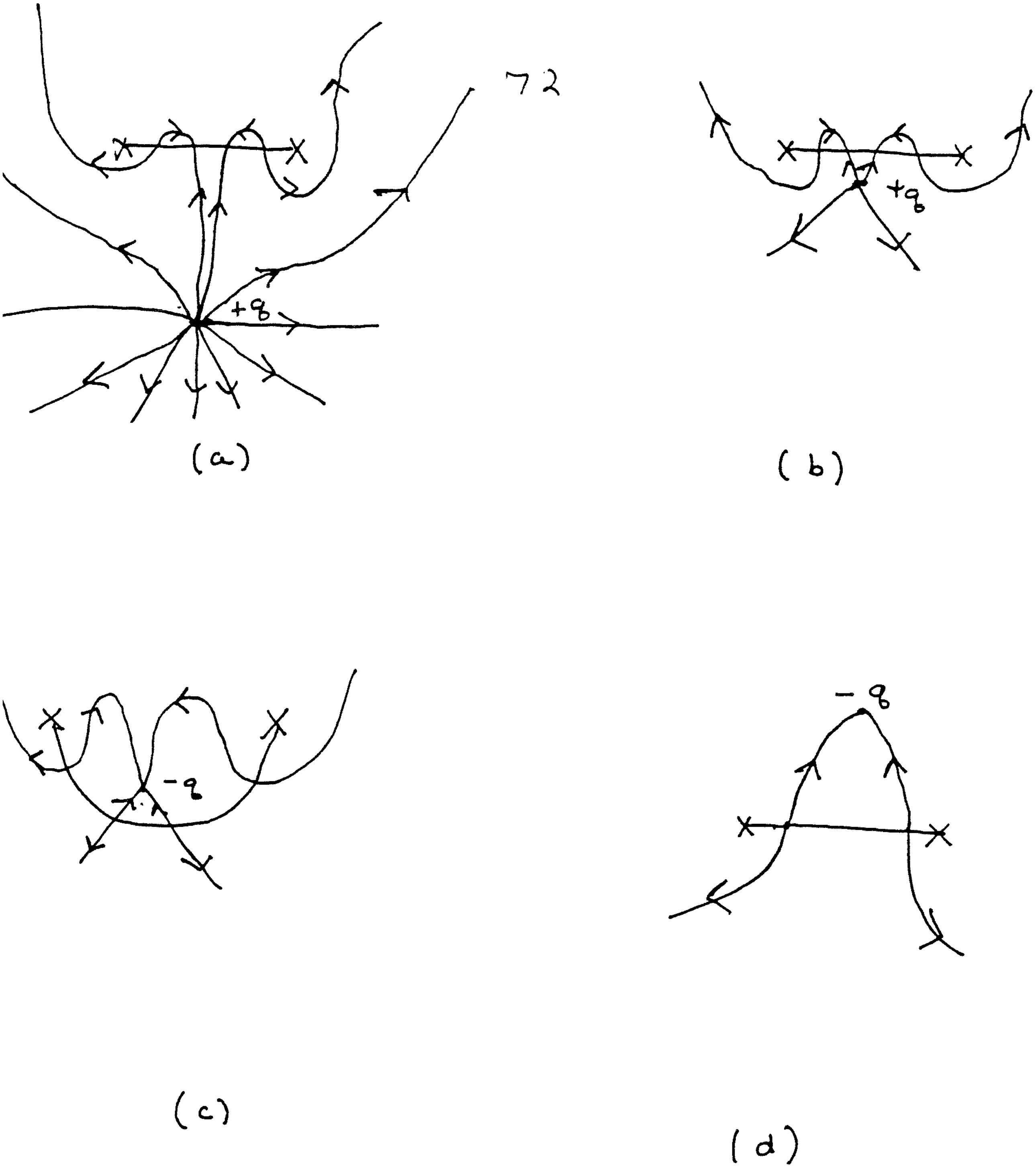


Figure 5.5

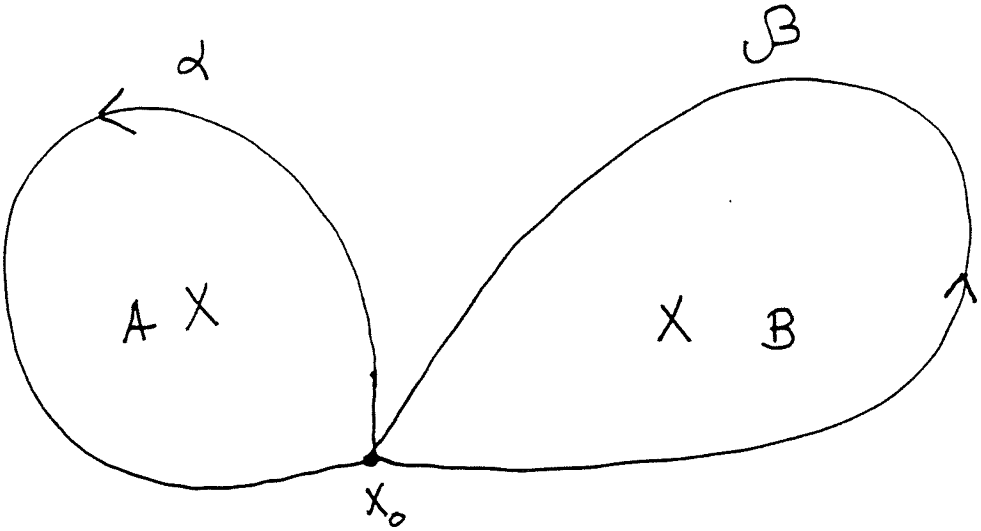


Figure 5.6

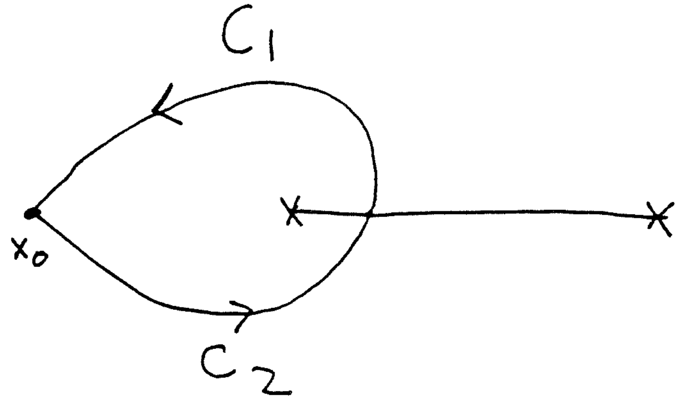


Figure 5.7

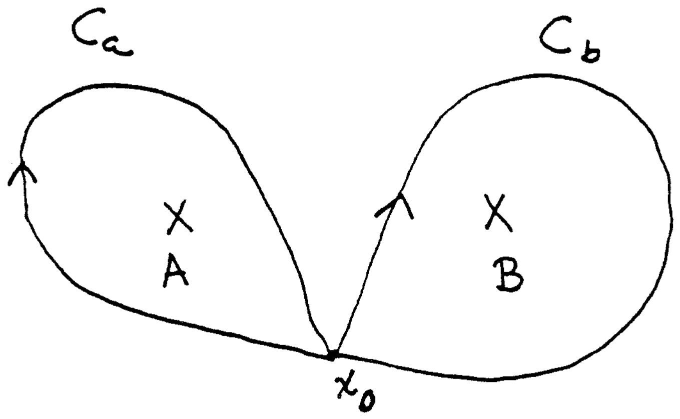


Figure 5.8