ANOMALOUS QUANTUM NUMBERS AND TOPOLOGICAL PROPERTIES OF FIELD THEORIES

Thesis by
Alexios P. Polychronakos

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

California Institute of Technology
Pasadena, California.

1987
(Submitted May 14, 1987)

ACKNOWLEDGEMENTS

A lot of people have contributed with their presence or memory to my carrying on in life and science to this point. I am grateful and obliged to all of them. Trying to name them individually would only create the risk of possible omissions, due to my desperatingly feeble memory. I will only express my explicit gratitude to my advisor in Caltech, Prof. John P. Preskill, as a representative of the pleiade of people worthy of such a mention, whose intellectual, moral and material assistance was invaluable in the pursuit and completion of the present work.

ABSTRACT

We examine the connection between anomalous quantum numbers, symmetry breaking patterns and topological properties of some field theories. The main results are the following: In three dimensions the vacuum in the presence of abelian magnetic field configurations behaves like a superconductor. Its quantum numbers are exactly calculable and are connected with the Atiyah-Patodi-Singer index theorem. Boundary conditions, however, play a nontrivial role in this case. Local conditions were found to be physically preferable than the usual global ones. Due to topological reasons, only theories for which the gauge invariant photon mass in three dimensions obeys a quantization condition can support states of nonzero magnetic flux. For similar reasons, this mass induces anomalous angular momentum quantum numbers to the states of the theory. Parity invariance and global flavor symmetry were shown to be incompatible in such theories. In the presence of massless flavored fermions, parity will always break for an odd number of fermion flavors, while for even fermion flavors it may not break but only at the expense of maximally breaking the flavor symmetry. Finally, a connection between these theories and the quantum Hall effect was indicated.

Table of Contents

INTRODUCTION

Ever since quantum field theories established themselves as good and promising mathematical frameworks for expressing the properties of elementary particles, and especially after the discovery and application of renormalization, a lot of work has been done scrutinizing their mathematical structure. At the same time, the classical field theories that underlie them have also been given a closer look, in view of their possible relevance to their quantum descendents.

The main problem, though, with these otherwise beautiful and rich theories is that, with very few exceptions, it is from very hard to impossible to derive any exact results from them, or even to perform any analytical calculations at all. As a matter of fact, the only way to squeeze any concrete numerical results out of them was, until recently, perturbation theory. The surprising and ingenious exact solution of quantum electrodynamics in one spatial and one temporal dimension (1+1 QED) given by Schwinger was the only known exception to this rule, that only served to make the lack of more physically relevant such results even more depressing.

This, however, should not lead us in underestimating the successes and even triumphs of perturbation theory, as well as the difficulties and technical obstacles that had to be overcome in order to clear its path. The amazing agreement between the predictions of perturbative quantum electrodynamics (QED) and experimental spectroscopic data is, to this day, a textbook example of what an exact science should ideally be. On the other hand, the very idea of renormalization is by itself one of unusual charm and originality (at least in the opinion of the author), and the proof that a large class of quantum field theories can benefit from it was one of the greater achievements of physics in the last few decades.

Despite all that, the issue remains, unfortunately, that quantum field theories (QFT) have until now been very stingy in providing us with exact results, or intuition about their more exotic properties. Even if one disregards the worries of axiomaticians that the very foundations of these theories are, in most cases, mathematically shaky and that the perturbation series themselves may be nonconvergent and thus ill-defined, the question of the applicability of perturbation theory at all when the perturbation parameter is not small, and of the ability of such an approach to unmask *all* the qualitative features of the full theory, still remains an important one. The canonical example is quantum chromodynamics (QCD): There, due to the fact that the coupling constant between quarks and gluons is large, perturbation theory is very unlikely to give reliable answers at any level. Moreover, phenomena such as confinement of quarks in color singlet combinations and chiral symmetry breaking are hardly likely to be reproduced or even hinted at by ordinary perturbative methods.

There was, thus, plenty of motivation for physicists to look for nonperturbative approaches to QFT. These approaches can be roughly classified into two categories: Methods that give some approximate or qualitative answers to questions beyond the range of perturbation theory, and methods that provide us with an *exact* answer to a problem. The distinction is not razor-sharp, naturally, and some methods may partially produce answers of both types. Further, some of these methods may use perturbative techniques alongside with any other gadgetry they may involve. The combination, of course, should give something more than perturbation alone. Examples of such methods include lattice calculations and simulations, $1/N$ techniques, exact solutions in two dimensions, either of the back-scattering (Bethe ansatz) or the operator variety, solutions of the Schwinger-Dyson equations and, last but not least, topological considerations.

This last approach, as suggested by the title, will mostly concern us in this treatise. Its power and beauty stem from its ability to relate and connect topological properties of the underlying classical theories with analytical (exact) aspects of the corresponding QFT. In this connection, topological methods have not only unmasked such nonperturbative peculiarities as monopoles and instantons, quantization conditions and global anomalies, but also contributed to our intuition and understanding of otherwise perturbative QFT subtleties such as chiral anomalies and charge fractionization. For this, the contribution of topology to our grasp of basic features of QFT is (again, in the author's opinion) all but invaluable.

Among the previously mentioned phenomena, the one concerning the anomalous and unusual quantum numbers induced by the presence of a soliton (this meaning any topologically stabilized configuration of external fields) and the resulting charge fractionization, is, perhaps, the most peculiar and intriguing. It is, further, no less physically interesting: such soliton configurations, apart from the celebrated Dirac monopole, include flux tubes, cosmological strings and domain walls. In addition, a phenomenologically interesting and quite successful model of baryons is the Skyrme model. In this, only fields corresponding to the meson degrees of freedom are assumed (and no quark fields whatsoever), and baryons are simply localized soliton configurations of these meson fields. This immediately solves the confinement problem and reproduces some of the fundamental properties of baryons (most notably, the I=J rule), but faces the problem of accounting for the fermionic properties of an object made purely of boson fields. This is readily and beautifully solved if one considers the fermion number of the baryons to be induced by their nontrivial topology, in the fashion that will be exposed in the next section. Finally, there are some analogies that suggest more than a superficial similarity between the induced vacuum current in a specific case (planar QED) and the quantum Hall effect, that itself is one of the most intriguing and amazing recently discovered physical phenomena.

Anomalous quantum numbers and topological properties of QFT are, thus, the topics that will be examined, and hopefully connected, in this treatise. Since the literature on these subjects is already quite extensive, no serious effort for a review or presentation of a complete list of references will be attempted here. Rather, we will be limited to the author's own modest contribution to the subject. As a result, the material closely parallels and tracks the contents of the set of papers that the latter has produced on the matter. These will constitute the null reference ([0]) in each relevant chapter to follow. The material to be covered is, roughly, as follows:

In chapter 1, we will give a physical exposé of the induced vacuum fermionic charge, concentrating on its basic features and giving a simple derivation of its value, and examine its dependence on local features of the inducing field configuration and its connection with anomalies and the bosonization method.

In chapter 2, we concentrate on the quantum numbers induced by a magnetic flux tube configuration. We calculate the induced charge and angular momentum, and discuss the several interesting qualitative properties of these quantities.

In chapter 3, we "take off" a bit mathematically and examine the effect of the fermion field boundary conditions on induced vacuum quantities and the celebrated Atiyah-Patodi-Singer index theorem. We derive a version of the index theorem using local boundary conditions, physically preferable to the global ones used so far in the literature, and clarify some issues concerning the observability of the Dirac string and the proper definition of the angular momentum operator.

In chapter 4, we concentrate on purely topological aspects and derive a quantization condition for the so-called topological mass term of the photons in odd dimensional QED, such that states with nonzero magnetic flux exist in the theory. We also give a topological argument demonstrating that parity will be broken in such theories. As a byproduct of our reasoning, we point out an incompleteness in the standard proof of the nonperturbative SU(2) anomaly in four dimensions and provide an argument that completes this proof.

In chapter 5, we exploit a topological construction to show that the topological mass term of abelian gauge bosons induces nontrivial angular momentum and statistics on flux tubes, a result missed by careless reasoning.

Finally, in chapter 6, we examine the global symmetry breaking patterns of odd dimensional QED and the relation of QED_3 with the quantum Hall effect. We show that, for even number of fermion flavors, either parity or flavor symmetry must break, and, for odd number of flavors, flavor symmetry may or may not break but parity must break. In the same context, we demonstrate that the mathematical settings of the two phenomena (parity breaking and the quantum Hall effect), if not their specific physical mechanisms, are identical.

We sincerely hope that reading this manuscript will be a painless, if not enjoyable, experience.

CHAPTER 1

A Physical Look at Vacuum Fermionic Charge

I. Introduction.

The phenomenon of vacuum charge induced by background fields with nontrivial topology, and of the resulting fermion number fractionization, is by now one of the most widely studied in the literature [1]. In particular, the 1+1 dimensional case has practically been flogged to death. The appearance and value of the vacuum charge has been understood in terms of vacuum degeneracy [2] (in a special case), diagrammatic techniques [3], the η-invariant of the Dirac hamiltonian [4], anomalies [5,6], Levinson's theorem and scattering phase shifts [7], time-delay of fermions scattered off the external fields [8] etc.

What makes the induced vacuum charge both interesting and amenable to a complete solution is its connection with the topology of the field that induces it. All the previously mentioned approaches exploit, one way or the other, the connection between the analytical properties of the quantum theory and the topology of the background fields. Nontopological aspects, like the dependence of the charge on local features of the inducing field and level crossings of the hamiltonian, have also been studied extensively [9].

The purpose of this chapter is to give a simple and physical demonstration of the existence of the vacuum charge and calculation of its value, as well as to examine and understand its connection with some other aspects of the field theories that support it. Specifically, it is shown that some very general principles, like locality and charge conservation, together with surprisingly little information about the specific properties of the theory, are enough to find an expression for the vacuum charge. Also, the connection with anomalies and the relevance of the regularization used, the effects of the size of the background field configuration, and the bosonization method are given a closer look. Most of the analysis is done in the context of a 1+1 dimensional model, but a generalization to higher-dimensional cases is given in the last section.

The organization of this chapter is as follows: in section II we give a simple physical derivation of the vacuum charge formula using general principles and examine the vacuum charge density in some specific cases of interest; in section III we give a simple derivation of the same formula that demonstrates the connection with the anomaly in the most direct possible way; in section IV we examine the relevance of the size of the charge-inducing soliton and give some exact criteria for the validity of the standard formula; in section V we expound on the connection with the anomaly and examine the differences between different regularization schemes and their relevance for the definition and conservation properties of the charge; in section VI we examine the problem in its bosonized form and show how the level-crossing phenomenon manifests itself in this picture; finally, in section VII we conclude with some generalizations of the previous discussions to a slightly more general class of fields, as well as to the 2+1 and 3+1 dimensional cases.

II. A simple physical derivation of the vacuum charge formula.

Let us examine the simplest case of induced vacuum charge due to vacuum polarization. Consider 2-component Dirac fermions in 1+1 dimensions interacting with an external chiral U(1) field. The lagrangian density is:

$$L = \overline{\psi} i \partial\!\!\!/ \psi - m \overline{\psi} e^{i\alpha\gamma_5} \psi. \tag{2.1}$$

Here, "γ^5" denotes the axial γ-matrix in 1+1 dimensions and α is a real scalar field. Note that the sign of the interaction term is a matter of convention, since it can be changed under the redefinition $\psi \to \gamma^5 \psi$.

We will be concerned with external static α-field configurations that reach asymptotically constant values as the spatial coordinate x goes to $\pm$ infinity:

$$\alpha(x = \pm\infty) = \alpha_{\pm}. \tag{2.2}$$

This corresponds to a localized "soliton" configuration, with fractional winding number equal to $\frac{\Delta\alpha}{2\pi} \equiv \frac{\alpha_+ - \alpha_-}{2\pi}$. (If we compactified our space into a circle S^1, this winding number would have to be integer and would measure $\pi_1(\mathrm{U}(1))$.) The question addressed is: Is there vacuum charge induced by this soliton, and how much?

The basic assumption in answering this question here, will be that there is a regularization of the charge operator such that the total charge is conserved, even under the presence of time-varying chiral field configurations $\alpha(x,t)$ (such an explicit regularization will be demonstrated later). Under this assumption, the expression for the vacuum charge immediately follows!

To see that, consider first a static configuration of many widely separated infinitesimal "steps" of the field α, of size $\pm\delta\alpha$ (fig. 1a). Notice that, for each value of α, under the chiral rotation

$$\psi \to e^{-\frac{i}{2}\alpha\gamma_5}\psi \tag{2.3}$$

the lagrangian (2.1) can be transformed into the lagrangian of a fermion with mass m. So, around each step we deal with a massive fermion interacting with an infinitesimal chiral field $\pm\delta\alpha$. Assuming that the distance between the steps is large compared with the correlation length of the fermions $\frac{1}{m}$, we conclude that the different steps do not "see" each other, and so the total induced charge is the sum of the charges each one of the steps would induce alone. Moreover, since both the charge and the field α are odd under charge conjugation, steps of size $-\delta\alpha$ induce the opposite charge than steps $+\delta\alpha$. We denote this charge by

$$\delta Q \equiv k\delta\alpha, \tag{2.4}$$

with k some coefficient, defined by (2.4), that could actually vanish. Adding the charges of each step, we get

$$Q = \sum \pm \delta Q = k \sum \pm \delta\alpha = k\Delta\alpha. \tag{2.5}$$

Now we can think of gradually bringing together the steps, so as to create a (sawtooth-like) given soliton profile (fig. 1b). Finally, we deform each step so as to smooth out the profile and end up with a desired soliton configuration (fig. 1c). Doing these transformations slowly enough (adiabatically), ensures that we do not excite any states in the process, but we continuously follow the evolution of the initial vacuum state. During neither of these transformations is there a net charge generation, since they overall constitute a time-dependent field $\alpha(x,t)$, and the vector current is conserved under such a process. The only possibility is that there may have been a net charge influx from infinity. However, we did not touch the asymptotic values of the field at infinity, and so there can be no currents generated there. So we conclude that formula (2.5) correctly gives the charge induced by an arbitrary soliton, up to an unknown coefficient. To determine this coefficient, we consider the "step function" configuration of figure 2a. Obviously, a step with $\Delta\alpha = 2\pi$ is invisible when exponentiated in (2.1) and corresponds to $\Delta\alpha = 0$. The induced charge of this configuration then should vanish. However, formula (2.5) gives $Q = 2\pi k$. The only explanation of this discrepancy, other than $k = 0$, is that, during the adiabatic buildup of this configuration, there has been some levels of the Dirac hamiltonian crossing zero, and so the final state reached is not the vacuum but an excited state, with some positive-energy levels filled (and possibly some negative-energy levels empty) (fig. 2b). This state has an integral charge, corresponding to the number of positive energy levels that are filled or (minus) the number of negative energy levels that are empty. This already tells us that $2\pi k$ has to be an integer. Examining the Dirac hamiltonian for the field of fig. 2a, we can easily find with an explicit calculation of the energy levels of the hamiltonian that there is only one upwards level crossing in the process of varying $\Delta\alpha$ from 0 to 2π, corresponding to the self-charge-conjugate value $\Delta\alpha = \pi$. So we conclude that $2\pi k = 1$, and the final formula is

$$Q = \frac{\Delta\alpha}{2\pi}, \tag{2.6}$$

with possible jumps of ± 1 at configurations where level crossings occur. For the step function configuration of figure 1d, the vacuum charge as a function of $\Delta\alpha$ is given in figure 2c.

For the special value $\Delta\alpha = \pi$, we can even calculate exactly the vacuum charge *density*. Since $\pi \sim \pi - 2\pi = -\pi$, as already mentioned this configuration is self-charge-conjugate. Due to the presence of the zero-energy mode, the vacuum for this value is doubly degenerate, the two states being connected with charge conjugation:

$$C|+>=|->, \tag{2.7}$$

where $|+>$, $|->$ are the two degenerate vacua and C is the charge conjugation operator. Since the charge density operator j^0 is odd under C:

$$Cj^0C = -j^0, \tag{2.8}$$

we conclude that

$$<-|j^0|->= -<+|j^0|+> . \tag{2.9}$$

On the other hand, we know that the two vacua differ only in the fact that the one (say, $|+>$) has the zero-mode filled, while the other has it empty. So, their vacuum currents differ exactly by the currents of this zero-mode:

$$<+|j^0|+>=<-|j^0|-> +\rho_0 \tag{2.10}$$

where $\rho_0 = \psi_0^\dagger \psi_0$ is the charge density of the zero-mode. So we conclude that

$$<\pm|j^0|\pm>= \pm\tfrac{1}{2}\rho_0. \tag{2.11}$$

It is easy to find that the zero-mode solution has the form

$$\psi_0 = m^{\frac{1}{2}} e^{-m|x|} u, \quad \text{with } u^\dagger u = 1 \tag{2.12}$$

(the step being at $x = 0$). The spinor u depends on the specific representation of γ-matrices that we use. We thus find

$$< \pm|j^0|\pm >= \pm\tfrac{1}{2} m e^{-2m|x|}. \tag{2.13}$$

So we see that the charge is localized around the (zero-size) soliton, within a few correlation lengths of the fermions (fig. 2d).

A possible puzzle may arise here: Consider two widely separated step function solitons, with $\Delta\alpha = \pm\pi$ (fig 3a). From what we said in the beginning of this section, we expect the vacuum charge density to be the superposition of the densities induced by each step alone, i.e., two functions of the form (2.13), centered around each step with opposite signs (the sign of the left one being + or − depending on whether α reaches the value π from below or above, respectively). However, this configuration is still self-charge-conjugate, and it is easy to check that, for any finite separation between the steps, there is no zero-mode of the hamiltonian. So the vacuum is nondegenerate, and repeating the steps (2.8-9) with $|+ >= |- >$ we see that $j^0 = 0$, in contradiction with our previous conclusion.

The reason for this discrepancy is that, the configuration of the figure 3a corresponds to fermions with mass m outside of the steps and $-m$ between the steps (or vice versa). Thus, around each step the fermions can *not* be thought of as having a mass m plus a (small) chiral perturbation, and the localization argument of the discussion following eq. (2.3) breaks down. What in fact happens is that the two would-be zero-modes, produced by each step if alone, due to the finite separation of the steps split and acquire energies $\pm\epsilon = \pm m e^{-mL}$, for $mL >> 1$. So, there are now three states that are *almost* degenerate with the true vacuum $|0 >$, call them $|+ >$, $|- >$ and $|\bar{0} >$, with energies $+\epsilon$, $+\epsilon$, $+2\epsilon$ and charge $+1$, -1 and 0, correspondingly (fig. 3b). Denoting by $a^\dagger$ the operator creating a fermion of energy $+\epsilon$ and by $b^\dagger$ the operator annihilating a fermion of energy $-\epsilon$, these states can be expressed as

$$|+>= a^\dagger|0>, \quad |->= b^\dagger|0>, \quad |\bar{0}>= a^\dagger b^\dagger|0>\,. \tag{2.14}$$

The normal-ordered charge density operator can be written

$$j^0 =: \psi^\dagger\psi := a^\dagger a\,\psi_a^\dagger\psi_a - b^\dagger b\,\psi_b^\dagger\psi_b + a^\dagger b^\dagger\,\psi_a^\dagger\psi_b + ba\,\psi_b^\dagger\psi_a + \cdots \tag{2.15}$$

where the dots stand for terms containing creation or annihilation operators of modes other than the ones under consideration. The wavefunctions ψ_a and ψ_b, for $mL >> 1$, are

$$\psi_a = \frac{1}{\sqrt{2}}(\psi_L + \psi_R), \quad \psi_b = \frac{1}{\sqrt{2}}(\psi_L - \psi_R) \tag{2.16}$$

with ψ_L and ψ_R wavefunctions of the form (2.12) centered around the left and right step correspondingly.

If we now define the normalized states

$$\begin{aligned} |+,+> &= |+>, \quad |+,->= \frac{1}{\sqrt{2}}(|0> + |\bar{0}>) \\ |-,-> &= |->, \quad |-,+>= \frac{1}{\sqrt{2}}(|0> - |\bar{0}>) \end{aligned} \tag{2.17}$$

it is straightforward to calculate the charge density for each one of these states to be

$$\begin{aligned} <+,+|j^0|+,+> &= \tfrac{1}{2}(\rho_L + \rho_R) \\ <+,-|j^0|+,-> &= \tfrac{1}{2}(\rho_L - \rho_R) \\ <-,+|j^0|-,+> &= \tfrac{1}{2}(-\rho_L + \rho_R) \\ <-,-|j^0|-,-> &= \tfrac{1}{2}(-\rho_L - \rho_R) \end{aligned} \tag{2.18}$$

where $\rho_{L,R}$ have the obvious connotation (we used the fact that ψ_L and ψ_R have negligible overlap). So the vacuum for finite L is a linear combination of the (degenerate) vacua for $L = \infty$, with vanishing charge density. For two general solitons separated by a great distance, that do not have the property that each one of them alone creates a zero-mode, this (almost) degeneracy of the vacuum does not arise and the vacuum charge density is just the sum of the densities created by each soliton alone.

III. A simple calculation based on the anomaly.

We proceed now to give another simple proof of formula (2.6), demonstrating in a direct way the anomalous nature of the vacuum currents.

We will again consider the step function potential of fig. 2a, and we will imagine that we adiabatically turn on the parameter $\Delta\alpha$ from 0 to its final value. We will adopt a timelike point-splitting regularization for the current operator. (This regularization has the property assumed in section II, as will be demonstrated in section V.) So, in particular, the spatial component of the current is defined

$$j^1_\tau = \overline{\psi}(t+\tau)\gamma^1\psi(t), \tag{3.1}$$

with τ the regularization parameter that will eventually be drawn to zero. If we perform now the chiral rotation

$$\psi \to e^{-\frac{i}{2}\alpha(x)\gamma^5}\psi \tag{3.2}$$

we see that the α field totally disappears, and the new fermion field satisfies a free massive Dirac equation. Moreover, since all external fields vanish at infinity, no currents are induced there in this "rotated" version of the problem, as we adiabatically switch on $\Delta\alpha$.

The only remnant of the original field $\alpha(x)$ in the rotated version is, in fact, a nontrivial boundary condition at the origin. Indeed, since the original field was continuous at $x = 0$, the new field satisfies the boundary condition

$$\psi(x = 0-) = e^{-\frac{i}{2}\Delta\alpha\gamma^5}\psi(x = 0+). \tag{3.3}$$

As we will show, this boundary condition acts as a source of charge, during the switching on of $\Delta\alpha$. To see this, use (3.3) to connect the current on the right of $x = 0$ with the current on the left:

$$\begin{aligned} j_{\tau,R} &= \overline{\psi}(t+\tau)\gamma^1\psi_R(t) \\ &= \overline{\psi}_L(t+\tau)e^{\frac{i}{2}\Delta\alpha(t+\tau)\gamma^5}\gamma^1 e^{\frac{i}{2}\Delta\alpha(t)\gamma^5}\psi_L(t) \\ &= \overline{\psi}_L(t+\tau)\gamma^1 e^{-\frac{i}{2}(\Delta\alpha(t+\tau)-\Delta\alpha(t))\gamma^5}\psi_L(t) \\ &= \overline{\psi}_L(t+\tau)\gamma^1\psi_L(t) - \frac{i}{2}\tau\Delta\dot{\alpha}\overline{\psi}_L(t+\tau)\gamma^1\gamma^5\psi_L(t) + \cdots \end{aligned} \tag{3.4}$$

where overdot denotes time derivative. So, the overall outflowing current from the origin is

$$j^1_{\tau,out} = j^1_{\tau,R} - j^1_{\tau,L} = -\frac{i}{2}\tau\Delta\dot{\alpha}\overline{\psi}_L(t+\tau)\gamma^1\gamma^5\psi_L(t) + \cdots \tag{3.5}$$

But, from operator product expansion, we know that

$$\overline{\psi}(t+\tau)\gamma^1\gamma^5\psi(t) = \langle\overline{\psi}(t+\tau)\gamma^1\gamma^5\psi(t)\rangle + \cdots = \frac{i}{\pi\tau} + O(1). \tag{3.6}$$

From (3.5) and (3.6) we thus see

$$j^1_{out} = \lim_{\tau\to 0} j^1_{\tau,out} = \frac{\Delta\dot{\alpha}}{2\pi}. \tag{3.7}$$

Since, as we said, no charge comes from infinity, the whole charge of the system comes from this anomalous charge outflow. So

$$Q = \int_0^t j^1_{out}dt = \frac{\Delta\alpha}{2\pi} \tag{3.8}$$

which is the formula found earlier. (Of course we still have to take into account the level crossings.) Going now back to the unrotated version, we can deform this step function profile to reach any configuration we wish with the same asymptotic values, and, due to the conservation of charge and the fact that nothing comes from infinity during the deformation process, the charge of the resulting soliton will be given by (3.8), modulo an integer.

In the previous derivation, there was the implicit assumption that the charge of the vacuum in the original and the rotated case are equal. A justification of this fact and some more comments will be given in the section about regularization.

The previous construction contains the nucleus of the idea behind the connection between vacuum charge and chiral anomaly, as will be elaborated in section V. Specifically, we have isolated the generation of charge in a single anomalous current source, coming from a nontrivial (chiral) boundary condition imposed on the fermion field. This should be reminiscent of the 3+1 dimensional chiral bag case [10], where again a nontrivial boundary condition at the surface of the bag "pumps" charge into the bag as we turn the chiral parameter of the boundary on [11].

IV. Criteria for the validity of the standard formula for the vacuum charge.

We saw previously that the standard formula (3.8) for the vacuum charge may actually not be strictly valid, because it may apply to a state that is not the vacuum, but rather a state with one or more positive energy levels filled or negative energy levels empty. The criterion for the validity of this formula is the "smoothness" of the soliton profile. We expect that, if the length scale of variation of the soliton is large compared with the Compton wavelength of the fermions, then the previous formula is valid. On the other hand, if the size of the soliton is much smaller than the Compton wavelength, we expect it to be essentially equivalent to a step potential soliton, in which case one has to subtract the nearest integer in order to find the true vacuum charge. This section makes these remarks quantitative.

To find a criterion for smoothness of the soliton, we first prove the following theorem:

Theorem 1: If $|\partial_x \alpha(x)| \leq 2m$ for all x, then the Dirac hamiltonian has no zero energy levels.

Proof: Let us first choose the explicit γ-matrix representation

$$\gamma^0 = \sigma_1, \quad \gamma^1 = i\sigma_3, \quad \gamma^5 = \sigma_2 \tag{4.1}$$

and define

$$\psi \equiv \begin{bmatrix} u \\ v \end{bmatrix}, \quad \tfrac{1}{2}\partial_x \alpha \equiv h. \tag{4.2}$$

Then a zero-energy (time-independent) solution of the Dirac equation after performing the chiral rotation (3.2) satisfies

$$\frac{d}{dx}\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} -m & h \\ -h & m \end{bmatrix}\begin{bmatrix} u \\ v \end{bmatrix}. \tag{4.3}$$

Notice that eq.(4.3) is real and so u and v can be chosen to be real for all x.

In the far left ($x \to -\infty$), where $\alpha \to$ constant $\Rightarrow h \to 0$, the wavefunction takes the form

$$\psi_- = Ae^{mx}\begin{bmatrix} 0 \\ 1 \end{bmatrix}. \tag{4.4}$$

Similarly, in the far right, the wavefunction becomes

$$\psi_+ = Be^{-mx}\begin{bmatrix} 1 \\ 0 \end{bmatrix}. \tag{4.5}$$

(These are the only forms that do not blow up at $x \to \pm\infty$.) So, equation (4.3) has to be able to "drive" the spinor from its "down" form at $x = -\infty$ into its "up" form at $x = +\infty$. But we can show that if $v \geq |u|$ at some $x = x_0$, then $v \geq |u|$ for *all*

$x > x_0$. Indeed, using (4.3) and the condition $|h| \leq m$, we have

$$\begin{aligned} \frac{d}{dx}(v-u)a &= -hu + mv + mu - hv = (m-h)(v+u) \geq 0, \\ \frac{d}{dx}(v+u) &= -hu + mv - mu + hv = (m+h)(v-u) \geq 0 \end{aligned} \quad . \tag{4.6}$$

So conditions $v \geq u$ and $v \geq -u$ (i.e., $v \geq |u|$) are preserved and remain valid for all $x \geq x_0$.

Now since at the far left, because of (4.4), the above condition holds (we can always choose $A > 0$), we conclude that the form (4.5) can never be reached, which proves our statement that (4.3) has no normalizable solutions.

Let us remark that, if we tried to reduce (4.3) into a single, Schroedinger-like second order equation, depending on the choice of variable, we would end up either with velocity-dependent potentials (terms containing first derivatives of the wavefunction), involving derivatives of h, or with imaginary (absorptive) potential terms. In either case, the derivation of the above result would not be obvious.

With the help of the previous theorem we can conclude now that a soliton with $|\partial_x \alpha| \leq 2m$ everywhere induces a vacuum charge given exactly by formula (3.8). Indeed, we can think of adiabatically building up this soliton by interpolating between $\alpha(x) = 0$ and the final soliton shape. The charge that comes from infinity during this process is just the one given by (3.8). Since for every intermediate shape of the soliton the previous condition on the derivative of α keeps holding, we see that there can be no energy levels crossing zero, and so the final state reached is the true vacuum.

To deal with sharp solitons, we prove the next theorem:

Theorem 2: If $\alpha(x)$ is nonconstant only in a region of length L, then the Dirac hamiltonian has a zero mode for $\Delta\alpha = (2n+1)\pi + \delta$, where n is an integer and $|\delta| < 2mL$.

Proof: Define $u = \rho \cos\frac{\theta}{2}$, $v = \rho \sin\frac{\theta}{2}$. Then, from (4.3) we can derive that $\theta(x)$

for a zero-energy solution satisfies

$$\frac{d\theta}{dx} = \frac{d\alpha}{dx} - 2m\sin\theta. \tag{4.7}$$

Taking into account (4.4) and (4.5), we have that θ must have the limiting behaviour

$$\theta(-\infty) = 0, \quad \theta(+\infty) = (2n+1)\pi. \tag{4.8}$$

Integrating (4.7) from $-\infty$ to $+\infty$ we get

$$\Delta\alpha - 2m\int_{-\infty}^{+\infty} \sin\theta dx = (2n+1)\pi \tag{4.9}$$

and since

$$|\delta| \equiv \left|2m\int_{-\infty}^{+\infty} \sin\theta dx\right| \leq 2m\int_{-\infty}^{+\infty} |\sin\theta| dx \leq 2mL \tag{4.10}$$

we obtain the desired expression. In (4.10), the range of integration of $|\sin\theta|$ has been reduced to the interval where $\alpha(x)$ is nonconstant, because we know that for a constant α the solution must have exactly the form (4.4) or (4.5) for which $\sin\theta$ vanishes.

If now we again build up this soliton adiabatically by interpolating from the $\alpha = 0$ configuration, every time that $\Delta\alpha$ passes from the neighborhood of an odd multiple of π, that is, within a distance smaller than $2mL$ (we assume that $2mL \ll \pi$), a level crossing will arise, and so a unit of charge will have to be subtracted (or added, if $\Delta\alpha < 0$). If the final $\Delta\alpha$ is not too near an odd multiple of π, i.e., within $2mL$, the vacuum charge of the soliton will be given by the distance of $\frac{\Delta\alpha}{2\pi}$ from the nearest integer.

As an example, and to demonstrate that the criterion derived from theorem 1 is the strictest one based upon an overall upper bound of $|\partial_x\alpha|$ alone, consider a linear solitonic profile of the form (fig. 4)

$$\alpha(x) = \frac{\Delta\alpha}{L}x, \text{ for } -\frac{L}{2} < x < \frac{L}{2}, \text{ otherwise } = 0. \tag{4.11}$$

It is easy to see that, for $h = \frac{\Delta\alpha}{2L} < m$, (4.3) has no normalizable solutions. For $h > m$, imposing continuity at $x = \pm\frac{L}{2}$, we find that the solutions have oscillatory behavior between $-\frac{L}{2}$ and $\frac{L}{2}$ with wavenumber $k = \sqrt{h^2 - m^2} > 0$, satisfying

$$\begin{aligned} m - ik &= \pm h e^{ikL}, \text{ or} \\ -\frac{k}{m} &= \tan kL, \end{aligned} \tag{4.12}$$

where + (−) corresponds to even (odd) in x solutions. We notice that for each interval $[n\pi, (n+1)\pi]$ of kL, (4.12) has exactly one solution. In particular, this means that there is a zero-mode for

$$0 < kL < \pi \;\Rightarrow\; h^2 - m^2 < \frac{\pi^2}{L^2}. \tag{4.13}$$

So, by choosing L large enough, the difference between h and m can be made arbitrarily small, and the bound of theorem 1 can be saturated. For $mL \ll 1$, the solutions for $\Delta\alpha$ approach the values

$$\Delta\alpha = (2n+1)\pi + \frac{2mL}{(n+\frac{1}{2})\pi} \tag{4.14}$$

that are in accordance with theorem 2. Also notice that, varying L between 0 and $\frac{\Delta\alpha}{2m}$, kL varies from $\frac{\Delta\alpha}{2}$ to 0, and since for L small the solutions for kL are close to half-integer multiples of π, we have overall $\left[\frac{\Delta\alpha}{2\pi} + \frac{1}{2}\right]$ level crossings ($[\cdots]$ stands for integer value), accounting for the difference between the vacuum charges of a thin and a broad soliton, as calculated earlier.

V. Relevance of the regularization in the definition of vacuum currents.

So far, we have just assumed the existence of an appropriate regularization for the current, that conserves the total charge in the presence of arbitrary time-varying solitons, and claimed that timelike point-splitting is such a regulator, without actually proving it. In this section we clarify this matter and demonstrate the relevance of the regularization procedure in the properties of the currents and the exact value of the vacuum charge.

As an example and initial motivation for this scrutiny, we examine the results of reference 6, where the Schwinger anomalous commutator is used to calculate the vacuum charge. We repeat here the main line of the arguments.

Let us start with an initial lagrangian of the form (2.1), and perform a chiral rotation as in (2.3). Due to the existence of an anomalous commutator between the vector and axial vector currents, the charge operator transforms nontrivially under this chiral rotation. Specifically, using the Schwinger result

$$[j^0(x), j_5^0(y)] = \delta'(x-y) \tag{5.1}$$

for the equal-time commutator of the charge density with the generator of chiral rotations, we can see that

$$Q_{original} = Q_{rotated} + \frac{1}{2\pi}\int \partial_x \alpha dx. \tag{5.2}$$

By arguing that in the rotated case the vacuum charge is *zero*, one then concludes that the charge of the original vacuum is given by the standard formula.

The assumption, however, of vanishing vacuum charge in the rotated case contradicts not only our assumption of section III that the charge in the rotated version is the same as the charge in the original version, but also the well-known expression of the vacuum charge in terms of the spectral asymmetry of the Dirac hamiltonian. Specifically, if we represent the vacuum charge as the (regulated) difference between the Dirac sea charge of the hamiltonian and the Dirac sea charge of the trivial ($\alpha(x) \equiv 0$) hamiltonian, properly symmetrized in order to have the correct transformation properties under charge conjugation, we end up with the expression

$$Q = -\frac{1}{2} \lim_{\epsilon \to 0+} \sum_n sign(E_n) e^{-\epsilon |E_n|} \qquad (5.3)$$

where E_n are the eigenvalues of the Dirac hamiltonian. The same expression is obtained using (euclidean) timelike point-split regularization for the charge operator. In this expression, the vacuum charge is a functional of only the spectrum of the hamiltonian.

It is obvious now that the original and the rotated Dirac hamiltonian have identical spectra, since the chiral rotation just maps an eigenfunction of the one into an eigenfunction of the other with the same eigenvalue. Thus, formula (5.3) tells us that the original and the rotated situations have *equal* vacuum charges, in contradiction with formula (5.2). What happens?

The answer to this puzzle is that the charge, as well as the anomalous commutator, are regularization-prescription-dependent. In fact, in the derivation of formula (5.1), a specific point-splitting regularization is used, different than the one we used previously and the one that leads to formula (5.3). With this regularization one should check what are the conservation properties of the vector current before concluding anything about the vacuum charge of any configuration involving soliton and axial gauge fields.

Instead of dealing right now with this prescription, we return to our own timelike point-splitting and find what the properties of the current are and how the original and rotated versions are connected, showing that the two pictures are compatible. In the original lagrangian, there are no axial gauge fields but there exists a chiral U(1) field. We will denote the current in this version $\overline{\psi}_o \gamma^\mu \psi_o = j_o^\mu$. In the rotated version the lagrangian is

$$L = \overline{\psi}_r \gamma^\mu (\partial_\mu - A_{\mu 5} \gamma^5) \psi_r - m \overline{\psi}_r \psi_r. \qquad (5.4)$$

In this version there is no U(1) field (just a mass term), but there exists an axial gauge field given by

$$A_{\mu 5} = -\tfrac{1}{2}\partial_\mu \alpha. \tag{5.5}$$

We will denote the current by $\overline{\psi}_r \gamma^\mu \psi_r = j_r^\mu$. Using the relation between original and rotated fermion fields and the same operator product expansion procedure as in section III, we find that the two point-split currents are connected by

$$j_o^\mu = j_r^\mu - \frac{1}{2\pi}\delta_1^\mu \partial_0 \alpha. \tag{5.6}$$

So in static cases, where $\partial_0\alpha = 0$, the two currents are the same. In particular, the charge densities are always the same, which proves our assumption that the vacuum charge of the two situations is the same.

In order to find the conservation properties of these currents, we evaluate the divergence of their point-split expressions. We will use the equations of motion of ψ_o and ψ_r

$$i\gamma^\mu \partial_\mu \psi_o = m e^{i\alpha\gamma^5}\psi \tag{5.7a}$$

$$\gamma^\mu (i\partial_\mu + \tfrac{1}{2}\partial_\mu \alpha \gamma^5)\psi_r = m\psi_r \tag{5.7b}$$

(that still hold, coming from a point-split regulated action), as well as the results

$$\left\langle \overline{\psi}(\vec{x}+\vec{\tau})\gamma^\mu \psi(\vec{x})\right\rangle = \frac{i}{\pi}\frac{\tau^\mu}{\tau^2} + O(1) \tag{5.8a}$$

$$\left\langle \overline{\psi}(\vec{x}+\vec{\tau})\gamma^\mu \gamma^5 \psi(\vec{x})\right\rangle = \frac{i}{\pi}\epsilon^{\mu\nu}\frac{\tau_\nu}{\tau^2} + O(1) \tag{5.8b}$$

(where arrow indicates a spacetime vector) that hold for both ψ_o and ψ_r, since the most singular terms in $|\tau|$ are independent of the external fields. The expectation values $\left\langle \overline{\psi}(\vec{x}+\vec{\tau})\psi(\vec{x})\right\rangle$ and $\left\langle \overline{\psi}(\vec{x}+\vec{\tau})\gamma^5\psi(\vec{x})\right\rangle$ do not contain a term of order $|\tau|^{-1}$.

Then, a straightforward calculation, using operator product expansion, and taking τ^μ purely timelike, gives

$$\partial_\mu j_o^\mu = 0 \tag{5.9a}$$

$$\partial_\mu j_r^\mu = \frac{1}{2\pi}\partial_1\partial_0\alpha. \tag{5.9b}$$

Notice that equations (5.9) are compatible with (5.6). This verifies our assumption that a timelike point-split regularization conserves the charge in the presence of arbitrary α-fields (but not in the presence of axial gauge fields).

Now we can build up our soliton, starting from a vanishing α-field and see what happens. In the original version charge is conserved, and so all the induced charge has to come from infinity. The current at infinity can only depend on the time derivatives of α, since a constant value of α can always be subtracted with an *a priori* global chiral rotation. Also, in an adiabatic calculation, only the first order term in the first derivative of α can contribute to the vacuum charge. From dimensions, such a term can only have the form $k\dot{\alpha}$, with k some numerical coefficient. This coefficient can easily be calculated *à la* Goldstone and Wilczek, or alternatively can be deduced from the level-crossing arguments of section II, and we obtain the standard result.

In the rotated version, however, the axial gauge field vanishes at infinity throughout the whole process, since α becomes a constant there. So, no currents are induced at infinity. The whole charge of the vacuum is thus due to the nonconservation of the vector current (5.9b). Integrating this anomalous contribution, we have

$$Q_r = \int_0^T dt \int_{-\infty}^{+\infty} dx\partial_\mu j_r^\mu = \frac{\alpha_+ - \alpha_-}{2\pi}. \tag{5.10}$$

So we see that, indeed, the charge of the two versions is the same.

We now return to a regularization procedure that conserves the vector current in both rotated and unrotated cases [6]. To do this, notice that, in 1+1 dimensions, since

$$\gamma^\mu \gamma^5 = -\epsilon^{\mu\nu}\gamma_\nu, \tag{5.11}$$

an axial gauge field is equivalent to a vector gauge field

$$A_\mu = \epsilon_{\mu\nu} A_5^\nu. \tag{5.12}$$

If we now define the point-split current in the usual gauge-invariant way, by inserting a path-ordered gauge field exponential between the fermion fields at $\vec{x}$ and at $\vec{x}+\vec{\tau}$, for infinitesimal τ we have

$$\begin{aligned} j_r^\mu &= \overline{\psi}_r(\vec{x}+\vec{\tau})\gamma^\mu e^{i\tau^\nu A_\nu}\psi_r \\ &= \overline{\psi}_r(\vec{x}+\vec{\tau})\gamma^\mu \psi_r(\vec{x}) - \frac{i}{2}\epsilon^{\nu\rho}\tau_\nu \partial_\rho \alpha \overline{\psi}_r(\vec{x}+\vec{\tau})\gamma^\mu \psi_r(\vec{x}). \end{aligned} \tag{5.13}$$

j_o^μ does not get any extra contributions. When calculating the difference between j_r^μ and j_o^μ, we get, now, two terms: one from the point-splitting of the current itself and one from the extra term. Then we average over all possible directions of $\vec{\tau}$ and take the limit $|\tau| \to 0$. The two terms contribute equally and the final result is

$$j_o^\mu = j_r^\mu + \frac{1}{2\pi}\epsilon^{\mu\nu}\partial_\nu \alpha. \tag{5.14}$$

The divergence of j_o^μ still vanishes and, since the extra term in (5.14) is identically conserved, so does the divergence of j_r^μ. So, in the rotated version, there is neither an influx of charge from infinity nor a net charge generation and thus the total charge vanishes.

This regularization, although it has the advantage of always conserving the current, has some obvious drawbacks, coming from the fact that it is not compatible with the expression of the vacuum charge in terms of the spectral asymmetry of the hamiltonian (5.3). To see that, consider the adiabatic buildup of a very narrow soliton, starting with the trivial vacuum. As has already been said, the final fermionic state with $\Delta\alpha = 2\pi$ is essentially the same as the original vacuum state, but with a positive energy state filled, and it is reasonable to describe it as a $Q = 1$ state. The charge of the rotated version, however, with the same spectrum, is supposed to vanish.

This discrepancy can be understood if we notice that, in the present regularization scheme, the field A_μ is regarded as a true gauge field, and thus a gauge-invariant definition of the energy operator is to be adopted, namely $i\partial_0 - A_0$, that, in general, yields a different spectrum than the previous definition $i\partial_0$. To make the point clear, consider a constant A_0 field, corresponding to $\alpha = 2A_0 x$ in the unrotated version. As we switch this field on, our previous timelike regularization gives

$$\partial_\mu \langle j_r^\mu \rangle = \langle \dot{\rho} \rangle = \frac{1}{\pi} \dot{a}_0 \tag{5.15}$$

(we used the homogeneity of the problem), while the present regularization gives vanishing charge generation. The explanation is that, in the previous case, the addition of A_0 shifted *all* the energy levels of the hamiltonian by A_0. Since, for high enough energy, the successive energy levels of the hamiltonian differ by $\frac{\pi}{L}$ (where L is the size of a box into which we enclose the system in order to make the spectrum of the hamiltonian discrete), and in the spectral asymmetry essentially only the high-energy part of the spectrum contributes, an easy explicit calculation of (5.3) gives $Q = \frac{L}{\pi} A_0$. (Indeed, a shift of the spectrum by $\frac{\pi}{L}$, "creates" one extra filled level in the Dirac sea, leaving everything else unchanged.) So the density of vacuum charge is $-\frac{1}{\pi} A_0$, in accordance with (5.15). On the other hand, with the present regularization and definition of the hamiltonian, the spectrum remains invariant, since a constant A_0 can be gauged away, and thus no charge is generated. Actually, it is easy to see that, in the presence of *any* gauge field configuration, the vacuum charge has to vanish, else the effective action $W(A)$ after integrating out the fermions would not be gauge invariant under the transformation $A_0 \to A_0 + \epsilon$ (remember that $\langle j^\mu \rangle = \frac{\delta W}{\delta A_\mu}$ and thus $\langle Q \rangle = \frac{dW}{d\epsilon}$).

Which procedure is best depends on interpretation. If the potential is supposed to be generated by a classical external source, that does give (or take) energy to the fermions that fall into it, then the previous timelike regularization is more appropriate. If, on the other hand, the gauge field configuration is supposed to be in the integrand of a path integral, with the gauge field to be eventually fully quantized, then the present gauge invariant prescription is the correct one.

Let us conclude this section by pointing out that the charge due to vacuum polarization from a classical electric potential $a_0(x)$ that vanishes at infinity, can be found by backtracking the steps that connect it with the soliton configuration (and use timelike point-splitting). The answer is

$$Q = \frac{1}{\pi} \int_{-\infty}^{+\infty} A_0(x) dx. \tag{5.16}$$

This has a simple physical interpretation: a positive potential repels virtual fermions and attracts virtual antifermions. Thus, a high-speed fermion that approaches the region of the potential slows down, as it climbs the potential, and so it stays around that region longer than it would in the free case. Similarly, an antifermion gains kinetic energy and speeds up near the potential, and so it stays around less. This creates a net surplus of fermions over antifermions in the region of the potential, inducing a vacuum charge.

VI. Level crossings and the bosonization method.

In their original paper about charge fractionization [3], Goldstone and Wilczek used bosonization [12] in 1+1 dimensions as an alternative way to derive the standard formula for the vacuum charge. We know, however, that this formula is not strictly correct, because it may apply to a state that is not the vacuum, and the correct vacuum charge may differ by an integer. This section deals with the question of how this phenomenon manifests itself in the bosonized version of the problem.

We summarize here the basic facts about the bosonization of a lagrangian of the form (2.1). We know that a theory of a fermion field ψ in 1+1 dimensions is equivalent to a theory of a boson field ϕ. The expression that gives ϕ in terms of ψ is nonlocal. Some fermion bilinears, however, transform into a simple local form in the bosonic version. Specifically

$$\overline{\psi} i \partial\!\!\!/ \psi \rightarrow \tfrac{1}{2}(\partial_\mu \phi)^2 \tag{6.1a}$$

$$\overline{\psi}\gamma^\mu \psi \rightarrow \frac{1}{\sqrt{\pi}}\epsilon^{\mu\nu}\partial_\nu \phi \tag{6.1b}$$

$$\overline{\psi} e^{i\alpha\gamma^5}\psi \rightarrow \mathrm{M}\cos(2\sqrt{\pi}\phi - \alpha) \tag{6.1c}$$

where M is an arbitrary mass scale. So, the bosonized version of (2.1) can be written:

$$L = \tfrac{1}{2}(\partial_\mu \phi)^2 + \mathrm{Mm}\cos(2\sqrt{\pi}\phi - \alpha). \tag{6.2}$$

(The sign convention followed here is the opposite than before.) The argument of ref. 3 now is that, the vacuum expectation value of ϕ is the value that minimizes the potential in (6.2), i.e.,

$$2\sqrt{\pi}\phi = \alpha \tag{6.3}$$

and substituting this value into the formula (6.1b) that gives the bosonized version of the current and integrating over space we find the standard formula for the vacuum charge.

The point is, however, that one should minimize the *whole* energy, to find the classical vacuum configuration, not just the potential term. This means that one should also take into account the kinetic term in (6.2). For a soliton profile α that varies smoothly enough compared with $\frac{1}{\sqrt{\mathrm{Mm}}}$, this term is very small and can be neglected, so, indeed, for a smooth soliton, the standard formula holds true. For a sharp soliton, though, this term becomes appreciable and cannot be neglected. Actually, it is obvious that at $\pm$ infinity the potential term *has* to be minimized, since there α becomes a constant and the kinetic term vanishes. This means that equation (6.3) has to hold, modulo a multiple of 2π. During the transition from $-\infty$

to $+\infty$, however, the field ϕ may be "better off" not following faithfully the profile of the soliton, but rather "loosing" a few cycles, thus achieving a smoother form that gives a smaller kinetic term at the expense of a somewhat bigger potential term. So, substituting this value into the formula for the charge we will get an answer that differs from the standard one by an integer, exactly like in the fermionic case. In fact, for a very sharp soliton, it will be energetically favorable for ϕ to interpolate between the asymptotic values using the "shortest" possible way, giving a result compatible with the one stated in section IV.

It is interesting to note that different "vacua" of the bosonic case, corresponding to same (modulo 2π) asymptotic values of ϕ but different winding numbers in between, are separated by an infinite potential barrier. In the fermionic case, these "vacua" correspond to states with some positive energy levels filled or some negative ones empty. So, a quantum selection rule, namely conservation of charge, that forbids transition between these states in the fermionic case, is ensured by an infinite energy barrier in the bosonic case.

To illustrate the above remarks, we solve in the bosonic case the step potential example of fig. 2a. The equation of motion for ϕ is

$$\partial^\mu \partial_\mu \phi + 2\sqrt{\pi}\mathrm{Mm}\sin(2\sqrt{\pi}\phi - \alpha) = 0. \tag{6.4}$$

In the regions $x < 0$ and $x > 0$, (6.4) becomes essentially free ($\alpha = 0$), after a suitable shift of ϕ. So the equation of motion for a static case and $\alpha = 0$ becomes

$$-\partial_x^2 \phi + 2\sqrt{\pi}\mathrm{Mm}\sin(2\sqrt{\pi}\phi) = 0. \tag{6.5}$$

This, upon the redefinition $2\sqrt{\pi}\phi = \theta$, is the equation of motion of a pendulum of unit length in a gravitational field $4\pi\mathrm{Mm}$, with θ the angle measured from the top (fig. 5a). Solutions of (6.5) that asymptotically go to their relaxation value, correspond to motions of the pendulum starting from the top and, after infinite time, "falling" down until some point on the circle. These motions are just parts of the "tail" of the famous soliton solution of eq. (6.5), having the expression

$$\theta = 4\tan^{-1} e^{2\sqrt{\pi \mathrm{Mm}}(x-x_o)}. \tag{6.6}$$

Continuity of ϕ at $x = 0$ and the symmetry of the motion with respect to $x = 0$ tells us that we have to take the portion of the motion from $\theta = 0$ to $\theta = \frac{\Delta\alpha}{2}$, on the left, and match it with the same shape, reversed, on the right (fig. 5b).

Now we see that, since $\Delta\alpha \sim \Delta\alpha - 2\pi$, there are two points on the circle that the pendulum could go. The one that is energetically favorable is the "highest" one. So, as we increase $\Delta\alpha$ from zero, at $\Delta\alpha = \pi$ both points become equivalent and the vacuum becomes degenerate. As $\Delta\alpha$ increases beyond π, the point corresponding to $\Delta\alpha - 2\pi$ becomes preferable and the corresponding vacuum charge decreases one unit, reproducing exactly the results of section II. Let us also notice that the behaviour of the solutions (6.6) for large $|x|$ is exponential with a fall $e^{-2\sqrt{\pi \mathrm{Mm}}|x|}$. This is compatible with the exact formula for the vacuum charge density calculated in section II if

$$\mathrm{M} = \frac{\mathrm{m}}{\pi}. \tag{6.7}$$

This choice makes the scales of the fermionic and the bosonic versions identical. Then, the expression for the charge density obtained from the above solution is

$$j^0 = \frac{\frac{4m}{\pi}\tan\frac{\Delta\alpha}{8} e^{-2m|x|}}{1 + \tan^2\frac{\Delta\alpha}{8} e^{-4m|x|}}. \tag{6.8}$$

This agrees quite well, but not exactly, with the exact formula (2.13). The reason for this is that (6.8) is based on the classical solution of the equation of motion, ignoring quantum corrections. We see here the peculiar "mixing of levels" between the fermionic and the bosonized version: The vacuum charge, which is an essentially one-loop effect in the fermionic version, is reproduced at the classical level in the bosonic version, and in the one case where the exact answer for the charge density is known in the fermionic version, a full field-theoretical calculation to all orders is required in the bosonic version in order to get the exact result.

Finally, for very small $\Delta\alpha$ the bosonized version gives essentially the correct answer, since then quantum corrections of $\langle\phi\rangle$ become negligible. The formula for the charge density obtained then is

$$j^0 = \frac{\Delta\alpha}{2\pi} m e^{-2m|x|} \tag{6.9}$$

which has the same form as (2.13) differing only in the overall coefficient.

VII. Generalizations and conclusions.

In all the previous discussions, we assumed, for simplicity, that the magnitude of the chiral field was constant (and equal to m). Here we examine the general case, where the chiral field is

$$\tilde{\phi} = \phi_1 + i\gamma^5\phi_2 \equiv m(x)e^{i\alpha\gamma^5}, \text{ with } m(x) > 0. \tag{7.1}$$

This version does not really present any more difficulty than the one with a constant m. Specifically, the simple proof of section II can be repeated, where now we start with a static configuration of widely separated infinitesimal steps of both $m(x)$ and $\alpha(x)$ (fig.6), paying attention so that no two steps of either kind are too close compared to the correlation length $\frac{1}{m(x)}$ in that region (the steps of $m(x)$ do not need to be equal). The same localization arguments tell us that the total charge is the sum of the charge that each step would induce alone. But, for a constant α, after a suitable global chiral rotation that eliminates α, the lagrangian is charge-conjugation invariant, and so the charge of the vacuum vanishes (for positive $m(x)$ there do not exist any zero modes, that would be the only source of vacuum charge). So, each infinitesimal mass step induces no charge and, from the known answer for the α steps, we recover the previous result. Then we can "collect together" the steps and smooth them out to make any configuration we want, for which the same formula holds.

All the derivations and results of the other sections go through, easily generalizable when needed. For example, in section IV, the criterion for smoothness becomes

$$|\partial_x \alpha(x)| < 2m(x), \quad \text{for all } x, \tag{7.2}$$

and the bound on δ in theorem 2 becomes

$$|\delta| < \int_0^L 2m(x)dx. \tag{7.3}$$

Similar changes undergo the form of the zero-mode for a $\Delta\alpha = \pi$ sharp soliton with arbitrary $m(x)$ as well as the bosonization formulae.

The next, and more important, generalization to be made is in the number of dimensions. We will briefly discuss the relevant phenomena in 2+1 and 3+1 dimensions.

In 2+1 dimensions there is no analogous U(n) (or SU(n)) field that, because of its topology, can induce vacuum charge, since $\pi_2(\mathrm{U(n)}) = 0$. On the other hand, an abelian gauge field *can* have a nontrivial topology, namely its monopole number, corresponding to the total flux running through a spatial two-surface in units of 2π. If we compactify our space into S^2, this number has to be an integer. Notice that the total flux is given by the circulation of the gauge field at the circle at spatial infinity

$$\Phi = \frac{1}{2\pi} \oint_{r=\infty} \vec{A} \cdot d\vec{x} \tag{7.4}$$

and thus depends only on the asymptotics of $\vec{A}$. Such static magnetic configurations in principle can induce vacuum charge. We will be concerned with such configurations extensively in chapters 2 and 3. It is, however, interesting and important to see how the arguments of this chapter apply to the present case and where they fail or need to be modified.

The same locality and additivity arguments of section II can be used presently (we assume the fermions to be massive) to show that

$$Q = k\Phi, \tag{7.5}$$

since we can start with many distant infinitesimal "flux tubes" $\delta\Phi$, and bring them together to a configuration with the same Φ, without touching $\vec{A}$ at infinity. The novel feature of this case is that the calculation of k *à la* section II has in store a surprise!

The analogous thing to a step function soliton here is an infinitely thin flux tube, i.e., a Dirac string of flux Φ piercing our space. Such a string with integer flux is unobservable, since the action of a fermion going around it is an irrelevant multiple of 2π (alternatively, it can be gauged away with a single-valued gauge transformation). So, a level-crossing picture is suspected and we expect, on these grounds, k to be integer. However, it is easy to check that here, during the process of switching on Φ from 0 to 1, there is *no* level crossing. What happens instead is that a fermion energy level at $E = m$, that is well-behaved for any finite-size flux tube, becomes singular at the limit of vanishing tube size and thus decouples from the physical spectrum of the theory. The exact mechanism and value of Φ at which this happens depend nontrivially on the boundary condition imposed on the fermion field at the position of the string [13], that has to be included as an extra input in the theory, and will be elaborated in chapter 3. The important point is that a *single* level of the hamiltonian disappears. This $E = m$ level would contribute $-\frac{1}{2}$ to the spectral asymmetry expression (5.3) of the vacuum charge, if $m > 0$, and $+\frac{1}{2}$ if $m < 0$, and has to account for the difference between the result of formula (7.5) ($Q = k$) and the true result ($Q = 0$). Putting everything together we get

$$Q = -\tfrac{1}{2} sign(m)\Phi. \tag{7.6}$$

So the derivation of section II is in this respect somewhat incomplete, since we should have checked there, too, for modes that misbehave for an infinitely thin soliton.

However, such modes do not arise in the 1+1 dimensional case and the arguments given there are, indeed, valid.

Practically nothing else from the previous sections generalizes to the 2+1 dimensional case. The reason is that in odd-dimensional spacetimes there are no chiral anomalies (no γ^5) and, although there is a connection between the *two*-dimensional anomaly and the vacuum charge [14], it is of a quite different nature.

In 3+1 dimensions, now, the familiar chiral anomaly is back and so the general similarity with the 1+1 dimensional case is restored. A model analogous to (2.1) that has nontrivial topology and can induce charge is a chiral SU(n) field coupled to Dirac fermions in the fundamental representation of SU(n):

$$L = \overline{\psi} i \partial\!\!\!/ \psi + m \overline{\psi} e^{i\phi_a \lambda^a \gamma^5} \psi \tag{7.7}$$

where ϕ_a are real scalar fields parametrizing SU(n) and λ^a are the relevant Gell-Mann matrices. Since $\pi_3(\mathrm{SU(n)}) = Z$, this field can have topologically nontrivial configurations (solitons) with integer winding number if its behaviour at spatial infinity is trivial.

The locality and conservation of charge arguments of section II again tell us that the vacuum charge can only depend linearly on the winding number of the configuration. Moreover, a soliton with winding number 1 has trivial behaviour at infinity ($\phi_a = 0$), and so can be shrunk to an arbitrarily small size, until it becomes "invisible". It is possible to check again that in the process of shrinking the soliton no modes behave singularly while there is one level crossing. So we obtain

$$Q = \frac{1}{48\pi^2} \int \epsilon^{ijk} \mathrm{tr} \left[(U^\dagger \partial_i U)(U^\dagger \partial_j U)(U^\dagger \partial_k U) \right] d^3x \tag{7.8}$$

where $U = e^{i\phi_a \lambda^a}$ and we used the expression for the winding number of the configuration.

The discussions of sections III-V remain qualitatively valid, although the actual labor of the derivations increases. Also, in section V, a chiral rotation of the form

$$\psi_L \rightarrow U^\dagger \psi_L, \;\; \psi_R \rightarrow \psi_R \tag{7.9}$$

is more appropriate than the analog of (2.3), in order to avoid possible complications in defining the square root of a nontrivial SU(n) field.

An interesting remark applies to section VI: In 3+1 dimensions there is, of course, no bosonization. There is, however, a "poor man's" version of it, namely the method of an effective lagrangian. Specifically, we substitute the fermion field by an SU(n) (bosonic) field V, that parametrizes the degrees of freedom of the (spontaneously broken) chiral SU(n) invariance of the model and reproduces the low-energy phenomena of the theory [15]. The coupling of this new field V with the field U in this chiral lagrangian now becomes

$$S = \int d^4x \left\{ \tfrac{1}{2} \mathrm{tr} \left[(\partial_\mu V^\dagger)(\partial^\mu V) \right] + m \mathrm{tr}(U^\dagger V + V^\dagger U) \right\} + S_{WZ}(V) + \cdots \tag{7.10}$$

where the dots stand for possible higher order in V terms and $S_{WZ}(V)$ is the Wess-Zumino action [16], required to correctly reproduce the symmetries and the anomalies of the original fermionic action. This term is first-order in the time derivative and, as has been demonstrated, its properties under rotations turn a V-soliton into a fermion [17].

In the presence, now, of an external (smooth) U-soliton, the potential term in (7.10) is minimized for $V = U$, and so the (classical) ground state, due to the WZ term, has a fermion number equal to the winding number of U. For a sharp soliton, of course, we also have to consider the kinetic term. Again, the two fields U and V have to match at infinity, but this leaves the possibility that their winding numbers differ by an integer, thus reproducing the phenomenon of level crossings, and the fermion number (the winding number of V) will differ from the winding number of U by an integer.

A special remark applies for SU(2): Strictly speaking, there is no WZ term in this case. Due, though, to the nontriviality of $\pi_4(\mathrm{SU}(2)) = \mathrm{Z}_2$, the configuration space breaks into two disjoint parts and there is the possibility of adding an extra (discrete) term in the lagrangian, being 0 for trivial configurations and π for nontrivial ones (since two nontrivial ones give again a trivial one, the value of this extra term has to be a half-multiple of 2π). In fact, this term *has* to be included, in order to account for the nonperturbative SU(2) anomaly [18]. This term again turns V-solitons into fermions and the previous discussion applies.

In conclusion, we saw that many of the properties of induced vacuum charge can be understood in simple physical terms. The peculiarity of this phenomenon (fractional fermion number), and its connection with topology, make it one of the most interesting and intriguing in field theory.

REFERENCES

[0] A. P. Polychronakos, Phys. Rev. **D35** (1987) 1417.

[1] For a review and a comprehensive list of references see A. J. Niemi and G. W. Semenoff, Phys. Rept. **135** (1986) 99.

[2] R. Jackiw and C. Rebbi, Phys. Rev. **D13** (1976) 2298.

[3] J. Goldstone and F. Wilczek, Phys. Rev. Lett. **47** (1981) 986.

[4] M. B. Paranjape and G. W. Semenoff, Phys. Lett. **132B** (1983) 369.

[5] W. A. Bardeen, S. Elitzur, Y. Frishman and E. Rabinovici, Nucl. Phys. **B218** (1983) 445.

[6] Y. Frishman, D. Gepner and S. Yankielowicz, Phys. Lett. **130B** (1983) 66.

[7] R. Blankenbecler and D. Boyanovsky, Phys. Rev. **D31** (1985) 2089. and *ibid.* **31** (1985) 3234; Zhong-Qi Ma, H. T. Nieh and R. K. Su, Phys. Rev. **D32** (1985) 3268.

[8] B. Grossman, Phys. Rev. Lett. **51** (1983) 959.

[9] R. MacKenzie and F. Wilczek, Phys. Rev. **D30** (1984) 2194; A. J. Niemi, Phys. Rev. Lett. **54** (1985) 631.

[10] A. Chodos, R. L. Jaffe, K. Johnson, C. B. Thorn and V. F. Weisskopf, Phys. Rev. **D9** (1974) 1371.

[11] J. Goldstone and R. L. Jaffe, Phys. Rev. Lett. **51** (1983) 1518.

[12] S. Coleman, Phys. Rev. **D11** (1975) 2088; S. Mandelstam, Phys. Rev. **D11** (1975) 3026.

[13] A. P. Polychronakos, Nucl. Phys. **B283** (1987) 268.

[14] A. J. Niemi and G. W. Semenoff, Phys. Rev. Lett. **51** (1983) 2077.

[15] E. Witten, Nucl. Phys. **B223** (1983) 422.

[16] J. Wess and B. Zumino, Phys. Lett. **37B** (1971) 95.

[17] E. Witten, Nucl. Phys. **B223** (1983) 433.

[18] E. Witten, Phys. Lett. **117B** (1982) 324.

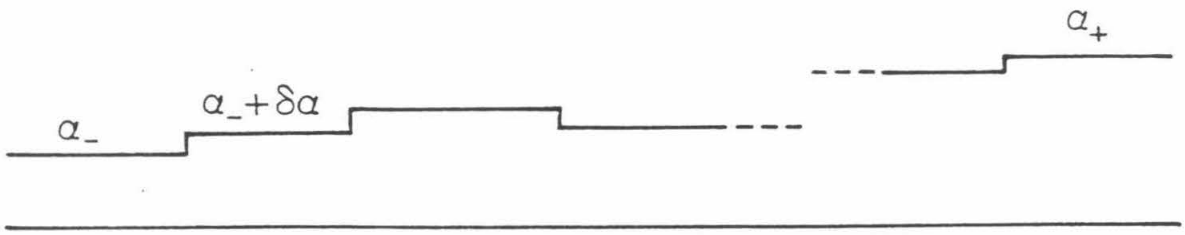

Fig. 1a

Fig. 1a: A collection of widely separated infinitesimal solitons.

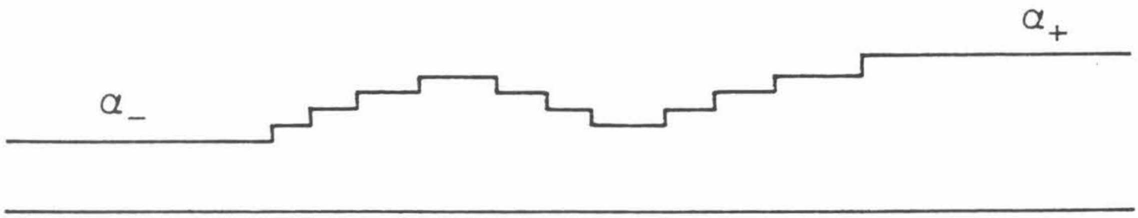

Fig. 1b

Fig. 1b: The infinitesimal solitons are brought together to form a sawtooth like profile.

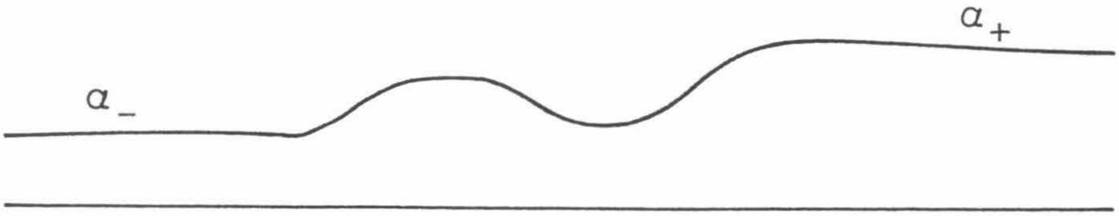

Fig. 1c

Fig. 1c: Smoothing out the profile of fig. 1b gives a general soliton profile.

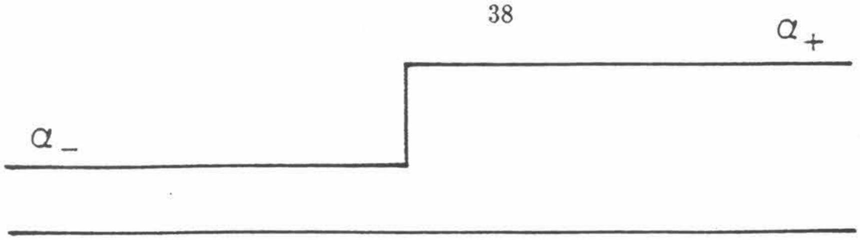

Fig. 2a

Fig. 2a: A step–function soliton.

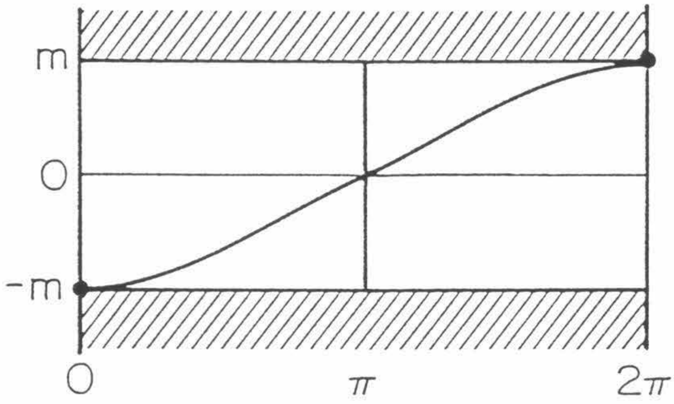

Fig. 2b

Fig. 2b: Spectral flow of the hamiltonian during the switching on of the step function soliton. At $\Delta\alpha = \pi$, a filled negative energy level becomes positive, eventually to join the continuum at $\Delta\alpha = 2\pi$. So the final state is not the vacuum but a $Q = 1$ excited state.

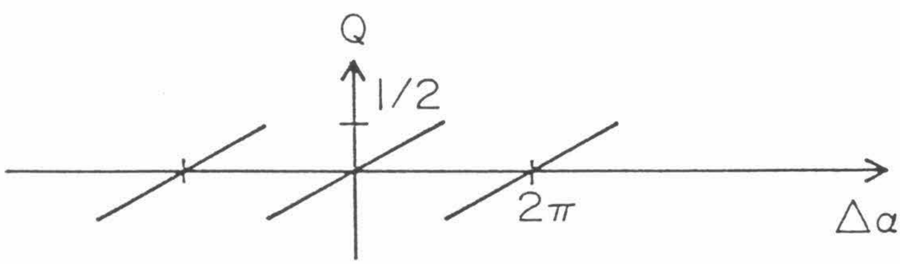

Fig. 2c

Fig. 2c: The vacuum charge of the step function soliton as a function of $\Delta\alpha$.

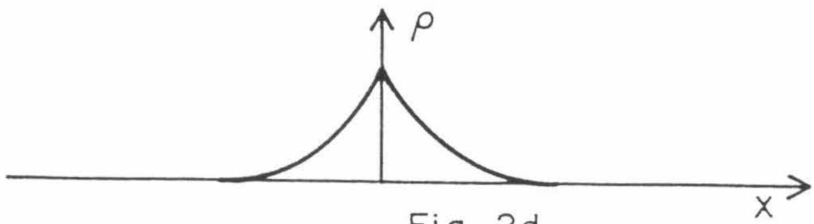

Fig. 2d

Fig. 2d: The vacuum charge density of the step function soliton with $\Delta\alpha = \pm\pi$.

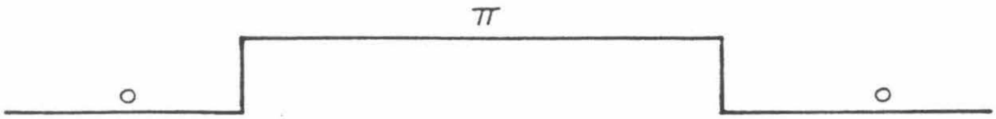

Fig. 3a

Fig. 3a: Two widely separated step function solitons with $\Delta\alpha = \pm\pi$.

Fig. 3b

Fig. 3b: The three almost-degenerate-with-the-vacuum states are obtained by filling or emptying the almost degenerate energy levels at $E = \pm\varepsilon$.

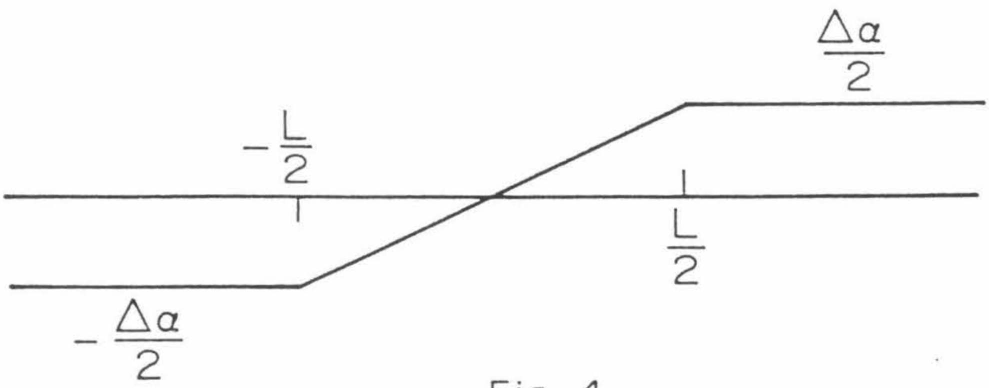

Fig. 4

Fig. 4: A finite size, piecewise−linear soliton.

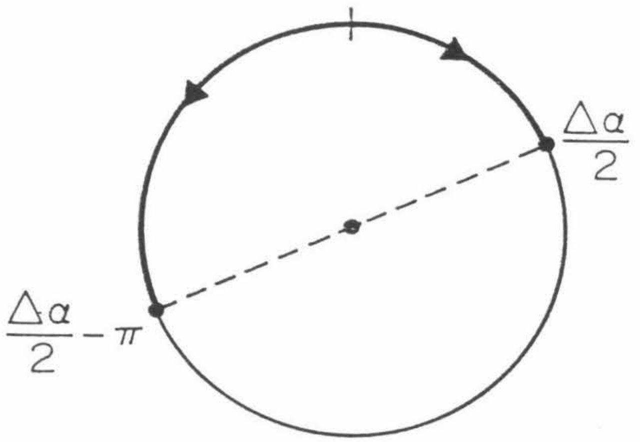

Fig. 5a

Fig. 5a: The vacuum configuration of the bosonized version of the step function soliton corresponds to the motion of a pendulum from the top, down to an angle $\frac{\Delta\alpha}{2}$, or alternatively $\frac{\Delta\alpha}{2} - \pi$. The true vacuum is the one corresponding to the "shortest" path. So, for $\Delta\alpha = \pi$, the two vacua are degenerate and we have a level crossing.

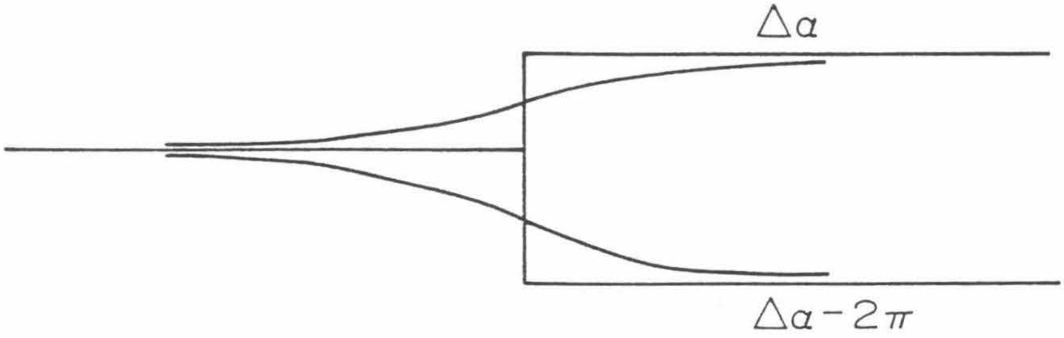

Fig. 5b

Fig. 5b: The θ−field configuration of the two possible vacua of fig. 5a.

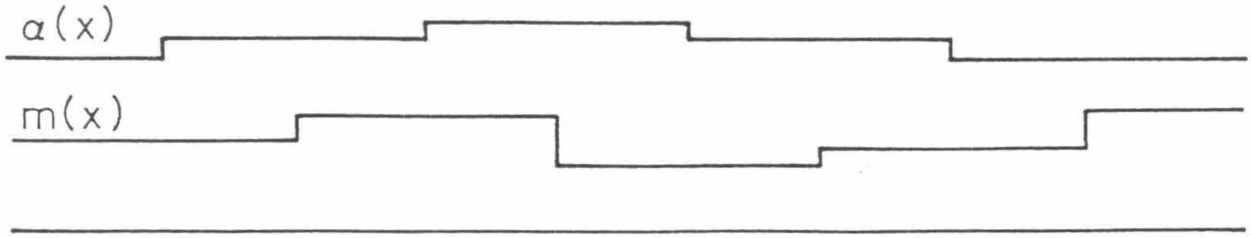

Fig. 6

Fig. 6: A collection of widely separated infinitesimal steps of $\alpha(x)$ and $m(x)$.

CHAPTER 2

Induced Vacuum Quantum Numbers in 2+1 Dimensions

I. Introduction.

As was briefly previewed in the last section of the first chapter, the problem of finding the induced vacuum charge in 2+1 dimensions has some qualitatively different features than its analog in 1+1 and 3+1 dimensions. Specifically, the nature of the induced charge is not anomalous *per se*, but is intimately connected with the anomalous chiral charge in *two* (euclidean) dimensions. Also, the limit of infinitesimally thin gauge field configurations is singular, and leads to *non*integer jumps of the charge in terms of the inducing magnetic flux. All these features make it worth a closer look.

One might object that a 2+1 dimensional situation is not very interesting. However, this is far from so. First of all, we should remember that, a field theory at a finite temperature can be formulated as a theory with a periodic euclidean time dimension, the period corresponding to $\beta = (\text{temperature})^{-1}$. For very large temperatures, the time period becomes very small and the time dimension collapses to nothing. Thus, three dimensional theories can be thought of as the high temperature limit of four dimensional field theories. But apart from this understandably indirect usefulness of these theories, there is a much more physical and practical one: 2+1 dimensional situations are realized whenever particles are more or less tightly bound to a planar surface. Such is the case, in particular, with the electrons at the interface of two different types of semiconductors. But exactly in that case one of the most interesting lately discovered phenomena takes place, namely the quantum Hall effect. As will become apparent in this and the last chapter, this phenomenon has a lot to share with the induced vacuum currents in 2+1 dimensional gauge theories. A better understanding of the quantum Hall effect is, thus, plenty of motivation for the study of these theories to be exposed in this chapter.

The peculiar and interesting properties of 2+1 dimensional gauge theories have been noticed quite some time ago [1], and are by now well established. As is the case with all odd-dimensional gauge theories, there is the possibility to add a Chern-Simons term in the lagrangian that renders the gauge field excitations massive [1]. The topology and properties of this term will be exhaustively studied in chapters 4 and 5, and thus we will not expound on it here. Let us only remark that this term will arise by radiative corrections when the gauge fields are coupled to fermions [2] and in turn contributes a parity-violating term to the vacuum expectation value of the fermionic current. In the abelian case:

$$\langle J^\mu \rangle = \pm \frac{1}{8\pi} \epsilon^{\mu\nu\rho} F_{\nu\rho} \,, \tag{1.1}$$

the sign depending on the sign of the mass of the fermions. In particular, the vacuum charge implied by (1.1) is

$$Q = \pm \frac{1}{4\pi} \int B d^2 x = \pm \frac{\Phi}{2} \,, \tag{1.2}$$

Φ being the total flux of the magnetic field. Although (1.1) is not the exact expression for the total vacuum current, (1.2) is exact and can be derived using the 1+1 dimensional anomaly [3]. On the other hand the vacuum charge can be expressed in terms of the (regulated) spectral asymmetry of the fermionic hamiltonian [4], and is therefore a function of the spectrum of the theory. Specifically, for time-independent background fields:

$$Q = -\tfrac{1}{2} \lim_{\epsilon \to 0+} \sum_{E_n} sign(E_n) e^{-\epsilon |E_n|} \,. \tag{1.3}$$

Moreover, if there is symmetry between positive and negative energy levels of the spectrum, only the zero modes contribute and we have degeneracy of the vacuum and integral or half-integral vacuum charge, as in the original example of charge fractionization [5]. In the case of massless 2+1 QED, the fermionic hamiltonian has indeed such a symmetry and the number of bound zero modes is given by the integer part of the flux [6]. Thus, an apparent puzzle arises here: consistent with the earlier remark, the fermionic charge should be (plus or minus) half the number of zero modes, and thus integer or half-integer. So for Φ noninteger one has trouble accounting for the fractional part of eq. (1.2). This puzzle persists in the massive case, as explained in section IV.

The standard explanation is that the fractional contribution comes from the continuum zero modes, since in this case there is no gap between the bound zero modes and the continuum. One of the purposes of this chapter is to deal with this puzzle and show how exactly the fractional part of the charge arises. More importantly, we are going to derive the expressions for the induced vacuum charge and angular momentum using a relatively straightforward method. (The canonical angular momentum has recently been calculated using trace-identities techniques [7]. Our results disagree by a factor of two.) The qualitative behavior of the results is quite interesting and will be discussed in the last section.

II. General setup and calculation of charge.

Let us consider Dirac fermions in 2+1 dimensions coupled to abelian gauge fields, with the lagrangian density

$$L = \overline{\psi}(i\partial\!\!\!/ + A\!\!\!/ - m)\psi - \frac{1}{4g^2}F^{\mu\nu}F_{\mu\nu}\,. \tag{2.1}$$

In 2+1 dimensions g^2 has dimensions of mass. We will use the explicit representation of the γ-matrices

$$\gamma^0 = \sigma_3\,, \quad \gamma^1 = i\sigma_1\,, \quad \gamma^2 = i\sigma_2\,. \tag{2.2}$$

Presently, all the calculations are going to be performed in the $m \to 0$ limit, where it is possible to derive explicit expressions for the vacuum density of charge and angular momentum. The generalization for nonzero mass is done in section IV.

We will be concerned with external static and rotationally symmetric gauge field configurations in the $A_0 = 0$ gauge that become asymptotically pure gauge for $\rho > \rho_0$ (ρ is the polar radius and ϕ is the polar angle), representing a magnetic flux "tube" of radius ρ_0 centered around $\rho = 0$. In an appropriate gauge the gauge field can be written:

$$\vec{A} = A(\rho)\hat{e}_\phi\,, \quad A(\rho) \equiv \frac{da}{d\rho} \equiv a'$$

$$B = \frac{1}{\rho}(\rho a')'\,, \quad a \to \Phi \ln \rho\,, \quad \text{for } \rho \to \infty. \tag{2.3}$$

The total magnetic flux running through our two-dimensional space is just the winding number of the $\vec{A}$ field at infinity and is not quantized as long as the topology of the two-dimensional space is taken to be that of a plane (if we compactified our space into a two-sphere, the flux would have to be quantized to an integer, with a Dirac string "bringing in" the flux). In this case it is given by

$$\Phi = \frac{1}{2\pi}\int B d^2x = (\rho a')\big|_{\rho=\infty}\,. \tag{2.4}$$

We will ignore here the electromagnetic field due to induced vacuum fermionic currents. Its effect on local densities (with the exception of angular momentum) is of order $g^2\rho_0$, which we shall assume $\ll 1$. In section IV the effect of this induced field on vacuum numbers is discussed.

The Dirac hamiltonian is

$$H = \begin{bmatrix} m & -\overline{\nabla} - \overline{\nabla}a \\ \\ \nabla - \nabla a & -m \end{bmatrix}, \quad \begin{array}{l} \nabla \equiv \partial_1 - i\partial_2 = e^{-i\phi}(\partial_\rho - \frac{i}{\rho}\partial_\phi) \\ \\ \overline{\nabla} \equiv \partial_1 + i\partial_2 = e^{i\phi}(\partial_\rho + \frac{i}{\rho}\partial_\phi). \end{array} \tag{2.5}$$

As m goes to zero, several energy levels of the fermion may become zero, thus introducing a well-known ambiguity in defining vacuum quantities due to the degeneracy of the resulting fermionic ground state. This ambiguity is, of course, resolved by considering the states that reach zero from above as empty, and the ones from below as filled, thus picking a unique vacuum. Since the sign of the energy of these states will depend on the sign of the mass, all the results will depend on $sign(m)$. We consider the case $m \to 0+$, understanding that taking the opposite limit flips the signs of all calculated expressions.

We now enclose the whole system in a box, in order to render the spectrum of the hamiltonian discrete and the level-counting procedure to follow well-defined. We choose a circular box of radius R, to preserve the rotational symmetry of our configuration, and assume $R \gg \rho_0$ (fig. 1). In order to impose the appropriate boundary conditions on the fermion field, we demand that the hamiltonian be Hermitian in the given volume V,

$$\int_V \chi^\dagger (H\psi)\, d^2x = \int_V (H\chi)^\dagger \psi\, d^2x\,, \tag{2.6}$$

which amounts to demanding

$$-i \int_{\partial V} \overline{\chi}\vec{\gamma}\psi \cdot d\vec{\sigma} = -i \int_0^{2\pi} \overline{\chi}\gamma^\rho \psi|_{\rho=R} \cdot R d\phi = 0. \tag{2.7}$$

The most general set of local boundary conditions that satisfies (2.7) is

$$u = \lambda v e^{-i\phi}\,, \quad \text{with } \psi \equiv \begin{pmatrix} u \\ v \end{pmatrix}, \tag{2.8}$$

where λ is any real constant.

We now notice that our hamiltonian possesses two discrete symmetries: charge conjugation C and "parity"* P. Thus, if $\binom{u}{v}$ is a solution of the energy eigenvalue

---

* The true parity transformation is

$$\psi \to \gamma^2 \psi\,, \ \phi \to -\phi\,, \ E \to -E\,, \ a \to -a\,, \ m \to -m.$$

Our P is more like true CT. The abuse of the language is due to the fact that a mass term breaks both symmetries.

equation with energy E, charge density j^0 and angular momentum density J, then C and P produce new solutions under the transformations:

$$C: \begin{pmatrix} u \\ v \end{pmatrix} \to \begin{pmatrix} v^\star \\ u^\star \end{pmatrix}, \ E \to -E\,,\ j^0 \to j^0\,,\ J \to -J\,,$$

$$a \to -a\,,\ m \to m\,,\ \lambda \to \frac{1}{\lambda}\,. \tag{2.9a}$$

$$P: \begin{pmatrix} u \\ v \end{pmatrix} \to \begin{pmatrix} u \\ -v \end{pmatrix}, \ E \to -E\,,\ j^0 \to j^0\,,\ J \to J\,,$$

$$a \to a\,,\ m \to -m\,,\ \lambda \to -\lambda\,. \tag{2.9b}$$

Thus we notice that there is no choice of λ that preserves both symmetries, even in the $m \to 0$ limit. The choices $\lambda = \pm 1$ preserve C and the choices $\lambda = 0$ or $\lambda = \infty$ ($u(R) = 0$ or $v(R) = 0$) preserve P. For the purposes of this calculation we choose to preserve P. For concreteness, we work with positive flux which we write in terms of its integer and fractional part

$$\Phi = N + \epsilon > 0\,,\ \ N = \text{integer}\,,\ 0 \le \epsilon < 1\,. \tag{2.10}$$

Then the choice of λ that preserves both P and the existence of localized zero modes is $\lambda = \infty$, or $v(R) = 0$, which we shall call "up" (U).

As noticed in reference [6], the system for $m = 0$ has zero modes. For our case and $m > 0$ these modes become threshold modes:

$$E = m\,,\ u = e^{-a} r^n e^{in\phi}\,,\ \ n = 0, 1, 2, \ldots \tag{2.11a}$$

For $m \to 0+$ they become zero modes from above. Notice that there is an infinite number of zero modes, but only N of them are "bound" and stick around the flux tube, while the rest are "unbound" and stick to the boundary. The charge-conjugated situation corresponds to $\Phi \to -\Phi < 0$, $\lambda = 0$ (we shall call it "down," D) and the threshold modes for $m > 0$ become

$$E = -m\,, \;\; u = 0\,, \;\; v = e^{a} r^{n} e^{-in\phi}\,, \;\; n = 0, 1, 2, \ldots. \qquad (2.11b)$$

For $m \to 0+$ they become zero modes from below (filled).

We can now proceed in the calculation of the charge. Each filled zero mode contributes to the charge inside a circle of radius ρ an amount

$$Q_n(\rho) = \int_0^{\rho} \psi_n^{\dagger} \psi_n 2\pi \rho d\rho. \qquad (2.12)$$

Assuming $\rho \gg \rho_0$, we see that the bound states, that fall off like $\rho^{n-N-\epsilon}$ for $\rho > \rho_0$, contribute mostly for $\rho \gtrsim \rho_0$, while the unbound states contribute mostly for $\rho \lesssim \rho_0$, where they can be approximated by their asymptotic form. So

$$Q_n(\rho) = \begin{cases} 1, & 0 \le n < N \\ x^{n+1-\Phi}, & n \ge N\,, \end{cases} \qquad (2.13)$$

where $x \equiv (\rho/R)^2$. The charge induced in our space due to the presence of the flux tube is the charge of the Dirac sea in the presence of the tube minus the charge of the Dirac sea in its absence:

$$Q(\rho) = \sum_{E_n<0} Q_n^U(\rho, \Phi) - \sum_{E_n<0} Q_n^U(\rho, 0), \qquad (2.14)$$

where the superscript U reminds us of the boundary conditions used. (For all mode-summation expressions a high-energy cutoff, as in (1.3), is implied. This is equivalent to an euclidean timelike point-splitting regularization for the corresponding operators, which is gauge-invariant in the $A_0 = 0$ gauge. In the massless case the cutoff will not play any role, because exact cancellation of the contributions of nonzero energy states will occur). Notice that we do not count the zero mode contribution, because these modes reach zero from above in the $m \to 0+$ limit and remain empty. Since the charge operator is odd under charge-conjugation, we define Q_{VAC} by the charge-conjugation odd part of Q:

$$Q_{VAC} = \tfrac{1}{2}[Q(\rho) - Q^C(\rho)] \tag{2.15}$$

where

$$Q^C(\rho) = \sum_{E_n<0} Q_n^D(\rho,-\Phi) + \sum_{E_n=0} Q_n^D(\rho,-\Phi) - \sum_{E_n<0} Q_n^D(\rho,0) - \sum_{E_n=0} Q_n^D(\rho,0). \tag{2.16}$$

Here we include the zero mode contribution, since these modes are filled. Also notice that the "up" vacuum with $\Phi = 0$ differs from the "down" $\Phi = 0$ vacuum by having its zero modes empty rather than filled, signaling the breakdown of charge conjugation due to the presence of these zero modes.

Combining formulae (2.14), (2.15) and (2.16) we have

$$\begin{aligned} Q_{VAC}(\rho) = \tfrac{1}{2}\Bigg\{ & \sum_{E_n<0} Q_n^U(\rho,\Phi) - \sum_{E_n<0} Q_n^D(\rho,-\Phi) \\ & - \sum_{E_n<0} Q_n^U(\rho,0) + \sum_{E_n<0} Q_n^D(\rho,0) \\ & - \sum_{E_n=0} Q_n^D(\rho,-\Phi) + \sum_{E_n=0} Q_n^D(\rho,0)\Bigg\} . \end{aligned} \tag{2.17}$$

Using C and P transformations we get

$$\begin{aligned} \sum_{E_n<0} Q_n^U(\rho,\Phi) &= \sum_{E_n>0} Q_n^D(\rho,-\Phi) = \sum_{E_n<0} Q_n^D(\rho,-\Phi) \\ \sum_{E_n<0} Q_n^U(\rho,0) &= \sum_{E_n>0} Q_n^D(\rho,0) = \sum_{E_n<0} Q_n^D(\rho,0) , \end{aligned} \tag{2.18}$$

so (2.17) becomes

$$Q_{VAC}(\rho) = -\tfrac{1}{2}\left\{ \sum_{E_n=0} Q_n^D(\rho, -\Phi) - \sum_{E_n=0} Q_n^D(\rho, 0) \right\} . \tag{2.19}$$

Using the explicit expressions (2.13) for the contribution of the zero modes we find

$$\begin{aligned} Q_{VAC}(\rho) &= -\tfrac{1}{2}\left\{ N + \sum_{n=N}^{\infty} x^{n+1-N-\epsilon} - \sum_{n=0}^{\infty} x^{n+1} \right\} \\ &= -\tfrac{1}{2}(N + \frac{x^{1-\epsilon} - x}{1-x}) . \end{aligned} \tag{2.20}$$

Now we can take the limit $\rho \to R$ $(x \to 1)$ to calculate the charge of the whole space. The result is finite and equal to

$$Q_{VAC} = -\frac{N+\epsilon}{2} = -\frac{\Phi}{2} , \tag{2.21}$$

which is the standard result.

So the picture of the situation that emerges from this analysis is the following: in the massless case and a finite volume, the theory possesses an infinite number of zero modes, some of them bound to the flux tube and some sticking to the boundary. The bound states contribute the (half-) integral part of the charge while the "tails" of the infinite boundary states contribute the fractional part. As Φ increases, more and more boundary states reach the $\rho = 0$ region and eventually "peel off" the boundary to become bound states. Thus the boundary plays the role of the infinity, where charge comes from in the perturbative calculations. The role of the boundary conditions in the behavior of vacuum quantities is analyzed more carefully in chapter 3.

III. Calculation of angular momentum.

Before we get involved in the calculation of the induced angular momentum, it is appropriate to discuss the different possible definitions used for it: the "canonical", the "kinematical" and the total angular momentum.

Due to the rotational symmetry of our field configuration, Noether's theorem predicts the existence of a conserved quantity that corresponds to the field-theoretical operator

$$\begin{aligned} J_C &= \int \psi^\dagger (-\frac{i}{2}\vec{\rho} \times \overleftrightarrow{\partial} + \frac{i}{4}\vec{\gamma} \times \vec{\gamma})\psi \, d^2x \\ &= \int \left\{ \tfrac{1}{2}[\psi^\dagger(-i\partial_\phi \psi) + h.c.] + \psi^\dagger \frac{\sigma_3}{2}\psi \right\} d^2x \, . \end{aligned} \tag{3.1}$$

Apparently, this operator is not gauge invariant. However, if we pick a nonsingular gauge with explicit rotational invariance, any residual gauge freedom does not affect J_C, so it can be defined in a gauge-independent fashion. (Actually, a little more can be shown: The vacuum expectation value of J_C is independent of *any* choice of gauge, because, due to the rotational symmetry of the vacuum expectation value of the charge, its variation under a gauge transformation integrates to zero). On the other hand, the proper kinematical definition for the angular momentum of the fermions leads to the gauge-invariant expression

$$\begin{aligned} J_K =& J_C - \int A_\phi \psi^\dagger \psi d^2x \\ =& J_C - \int \rho a' \psi^\dagger \psi \, d^2x \, . \end{aligned} \tag{3.2}$$

An amount of disagreement has arisen in the past on whether the physical definition of the angular momentum should be the canonical or the kinematical one [8,9]. The discrepancy between the two definitions can be understood and reconciled if we consider the total angular momentum carried by fermions and gauge fields [9,10] :

$$J = J_K + \frac{1}{g^2}\int \vec{\rho}\times(\vec{E}\times\hat{e}_z B)d^2x \tag{3.3}$$

The magnetic field is the field of the flux tube, while the electric field is due to the charged fermions. The effect of this induced electric field on the induced vacuum currents is of order $g^2\rho_0$ and can be neglected. However, the effect on the angular momentum is of order zero in $g^2\rho_0$. Indeed, for a rotationally symmetric tube, using (2.3) and the source equation (Gauss law) for $\vec{E}$

$$\mathrm{div}\vec{E} = g^2\psi^\dagger\psi \tag{3.4}$$

it is easy to see that

$$\begin{aligned} J &= J_K + \int \rho a'\psi^\dagger\psi\, d^2x - \Phi\int \psi^\dagger\psi\, d^2x \\ &= J_C - \Phi Q\,. \end{aligned} \tag{3.5}$$

So the total angular momentum is the canonical one subtracted by the product of the total flux times the total charge. If we vary Φ, J is not conserved while J_C is.

The difference between J_C and J can be traced to the contribution of a "return" flux at infinity equal to $-\Phi$. J does not account for the angular momentum carried by that flux, while J_C does. A nice way to look at the situation is the following: Imagine that we turn on the flux tube gradually, starting with zero flux. Since the gauge potential far away from the tube cannot change instantaneously, there is a circular "wavefront" of flux equal to $-\Phi$, propagating with the velocity of light. Due to the interaction of its magnetic field with the electric field produced by the charge inside the wavefront, it carries away to infinity angular momentum equal to ΦQ, that is included in J_C but not J. So the situation should now be fairly clear: J_K is just the angular momentum of the fermions. J is the total angular momentum of fermions plus local gauge fields, while J_C is a constant of motion but includes contributions at infinity and thus is not a local measurable.*

* More support for the fact that J rather than J_C is the physically relevant quantity comes from our investigation of the infinitely thin Dirac string, exposed in the next chapter.

Since J_K is the angular momentum of the fermions alone, we can in principle identify its vacuum expectation value as the sum of the angular momenta of the fermions in the Dirac sea subtracted by the same sum over the trivial Dirac sea. The same is not true, however, for the expectation values of J and J_C, since they include contributions coming from the gauge fields. For example, let us try to identify the expectation value of J_C as the sum over the Dirac sea of the canonical angular momenta of each single-particle state minus the same sum over the trivial Dirac sea. Then

$$\begin{aligned}\langle J_C \rangle_{VAC} &= \sum_{E_n<0} J_{Cn}(\vec{A}) - \sum_{E_n<0} J_{Cn}(0) \\ &= \sum_{E_n<0} [J_{Kn}(\vec{A}) + \langle A_\phi \rangle_n] - \sum_{E_n<0} J_{Kn}(0) \\ &= \langle J_K \rangle_{VAC} + \sum_{E_n<0} \langle A_\phi \rangle_n \ .\end{aligned} \tag{3.6}$$

Thus we see that the naive subtracted Dirac sea sum for J_C contains an infinite non-subtracted sum over negative energy states of $\langle A_\phi \rangle$. This is because the contribution from the gauge field includes the electric field produced by each negative-energy state and thus "sees" the infinite total charge of the Dirac sea, in contradiction with the normal ordering required for the charge operator. This means that $\langle J \rangle_{VAC}$ and $\langle J_C \rangle_{VAC}$ cannot be calculated as naive subtracted Dirac sea sums of the relevant one-particle angular momenta . (Later, a modification of the procedure so as to correctly calculate these quantities will be presented.)

Notice that the zero modes are eigenstates of J_C but not J_K. The eigenvalues of J_C for these states are

$$J_{Cn}^{U} = n + \tfrac{1}{2}\, , \ \ n = 0, 1, 2... \tag{3.7a}$$

and for the charge-conjugated situation

$$J_{Cn}^{D} = -n - \tfrac{1}{2}\, , \ \ n = 0, 1, 2, ... \tag{3.7b}$$

To find the contribution of each of these states to J_K inside a circle of radius ρ

we have to calculate

$$J_{K}{}_n^U(\rho) = \int\limits_0^\rho \psi_n^\dagger(n + \tfrac{1}{2} - \rho a')\psi_n \, d^2x, \tag{3.8}$$

and correspondingly with negative sign for the charge-conjugated case. The part proportional to J_C can be calculated in analogy with the contribution to the charge, distinguishing between $n < N$ and $n \geq N$, and is just $(n + \frac{1}{2})$ times the charge. The gauge-field-dependent part is

$$\begin{aligned} \langle \rho a' \rangle_n^U &= \frac{\int_0^\rho \rho a' e^{-2a} \rho^{2n} 2\pi \rho d\rho}{\int_0^R e^{-2a} \rho^{2n} 2\pi \rho d\rho} \\ &= (n+1) \frac{\int_0^\rho e^{-2a} \rho^{2n+1} d\rho}{\int_0^R e^{-2a} \rho^{2n+1} d\rho} - \frac{\rho^{2(n+1-\Phi)}}{2\int_0^R e^{-2a} \rho^2 n + 1 d\rho} . \end{aligned} \tag{3.9}$$

For $n < N$ the second term becomes negligible and the first term becomes $n + 1$ (remember $\rho \gg \rho_0$). For $n \geq N$ we can approximate the integrals with their $\rho > \rho_0$ parts, and the result is just $\Phi x^{n+1-\Phi}$. So overall

$$J_{K}{}_n^U(\rho) = \begin{cases} -\frac{1}{2}, & 0 \leq n < N \, , \\ (n + \frac{1}{2} - \Phi) x^{n+1-\Phi}, & n \geq N \, , \end{cases} \tag{3.10}$$

and $J_{K}{}_n^D(\rho) = -J_{K}{}_n^U(\rho)$ for the charge-conjugated case. It is interesting that all the bound states have J_K equal to $-\frac{1}{2}$ (or $+\frac{1}{2}$ for $\Phi < 0$), despite their different ϕ-dependences.

Now we can proceed with the calculation of $J_{K_{VAC}}$. Since the J_K operator is even under charge conjugation, we define

$$J_{K_{VAC}}(\rho) = \tfrac{1}{2}[J_K(\rho) + J_K^C(\rho)] \tag{3.11}$$

where

$$J_K(\rho) = \sum_{E_n < 0} J_{K}{}_n^U(\rho, \Phi) - \sum_{E_N < 0} J_{K}{}_n^U(\rho, 0) \tag{3.12a}$$

$$J_K^C(\rho) = \sum_{E_n<0} J_{K_n}^{D}(\rho, -\Phi) + \sum_{E_n=0} J_{K_n}^{D}(\rho, -\Phi)$$
$$- \sum_{E_n<0} J_{K_n}^{D}(\rho, 0) - \sum_{E_n=0} J_{K_n}^{D}(\rho, 0) \,. \tag{3.12b}$$

(The remarks about filling the zero modes in the D case but not in the U case are relevant here.) Combining (3.8) and (a-b) we have

$$J_{K_{VAC}}(\rho) = \tfrac{1}{2}\Bigg\{ \sum_{E_n<0} J_{K_n}^{U}(\rho, \Phi) + \sum_{E_n<0} J_{K_n}^{D}(\rho, -\Phi)$$
$$- \sum_{E_n<0} J_{K_n}^{U}(\rho, 0) - \sum_{E_n<0} J_{K_n}^{D}(\rho, 0) \tag{3.13}$$
$$+ \sum_{E_n=0} J_{K_n}^{D}(\rho, -\Phi) - \sum_{E_n=0} J_{K_n}^{D}(\rho, 0) \Bigg\} \,.$$

Using C and P, the terms with $E_n < 0$ cancel and we are left with

$$J_{K_{VAC}}(\rho) = \tfrac{1}{2}\Bigg\{ \sum_{E_n=0} J_{K_n}^{D}(\rho, -\Phi) - \sum_{E_n=0} J_{K_n}^{D}(\rho, 0) \Bigg\} \,. \tag{3.14}$$

Substituting the explicit expressions (3.7) and summing we finally have

$$J_{K_{VAC}}(\rho) = \frac{1}{4}\left[N + \frac{(2\epsilon - 1)x^{1-\epsilon} + x}{1-x} + 2\frac{x^2 - x^{2-\epsilon}}{(1-x)^2} \right] \,. \tag{3.15}$$

Then we let $\rho \to R$ to evaluate the whole angular momentum. The result is

$$J_{K_{VAC}} = \frac{N + \epsilon^2}{4} \,. \tag{3.16}$$

So we obtain a linear dependence on the integer part of Φ and a quadratic one on its fractional part, due to the differing behavior of bound and unbound modes near the flux tube (fig. 2).

In order to calculate $J_{C_{VAC}}$ we cannot use the same procedure, because, as was pointed out, it is not valid. However if we take the charge-conjugation-even part of formula (3.6) and define: $q_n = \psi_n^\dagger \psi_n$, we get

$$\begin{aligned}
\langle J_C \rangle_{VAC} =& \langle J_K \rangle_{VAC} + \tfrac{1}{2} \left\{ \sum_{E_n<0} \langle A_\phi \rangle_n^U - \sum_{E_n \le 0} \langle A_\phi \rangle_n^D \right\} \\
=& \langle J_K \rangle_{VAC} + \int A_\phi \cdot \tfrac{1}{2} \left\{ \sum_{E_n<0} q_n^U(\Phi) - \sum_{E_n<0} q_n^U(0) - \sum_{E_n\le 0} q_n^D(-\Phi) + \sum_{E_n \le 0} q_n^U(0) \right\} d^2x \\
& + \int A_\phi \cdot \tfrac{1}{2} \left\{ \sum_{E_n<0} q_n^U(0) - \sum_{E_n \le 0} q_n^D(0) \right\} d^2x \\
=& \langle J_K \rangle_{VAC} + \int A_\phi \cdot \left\langle \psi^\dagger \psi \right\rangle_{VAC} - \tfrac{1}{2} \int A_\phi \sum_{E_n=0} q_n^D(0) ,
\end{aligned} \tag{3.17}$$

where we used the symmetries of the problem. The first two terms in the last line of (3.17) are the true vacuum expectation value of J_C. So we see that if the up and down vacua were the same, the last term in (3.17) would be absent and we could calculate $J_{C_{VAC}}$ by taking the charge-conjugation-even part of its subtracted Dirac sea sum. But due to the presence of the zero modes one has to add a correction to the LHS of (3.17), (i.e. to the Dirac sea sum), in order to get the true $J_{C_{VAC}}$. The calculations are similar with the previous ones and we get:

$$J_{C_{VAC}}(\rho) = -\tfrac{1}{2} \left[\frac{x^{2-\Phi} - x^2}{(1-x)^2} + \tfrac{1}{2} \frac{x^{1-\Phi} - x}{1-x} \right] + \tfrac{1}{2} \Phi \frac{x}{1-x}. \tag{3.18}$$

The term in the bracket comes from the naive Dirac sea sum and is not finite in the $\rho \to R$ limit. However, the last term makes the limit finite, and we obtain*

$$J_{C_{VAC}} = -\frac{\Phi^2}{4}. \tag{3.19}$$

Using (2.21), (3.5) and (3.19) the vacuum expectation value of J can now be calculated

---

* Note that this disagrees with the result of ref. [7] by a factor of two. Also, the sign of our results is the correct one for the chosen conventions for the sign of the mass and the gauge coupling terms in the lagrangian.

to be

$$J_{VAC} = \frac{\Phi^2}{4}. \qquad (3.20)$$

The same answer is obtained if we repeat the procedure used for the calculation of J_C, that is, if we take the charge-conjugation-even part of the naive subtracted Dirac sea sum, where now the zero mode contribution to J for $\rho \gg \rho_0$ is

$$J_n^U(\rho) = (n + \tfrac{1}{2} - \Phi)x^{n+1-\Phi} \qquad (3.21)$$

and opposite for the charge-conjugated case . (Note that now the difference between up and down trivial vacua is inconsequential, since the contribution of their zero modes to the electric field in the region near the tube is vanishingly small). Reassuringly, a direct computation of the electromagnetic part of the angular momentum, using the electric field produced by the (known) vacuum charge density, correctly yields the difference between J_{VAC} and J_{KVAC}.

Since J_C and J do not explicitly depend on the details of the flux tube, their vacuum expectation values depend in a simple analytical way on the total flux.

Finally, it is very easy to calculate the total induced spin [11] $S = \frac{1}{2}\psi^\dagger\sigma_3\psi$ (which is part of J_K, J_C and J) with our method. S is even under both C and P and the zero modes contribute $\pm\frac{1}{2}$. A calculation similar with the previous ones gives:

$$S_{VAC} = -\frac{|\Phi|}{4}. \qquad (3.22)$$

The induced vacuum spin is proportional to the absolute value of the vacuum charge.*

---

* This result has also been derived in ref. [11], but without the absolute value.

IV. Generalization to nonzero mass and discussion of the results.

So far we dealt exclusively with the massless case. In the massive case, there is no parity symmetry in the way defined in (2.9b). There is, however, a symmetry of the spectrum, which we shall call "modified parity" $\tilde{C}$, that maps positive energy states of the hamiltonian into negative energy ones:

$$\tilde{C}: \begin{pmatrix} u \\ v \end{pmatrix} \to \begin{pmatrix} (E-m)u \\ (-E-m)v \end{pmatrix}, \quad E \to -E\,, \quad a \to a\,, \quad m \to m\,, \quad \lambda \to \left(\frac{E-m}{-E-m}\right)\lambda\,. \tag{4.1}$$

For $m = 0$, $\tilde{C}$ is the same as C, and for $\lambda = 0$ or $\lambda = \infty$ it becomes a good symmetry.

The only states that are not mapped into a $\tilde{C}$-conjugate state are states with $E = m$, $v = 0$, or $E = -m$, $u = 0$, the so-called threshold modes (formulae 2.11a,b). (Modes with $E = 0$ would be self-conjugate under $\tilde{C}$, but they are absent in our case due to the presence of a mass gap). So, in the infinite-volume limit the puzzle mentioned in the introduction persists: only threshold modes can contribute to the spectral asymmetry (1.3) and so the vacuum charge should be integer or half-integer. A generalization of the treatment of section II resolves the puzzle. In a space with boundary, there is an infinite number of threshold modes, N of them being bound to the flux tube and the rest of them sticking to the boundary. Although these last ones disappear at the infinite-volume limit, the contribution of their "tails" to the vacuum charge remains finite, thus accounting for its fractional part. Notice, though, that local densities do not transform in any particularly simple way under $\tilde{C}$, since the relative strengths of the up and down components of ψ are changed. So, it is not possible any more to derive expressions analogous to (2.20), (3.15), (3.18). Yet it should be clear that formulae (2.21), (3.19) and (3.20) for the total values of Q, J_C and J remain valid, since a one-particle eigenstate of J_C is transformed under $\tilde{C}$ into a one-particle eigenstate with the same eigenvalue, and so repetition of the procedure of sections II and III, taking into account the high-energy cutoff, leads to finite corrections to formulae (2.20) and (3.18) that go to zero as ρ goes to R. Then from (3.5) the value of J_{VAC} follows. On the other hand, it is not possible to calculate the vacuum value of J_K since it depends explicitly on the local behavior of each mode around the flux tube.

The first thing to be noticed about the calculated expressions for the massless case is that they only depend on ρ/R. For the charge this means that, apart from the part $-N/2$ that remains near the tube, in the $R \to \infty$ limit the density of the fractional part of the charge goes to zero. So, in the infinite volume limit a local measurement of charge always gives an integer or half-integer result, while the infinitely diluted fractional part is unobservable. This behavior is not surprising: in the massless case the theory has no scale other than the size of the tube (since we ignore, so far, the effect of the fermions on the gauge fields, g^2 does not play any role). The tube can support charge equal to half the number of its bound zero modes, and the rest of the charge scales with the size of the whole space. In the massive case we expect that for $\rho_0 \ll \frac{1}{m}$ the fractional part lies within a few Compton wavelengths from the tube, while the (half-) integral part is still within $\rho_{\sim}\rho_0$. For $\rho_0 \gg \frac{1}{m}$, formula (1.1) becomes valid (fig. 3). Similar remarks hold for the angular momentum.

Another property of the vacuum charge at $m = 0$ is that it is highly "volatile." The presence of an "antitube" of flux $-\Phi$ at arbitrarily large distance from our tube would make not only the nonlocal fractional part but also the localized around our tube (half-) integral part to vanish. Indeed, in our method, the existence of an antitube "kills" all local zero modes and gives to the boundary modes the same large-ρ dependence as in the absence of any flux, thus making all local charges to vanish in the $R \to \infty$ limit. This may at first sight look paradoxical, since one could imagine simultaneously switching on adiabatically a tube and an antitube at a great distance from each other and, certainly, one does not expect the charge induced around one of them to be influenced by the presence of the other. We should remember, though, that the calculated charge is in a steady-state configuration. In general, the state reached after switching on the tubes in finite time is not the vacuum. Since at the massless case the correlation length of the fermions is infinite, charge will slowly start to flow from one tube to the other until all local charges vanish. If on the other hand we switch the fluxes on slowly enough so that the adiabatic approximation hold well, there will be enough time during the process for charge to flow between the two *tubes* and in the final (vacuum) state all charges will be zero. We would expect this behavior to be rectified in the $m \neq 0$ case, as long as the tubes are far apart compared with the Compton wavelength of the fermions.

Probably the most peculiar expression derived in this paper is (3.16). Yet we can account physically for the form of this expression in the following way: When Φ is integer, all the fermionic angular momentum is due to the bound zero modes and, as mentioned in section III, all bound zero modes have J_K equal to $\pm\frac{1}{2}$. When Φ increases further, more charge and angular momentum is induced. Since most of this extra charge lies outside of the tube, its contribution to the gauge-field-dependent part of J_K is simply Φ times the extra charge. When we add to this the extra J_C induced, from formula (3.19), we find exactly the fractional part of (3.16).

It should be clear that formulae (2.21) and (3.20) remain valid even in the case of a rotationally nonsymmetric configuration, as long as the gauge field becomes asymptotically pure gauge at infinity. Since Q and J are conserved, their vacuum value is due only to the contribution from infinity, when we adiabatically switch on the flux. J_C is not well defined in the form (3.1) for rotationally nonsymmetric configurations, since a gauge transformation can now alter it, and, although it can still be defined as $J + \Phi Q$, it is not an interesting quantity. On the other hand, it is clear from our procedure that (3.16) still holds for rotationally nonsymmetric configurations in the $m \to 0$ limit. In general, though, since J_K is not conserved, (3.16) does not hold for $m \neq 0$. In fact, it is easy to see that J_K, as well as any short-ranged quantity with nonlinear dependence on the flux, cannot be a topological invariant for $m \neq 0$. Indeed, if we think of two flux tubes separated by a distance large compared to the Compton wavelength of the fermions, it is clear that all the short-range vacuum numbers they induce are just the sum of the quantum numbers each one of them would induce alone, since they don't "see" each other at that distance, which is incompatible with a nonlinear dependence on Φ alone. The total angular momentum is, however, long-range and the above argument breaks down. In fact, we can see that, apart from the part $\frac{\Phi_1^2}{4} + \frac{\Phi_2^2}{4}$, which is the sum of the individual angular momenta ("spins") of each tube, we also have an "orbital" part due to the interaction of the magnetic field of each one of the tubes with the electric field of the charge around the other, equal to $-\Phi_1 Q_2 = -\Phi_2 Q_1 = \frac{\Phi_1\Phi_2}{2}$. Adding the "spin" and "orbital" parts we have $J = \frac{(\Phi_1+\Phi_2)^2}{4}$, as in (3.20). In the massless limit, due to long-range correlations, local quantities like J_K with nonlinear Φ-dependence can still be topological invariants.

In the $m \gg \frac{1}{\rho_0}$ limit we can still evaluate $J_{K_{VAC}}$ for a rotationally symmetric tube, by adding to $J_{C_{VAC}}$ the contribution from the $-\int A_\phi \left\langle \psi^\dagger \psi \right\rangle$ part. By using formula (1.1) it is easy to calculate

$$\int A_\phi \left\langle j^0 \right\rangle d^2x = -\frac{\Phi^2}{4} \text{ and thus } J_{K_{VAC}} = 0. \tag{4.2}$$

Thus, a tube with diameter large compared with the Compton wavelength of the fermions does not induce any kinematical angular momentum. Let us notice that, during the adiabatic switching on of the flux, an azimuthal electric field is necessarily induced, and it is possible to use the equation of motion for J_K and formula (1.1) to calculate the induced angular momentum due to the action of the electromagnetic field on the induced vacuum current. If we put

$$\vec{A}(t) = f(t) A_\phi \hat{e}_\phi\,, \;\; f(0) = 0\,, \;\; f(T) = 1$$

$$\vec{E}(t) = \dot{f}(t) A_\phi \hat{e}_\phi\,, \;\; B(t) = f(t) \frac{1}{\rho} (A_\phi \rho)'\,, \tag{4.3}$$

and use

$$\left\langle \dot{J}_K \right\rangle = \int \vec{\rho} \times \left\langle \vec{F} \right\rangle d^2x = \int \vec{\rho} \times \left[\left\langle j^0 \right\rangle \vec{E} + \left\langle \vec{j} \right\rangle \times \hat{e}_z B \right] d^2x \tag{4.4}$$

we can easily see that the induced J_K vanishes:

$$J_{K\,ind} = \int_0^T \left\langle \dot{J}_K \right\rangle dt = 0\,. \tag{4.5}$$

This, together with (4.2), tells us that there is no influx of J_K from infinity during the adiabatic switching on of the flux. In the large m limit we can explicitly calculate the angular momentum of the gauge fields, to lowest order in g^2. From (1.1), (2.3) and (3.4) we see that

$$\vec{E} = \frac{g^2}{4\pi}\vec{A} \times \hat{e}_z = -\frac{g^2}{4\pi}a'\hat{e}_\rho \tag{4.6}$$

and so

$$J_{EM} = \frac{1}{g^2}\int \vec{\rho} \times (\vec{E} \times \hat{e}_z B)\, d^2x = \frac{\Phi^2}{4} \tag{4.7}$$

in agreement with (3.20) and (4.2).

We come now to the question of the action of the fermions on the gauge fields. Due to the induced charge, a long-range electric field will be produced, that, according to eq. (1.1) will give rise to currents in the ϕ-direction. These currents, in turn, create a magnetic field with a seemingly divergent total flux. Before that happens, of course, this magnetic field will induce enough additional vacuum charge so as to completely screen the charge around the tube and make all long-range fields vanish. This screening charge will be distributed over a length scale of order g^{-2} (see eq. 3.4). So, the gauge field has lost its long-range nature and has acquired a finite correlation length of order g^{-2}, which means that it is effectively massive with a mass of order g^2 [1]. With the same argument we see that all external charges are screened by the vacuum, with the simultaneous creation of a magnetic flux around them [11], needed to induce the equal and opposite screening charge. What all that means for our results is that the fractional part of the induced charge is completely unobservable in the $m < g^2$ case, since it would lie outside of the screening length. The (half-) integral part is still observable, provided we can probe distances shorter than the screening length. The same is true for the whole charge in the $m \gg g^2$ case, since it lies well within the screening length. Finally, if the condition $g^2\rho_0 \ll 1$ is not satisfied, the magnetic field as well as all local charges are completely screened. Similar remarks hold for the angular momentum.

The above properties of the vacuum (screening of external charges and magnetic fluxes, finite range of electromagnetic interactions) essentially turn it into a superconductor. Electric currents are spontaneously generated whenever needed in order to screen magnetic fluxes. Notice, though, that, as formula (1.1) implies, they are also generated in the presence of electric fields, but with a direction *normal* to the field. This is exactly the situation encountered in the quantum Hall effect. There, again, a current normal and proportional to the electric field is generated (in the presence of a strong magnetic field as a catalyst). The quantum Hall conductance (the off diagonal component of the 2×2 conductance tensor) is extremely insensitive to the size of the fields and impurities of the system and is equal to an integer multiple of $1/2\pi$ in natural units ($\hbar = c = e = 1$). In our case, from formula (1.1) we have

$$j^i = \frac{1}{4\pi}\epsilon^{ij}E_j \tag{4.8}$$

and thus the vacuum Hall conductance is $1/4\pi$, that is *half* the quantum hall conductance. The explanation of this fact and the connection of the two phenomena will further be analyzed in the last chapter.

As a final remark, we point out that the induced fermionic vacuum energy in the large m case vanishes, since $\vec{E}$ and $\left\langle \vec{j} \right\rangle$ are normal and so no energy is produced during the adiabatic switching on of the flux. However the question of whether there is a nonzero fermionic vacuum (Casimir) energy for arbitrary m remains open.

REFERENCES

[0] A. P. Polychronakos, Nucl. Phys. **B278** (1986) 207.

[1] S. Deser, R. Jackiw and S. Templeton, Phys. Rev. Lett. **48** (1982) 975 and Ann. Phys. (N.Y.) **140** (1982) 372.

[2] A. N. Redlich, Phys. Rev. Lett. **52** (1984) and Phys. Rev. **D29** (1984) 2366.

[3] A. J. Niemi and G. W. Semenoff, Phys. Rev. Lett. **51** (1983) 2077.

[4] M. B. Paranjape and G. W. Semenoff, Phys. Lett. **132B** (1983) 369.

[5] R. Jackiw and C. Rebbi, Phys. Rev. **D13** (1976) 3398.

[6] R. Jackiw, Phys. Rev. **D29** (1984) 2375.

[7] M. B. Paranjape, Phys. Rev. Lett. **55** (1985) 2390.

[8] F. Wilczek, Phys. Rev. Lett. **48** (1982) 1144.

[9] R. Jackiw and A. N. Redlich, Phys. Rev. Lett. **50** (1983) 555.

[10] A. S. Goldhaber, Phys. Rev. Lett. **49** (1982) 905.

[11] D. Boyanovsky, R. Blankenbecler and R. Yahalom, SLAC-PUB-3764.

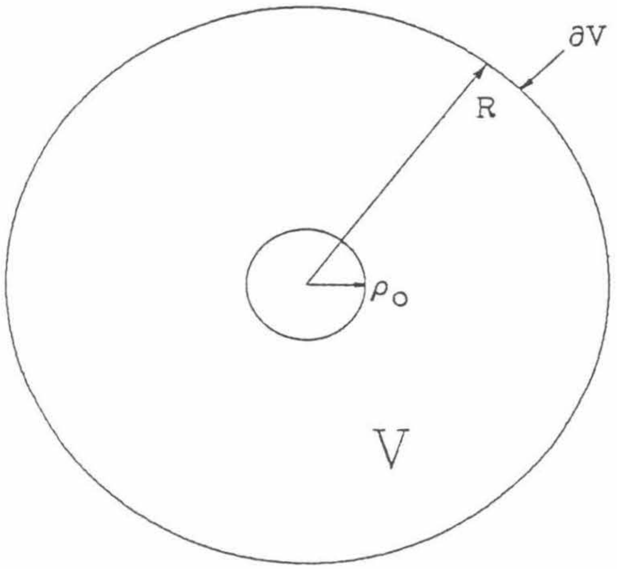

Figure 1.

Fig. 1: Our compactified space is taken to be circular with local boundary conditions imposed on the fermion fields at the boundary ∂V. The region of radius ρ_0 around the origin represents the cross section of the flux tube.

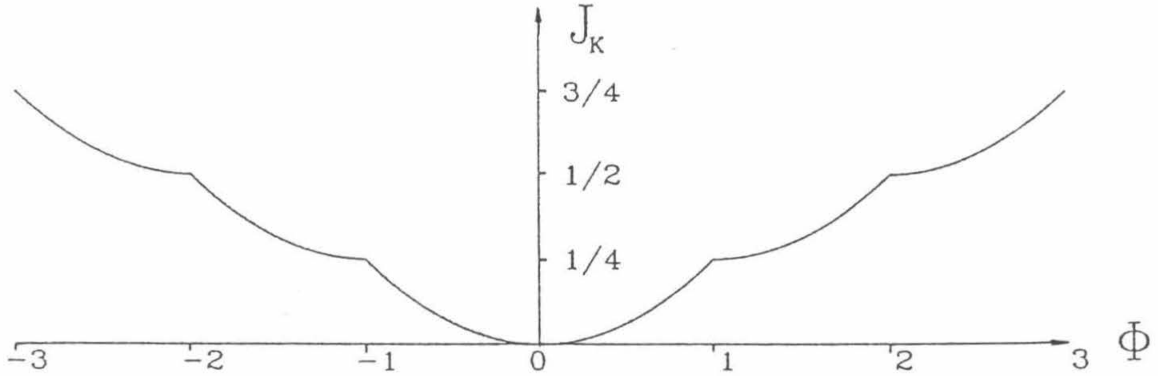

Figure 2.

Fig. 2: Plot of the fermionic angular momentum J_K versus the total flux Φ.

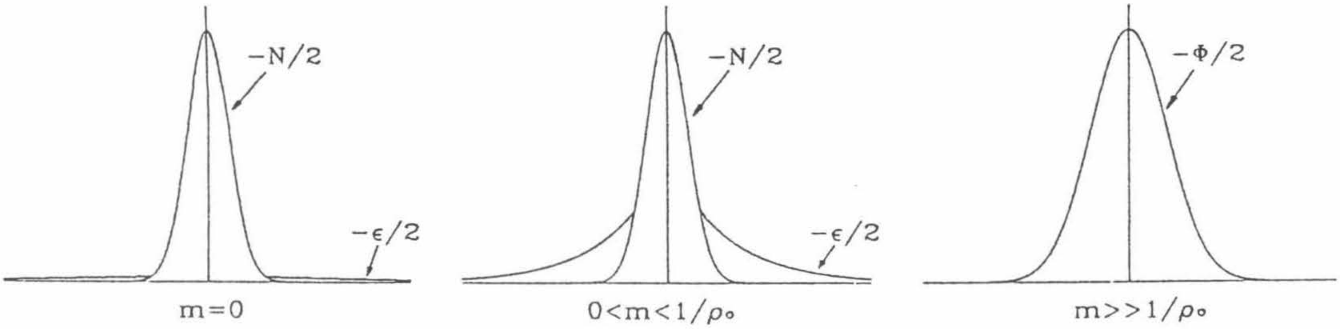

Figure 3.

Fig. 3: Qualitative plots of the distribution of charge around a flux tube for different values of m.

CHAPTER 3

Boundary Conditions, Vacuum Quantum Numbers and the Index Theorem

I. Introduction.

From the analysis presented in the previous two chapters, it should be obvious that topologically nontrivial field configurations [1] do indeed possess peculiar properties, connected with anomalies and irregular quantum numbers. In spite of that, no topological tools were used so far in calculating these quantum numbers, and the connection with topology may have remained somewhat unclear. In this chapter, we will apply such a tool and examine what modifications to its standard form are needed in order to faithfully reproduce the desired physical results.

The most powerful topological tool used by physicists has proven to be the Atiyah-Singer index theorem [2]. Its ability to relate analytical properties of the field theory and topological properties of the field configurations, apart from providing us intuition about the behavior of the field theory, is also invaluable in unmasking such nonperturbative phenomena as global anomalies [3] and parity anomalies [4]. Moreover, the index theorem can be used even to obtain analytical results, with most notable example the one of interest here, namely the evaluation of the vacuum fermion number induced by topologically nontrivial background fields [5]. For example, if the problem has some form of conjugation symmetry, the fermion number is ($\pm\frac{1}{2}$ times) the number of zero modes of the corresponding Dirac hamiltonian [6], and this can usually be provided by an index theorem.

Index theorems take their simplest form on compact boundaryless manifolds. Suitable generalizations have been derived, for manifolds with boundaries [7] as well as for open infinite manifolds [8]. These generalizations have proved very useful, especially in the context of fermion number fractionization on infinite spaces [6,8]. However, the derivation of index theorems on manifolds with boundaries requires the use of the so-called spectral boundary conditions for the fermion fields [7,9]. Although these conditions allow for a well-defined integer index and a corresponding index theorem, they are physically undesirable, because of their nonlocal, and thus acausal nature.

Here, we present the alternative of using *local* boundary conditions, which are more palatable from the physics point of view. The identification of an index, under such boundary conditions, becomes nontrivial. However, we show here that, under a suitable and quite natural regularization procedure, we are still able to define an index, which in this case turns out to be *fractional*, despite its usual interpretation as a difference of zero modes of an operator. This definition is physically motivated, and the results relate naturally to the results on a corresponding infinite space.

We will concentrate, throughout this chapter, on the special case of a two-dimensional abelian operator, which can be interpreted either as a 1+1 dimensional euclidean Dirac operator, or as a 2+1 dimensional Dirac hamiltonian in the presence of time-independent background fields. The two interpretations are completely equivalent [10]. The relevant physical issues are the axial anomaly, in the 1+1 case, and the induced vacuum charge, in 2+1 dimensions. We work, here, in the context of the 2+1 dimensional hamiltonian, since it relates more directly to a physical observable, i.e., the vacuum charge (in the 1+1 case it would correspond to the expectation value in euclidean spacetime of the chiral charge). Further, it is possible in this case to define other physical quantities, connected with the index, like the vacuum expectation value of the angular momentum [11,12].

The main physical distinction between the results using different boundary conditions is the behavior of vacuum quantities near the boundaries. Of course we do not expect the boundaries to influence local quantities far away from them, if the theory has finite correlation length. There is, however, the possibility of altering the results at their vicinity, with respect to the results in their absence, thus giving rise to spurious contributions to vacuum quantities. These contributions remain near the boundaries and follow them as they go to infinity, thus decoupling from local physics in the infinite-volume limit. The index theorem, though, counts the *total* amount of vacuum quantities, and thus may give misleading results. This is indeed the case with global boundary conditions. On the other hand, local boundary conditions are shown to produce *no* such extra contributions, thus giving reliable results and a smooth transition to the infinite-volume case. This is their main virtue, together with their locality and causality.

The organization of this chapter is as follows: In section II, we solve the problem of the induced charge inside a bounded region using spectral boundary conditions and we recover the standard Atiyah-Patodi-Singer index theorem. (A similar example has also been worked out in [13], in the context of the chiral anomaly. Our procedures are inequivalent.) It is shown that the boundary induces extra unwanted (i.e., localized at the boundary) contributions to the charge, compared to the infinite-space case. In section III, local boundary conditions are introduced, and we define and calculate our fractional index. The results are shown to agree with those on an open infinite space. In section IV, we deal with the problem of the vacuum numbers induced by an infinitely thin flux tube (a Dirac string with nonquantized flux) piercing our space, as a problem related to the boundary conditions of the fermion field at the position of the string. It is shown that, depending on the nature of conditions imposed, the charge (and other vacuum quantities) behaves discontinuously in the flux of the string, with *half-integer* jumps at values of the flux equal to a half-integer number of quanta, or vanishes identically for all values of the flux. In this context, we resolve a puzzle connected with the apparent nonperiodicity of the induced quantum numbers in the flux of the string. Finally, in section V, we solve the same problem on a sphere, where no boundary conditions are required, and demonstrate how the Atiyah-Singer theorem works in this case. It is shown that the Dirac string is indeed unobservable. The vacuum charge and angular momentum density are calculated in a special case and shown to have the symmetries of the problem, irrespective of the position of the string.

II. Global boundary conditions.

We start our analysis by examining the induced fermionic charge inside a flat compact two-dimensional space, with standard spectral boundary conditions imposed on the fermions at the boundary. The mass of the fermions is m, and can be either positive or negative. For convenience, we include the gauge coupling, which has dimensions of mass dimension $\frac{1}{2}$, in the definition of the gauge fields.

As in ref. [12], our space is assumed to be a flat circular disc D_2 of radius R. The Dirac hamiltonian for fermions interacting with external static abelian gauge fields is

$$H = \begin{bmatrix} m & -\nabla - \nabla a \\ \overline{\nabla} - \overline{\nabla} a & -m \end{bmatrix} \equiv \begin{bmatrix} m & D \\ D^\dagger & -m \end{bmatrix} , \quad \begin{aligned} \nabla &\equiv \partial_1 - i\partial_2 = e^{-i\phi}(\partial_\rho - \frac{i}{\rho}\partial_\phi) \\ \overline{\nabla} &\equiv \partial_1 + i\partial_2 = e^{i\phi}(\partial_\rho + \frac{i}{\rho}\partial_\phi) \end{aligned} \tag{2.1}$$

where ρ and ϕ are the polar radius and polar angle, respectively, and we adopted the Coulomb gauge for the gauge field

$$A_0 = 0 \ , \quad A_i = -\epsilon_{ij}\partial_j a. \tag{2.2}$$

D and $D^\dagger$ are formally adjoint. However, for their true adjointness to be established, and so the hermiticity of H to be secured, suitable boundary conditions have to be imposed on the fermion fields on the boundary of the disc S^1, such that

$$\langle \psi_1, H\psi_2 \rangle = \langle H\psi_1, \psi_2 \rangle \tag{2.3}$$

or equivalently

$$\langle u, Dv \rangle = \left\langle D^\dagger u, v \right\rangle \tag{2.4}$$

for *all* u, v, where we wrote $\psi = \begin{bmatrix} u \\ v \end{bmatrix}$. One way to secure this is to use "spectral" boundary conditions. To define them, we first define the kinematical angular momentum operator:

$$J_K = -i\partial_\phi - A_\phi + S = -i\partial_\phi - \partial_\rho a + \tfrac{1}{2}\sigma_3 \tag{2.5}$$

We assume that the gauge field becomes locally pure gauge near S^1, and a approaches the asymptotic value

$$a \to \Phi \ln \rho \ , \quad \rho \to R \tag{2.6}$$

where Φ is the total magnetic flux running through the disc. Then the restriction of J_K on S^1 takes the form

$$J_K|_R = -i\partial_\phi - \Phi + \tfrac{1}{2}\sigma_3. \tag{2.7}$$

Define now the operator P as the projection operator onto the positive eigenvalues of $J_K|_R$. Then the spectral boundary conditions are given by

$$(1+\sigma_3)P\psi|_R = 0 \tag{2.8a}$$

$$(1-\sigma_3)(1-P)\psi|_R = 0. \tag{2.8b}$$

In terms or u, v, (2.8) can be written

$$P_U\, u|_R = 0 \ , \quad (1-P_V)\, v|_R = 0 \tag{2.9}$$

where P_U and P_V are projection operators onto the positive eigenvalues of the operators

$$J_U = -i\partial_\phi - \Phi + \tfrac{1}{2} \ , \quad J_V = -i\partial_\phi - \Phi - \tfrac{1}{2} = J_U - 1 \tag{2.10}$$

respectively. With these boundary conditions we see that

$$\int u^* v|_R \, d\phi = 0, \tag{2.11}$$

since u and v restricted on S^1 belong to disjoint eigenspaces of the hermitean operator J_U, which is enough to secure (2.4).

The hamiltonian (2.1) has an eigenvalue-reversing discrete symmetry. Specifically, if $\begin{bmatrix} u \\ v \end{bmatrix}$ is an eigenvector of H with eigenvalue E, the transformation

$$\begin{bmatrix} u \\ v \end{bmatrix} \to \begin{bmatrix} (E-m)u \\ (-E-m)v \end{bmatrix} \tag{2.12}$$

gives an eigenvector of eigenvalue $-E$, for all $E \neq \pm m$. Since the boundary conditions (2.9) are invariant under (2.12), this remains a good symmetry of the spectrum.

The vacuum charge induced by the external gauge fields is given by the spectral asymmetry of H [14]

$$\langle Q \rangle = -\tfrac{1}{2}\eta_H(0) = -\tfrac{1}{2} \lim_{s \to 0+} \sum_n sign(E_n)|E_n|^{-s} \tag{2.13}$$

Due to the symmetry (2.12), The contributions to $\eta_H(0)$ from all $E_n \neq \pm m$ cancel, and we are left with

$$\langle Q \rangle = -\tfrac{1}{2}(n_+ - n_-) \tag{2.14}$$

where n_+ (n_-) is the number of states with eigenvalue $|m|$ ($-|m|$). In this (abelian) case, and with the boundary conditions (2.9), the modes with energy m ($-m$) acquire the form $[u, v=0]$ ($[u=0, v]$) respectively. These are the zero modes of the operators $D^\dagger$ and D, and so the induced charge can be expressed as $-\frac{1}{2}sign(m)$ times the index of the operator

$$H_0 \equiv H(m=0) = \begin{bmatrix} 0 & D \\ D^\dagger & 0 \end{bmatrix} \tag{2.15}$$

defined as

$$ind(H_0) \equiv \dim \ker D^\dagger - \dim \ker D \tag{2.16}$$

The general solution for the zero modes of $D^\dagger$ and D respectively is [15]

$$u_n = e^{-a}\rho^n e^{in\phi} \ , \quad D^\dagger u_n = 0 \ , \quad n = 0, 1, 2, \ldots \tag{2.17a}$$

$$v_n = e^a \rho^n e^{-in\phi} \ , \quad Dv_n = 0 \ , \quad n = 0, 1, 2, ... \tag{2.17b}$$

(only positive values of n are allowed because of the requirement of square integrability of the wavefunctions around $\rho = 0$), and on S^1 they become

$$u_n|_R = R^{n-\Phi} e^{in\phi} \tag{2.18a}$$

$$v_n|_R = R^{n+\Phi} e^{-in\phi} \tag{2.18b}$$

It is now easy to see that the requirement (2.9) allows only the u_n with

$$n - \Phi + \tfrac{1}{2} \leq 0 \ \Rightarrow \ 0 \leq n \leq \Phi - \tfrac{1}{2} \tag{2.19a}$$

and only the v_n with

$$-n - \Phi - \tfrac{1}{2} > 0 \ \Rightarrow \ 0 \leq n < -\Phi - \tfrac{1}{2} \tag{2.18b}$$

So we now see the logic behind the conditions (2.9). In order to have a well-defined index, a way to discard all but a finite number of the solutions (2.17) has to be found, and the conditions (2.9) do just that. This is analogous to the case of an infinite space, where the requirement of square integrability of (2.17) for $\rho \to \infty$ would allow only the modes with $n \leq [|\Phi|] - 1$, of the appropriate spin, depending on the sign of Φ ($[\cdots]$ denotes the integer part).

Now, the expression for the index, and correspondingly for the charge, becomes (fig. 1)

$$ind(H_0) = [\Phi + \tfrac{1}{2}]. \tag{2.20a}$$

$$\langle Q \rangle = -\tfrac{1}{2} sign(m) [\Phi + \tfrac{1}{2}]. \tag{2.20b}$$

In this case it is easy to calculate explicitly the spectral asymmetry of the boundary operator $J_K|_R$ (or J_V or J_U). The result is

$$\eta_J(0) = -2\left\langle\Phi + \tfrac{1}{2}\right\rangle + 1 \tag{2.21}$$

where $\langle\cdots\rangle$ denotes the fractional part. So the following relation holds:

$$ind(H_0) = \Phi + \tfrac{1}{2}\eta_J(0). \tag{2.22}$$

Given that

$$\frac{1}{2\pi}\int\limits_{D_2} F = \frac{1}{2\pi}\int\limits_{D_2} Bd^2x = \Phi \tag{2.23}$$

we see that (2.22) is nothing but the Atiyah-Patodi-Singer index theorem.

On the other hand, the expression for the anomaly [16], or equivalently for the vacuum charge [10] in the case of an open infinite space is known to be

$$\langle Q\rangle = -\tfrac{1}{2}sign(m)\Phi \tag{2.24}$$

Thus we see that we have a discrepancy for the charge between the open and the finite space. If the fermions are massive, and the region with nonvanishing magnetic field is small and isolated from the boundary (a "flux tube"), we expect the charge around the flux tube to be the same in both cases, and equal to the open space value. So, the only explanation for the discrepancy is that, due to the spectral boundary conditions, there is an extra contribution to the charge, localized near the boundary, equal to

$$\left\langle Q_{boundary}\right\rangle = -\frac{1}{4}sign(m)\eta_J(0) = \frac{1}{2}sign(m)\left(\left\langle\Phi + \tfrac{1}{2}\right\rangle - \tfrac{1}{2}\right). \tag{2.25}$$

Notice also that, as we increase Φ, although the charge near the flux tube increases in value, the total charge remains constant, except from abrupt jumps at Φ =integer+$\frac{1}{2}$. What happens is that charge localized near the boundaries "migrates" to the region near the flux tube, while extra charge appears when the boundary conditions allow the existence of an extra zero mode.

Overall, we see that the spectral boundary conditions, apart from their unphysical nature (nonlocality), create extra boundary contributions to physical quantities and thus are not a good way to get the infinite-volume (boundaryless) limit.

III. Local boundary conditions.

We now turn our attention to boundary conditions that can be imposed pointwise on the boundary, i.e. local, and guarantee the hermiticity of H (eq. (2.3)). Obviously, we cannot demand that the fermion field ψ vanish at the boundary. Since the hamiltonian is first-order in derivatives, such boundary conditions would make not only the zero modes, but also the *whole* Hilbert space disappear. Vanishing of a two-component Dirac spinor is equivalent to two complex conditions. A milder boundary condition is necessary, equivalent to *one* complex condition.

Such a condition can be found by noticing that (2.3), after integrating by parts, is equivalent to

$$\int \overline{\psi}_1 \gamma^\rho \psi_2 \big|_{\rho=R} R d\phi = 0 \tag{3.1}$$

where γ^ρ is the radial γ-matrix. For this to be satisfied with a *local* boundary condition we must impose

$$\overline{\psi}_1 \gamma^\rho \psi \big|_{\rho=R} = 0 \tag{3.2}$$

everywhere on the boundary. This can be written

$$iu_1^* e^{-i\phi} v_2 - iv_1^* e^{i\phi} u_2 = 0 \quad \text{on } S^1. \tag{3.3}$$

By rewriting this condition in the form

$$\left(\frac{u_1}{v_1} e^{i\phi} \right)^* = \frac{u_2}{v_2} e^{i\phi} \tag{3.4}$$

and choosing $\psi_1 = \psi_2$, we see that the quantity in (3.4) must be real for all ψ. Then, restoring $\psi_1 \neq \psi_2$, we conclude that it is also independent of ψ. Thus, the most general boundary condition that can be imposed globally on the Hilbert space and satisfies (3.3) is

$$u = \lambda(\phi) e^{-i\phi} v \tag{3.5}$$

with λ a real constant that can, in general, vary over the boundary. So we have a one-parameter family of possible conditions for each point of the boundary. If we also want to preserve the discrete symmetry (2.13) of the theory, the only choices are

$$u|_R = 0 \tag{3.6a}$$

$$\text{or } \mathrm{v}|_\mathrm{R} = 0 \tag{3.6b}$$

and for concreteness we shall choose the latter one.

We should first realize that (3.6b) leads to a well-defined eigenvalue problem for H. Indeed, from the eigenvalue equation

$$(E - m)u = Dv \tag{3.7a}$$

$$(E + m)v = D^\dagger u \tag{3.7b}$$

we get for v

$$D^\dagger D v = (E^2 - m^2)v \tag{3.8}$$

which is a Laplace-like equation and, under the boundary condition (3.6b), it has a well-defined discrete positive spectrum, guaranteeing $|E| > m$. Then, for $E \neq m$, u is completely determined from (3.7a). Notice that for each solution of (3.8) we obtain *two* different u's, corresponding to the positive and negative value of E. If, on the other hand, we put $v = 0$, then $E = m$ and from (3.7b)

$$E = m \ , \quad D^\dagger u = 0 \tag{3.9}$$

which has an infinite number of solutions. We can then define our Hilbert space as the set of all (normalizable) spinors that can be written as an infinite superposition of solutions of (3.8-3.7a) and (3.9).

From (3.7) and with the condition (3.6b), it follows that

$$\int |u|^2 d^2x = \frac{E+m}{E-m} \int |v|^2 d^2x \tag{3.10}$$

and so

$$\langle S \rangle_E \equiv \left\langle \tfrac{1}{2}\sigma_3 \right\rangle_E = \frac{m}{2E} \,. \tag{3.11}$$

So the expectation value of the spin is a function of the energy only.

The most notorious feature of the theory with the boundary conditions (3.6b) is that it acquires an infinite number of threshold modes ($E = m$), that become zero modes in the $m = 0$ case, namely the ones given in (2.17a). Thus the index of H_0 becomes meaningless, since it would have to be infinite.

There is, however, a procedure that leads to a well-defined, yet fractional, index for H_0. Define $P_0(A)$ to be the projection operator onto the (infinite-dimensional) null space of H_0 in the presence of the gauge potential A, and P_V the operator that multiplies by 1 inside a volume $V \subset D_2$ that nowhere touches the boundary, and by 0 outside V. Then define

$$ind(H_0) \equiv \lim_{\epsilon \to 0+} \{Tr[P_V P_0(A)] - Tr[P_V P_0(0)]\} \tag{3.12}$$

where ϵ is the maximum distance of the boundary of V from the boundary of D_2 and Tr denotes the functional trace in the Hilbert space of the Hamiltonian. The traces, before taking the limit, are *finite*, because the zero modes (2.17a) for large n are highly localized near the boundary and their contribution inside the volume V goes fast enough to zero.

This regularization procedure is very natural. Since (2.12) remains a good symmetry, the induced vacuum charge is still given by ($-\frac{1}{2}sign(m)$ times) the index of H_0. So eq.(3.12) can be interpreted as the charge induced inside the volume V, subtracted by the charge of the trivial ($A = 0$) hamiltonian. The peculiar feature of the present case is that, even for $A = 0$, the hamiltonian has a spectral asymmetry, caused by the existence of the modes (2.17a), that has to be subtracted from the asymmetry of $H(A)$ in order to find the induced charge.

The expression (3.12) can be calculated easily in the special case where V is a circular disc concentric with D_2, of radius $R-\epsilon$. Then

$$ind(H_0) = \lim_{\epsilon \to 0+} \left(\sum_{n=0}^{\infty} \int_0^{R-\epsilon} |u_n(A)|^2 d^2x - \sum_{n=0}^{\infty} \int_0^{R-\epsilon} |u_n(0)|^2 d^2x \right) \tag{3.13}$$

We remark that the modes (2.17a) are not orthogonal on the disc, unless a is independent of ϕ. However, since each term in the summations in the limit $\epsilon \to 0$ goes to 1, we can neglect any finite number of initial terms from both summations. But, for high enough n, the zero modes become highly localized near the boundary, and a can be approximated with its asymptotic expression (2.6), giving

$$u_n \to \rho^{n-\Phi} e^{in\phi} \ , \quad R-\epsilon < \rho < R. \tag{3.14}$$

These modes are now orthogonal, and normal to the subspace of modes we omitted. We can, thus, normalize them and use them for the evaluation of (3.13). The integrals in (3.13) can be calculated explicitly, giving

$$ind(H_0) = \lim_{\epsilon \to 0+} \left[\sum_{n=N}^{\infty} \left(1 - \frac{\epsilon}{R}\right)^{2(n-\Phi+1)} - \sum_{n=N}^{\infty} \left(1 - \frac{\epsilon}{R}\right)^{2(n+1)} \right] \tag{3.15}$$

where N is large enough so that the approximation (3.14) hold well. The sums now can be done explicitly, and taking the limit we find

$$ind(H_0) = \Phi \tag{3.16}$$

From this we imply that the vacuum charge is

$$\langle Q \rangle = -\tfrac{1}{2} sign(m)\, \Phi \tag{3.17}$$

which is the same as the result (2.24) for the open infinite space.

It becomes apparent thus that the boundary condition (3.6b) did not introduce any spurious boundary contributions to the charge. It should be obvious that boundary condition (3.6a) would lead to the exact same expression (3.16). The correct infinite-space limit is reproduced with boundary conditions (3.6b) for $\Phi > 0$, and with (3.6a) for $\Phi < 0$. These choices will secure the existence of the corresponding localized zero modes around the flux tube, otherwise all the charge will be concentrated near the boundary.

Notice that, as we increase Φ, the induced charge inside D_2 increases in value continuously. Since the fermionic current is conserved, this means that there is an "influx" of charge from the boundary. This influx of charge from the boundary is quite peculiar, since the boundary condition (3.2) naively implies that the radial component of the current on the boundary should vanishes. What we have here is a phenomenon of breakdown of hermiticity of H, during the switching on of topologically nontrivial background fields. In our case the boundary has assumed the role of infinity in an open infinite space.

This "pumping in" of charge from the boundary is reminiscent of the case of the chiral bag, [17], where again, due to the (local) chiral boundary conditions, an influx of charge appeared as we turned on the chiral parameter of the boundary conditions [18]. There, however, the influx was due to an explicit anomalous component of the current at the boundary, while here it is a result of the nontrivial two-dimensional geometry of the gauge potential (in odd dimensions there are no chiral anomalies).

We note that we can also calculate the vacuum expectation values of other operators, as the spectral asymmetry of H weighted with the expectation value of the operators in each one-particle energy state:

$$\langle \Pi \rangle = -\tfrac{1}{2} \lim_{\epsilon \to 0+} \lim_{s \to 0+} \sum_n sign(E_n) \, \langle P_V \Pi \rangle_n \, |E_n|^{-s} \Bigg|_{A=0}^{A} \tag{3.18}$$

where again the limiting procedure and vacuum subtraction of (3.12) was used. If Π possesses a symmetry such that the contributions from nonzero (or nonthreshold) modes vanish, $\langle \Pi \rangle$ can be explicitly calculated. For details see [12]. We state here the results:

$$\langle J \rangle = sign(m)\frac{\Phi^2}{4} \tag{3.19a}$$

$$\langle J_K \rangle = sign(m)\frac{[\Phi] + \langle \Phi \rangle^2}{4} \quad (m \to 0) \tag{3.19b}$$

$$(3.19c)\, \langle S \rangle = -sign(m)\frac{|\Phi|}{4} \quad (m \to 0) \tag{3.19c}$$

where J_K is the "kinematical" angular momentum (the one of the fermions alone) and J is the total angular momentum of fermions plus gauge fields. Notice that, from (3.11), S is itself an odd function of the energy, and so the *even* part of the energy spectrum is relevant in (3.18), rather than the odd [19], which is not a topological invariant. However, for $m \to 0$, the contributions from $E_n \neq 0$ vanish, and we can still obtain the expression (3.19c). Similarly for J_K, the contributions from $E_n \neq 0$ cancel only in the limit $m \to 0$, when (3.19b) holds. For $m \to \infty$, $\langle J_K \rangle$ can be shown to vanish.

Overall we see that the boundary conditions (3.6) are physically more satisfying than spectral boundary conditions, both because of their locality and the fact that they do not induce spurious boundary contributions to the anomaly of H_0 (i.e., to the vacuum charge). The index of H_0 can still be defined, suitably regularized, and the infinite-space limit is reached more naturally.

IV. The Dirac string.

The expression (2.24) for the induced charge is exact, and holds for static magnetic configurations of arbitrary shape and size. In particular, it should hold when all the magnetic flux is concentrated in a vanishingly small region of space, i.e., to the case of an infinitely thin Dirac string piercing our 2-d spatial surface. Thus, an apparent puzzle arizes here: As it is well established, a Dirac string with flux Φ quantized to an integer is quantum mechanically unobservable. More generally, the quantum mechanics of this string are independent of a shift of the flux by an integer, that is, they are periodic in Φ with period 1. There are several ways to see this. Firstly, the quantum mechanical phase picked up by a wave going around such a string (the one producing the Aharonov-Bohm effect) is

$$\oint \vec{A}\cdot d\vec{l} = \int B d^2x = 2\pi\Phi \tag{4.1}$$

and for integer shifts of Φ the phase changes by an unobservable multiple of 2π. Equivalently, the gauge potential of the string

$$\vec{A} = \frac{\Phi}{2\pi\rho}\hat{e}_\phi \tag{4.2}$$

(in a rotationally symmetric gauge singular at $\rho = 0$) can be transformed into the potential of a string with flux $\Phi + N$ with the gauge transformation

$$U = e^{iN\phi} \tag{4.3}$$

which is well-defined (single-valued) for integer N. Finally, the spectrum of the Dirac hamiltonian in the presence of the string is identical to the spectrum of a hamiltonian with flux $\Phi + N$. Indeed, if ψ_n are the energy eigenstates of $H(\Phi)$

$$H(\Phi)\psi_n = E_n\psi_n \tag{4.4}$$

then the eigenstates of $H(\Phi + N)$ are just

$$\chi_n = e^{iN\phi}\psi_n$$

$$H(\Phi + N)\chi_n = E_n\chi_n \tag{4.5}$$

So the spectral asymmetry of $H(\Phi + N)$ should be the same as that of $H(\Phi)$ and, in particular, the vacuum charge of a string with integer Φ should vanish, in contradiction with (2.24). What happens?

The standard mechanism that leads to dependence of the induced vacuum charge on the local details of the background fields, while otherwise it should be topologically invariant, is level crossings [20]. If, in the process of shrinking the flux tube to small size, a level of the hamiltonian crosses zero, then the state reached is not the vacuum any more, and the true vacuum charge will differ from the calculated value by an integer. Here, however, it easy to see that we have *no* energy levels crossing zero, due to the presence of a mass gap in the spectrum. Moreover, it should be immediately obvious that level crossings could not save the situation: At $\Phi = 1$, the induced vacuum charge is equal to $\pm\frac{1}{2}$, and level crossings account only for *integer* jumps. A different (and subtler) mechanism should be sought for, in order to account for the periodicity of vacuum quantities in Φ.

In order to deal with the problem properly, we should examine the question of the boundary conditions satisfied by the fermion field near the position of the string (which we shall take it to be at $\vec{x} = 0$). Indeed, the gauge potential (4.2) becomes infinite at $\vec{x} = 0$, and thus we have to remove this point from our space. In fact, we shall remove a *finite* small disc of radius ρ_0 around the string, in order to avoid various singularities that appear when we excise a single point from the space. This will create a small boundary around the string, on which appropriate boundary conditions have to be imposed on the fermion field. Again, either spectral or local conditions can be chosen, leading to different results even in the limit $\rho_0 \to 0$. Of course, a zero-size flux tube is something unphysical, and so it shouldn't be discomforting that the results depend on the boundary conditions. The only physically "allowed" string is the Dirac string, with $\Phi =$ integer, which is really an unobservable artifact of the gauge, and in that case both boundary conditions should give compatible results. The results for noninteger Φ, however, differ drastically. As we will show, with "global" boundary conditions the charge behaves more or less as (2.24) predicts, but with abrupt jumps at half-integer values of Φ. On the contrary, with "local" boundary conditions the vacuum charge vanishes for all values of Φ.

For concreteness, we shall again consider our space to be a circular disc of radius R, (in fact an annulus with internal radius ρ_0), with local boundary conditions of the type (3.6b) imposed at $\rho = R$.

The solutions (2.17a), which are the ones allowed by the boundary conditions at $\rho = R$, take here the form

$$u_n = \rho^{n-\Phi} e^{in\phi} \ , \quad v_n = 0 \tag{4.6}$$

The first thing that should be observed is that the requirement $n \geq 0$ for the threshold modes (4.6) does *not* have to hold any more. Indeed, that condition had to do with the square integrability of the wavefunction near the origin, and since here we have excised a finite neighborhood around the origin, the solutions (4.6) are normalizable for *all* n. Any further restriction on n will have to come from the boundary conditions at $\rho = \rho_0$.

In analogy with (3.1), the condition for hermiticity of H at $\rho = \rho_0$ is

$$\int \overline{\psi}_1 \gamma^\rho \psi_2 \big|_{\rho=\rho_0} \rho_0 d\phi = 0. \tag{4.7}$$

Again, one way to satisfy it is to impose spectral boundary conditions, of the type (2.8, 2.9). Specifically, we impose

$$(1 - P_U)u|_{\rho_0} = 0 \tag{4.8a}$$

$$P_V v|_{\rho_0} = 0 \tag{4.8b}$$

that ensure (4.7) as well as the symmetry of the spectrum (2.12). The reason why this time we chose to project onto the *positive* spectrum of J_U and the *negative* spectrum of J_V is similar to the one that led to the condition (2.9): We need to bound n from *below*, and (4.8) does just that. Indeed, for the threshold modes (4.6), (4.8b) is automatically satisfied, since $v = 0$, but (4.8a) implies

$$n - \Phi + \tfrac{1}{2} \geq 0 \tag{4.9}$$

So, as we increase Φ, one mode disappears every time that Φ becomes a half-integer. Since n is unbounded from above, we see that the spectrum of threshold modes of

H (or of zero modes of H_0) is indeed periodic in Φ. The calculation of the induced charge is now identical to the one in section III, with the exception that we need to *exclude* $[\Phi + \frac{1}{2}]$ modes from the first sum in (3.13), if $\Phi > 0$, or to *include* $[|\Phi| + \frac{1}{2}]$ extra modes, if $\Phi < 0$. Thus the result is (fig.2a)

$$\langle Q \rangle = -\tfrac{1}{2} sign(m) \left(\langle \Phi + \tfrac{1}{2} \rangle - \tfrac{1}{2} \right). \tag{4.10}$$

As expected, $\langle Q \rangle$ is periodic in Φ and for $\Phi =$ integer it vanishes. Also, it is an odd function of Φ, consistent with the fact that Q and Φ are odd under charge conjugation.

The behavior of $\langle Q \rangle$ in (4.10) is typical of the boundary conditions used: As Φ increases, the induced charge increases because of the influx of charge at $\rho = R$. At $\Phi =$integer $+\frac{1}{2}$, $\langle Q \rangle$ jumps by *half* a unit, due to the appearance (or disappearance) of an extra mode. This is quite different from the case of level crossings: There, a jump from one *state* of the Hilbert space into another happened (by emptying or filling a level), causing an *integer* jump in $\langle Q \rangle$. Here, at half-integer Φ, we jump from one *Hilbert space* into another, and the corresponding charges do *not* have to differ by an integer.

The calculation of other vacuum quantities can be done now, with a procedure analogous to the one in [12], taking into account the present threshold mode spectrum of H. We do not give the details of the calculations here, but just state the results (figs. 2b-d):

$$\langle J \rangle = \langle J_K \rangle = sign(m) \frac{\epsilon^2}{4} \tag{4.11a}$$

$$\langle S \rangle = -sign(m) \frac{|\epsilon|}{4}, \quad (m \to 0) \tag{4.11b}$$

$$\langle J_C \rangle = -sign(m) \frac{2N\epsilon + \epsilon^2}{4} = -sign(m) \frac{\Phi^2 - N^2}{4} \tag{4.11c}$$

where we defined

$$\Phi = N + \epsilon, \quad N = \text{integer}, \quad -\tfrac{1}{2} < \epsilon < \tfrac{1}{2}. \tag{4.12}$$

Here, since there is no magnetic field outside the string, the total angular momentum is equal to the kinematical (fermionic) one. Also, the formula for $\langle S \rangle$ is again exact only in the massless limit.

Formula (4.11c) needs special commenting: J_C is the so-called canonical angular momentum, defined

$$J_C = \int \tfrac{1}{2}\psi^\dagger(-i\partial_\phi + \tfrac{1}{2}\sigma_3)\psi d^2x + h.c. \tag{4.13}$$

This quantity is not gauge invariant, but it can be defined for a rotationally symmetric configuration. In [12] it was argued that J_C is not a physical observable, but rather that the total angular momentum J, defined as the angular momentum of fermions plus local gauge fields, is the physical quantity. J and J_C are related by

$$J = J_C - \Phi Q \tag{4.14}$$

As we see from (4.11a), $\langle J \rangle$ is a periodic function of Φ. On the contrary, $\langle J_C \rangle$ is *not* periodic in Φ, and thus, by the arguments given in the beginning of this section, it is not a physical observable, in agreement with the earlier claim. The reason for this nonperiodicity can be traced to the fact that, although J_C can be defined so that it be invariant under small gauge transformations, it is still not invariant under gauge transformations of the form (4.3), with nonzero winding number around the origin. Also, as was pointed out in [12], $\langle J_C \rangle$ cannot be defined as a naive subtracted Dirac sea sum, because it is not a purely fermionic quantity. A special subtraction procedure is needed in order to obtain (4.11b), else the obtained value is infinite.

We examine now the problem of the infinitely thin string using local, rather than spectral, boundary conditions at $\rho = \rho_0$. The conditions that are compatible with the ones at $\rho = R$ are

$$v|_{\rho=\rho_0} = 0 \tag{4.15}$$

With these conditions, there are *no* restrictions on n and all the modes in (4.6) are acceptable. Since n is unbound from both above and below, we see that again the spectrum of threshold modes of H is periodic in Φ, since shifting Φ by one corresponds to shifting n by one.

In order to define the index of H_0, and so the vacuum charge, we need to use the same limiting procedure and subtraction near $\rho = \rho_0$ as we did near $\rho = R$ in (3.11-3.12). Specifically, we define

$$ind(H_0) = \lim_{\delta \to 0+} \lim_{\epsilon \to 0+} \left(\sum_{n=-\infty}^{\infty} \int_{\rho_0+\delta}^{R-\epsilon} |u_n(\Phi)|^2 d^2x - \sum_{n=-\infty} \int_{\rho_0+\delta}^{R-\epsilon} |u_n(0)|^2 d^2x \right). \quad (4.16)$$

The order of the limits is immaterial, since the integrals can be broken into two parts, one near $\rho_0 + \delta$ and another near $R - \epsilon$, and the limit of the sums of each part is separately finite. Normalizing the modes (4.6) and plugging in (4.16) we get

$$\begin{aligned} ind(H_0) = \lim_{\delta \to 0+} \lim_{\epsilon \to 0+} \Bigg(\sum_{n=-\infty}^{\infty} & \frac{(R-\epsilon)^{2(n-\Phi+1)} - (\rho_0+\delta)^{2(n-\Phi+1)}}{R^{2(n-\Phi+1)} - \rho_0^{2(n-\Phi+1)}} \\ - \sum_{n=-\infty}^{\infty} & \frac{(R-\epsilon)^{2(n+1)} - (\rho_0+\delta)^{2(n+1)}}{R^{2(n+1)} - \rho_0^{2(n+1)}} \Bigg) \end{aligned} \quad (4.17)$$

Assume that $R - \epsilon > \rho_0 + \delta$, which is certainly true for small enough ϵ and δ. Then, for n large enough, such that $n - \Phi + 1 \gg 0$ and $n + 1 \gg 0$, the terms $(R - \epsilon)$ in the numerator and R in the denominator completely dominate the $(\rho_0 + \delta)$ and ρ_0 terms. Correspondingly, for $-n$ large enough, such that $n - \Phi + 1 \to 0$ and $n + 1 << 0$, the terms $(\rho_0 + \delta)$ and ρ_0 dominate. Since we can always drop a finite number of terms from both sums in (4.17), we can write

$$\begin{aligned} ind(H_0) = & \lim_{\epsilon \to 0+} \sum_{n=N}^{\infty} \left[(1 + \frac{\epsilon}{R})^{2(n-\Phi+1)} - (1 - \frac{\epsilon}{R})^{2(n+1)} \right] \\ & + \lim_{\delta \to 0+} \sum_{n=M}^{\infty} \left[(1 + \frac{\delta}{\rho_0})^{2(-n-\Phi+1)} - (1 + \frac{\delta}{\rho_0})^{2(-n+1)} \right] \end{aligned} \quad (4.18)$$

where N and M are numbers large enough so that the previous approximations become valid. The sums can now be evaluated exactly and the limits can be taken. The first sum in the limit gives Φ, as calculated in section III. The second sum, however, the one involving ρ_0, gives in the limit $-\Phi$, and so overall

$$ind(H_0) = 0 \tag{4.19}$$

With local boundary conditions, *no* charge is induced around a Dirac string for *all* values of Φ. We can verify as well that the other vacuum quantities also vanish. So, the requirement of periodicity of the vacuum numbers in Φ and vanishing at Φ =integer is in this case trivially satisfied.

This result should not be surprising, in view of the behavior of local boundary conditions demonstrated in section III. In the same way that, as we turn the flux on, the boundary at $\rho = R$ becomes a source of charge, the boundary at $\rho = \rho_0$ becomes a sink of charge. As Φ increases, there is a continuous radial current bringing charge from the boundary and dumping it in the string (or vice versa). The total charge, however, remains constant and equal to zero.

One question that arises is whether the boundary conditions (4.8) or (4.15), at $\rho = \rho_0$, would change the results of sections II and III, when used with nonsingular (stringless) potentials. It is easy to see that nothing changes. For such potentials, A_ϕ becomes zero near the origin, and so the condition (4.9) (with $\Phi = 0$) gives exactly the same modes as the requirement for square integrability ($n \geq 0$). Also, we can check that the modes (2.17) for a regular gauge field have the same small-ρ behavior as the ones for zero field, and so the contribution of the second sum in (4.18) vanishes, yielding the standard result.

Finally, we point out that a third boundary condition, which is usually adopted in well-defined (stringless) two-dimensional problems, when solved in polar coordinates, could have been used. If we impose

$$\lim_{\rho \to 0}(\rho^{\frac{1}{2}}\psi) = 0 \tag{4.20}$$

then we see that (4.7) is satisfied in the limit $\rho_0 \to 0$. It is easy to check that this boundary condition is actually equivalent to the spectral boundary conditions (4.8), leading to the same threshold mode spectrum and identical results.

Concluding this section, we see that in the case of the (fractional) Dirac string, different boundary conditions lead indeed to different field theories. Which conditions are to be used is a matter of preference in this case, given the unphysicality of a string with fractional Φ. A "sink" in our space is as peculiar as half-integral "jumps" in the charge. The reassuring fact is that the results agree, as they should, for Φ =integer, and correspond to a free hamiltonian.

V. The case of a sphere.

The main advantage of working with a finite space is that the spectrum of H becomes discrete, and so the counting of states becomes straightforward and the calculation of vacuum quantities relatively easy. One should be cautious, however, because the presence of the boundary may introduce spurious contributions to these quantities, as is the case with spectral boundary conditions, and thus lead to the wrong infinite-space limit. On the other hand, an infinite space frees us from the boundaries, but counting arguments in general cannot be used there, because they may be misleading. For example, we point out that, for an infinite space, the hamiltonian (2.1) has *extra* localized normalizable threshold modes, other than the ones given in (2.17). Specifically, for an infinite space (and $\Phi > 0$ for concreteness), there are $[\Phi]$ in number normalizable threshold modes with energy $E = m$ of the form (2.17a), with $0 \leq n < [\Phi] - 1$ (required for the square integrability of the wavefunctions for $\rho \to \infty$). In addition to these modes, we have $[\Phi] - 1$ modes with *opposite* energy $E = -m$, having the form

$$u_n = e^a \rho^n e^{in\phi} \ , \quad 0 \leq n \leq [\Phi] - 2$$

$$v_n = -2me^a \rho^{-n-1} e^{i(n+1)\phi} q_n \tag{5.1a}$$

where q_n is the solution of the equation

$$\nabla q_n = e^{-2a}\rho^{2n+1} \tag{5.1b}$$

that behaves like $\rho^{2(n+1)}$ at $\rho = 0$ and falls off like $\rho^{2(n-\Phi+1)}$ at infinity. That such a solution of (5.1b) exists is guaranteed by the $\bar{\partial}$-Poincaré lemma. In the special case of a rotationally symmetric configuration, where a does not depend on ϕ, the extra modes take the explicit form

$$v_n = -2me^a\rho^{-n-1}e^{i(n+1)\phi}\int_0^\rho e^{-2a}r^{2n+1}dr\ , \quad n = 0, 1, ...[\Phi] - 2. \tag{5.2}$$

Thus we see that, if we considered the number of normalizable threshold modes of H as an indication of the amount of charge induced by the gauge field, motivated by (2.14), we would get a completely wrong answer, even in the case of integer Φ, since $n_+ - n_- = 1$ for *all* Φ. Obviously, the symmetry (2.12) of the discrete spectrum does not leave invariant the spectral density of the continuum spectrum. Also, the *continuum* (scattering) zero modes have to be considered, accounting for the fractional part of $\langle Q\rangle$ in the limit $m \to 0$.

The local boundary conditions (3.5) were shown to induce no boundary contributions to vacuum quantities, and so they can be used as a reliable calculational tool to obtain the infinite-space limit. It is, however, instructive to solve the problem on a compact *boundaryless* space, where counting arguments are still valid, due to the discreteness of the spectrum of H, and one does not have to worry about boundary contributions. Of course, some limitations are implied by the structure of the space. The most obvious one is the fact that the total flux running out of the space has now to be quantized into an integer. This is true for closed two-dimensional manifolds of arbitrary genus (i.e., number of handles), and the most physical way to see it is by noticing that, if one tried to write the gauge potential in a globally defined gauge, there would have to exist one or more singular points, corresponding to Dirac strings "bringing in" the flux. For these strings to be unobservable, their flux should be an integer, thus giving the quantization condition. In more rigorous terms, the quantization is a consequence of the fact that the second Cech cohomology class with integer coefficients of any closed two-dimensional manifold contains the integers.

Another peculiarity of compact spaces is that the definitions of linear and angular momentum may mix, if we choose different points of the space as the origin, or may even not exist at all, if the manifold does not have the corresponding invariances.

For the purposes of this calculation, we choose our space to be a sphere of radius R. This is the most natural compactification of R^2, since it is homogeneous and isotropic. Fixing a point on the sphere, the translational and rotational invariance of the original R^2 space becomes the $O(3)$ invariance on the sphere, with rotations around axes normal to the major axis passing through this point corresponding to translations. The fact that these rotations do not commute is a "finite volume" correction to the commutations of the generators of translations on a flat space.

What we will achieve with this calculation is an explicit demonstration of the Atiyah-Singer theorem at work. The existence of string singularities in the gauge field will prove to be harmless, and actually beneficial, for the solution of the problem. Calculation of some vacuum quantities on the sphere, then, will show that the physical results are well-behaved, in spite of the existence and the arbitrariness of the position of the strings, and respect the geometric symmetries of the problem.

Defining θ and ϕ to be the polar and azimuthal angles on the sphere, the metric of spacetime becomes

$$ds^2 = -dt^2 + R^2(d\theta^2 + \sin^2\theta d\phi^2). \tag{5.3}$$

In order to be able to write the action for fermions on this space, we need to define the dreibein $e^m{}_\mu$, referring to a local orthonormal frame at each point of spacetime, such that

$$g_{\mu\nu} = \eta_{mn} e^m{}_\mu e^n{}_\nu$$

$$\eta^{mn} = g^{\mu\nu} e^m{}_\mu e^n{}_\nu \tag{5.4}$$

where $g_{\mu\nu}$ is the metric tensor implied by (5.3) and η_{mn} is a flat metric of the form $(-1,1,1)$. From $e^m{}_\mu$ we can calculate the spin connection $\omega_{mn\mu}$, defined by

$$e^m{}_{\mu,\nu} - e^m{}_{\nu,\mu} = \omega^m{}_{n\mu} e^n{}_\nu - \omega^m{}_{n\nu} e^n{}_\mu$$

$$\omega_{mn\mu} = -\omega_{nm\mu} \tag{5.5}$$

where comma denotes ordinary differentiation and greek (latin) indices are lowered and raised with $g_{\mu\nu}$ (η_{mn}) correspondingly. Now the Dirac lagrangian can be written

$$L = \overline{\Psi}(i\gamma^m e_m{}^\mu D_\mu - m)\Psi \tag{5.6}$$

where γ^m are usual (flat) three-dimensional Dirac matrices, and the covariant derivative D_μ contains both a gauge and a spin connection part:

$$iD_\mu = i\partial_\mu - \tfrac{1}{2}\omega_{mn\mu}\sigma^{mn} + A_\mu \tag{5.7}$$

with σ^{mn} being the commutators of the γ-matrices

$$\sigma^{mn} = \frac{i}{4}[\gamma^m, \gamma^n]. \tag{5.9}$$

Choosing our local frame to be $(\hat{e}_t, \hat{e}_\theta, \hat{e}_\phi)$, we can calculate the dreibein $e^m{}_\mu$ for the specific metric (5.3) explicitly:

$$e^0{}_0 = e^1{}_1 = 1 \ , \quad e^2{}_2 = \sin\theta \ , \quad \text{all others} = 0, \tag{5.9}$$

the indices $\mu = 0,1,2$ corresponding to t, θ, ϕ in this order. From this and (5.5) $\omega_{mn\mu}$ is calculated to be

$$\omega_{12\,2} = -\omega_{21\,2} = -\cos\theta \ , \quad \text{all others} = 0. \tag{5.10}$$

As expected, the 0-component of all quantities decouples from the spatial components.

Making the specific choice of γ-matrices

$$\gamma^0 = \sigma_3 \ , \quad \gamma^1 = i\sigma_1 \ , \quad \gamma^2 = i\sigma_2 \tag{5.11}$$

and choosing the gauge $A_0 = 0$, (5.6) becomes

$$L = \overline{\Psi}(i\gamma^0\partial_0 - \gamma^0 H)\Psi \tag{5.12}$$

where the hamiltonian H acquires the form

$$H = \frac{1}{R}\begin{bmatrix} mR & D \\ & \\ D^\dagger & -mR \end{bmatrix} \ , \quad \begin{aligned} D &\equiv \partial_\theta - \frac{i}{\sin\theta}\partial_\phi - iA_\theta - \frac{1}{\sin\theta}A_\phi + \tfrac{1}{2}\cot\theta \\ D^\dagger &\equiv -\partial_\theta - \frac{i}{\sin\theta}\partial_\phi + iA_\theta - \frac{1}{\sin\theta}A_\phi - \tfrac{1}{2}\cot\theta \end{aligned} \tag{5.13}$$

The spinor Ψ is defined in the local orthocanonical frame in space $(\hat{e}_\theta \ , \ \hat{e}_\phi)$. We should notice that, when we transcribe the path $\theta =$ constant, ϕ from 0 to 2π on the sphere, our local frame undergoes a full rotation around itself, and, correspondingly, the spinor Ψ should pick up a minus sign. This means that Ψ satisfies

$$\Psi(\phi = 2\pi) = -\Psi(\phi = 0). \tag{5.14}$$

If we make the substitution

$$\Psi = e^{\frac{i}{2}\sigma_3\phi}\psi \tag{5.15}$$

we see that the new spinor ψ satisfies periodic boundary conditions in ϕ and is thus single-valued on the sphere.

We impose now the Coulomb gauge on the gauge field:

$$\partial_i A^i = \frac{1}{R\sin\theta}(\sin\theta A_\theta)_{,\theta} + \frac{1}{R\sin^2\theta}A_{\phi,\phi} = 0 \tag{5.16}$$

and write it in the form

$$A_\theta = -\frac{1}{R\sin\theta}\partial_\phi a \, , \quad \frac{1}{\sin\theta}A_\phi = \frac{1}{R}\partial_\theta a \tag{5.17}$$

that trivially satisfies (5.16). In terms of the new spinor ψ and the field a, the hamiltonian becomes

$$H = \frac{1}{R}\begin{bmatrix} mR & \nabla - \nabla a - \frac{1}{2}e^{-i\phi}\tan\frac{\theta}{2} \\ -\overline{\nabla} - \overline{\nabla} a + \frac{1}{2}e^{i\phi}\tan\frac{\theta}{2} & -mR \end{bmatrix},$$

$$\text{where} \quad \begin{aligned} \nabla &\equiv e^{-i\phi}(\partial_\theta - \frac{i}{\sin\theta}\partial_\phi) \\ \overline{\nabla} &\equiv e^{i\phi}(\partial_\phi + \frac{i}{\sin\theta}\partial_\phi) \end{aligned} . \tag{5.18}$$

This is exactly of the form (2.1), expressed in polar coordinates, with the substitution of $\frac{1}{\rho}\partial_\phi$ by the covariant derivative on the sphere $\frac{1}{R\sin\theta}\partial_\phi$, and the addition of the extra terms proportional to $\frac{1}{R\tan\frac{\theta}{2}}$ that come from the nontrivial spin connection.

We examine now how the threshold solutions of (5.18) arise. These modes satisfy

$$(\overline{\nabla} + \overline{\nabla} a - \tfrac{1}{2}e^{i\phi}\tan\frac{\theta}{2})u = 0 \tag{5.19a}$$

$$(\nabla - \nabla a - \tfrac{1}{2}e^{-i\phi}\tan\frac{\theta}{2})v = 0 \tag{5.19b}$$

for energies $+m$ and $-m$ respectively. The general form of the solutions of these equations is

$$u = \left(\cos\frac{\theta}{2}\right)^{-1} e^{-a} f \quad \text{with} \quad \overline{\nabla} f = 0 \tag{5.20a}$$

$$v = \left(\cos\frac{\theta}{2}\right)^{-1} e^{a} g \quad \text{with} \quad \nabla g = 0. \tag{5.20b}$$

Equations (5.20a,b) tell us that f and g^* are meromorphic functions on the sphere, considered as a complex manifold. Such functions are either constants or have one or more singularities, of the form $\frac{1}{z^n}$ locally around the singularity. Moreover, functions with a singularity of the same order n at the same point, are linearly dependent modulo less singular functions.

On the other hand, if we write the gauge field in a globally defined gauge, it will have one or more singular points, corresponding to Dirac strings piercing the sphere. The position of these points can be moved around with a singular gauge transformation. If we arrange for the singularities to occur at a single point, the azimuthal component of the gauge potential near the point will behave like

$$A \sim -\frac{\Phi}{2\pi\rho} \quad \Rightarrow \quad a \sim -\Phi a \ln \rho \tag{5.21}$$

(Φ is the total magnetic flux on the sphere and $-\Phi$ is the flux of the string), ρ being the distance from the point.

We can now choose the position of both the string and the singularity of f or g to arise at $\theta = \pi$. Then, the modulus of the wavefunctions (5.20) around $\theta = \pi$ will behave like

$$|u| \sim \theta^{\Phi-n-1}\,, \quad |v| \sim \theta^{-\Phi-n-1}. \tag{5.22}$$

So, for $\Phi > 0$, $|u|$ will be regular for all values of n such that

$$\Phi - n - 1 \geq 0 \quad \Rightarrow \quad 0 \leq n \leq \Phi - 1 \tag{5.23a}$$

and for $\Phi < 0$, $|v|$ will be regular for all values of n such that

$$-\Phi - n - 1 \geq 0 \quad \Rightarrow \quad 0 \leq n \leq -\Phi - 1. \tag{5.23b}$$

(The condition $n \geq 0$ is needed because else f or g will not have a singularity at $\theta = 2\pi$, and so they will have a singularity somewhere else.) Since at each level of singularity n the functions f and g are essentially unique, we conclude that we have overall $|\Phi|$ modes of the appropriate spin, in accordance with the index theorem.

The above construction is not unique. We could have spread all $|\Phi|$ Dirac strings, of strength $-sign(\Phi)$ each, at different points of the sphere. Then, for each string, there is a unique analytic function f (or g^*, for $\Phi < 0$) with exactly one pole of unit strength at the position of the string, and one zero at $\theta = 0$ (we must always have # of poles = # of zeros, the poles counted with accordance to their strength). The corresponding wavefunctions (5.20) are regular, since the string "eats" the singularity of f (or g) and the zero "kills" the singularity of (5.20) at $\theta = \pi$. Overall, we have $|\Phi|$ solutions of proper sign, as before.

We see that the Dirac string singularity here not only is not harmful, but it actually collaborates with the analytical properties of the wavefunctions to produce the correct threshold modes. The explicit form of the modes depends on the position of the strings, but the physical results should be independent of the strings. Note that the -1 in the exponents of (5.22), due to the $\left(\cos\frac{\theta}{2}\right)^{-1}$ factor in (5.20), comes from the nontrivial spin connection on the sphere and is essential for the correct counting of modes.

To illuminate the previous arguments, we solve explicitly the problem of a constant magnetic field all over the sphere, of strength

$$B = \frac{2\pi\Phi}{4\pi R^2} = \frac{\Phi}{2R^2} = \text{constant}. \tag{5.24}$$

The azimuthal component of the gauge field A and the quantity a can be chosen to be

$$A = \frac{\Phi}{2R}\tan\frac{\theta}{2} \quad \Rightarrow \quad a = -\Phi\ln\cos\frac{\theta}{2} \tag{5.25}$$

which has a string singularity at the south pole. From now on we adopt the shorthands

$$s \equiv \sin\frac{\theta}{2}\,, \quad c \equiv \cos\frac{\theta}{2}\,, \quad t \equiv \tan\frac{\theta}{2}. \tag{5.26}$$

The form of the general meromorphic function $f \sim z^n$ on the sphere can be found by conformally projecting onto the plane tangent at the north pole, giving

$$z \equiv x + iy = 2Rte^{i\phi} \quad \Rightarrow \quad f \sim t^n e^{in\phi} \tag{5.27}$$

which has an n^{th} order singularity at the south pole. So, the solutions (5.20) become

$$u_n = s^n c^{\Phi-n-1} e^{in\phi} \tag{5.28a}$$

$$v_n = s^n c^{-\Phi-n-1} e^{-in\phi}. \tag{5.28b}$$

Since s and c become zero at the north and south pole respectively, we see that, for $0 \le n \le \Phi - 1$, u_n are regular and, for $0 \le n \le -\Phi - 1$, v_n are regular, as derived earlier. The spin connection term has exactly the form of a uniform magnetic field with strength -1 for the upper component and $+1$ for the lower one.

An alternative way to find (5.28), avoiding the string singularities, is to cover the sphere with two patches, one covering the northern hemisphere and one covering the southern, and choosing regular gauges at each patch. The gauges are related at the equator with the gauge transformation

$$U = e^{i\Phi\phi} \quad \Rightarrow \quad \psi_N = e^{i\Phi\phi}\psi_S \quad \text{at } \theta = \frac{\pi}{2} \tag{5.29}$$

$\psi_{N,S}$ being the Dirac fields on the northern and southern hemisphere respectively. We see that, for the gauge transformation to be well-defined and the Dirac fields to be single-valued, we must indeed have the quantization condition Φ =integer. For simplicity suppose $\Phi > 0$. Then, the form of the solutions on each hemisphere is

$$u_{N,n} = s^n c^{\Phi-n-1} e^{in\phi} \quad n \ge 0 \tag{5.30a}$$

$$u_{S,m} = s^{\Phi-m-1} c^m e^{-i(m+1)\phi} \quad m \ge 0. \tag{5.30b}$$

The condition (5.29) leads to the requirement

$$e^{in\phi} = e^{i\Phi\phi} e^{-i(m+1)\phi} \quad \Rightarrow \quad n + m = \Phi - 1. \tag{5.31}$$

There are exactly Φ combinations of non-negative values (n, m) satisfying (5.31), so we recover the same modes as before. The treatment for $\Phi < 0$ is similar.

It is easy to check that, here, the modes analogous to (5.1) behave singularly at $\theta = \pi$ and so they are indeed, as expected, only a peculiarity of the infinite R^2 space.

In the limit of massless fermions, the contributions of positive and negative energy states in the expression for the vacuum charge *density* cancel, and thus the density can be evaluated as

$$\langle j^0 \rangle = -\tfrac{1}{2} sign(m) \sum_n |u_n|^2 \tag{5.32}$$

where $|u_n|$ are the (orthogonal) modes (5.28) (assume $\Phi > 0$ for simplicity). In the case of constant B, this density should be constant, due to the spherical symmetry of the problem, and equal to the total charge over the area of the sphere.

To verify it, we first normalize the states (5.28):

$$u_n = \left[\frac{\Phi}{4\pi R^2} \binom{\Phi - 1}{n} \right]^{\frac{1}{2}} s^n c^{\Phi - n - 1} e^{in\phi} \tag{5.33}$$

where we used the identity

$$\int_0^1 x^n (1-x)^m dx = \frac{1}{(n+m+1)\binom{n+m}{n}}. \tag{5.34}$$

Then (5.33) gives

$$\begin{aligned} \langle j^0 \rangle &= -\tfrac{1}{2} sign(m) \frac{\Phi}{4\pi R^2} \sum_{n=0}^{\Phi - 1} \binom{\Phi - 1}{n} s^{2n} c^{2(\Phi - n - 1)} \\ &= -\tfrac{1}{2} sign(m) \frac{\Phi}{4\pi R^2} (s^2 + c^2)^{\Phi - 1} = -\tfrac{1}{2} sign(m) \frac{\Phi}{4\pi R^2} \end{aligned} \tag{5.35}$$

in agreement with our expectations.

We can also calculate the angular momentum induced on the sphere. Due to the symmetry of the problem, there is no electric field produced by the vacuum charge, and thus no electromagnetic angular momentum. So the total angular momentum *equals* the kinematical (fermionic) one. (Notice that the canonical angular momentum is not well-defined here, since it depends on the position of the Dirac strings, further indicating its nonphysical nature.) The density of fermionic angular momentum j_K can, in the $m \to 0$ limit, be calculated from the zero mode contribution. Note that, although j_K is a local quantity, it is defined with respect to a fixed major axis of the sphere, in this case the polar axis:

$$j_K = \tfrac{1}{2}\psi^\dagger K\psi + h.c. \ , \quad \text{where}$$

$$K \equiv -i\partial_\phi - A_\phi + \tfrac{1}{2}\sigma_3 = -i\partial_\phi - A_\phi + s^2\sigma_3 + \cos\theta\,\tfrac{1}{2}\sigma_3. \tag{5.36}$$

A_ϕ is the covariant ϕ-component of the gauge field and equals $R\sin\theta$ times the magnitude of the azimuthal component of A. The last term in the second expression for K is the projection of the spin on the direction of the polar axis. So, the previous terms can be interpreted as the orbital angular momentum of the fermions around this axis. The term proportional to s^2 is a finite volume correction (of order R^{-2}) to the orbital part, coming from the nontrivial geometry of the sphere.

We can anticipate the result for $\langle j_K\rangle$ by noticing that, because of the symmetry of the problem, the orbital part of $\langle j_K\rangle$ should vanish, else it would indicate a preferred direction on the sphere. The spin part is proportional to $\frac{1}{2}sign(\Phi)$ times the charge density, so we expect

$$\langle j_K\rangle = -\frac{1}{4}sign(m)\cos\theta\frac{|\Phi|}{4\pi R^2}. \tag{5.37}$$

To check this result, we plug the modes (5.33) (for $\Phi > 0$) in the expression

$$\begin{aligned}\langle j_K\rangle &= -\tfrac{1}{2}sign(m)\sum_{n=0}^{\Phi-1}\langle j_K\rangle_n \\ &= -\tfrac{1}{2}sign(m)\sum_{n=0}^{\Phi-1}\left[n-(\Phi-1)s^2+\tfrac{1}{2}\cos\theta\right]\frac{\Phi}{4\pi R^2}\binom{\Phi-1}{n}s^{2n}c^{2(\Phi-1-n)} .\end{aligned} \tag{5.38}$$

By writing $n-(\Phi-1)s^2 = nc^2-(\Phi-1-n)s^2$ and making the change of variable $\Phi-1-n = k$ in the second term, we can show that the contributions of these two terms in the sum cancel. The remaining is just the spin part, and we recover (5.37). So in this case *all* the angular momentum is due to spin.

Concluding, we see that we can work with a singular gauge on the sphere, with string singularities, and still recover the expected physical results.

VI. Conclusions.

We showed, in a special case, that local boundary conditions can be used with fermions, and give reasonable and desirable physical results. The standard definition of the index of an operator has to be modified in order to accommodate the new conditions into an index theorem. With these conditions, the boundaries do not accumulate extra contributions of vacuum quantities, but become "transparent", allowing currents to flow out of them as we turn on background fields of nontrivial topology.

We also calculated the induced vacuum quantum numbers in some cases of interest, and clarified some issues concerning the proper definition of the angular momentum.

Although the derivation of our index theorem with local boundary conditions was done under several simplifying technical assumptions, namely on a flat circular disc with rotationally symmetric gauge field configurations becoming pure gauge near the boundary, we have extended this derivation to the completely general case of an arbitrary curved two dimensional manifold with holes and arbitrary gauge fields. The generalization is straightforward but messy, involving a few technical tricks, and will not be presented here.

It is quite plausible that our procedures can also be generalized to higher dimensional cases and nonabelian operators (the generalization is immediate when the problem can be reduced to a direct product of two dimensional abelian operators). This, however, has not been done yet and still constitutes an open question.

REFERENCES

[0] A. P. Polychronakos, Nucl. Phys. **B283** (1987) 268.

[1] For a review see T. Eguchi, R. Gilkey and A. Hanson, Phys. Reports **66** (1980) 213.

[2] M. F. Atiyah and I. M. Singer, Ann. Math. **87** (1968) 485.

[3] E. Witten, Phys. Lett. **117B** (1982) 324.

[4] A. N. Redlich, Phys. Rev. Lett. **52** (1984) 18; Phys. Rev. **D 29** (1984) 2366.

[5] A. J. Niemi and G. W. Semenoff, Phys. Rev. **D3** (1984) 2366.

[6] R. Jackiw and C. Rebbi, Phys. Rev. **D13** (1976) 3398.

[7] M. F. Atiyah, V. K. Patodi and I. M. Singer, Math. Proc. Camb. Phil. Soc. **77** (1975) 43; **78** (1975) 405; **79** (1976) 71; Bull. London Math. Soc. **5** (1973) 229.

[8] C. Callias, Comm. Math. Phys. **62** (1978) 213; R. Bott and R. Seeley, Comm. Math. Phys. **62** (1978) 235.

[9] M. F. Atiyah and R. Bott, in Differential Analysis, Bombay Colloquium, Oxford, 1964.

[10] A. J. Niemi and G. W. Semenoff, Phys. Rev. Lett. **51** (1983) 2077.

[11] M. B. Paranjape, Phys. Rev. **55** (1985) 2390.

[12] A. P. Polychronakos, Nucl. Phys. **B278** (1986) 207.

[13] M. Ninomiya and C-I. Tan, Nucl. Phys. **B257** (1985) 199.

[14] M. B. Paranjape and G. W. Semenoff, Phys. Lett. **132B** (1983) 369.

[15] R. Jackiw, Phys. Rev. **D16** (1984) 2375.

[16] J. Kiskis, Phys. Rev. **D15** (1977) 2329.

[17] A. Chodos, R. L. Jaffe, K. Johnson, C. B. Thorn and V. F. Weisskopf, Phys. Rev. **D9** (1974) 1371.

[18] J. Goldstone and R. L. Jaffe, Phys. Rev. Lett. **51** (1983) 1518.

[19] A. J. Niemi, Nucl. Phys. **B253** (1984) 14.

[20] R. M^c Kenzie and F. Wilczeck, Phys. Rev. **D30** (1984) 2194.

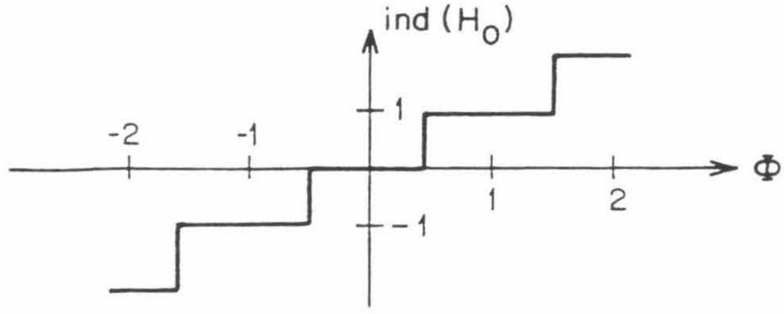

Fig. 1

Fig. 1: The index of the two–dimensional Dirac operator with the global boundary conditions (2.8).

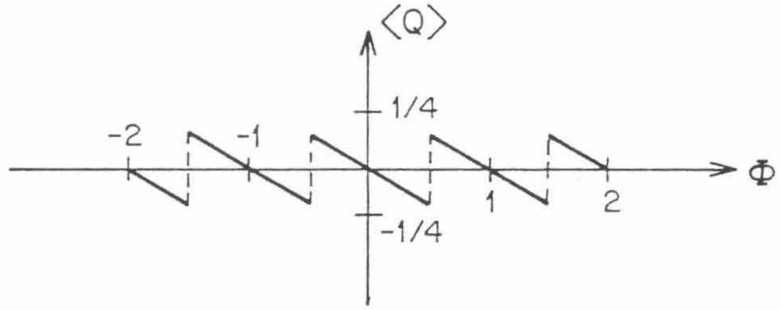

Fig. 2a

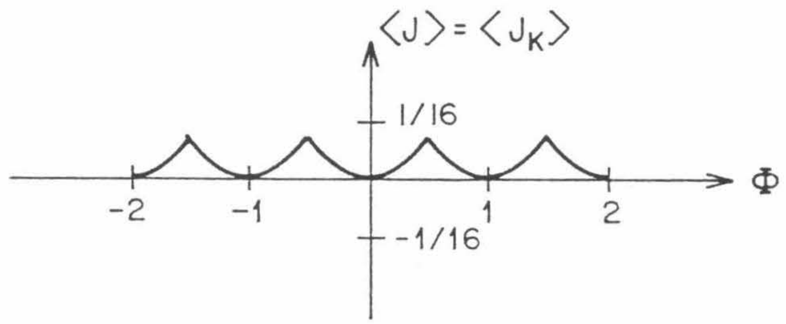

Fig. 2b

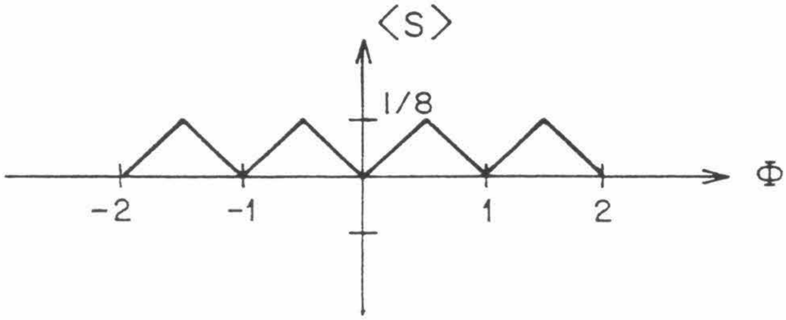

Fig. 2c

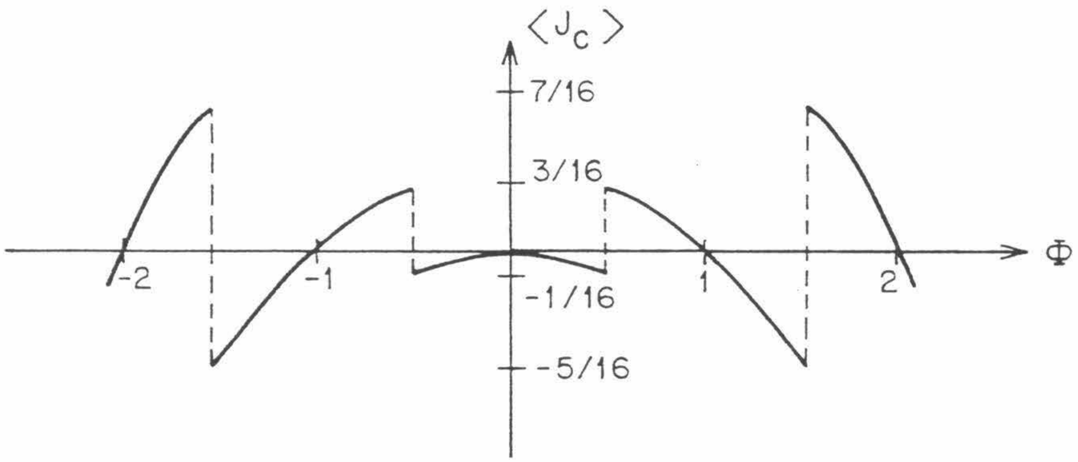

Fig. 2d

Fig. 2a–d: The charge, angular momentum, spin and "canonical" angular momentum induced around a Dirac string with flux Φ, with spectral boundary conditions imposed at the position of the string.

CHAPTER 4

Topological Mass Quantization and Parity Violation in Odd Dimensional QED

I. Introduction.

As has already been previewed in chapter 2, in 2+1 dimensions, gauge theories acquire rather peculiar and surprising properties. A Chern-Simons term can be included in the action, playing the role of a mass term for the gauge fields [1]. The lagrangian density of this term is not gauge invariant under gauge transformations, but rather changes by a total derivative. So, the action, being its integral over all spacetime, changes by a surface term and, provided that the transformation is "small" (this meaning that the surface term vanishes), it is gauge invariant. There are, however, "large" gauge transformations that do not leave the action gauge invariant. If we compactify our spacetime into a three sphere S^3, in order that its volume be finite, these "large" gauge transformations are topologically nontrivial maps from S^3 to the gauge group G, and are classified by the third homotopy group of the gauge group $\pi_3(G)$. This group, for G anything but U(1), is nontrivial and equal to Z. For such gauge transformations, the action changes by a constant, proportional to the coefficient of the Chern-Simons term, times the winding number of the transformation (the element of Z). Since the action S appears in the path integral in the form $exp(iS)$, for this exponential to be gauge invariant the action must change by an integer multiple of 2π and thus the coefficient of the Chern-Simons term must be quantized [1].

On the other hand, although in odd dimensional spacetimes there are no perturbative anomalies, due to the nonexistence of the analog of γ^5, a theory of massless fermions coupled to gauge fields breaks parity and time reversal. This has been shown by demonstrating that, if the theory is regulated in a way that preserves parity, the effective action after integrating out the fermions will not be invariant under nontrivial gauge transformations with an odd winding number [2]. A gauge invariant (e.g., Pauli-Villars) regularization induces in the one-loop level a topological mass term for the gauge fields, but with a coefficient equal to *half* the quantization unit, whose global gauge noninvariance is compensated by the noninvariance of the rest of the effective action. This term is a pseudoscalar and thus breaks parity and induces a parity-violating part in the vacuum fermionic currents in the presence of external gauge fields [2,3]. The above properties can be generalized to higher odd dimensions, although there the Chern-Simons term is not a mass term and a higher homotopy group of the gauge group will be relevant.

It should be noted that, although the above-mentioned nontrivial gauge transformations cannot be continuously connected to the unity (the gauge transformation $U(x) = 1$), a gauge field configuration *is* continuously connected to its nontrivial gauge transform (through configurations that are, of course, gauge nonequivalent with the initial one). So, there is no consistent way of restricting the range of integration of the gauge fields such as to include only *one* copy out of each family of nontrivial gauge transformations of a gauge field and thus to evade the previous arguments.

In the abelian case, on the other hand, all homotopy groups other than π_1 vanish, and thus one does not get a quantization of the coefficient of the gauge mass term. Moreover, although it is known that parity breaking does occur, with the fermionic vacuum acquiring anomalous quantum numbers [3,4,5], there is no corresponding topological argument to demonstrate the necessity of parity violation. We provide here such an argument, as well as a quantization condition for the coefficient of the mass term in the U(1) case. As a subproduct, we point out and rectify an incompleteness in the standard proof of the global SU(2) anomaly in four dimensions. Our argument for the quantization of the mass term is essentially equivalent with the corresponding cohomological argument outlined in ref. [6], with some subtleties with the normalization at the lagrangian level straightened out.

A quantization condition is worked out also in ref. [7], in the presence of an "instanton" configuration, i.e., in the presence of a monopole point in three-dimensional spacetime that creates a transition between different total magnetic fluxes running through the two-dimensional spatial surface, as this surface "sweeps" the instanton. However, in the derivation it was assumed that, due to the topological mass term, the instanton produces a point electric charge, whose Dirac quantization condition then gives us the quantization of the mass term. What in fact happens is that the instanton field produces a *continuous* charge density distribution, proportional to the strength of the magnetic field on the spatial surface, and thus it is not necessary that we have a quantization condition at the lagrangian level. Our procedure is independent of such an assumption. Also, the "monopole" that we mention later in the text lies *outside* our spacetime and simply indicates a nonvanishing total magnetic flux, not a transition between different fluxes.

II. The quantization condition.

As we mentioned already, all the topological arguments for the nonabelian case assume a compactification of spacetime into a three-sphere S^3. The key point, here, is that we will consider a compactification of spacetime into a product of spheres $S^2 \times S^1$ ($T^3 = S^1 \times S^1 \times S^1$ would do just as well). This compactification is necessary when, for example, we want our spatial section to have a nonzero total magnetic flux. The topological mass term, in differential form notation, is

$$I_{CS} = \kappa \int_{S^2 \times S^1} AF \tag{2.1}$$

(κ is a dimensionless combination of the gauge field mass and the fermion coupling constant). The field strength $F = dA$ is a closed form but its integral over the space S^2 need not vanish. If S^2 contains a "monopole", the gauge field will be a connection of a nontrivial U(1) bundle over S^2 and the integral of F will give the "monopole" number

$$\int_{S^2} F = 2\pi\Phi \tag{2.2}$$

with Φ an integer. We now notice that on $S^2 \times S^1$ there *are* nontrivial gauge transformations, those with a nonzero winding number n around S^1. (Since $\pi_2(U(1)) = \pi_3(\mathrm{U}(1)) = 0$ this is the *only* type of nontrivial gauge transformations on $S^2 \times S^1$.) Performing such a gauge transformation Ω on a gauge field configuration with non-vanishing Φ, the mass term transforms:

$$I_{CS} \to I_{CS} + \kappa \int_{S^2\times S^1} \Omega^{-1} d\Omega F \tag{2.3}$$

Writing $\Omega^{-1}d\Omega = d\omega$, with ω an angle winding n times around S^1, one is tempted to write

$$\int_{S^2\times S^1} \Omega^{-1} d\Omega F = \int_{S^1} d\omega \cdot \int_{S^2} F = 2\pi n \cdot 2\pi\Phi. \tag{2.4}$$

However, this is the wrong result. The reason is that A cannot be globally defined, since dA belongs to a nontrivial element of the DeRham cohomology class $H^2_{DR}(S^2, Z)$, and thus $\int AF$ can only be defined as a sum over patches. For this sum to be independent of the patching, correction terms have to be included. (For a review of the relevant cohomology notions see reference [6].)

Specifically, we cover our spacetime with patches U_α, such that any finite intersection of them be simply connected (fig.1). Then the expression for $\int AF$ becomes

$$\int AF = \sum_\alpha \int A_\alpha F - \sum_{\alpha\beta} \int J_{\alpha\beta} + \sum_{\alpha\beta\gamma} \int K_{\alpha\beta\gamma} - \sum_{\alpha\beta\gamma\delta} H_{\alpha\beta\gamma\delta} \tag{2.5}$$

where J, K and H are 2-, 1- and 0-forms respectively, satisfying

$$A_\alpha F - A_\beta F = dJ_{\alpha\beta}$$

$$J_{\alpha\beta} + J_{\beta\gamma} + J_{\gamma\alpha} = dK_{\alpha\beta\gamma} \tag{2.6}$$

$$K_{\alpha\beta\gamma} - K_{\beta\gamma\delta} + K_{\gamma\delta\alpha} - K_{\delta\alpha\beta} = dH_{\alpha\beta\gamma\delta}$$

and the integrals in (2.5) are done over the boundaries of corresponding dimensionality lying in the intersection indicated by the indices. Knowing the transition functions for the gauge field A:

$$A_\alpha - A_\beta = d\psi_{\alpha\beta}$$

$$\Psi_{\alpha\beta} + \Psi_{\beta\gamma} + \Psi_{\gamma\alpha} = c_{\alpha\beta\gamma} \tag{2.7}$$

with $c_{\alpha\beta\gamma}$ being an element of the integer Čech cohomology class $H^2_C(S^2, Z)$ satisfying

$$c_{\alpha\beta\gamma} - c_{\beta\gamma\delta} + c_{\gamma\delta\alpha} - c_{\delta\alpha\beta} = 0$$

$$\sum_{\alpha\beta\gamma} c_{\alpha\beta\gamma} = \int_{S^2} F = 2\pi\Phi \tag{2.8}$$

(where the sum is over the intersection of the indicated one-dimensional boundaries with a given S^2) we can explicitly construct the transition functions for AF, in the fashion:

$$J_{\alpha\beta} = \Psi_{\alpha\beta}F$$

$$K_{\alpha\beta\gamma} = c_{<\alpha\beta\gamma}A_{\gamma>} \tag{2.9}$$

$$H_{\alpha\beta\gamma\delta} = c_{<\alpha\beta\gamma}\Psi_{\gamma\delta>}.$$

The symbol $< \cdots >$ means: Put the indices in increasing order, with repeated indices matching according to their *position*, not value, and multiply by the parity of the permutation. For example:

$$c_{<132}A_{2>} = -c_{123}A_3. \tag{2.10}$$

With these choices and the help of (2.7), (2.8), we may show that (2.6) are satisfied.

If we now perform a nontrivial gauge transformation around S^1, $\Psi_{\alpha\beta}$ and $c_{\alpha\beta\gamma}$ do not change, since we are changing A by a globally well-defined pure gauge field, and so the changes in J, K and H are

$$\Delta J_{\alpha\beta} = 0$$

$$\Delta K_{\alpha\beta\gamma} = c_{\alpha\beta\gamma} d\omega \tag{2.11}$$

$$\Delta H_{\alpha\beta\gamma\delta} = 0$$

and the corresponding change in (2.5) is

$$\sum_{\alpha} \int d\omega \cdot F + \sum_{\alpha\beta\gamma} \int c_{\alpha\beta\gamma} d\omega = \sum_{\alpha} \int_{S^1} d\omega \cdot \int_{S^2} F + \sum_{\alpha\beta\gamma} c_{\alpha\beta\gamma} \cdot \int_{S^1} d\omega$$
$$= 2\pi n \cdot 2\pi\Phi + 2\pi\Phi \cdot 2\pi n = 2 \cdot 2\pi n \cdot 2\pi\Phi \ . \tag{2.12}$$

The correct quantization condition can now be derived by demanding that the action change by an integer multiple of 2π under a nontrivial gauge transformation, which leads to the requirement:

$$4\pi\kappa = \text{integer} . \tag{2.13}$$

If this quantization condition holds, then the path integral is gauge invariant and all flux sectors give a nonzero contribution. If, on the other hand, the coefficient is an irrational multiple of the quantization unit, then each nonzero flux sector contributes *zero* to the path integral. The zero flux sector, though, still contributes a nonzero amount and so the overall path integral does not vanish and the quantum theory is not inconsistent. Thus, the quantization condition needs to hold only if we want to quantize the theory in a nontrivial flux background alone. In the opposite case, the theory simply does not contain any states of nonzero total flux.

It should be clear, now, why we have no quantization condition for S^3. The gauge field configurations that may lead to gauge noninvariance are those with nonzero magnetic flux running through a two-dimensional subsurface. On S^3 no such configurations are allowed. If we want them to be included (and to contribute a nonzero amount to the path integral), condition (2.13) is necessary. For T^3, configurations with nonzero flux running out of a spatial two-torus T^2 exist, and so (2.13) has to hold as well.

It should be realized that our quantization condition is equivalent to the cohomological argument [6]. There, if we assume that $S^2 \times S^1$ is a section of an $S^2_a \times S^2_b$ manifold with magnetic monopoles in each S^2 sector, and the gauge field is the appropriate three-dimensional restriction of the four-dimensional one-form, the topological lagrangian is defined only modulo the integral

$$\kappa \int\limits_{S^2_a \times S^2_b} F^2 \tag{2.14}$$

By decomposing A into its components A_a and A_b, lying on the first and second sphere respectively, and similarly the total derivative d into d_a and d_b, we have

$$F^2 = [(d_a + d_b)(A_a + A_b)]^2 = 2 d_a A_a d_b A_b + \text{ mixed terms} \tag{2.15}$$

where "mixed" terms are terms of the form $d_a A_b d_b A_b$ etc. A_b is not globally defined on S^2_b but, for a fixed point of that sphere, it is globally defined on all of S^2_a. So this term, for this point of S^2_b, can be written $d_a(A_b d_b A_b)$ (the two-form $d_b A_b$ is globally defined and closed) and integrated over S^2_a it vanishes, as do all the mixed terms. So overall

$$\kappa \int\limits_{S^2_a \times S^2_b} F^2 = 2\kappa \int\limits_{S^2_a \times S^2_b} d_a A_a d_b A_b = 2\kappa \cdot 2\pi\Phi \cdot 2\pi n \tag{2.16}$$

(where we assumed that the monopole numbers in S^2_a and S^2_b are Φ and n respectively) which leads to the same condition (2.13).

III. Parity violation.

To show that parity violation has to occur, we still assume the same compactification of spacetime as in section II, and closely follow the construction of references [2] and [8]. We construct an augmented Dirac operator

$$\not{D} = \gamma^i(\partial_i + A_i)$$

$$\gamma^i = \begin{bmatrix} i\sigma_i & 0 \\ & \\ 0 & -i\sigma_i \end{bmatrix}, \quad i = 1,2,3 \tag{3.1}$$

where $i\sigma_i$ are the (antihermitian) euclidean γ-matrices. Obviously, if λ_n are the eigenvalues of the Dirac operator $\not{D}_3 = i\sigma_i(\partial_i + A_i)$, the eigenvalues of $\not{D}$ are $\pm\lambda_n$.

We wish now to evaluate the determinant of $\not{D}_3$, which is the effective action obtained by integrating out the fermions, in a parity invariant way. A parity transformation inverts the spectrum of $\not{D}_3$ but obviously leaves the spectrum of $\not{D}$ invariant. Thus, $\det \not{D}$ can be regulated with a parity-invariant Pauli-Villars mass:

$$\det{}_M \not{D} = \prod_{\pm\lambda_n} \frac{\lambda_n}{\lambda_n + M} = \prod_{\lambda_n} \frac{\lambda_n^2}{\lambda_n^2 - M^2} \tag{3.2}$$

and a parity conserving definition of the determinant of $\not{D}_3$ is as the square root of $\det \not{D}$. Either the positive or negative value can be chosen, but the evolution of the value of the square root has to be done smoothly in the space of gauge potentials. One way to do this is to define $\det \not{D}_3$ as the product of the positive eigenvalues of $\not{D}$ and follow the evolution of these eigenvalues as we continuously move to other gauge field configurations.

We consider again an initial configuration with nonzero Φ. Then we construct a one-parameter family of configurations $A(\tau)$, interpolating adiabatically between the original gauge configuration, at $\tau = -\infty$, and its gauge transform A^{Ω}, at $\tau =$

$+\infty$, with Ω having a winding number n around S^1. Following the same adiabatic arguments of ref. [8], we deduce that the number of eigenvalues of $\not{D}$ that cross zero an odd number of times (i.e. change sign) during this process equals the number of zero modes of the four-dimensional Dirac operator

$$\not{D}_4 = \gamma^\mu(\partial_\mu + A_\mu),$$

$$A_4 = 0 \ , \quad \gamma^4 = i \begin{bmatrix} 0 & I \\ I & 0 \end{bmatrix} \ , \quad x^4 = \tau \tag{3.3}$$

defined on the space $S^2 \times S^1 \times R$. If this number is odd, then an odd number of positive eigenvalues of $\not{D}$ will evolve into negative ones, and so the determinant of $\not{D}_3$, defined as the product of the positive eigenvalues of $\not{D}$, will change sign (since the initial and final configurations are gauge equivalent, the respective spectra of $\not{D}$ are identical and thus the variation of $\det \not{D}_3$ is limited to a possible change of sign).

The number of zero modes of $\not{D}_4$ can be found using the Atiyah-Singer index theorem (in the abelian case *all* the zero modes have the same chirality), which states that

$$\text{ind}\not{D}_4 = \frac{1}{8\pi^2} \int\limits_{S^2 \times S^1 \times R} F^2 \tag{3.4}$$

The calculation of $\int F^2$ is similar to the one in (2.14-2.16) and we get

$$\text{ind}\not{D} = \Phi \cdot n \tag{3.5}$$

(The monopole number of $S^1 \times R$ is the difference of the integrals of A over S^1 at $\tau = +\infty$ and $\tau = -\infty$.) By choosing both Φ and n odd, we see that an odd number of eigenvalues of $\not{D}$ will change sign and so the parity-invariant defined $\det \not{D}_3$ will flip sign under this nontrivial gauge transformation. Thus we see that a parity-invariant definition of $\det \not{D}_3$ is inconsistent. We can obtain a gauge invariant $\det \not{D}_3$ by adding to the previous one the Chern-Simons form with a half-integral coefficient, but, of course, this term will break parity.

It is instructive to give here an explicit illustration of the spectral flow of $\not{D}$ and a construction of the zero modes of $\not{D}_4$. This will gain us intuition about the behavior of the spectrum of $\not{D}_3$ and will be useful in our considerations of the SU(2) anomaly in the next section.

We first show how we can explicitly find the zero modes of $\not{D}_4$ by exploiting the nontrivial topology of the gauge configuration in each two-dimensional component of our space. For simplicity, we assume that $A_{1,2}$ depend only on $x^{1,2}$, while $A_{3,4}$ depend only on $x^{3,4}$. Then we notice that, with an appropriate redefinition, the four-dimensional γ-matrices can be expressed in the representation:

$$\begin{aligned} \gamma^1 &= i\sigma_1 \otimes \sigma_3 \quad \gamma^3 = iI \otimes \sigma_1 \\ \gamma^2 &= i\sigma_2 \otimes \sigma_3 \quad \gamma^4 = iI \otimes \sigma_2 \end{aligned} \tag{3.6}$$

or for simplicity

$$\vec{\gamma} = i\vec{\sigma} \otimes \sigma_3 \quad \tilde{\gamma} = iI \otimes \tilde{\sigma} \tag{3.7}$$

(where vector is used for the 1,2 components and tilde for the 3,4 components). The Dirac spinor ψ is now a 2×2 matrix, with the first (second) matrix in the direct products in (3.7) acting on the first (second) index of ψ respectively. This would correspond to the conventional representation

$$\gamma^{1,2} = \begin{bmatrix} i\sigma_{1,2} & 0 \\ 0 & -i\sigma_{1,2} \end{bmatrix}, \quad \gamma^3 = i \begin{bmatrix} 0 & I \\ I & 0 \end{bmatrix}, \quad \gamma^4 = i \begin{bmatrix} 0 & -iI \\ iI & 0 \end{bmatrix}$$

$$\text{with } \psi^T = [\psi_{11}\ \psi_{21}\ \psi_{12}\ \psi_{22}] \tag{3.8}$$

Then the zero-eigenvalue equation for $\not{D}_4$ becomes

$$[\vec{\gamma}(\vec{\partial} + \vec{A}(\vec{x})) + \tilde{\gamma}(\tilde{\partial} + \tilde{A}(\tilde{A}))]\psi = 0. \tag{3.9}$$

By writing ψ in the decoupled form

$$\psi = \phi(\vec{x}) \otimes \chi(\tilde{x}) \tag{3.10}$$

ϕ and χ being two-columns, we get

$$\vec{\gamma} \cdot \vec{D}\phi \otimes \sigma_3 \chi + \phi \otimes \tilde{\gamma} \cdot \tilde{D}\chi = 0. \tag{3.11}$$

We notice now that $\vec{\gamma} \cdot \vec{D}$ and $\tilde{\gamma} \cdot \tilde{D}$ are two-dimensional Dirac operators on the spaces S^2 and $S^1 \times R$ respectively. The Atiyah-Singer index theorem tells us that these operators will have normalizable zero modes equal in number to $\frac{1}{2\pi}$ times the integral of the corresponding gauge field strengths over the respective spaces. Since we supposed that the gauge configurations on each component space are nontrivial (they contain a monopole), $\vec{\gamma} \cdot \vec{D}$ and $\tilde{\gamma} \cdot \tilde{D}$ *will* have zero modes, say n_1 and n_2 in number, where $n_{1,2}$ are the monopole numbers of the component spaces. By choosing ϕ to be any of the zero modes of $\vec{\gamma} \cdot \vec{D}$ and χ any of the zero modes of $\tilde{\gamma} \cdot \tilde{D}$, we can construct all solutions of eq. (22), in total $n_1 \cdot n_2$ solutions.

For an explicit illustration of the spectral flow of $\not{D}$ consider the configuration $\vec{A} = \vec{A}(\vec{x})$, A_3 =constant. Then the eigenvalue equation for $\not{D}$ is

$$[\gamma^3(\partial_3 + A_3) + \vec{\gamma} \cdot \vec{D}]\psi = \lambda\psi \tag{3.12}$$

with ψ satisfying antiperiodic boundary conditions in S^1:

$$\psi(x^3 = 0) = -\psi(x^3 = T), \tag{3.13}$$

T being the length of S^1. By redefining ψ as

$$\psi = e^{-iA_3 x^3}\phi\ , \quad \phi(x^3 = T) = -e^{iA_3 T}\phi(x^3 = 0) \tag{3.14}$$

we can get rid of A_3. Decomposing ϕ in terms of eigenstates of $\vec{\gamma} \cdot \vec{D}$

$$\phi = \sum_n a_n(x^3) \cdot \phi_n(\vec{x}) \ , \quad \text{with } \vec{\gamma} \cdot \vec{D}\phi_n = e_n \cdot \phi_n \tag{3.15}$$

and using the relations

$$\gamma^3 \psi_n = \psi_{-n} \ , \quad e_{-n} = -e_n \tag{3.16}$$

(we used the fact that γ^3 anticommutes with $\vec{\gamma} \cdot \vec{D}$), we end up with the equations

$$i\partial_3 a_{-n} + e_n a_n = \pm \lambda a_n$$

$$i\partial_3 a_n - e_n a_{-n} = \pm \lambda a_{-n} \tag{3.17}$$

for every n such that $e_n \neq 0$. The solution of these equations is:

$$a_n = e^{ikx^3} a_n(0) \tag{3.18}$$

$$\lambda = \pm \sqrt{e_n^2 + k^2}. \tag{3.18}$$

So, for nonzero e_n, λ cannot vanish. However, for $e_n = 0$ we get

$$i\partial_3 a_0 = \pm \lambda a_0 \tag{3.19}$$

and the solution of this, taking into account the boundary conditions (3.14), is

$$a_0 = e^{ikx^3} a_0(0) \ , \quad k = A_3 + \frac{2m+1}{T}\pi$$

$$\lambda = \pm k \ , \quad m = \text{integer}\,. \tag{3.20}$$

A nontrivial gauge transformation of the form $\Omega = e^{i2\pi n \frac{x^3}{T}}$ changes A_3 into $A_3 + \frac{2\pi n}{T}$. So, we see that as we smoothly vary A_3 from the original to its gauge transformed value, n eigenvalues of $D\!\!\!/$ for each zero eigenvalue of $i\vec{\gamma} \cdot \vec{D}$ will cross zero, overall $\Phi \cdot n$ eigenvalues, as calculated earlier.

IV. The SU(2) global anomaly.

As we saw in the previous section, a parity invariant regularization of $\det \not{D}_3$ (as the positive square root of $\det \not{D}$) is not gauge invariant. This is reminiscent of the four dimensional SU(2) global anomaly, where again a gauge invariant definition of the square root of the determinant of a Dirac operator is impossible. A crucial remark, however, is that, in contradistinction to the four dimensional SU(2) case, there is in the present case an operator that *commutes* with $\not{D}$, namely

$$\Gamma = \begin{bmatrix} I & 0 \\ 0 & -I \end{bmatrix} \tag{4.1}$$

with eigenvalues ± 1, and so each eigenstate of $\not{D}$ can be assigned an "index" (± 1) depending on its Γ eigenvalue. States with opposite $\not{D}$ eigenvalues have opposite indices and eigenvalues with $\Gamma = +1$ correspond to eigenvalues of $\not{D}_3$. One possible choice (in fact the correct choice) for the definition of $\det \not{D}_3$ is as the product of the eigenvalues of $\not{D}$ with index $+1$. This choice does not conserve parity, since indices flip sign under a parity transformation, but is obviously globally gauge invariant, since the initial and final spectra of $\not{D}$ are identical, with the *same* index assignments (fig. 2). Thus we see that the level-crossing picture alone does not guarantee that $\det \not{D}$ will flip sign, since there may be (and there is, in this case) a choice that preserves gauge invariance. In the four-dimensional SU(2) case there is no such obvious choice, but one should *prove* that such a clever choice is not possible in order that the existence of an SU(2) anomaly be established. We provide here such a proof.

The essential feature of the SU(2) case is that the classifying homotopy group is a finite group, namely Z_2, while in our case it is Z. To exploit this fact for the SU(2) case, we construct a continuous path of gauge field configurations depending on a parameter $\tau \in [0,1]$, such that $A(0) = A$, $A(\tau + \frac{1}{2}) = A(\tau)^\Omega$, Ω being the nontrivial element of $\pi_4(SU(2))$ (fig. 3). Thus $A(\frac{1}{2}) = A^\Omega$ and $A(1) = (A^\Omega)^\Omega = A$, due to the Z_2 nature of $\pi_4(SU(2))$. This is a closed path in the space of gauge field configurations and, due to the trivial topology of this space, it is contractible. In particular, since a specific eigenvalue of $\not{D}(A)$ is a continuous functional of A, its evolution $\lambda(\tau)$, with $\lambda(0)$ and $\lambda(1)$ fixed, is contractible to the trivial evolution $\lambda(\tau) = \lambda(0)$. This means that $\lambda(0) = \lambda(1)$. On the other hand, since $A(\tau + \frac{1}{2}) = A(\tau)^{\Omega}$, the $[\frac{1}{2}, 1]$ part of the flow of eigenvalues of $\not{D}$ is an exact replica of the $[0, \frac{1}{2}]$ part. These facts imply that, in the $A \to A^{\Omega}$ process, eigenvalues can only mix in *pairs*. No more complicated rearrangements are allowed.

Consider now for simplicity the case where only one eigenvalue crosses zero in the interval $[0, \frac{1}{2}]$. Correspondingly, there is only one eigenvalue (the opposite one) crossing in the opposite direction. Consistent with the previous remarks, and the symmetry between positive and negative eigenvalues of $\not{D}$, the only possibility is that λ is mapped into $-\lambda$ at $\tau = \frac{1}{2}$. Thus, if we started by choosing to include λ, and not $-\lambda$, in the definition of the square root of $\det \not{D}$, continuously following this choice leads us to include $-\lambda$, rather than λ, at $\tau = \frac{1}{2}$. This shows that there is no choice of λ's that remains invariant, like the one in fig. 2. In the more general case of an odd number of eigenvalue crossings the reasoning is similar, and is based on the fact that the set of eigenvalues that cross downwards is mapped, at $\tau = \frac{1}{2}$, into the set of their negatives, and so at least one eigenvalue will be mapped onto its negative, rendering again a consistent choice impossible. The proof can easily be extended to a general finite homotopy group Z_n.

We can also regard the nonexistence of a gauge invariant choice as a manifestation of the fact that different choices correspond to different regularizations and the corresponding actions should be connected with local counterterms. If a gauge invariant choice existed, there should be a local counterterm that would connect it with the gauge noninvariant action, i.e., the one changing by π under the nontrivial gauge transformation. Such a term would essentially "count" the Z_2 winding number of the gauge transformation around spacetime. However, such a local form does not exist, since Z_2 is pure tortion. On the contrary, in the three-dimensional case the homotopy group is Z and such a term does exist, namely the Chern-Simons term.

V. Generalizations to arbitrary odd dimensions and conclusions.

The above arguments for the quantization of the topological term and the parity violation can be generalized to $2n+1$ dimensions, compactified into $S^2 \times ... \times S^2 \times S^1$. There again we can construct the transition functions of the Chern-Simons term AF^n in terms of the transition functions of A, in the fashion

$$J^{2n}_{\alpha_1\alpha_2} = \Psi_{\alpha_1\alpha_2} F^n$$

$$J^{2n-1}_{\alpha_1\alpha_2\alpha_3} = c_{<\alpha_1\alpha_2\alpha_3} A_{\alpha_3>} F^{n-1}$$

$$. \quad . \quad . \quad . \quad . \quad . \quad .$$

$$J^0_{\alpha_1 ... \alpha_{2n+2}} = c_{<\alpha_1\alpha_2\alpha_3} \cdots c_{\alpha_{2n-1}\alpha_{2n}\alpha_{2n+1}} \Psi_{\alpha_{2n+1}\alpha_{2n+2}>} \tag{5.1}$$

Performing a nontrivial gauge transformation of winding number N in the S^1 direction, only the transition functions containing A explicitly will change, in total $n+1$ terms, each one contributing a change in the lagrangian

$$n! \cdot 2\pi N \cdot 2\pi\Phi_1 \cdots 2\pi\Phi_n \tag{5.2}$$

some of the $2\pi\Phi$ terms coming from integrals of F's and some from sums of c's, and $n!$ coming from combinatorics of c's and powers of F's, like the factor 2 in (2.15). So, for the exponentiated action to be well-defined, we must have

$$(n+1)!(2\pi)^n \kappa = \text{integer}\,. \tag{5.3}$$

Similar arguments hold for the proof of parity violation, where, now, the Atiyah-Singer index theorem in $2n+2$ dimensions will be relevant.

In conclusion, we saw that, even in the abelian case, one can obtain a quantization condition (although its interpretation is different than in the nonabelian case) and give a topological argument demonstrating the necessity of parity violation. Note, however, that our results hold for *one* fermion flavor only. In the case of N flavors, the previous results may not hold, while there is the extra possibility of a breakdown of the global $SU(N)$ flavor symmetry. This case is examined in chapter 6.

REFERENCES

[0] A. P. Polychronakos, Nucl. Phys. **B281** (1987) 241.

[1] S. Deser, R. Jackiw and S. Templeton, Phys. Rev. Lett. **48** (1982) 975 and Ann. Phys. (N.Y.) **140** (1982) 372.

[2] A. N. Redlich, Phys. Rev. Lett. **52** (1984) and Phys. Rev. **D29** (1984) 2366.

[3] A. J. Niemi and G. W. Semenoff, Phys. Rev. Lett. **51** (1983) 2077.

[4] M. B. Paranjape, Phys. Rev. Lett. **55** (1985) 2390.

[5] A. P. Polychronakos, Nucl. Pys. **B278** (1986) 207.

[6] O. Alvarez, Comm. Math. Phys. **100** (1985) 279.

[7] M. Henneaux and C. Teitelboim, Phys. Rev. Lett. **56** (1986) 689.

[8] E. Witten, Phys. Lett. **117B** (1982) 324.

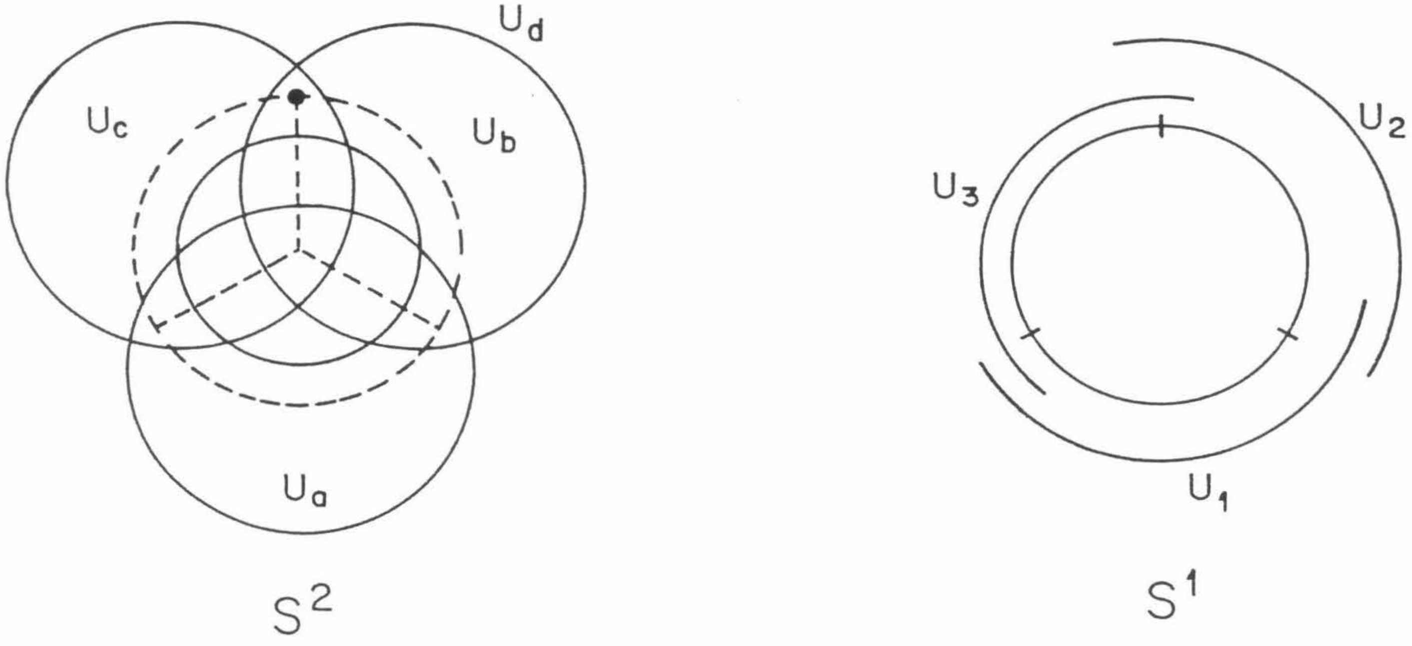

Fig. 1

Fig. 1: A possible patch covering and partition of $S^2 \times S^1$. For S^2, four patches are used, drawn with thick lines (patch U_d is the region outside the central thick circle). The partition is drawn with broken lines. The transition functions can be chosen to be: $\Psi_{ab} = \Psi_{bc} = \Psi_{ca} = 0$, $\Psi_{ad}, \Psi_{bd}, \Psi_{cd} = \Phi \, d\phi$, $c_{bcd} = 2\pi\Phi$, ϕ being an angle measured along the dotted circle, with $\phi = 0$ at the dotted point in U_{bcd}. For S^1, three patches are used (the partition is denoted with marks) and all transition functions can be chosen to vanish. The overall patching is then defined: $U_\alpha = U_A \times U_i$, with $\alpha = (A, i)$, $A = a, b, c, d$, $i = 1, 2, 3$, and $\Psi_{\alpha\beta} = \Psi_{AB}$, provided U_α and U_β overlap. Similarly, the overall partition is the cartesian product of the partitions of S^2 and S^1.

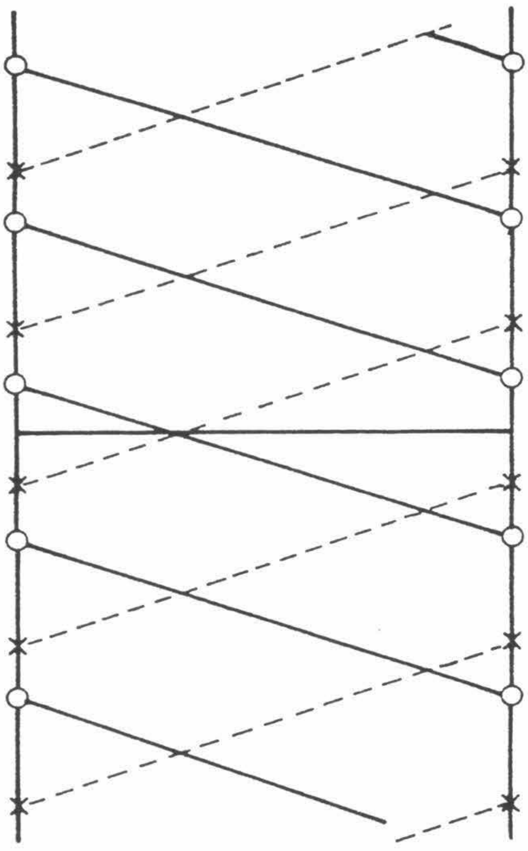

Fig. 2

Fig. 2: A possible spectral flow for $\not{D}$ with one level crossing (see appendix). Dots correspond to $\Gamma = +1$ and crosses to $\Gamma = -1$. A manifestly gauge invariant definition of the square root of $\det\not{D}$ is as the product of dots. In fact, the spectral flow depicted in this figure is the flow of the eigenvalues (3.20) for the gauge configuration $A_3(\tau) = A_3(0) + \frac{2\pi\tau}{T}$, with $\tau \in [0,1]$ and $\Phi = 1$.

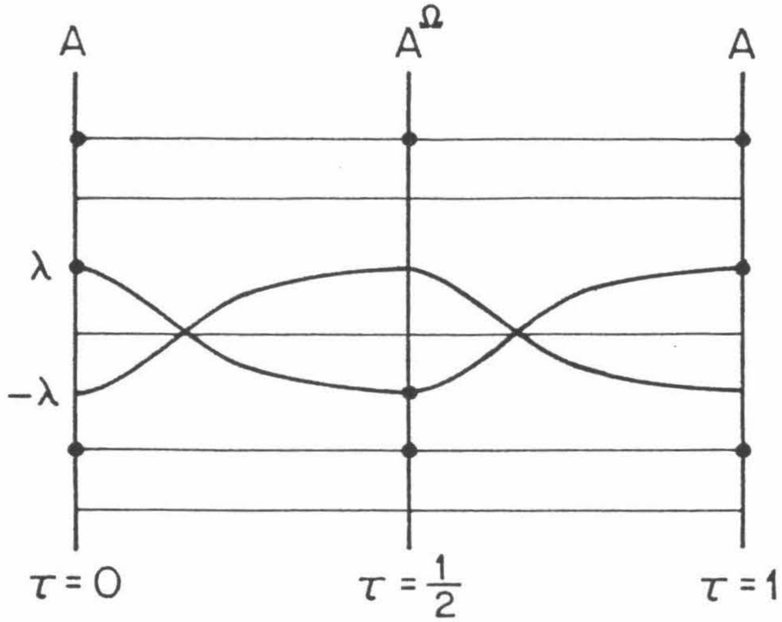

Fig. 3

Fig. 3: The construction of $A(\tau)$ for $SU(2)$. One level crossing is shown.

CHAPTER 5

Induced Angular Momentum and the Topology of the Chern-Simons Form

I. Introduction.

As has already been described in the previous chapter, in odd dimensional gauge theories we can write down an unusual, and yet gauge invariant term in the action, namely the Chern-Simons form. In 2+1 dimensions, in particular, this term is acting as a mass term for the gauge bosons. For the theory to be well-defined in the presence of this term, then, its coefficient has to be quantized in the nonabelian case [1]. In the abelian case, on the other hand, the gauge field configuration space decomposes into topologically disjoint sectors, classified by the value of the total magnetic flux. If one wants the sectors with nonzero flux to contribute to the path integral, then the coefficient of the Chern-Simons term has to be quantized in this case too [2,3]. In the opposite case of an arbitrary coefficient, equal to an irrational multiple of the quantization unit, the theory only contains states of zero total magnetic flux and quantization in a sector of nonzero flux alone is inconsistent.

The difference between nonabelian and abelian theories can be demonstrated in terms of different compactifications of spacetime: one could compactify it into either S^3 or $S^2 \times S^1$ (the case $S^1 \times S^1 \times S^1$ is essentially the same as $S^2 \times S^1$). Then, in the nonabelian case, the gauge field configuration space contains always only one connected component. Since, moreover, $\pi_1(\mathrm{SU}(n)) = \pi_2(\mathrm{SU}(n)) = 0$, nontrivial gauge transformations are classified by $\pi_3(\mathrm{SU}(n)) = Z$ for both compactifications, that leads to the quantization condition. In the abelian case, though, S^3 compactification restricts us to the zero-flux sector, while $S^2 \times S^1$ allows for all possible values of the flux. Moreover, since only $\pi_1(\mathrm{U}(1)) = Z$ is nontrivial, there are *no* nontrivial gauge transformations in the S^3 case, while there exist such transformations in the

$S^2 \times S^1$ case. The path integral, then, in a sector of nonzero flux is nonzero only if the coefficient of the Chern-Simons term obeys a quantization condition.

In the case of the abelian theory, in particular, in addition to making gauge bosons massive, the Chern-Simons term gives to the states of the theory unusual quantum numbers. Specifically, it gives to "flux tubes" (states of nonzero localized magnetic flux) nontrivial fermion charge, angular momentum and statistics [4-9]. It is quite straightforward to see that magnetic flux implies also electric charge in this theory, either through the equations of motion, or through the definition of the expectation value of charge in the quantum theory. It is nontrivial, though, to see and calculate correctly the induced angular momentum and statistics, and thus deduce that the charge is actually fermionic. This question is dealt with in this chapter.

II. Calculation of the angular momentum.

For abelian gauge fields, the Chern-Simons term has the form

$$S = -\frac{n}{4\pi}\Omega = -\frac{n}{4\pi}\int AdA = -\frac{n}{4\pi}\int \epsilon^{\mu\nu\rho}A_\mu \partial_\nu A_\rho d^3x \tag{2.1}$$

where n has to be quantized to an integer for e^{iS} to be globally gauge invariant in nonzero flux sectors. Then the charged current of the theory due to this term is

$$j^\mu = -\frac{\delta S}{\delta A_\mu} = \frac{n}{2\pi}\epsilon^{\mu\nu\rho}\partial_\nu A_\rho. \tag{2.2}$$

A flux tube is a localized magnetic gauge field configuration with total flux

$$\Phi = \frac{1}{2\pi}\int B = \frac{1}{2\pi}\int dA. \tag{2.3}$$

If the space is assumed to be compact, then the total magnetic flux of the space is quantized to an integer (the "monopole number" enclosed by the space). From (2.2) we see that a flux tube with one quantum of flux carries n units of charge. Moreover, as we will show, the tube carries angular momentum equal to

$$J = n \cdot \tfrac{1}{2}\Phi^2. \tag{2.4}$$

So, because of the coefficient $\frac{1}{2}$ in (4) (and also from the fact that the Chern-Simons form is induced when the gauge field is coupled to fermions), we can interpret the current (2.2) as a fermionic current.

There are several ways to understand physically the origin of the angular momentum of the flux tube. One way to think of it is as due to the interaction of the magnetic field of the tube with the electric field produced by the fermion charge of the tube. For a rotationally symmetric tube this (radial) electric field is

$$E(r) = \frac{Q(r)}{r} = n\frac{\Phi(r)}{r} \tag{2.5}$$

(r, ϕ are polar coordinates and $Q(r)$, $\Phi(r)$ are the charge and flux inside a circle of radius r). Then, an easy calculation of the electromagnetic angular momentum $J = \int \vec{r} \times (\vec{E} \times \vec{B}) = \int B\vec{r} \cdot \vec{E}$ yields the result (2.5).

Another way to understand it is as due to the Aharonov-Bohm phase picked up by the wavefunction of the tube after rotating it by 2π, because of the rotation of the charge of the tube around its magnetic flux. Since an infinitesimal element of the tube $d\Phi_1 = B(x)d^2x$ has charge equal to $dQ_1 = nd\Phi_1$, after it gets transported around another infinitesimal element $d\Phi_2$ picks up an Aharonov-Bohm phase equal to $dQ_1 \cdot 2\pi d\Phi_2 = 2\pi \cdot nd\Phi_1 \cdot d\Phi_2$. So the total phase picked up is

$$S = 2\pi n \cdot \frac{1}{2} \iint d\Phi_1 \cdot d\Phi_2 = n \cdot 2\pi \cdot \tfrac{1}{2}\Phi^2 \tag{2.6}$$

(the coefficient $\frac{1}{2}$ is put in the integral because each pair of infinitesimal elements ($d\Phi_1$, $d\Phi_2$) is counted twice). Equating this with $2\pi J$, we again obtain the standard result.

Finally, we can attribute the angular momentum of the flux to the fermionic degrees of freedom it carries. Let us consider the case $n = 1$ first. Since a tube with Φ units of flux carries also Φ units of fermion charge, we can think of putting these fermions in successive angular momentum eigenstates, with eigenvalues $\frac{1}{2}$, $\frac{3}{2}$, ..., $\frac{2\Phi-1}{2}$ (the exclusion principle forbids us to put them in the same state). Then the total angular momentum is the sum of all the individual eigenvalues, which equals $\frac{1}{2}\Phi^2$. (The Chern-Simons term being a pseudoscalar, it "polarizes" the spins of all the fermions in the same direction.) The fact, now, that n multiplies the *whole* angular momentum, and not Φ itself, means that we have to interpret it as a number of fermion flavors, since n fermions can be put in the same state. Indeed, in ref. [10] it was shown that if an S^2 nonlinear sigma model is coupled to fermions, then integrating out the fermions induces the Hopf term (which is similar in form with the Chern-Simons term), with coefficient $n = 1$. Obviously, coupling it with n flavors of fermions gives the same term with a coefficient n. The case $n < 0$ corresponds to opposite sign of the coupling term and so to opposite polarization of the (anti)fermions around the tube.

The situation is slightly more complicated when the gauge field itself is coupled to fermions [11,12,13]. Then, a Chern-Simons term is again induced, but with a coefficient quantized to *half* the unit, in order to cancel the global gauge anomalies of the fermionic determinant [2,4]. In the case of massless fermions, this anomaly can be understood as due to the existence of zero-modes of the fermionic hamiltonian in a nonzero flux background. These states contribute a fermion number equal to a half each, this being the case since the hamiltonian has a symmetric spectrum. This accounts for the half-integral n and thus the half-integral fermion charge of flux tubes. The angular momentum, in this case, can be attributed partly to the fermions and partly to the gauge field itself [12].

The previous arguments, however, can be at most heuristic. To really see whether a flux tube acquires any nontrivial statistics and angular momentum, we should adiabatically rotate it through 2π, calculate the action S_{rot} associated with this rotation and equate the phase $e^{iS_{rot}}$ that the state picks up after this rotation with $e^{i2\pi J}$, where J is the angular momentum of the tube. Since S is the only term in the action that is first-order in time derivatives, it is the only one that could contribute to J. However, a naive calculation of S immediately leads to trouble, since it gives a vanishing result, in contradiction to the expectation that the tube should possess the angular momentum of the fermion number it carries, which is in general nonzero. Indeed, if we write the gauge field in the gauge

$$A_0 = 0 \ , \quad A_i = \epsilon_{ij}\partial_j a, \tag{2.7}$$

where a is a scalar field, and consider gauge field configurations that become pure gauge at spatial infinity, for which $a \to \Phi \ln r$ for $r \to \infty$, we get

$$S \sim \iint \epsilon_{ij}\partial_i a \cdot \partial_j \dot{a}\, d^2x dt = \oint \partial_i a \cdot \dot{a} dx^i dt = 0, \tag{2.8}$$

leading to believe that there are no nontrivial spin and statistics. (A similar calculation in ref. [8] that gave a nonzero result is incorrect in that the discontinuities of the angular variables used in the integrand were handled improperly.) The problem can be made more explicit if we consider a rotationally symmetric tube with gauge field

$$A_0 = 0 \ , \quad \vec{A} = \hat{e}_\phi \frac{\Phi(r)}{r} . \tag{2.9}$$

Then, a rotation of this tube leaves the gauge field invariant, and thus S trivially vanishes.

This difficulty can be temporarily overcome if we interpret the rotation of the tube as a transformation of the gauge field at each point into its Lorenz boosted one, with a velocity ωr in the ϕ-direction (ω being the angular velocity of rotation). For small ω, $\vec{A}$ does not change, but there is a nonzero A_0 generated:

$$A_0 = \omega \Phi(r). \tag{2.10}$$

This ensures, for instance, that there is a ϕ-component of the current (2.2) generated, due to the rotation of the charge density j^0, equal to $\omega r j^0$. However, it is easy to see that, even for this new field, the integrand in (2.1) vanishes. In fact, since the Chern-Simons term is a differential form, it is invariant under a general (local) coordinate transformation, and the previous boost is just such a transformation.

We are going to explain where the previous calculations fail in a minute. For the moment, let us circumvent the difficulty by performing a gauge transformation of the form

$$U = e^{i\theta(t)\Phi} \quad \Rightarrow \quad A_0 \rightarrow \omega\,(\Phi(r) - \Phi)\ ,\ \ \dot{\theta} = \omega\,. \tag{2.11}$$

In this new gauge, the 0-component of the gauge field far away from the tube vanishes. This choice, although gauge equivalent with the previous one, seems more palatable, since we avoid possible extra complications due to the interaction of the field A_0 with other fluxes that our space may contain. Calculating now S again, we get an additional contribution due to the extra term in A_0, equal to

$$S = -\frac{n}{4\pi}\int_0^{\mathrm{T}} 2\pi\Phi \cdot (-\omega\Phi)\cdot dt = n\pi\Phi^2 \tag{2.12}$$

and, equating this with $2\pi J$, we obtain the correct result (2.4). This should teach us that naive calculations may give misleading answers and that something is going on that we are missing. Indeed, gauge invariance implies that the two gauge-equivalent choices of the previous calculation should give equivalent results, which does not seem to be the case.

The reason why the previous calculations fail is that, as explained in the previous chapter, the gauge field for a nonzero flux configuration cannot be globally defined over a compact space, and thus the Chern-Simons term can only be written as a sum over patches. For this sum to be independent of the patching, appropriate correction terms have to be included [14]. Following the construction of the previous chapter, if we know the transition functions between patches for the gauge potential A

$$A_\alpha - A_\beta = d\psi_{\alpha\beta}$$

$$\psi_{\alpha\beta} + \psi_{\beta\gamma} + \psi_{\gamma\alpha} = c_{\alpha\beta\gamma} \tag{2.13a}$$

(where greek indices enumerate the patches), with c's satisfying the relations

$$c_{\alpha\beta\gamma} - c_{\beta\gamma\delta} + c_{\gamma\delta\alpha} - c_{\delta\alpha\beta} = 0, \tag{2.13b}$$

then the full expression for Ω, including the correction terms, is, in differential form notation

$$\Omega = \sum_{\alpha} \int A_\alpha dA + \sum_{\alpha\beta} \int \psi_{\alpha\beta} dA - \sum_{\alpha\beta\gamma} \int c_{<\alpha\beta\gamma} A_{\gamma>} + \sum_{\alpha\beta\gamma\delta} c_{<\alpha\beta\gamma}\psi_{\gamma\delta>} \tag{2.14}$$

where the integrals are over the boundaries of corresponding dimensionality lying in the intersection of the patches denoted by the indices (the field strength dA is, of course, globally well-defined). The symbol $< \cdots >$ is the same as defined in the previous chapter and means: put the indices in increasing order, with indices in the positions of the initial repeated indices matching, and multiply by the parity of the permutation. For example,

$$c_{<412}\psi_{23>} = -c_{123}\psi_{34}\,. \tag{2.15}$$

Considering our space to be compactified into a sphere S^2, a possible partition of the space and choice of the gauge field and transition functions around a flux tube with flux Φ is shown in fig. 1.

Now we can calculate S again, using the gauge field (2.11). Since at times 0 and T the configurations are identical, we consider our time to be periodic. For the field (2.11), the first term in (2.14) gives $-2\pi\Phi \cdot 2\pi\Phi$. The second term gives zero, since dA has no component in the radial direction. The third term gives zero too, since $A_0 = 0$ outside of the flux tube. The fourth term does not contribute, since there are no nontrivial 0-dimensional boundaries in this special case. So, we get the previous result. This explains why the naive calculation for the field (2.11) was correct: All the correction terms happened to vanish.

This is not so, however, for the field (2.10). There, the second and fourth terms also vanish, for the same reasons as before. The first term gives zero too, which is the result of the naive calculation. The third term, however, is now nonzero, since $A_0 = \omega\Phi$ at the dotted point in fig. 1, and contributes $2\pi\Phi \cdot 2\pi\Phi$. This is the opposite result than before. However, remember that, in nonzero total flux backgrounds, the Chern-Simons term is gauge invariant only up to a multiple of 2π. Thus, if Φ is the total flux of the space, it must be an integer and thus the two results for S differ only by an irrelevant multiple of 2π, (given that $n \cdot \Phi$ is integer, else no quantum states with flux Φ exist at all). To unambiguously decide which is the contribution of the flux tube alone to the angular momentum, we observe that the calculation for (2.11) holds even if the space contains other fluxes. On the other hand, if the flux of the rest of the space is Φ_{rest} (and so $\Phi_{total} = \Phi + \Phi_{rest}$), the calculation based on (2.10) gives

$$\Omega = 2\pi\Phi \cdot 2\pi\Phi + 2 \cdot 2\pi\Phi \cdot 2\pi\Phi_{rest} = -2\pi\Phi \cdot 2\pi\Phi + 2 \cdot 2\pi\Phi \cdot 2\pi\Phi_{total} \, . \qquad (2.16)$$

(The extra factor of 2 in the term involving Φ_{rest} is due to the interplay between A_0 and the correction terms around the rest of the fluxes in the space. The details are explained in the previous chapter.) The term involving Φ_{total} comes from the gauge transformation connecting (2.10) and (2.11), and so the contribution of Φ alone is the same as before.

The previous facts can be clarified if we work in the zero-flux sector of the theory (which exists for arbitrary nonquantized n), by putting $\Phi_{rest} = -\Phi \Rightarrow \Phi_{total} = 0$. In this case no correction terms are necessary (although they can still be introduced, if we want the gauge field in regions outside of flux tubes to be the same as in the absense of the tubes). Then, either of the previous gauges gives the same result, as long as we rotate *only* Φ (and not Φ_{rest}, in which case the total action vanishes).

III. The topological construction.

The question arises: Could we have seen the nontrivial phase (2.12) without the somewhat *ad hoc* interpretation of the rotation as an appropriate Lorenz boost of the gauge field at each point? The answer is yes, but in an intricate way.

Let us imagine that we rotate our flux tube by rigidly rotating the *whole* patching of fig. 1 through 2π. This means that at each (rotated) position of the patches the transition functions are the same as the original ones, but rotated by the same angle, and that the dotted point in $U_{\alpha\gamma\delta}$ winds around the world-cylinder transcribed by the circle once and comes back to its original position (fig. 2). So, for an intermediate position, the transition functions are

$$\psi_{\alpha\beta} = \psi_{\beta\gamma} = \psi_{\gamma\alpha} = 0 \ , \quad \psi_{\alpha\delta} = \psi_{\beta\delta} = \psi_{\gamma\delta} = \Phi \cdot (\phi - \theta) \tag{3.2}$$

where θ is the angle that the patching has been rotated. From this we conclude that the 0-components of the gauge field inside and outside of the circle, previously both zero, now have to satisfy

$$A_{0,\mu} - A_{0,\delta} = \frac{\partial}{\partial t}\psi_{\mu\delta} = -\Phi\dot{\theta} = -\Phi\omega, \tag{3.2}$$

where μ stands for α, β, γ. Thus, if $A_{0,\delta} = 0$, we conclude that there must be a nonzero A_0 inside the circle, equal to $-\Phi\dot{\theta}$. (The choice $A_{0,\delta} = \Phi\theta dot$, $A_{0,\mu} = 0$, would produce extra contributions due to the interactions of the nonzero $A_{0,\delta}$ with other fluxes that may exist in the space. At any rate, the two choices are connected with a timelike gauge transformation, and if the coefficient n and the total flux of the space are properly quantized, they will give phases differing by an integer multiple of 2π).

It is easy now to see that, due to this new value of $A_{0,\mu}$, the term $A_\alpha dA$ in (2.14) produces an extra contribution equal to

$$S = -\frac{n}{4\pi}\int_0^T (-\Phi\dot{\theta}) \cdot 2\pi\Phi dt = n\pi\Phi^2 \tag{3.3}$$

while the correction terms do not contribute, since they all contain $A_{0,\delta}$. So we again recover the same nontrivial phase factor that gives to the tube nonzero angular momentum. Note that, even if we had chosen a different ordering of the patches, say $\delta < \alpha < \beta < \gamma$, we would still get the same result. In fact, the term $c_{\delta\alpha\gamma}A_\gamma$ would now appear, due to the definition of the $< \cdots >$ symbol, that would give one contribution $2\pi\Phi \cdot (-\Phi\dot{\theta})dt$, due to $A_{0,\mu}$ and one $2\pi\Phi \cdot \frac{\Phi}{r} \cdot \dot{\theta}rdt$, due to $A_{\phi,\mu}$, that cancel each other.

Thus we see that the patching creates a sort of "frame of reference" for the flux tube, whose rotation produces the nontrivial phase, if we correctly account for the correction terms in (2.14). This patching, twisting by 2π and coming back to itself, constitutes a nontrivial mapping of $S^2 \times S^1$ (our spacetime) into S^2 (the patching, or space itself), that is classified by $\pi_3(S^2) = Z$. For such a mapping with winding number N, we pick up a phase equal to $2\pi N \cdot n \cdot \frac{1}{2}\Phi^2$. Note, also, that the previous (topological) derivation holds for arbitrary shape tubes, not just rotationally symmetric ones.

It is easy to see that the nontrivial spin of flux configurations implies also nontrivial statistics. Since the total angular momentum depends only on the total flux, the total phase picked up by a system of two well-separated tubes after a 2π rotation is $2\pi n \cdot \frac{1}{2}(\Phi_1 + \Phi_2)^2$. The parts proportional to $\frac{1}{2}\Phi_1^2$ and $\frac{1}{2}\Phi_2^2$ are due to the rotation of the tubes themselves. So, the remaining part $2\pi n \cdot \Phi_1\Phi_2$ is due to their interaction and, for a rotation of π (corresponding to exchanging the tubes), we get a phase $n\pi \cdot \Phi_1\Phi_2$. For odd fluxes and odd integer n, when both tubes have odd charges, the phase factor is just -1, corresponding to fermion statistics. For intermediate values of n, tubes obey fractional statistics.

It is interesting that a theory of intrinsically bosonic objects of solitonic nature (flux tubes) can describe half-integral spin and statistics. This is reminiscent of bosonization in two dimensions. One could conjecture that in three dimensions, too, there is a bosonization (or fermionization) procedure, that allows a description of the tubes in terms of intrinsically fermionic fields (spinors). If this description turned out to be local and renormalizable, it would be a remarkable extension of bosonization techniques to more than two dimensions. On general grounds, however, one generically does not expect this to happen. Such a construction, of course, has yet to be attempted.

IV. Conclusions.

We showed, using several methods: heuristic, analytical and topological, that the Chern-Simons term induces unusual quantum numbers on flux tubes. The interpretation of nontrivial spin and statistics given in the previous section, though, in terms of patchings, may seem somewhat unnatural. There are two ways to view things: In three dimensions (compactified into $S^2 \times S^1$), there are several inequivalent ways to choose the patchings of spacetime and the transition functions for Ω, classified by $\pi_3(S^2)$. If we impose that all these patchings give the same exponentiated action, then n has to be even and flux tubes are bosons. We can, however, interpret the different patchings as corresponding to rotations of the tube by a multiple of 2π indicated by the element of $\pi_3(S^2)$. This, for arbitrary n, gives to the tubes fractional statistics. In order, however, that states with nonzero total flux exist in the quantum theory, n has to be an integer. Thus, we see that, in nonzero flux states, flux tubes with integer flux can have only ordinary (fermion or boson) statistics.

REFERENCES

[0] A. P. Polychronakos, CALT-68-1397.

[1] S. Deser, R. Jackiw and S. Templeton, Phys. Rev. Lett. 48 (1982) 975 and Ann. Phys. (N.Y.) **140** (1982) 372.

[2] A. P. Polychronakos, Nucl. Phys. **B281** (1987) 241.

[3] O. Alvarez, Comm. Math. Phys. **100** (1985) 279.

[4] A. N. Redlich, Phys. Rev. Lett. **52** (1984) 2366 and Phys. Rev. **D29** (1984) 2366.

[5] A. J. Niemi and G. W. Semenoff, Phys. Rev. Lett **51** (1983) 2077.

[6] R. Jackiw, Phys. Rev. **D29** (1984) 2375.

[7] F. Wilczek, Phys. Rev. Lett. **48** (1982) 1144; F. Wilczek and A. Zee, Phys. Rev. Lett. **51** (1983) 2250; Yong-Shi Wu and A. Zee, Phys. Lett. **147B** (1984) 325; M. Bowick, D. Karabali and L.C.R. Wijewardhana, Nucl. Phys. **B271** (1986) 417.

[8] D. Arovas, J.R. Schrieffer, F. Wilczek and A. Zee, Nucl. Phys. **B251** [FS13] (1985) 117; R. Hagen, Phys. Rev. **D31** (1985) 2135.

[9] G.W. Semenoff and L.C.R. Wijewardhana, E.T.H./Yale preprint.

[10] T. Jaroszewicz, Phys. Lett. **146B** (1984) 337 and Phys. Lett. **159B** (1985) 299.

[11] M. B. Paranjape, Phys. Rev. Lett. **55** (1985) 2390; L. S. Brown, Phys. Rev. Lett. **57** (1986) 499; M. B. Paranjape, Phys. Rev. Lett. **57** (1986) 500.

[12] A. P. Polychronakos, Nucl. Phys. **B278** (1986) 207.

[13] R. Blankenbecler and D. Boyanovsky, Phys. Rev. **D34** (1986) 612.

[14] For a review of the relevant cohomological notions in a physics context see reference [3].

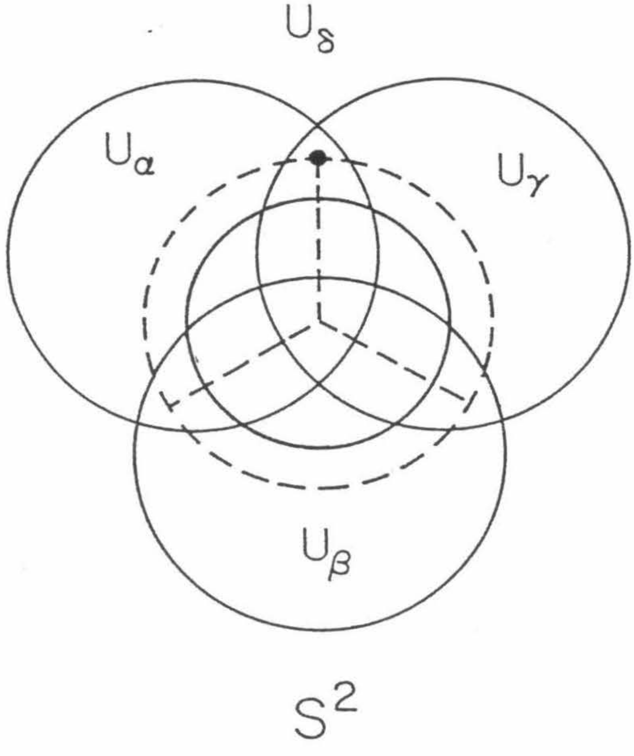

Fig. 1

Fig.1: A possible patch covering and partition of space around a flux tube of flux Φ. Four patches U_α are used, drawn with solid lines and put in the order $\alpha < \beta < \gamma < \delta$ (the patch U_δ is the region outside the central solid circle). The boundaries are denoted with broken lines. The gauge field is equal to the one in (9) in $U_{\alpha,\beta,\gamma}$ and zero in U_δ (all the flux is inside the circular boundary). The transition functions are $\psi_{\alpha\beta} = \psi_{\beta\gamma} = \psi_{\gamma\alpha} = 0$, $\psi_{\alpha\delta}$, $\psi_{\beta\delta}$, $\psi_{\gamma\delta} = \Phi\, d\phi$, $c_{\alpha\gamma\delta} = 2\pi\Phi$, where ϕ is the polar angle measured around the circular boundary, with $\phi = 0$ at the dotted point.

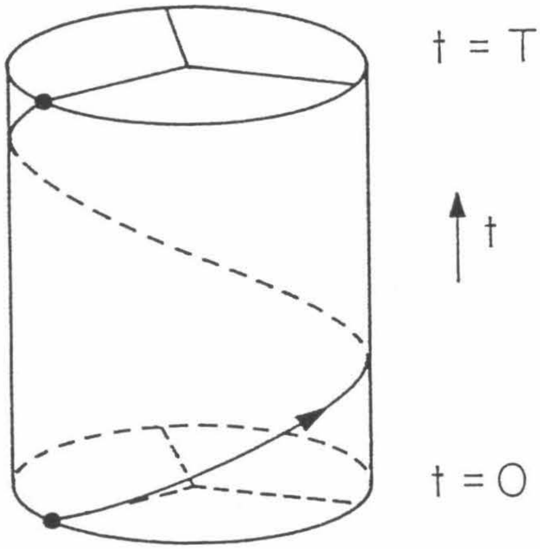

Fig. 2

Fig.2: The world cylinder of the circular boundary and the world line of the dotted point for the patching of fig.1 rotated in time through 2π. The $t = 0$ and $t = \mathrm{T}$ slices are identified. The whole process constitutes a nontrivial element of $\pi_3(S^2)$ with winding number one.

CHAPTER 6

Global Symmetry Breaking in Planar Systems and the Quantum Hall Effect

I. Introduction.

From the analysis and the results presented in the previous chapters, some general patterns in the behaviour and the properties of 2+1 dimensional gauge theories should have become obvious. To summarize, both nonabelian and abelian theories with one fermion flavor break parity [1,2], some quantization condition is either required or desired, respectively, for the coefficient of the Chern-Simons term [2,3], and, in the abelian case, the vacuum behaves like a superconductor and has a quantum Hall property with a conductance strictly independent of the electromagnetic field and equal to *half* the quantum of conductance in the condensed matter case [1,4]. Moreover, all of these phenomena seem to be intimately related with the topology of the configuration space of the gauge fields.

Although the derivation of the quantization condition and the demonstration of the parity breaking, as presented in chapter 4, seem to be already quite similar, their connection is still superficial, and the similarity with the quantum Hall effect (QHE) has only been hinted at, so far, with no justification or deeper connection between this and the previous phenomena. One would like, thus, to have a unifying picture of the situation, where everything lucidly follows from a generic feature of the theories under scrutiny, if only for one's own clarity of mind and understanding of the subject. This is attempted in this chapter. As it turns out, in addition to an improved intuition, we will obtain some extra and more general results, concerning the patterns of global symmetry breaking in multiflavor theories. On this subject, an amount of disagreement and confusion has arisen in the literature [5-7]. It is, thus, of importance to have a relatively lucid and solid argument on what the actual situation is. Moreover, we will establish a painless way to evaluate the anomalous *quantum* numbers of the vacuum due to parity breaking, as well as a generalization of the charge formula for nonzero temperature.

These results would be very important would they be extendable to the 3+1 dimensional case, since the question of how gauge theories realize their global symmetries still is one of the most interesting open ones in non-string physics. It seems, however, that the considerations exposed in the next sections are quite particular to the topology of three dimensional theories, and so, one should be more skeptical than panegyric about their usefulness for "real" physical theories.

The fact remains, at any rate, that the analysis to follow, apart from constituting a satisfying and, hopefully, elucidating wrapup of the work exposed so far, will also be relevant and useful to the study of planar condensed matter systems, and, the mathematical connection to be realized between such systems and the present theories will, at least, establish some common workground between the two fields, facilitating any further exchange of ideas or experience that would be profitable for either subject.

II. Symmetry breaking in multiflavor 2+1 QED.

Let us initially concentrate our attention to abelian theories in three dimensions, but with many fermion flavors. The general lagrangian for such a model is

$$L = \overline{\psi}_i \not{D} \psi_i + m_i \overline{\psi}_i \psi_i \ , \tag{2.1}$$

where $\not{D}$ is the usual covariant derivative containing the U(1) gauge field A and $i = 1, \ldots N$ is a flavor index, assumed always to be summed in the expression for L. Notice that L is a lagrangian of N two-component fermions interacting only through the gauge field A.

In addition to the gauge U(1) symmetry, for the special case $m_1 = \ldots = m_N$ the theory is invariant under a global SU(N) symmetry that mixes the fermion flavors.

On the other hand, as is well known, in odd spacetime dimensions a mass term is a pseudoscalar and, thus, explicitly breaks parity. Specifically, a parity transformation that reflects space with respect to x^1 is defined as

$$\psi_i \to \gamma^2 \psi_i \,, \; x^2 \to -x^2 \,, \; A_2 \to -A_2 \,, \; m_i \to -m_i \,. \tag{2.2}$$

Under (2.2) the action remains invariant. Due to the nontrivial transformation of the mass, though, parity is an exact symmetry classically only in the limit $m_i \to 0$.

If, however, N is even and, moreover, the relations $m_1 = -m_2$, $m_3 = -m_4$, ... $m_{N-1} = -m_N$ hold (where, of course, the ordering of m_i is immaterial), we can modify our definition of parity transformation such that, in addition to flipping one space dimension, it also interchanges the $2j-1$ and $2j$ flavors. This modified parity, now, is obviously a symmetry of the lagrangian. Physically, what happens is that the reflection of a state into a mirror corresponds to a state with the names of the flavors interchanged in pairs. Since these names were arbitrary to begin with, the reflected state is a physical state just as well.

We see, thus, that flavor SU(N) and parity become good symmetries of the classical theory only at the limit of all masses going to zero. The question addressed here is: Do these symmetries survive in the quantum theory, and, if not, how do they break?

In order to answer this question, we will confine our attention to external (classical) field configurations in the $A_0 = 0$ gauge that are independent of time, that is, to purely magnetic configurations. The space will again be assumed to be compactified into a compact boundaryless manifold, in order to have a well-defined discrete spectrum of the Dirac hamiltonian (this is purely a matter of convenience). Moreover, we shall consider not parity itself, but rather CT. That is because, since magnetic fields are odd under both C and T, this transformation (from now on called M, since it also flips the sign of mass terms) leaves magnetic configurations invariant. From the CPT theorem we know that PM cannot break. Thus, if one breaks so does the other, and if one is unbroken the other is too.

Let us now consider the properties of the fermionic Fock vacuum $|A(\vec{x})>$ for vanishing masses m_i under this transformation. If this vacuum state were nondegenerate, then, due to the fact that the charge operator Q is odd under M, we would have

$$M|A(\vec{x})>=|MA(\vec{x})>=|A(\vec{x})> \text{ , and so}$$

$$\langle Q\rangle = -\langle MQM\rangle = -\langle Q\rangle = 0\,. \tag{2.3}$$

So, any nonzero expectation value of Q indicates not only that the vacuum is degenerate, but also that M maps a specific vacuum state into a *different* one, else the previous argument would again be applicable. Since we know, from the previous chapters, that $\langle Q\rangle$ is, indeed, nonzero, we understand that, indeed, we should look for such a situation.

Notice, now, that the Dirac hamiltonian for each flavor is identical with a two dimensional euclidean Dirac operator. The Atiyah-Singer index theorem immediately tells us that, if the total flux of the space Φ is nonzero (and equal to an integer, since the space is compact), then such an operator has $|\Phi|$ in total zero modes. These zero modes can be either filled or emptied, since that cannot change the energy of the vacuum. So, we see that the vacuum is, indeed, degenerate, with degeneracy $2^{|\Phi|N}$. Actually, we will mostly consider vacua where the $|\Phi|$ modes of each flavor are either all filled or all empty, for reasons having to do with how these vacua can be achieved as limits of massive theories, that will be explained later. There are only 2^N such vacua. In fig. 1, the case $\Phi=1$, $N=2$ has been drawn schematically.

Let us now realize the action of the flavor SU(N) and parity transformations on the vacuum in terms of the spectrum of the Dirac hamiltonian. The prescriptions are: For SU(N), simply substitute the states of the theory at some level E (common for all flavors, due to the classical SU(N) symmetry of the hamiltonian) with an SU(N) linear combination of them. In the special SU(2) case of fig. 1a, and for the special element $i\sigma_1$ of SU(2), this is just an exchange of the two states at energy E (times a multiplication by i, which can be absorbed into a global gauge transformation). For

M, on the other hand, first map the whole spectrum into its negative one (reflect it with respect to 0), then substitute all filled states with empty ones, and *vice versa*, and, finally, exchange the states of the $2i-1$ fermion with the ones of the $2i$ fermion. This can be shown to produce the correctly transformed Fock state, the anticommuting nature of the fermion field operators taken into account. The fact that the spectrum is, actually, invariant under the first operation (reflection) is a corollary of the statement that the classical theory has M as an exact symmetry.

We are now ready to see whether any of the possible $2^{|\Phi|N}$ (or 2^N) vacua is a singlet under both of these transformations. In order for SU(N) to be a symmetry of the vacuum, either *all* zero modes should be filled, or *all* should be empty (fig. 1a). Indeed, in the opposite case an SU(N) transformation would mix empty with filled states, thus giving a vacuum state different than any of the original ones. In order, on the other hand, for a vacuum state to remain invariant under M, in each $\{2i-1, 2i\}$ pair of groups of zero modes, one group should be filled (say, the $|\Phi|$ modes of $2i-1$ flavor) and one should be empty (say, the $|\Phi|$ modes of $2i$ flavor, or *vice versa*). Indeed, the first two steps involved in an M transformation will reverse the occupation number of these states, while the final exchange of the two flavors will restore the original state (fig. 1b).

It should be obvious, thus, that no vacuum state is a singlet under both transformations (with the exception of the trivial $\Phi = 0$ nondegenerate state). In fact, we can see that, for N even, some choices of vacuum conserve M ($N!/(N/2)!^2$ in number) but break SU(N), some choices conserve SU(N) (two in number) but break M, and the remaining choices break both symmetries. So, no vacuum state conserves both symmetries, which means that the effective action of the gauge bosons after integrating out the fermions will be noninvariant under either or both of M and SU(N). For N odd, there are two choices of vacuum that conserve SU(N) but there is *no* choice of vacuum preserving M (fig. 2). Thus, flavor symmetry may or may not break, but parity *must* break.

It is easy to see what limits of massive theories would lead to the previous parity preserving or flavor symmetry preserving massless theories. If *all* the fermion masses are drawn to zero from positive values, or *all* from negative values, then either all of the zero modes are shifted above zero, and so remain empty in the massless limit, or all are shifted below zero and remain full, thus leading to the two SU(N) preserving vacuum states. If, again, for even N, the masses go to zero in positive-negative pairs, the zero modes will be empty-filled in pairs and we will end up with one of the parity preserving states.

One may wonder why, in the previous analysis, we only considered the $2^{|\Phi|N}$ vacuum states and not any linear combination of them. The reason is that such a combination would not be gauge invariant and would lead to a nonunitary effective action for A. To see that, remember that j^0 is the generator of gauge transformations and so Q is the generator of global gauge transformations. As explained in the previous chapters, a state with symmetric energy spectrum and some zero modes is an eigenstate of Q with eigenvalue equal to half the number of the filled zero modes. Thus, the previously considered vacua are eigenstates of Q with (half-) integer eigenvalues. Under the global gauge transformation $e^{i\alpha}$, each state transforms as

$$|Q> \rightarrow \; e^{iQ\alpha}|Q> \; . \tag{2.4}$$

Apparently, these states are not gauge invariant. One may correct that, however, by shifting the gauge transformation generator by a gauge field dependent constant, equal to $-\langle Q \rangle$, thus leading to invariant states. This is readily and easily accompliced at the level of the gauge field effective action by adding the Chern-Simons term, with coefficient $-\frac{1}{2}$ for each fermion flavor whose zero modes are taken to be filled and $+\frac{1}{2}$ for each flavor whose zero modes are taken to be empty, that is, with a total coefficient

$$n = +\tfrac{1}{2} \sum_{i=1}^{N} sign(m_i) \, , \tag{2.5}$$

$sign(m_i)$ being the sign of the mass of the i^{th} flavor as it is drawn to zero. So we see that we are naturally led to the fact that a gauge invariantly defined effective action includes a Chern-Simons term. Since this term is a pseudoscalar, it breaks parity. Only if n is zero the effective action is parity invariant, that happens only if $sign(m_i)$ are opposite in pairs, as found before.

A state, now, which is not a charge eigenstate but a linear combination of eigenstates with different eigenvalues, is hopelessly gauge noninvariant. Indeed, a global gauge transformation shifts the *relative* phase between the different charge eigenstates, that cannot be compensated with a redefinition of the gauge generator. Moreover, it leads to a nonunitary effective action. Consider, for example the state

$$|vac>= \frac{1}{\sqrt{2}}|1> + \frac{1}{\sqrt{2}}|2> \tag{2.6}$$

with $|1>$ and $|2>$ being the two SU(N) conserving vacuum states of the theory. The resulting state is a singlet under *both* SU(N) and M, since M interchanges $|1>$ and $|2>$. The effective action obtained from this vacuum state, however, has the unfortunate form

$$e^{iS_{eff}} = \frac{1}{\sqrt{2}}e^{iS_1} + \frac{1}{\sqrt{2}}e^{iS_2} \tag{2.7}$$

where S_1 and S_2 are the effective actions resulting from the two component states alone. Thus it is obvious that, not only is S_{eff} not gauge invariant, but also it is nonreal, and thus the resulting quantum theory is nonunitary.

From the previous considerations, we can find in what specific patterns flavor SU(N) will break. Denote with N_- the number of flavors with filled zero modes (i.e., the number of negative $sign(m_i)$), and with $N_+ = N - N_-$ the number of flavors with empty zero modes (the number of positive $sign(m_i)$). Obviously, an SU(N) transformation that only mixes levels with the same occupancy is still a symmetry. Moreover, we can still perform phase transformations to all flavors. So, the breaking pattern is

$$\mathrm{SU}(N) \ \rightarrow \ \mathrm{SU}(N_-) \otimes \mathrm{SU}(N_+) \otimes \mathrm{U}(1)_A \ . \tag{2.8}$$

where $\mathrm{U}(1)_A$ is an "axial" global transformation that rotates the phases of the N_+ and of the N_- fermions by opposite amounts. From (2.8) it follows that there are

$2N_+N_-$ broken generators, and thus $2N_+N_-$ Goldstone bosons. In the particular case where parity is preserved, we have

$$\mathrm{SU}(2N) \;\rightarrow\; \mathrm{SU}(N) \otimes \mathrm{SU}(N) \otimes \mathrm{U}(1)_A \,. \tag{2.9}$$

Of all the patterns with constant $N_+ + N_- = 2N$, this is the one with the largest $2N_+N_-$. So, when parity is conserved, we have the maximum possible breaking of flavor symmetry.

One fact that has created some confusion and led to misleading conclusions some authors [7] is the following: One can regulate 2+1 dimensional theories by dimensional regularization. This regularization conserves *both* flavor and parity symmetries. So, it appears that it is possible to define the theory in a way preserving these symmetries. The thing remains, however, that the gauge anomaly inherent in any parity invariant definition of such a theory is a *global* anomaly, and thus it cannot reliably be detected in a perturbative regularization scheme. One possible failure of the dimensional regularization scheme could be, for example, that there is no generalization of the Levi-Civita tensor in d+ϵ dimensions, and thus such a prescription would naturally fail to give rise to the Chern-Simons term. Such a scheme would fail to produce an overall regulated effective action, although it regulates all individual diagrams. Dimensional regularization, incidentally, apparently works and conserves parity even in *non*abelian theories, where breakdown of parity is established beyond doubt.

A Pauli-Villars regularization, on the other hand, properly regulates the action and produces the parity violating Chern-Simons term. The discrepancy between the two regularizations has been accounted in the possibility of adding a local counterterm, namely the Chern-Simons term with a *half*-integral coefficient. Such a term is nonacceptable in nonabelian theories, due to the quantization condition, but supposedly acceptable in abelian theories. We know, however, that, if we want our theory to include states of nonzero total flux (which are the ones creating all the trouble anyway), then we also must have a quantization condition. So, we conclude that the reliable regularization scheme, that cannot lead to possible inconsistencies, is the Pauli-Villars scheme.

A rather amusing confusion has arisen also in the context of this regularization scheme [6]: A calculation of the induced Chern-Simons term for massive fermions in the first couple of orders in perturbation theory gives a vanishing result! The contribution from the regulator fermions actually cancels the contribution of the physical fermions. This puzzle is related to the fact that, after adding the Chern-Simons term, the charge of the vacuum, defined as the generator of global gauge transformations, appears to vanish.

The situation is analogous to a possible puzzle that may arise in the four dimensional axial anomaly case: There, a mass term explicitly breaks axial symmetry. In the limit of massless fermions, however, instead of having restoration of the symmetry, one has an anomaly, and instead of the total axial charge to be conserved, it changes by (twice) the instanton number of the gauge field configuration. What is probably not appreciated by everybody is that, in the *massive* case, where one expects to have an extra explicit nonconservation due to the mass term, the total axial charge at times $-\infty$ and $+\infty$ is the *same*! To see this, notice that the total change of the axial charge ΔQ_5, calculated in the standard path integral way, is

$$\Delta Q_5 = 2\,\mathrm{Tr}\,\gamma^5 \left\{ \frac{\mathrm{M}}{\not{D} + \mathrm{M}} - \frac{m}{\not{D} + m} \right\} . \tag{2.10}$$

Here, M is the mass of the Pauli-Villars fermion and m the mass of the physical fermion of the theory. The difference in sign of the two terms is due to the opposite statistics that Pauli-Villars fermions obey. Due to the fact that $\not{D}$ anticommutes with γ^5, in the basis of eigenvectors of $\not{D}$ only the zero modes contribute to the trace. So, since $\not{D} = 0$ in the subspace of its zero modes, ΔQ_5 takes the form

$$\Delta Q_5 = 2\,\mathrm{Tr}_0\,\gamma^5 \left\{ \frac{\mathrm{M}}{\mathrm{M}} - \frac{m}{m} \right\} = 0\,, \tag{2.11}$$

independently on the size of the masses. One, thus, would expect this to hold also in the limit $m \to 0$. Where is, then, the anomaly?

The explanation is given in fig. 3. The *density* of ΔQ_5 (that is, $\partial_\mu j_5^\mu$) contains one contribution coming from the physical fermions and one coming from the regulator fermions. The contribution of physical fermions spreads over a distance scale of order $\frac{1}{m}$ around the anomaly-producing instanton. The contribution of regulator fermions, however, in the limit M $\rightarrow \infty$, gives exactly the anomaly distribution $F\tilde{F}$. In the limit, thus, of massless physical fermions, their contribution to ΔQ_5 spreads all over spacetime, and so only the density due to M (the anomaly) survives. Put in a different way, if V is the volume over which we integrate $\partial_\mu j_5^\mu$, the limits $m \rightarrow 0$ and $V \rightarrow \infty$ do not commute.

Exactly the same happens in the 2+1 dimensional case. Again, the total charge for massive fermions is zero, but, in the $m \rightarrow 0$ limit, the contribution of the physical fermions spreads all over space (see chapter 2) while the Pauli-Villars fermions contribute exactly the Chern-Simons term in the M$\rightarrow \infty$ limit. The contribution of the physical fermions is a highly nonlocal functional of the gauge fields. Since, however, its integral over space always gives the result $-\frac{1}{2}sign(m)\Phi$, we know that his perturbative expansion in terms of local functionals of A will *always* contain a Chern-Simons form (being the only local pseudoscalar term that integrates to Φ), that will cancel the Chern-Simons form coming from the regulator fermions. This explains the puzzle.

Finally, let us point out that, by taking advantage of the vacuum degeneracy, we can easily calculate the vacuum value of *any* operator that is odd under M. Such operators include the charge Q, the angular momentum J and the spin S, as well as their densities. Take the vacuum to be the one obtained by a specific choice of signs of m_i as they go to zero, denoted by $|m_1 \ldots m_N> \equiv |m_i>$ (the dependence on A is suppressed). Then, from the action of M on vacuum states we have

$$M|m_1 \ldots m_N> = |-m_1 \ldots -m_N> \ . \qquad (2.12)$$

So, for an M-odd operator Π, we have

$$<m_i|\Pi|m_i> = -<m_i|M\Pi M|m_i> = -<-m_i|\Pi|-m_i> \ . \qquad (2.13)$$

But we know that the two vacua differ only in the quantum numbers of their zero modes, since they have the ones with $m_i\Phi > 0$ empty and the ones with $m_i\Phi < 0$ filled. So

$$< m_i|\Pi|m_i > = < -m_i|\Pi| - m_i > - \sum_i sign(m_i\Phi)\,\langle\Pi\rangle_i \ , \tag{2.14}$$

where the expectation value of Π in the sum is in the $|\Phi|$ one-particle states of flavor i. Combining (2.13) and (2.14) we obtain

$$< m_i|\Pi|m_i > = -\tfrac{1}{2}\sum_i sign(m_i\Phi)\,\langle\Pi\rangle_i \ . \tag{2.15}$$

So the problem is reduced to calculating Π only for the zero modes. For Q, since its eigenvalue in any one-particle state is one, $\langle Q\rangle_i = |\Phi|$ and we obtain

$$\langle Q\rangle = -\tfrac{1}{2}\sum_i sign(m_i)\Phi \ . \tag{2.16}$$

For S, the spins of the zero modes are polarized along Φ, and so $\langle S\rangle_i = \frac{1}{2}\Phi$. So [8]

$$\langle S\rangle = -\frac{1}{4}\sum_i sign(m_i)|\Phi| \ . \tag{2.17}$$

For J the calculation over the zero modes is more complicated, and we also have to deal with the different definitions for it, but the results agree with the ones of chapters 2 and 3. Notice that the expression of the vacuum value of *any* M-odd operator contains the factors $sign(m_i)$, as a signal that its anomalous expectation value is due to the breaking of parity in the theory.

To summarize, the topology of gauge fields in three dimensions implies, through the Atiyah-Singer index theorem, the existence of degenerate vacua, which in turn implies that flavor and parity symmetries will be broken in some possible patterns. The appearance of the Chern-Simons term is naturally concluded in this picture. A partial verification of the above conclusions has been provided by some lattice simulations for the SU(2) flavor case [9]. Due to the method of handling the fermion fields used, the situation corresponded to a parity preserving regularization. Flavor symmetry was, then, observed to be broken, with the states of the full theory not grouping into irreducible representations of SU(2).

III. The connection with the quantum Hall effect.

Armed, now, with the intuition on abelian three dimensional theories provided by the previous section, we come to the connection of such theories with the quantum Hall effect. The most prominent and characteristic common feature of the present theories and QHE is the strict quantization of the off-diagonal component of the conductance tensor [10,11]. Their most prominent difference is that the quantization unit in QHE has double the value of that in three dimensional QED.

Some indication why there is this discrepancy should already be available to us, by examining the induced charge in the QED case. For integer flux Φ, this charge is quantized to *half*-integer values. This fact is related to the coefficient $\frac{1}{2}$ in the formula expressing the vacuum charge in terms of the spectral asymmetry of the Dirac hamiltonian

$$\langle Q \rangle = -\tfrac{1}{2} \lim_{\epsilon \to 0} \sum_n sign(E_n) e^{-\epsilon |E_n|} , \tag{3.1}$$

which, in turn, is due to the fact that the spectrum of this hamiltonian is not bounded from below. Indeed, if there were some lowest energy state, then the charge of the vacuum, defined as the difference of the charge of the filled levels minus the same charge in the trivial (free) Dirac hamiltonian, would always turn out integer, since it would be the difference of two finite integers. This would actually be the case in a condensed matter situation, where there is indeed a Fermi (rather than Dirac) sea to fill, with a finite depth. Now that such finite integers do not exist, a regularization procedure is needed, as well as a symmetrization with respect to positive energy levels, in order that the charge be odd under charge conjugation. This leads to the coefficient $\frac{1}{2}$, that, for reflection symmetric spectra, gives a half-integral Q.

We shall try to explore this simultaneous difference and similarity by working in a setting that exhibits the common features of the two phenomena as clearly as possible. But before that, as a warmup, we will rederive the formula for Q in a way that, although now may seem unnecessarily indirect, it paves the way for the future connection and derives some important intermediate results.

We shall, actually, give an explicit expression for the path integral over fermions of the action for external gauge fields. For that, we will need the time, as well as the space, to be compact. So, assume that time is periodic with period T and has, thus, the topology of S^1. The space can be any compact boundaryless manifold.

In this minkowskian spacetime, the path integral can be written

$$W = \int \mathcal{D}\psi^\dagger \mathcal{D}\psi e^{i\int d^3x \psi^\dagger [i\partial_0 + H(A)]\psi} = \det[i\partial_0 + H] . \tag{3.2}$$

$H(A)$ is the Dirac hamiltonian of the fermions at time t as a functional of the gauge fields A. If we consider, as usual, A to be static, then H is independent of time. Assuming that ψ satisfies antiperiodic boundary conditions in time,

$$\psi(\mathrm{T}) = -\psi(0) , \tag{3.3}$$

the eigenvalues of the operator $i\partial_0 + H(A)$ are

$$[i\partial_0 + H(A)]\psi_{n,k} = \lambda_{n,k}\psi_{n,k} \quad \text{with}$$

$$\lambda_{n,k} = E_n + A_0 + \frac{(2k+1)\pi}{\mathrm{T}} , \quad \psi_{n,k} = e^{\frac{(2k+1)\pi}{\mathrm{T}}t}\psi_n , \quad k = \text{integer}, \tag{3.4}$$

where ψ_n are eigenfunctions of H with eigenvalues E_n. Thus the path integral decomposes into a product of partial determinants, one for each eigenvalue of the energy

E_n. Each such determinant can be formally calculated to be [12,13]

$$\prod_k \lambda_{n,k} = \prod_k \lambda_k \cdot \prod_k \frac{\lambda_{n,k}}{\lambda_k} \quad \left(\lambda_k \equiv \frac{(2k+1)\pi}{\mathrm{T}}\right)$$
$$= N_0 \prod_k \left(1 + \frac{(E_n + A_0)\mathrm{T}}{(2k+1)\pi}\right) = N_0 \cos\frac{(E_n + A_0)\mathrm{T}}{2}, \tag{3.5}$$

where N_0 is an infinite constant normalization factor. The total path integral is, then

$$W = N' \prod_n \cos\frac{(E_n + A_0)\mathrm{T}}{2} = N \prod_n \left(e^{i\frac{(E_n+A_0)\mathrm{T}}{2}} + e^{-i\frac{(E_n+A_0)\mathrm{T}}{2}}\right). \tag{3.6}$$

Again, N', N are infinite constants.

Notice that the original path integral was invariant under the shift $A_0 \to A_0 + \frac{2\pi}{\mathrm{T}}$, since this can be undone with a global single-valued gauge transformation of the fermion field in the T direction. Formula (3.6), however, does not reflect this fact. Actually, instead of (3.5) being invariant under this shift, as it should, it transforms into minus itself. This is a manifestation of the global anomaly that we extensively talked about so far: our regularization breaks global gauge invariance.

It should be obvious, though, that this noninvariance of (3.5) is harmful only if H has zero modes: since the spectrum of H is reflection symmetric, if in the regularization of the infinite product in (3.6) we use a cutoff symmetric in E, then the terms with nonzero E_n will always appear in the combination $\cos\frac{(E_n+A_0)\mathrm{T}}{2}\cos\frac{(-E_n+A_0)\mathrm{T}}{2}$, which is invariant under the previous shift of A. Only unpaired zero modes can lead to a noninvariance. We again, then, trace down the global anomaly to the existence of zero modes. This global anomaly can be taken care of by adopting a gauge invariant regularization, or, equivalently, by adding the Chern-Simons term, which, in this case, amounts to define the determinant (3.5) as

$$\det_{E_n}[i\partial_0 + H] = N_0 \left(e^{i(E_n+A_0)\mathrm{T}} + 1\right). \tag{3.7}$$

Since we are not interested in calculating the (known) contribution of the Chern-Simons term to the vacuum charge, we will work with the previous definition (3.6), as long as we remember what it means.

We can do now a Wick rotation into imaginary time, in order to see how the path integral behaves asymptotically for large (euclidean) time. This corresponds to the substitution T $\rightarrow i$T. For positive E_n only the second term in (3.6) survives, and for negative E_n only the first term survives. This corresponds to the fact that, at zero temperature, the occupation number of all negative energy states is one and of all positive energy states is zero. Rotating back to minkowskian time, we obtain the large-T path integral as

$$W = N \prod_n e^{-i\frac{|E_n + A_0|\mathrm{T}}{2}} = N e^{-\frac{i}{2}\sum_n |E_n + A_0|\mathrm{T}} . \tag{3.8}$$

Now, from the formula

$$\frac{1}{W}\frac{dW}{dA_0} = i\langle Q\rangle\,\mathrm{T} \tag{3.9}$$

we can calculate $\langle Q\rangle$ as

$$\langle Q\rangle = \frac{1}{i\mathrm{T}}\frac{1}{W}\left.\frac{dW}{dA_0}\right|_{A_0=0} . \tag{3.10}$$

It is easy to see that, due to the absolute value, A_0 appears in the positive energy terms in the exponent of W with opposite sign than it appears in the negative energy terms. Since these terms are exactly paired, their contribution to $\langle Q\rangle$ vanishes. Only zero energy terms can contribute. For these terms, however, there is an ambiguity to the sign of the term at $A_0 = 0$. This ambiguity is resolved if we shift them slightly above or slightly below zero. This can be done by giving to the fermions a nonzero mass m. The symmetry of the spectrum, then, still persists, but the unpaired zero modes become unpaired threshold ($E = \pm m$) modes. The final result is

$$\langle Q\rangle = -\tfrac{1}{2}(n_+ - n_-) \tag{3.11}$$

n_+ (n_-) being the positive (negative) shifted zero energy modes.

It is obvious, then, that what we did is rediscover the well-known formula for the vacuum charge, with a method that connects it with the path integral and the effective action. As an aside, notice that we could have calculated $\langle Q \rangle$ in the case of finite euclidean T and nonzero mass in exactly the same way. The result is

$$\langle Q(\mathrm{T}) \rangle = \langle Q(0) \rangle \tanh \frac{m\mathrm{T}}{2} \, . \tag{3.12}$$

This is the exact result for the vacuum charge at a finite temperature $\Theta = \frac{1}{\mathrm{T}}$.

We come, finally, to the situation paralleling QHE *per se*. We consider space to be a flat torus with periods L_1 and L_2. This simply means that it is a parallelogram with periodic boundary conditions imposed on all fields in both directions (fig. 4). We will assume a constant electric field applied in the L_1 direction, and will concentrate on the current flowing in the L_2 direction. This is the simplest homogeneous configuration with finite volume that exhibits the basic QHE morphology.

In the $A_0 = 0$ gauge, an electric field can be generated only from a time dependent A. Specifically

$$\vec{E} = \hat{e}_1 E = \hat{e}_1 \dot{A}_1 = \frac{d}{dt}(\hat{e}_1 E t) \, . \tag{3.13}$$

So, a static configuration in terms of field strengths becomes a time dependent situation in terms of the gauge fields. Notice that, due to gauge invariance, the system acquires a natural periodicity: every time that A_1 becomes an integer multiple of $\frac{2\pi}{\mathrm{L}_1}$ it is gauge equivalent to zero, and thus the period is

$$\mathrm{T} = \frac{2\pi}{E\mathrm{L}_1} \tag{3.14}$$

(in natural units).

To probe the current in the L_2 direction, we introduce a constant gauge field in that direction A_2. Then

$$i \int dt d^2x \, \langle j^2 \rangle = \frac{1}{W} \frac{dW}{dA_2} \tag{3.15}$$

where the spacetime integral of j^2 and the path integral are calculated over the fundamental period of the system. From homogeneity, $\langle j^2 \rangle$ is spacetime independent, and so

$$\int dtd^2x \left\langle j^2 \right\rangle = \mathrm{TL_1L_2} \left\langle j^2 \right\rangle = \mathrm{TL_2} I_2 \ . \tag{3.16}$$

If we also call $\mathrm{L}_2 A_2 = \alpha_2$ and use (3.14), formula (3.15) becomes

$$i2\pi \frac{\left\langle j^2 \right\rangle}{E} \equiv i2\pi\sigma = \frac{1}{W}\frac{dW}{d\alpha_2} \ . \tag{3.17}$$

σ is the (vacuum) quantum Hall conductance. Finally, if we are only interested in the value of σ averaged over all values of A_2 (following a standard practice among mathematical and condensed matter physicists), which amounts to average over one natural period of α_2 of length 2π (again, gauge invariance implies this periodicity), we get

$$i2\pi \left\langle \sigma \right\rangle = \frac{1}{2\pi} \int_0^{2\pi} \frac{1}{W}\frac{dW}{d\alpha_2} d\alpha_2 = \ln W|_{\alpha_2=0}^{2\pi} \ . \tag{3.18}$$

Since $\alpha_2 = 2\pi$ is gauge equivalent to $\alpha_2 = 0$, W should be the same for both values. So, its logarithm can only differ by a change of Riemann sheet, that is, by an integer multiple of $i2\pi$. This implies

$$2\pi \left\langle \sigma \right\rangle = \text{integer} \, . \tag{3.19}$$

This is the famous quantization of the quantum Hall conductance.

In our case, however, we should be careful. Remember that our regularization was not globally gauge invariant, and thus the previous argument may fail. In fact, if a global anomaly is present, $W(2\pi)$ will be equal to $-W(0)$, and so $\ln W$ may change by a *half* integer multiple of $i2\pi$. This, then, would lead to a quantization of σ to *half* the previous quantization unit.

In order to see if, indeed, we have an anomaly in this case and to find the exact change in the effective action ($\ln W$), we exploit our knowledge of the situation in the case of static fields. There, as we have repeatedly demonstrated, the above defined effective action will change exactly by $\pm n\Phi\pi$, where n is the winding number of the nontrivial transformation in T and Φ is the total magnetic flux running through our two-space (the sign, as usual, depends on the sign of the mass of the fermions as it is drawn to zero). In the present case, since spacetime has the topology of $S^1 \times S^1 \times S^1$, we may regard L_2 as time and $(\mathrm{T}, \mathrm{L}_1)$ as our spatial section. A change of α_2, then, by 2π is indeed a nontrivial gauge transformation of winding number 1. The flux running through the $(\mathrm{T}, \mathrm{L}_1)$ section is

$$\frac{1}{2\pi} \int dt dx^1 F_{01} = \frac{1}{2\pi} \mathrm{T} \mathrm{L}_1 E = 1 \,. \tag{3.20}$$

So we see that $\Delta \ln W = \pm\pi$ and

$$\langle \sigma \rangle = \pm \frac{1}{4\pi} \,. \tag{3.21}$$

The connection of QHE and three dimensional QED should now be transparent: The derivation that led to (3.18) is exactly applicable to the condensed matter physics situation. There, the (ordinary quantum mechanical) path integral with a *finite* number of states filled can be calculated along the same lines, with no possible anomaly arising, due to the finiteness of filled levels, and the quantization follows from gauge invariance. Remember that gauge invariance was central in the original explanation of the phenomenon by Laughlin. In fact, our procedure is the translation into field theory language of the manipulations that mathematical physicists use [11,14], in terms of the many-particle state of the condensed matter system, to show the connection of the quantum Hall conductance with topology and the famous TKN^2 integers [15]. The same mathematical structure leads to the vacuum quantum Hall conductance in our case. Its field theoretical nature manifests in the infinity of filled levels, that leads to the global anomaly, that leads to the difference of half between the two conductances.

There remain a couple of important phenomena to relate and explain. The first is the fact that, in the QHE case, as we increase the magnetic field, the conductance jumps from one quantized value to the next, while in the present case it is strictly constant and totally independent of the magnetic field.

The explanation of this phenomenon in the QHE case has to do with level crossings. In ordinary quantum mechanics, what produces the "winding number" of the phase of the vacuum state that gives rise to (3.18) is degeneracies. These degeneracies have codimension three, which means that we need to control three parameters in order to achieve one. So, generically, as α_1 and α_2 vary from 0 to 2π there are no energy levels of the system crossing each other, that is to say, there are no values of α's where a degeneracy occurs. As we vary the other parameters of the system, though, and namely the magnetic field B, such degeneracies may be achieved. For each such value of B the number of degeneracies "contained" in the torus (α_1, α_2) increases by one, and thus the integer in (3.19) changes by one. In the present case, this "winding number" of the path integral happens to be independent of B, and so this phenomenon does not happen.

The other phenomenon is that of the fractional QHE, where $2\pi\sigma$ is quantized to a *fraction*. The irreducible denominator of this fraction is, as a rule, odd. This phenomenon is not yet perfectly understood. The proposed explanation [16] is that the ground state is "continuously" degenerate, with only a fraction of its levels filled, and the state reached at $\alpha_2 = 2\pi$ is the same as the original one but with *different* levels filled, and thus requiring α_2 to change by a multiple of 2π before we reach the identical initial state. So, $\frac{1}{2\pi}$ in (3.18) is divided by an integer, that becomes the denominator of the fraction. This explanation has some obvious weaknesses, namely it provides no explanation why such a continuous degeneracy should occur and it gives no clue for the odd denominator rule. In our case, this simply does not happen. It is true that we need a transformation of 4π in order to have the same path integral, due to the global anomaly, which gives rise to the extra factor $\frac{1}{2}$, but the similarities stop here. There is no continuous degeneracy, and the odd denominator rule is violated. It is conceivable that the two phenomena may be more alike than they appear presently, but this has yet to be demonstrated.

IV. Conclusions.

We provided, in this last chapter, some fairly simple hamiltonian arguments that demonstrate the necessity of breaking either parity or global flavor symmetry in three dimensional QED. We also examined the possible patterns into which these symmetries can break, and showed that, when parity does not break flavor symmetry breaks maximally, and that for odd number of flavors parity necessarily breaks.

It should be pointed out, however, that we offered of no argument as to determine which of the possible breaking patterns will be realized in a theory that starts with (strictly) massless fermions. It is plausible, though, that *any* pattern may be obtained, and that taking the limit of a massive theory is a necessary procedure in order to have a well-defined quantum theory. In this connection, we are skeptical about the validity of some recent numerical calculations [17] indicating that the Schwinger-Dyson equations of a theory with zero fermionic bare masses do not possess any solutions with nonzero mass of the photon or the fermions, thus suggesting that parity is unbroken. We believe that starting with strictly massless fermions creates a statistical mixture of two theories that break parity in opposite patterns, and thus shows no sign of the breaking. This can be seen from the fact that the path integral, in the presence of exact zero modes, does not relax into a vacuum configuration with a sharp charge, in the limit of large euclidean time (large temperature), but gives an average of the two path integrals obtained with the zero modes filled or empty. However, as we showed, such a theory is nonunitary. Along these lines, it may also be that the treatment of ref. [5] is not rigorous in this sense. There, QED_3 was examined for massless fermions in the limit of large even N, and was concluded that parity does not break, but flavor $U(2N)$ breaks into $U(N) \otimes U(N)$. This is, at any rate, in agreement with our conclusions.

We also showed that there is a general connection between the behavior of the vacuum in such theories and the quantum Hall effect. It is true, that what was shown was more a similarity between the basic mathematical framework that creates the two phenomena rather than between their specific physical mechanisms. In fact, the argument exposed works for any planar field theoretical system. It is, however, the specific properties of three dimensional QED, and namely the global anomaly, that drives the right hand side of (3.19) into nonzero, and in fact noninteger values.

The fact remains, after all this, that there is still progress to be made towards really understanding the mechanism of QHE, in particular the fractional one, and determining what exact properties of the planar system are crucial for the phenomenon to manifest. Whether the connection with three dimensional QED will contribute in this direction and remain a useful tool is, yet, to be seen.

Finally, let us conclude with some remarks about nonabelian theories. Although fermionic zero modes *do* occur in such theories, as can be seen by starting with a nonzero flux configuration of an abelian theory and "nonabelianizing" it (which, incidentally, also gets rid of the string singularities), they are not dominant, in the sense that the gauge field configuration space does not decompose into components characterized by a fixed number of zero modes each. Moreover, the total fermion number of the vacuum even in the presence of zero modes vanishes, since they always come in empty-filled combinations (this is a corollary of the fact that the index of any nonabelian Dirac operator in two dimensions vanishes, due to the tracelessness of the group generators). So, it is clear that we cannot use the reasoning of section II to determine how parity and/or flavor symmetry break.

The hamiltonian explanation of parity breaking in the single flavor case is that, in a parity-invariant regulated theory, the *phase* of the ground state cannot be globally well defined. As we adiabatically transport the vacuum state around a nontrivial loop of gauge equivalent configurations (which is a map from $S^2 \times S^1$ to the gauge group, classified by $\pi_3(G) = Z$), it picks up a phase equal to π, which prohibits a definition of its phase even patchwise in the space (due to the noninteger winding number). This happens exactly because this path is "entangled" with regions of vacuum degeneracy (zero modes), where the phase is not defined. To remedy that, we need to redefine the phase of the ground state so as to have an integer winding number, which changes the dynamics of the theory exactly by adding to the action the Chern-Simons form with a half-integral coefficient. This term, however, breaks parity.

It is already clear that in a theory with an odd number of flavors parity must still break, since the vacuum state again picks up a phase equal to the number of flavors times π, which is still nontrivial. It appears plausible that flavor symmetry will break in a way similar to the one in abelian theories. However, we have not yet thought on whether and how the above picture can be used, in analogy with section II, in order to further probe the global symmetry breaking patterns of a multiflavor nonabelian theory.

REFERENCES

[1] A. N. Redlich, Phys. Rev. Lett. **52** (1984) 18 and Phys. Rev. **D29** (1984) 2366.

[2] A. P. Polychronakos, Nucl. Phys. **B281** (1987) 241.

[3] S. Deser, R. Jackiw and S. Templeton, Phys. Rev. Lett. **48** (1982) 975 and Ann. of Phys. **140** (1982) 372.

[4] A. P. Polychronakos, Nucl. Phys. **B278** (1986) 207.

[5] R. D. Pisarski, Phys. Rev. **D29** (1984) 2423.

[6] Y-C. Kao and M. Suzuki, Phys. Rev. **D31** (1985) 2137.

[7] S. Rao and R. Yahalom, Phys. Lett. **B172** (1986) 227.

[8] Note that this is in disagreement with the result of D. Boyanovsky, Phys. Rev. **D34** (1986) 612, where the expression $\langle S \rangle = \frac{1}{4}\Phi$ is found. We point out that this expression has the wrong transformation properties under both C and P.

[9] K. Farakos and G. Koutsoumbas, Technical Univ. of Athens report, unpublished.

[10] R. B. Laughlin, Phys. Rev. **B23** (1981) 5632.

[11] D. J. Thouless, M. Kohmoto, M. P. Nightingale and M. den Nijs, Phys. Rev. Lett. **49** (1982) 405.

[12] R. Dashen, B. Hasslacher and A. Neveu, Phys. Rev. **D12** (1975) 2443.

[13] A. J. Niemi, G. W. Semenoff and Y-S. Wu, Nucl. Phys. **B276** (1986) 173.

[14] J. E. Avron and R. Seiler, Phys. Rev. Lett. **54** (1985) 259.

[15] B. Simon, Phys. Rev. Lett. **51** (1983) 2167.

[16] R. Tao and Y-S. Wu, Phys. Rev. **B30** (1984) 1097.

[17] S. Rao and R. Yahalom, Phys. Rev. **D34** (1986) 1194.

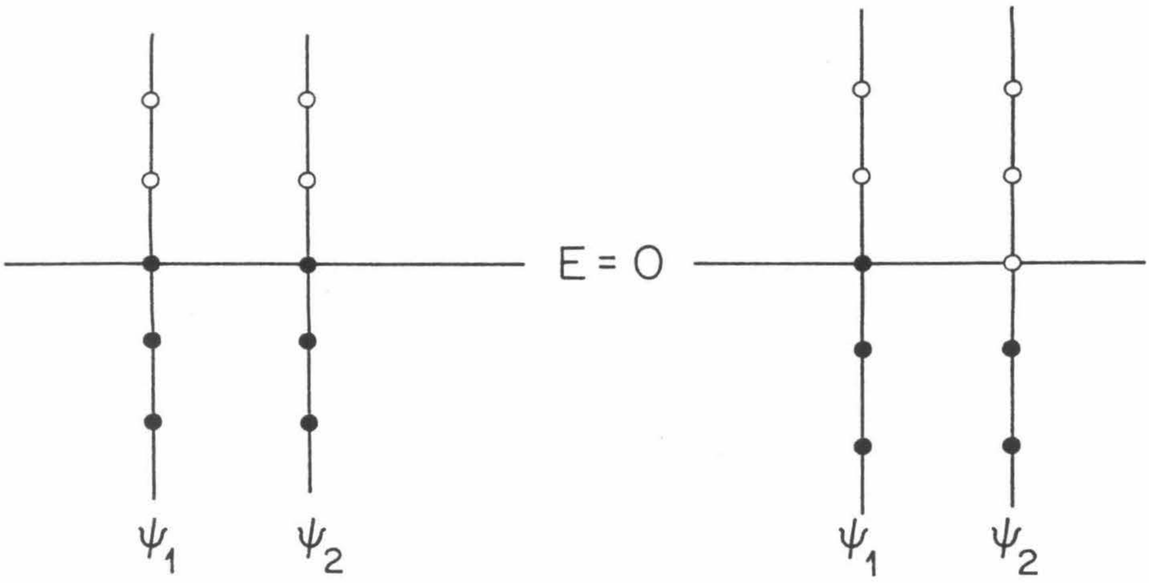

Fig. 1a–b: The Dirac spectrum of a theory with two fermion flavors in the presence of a flux $\Phi=1$ is represented schematically. In fig. 1a, only one zero energy level is filled, which creates a parity invariant state. In fig 1b, both zero modes are filled, resulting to an SU(2) invariant state.

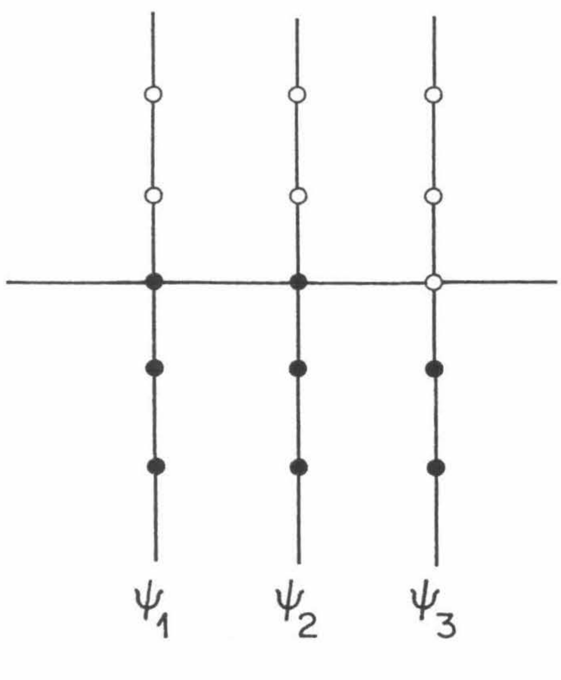

Fig. 2: The Dirac spectrum of a theory with three flavors is represented, again for $\Phi=1$. The state shown is neither parity nor SU(2) invariant.

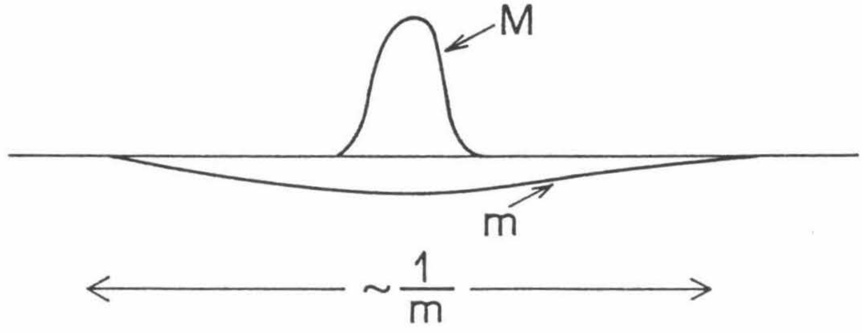

Fig. 3: Qualitative plot of the distribution in spacetime of the density of $\partial_\mu j_5^\mu$ in four dimensions. The area under both curves is the same. The lower curve, however, for m going to zero "dilutes" infinitely.

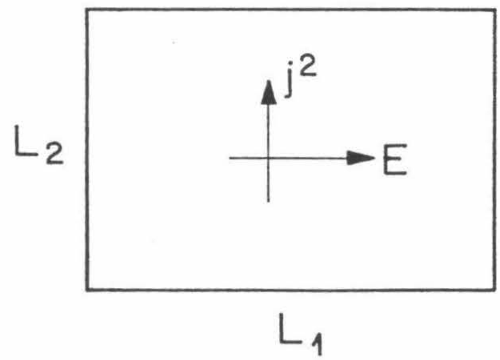

Fig. 4: The basic configuration of space, electric field and current for the topology of the quantum Hall effect.

EPILOGUE

This chapter is not headed "conclusions" simply because it is not supposed to contain any. Rather, it is merely meant to be an editorial of the opinions of the author on questions concerning the prospects of physics in the few years to come, as well as a short guideline of possible future directions in which research related to the subjects exposed in this treatise might be continued. However, for the benefit of those who, for reasons with which we fully sympathize, gave up reading this manuscript before its long awaited end and jumped directly to this chapter, we give a brief summary of the results and conclusions already stated in the relevant sections of the preceding chapters.

The main points made in chapters 1 to 3 are, firstly that a quite intuitive and physical picture of vacuum charge and charge fractionization can be obtained by looking at general properties of the field theories that create them and establishing connections with other well-studied quirks of these theories, like anomalies and bosonization, and secondly that this charge, as well as other quantities like angular momentum and spin, can be exactly calculated in three dimensional QED. The problem of what boundary conditions are to be used for the fermion fields, in the case of bounded space, turned out to be nontrivial. A local variant of boundary conditions was introduced and used, which turned out to be more physical than the spectral conditions that people have carried over from mathematics and used in physics contexts so far. The Atiyah-Patodi-Singer index theorem was suitably modified in order to be reconciled with the new conditions, and the "index" turned out fractional. Finally, questions concerning the definition of angular momentum and the behavior of Dirac strings were examined and clarified.

Chapters 4 and 5 were devoted to topological considerations. A quantization condition for the gauge invariant photon mass in three dimensional QED was derived, required if the theory is to possess states of nonzero magnetic flux. It was further shown that the nontrivial topological properties of the Chern-Simons term, that creates the photon mass, are responsible for the anomalous properties that the theory possesses under rotations, and thus for the anomalous values of angular momentum. Finally, a topological argument was given to demonstrate that parity must break in this theory. An isolated but cute byproduct of this analysis was the discovery and rectification of a flaw in the argument for the four dimensional global SU(2) anomaly.

Finally, chapter 6 gave a hamiltonian picture of the parity anomaly that permitted to generalize the results to the many-flavor case and determined the patterns into which flavor symmetry and parity break. It turned out that the two symmetries are quantum mechanically incompatible. Conservation of parity is possible only for even number of flavors and results in maximal breaking of flavor symmetry. It was shown that parity breaking was responsible for all exactly computable vacuum quantities in this case and a simple way to derive them was given, as well as a generalization for nonzero temperature. The parity anomaly of three dimensional QED and a path integral argument were, then, combined to exhibit a formal connection between this theory and the quantum Hall effect and to explain the difference in the unit of quantization of the Hall conductance in the two cases.

Future work on the subject could continue in several directions. In the era of purely calculational (but nontrivial) tasks, one could attempt to evaluate the *densities* of vacuum quantities in the case where a symmetry argument is not available. The identity of the results for a very shallow and a maximally tall thin soliton in two dimensions suggests that this may actually be achievable, despite the fact that such densities are generally expected to be nonanalytic functions in space. Similar results can be sought for the Casimir energy of the vacuum in several situations, which again is not provided by topological arguments.

A much more interesting question is the one concerning the realization of global symmetries in situations other than the ones examined so far. It would actually be extremely important to have a solid argument for chiral symmetry breaking in QCD using techniques similar to the ones of chapter 6. These arguments do not *a* *priori* work, since the topological considerations that drive them are absent in the four dimensional case. From the recent work of G. Tiktopoulos, however, where a mechanism involving fermion zero modes was called for in order to create the chiral symmetry breaking, it may turn out that such an approach *may* work if properly extended. This is the most intriguing possibility.

It would be interesting, from the mathematical point of view, to examine further the properties of the local boundary conditions introduced in this work and see to what extent the "breakdown of unitarity" due to the switching on of topologically nontrivial gauge fields can be a fruitful notion with mathematically rich consequences. We do not personally think, though, that this is a physically very interesting pursuit, although the work of chapter 3 has recently aroused some interest. We believe that the results of this chapter can be generalized to the many-dimensional case and for nonabelian operators. The local boundary conditions to be used there are more varied and, possibly, more interesting. Further, one could examine the most general class of local boundary conditions available in two dimensions, and not just the ones preserving parity. The index would then almost always vanish, but we conjecture that the *spectral asymmetry* of the Dirac operator will remain invariant and equal to the index of the previous (symmetric) case. Out of sheer lack of time and immediate interest we have not pursued this investigation any further.

Finally, further work can be done towards the connection with the quantum Hall effect (QHE). It is plausible that the mechanisms of QHE and superconductivity are closely connected. This is suggested by the fact that the vacuum in three dimensional QED has properties mimicking both phenomena, and, what is more, both seem to stem from the same basic features of the theory. In view of the physical interest of QHE and the renewed excitement about superconductivity, due to the discovery of high temperature superconductors, such a connection would be extremely useful and welcome. Given that the work of mathematical physicists had, so far, practically zero impact in the physics of the field, this would be the first significant contribution of the kind.

Let us conclude with some remarks of a highly personal and opinionate nature on the track that physics seems to be following lately. Surely enough, all theoretically interesting recent advances seem to be connected with string theory. This is to a certain degree understandable, since this appears to be the only theory that is not *a priori* hopeless as far as unifying gravity with the rest of fundamental interactions is concerned. It is, unfortunately, also to a significant degree hopelessly remote from any phenomenologically relevant results. One can not exclude that, by some remarkable discovery or spark of intuition, somebody could come up with a result of such generality and breadth of range that it would contribute something of significance to our desperatingly low-energy experimental world (the connection between the compactification of extra spacetime dimensions and some Yukawa couplings is the only result along this line that the author is aware of). This, however, would be unexpected, to say the least. Moreover, most of the work on the field nowadays is directed either towards better understanding its basic features or formulating it in terms of the even more high browed and exotic string field theory.

It is the author's personal opinion that part of the reason why strings enjoyed such immediate popularity after their internal consistency was established is due to the fact that their mathematical structure is exceedingly rich and one has the feeling that one gets out in terms of results more than one puts in in terms of effort. This is, simultaneously, the great peril of the subject. Young physicists who get immediately involved in the field, and after having indulged in the mathematical sybaritism of string theory, run the risk of never quite appreciating the challenges and importance of ordinary field theory. There are, however, a lot of open and important questions there and effort on them should by no means be abandoned. Will people have the courage and will to go back to a "lower standards of living" physics if reality makes it clear that this is the only way to progress?

The author's involvement in the subject, so far, of insignificant contribution as it may have been, was one resembling Ulysses' mythical trip past the isle of Seirenes: Tightly bound to his mast with the ropes of physical reality, he enjoyed their songs of incomparable mathematical lyricism without running the risk of being terminally seduced and captured into their surreal world. One cannot be too sure about oneself, however. If the pleasures of working on the subject grow uncontrolled and the challenge of remaining in business while doing "classical" physics keeps increasing, his involvement may become total. And, after all, one should not underestimate the romanticism that any physicist carries in himself to a lesser or greater degree: the prospect of making progress towards the ultimate unification of all interactions, however utopic, is enough to make anyone abandon any set of deeply-rooted principles, and the author cannot hope to be an exception.

May all four forces be with us soon!