ANNALS OF PHYSICS 159, 220–251 (1985)

The Schwinger Model and Its Axial Anomaly

N. S. MANTON

Institute for Theoretical Physics, University of California, Santa Barbara, California 93106

Received May 16, 1984

The Schwinger model (quantum electrodynamics with massless fermions in one spatial dimension) is solved, supposing that space is a circle. This clarifies aspects of the usual version of the model, where space is a line, without changing the physics. The Hamiltonian formalism is used. On a circle, an abelian gauge field has one physical degree of freedom, and the gauge covariant Dirac operator, which couples the fermions to this degree of freedom, exhibits spectral flow. The relationship between the spectral flow and the axial anomaly is explained. Some variants of the Schwinger model are also discussed.

1. INTRODUCTION

The Schwinger model [1]—quantum electrodynamics in one spatial dimension with massless fermions—is a fascinating field theory. In its standard fermionic form it is not trivial, nor obviously soluble. However, by a study of Green's functions [1, 2], and later by explicit bosonization [3], it was completely solved. The model is in fact equivalent to the free field theory of a massive scalar field in one spatial dimension. The physical scalar particles have mass $m = e/\sqrt{\pi}$, where e is the electric coupling constant.

Because the fermions are massless, one might expect both the electric charge and the axial charge to be conserved quantities. However, there is an axial anomaly [4], and the divergence of the axial current is

$$\partial_\mu j_5^\mu = \frac{1}{\pi} E_x, \tag{1.1}$$

where E_x is the electric field in the spatial ($= x$) direction. The fact that the scalar particles are massive, even though the fundamental fermion and abelian gauge fields are massless, is closely related to this anomaly equation.

The purpose of this paper is to re-examine the Schwinger model in the canonical Hamiltonian formalism. We shall suppose that space is a circle of length 2π. It is a simple modification to let the length be $2\pi L$, and in the limit $L \to \infty$ one recovers the usual version of the model. As we shall see, working on a circle greatly clarifies the

0003-4916/85 $7.50

properties of the model that interest us, but it does not change the physics in any significant way.

What is particularly clear on a circle is that there is precisely one physical degree of freedom in the gauge potential, namely, the phase of the Wilson loop variable

$$\exp i\int_0^{2\pi} A_x(x)\,dx. \tag{1.2}$$

It is possible to choose a gauge where A_x is independent of x (Coulomb gauge) and $0 \leqslant A_x \leqslant 1$. The values $A_x = 0$ and $A_x = 1$ are gauge equivalent and should be identified. The true configuration space of pure electrodynamics is therefore itself a circle. On the infinite line, this is also true but harder to see. If the circle has length $2\pi L$, the analog of the gauge above becomes $0 \leqslant A_x \leqslant L^{-1}$, and this doesn't have a good limit as $L \to \infty$. A related observation is that on the circle, finite energy transverse electric fields are physically important. This remains true as $L \to \infty$, but in the limit, such electric fields become infinitesimal.

A more technical advantage of working on a circle is that momenta are discrete. This is helpful because various operators, which depend on the momentum variable p, have different behavior depending on whether $p = 0$ or $p \neq 0$.

The essentially novel part of this paper is Section 3. Here it is shown how the zero momentum scalar particles arise from the coupled fermion and gauge fields. It is also shown why the axial charge is not conserved. The underlying idea is that of spectral flow [5]. The one-dimensional covariant Dirac operator which occurs in the Hamiltonian is

$$\begin{pmatrix} -i\partial_x + A_x & 0 \\ 0 & i\partial_x - A_x \end{pmatrix}. \tag{1.3}$$

In the Coulomb gauge, the energy spectrum is $p + A_x$ for left-handed fermions (upper component) and $-p - A_x$ for right-handed fermions (lower component). The momentum p takes all integer values. As one orbits the configuration space once (i.e., as A_x increases from 0 to 1) the spectrum is permuted. The energies of left-handed fermions increase by one, those of right-handed fermions decrease by one. The spectral flow is two. While it has been recognized before that the spectral flow is related to the anomaly [6], we go further and find directly how the spectral flow influences the dynamics of the gauge field.

One consequence of the spectral flow is that the regularized axial charge depends on the background gauge potential. Naively, the axial charge is an integer—the difference between the numbers of left- and right-handed particles—but it is also ambiguous because the filled Fermi–Dirac sea has an infinite number of particles of each handedness. After regularization, the charge is no longer ambiguous, but neither is it an integer. This leads to the anomaly. In our analysis, therefore, it is "fractional charge" that causes the anomaly.

In other field theories, especially those with solitons coupled to Fermi fields, it has been shown that the solitons can acquire a fractional charge [7]. There, the analysis

indicated that an anomaly was responsible for the fractional charge. Clearly these ideas are very closely related.

The analysis of the Schwinger model presented here developed from our study of a simple quantum mechanical system— a particle moving on a torus in a background magnetic field. There, too, there is a spectral permutation and a type of anomaly. In fact, the restricted Schwinger model which we discuss at the end of Section 3, in which fermionic excitations are assumed to be absent, is mathematically equivalent to the particle on the torus, but the physics is quite different. We find it interesting that this system has an anomaly, but none of the complications, and infinities, of a quantum field theory.

The material of this paper is organized as follows. Section 2 introduces electrodynamics on a circle. Section 3 is the heart of the paper; this is where we discuss the physical consequences of the spectral flow in the Schwinger model. In Section 4, we complete the analysis of the model using the technique of bosonization. Sections 5 and 6 discuss some modifications of the Schwinger model. Finally, in Section 7, we outline the relevance of some of the ideas discussed here to quantum chromodynamics in three spatial dimensions.

The proof of a lemma fundamental to bosonization is in Appendix A. In Appendix B we discuss quantum mechanical particle motion on a torus in a background magnetic field [8]. In Appendix C we explain the connection between fractionally charged solitons and the anomaly of axial electrodynamics.

2. Electrodynamics on a Circle

Let us consider pure electrodynamics in one spatial dimension, where space is a circle of length 2π. This is a system with one degree of freedom, since at a given time, the set of gauge inequivalent field configurations is itself a circle. The only gauge invariant quantity that can be constructed from the spatial component of the gauge potential, $A_x(x)$, is the phase of the Wilson loop variable

$$\exp i \int_0^{2\pi} A_x(x)\, dx. \tag{2.1}$$

(The coupling constant e is absorbed in the gauge potential.)

We can make the configuration space explicit by choosing a gauge. A gauge transformation has the form

$$A_x' = A_x + i(\partial_x g)\, g^{-1}, \tag{2.2}$$

where $g(x) = \exp i\Lambda(x)$. g must be single-valued, so Λ is single-valued mod 2π. If $\Lambda(2\pi) = \Lambda(0)$ then the gauge transformation is topologically trivial, and if $\Lambda(2\pi) = \Lambda(0) + 2\pi n$, with $n \neq 0$, then it is nontrivial and has winding number n. By a topologically trivial transformation one can make A_x independent of x, and by a further nontrivial transformation of the form $g = \exp inx$, one can bring A_x to the interval $[0, 1]$. Since $A_x = 0$ and $A_x = 1$ are gauge equivalent configurations, we shall

identify them. (This identifies what could be regarded as topologically distinct vacua.) The configuration space is therefore a circle. From now on we shall always work in this gauge, so A_x depends only on time.

The kinetic energy of a time-dependent field is

$$\frac{1}{2e^2}\int_0^{2\pi} E_x^2\, dx, \tag{2.3}$$

where E_x is the electric field

$$E_x = \frac{dA_x}{dt} - \partial_x A_t. \tag{2.4}$$

The time component of the gauge potential is an auxiliary quantity, which is determined by imposing Gauss' law

$$\partial_x E_x = \partial_x \dot{A}_x - \partial_x \partial_x A_t = 0. \tag{2.5}$$

In our gauge, this implies $\partial_x A_t$ is independent of x, and since A_t is single-valued, $\partial_x A_t$ must vanish. So $E_x = \dot{A}_x$. In one spatial dimension there is no magnetic field, so no field potential energy, and the total field energy is therefore $\pi \dot{A}_x^2/e^2$. The classical equation of motion, $\ddot{A}_x = 0$, allows an arbitrary electric field, constant in both space and time.

The quantization of this system is elementary. The conjugate operators E_x and A_x obey the commutation relations

$$[E_x, A_x] = -\frac{ie^2}{2\pi} 1, \tag{2.6}$$

and the Hamiltonian operator is

$$H = \frac{\pi}{e^2} E_x^2. \tag{2.7}$$

The most general operator representing the electric field E_x, which satisfies (2.6), is

$$E_x = -\frac{ie^2}{2\pi}\left(\frac{d}{dA_x} + i\Theta\right). \tag{2.8}$$

Θ is an arbitrary real parameter, and physically $e^2\Theta/2\pi$ acts as a classical constant electric field. However, we shall suppose $\Theta = 0$, and postpone till Section 6 a discussion of the significance of nonzero values of Θ in pure electrodynamics and in the Schwinger model.

Since $2\pi A_x$ is an angular variable, the wavefunction $\psi(A_x)$ must be periodic with period 1. A stationary state has the form

$$\psi(A_x) = \exp 2\pi i n A_x, \qquad n \in \mathbf{Z}. \tag{2.9}$$

It is an eigenstate both of the electric field operator, with the now quantized eigenvalue e^2n, and of the Hamiltonian, with energy eigenvalue $\pi e^2 n^2$. If the circle were opened up, such an electric field would be that produced classically by charges $\pm n$ at the ends.

3. The Schwinger Model on a Circle

The Lagrangian density of the model is

$$L = \frac{1}{2e^2}(\partial_t A_x - \partial_x A_t)^2 + \bar{\psi} i\gamma^\mu (\partial_\mu + iA_\mu)\psi. \tag{3.1}$$

ψ is a 2-component massless Dirac spinor, and $\bar{\psi} = \psi^\dagger \gamma^0$. We work with the Dirac matrices $\gamma^0 = \sigma^1$, $\gamma^1 = -i\sigma^2$, and $\gamma^5 = \gamma^0\gamma^1 = \sigma^3$. σ_1, σ_2, and σ_3 are the Pauli matrices.

We choose the same gauge as before,

$$\partial_x A_x = 0, \qquad 0 \leqslant A_x \leqslant 1. \tag{3.2}$$

Gauss' law is now

$$-\partial_x \partial_x A_t = e^2 \psi^\dagger \psi, \tag{3.3}$$

where $\psi^\dagger\psi$ represents the electric charge density. This equation leads to an important constraint. For consistency, the total electric charge must be zero. Physically this is reasonable, as the sources and sinks of electric flux on a circle must balance; on a line, flux could escape to infinity.

Standard manipulations [9] lead to the quantum Hamiltonian

$$H = -\frac{e^2}{4\pi}\frac{\partial^2}{\partial A_x^2} - \int_0^{2\pi} \psi^\dagger i\sigma^3(\partial_x + iA_x)\psi \, dx + \frac{e^2}{2}\int_0^{2\pi} (\psi^\dagger\psi)\frac{1}{-\partial_x^2}(\psi^\dagger\psi)\, dx, \tag{3.4}$$

with the canonical anti-commutation relations

$$\begin{aligned} \{\psi_\alpha^\dagger(x), \psi_\beta(y)\} &= \delta(x-y)\,\delta_{\alpha\beta}, \\ \{\psi_\alpha(x), \psi_\beta(y)\} &= \{\psi_\alpha^\dagger(x), \psi_\beta^\dagger(y)\} = 0. \end{aligned} \tag{3.5}$$

The operator

$$E_x^{\mathrm{tr}} = -\frac{ie^2}{2\pi}\frac{d}{dA_x} \tag{3.6}$$

represents the transverse electric field, which is classically $\dot{A}_x$, and is independent of x. The first term in H is the energy in this part of the field. The second term is the kinetic energy in the Fermi field, and the third is the Coulomb energy which is induced when Gauss' law is used to eliminate A_t. The x-space Green's function for

$-\partial_x^2$ on the circle is slightly awkward to write down, but is not needed. We shall change to the momentum representation where $-\partial_x^2 = p^2$ as usual.

Important operators are the vector current $j^\mu = \bar{\psi}\gamma^\mu\psi$, which is the electric current, and the axial current $j_5^\mu = \bar{\psi}\gamma^\mu\gamma^5\psi$. They have components

$$j^0 = j_5^1 = \psi_1^\dagger\psi_1 + \psi_2^\dagger\psi_2, \tag{3.7}$$

$$j^1 = j_5^0 = \psi_1^\dagger\psi_1 - \psi_2^\dagger\psi_2. \tag{3.8}$$

Naively, both the electric charge

$$Q = \int_0^{2\pi} (\psi_1^\dagger\psi_1 + \psi_2^\dagger\psi_2)\, dx \tag{3.9}$$

and the axial charge

$$Q_5 = \int_0^{2\pi} (\psi_1^\dagger\psi_1 - \psi_2^\dagger\psi_2)\, dx \tag{3.10}$$

commute with the Hamiltonian, but the axial charge is not conserved, as we shall see.

The momentum space Fermi operators are defined by Fourier expansion

$$\begin{pmatrix}\psi_1(x)\\ \psi_2(x)\end{pmatrix} = \frac{1}{\sqrt{2\pi}} \sum_k \begin{pmatrix}a_{1,k}\\ a_{2,k}\end{pmatrix} e^{ikx}, \qquad k \in \mathbf{Z}; \tag{3.11}$$

$$\begin{aligned}\{a_{\alpha,k}^\dagger, a_{\beta,l}\} &= \delta_{kl}\delta_{\alpha\beta},\\ \{a_{\alpha,k}, a_{\beta,l}\} &= \{a_{\alpha,k}^\dagger, a_{\beta,l}^\dagger\} = 0.\end{aligned} \tag{3.12}$$

Let us define momentum space chiral charge density operators

$$\begin{aligned}\rho_\alpha(p) &= \int_0^{2\pi} \psi_\alpha^\dagger(x)\, \psi_\alpha(x)\, e^{ipx}\, dx\\ &= \sum_k a_{\alpha,k+p}^\dagger a_{\alpha,k},\end{aligned} \tag{3.13}$$

so $j^0(p) = \rho_1(p) + \rho_2(p)$ and $j^1(p) = \rho_1(p) - \rho_2(p)$.

The Hamiltonian, in the momentum representation, takes the form

$$\begin{aligned}H = &-\frac{e^2}{4\pi}\frac{\partial^2}{\partial A_x^2} + \sum_p a_{1,p}^\dagger a_{1,p}(p + A_x) + \sum_p a_{2,p}^\dagger a_{2,p}(-p - A_x)\\ &+ \frac{e^2}{4\pi} \sum_{p\neq 0} j^0(p)\frac{1}{p^2} j^0(-p).\end{aligned} \tag{3.14}$$

The fact that A_x is independent of x has been used here. Also, since the total electric charge is zero, the longitudinal part of the electric field $E_x^{\text{long}} = -\partial_x A_t$ has no $p = 0$

component, and the potentially singular $p = 0$ component of the Coulomb energy is absent. This is important later.

For fixed A_x, $a^\dagger_{1,p}$ and $a_{1,p}$ are creation and destruction operators for a positive chirality particle of momentum p and energy $p + A_x$. Similarly $a^\dagger_{2,p}$ and $a_{2,p}$ are creation and destruction operators for a negative chirality particle of momentum p and energy $-p - A_x$. We shall refer to these types of particles as left- and right-handed, respectively.

Normally in field theory one would distinguish the positive and negative energy particles and reinterpret the particle destruction operator $a_{\alpha,p}$ as an antiparticle creation operator if the particle's energy would be negative. Here we shall not do this, because as A_x changes, an energy level can change sign. In the standard language one would then have to admit the possibility of a particle appearing from nowhere, or an antiparticle disappearing, as A_x varied.

For a fixed value of A_x, a basis for the physical states consists of those states, subject to Fermi statistics, in which each energy level is specified as either filled (a particle) or empty (no particle), and in which almost all (i.e., all but a finite number) of the negative energy levels are filled and almost all of the positive energy levels are empty. A physical state is a linear combination of such basis states, with complex amplitudes. When A_x is allowed to vary, the amplitudes are functions of A_x. Of course, in a given basis state it is the particle momenta that remain constant as A_x changes, rather than the particle energies.

It is useful to define the Fermi surfaces for the left- and right-handed particles, which we denote by E^F_L and E^F_R. This is done at fixed A_x. Associated with any basis state there is an unexcited basis state, with the same numbers of left- and right-handed particles, in which no empty level has a lower energy than a filled one. The Fermi surfaces are defined to lie halfway between the highest filled and lowest empty levels of this unexcited state (see Fig. 1 for an example). Any basis state can be described in terms of excitations relative to its Fermi surfaces.

Let us now consider the spectral flow of the one-particle energy levels, and its consequences. The spectrum at $A_x = 0$ and $A_x = 1$ is the same, namely, the integers, as it must be since these values of A_x are gauge equivalent and are identified.

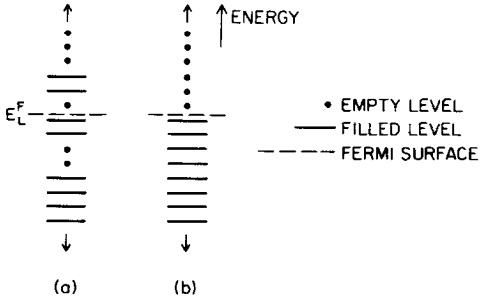

FIG. 1. An excited basis state (a) and its corresponding unexcited state (b). E^F_L is the Fermi surface. (Only the energy levels of left-handed particles are shown.)

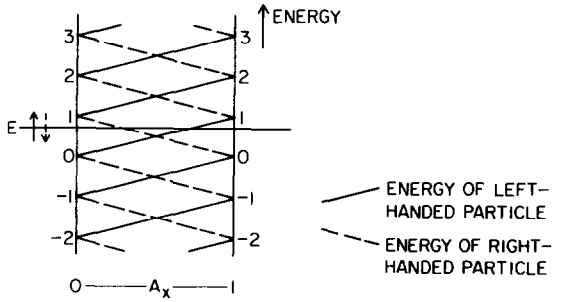

FIG. 2. The one-particle fermionic spectrum as a function of the background gauge potential A_x. The spectral flow at the arbitrarily chosen energy E is two.

However, as A_x increases from 0 to 1, the energies of the left-handed particles each increase by one while the energies of the right-handed particles decrease by one (see Fig. 2). The numbers of energy levels which pass through any given energy value, weighted by the direction of flow, are $+1$ and -1, respectively, for the left- and right-handed particles. The spectral flow [5] associated with the circle of one-dimensional Dirac operators parametrized by A_x is defined to be the difference of these numbers. It equals two.

The spectral flow induces a nontrivial periodicity on the amplitudes. To illustrate this, let us consider unexcited states. Let $\psi_{M,N}(A_x)$ denote the amplitude of the basis state in which left-handed particles of all momenta $\leqslant M$ and right-handed particles of all momenta $\geqslant N$ are present. The boundary conditions that should be imposed are

$$\psi_{M,N}(1) = \psi_{M+1,N+1}(0), \tag{3.15a}$$

$$\partial_{A_x}\psi_{M,N}(1) = \partial_{A_x}\psi_{M+1,N+1}(0). \tag{3.15b}$$

The topologically nontrivial gauge transformation $g(x) = \exp ix$ decreases A_x from 1 to 0, but at the same time it changes the eigenstate of the Dirac operator $\exp ipx$ to $\exp i(p+1)x$, thereby increasing the momenta of all particles by one. Equation (3.15a) therefore identifies the amplitudes for gauge equivalent states. More physically, the Fermi surface E_L^F (E_R^F) of a basis state increases (decreases) by one as A_x increases continuously from 0 to 1, because of the spectral flow, and Eq. (3.15a) identifies the amplitudes of basis states with the same Fermi surfaces. $\psi_{M,N}$ near $A_x = 1$ should actually join smoothly on to $\psi_{M+1,N+1}$ near $A_x = 0$, because there is nothing special about these values of A_x if $2\pi A_x$ is regarded as an angle. Equation (3.15b) matches first derivatives, and the Schrödinger equation will ensure that all higher derivatives match too.

The boundary conditions to be imposed on the amplitudes of arbitrary basis states are the obvious generalizations of (3.15). One must equate the amplitudes of gauge equivalent basis states at $A_x = 0$ and $A_x = 1$, bearing in mind that particle energies are globally gauge invariant, not particle momenta.

Let us now define the regularized chiral charges Q_L and Q_R of a basis state. Formally these charges are the eigenvalues of the charge operators

$$Q_L = \sum_p a^\dagger_{1,p} a_{1,p}, \qquad Q_R = \sum_p a^\dagger_{2,p} a_{2,p}, \tag{3.16}$$

which count the number of particles of each handedness. These numbers are infinite, of course, and the charges must be regularized. We shall use the method of heat-kernel regularization. We can simplify our discussion by again considering only unexcited states. Consider then a basis state $\psi_{M,N}$. Formally,

$$Q_L = \sum_{m \leqslant M} 1, \qquad Q_R = \sum_{n \geqslant N} 1. \tag{3.17}$$

The regularized charges are

$$Q^\lambda_L = \sum_{m \leqslant M} e^{\lambda(m + A_x)}, \qquad Q^\lambda_R = \sum_{n \geqslant N} e^{\lambda(-n - A_x)}, \tag{3.18}$$

where $\lambda > 0$. The exponential weights are gauge invariant, since they just depend on the particle energies. The regularized charges suppress the contribution of particles with large negative energies, but when $\lambda = 0$ they agree with the formal expressions (3.17).

Evaluating the sums (3.18) for small λ gives

$$Q^\lambda_L = \frac{1}{\lambda} + \left(M + A_x + \frac{1}{2}\right) + \frac{1}{2}\lambda \left(M + A_x + \frac{1}{2}\right)^2 - \frac{1}{24}\lambda + O(\lambda^2), \tag{3.19a}$$

$$Q^\lambda_R = \frac{1}{\lambda} + \left(-N - A_x + \frac{1}{2}\right) + \frac{1}{2}\lambda \left(-N - A_x + \frac{1}{2}\right)^2 - \frac{1}{24}\lambda + O(\lambda^2). \tag{3.19b}$$

Charge differences are well defined as $\lambda \to 0$, and we define the absolute regularized charges by subtracting the divergent constant λ^{-1}. Thus

$$Q^{\text{reg}}_L = M + A_x + \tfrac{1}{2} = E^F_L, \qquad Q^{\text{reg}}_R = -N - A_x + \tfrac{1}{2} = E^F_R. \tag{3.20}$$

For more general basis states, with excitations, the sums (3.18) are over all momenta m, n for which particles are present. It remains true that $Q^{\text{reg}}_L = E^F_L$ and $Q^{\text{reg}}_R = E^F_R$. Exciting a particle changes the energy but not the charges.

The reader may be surprised that these charges depend on the background field A_x, and are generally nonintegral. That this must be so is a consequence of the spectral flow. Q^{reg}_L (Q^{reg}_R) must increase (decrease) by one as A_x increases from 0 to 1, if gauge equivalent basis states are to have the same charges [cf. the boundary conditions (3.15)].

Recall that the electric charge $Q^{\text{reg}}_L + Q^{\text{reg}}_R$ of any physical state must be zero. This requires that

$$E^F_L = -E^F_R. \tag{3.21}$$

For unexcited states $N = M + 1$. Note that A_x cancels in these equations, so that there is a constraint on the basis states that can occur in a physical state, but no constraint on A_x. If A_x had been constrained, the theory would have been inconsistent (cf. Section 5).

The axial charge is

$$\begin{aligned} Q_5^{\text{reg}} &= Q_{\text{L}}^{\text{reg}} - Q_{\text{R}}^{\text{reg}} \\ &= 2E_{\text{L}}^{\text{F}}. \end{aligned} \tag{3.22}$$

For unexcited states

$$Q_5^{\text{reg}} = 2M + 2A_x + 1. \tag{3.23}$$

As A_x increases from 0 to 1, the axial charge changes by 2, as one left-handed particle is created and one right-handed particle is destroyed.

The $\frac{1}{2}$ that occurs in the definitions (3.20) of $Q_{\text{L}}^{\text{reg}}$ and $Q_{\text{R}}^{\text{reg}}$ may appear arbitrary, but is justified on symmetry grounds. One consequence is that an unexcited state with zero electric charge has zero momentum, because the momenta of the left- and right-handed particles together take all integer values precisely once. Furthermore, the momentum of an excited state, which is generally nonzero, is determined by the nature of the excitation, and is independent both of the value of A_x and of the location of the Fermi surfaces.

By considering its momentum, one can also see that a state with nonzero electric charge is unphysical. In an unexcited state of charge one, for example, the particle momenta take all integer values once, and one value occurs twice. The net momentum is this particular value, say, P. But it is not gauge invariant, since it can be changed to $P + 1$ by a topologically nontrivial gauge transformation. There is therefore a momentum anomaly, which is unacceptable.

Now, and for the remainder of this section, we shall make what may seem a drastic approximation. We shall suppose there are no fermionic excitations. We shall also ignore the Coulomb term in the Hamiltonian. It turns out that some of the states of the Schwinger model can be understood this way, namely, those involving physical scalar particles of zero momentum. The axial charge anomaly can also be understood. In Section 4, we shall allow for fermionic excitations, and show how the results obtained here fit into the complete theory.

If there are no excitations, and the electric charge is zero, the complete quantum state is specified by a set of wavefunctions $\{\psi_P(A_x): P \in \mathbf{Z}, 0 \leqslant A_x \leqslant 1\}$ subject to the boundary conditions

$$\psi_P(1) = \psi_{P+1}(0), \tag{3.24a}$$

$$\partial_{A_x}\psi_P(1) = \partial_{A_x}\psi_{P+1}(0). \tag{3.24b}$$

We have here introduced a new notation ψ_P for what was previously $\psi_{P,P+1}$. The evolution of ψ_P is determined by the Schrödinger equation

$$i\frac{\partial \psi_P}{\partial t} = \left(-\frac{e^2}{4\pi}\frac{\partial^2}{\partial A_x^2} + V_P^{\text{reg}}(A_x)\right)\psi_P, \tag{3.25}$$

where V_P^{reg} is the regularized kinetic energy of the fermions.

Formally, the kinetic energy of the fermions is

$$V_P(A_x) = \sum_{p \leqslant P} (p + A_x) + \sum_{p \geqslant P+1} (-p - A_x). \tag{3.26}$$

This is regularized in the same way that the charges Q_L and Q_R were. We define

$$V_P^\lambda(A_x) = \sum_{p \leqslant P} (p + A_x)\, e^{\lambda(p + A_x)} + \sum_{p \geqslant P+1} (-p - A_x)\, e^{\lambda(-p - A_x)}. \tag{3.27}$$

The sums can be evaluated by differentiating (3.19) with respect to λ, giving

$$V_P^\lambda(A_x) = -\frac{2}{\lambda^2} + \left(P + A_x + \frac{1}{2}\right)^2 - \frac{1}{12} + O(\lambda). \tag{3.28}$$

Subtracting the divergent constant, and taking the limit $\lambda \to 0$, one obtains the regularized energy

$$V_P^{\text{reg}}(A_x) = (P + A_x + \tfrac{1}{2})^2. \tag{3.29}$$

Note that $V_P^{\text{reg}}(A_x)$ near $A_x = 1$ smoothly joins on to $V_{P+1}^{\text{reg}}(A_x)$ near $A_x = 0$, which is what we expect from gauge invariance.

The expression (3.29) differs from the normal ordered energy, which is the total energy of the positive energy particles and positive energy antiparticles. Since there are finitely many such particles and antiparticles, and since each energy level varies linearly with A_x, the normal ordered energy is a linear approximation to the quadratic function (3.29). In this approximation, $V_p(1) = V_{P+1}(0)$, but $V_P(A_x)$ and $V_{P+1}(A_x)$ have different slopes. The smooth joining of the functions (3.29) is preferable because it respects the angular character of A_x. Physically, by adopting (3.29), we admit that changes of the energy levels in a filled Fermi–Dirac sea are important.

To solve the Schrödinger equations (3.25) it is convenient to work in an extended scheme. Let us define a potential $V(\tilde{A}_x)$ and wavefunction $\psi(\tilde{A}_x)$ over the entire interval $-\infty < \tilde{A}_x < \infty$, such that

$$V_P^{\text{reg}}(A_x) = V(A_x + P), \tag{3.30}$$

$$\psi_P(A_x) = \psi(A_x + P). \tag{3.31}$$

Then

$$V(\tilde{A}_x) = (\tilde{A}_x + \tfrac{1}{2})^2. \tag{3.32}$$

The Equations (3.25) reduce to the Schrödinger equation for a harmonic oscillator

$$i\frac{\partial \psi}{\partial t} = \left(-\frac{e^2}{4\pi}\frac{d^2}{d\tilde{A}_x^2} + \left(\tilde{A}_x + \frac{1}{2}\right)^2\right)\psi, \tag{3.33}$$

and the boundary conditions (3.24) are satisfied if $\psi(\tilde{A}_x)$ is everywhere smooth.

Technically, it is not surprising that the minimum of $V(\tilde{A}_x)$ occurs at $\tilde{A}_x = -\frac{1}{2}$, corresponding to $P = -1$, $A_x = \frac{1}{2}$. This is where the Fermi surfaces E_L^F and E_R^F are equal to zero. Here, too, the set of energy levels is $\{n + \frac{1}{2} : n \in \mathbf{Z}\}$ for both left- and right-handed fermions; all negative energy levels are filled and all positive energy levels are empty. Both charges Q_L^{reg} and Q_R^{reg} vanish as well, and this again justifies the $\frac{1}{2}$ in their definition.

Physically, it *is* surprising that the unexcited Fermi–Dirac sea has its lowest energy when $A_x = \frac{1}{2}$ and the phase factor (2.1) equals -1. However, when $A_x = 0$ there are two zero energy levels, one of each handedness, and a charge zero state of minimum energy is one where all negative energy levels, and just one zero energy level, are filled. This two-fold degeneracy suggests that a lower energy can be obtained by changing A_x, as is the case.

Stationary state wave functions of the harmonic oscillator (3.33) are well known. They are spread over the entire range of $\tilde{A}_x$, so the amplitude $\psi_P(A_x)$ is not identically zero for any value of P, although it is small if $|P|$ is large. The frequency of the oscillator is $e/\sqrt{\pi}$, so the spacing of the energy eigenvalues is $e/\sqrt{\pi}$ too.

Let us now recall the Hamiltonian $\mathscr{H}$ of a free, bosonic scalar field of mass m, on a circle of length 2π. In the momentum representation

$$\mathscr{H} = \frac{1}{2} \sum_p (\Pi^\dagger(p)\, \Pi(p) + (p^2 + m^2)\, \Phi^\dagger(p)\, \Phi(p)), \tag{3.34}$$

with

$$\begin{aligned} [\Phi(p), \Phi(p')] = [\Pi(p), \Pi(p')] &= 0, \\ [\Pi(-p), \Phi(p')] &= -i\delta_{pp'}, \\ \Pi(p) = \Pi^\dagger(-p), \qquad \Phi(p) &= \Phi^\dagger(-p), \end{aligned} \tag{3.35}$$

so that $\mathscr{H}$ describes an infinity of uncoupled harmonic oscillators. We shall establish in Section 4 that the Hamiltonian of the Schwinger model on a circle is equivalent to (3.34), with $m = e/\sqrt{\pi}$.

What we have shown so far, assuming that it is not misleading to ignore the Coulomb interaction, is that the states of the Schwinger model with no fermionic excitations correspond to multi-particle states of the free scalar field, where all particles have zero momentum. This is because the zero momentum field oscillator in $\mathscr{H}$ has frequency $e/\sqrt{\pi}$. It turns out, for reasons that will be clear later, that the correct way to identify the zero momentum scalar field operators with the operators of the Schwinger model is

$$\Pi(0) = \sqrt{2}\left(\tilde{A}_x + \frac{1}{2}\right), \qquad \Phi(0) = \frac{i}{\sqrt{2}} \frac{d}{d\tilde{A}_x}. \tag{3.36}$$

One might have guessed the conjugate pairing

$$\Pi(0) = -\frac{ie}{\sqrt{2\pi}} \frac{d}{d\tilde{A}_x}, \qquad \Phi(0) = \frac{\sqrt{2\pi}}{e}\left(\tilde{A}_x + \frac{1}{2}\right), \tag{3.37}$$

but this is wrong.[1]

Finally, let us look at the axial charge anomaly. The expectation value of Q_5^{reg} is

$$\langle Q_5^{\text{reg}} \rangle = \sum_P \int_0^1 (2P + 2A_x + 1)\, \psi_P^*(A_x)\, \psi_P(A_x)\, dA_x. \tag{3.38}$$

In general, this is not constant. With the time dependence of the wavefunctions determined by the Schrödinger equation (3.25), one deduces, after integrating by parts, that

$$\begin{aligned} \frac{d}{dt}\langle Q_5^{\text{reg}} \rangle &= \sum_P \int_0^1 \psi_P^*(A_x) \left(-\frac{ie^2}{\pi} \frac{d}{dA_x}\right) \psi_P(A_x)\, dA_x \\ &= 2\langle E_x^{\text{tr}} \rangle, \end{aligned} \tag{3.39}$$

where E_x^{tr} is the transverse electric field operator. Surface terms at $A_x = 0$ and $A_x = 1$ cancel, because of the boundary conditions (3.24). Equation (3.39) is the integrated form of the anomaly equation (1.1).

In the extended scheme

$$\begin{aligned} \langle Q_5^{\text{reg}} \rangle &= \int_{-\infty}^{\infty} (2\tilde{A}_x + 1)\, \psi^*(\tilde{A}_x)\, \psi(\tilde{A}_x)\, d\tilde{A}_x \\ &= \langle 2\tilde{A}_x + 1 \rangle, \end{aligned} \tag{3.40}$$

and (3.39) is an immediate consequence of the Heisenberg operator equation

$$\frac{d}{dt}\tilde{A}_x = i\left[-\frac{e^2}{4\pi}\frac{d^2}{d\tilde{A}_x^2} + V(\tilde{A}_x), \tilde{A}_x\right] = -\frac{ie^2}{2\pi}\frac{d}{d\tilde{A}_x} = 2E_x^{\text{tr}}. \tag{3.41}$$

(Note that d/dA_x and $d/d\tilde{A}_x$ are the same operators, even though A_x and $\tilde{A}_x$ only agree mod 1.)

It is often claimed that

$$\frac{d}{dt}\langle A_x \rangle = \langle E_x^{\text{tr}} \rangle, \tag{3.42}$$

implying that the charge $Q_5' = Q_5^{\text{reg}} - 2A_x$ is conserved, albeit not globally gauge invariant. In fact,

$$\langle Q_5' \rangle = \sum_P \int_0^1 (2P + 1)\, \psi_P^*(A_x)\, \psi_P(A_x)\, dA_x, \tag{3.43}$$

[1] I am grateful to Sidney Coleman for alerting me to this.

so Q_5' is the normal ordered version of the axial charge, which has integer eigenvalues. Were it not for the boundary conditions (3.24), Q_5' would be conserved, because the Schrödinger equation (3.25) does not explicitly mix the wave function ψ_P with $\psi_{P'}$, for any $P' \neq P$. However, this time the surface terms do not cancel, and one finds

$$\frac{d}{dt}\langle Q_5' \rangle = \frac{ie^2}{4\pi}\left[\sum_P (2P+1)\left(\psi_P^* \frac{d\psi_P}{dA_x} - \frac{d\psi_P^*}{dA_x}\psi_P\right)\right]_0^1$$
$$= -\frac{ie^2}{2\pi}\sum_P \left(\psi_P^* \frac{d\psi_P}{dA_x} - \frac{d\psi_P^*}{dA_x}\psi_P\right)\Big|_{A_x=0}. \tag{3.44}$$

Equation (3.42) is incorrect because $2\pi A_x$ is an angle. There is a surface correction which is half the right-hand side of (3.44). In the extended scheme, it is true that

$$\frac{d}{dt}\langle \tilde{A}_x \rangle = \langle E_x^{\mathrm{tr}} \rangle, \tag{3.45}$$

but to say that $Q_5^{\mathrm{reg}} - 2\tilde{A}_x$ is conserved is a triviality because $\tilde{A}_x$ is defined in terms of Q_5^{reg}.

Our conclusion is that neither Q_5^{reg} not Q_5' is conserved. Actually, we already knew that Q_5' was not conserved because the stationary states of the harmonic oscillator involve all values of P, so that they are not eigenstates of Q_5'.

4. Bosonization

A complete treatment of the Schwinger model requires a consideration of excited states of the fermions, and then one must work with the entire Hamiltonian (3.14). We shall show that this Hamiltonian can be remarkably simplified by the technique of bosonization. We shall also show how, in principle, one may obtain the exact stationary states of the model. The material of this section is not essentially new, but there are some technical differences (improvements) between our presentation and that of the literature.

Because space is a circle of length 2π, and therefore momenta are integers, we shall employ the boson operators used most frequently to solve one-dimensional condensed matter problems. These operators are basic in Mattis and Lieb's solution [10] of the Luttinger model, work that has been fully reviewed by T. Bohr [11]. For a more field theoretical approach to bosonization, see Refs. [3].

The bosonic operators are the momentum space, chiral charge density operators defined earlier

$$\rho_\alpha(p) = \sum_k a_{\alpha,k+p}^\dagger a_{\alpha,k}, \qquad \alpha = 1, 2, \qquad p \in \mathbf{Z}, \tag{4.1}$$

which have the hermiticity property

$$\rho_\alpha^\dagger(p) = \rho_\alpha(-p). \tag{4.2}$$

We shall only need these operators for $p \neq 0$.

Let us now recall the crucial fact that a physical state is a linear combination of physical fermion basis states in which all but a finite number of negative energy levels are filled and all but a finite number of positive energy levels are empty. Acting on such states, the operators ρ_α have the commutation relations

$$[\rho_1(-p), \rho_1(p')] = p\delta_{pp'}1, \tag{4.3a}$$

$$[\rho_2(-p), \rho_2(p')] = -p\delta_{pp'}1, \tag{4.3b}$$

$$[\rho_1(-p), \rho_2(p')] = 0. \tag{4.3c}$$

These are demonstrated in Appendix A. It follows from the arguments given there that for $p \neq 0$, $\rho_\alpha(p)$ is a finite operator, unlike the charges $Q_L = \rho_1(0)$ and $Q_R = \rho_2(0)$, which had to be regularized.

Let us next define, for $p \neq 0$,

$$\Phi(p) = -\frac{1}{\sqrt{2}ip}(\rho_1(p) + \rho_2(p)), \tag{4.4}$$

$$\Pi(p) = \frac{1}{\sqrt{2}}(\rho_1(p) - \rho_2(p)). \tag{4.5}$$

These operators satisfy the canonical Bose commutation relations and hermiticity properties (3.35). Clearly

$$j^0(p) = j_5^1(p) = \sqrt{2}ip\Phi(p), \tag{4.6a}$$

$$j^1(p) = j_5^0(p) = \sqrt{2}\Pi(p). \tag{4.6b}$$

Now we are ready to give the bosonized form of the operators which occur in the Hamiltonian (3.14). On substituting for $j^0(p)$, the Coulomb term becomes immediately

$$\frac{e^2}{2\pi}\sum_{p\neq 0}\Phi^\dagger(p)\,\Phi(p). \tag{4.7}$$

The fermion kinetic energy has two contributions. One is the energy of the unexcited Fermi–Dirac sea. This depends on $\tilde{A}_x$ and is what we computed in the last section. The other contribution is the excitation energy. The operator measuring this has the bosonized form

$$\frac{1}{2}\sum_{p\neq 0}(\Pi^\dagger(p)\,\Pi(p) + p^2\Phi^\dagger(p)\,\Phi(p)) + \text{infinite constant}. \tag{4.8}$$

Verifying this claim requires a number of steps. First, observe that in terms of the operators ρ_α, (4.8) becomes

$$\frac{1}{2}\sum_{p\neq 0}\rho_1(p)\rho_1(-p)+\frac{1}{2}\sum_{p\neq 0}\rho_2(-p)\rho_2(p)+\text{infinite constant.} \tag{4.9}$$

Using the commutation relations (4.3), this can be rewritten as $K_L + K_R$, where

$$K_L=\sum_{p>0}\rho_1(p)\rho_1(-p), \tag{4.10a}$$

$$K_R=\sum_{p>0}\rho_2(-p)\rho_2(p). \tag{4.10b}$$

The infinite constant has been arranged to cancel here. The ordering of the operators in K_L and K_R is desirable, because for $p \gg 0$ both $\rho_1(-p)$ and $\rho_2(p)$ annihilate physical basis states.

Let us consider next the action of K_L on a physical basis state. Only the momenta of the left-handed particles matter here. The operator $\rho_1(p)\rho_1(-p)$ lowers one particle's momentum by p, then raises another's by p. Suppose there are two distinct pairs of momenta $r+t$, r and $s+t$, s, with $r>s$ and $t>0$, such that there are particles with momenta $r+t$ and s present, and no particles with momenta r and $s+t$. Then the term in K_L (with $p=t$)

$$a^\dagger_{1,s+t}a_{1,s}a^\dagger_{1,r}a_{1,r+t} \tag{4.11}$$

exchanges the filled and empty levels, but so also does the term (with $p=r-s$)

$$a^\dagger_{1,r}a_{1,s}a^\dagger_{1,s+t}a_{1,r+t}. \tag{4.12}$$

The sum of (4.11) and (4.12) vanishes, because of the anti-commutation relations (3.12). A similar argument applies if $r<s$. So after all, K_L does not change the momenta of any particles. Its only nontrivial action is through terms of the form

$$a^\dagger_{1,r+p}a_{1,r}a^\dagger_{1,r}a_{1,r+p}, \qquad p>0, \tag{4.13}$$

which acts as a unit operator when there is a particle of momentum $r+p$ present, but no particle of momentum r. Otherwise this operator annihilates the state.

We conclude that on any physical basis state, K_L acts as a constant N_L times the unit operator, where N_L is the number of pairs of energy levels in which the upper level is occupied and the lower is empty. In fact, N_L is the excitation energy. For example, the excited basis state whose level occupation is shown in Fig. 1a has an excitation energy of 10 units, and N_L is also 10. To see that the excitation energy always equals N_L, consider raising any particle by one energy unit. This creates immediately one extra pair of levels where the upper is occupied and the lower is empty, but apart from this one the number of such pairs is unchanged. Existing pairs involving the previously filled (empty) level are replaced by pairs involving the newly

filled (empty) level. Any state can be obtained by exciting particles by one energy step at a time, and for each step N_L increases by one. Since in the unexcited state N_L is zero, it follows that N_L is the excitation energy.

This completes our argument that K_L measures the excitation energy of the left-handed particles. K_R similarly measures the excitation energy of the right-handed particles. (The reverse order of the operators reflects the fact that for right-handed particles, increasing momentum implies decreasing energy.)

Putting it all together, we find that the bosonized form of the Hamiltonian (3.14) is

$$H = -\frac{e^2}{4\pi}\frac{\partial^2}{\partial \tilde{A}_x^2} + \left(\tilde{A}_x + \frac{1}{2}\right)^2 + \frac{1}{2}\sum_{p\neq 0}\left(\Pi^\dagger(p)\,\Pi(p) + (p^2+m^2)\,\Phi^\dagger(p)\,\Phi(p)\right) + \text{infinite constant.} \tag{4.14}$$

If we now make the identification of the zero momentum field operators given by (3.36), then (apart from an irrelevant constant) we obtain the canonical Hamiltonian of a free massive scalar field (3.34).

It is now clear why we should identify $\Pi(0)$ with $\sqrt{2}(\tilde{A}_x + \frac{1}{2})$. The definition (4.5) of $\Pi(p)$ extends to $p = 0$, becoming

$$\Pi(0) = \frac{1}{\sqrt{2}}\,Q_5 \tag{4.15}$$

and the regularized axial charge Q_5^{reg} is precisely $2\tilde{A}_x + 1$.

The definition of $\Phi(0)$ also makes sense. Note that for $p \neq 0$

$$\Phi(p) = -\frac{1}{\sqrt{2}e^2}\,E_x^{\text{long}}(p), \tag{4.16a}$$

where $E_x^{\text{long}}(p)$ is the Fourier transform of the longitudinal part of the electric field $-\partial_x A_t$. This follows from Gauss' law (3.3). On the other hand, the definition (3.36) of $\Phi(0)$ is equivalent to

$$\Phi(0) = -\frac{1}{\sqrt{2}e^2}\,E_x^{\text{tr}}(0), \tag{4.16b}$$

where $E_x^{\text{tr}}(0) = 2\pi E_x^{\text{tr}}$ (i.e., the zero momentum Fourier transform of the constant transverse electric field).

Now, let us consider again the currents j^μ and j_5^μ. The Heisenberg equations for the free scalar field are

$$\frac{d}{dt}\Phi(p) = \Pi(p), \tag{4.17}$$

and

$$\frac{d}{dt}\Pi(p) = (-p^2 - m^2)\,\Phi(p). \tag{4.18}$$

Expressing $\Phi(p)$ and $\Pi(p)$ in terms of current components, using (4.6), Eq. (4.17) becomes

$$\frac{d}{dt} j^0(p) - ipj^1(p) = 0, \tag{4.19}$$

which is equivalent to the conservation law $\partial_\mu j^\mu = 0$. On the other hand, using (4.16) as well, Eq. (4.18) becomes

$$\frac{d}{dt} j_5^0(p) - ipj_5^1(p) = \frac{1}{\pi} E_x^{\text{total}}(p). \tag{4.20}$$

Taking the expectation value of the $p = 0$ term here reproduces the axial charge anomaly, Eq. (3.39). But now we have more, because the Fourier transform of (4.20) is the complete anomaly equation

$$\partial_\mu j_5^\mu = \frac{1}{\pi} E_x. \tag{4.21}$$

Finally, let us re-examine the unexcited states that we constructed in the last section. Their energies are correct, because the zero momentum part of the bosonized Hamiltonian is decoupled from the nonzero momentum part. However, they are not true stationary states of the full Hamiltonian. Since they have no fermion excitations, they are annihilated by $\rho_1(-p)$ and $\rho_2(p)$ for all $p > 0$. Equivalently, because of (4.4) and (4.5), they are annihilated by the operator

$$\Pi(p) - i|p|\Phi(p), \qquad p \neq 0, \tag{4.22}$$

which is proportional to the annihilation operator for a *zero mass* scalar field. A true stationary state with only zero momentum scalar particles is annihilated by

$$\Pi(p) - i\sqrt{p^2 + m^2}\Phi(p), \qquad p \neq 0. \tag{4.23}$$

Transforming from one type of state to the other requires a Bogoliubov transformation.

5. Axial Electrodynamics and Related Models

Here we show that certain models similar to the Schwinger model, of which chiral electrodynamics is an example, do not exist as well-defined quantum field theories. But axial electrodynamics is well defined. For an alternative derivation of these results, see Refs. [12]. We continue to suppose that space is a circle.

The Schwinger model has two independent massless Fermi fields, one left-handed and the other right-handed, both of charge 1. It is straightforward to write down a

model with n_L left-handed and n_R right-handed Fermi fields, having integer charges $\{y_L^{(r)}: 1 \leqslant r \leqslant n_L; y_R^{(s)}: 1 \leqslant s \leqslant n_R\}$. The Lagrangian density is

$$L = \frac{1}{2e^2}(\partial_t A_x - \partial_x A_t)^2 + \sum_r \bar{\psi}_L^{(r)} i\gamma^\mu (\partial_\mu + iy_L^{(r)} A_\mu)\psi_L^{(r)} + \sum_s \bar{\psi}_R^{(s)} i\gamma^\mu (\partial_\mu + iy_R^{(s)} A_\mu)\psi_R^{(s)}. \tag{5.1}$$

$\psi_L^{(r)}$ ($\psi_R^{(s)}$) is a 2-component spinor, in which the lower (upper) component vanishes identically.

A special case of (5.1) is axial electrodynamics, where $n_L = n_R = 1$ and $y_L^{(1)} = 1$, $y_R^{(1)} = -1$. Another is chiral electrodynamics, which has just a single left-handed field of charge 1.

Formally, one can derive a Hamiltonian just as for the Schwinger model, working again in the gauge where $\partial_x A_x = 0$ and $0 \leqslant A_x \leqslant 1$. However, there is a problem. To satisfy Gauss' law, the total electric charge must be zero. Now we know that the regularized charges of any fermionic state with prescribed occupation of the levels varies with A_x, so as to be gloally gauge invariant. The generalized form of the charges (3.20) is

$$Q_L^{(r)\text{reg}} = (M^{(r)} + y_L^{(r)} A_x + \tfrac{1}{2})\, y_L^{(r)}, \tag{5.2a}$$

$$Q_R^{(s)\text{reg}} = (-N^{(s)} - y_R^{(s)} A_x + \tfrac{1}{2})\, y_R^{(s)}. \tag{5.2b}$$

For the total electric charge to be zero, independently of A_x, we must have

$$\sum_r \left(M^{(r)} + \frac{1}{2}\right) y_L^{(r)} = \sum_s \left(N^{(s)} - \frac{1}{2}\right) y_R^{(s)}, \tag{5.3}$$

and

$$\sum_r (y_L^{(r)})^2 = \sum_s (y_R^{(s)})^2. \tag{5.4}$$

While the first equation is easy to satisfy by an appropriate choice for the integers $M^{(r)}$ and $N^{(s)}$, the second is a constraint on the charges of the particles in the model. The Schwinger model satisfies the constraint, and so does axial electrodynamics, but chiral electrodynamics does not.

The inconsistency of models which do not satisfy (5.4) was understood before in a more general, but less direct way, as a consequence of the gauge potential being coupled to an anomalous electric current [12]. Let $A_x(t)$ be a background, x-independent gauge potential, so $E_x = \dot{A}_x$ is the background electric field. Then directly from (5.2)

$$\frac{d}{dt}\langle Q\rangle = E_x \left(\sum_r (y_L^{(r)})^2 - \sum_s (y_R^{(s)})^2\right), \tag{5.5}$$

where Q is the total electric charge. Equation (5.5) is true not only when the fermionic state changes adiabatically, that is, with unchanging occupation of the energy levels, but also if particles are excited, since such excitations do not change the total charge. Clearly, if (5.4) is not satisfied, then the total charge is not conserved. So when there is a problem with Gauss' law, then the electric current is anomalous. Our argument that (5.4) must be satisfied in a sensible model is therefore consistent with the earlier results.

If Eqs. (5.3) and (5.4) are satisfied, then there is no momentum anomaly; unexcited states which are connected by the spectral flow and by gauge transformations have the same momentum. Moreover, in axial electrodynamics, the total momentum of an unexcited state is zero, as in the Schwinger model. The accounting is a little different in the two cases, though. In fact, an analysis of axial electrodynamics along the lines given here shows that it is dynamically equivalent to the Schwinger model, but the roles of the axial and vector currents are reversed. In axial electrodynamics, it is the vector current j^μ, and therefore also the fermion number, that is not conserved.

6. The Irrelevant Θ-Parameter in the Schwinger Model [13]

In pure electrodynamics on a circle, recall that the Hamiltonian is

$$H = \frac{\pi}{e^2} E_x^2 \tag{6.1}$$

and the wavefunction $\psi(A_x)$ must be periodic, with period 1. But there is not a unique canonical quantization, because of the arbitrary real parameter Θ in the representation (2.8) of the electric field operator E_x. For a general value of Θ, the Hamiltonian (6.1) has eigenvalues $\{\pi e^2(n + \Theta/2\pi)^2 : n \in \mathbf{Z}\}$, so the fractional part of $\Theta/2\pi$ is physically significant. The integer part can be removed by a gauge transformation.

Θ can be entirely removed from the Hamiltonian at the expense of changing the boundary conditions. Let

$$\psi'(A_x) = e^{i\Theta A_x}\psi(A_x). \tag{6.2}$$

This would be a gauge transformation if $\Theta/2\pi$ were an integer. The Schrödinger equation for ψ' is

$$i\frac{\partial\psi'}{\partial t} = -\frac{e^2}{4\pi}\frac{\partial^2\psi'}{\partial A_x^2} \tag{6.3}$$

and ψ' obeys the boundary condition

$$\psi'(1) = e^{i\Theta}\psi'(0), \tag{6.4}$$

so Θ does not disappear, unless $\Theta/2\pi$ is an integer.

In the Schwinger model, Θ is no longer a relevant parameter. To see this, it is sufficient to consider again unexcited states, since excitations are insensitive to the transverse electric field. When Θ is nonzero, the Schrödinger equation (3.25) is replaced by

$$i\frac{\partial \psi_P}{\partial t} = \left(-\frac{e^2}{4\pi}\left(\frac{\partial}{\partial A_x} + i\Theta\right)^2 + V_P^{\mathrm{reg}}(A_x)\right)\psi_P \tag{6.5}$$

but the boundary conditions (3.24) are unchanged. Let now

$$\psi'_P(A_x) = e^{i\Theta(P+A_x)}\psi_P(A_x). \tag{6.6}$$

As above, changing to the primed wavefunctions removes Θ from the Hamiltonian. In addition, because of the extra factor of $\exp iP\Theta$, the primed wavefunctions obey the original boundary conditions

$$\psi'_P(1) = \psi'_{P+1}(0). \tag{6.7}$$

So a change of variables removes Θ. Since $P + A_x = \frac{1}{2}(Q_5^{\mathrm{reg}} - 1)$, Eq. (6.6) is formally a chiral transformation, and one sees the intimate connection between the irrelevance of the Θ-parameter and the existence of the axial anomaly. Both follow from the boundary conditions (3.24) or, more fundamentally, from the spectral flow.

7. Discussion

When space is a circle, what we have discovered is an essential topological difference between pure electrodynamics and the Schwinger model. In pure electrodynamics, the configuration space is itself a circle, since the gauge field is specified, up to gauge equivalence, by the phase factor (2.1), and the wavefunction must be single-valued on this circle. In the Schwinger model, on the other hand, the circle of Dirac operators associated with this configuration space has a nonzero spectral flow. The immediate consequence is that the total wavefunction of a quantum state is no longer strictly periodic, but periodic "with a twist," as one orbits the configuration space. [For unexcited states the relevant boundary conditions are Eqs. (3.24).] This means that the configuration space is effectively replaced by its covering space—the line— when fermions are present. States of the Schwinger model are best described in an extended scheme, using the linear variable $\tilde{A}_x$, rather than the angular variable A_x. It is this change of topology that is responsible for the axial anomaly, and for the irrelevance of the Θ-parameter.

A particularly interesting dynamical property of the Schwinger model on a circle, of which only a vestige remains in the usual version of the model, is that the semi-classical vacuum, the gauge field that minimizes the regularized energy of the Fermi–Dirac negative energy sea, occurs when the phase factor (2.1) equals -1. It is unlikely one could have discovered this using perturbative methods.

Let us conclude by reviewing some well-known aspects of quantum chromodynamics (QCD) with massless quarks in three spatial dimensions, in the light of our Schwinger model results [14]. In QCD there are an infinity of physical degrees of freedom in the pure Yang–Mills field, and one of these is the phase factor

$$X = \exp -\frac{i}{8\pi}\int \varepsilon_{ijk}\,\mathrm{Tr}\left(F_{ij}A_k - \frac{2}{3}A_iA_jA_k\right) d^3x. \tag{7.1}$$

The integrand here is the Chern–Simons 3-form [15]. X plays a similar role to the phase factor (2.1) in the Schwinger model. This is not surprising, since A_x, the integrand there, is the abelian Chern–Simons 1-form. X is gauge invariant, because under a gauge transformation of topological winding number n, the phase of X changes by $2\pi n$. Topologically trivial gauge transformations do not change the integral of the Chern–Simons form at all. A closed loop in the Yang–Mills configuration space is noncontractible if (and only if) the phase of X changes (continuously) by $2\pi n$ along it, where $n \neq 0$.

An instanton is a gauge field which traverses, with Euclidean time as the parameter, such a noncontractible loop. This follows from the defining property of an instanton

$$-\frac{1}{16\pi}\int \varepsilon_{\mu\nu\sigma\tau}\,\mathrm{Tr}(F_{\mu\nu}F_{\sigma\tau})\,d^3x\,dt = 2\pi n, \qquad n \neq 0, \tag{7.2}$$

and from the identity

$$\frac{d}{dt}\int \varepsilon_{ijk}\,\mathrm{Tr}\left(F_{ij}A_k - \frac{2}{3}A_iA_jA_k\right) d^3x = \frac{1}{2}\int \varepsilon_{\mu\nu\sigma\tau}\,\mathrm{Tr}(F_{\mu\nu}F_{\sigma\tau})\,d^3x, \tag{7.3}$$

which is true provided the field tensor F decays sufficiently rapidly at spatial infinity, and one chooses a gauge where $A_t = 0$ there too.

In massless QCD, as in the Schwinger model, it is known that the gauge covariant three-dimensional Dirac operator has a nonvanishing spectral flow along noncontractible loops [5], so again the presence of fermions changes the topology, and one must pass to the simply connected covering of the Yang–Mills configuration space. As in the Schwinger model, this change of topology implies that the Θ-parameter of massless QCD has no physical effect [14], and that there is an axial anomaly. This time the anomaly equation is [16]

$$\partial_\mu j_5^\mu = -\frac{N_f}{16\pi^2}\,\varepsilon_{\mu\nu\sigma\tau}\,\mathrm{Tr}(F_{\mu\nu}F_{\sigma\tau}), \tag{7.4}$$

where j_5^μ is the color singlet axial current, and N_f is the number of quark flavors.

It would be very interesting to know the semi-classical vacuum of massless QCD, that is, the background static gauge potential for which the sum of the classical potential energy (the color magnetic field energy) and the regularized energy of the

Fermi–Dirac sea in that background field is minimized. Fluctuations of the gauge potential around this vacuum, and fermionic excitations, could perhaps be treated perturbatively. The observed properties of quarks—their masses in particular—might be a dynamical consequence of this vacuum being nontrivial.

There are a number of reasons why the three-dimensional theory is harder to study than the Schwinger model. As already mentioned, the Yang–Mills configuration space is much more complicated, and little is known about the spectrum of the gauge covariant Dirac operator which occurs in the Hamiltonian. Moreover, the eigenstates of the Dirac operator depend on the background gauge field, whereas in the Schwinger model (in the Coulomb gauge) only the eigenvalues did. It is not clear whether it is feasible to work with a fermionic Fock space constructed from such energy eigenstates, which vary with the background gauge field, rather than the usual Fock space constructed from unvarying plane wave states. Another problem in three dimensions is that the fermionic excitations cannot be handled by the bosonization technique described earlier.

Despite these difficulties, we hope that the ideas and results presented here can be used to gain a better understanding of QCD.

APPENDIX A: A BOSONIZATION LEMMA

Here we establish the commutation relations (4.3). Let p, p' be fixed nonzero momenta. Suppose also a physical basis state $|\varphi\rangle$ is given. We shall show that

$$[\rho_1(-p), \rho_1(p')]\,|\varphi\rangle = p\delta_{pp'}\,|\varphi\rangle, \tag{A.1}$$

where

$$\rho_1(p) = \sum_{k=-\infty}^{\infty} a^\dagger_{1,k+p} a_{1,k}. \tag{A.2}$$

We drop the subscript 1 in the following.

The idea is that for $|k|$ sufficiently large, $a^\dagger_{k+p}a_k$ annihilates $|\varphi\rangle$, because left-handed particles of momenta k and $k+p$ are either both present ($k \ll 0$) or both absent ($k \gg 0$). One can therefore restrict the range of summation in ρ and compute

$$\left[\sum_{k=-K}^{K} a^\dagger_{k-p}a_k, \sum_{l=-L}^{L} a^\dagger_{l+p'}a_l\right]|\varphi\rangle, \tag{A.3}$$

where K and L are both large. It is convenient technically to have $K \geqslant L + |p'| + |p|$. Now, using the anti-commutation relations (3.12), the expression above reduces to

$$\sum_{k=-K}^{K}\sum_{l=-L}^{L} (\delta_{k,l+p'}a^\dagger_{k-p}a_l - \delta_{k-p,l}a^\dagger_{l+p'}a_k)\,|\varphi\rangle. \tag{A.4}$$

Summing over k, this becomes

$$\left(\sum_{l=-L}^{L} a^{\dagger}_{l+p'-p} a_l - \sum_{l=-L}^{L} a^{\dagger}_{l+p'} a_{l+p} \right) |\varphi\rangle, \tag{A.5}$$

which equals

$$\left(\sum_{l=-L}^{L} a^{\dagger}_{l+p'-p} a_l - \sum_{l=-L+p}^{L+p} a^{\dagger}_{l+p'-p} a_l \right) |\varphi\rangle. \tag{A.6}$$

Only the terms in these sums near $\pm L$ do not cancel. Particles of all momenta near $-L$ are present, but there are none with momenta near L. Therefore, if $p' \neq p$ then (A.6) vanishes, and if $p' = p$ then (A.6) equals $p\,|\varphi\rangle$. This establishes (A.1). Note that (A.5) is ambiguous if L is infinite, and this is why one must be careful to specify the type of states on which $\rho(p)$ acts.

Equation (4.3a) follows because any physical state is a linear combination of basis states for which (A.1) is true. Equation (4.3b) is established similarly. The opposite sign occurs there because the right-handed Fermi–Dirac sea has particles of large positive momenta. Equation (4.3c) is true, because the analog of (A.4) vanishes.

APPENDIX B: AN ANOMALY IN A QUANTUM MECHANICAL SYSTEM

Consider a particle of unit mass and unit charge moving on a 2-torus of unit dimensions, and coupled to an abelian gauge field. The following gauge invariant data specifies the field: (1) the field strength, which acts as a magnetic field normal to the torus and whose total flux must be an integer multiple of 2π—the same flux as is emitted by an elementary Dirac monopole; (2) the holonomy (Wilson loop variable) on any two closed curves which are topologically equivalent to the two orthogonal circles which generate the torus—these curves are not the boundaries of any region of the torus. The holonomy on any other closed curve is determined using this data and Stokes' theorem.

Let us introduce coordinates $\{(x, y)\colon -\delta \leqslant x \leqslant 1 + \delta,\ 0 \leqslant y \leqslant 1\}$ on the torus. Identifying $y = 1$ with $y = 0$ gives a cylinder, and then identifying x with $1 + x$ $(-\delta \leqslant x \leqslant \delta)$ joins the ends of the cylinder, with an overlap, and gives the torus.

We select a gauge field with uniform magnetic field $B = 2\pi$, and holonomy equal to one on the two circles $x = 0$ and $y = 0$. Because the total flux is not zero, this is a gauge field on a nontrivial $U(1)$ bundle over the torus. To specify the bundle [17], we define the gauge transformation $g(x, y) = \exp - 2\pi i y$ to act as a $U(1)$ transition function between the overlapping cylinder ends. One choice of gauge potential for the field of interest is then

$$a_x = 0, \qquad a_y = 2\pi x. \tag{B.1}$$

The gauge potentials on the overlapping regions are correctly related by the gauge transformation g, since

$$a_\mu(1+x, y) = a_\mu(x, y) + i((\partial_\mu g)\, g^{-1})(x, y). \tag{B.2}$$

We shall suppose from now on that δ is infinitesimally small and effectively zero, and that g is effectively defined on the circle $x = 0$. This will cause no problems.

The holonomy on circles in the y-direction is

$$\begin{aligned} W_y(x) &\equiv \exp i \int_0^1 a_y \, dy \\ &= \exp 2\pi i x, \end{aligned} \tag{B.3}$$

and in the x-direction,

$$\begin{aligned} W_x(y) &\equiv \left(\exp i \int_0^1 a_x dx\right) g(y) \\ &= \exp - 2\pi i y. \end{aligned} \tag{B.4}$$

So $W_y(0)$ and $W_x(0)$ are both equal to one, as desired. The need for the extra factor $g(y)$ is explained in Ref. [17]. The difference in the phases of $W_y(x_1)$ and $W_y(x_2)$ is the flux through the cylinder bounded by the circles $x = x_1$ and $x = x_2$, and $W_y(x)$ is single-valued, because the total flux is quantized.

Let us now look for the stationary states of this system, by solving the Schrödinger equation

$$i \frac{\partial \psi}{\partial t} = H\psi \tag{B.5}$$

with Hamiltonian

$$H = -\frac{1}{2}\frac{\partial}{\partial x^2} - \frac{1}{2}\left(\frac{\partial}{\partial y} + 2\pi i x\right)^2. \tag{B.6}$$

Since a gauge transformation is needed to connect $x = 0$ with $x = 1$ the wavefunction is not single-valued in x, although it is in y. It must satisfy the boundary conditions

$$\psi(1, y, t) = e^{-2\pi i y}\psi(0, y, t), \tag{B.7a}$$

$$\partial_x \psi(1, y, t) = e^{-2\pi i y}\partial_x \psi(0, y, t). \tag{B.7b}$$

The Schrödinger equation separates if we write

$$\psi(x, y, t) = \sum_{n=-\infty}^{\infty} \psi_n(x, t)\, e^{2\pi i n y}. \tag{B.8}$$

The functions ψ_n satisfy the equations

$$i\frac{\partial\psi_n}{\partial t} = -\frac{1}{2}\frac{\partial^2\psi_n}{\partial x^2} + 2\pi^2(n+x)^2\,\psi_n, \tag{B.9}$$

$$\psi_{n+1}(0,t) = \psi_n(1,t), \tag{B.10a}$$

$$\partial_x\psi_{n+1}(0,t) = \partial_x\psi_n(1,t). \tag{B.10b}$$

These equations are equivalent to (3.24), (3.25), and (3.29) above, and can be solved in the same way. All stationary states are obtained by taking a stationary state u of the harmonic oscillator of frequency 2π, satisfying the equation

$$-\frac{1}{2}\frac{d^2u}{dx^2} + 2\pi^2x^2u = Eu, \qquad -\infty < x < \infty, \tag{B.11}$$

and setting $\psi_n(x,t) = e^{-iEt}u(n+x)$. The energy spectrum of the particle on the torus is therefore $\{2\pi(m+\frac{1}{2}): m \geqslant 0\}$. The ground state wavefunction is

$$\begin{aligned}\psi(x,y,t) &= Ne^{-\pi it}\sum_{n=-\infty}^{\infty} e^{-\pi(n+x)^2+2\pi iny} \\ &= Ne^{-\pi x^2-i\pi t}\Theta(-x+iy;1),\end{aligned} \tag{B.12}$$

where Θ is the third Jacobi theta function, and N is a normalization constant.

A comparison with the Landau levels of the same particle moving in the same magnetic field, but in a plane, is interesting [18]. There the energy levels are also those of a harmonic oscillator of frequency 2π but each level has a large multiplicity. If the particle is restricted to a large rectangle of area A then the multiplicity is estimated to be A. We see that when $A = 1$ and boundary conditions appropriate to a torus are imposed, then the multiplicity is exactly 1. In their study of the QCD vacuum, the Copenhagen group have made important use of quantum mechanical eigenstates for particles in background magnetic fields and have also been led to the wavefunction (B.12) (See, e.g., Refs. [19, 20].) Reference [20] in particular contains some useful information on the theta function. The Landau levels for a particle moving on a compact 2-manifold, for example, on a sphere or a torus, are also important in studies of the quantized Hall effect [21].

Less important than the solution of our model is the subtle separation of variables (B.8) that makes it possible, so let us review that. Because the momentum operator $p_y = -i\partial/\partial y$ commutes with the Hamiltonian, it is natural to expand ψ in eigenstates $e^{2\pi iny}$ of this operator. But the eigenstates of H are not eigenstates of p_y because of the boundary conditions (B.10). The underlying reason for these boundary conditions is a spectral permutation. The second term in the Hamiltonian

$$-\frac{1}{2}\left(\frac{\partial}{\partial y} + 2\pi ix\right)^2 \tag{B.13}$$

is an operator in which x appears as a parameter, and whose spectrum is $\{2\pi^2(n+x)^2 : n \in \mathbf{Z}\}$. If we follow one eigenvalue continuously from $x=0$ to $x=1$, it does not return to its original value. The spectrum is permuted by the operation of circling the torus in the x-direction. Consequently, the eigenfunction with eigenvalue $2\pi^2 n^2$ at $x=1$, namely, $\exp 2\pi i(n-1)y$, should be identified with the eigenfunction with the same eigenvalue at $x=0$, namely, $\exp 2\pi i n y$, and this is precisely the identification made by the gauge transformation $g(y)$. The boundary conditions represent the identification of the coefficients of these eigenfunctions.

A closely related consequence of the spectral permutation is that p_y is not conserved, even though it commutes with the Hamiltonian. Consider the partially integrated expectation value

$$\langle p_y \rangle_x = \int_0^1 dy \psi^* \left(-i\frac{\partial}{\partial y}\right)\psi$$

$$= 2\pi \sum_{n=-\infty}^{\infty} n\psi_n^* \psi_n. \tag{B.14}$$

The evolution of ψ_n given by (B.9) implies that

$$\frac{\partial}{\partial t}\langle p_y \rangle_x + \frac{\partial}{\partial x}\left[i\pi \sum_n n\left(\frac{\partial \psi_n^*}{\partial x}\psi_n - \psi_n^* \frac{\partial \psi_n}{\partial x}\right)\right] = 0, \tag{B.15}$$

and we deduce that the quantity of interest, the rate of change of the expectation value $\langle p_y \rangle$, is a difference of surface terms

$$\frac{d}{dt}\langle p_y \rangle = \frac{d}{dt}\int_0^1 dx \langle p_y \rangle_x$$

$$= \left[-i\pi \sum_n n\left(\frac{\partial \psi_n^*}{\partial x}\psi_n - \psi_n^* \frac{\partial \psi_n}{\partial x}\right)\right]_0^1. \tag{B.16}$$

Because of the boundary conditions (B.10) these surface terms do not cancel completely. We find

$$\frac{d}{dt}\langle p_y \rangle = i\pi \sum_n \left(\frac{\partial \psi_n^*}{\partial x}\psi_n - \psi_n^* \frac{\partial \psi_n}{\partial x}\right)\Bigg|_{x=0}. \tag{B.17}$$

The expression on the right-hand side has the same value if evaluated at $x=1$, and has the physical meaning of being 2π times the probability flux across the circle $x=0$. Also, because of surface terms, Ehrenfest's theorem has the anomalous form

$$\frac{d}{dt}\langle x \rangle = \langle p_x \rangle - \frac{i}{2}\sum_n \left(\frac{d\psi_n^*}{\partial x}\psi_n - \psi_n^* \frac{\partial \psi_n}{\partial x}\right)\Bigg|_{x=0}. \tag{B.18}$$

The anomaly here occurs simply because x is the coordinate on a circle, so there is a contribution to $d\langle x \rangle/dt$ equal to the product of the probability flux at $x=0$ and the

discontinuity of the coordinate there, which is unity. Combining (B.17) and (B.18) gives the simple result

$$\frac{d}{dt}\langle p_y + 2\pi x\rangle = \langle 2\pi p_x\rangle. \tag{B.19}$$

This is the quantum mechanical version of

$$\frac{dv_y}{dt} = 2\pi v_x, \tag{B.20}$$

one of the classical equations of motion for a particle in a magnetic field of strength 2π.

The gauge invariant operator $p_y + 2\pi x$ is the analog of Q_5^{reg} in the Schwinger model and Eq. (B.19) is the analog of the anomaly equation (3.39). p_y is the analog of Q_5'; it appears naively to be conserved, but is not because of surface terms. The more fundamental reason why p_y is not conserved is because rotations in the y-direction are not quite symmetries. Clearly the torus itself, the magnetic field, and even the gauge potential (B.1), are invariant under $y \to y - \alpha$ (mod 1). But $W_x(y)$, the holonomy on a circle in the x-direction, is not; it is multiplied by $\exp 2\pi i\alpha$. In the gauge that we are using, the mechanism for this is that the gauge transformation $g(y)$ changes from $\exp - 2\pi i y$ to $\exp - 2\pi i(y - \alpha)$.

There is another interesting gauge in which to describe the rotated gauge field. Namely, retain the gauge transformation $g(y) = \exp - 2\pi i y$ but change the gauge potential to

$$a_x = 2\pi\alpha, \qquad a_y = 2\pi x. \tag{B.21}$$

The constant potential a_x doesn't change the magnetic field, but the holonomy is now $W_x(y) = \exp - 2\pi i(y - \alpha)$. For a particle moving in one dimension a physical Bohm–Aharonov effect is produced by a potential like a_x that changes the holonomy but not the electromagnetic field. For example, on a unit circle, the spectrum of the Hamiltonian

$$-\frac{1}{2}\left(\frac{\partial}{\partial x} + 2\pi i\alpha\right)^2 \tag{B.22}$$

depends on α [22]. In our two-dimensional model, on the other hand, α can be eliminated by a rotation, so it is an irrelevant parameter, just like the Θ-parameter of the Schwinger model.

APPENDIX C: THE FRACTIONAL FERMION NUMBERS OF SCALAR SOLITONS

In the main part of this paper we have computed fermionic charges in a background gauge potential in one dimension, cf. Eqs. (5.2). These results may be used to rederive the fractional fermion numbers of one-dimensional scalar solitons.

The Goldstone–Wilczek Hamiltonian [7], expressed in a chiral basis, is

$$H = \int_{-\infty}^{\infty} [\psi_1^\dagger (-i\partial_x)\, \psi_1 + \psi_2^\dagger (i\partial_x)\, \psi_2 - \psi_2^\dagger \psi_1 \Phi - \psi_1^\dagger \psi_2 \Phi^*]\, dx. \tag{C.1}$$

Φ is a background complex scalar field, which has the property $|\Phi(-\infty)| = |\Phi(\infty)| = m$, $\Phi(\infty) = e^{-i\Theta}\Phi(-\infty)$. Earlier, Jackiw and Rebbi [7] considered the special case where Φ is real and $\Phi(-\infty) = -m$, $\Phi(\infty) = m$.

Formally, the fermion number of the soliton is the number of negative energy levels of the Hamiltonian H, since these are each filled by one fermion in the ground state. A regularized fermion number may be defined as follows. Suppose that space is a long interval $[-L, L]$, and that

$$\Phi(L) = e^{-i\Phi}\Phi(-L). \tag{C.2}$$

Then impose suitable boundary conditions on the Fermi field, to make the spectrum discrete, and use heat kernel regularization to control the ultraviolet divergence. With this prescription, which was used previously by Roy and Singh [23], the fermion number F is

$$F = \lim_{\lambda \to 0+} \left(\sum_{E_n < 0} e^{\lambda E_n} - \frac{2}{\lambda} \right), \tag{C.3}$$

where $\{E_n\}$ is the spectrum of H. F is independent of L for sufficiently large L, and its fractional part does not depend on L at all.

It makes no difference to the spectrum if we rewrite the Hamiltonian (C.1) in the gauge invariant form

$$H = \int_{-\infty}^{\infty} [\psi_1^\dagger (-i\partial_x + A_x)\, \psi_1 + \psi_2^\dagger (i\partial_x + A_x)\, \psi_2 - \psi_2^\dagger \psi_1 \Phi - \psi_1^\dagger \psi_2 \Phi^*]\, dx \tag{C.4}$$

and set $A_x = 0$. This is the Hamiltonian of axial electrodynamics, with an additional Yukawa coupling to a scalar field of charge -2. Gauge invariance now suggests the boundary conditions

$$\psi_1(L) = e^{(1/2)i\Theta}\psi_1(-L), \tag{C.5a}$$

$$\psi_2(L) = e^{-(1/2)i\Theta}\psi_2(-L) \tag{C.5b}$$

together with (C.2). All gauge invariant quantities, for example, $\psi_2^\dagger \psi_1 \Phi$, are equal at $-L$ and $+L$. One may therefore identify these points, obtaining a circle of length $2L$. In particular, the equality of $\psi_1^\dagger \psi_1 - \psi_2^\dagger \psi_2$ at $-L$ and $+L$ ensures that the Hamiltonian is self-adjoint, so the boundary conditions (C.5) are in the class considered by Roy and Singh. (Note that these boundary conditions will lead to a unique value for the fermion number, whereas the value found by Roy and Singh was ambiguous, because they *only* required self-adjointness of the Hamiltonian, which is not a restrictive enough condition.)

Next, we are free to perform the gauge transformation

$$\begin{aligned} &\psi_1 \to e^{i\alpha(x)}\psi_1, \qquad &&\psi_2 \to e^{-i\alpha(x)}\psi_2 \\ &\Phi \to e^{-2i\alpha(x)}\Phi, \qquad &&A_x \to A_x - \frac{d\alpha}{dx} \end{aligned} \tag{C.6}$$

where $\alpha(x) = -\Theta(x+L)/4L$. Such a transformation does not change the fermion spectrum, nor the fermion number of any state. In the new gauge, $A_x = \Theta/4L$; also, the Fermi and scalar fields are equal at $\pm L$, and therefore continuous on the circle. It is now possible to continuously reduce the scalar field to zero everywhere. The fermion spectrum changes in the process, but the fractional part of the fermion number does not—it is a topological invariant. The reason is that this fractional part depends only on the nature of the spectrum at very large negative energy. (If we define F_E by restricting the sum in (C.3) to energies $E_n < E$, with $E \ll 0$, then $F - F_E$ is an integer.) Reducing the scalar field to zero produces an $O(E_n^{-1})$ change in an energy eigenvalue $E_n \ll 0$, as can easily be verified in the WKB approximation, and this change is sufficiently small that the fractional fermion number is unaltered [24].

When Φ vanishes, the Hamiltonian is simply that of axial electrodynamics. The fermion number F is the difference $Q_{\mathrm{L}}^{\mathrm{reg}} - Q_{\mathrm{R}}^{\mathrm{reg}}$. From equations analogous to (5.2), appropriate for a circle of length $2L$ rather than 2π, one finds

$$Q_{\mathrm{L}}^{\mathrm{reg}} - Q_{\mathrm{R}}^{\mathrm{reg}} = M - N + 1 + \frac{2L}{\pi} A_x, \tag{C.7}$$

where M and N are the integers specifying the location of the Fermi surfaces. In the lowest energy state, where all negative energy levels are filled and all positive energy levels are empty, $M = -1$ and $N = 1$ if $0 < A_x < \pi/L$. Since $A_x = \Theta/4L$, the fermion number is

$$F = -1 + \frac{\Theta}{2\pi}. \tag{C.8}$$

From this we conclude that the fractional part of the fermion number of the original soliton is $\Theta/2\pi$, which agrees with Goldstone and Wilczek's result. When $\Theta = \pi$, corresponding to a real scalar soliton, the fractional fermion number is $\frac{1}{2}$, which is Jackiw and Rebbi's result.

To see that the boundary conditions (C.5) are reasonable, consider the case where $\Phi(x) = \tanh x$. There is precisely one eigenstate of the Dirac Hamiltonian whose energy ΔE approaches zero as $L \to \infty$; in the limit it becomes the Jackiw–Rebbi zero mode. A good approximation to this eigenstate is

$$\psi \approx \begin{pmatrix} \dfrac{1}{\cosh x} + i\,\Delta E \sinh x \\ \dfrac{i}{\cosh x} + \Delta E \sinh x \end{pmatrix}. \tag{C.9}$$

By imposing the boundary conditions on ψ, the energy is determined to be

$$\Delta E \approx 4e^{-2L}. \tag{C.10}$$

(Both (C.9) and (C.10) are correct to leading order in e^{-L}.) There is no low-energy mode hidden at infinity, because with the circular geometry, and in the right gauge, the scalar field is effectively constant there.

Roy and Singh pointed out that when Φ is real, and the boundary conditions on the Fermi field are fixed, then no fermion energies pass through zero as the scalar field is changed. We deduce from (C.8), therefore, that the Jackiw–Rebbi soliton has absolute fermion number $-\frac{1}{2}$. This is as expected, since the energy ΔE is positive and the corresponding level is unoccupied. In the limit $L \to \infty$, the zero mode is unoccupied.

To recover the Jackiw–Rebbi soliton with the zero mode occupied we must modify the boundary conditions. Provided the complex scalar field is never zero, the phase Θ is absolutely well defined, not just mod 2π, and the boundary conditions are unambiguous. However, for a real soliton, the scalar field must vanish somewhere, and it is then consistent to set $\Theta = -\pi$ in (C.5), rather than $\Theta = \pi$ as before. Now the absolute fermion number can be shown to be $\frac{1}{2}$. But also $\Delta E = -4e^{-2L}$, so the zero mode would be occupied in the limit $L \to \infty$.

To summarize, we have shown that the fractional fermion number of scalar solitons and the anomaly in the fermion number in axial electrodynamics are related. They have their origin in topologically equivalent fermionic spectra. For another view of this relationship, see Ref. [25].

Acknowledgments

This work was inspired by conversations with Iz Singer, and by his lectures at the ITP Santa Barbara Workshop on Functional Integrals. I am also grateful to Ian Affleck, Jan Ambjørn, Sidney Coleman, Frans Klinkhamer, Roman Jackiw, Emil Mottola, and Frank Wilczek for discussions, and would particularly like to thank Roman Jackiw for suggesting a re-examination of axial and chiral electrodynamics in one spatial dimension. An error in an earlier version of this paper led me to state incorrectly that axial, as well as chiral, electrodynamics was inconsistent; I am grateful to Sang Gyu Jo for pointing out this error. This material is based upon research supported in part by the National Science Foundation under Grant PHY77-27084, supplemented by funds from the National Aeronautics and Space Administration.

References

1. J. Schwinger, *Phys. Rev.* **128** (1962), 2425.
2. L. S. Brown, *Nuovo Cimento (10)* **29** (1963), 617; J. Lowenstein and A. Swieca, *Ann. Phys. (N.Y.)* **68** (1971), 172.
3. J. Fröhlich and E. Seiler, *Helv. Phys. Acta* **49** (1976), 889. This generalizes to the Schwinger model, and makes rigorous the discussion of the Sine–Gordon–Thirring model by S. Coleman, *Phys. Rev. D* **11** (1975), 2088. See also S. Mandelstam, *Phys. Rev. D* **11** (1975), 3026.

4. K. JOHNSON, *Phys. Lett.* **5** (1963), 253. For a review, see R. Jackiw, *in* "Relativity, Groups and Topology II, Les Houches 1983," edited by B. S. De Witt and R. Stora, North Holland, 1984. The anomaly in the lattice Schwinger model is discussed by M. Weinstein, *Phys. Rev. D* **26** (1982), 839.
5. M. F. ATIYAH, V. K. PATODI, AND I. M. SINGER, *Math. Proc. Cambridge Philos. Soc.* **79** (1976), 71.
6. C. G. CALLAN, JR., R. DASHEN, AND D. J. GROSS, *Phys. Rev. D* **17** (1978), 2717; J. KISKIS, *Phys. Rev. D* **18** (1978), 3690; N. H. CHRIST, *Phys. Rev. D* **21** (1980), 1591; J. AMBJØRN, J. GREENSITE, AND C. PETERSON, *Nuclear Phys. B* **221** (1983), 381; H. B. NIELSEN AND M. NINOMIYA, *Phys. Lett. B* **130** (1983), 389.
7. R. JACKIW AND C. REBBI, *Phys. Rev. D.* **13** (1976), 3398; J. GOLDSTONE AND F. WILCZEK, *Phys. Rev. Lett.* **47** (1981), 986. See also Refs. [23] and [25]. (These papers are of direct relevance here, but they represent just a small part of the literature on this subject.)
8. This appeared previously: N. S. Manton, ITP Santa Barbara preprint NSF-ITP-83-164, 1983, unpublished.
9. T. D. LEE, "Particle Physics and Introduction to Field Theory," Harwood Academic, Chur, 1981.
10. D. C. MATTIS AND E. H. LIEB, *J. Math. Phys.* **6** (1965), 304.
11. T. BOHR, The Luttinger model, Nordita preprint 81/4, 1981.
12. Axial and chiral electrodynamics in one spatial dimension are specifically discussed in R. Jackiw's lectures, Ref. [4]. See also D. J. GROSS AND R. JACKIW, *Phys. Rev. D* **6** (1972), 477; C. P. KORTHALS ALTES AND M. PERROTTET, *Phys. Lett. B* **39** (1972), 546.
13. S. COLEMAN, R. JACKIW, AND L. SUSSKIND, *Ann. Phys. (N.Y.)* **93** (1975), 267. See also J. LOWENSTEIN AND A. SWIECA, Ref. [2].
14. G. 'T HOOFT, *Phys. Rev. Lett.* **37** (1976), 8; *Phys. Rev. D* **14** (1976), 3432; C. G. CALLAN, JR., R. F. DASHEN, AND D. J. GROSS, *Phys. Lett. B* **63** (1976), 334; R. JACKIW AND C. REBBI, *Phys. Rev. Lett.* **37** (1976), 172.
15. S.-S. CHERN, "Complex Manifolds without Potential Theory," Appendix 3, Springer-Verlag, New York, 1979.
16. S. ADLER, *Phys. Rev.* **177** (1969), 2426; J. BELL AND R. JACKIW, *Nuovo Cimento A (10)* **60** (1969), 47. For a review see R. Jackiw, Ref. [4].
17. T. T. WU AND C. N. YANG, *Phys. Rev. D* **12** (1975), 3845.
18. L. D. LANDAU AND E. M. LIFSHITZ, "Quantum Mechanics," Pergamon, Oxford, 1977.
19. N. K NIELSEN AND P. OLESEN, *Nuclear Phys. B* **144** (1978), 376.
20. J. AMBJØRN AND P. OLESEN, *Nuclear Phys. B* **170** (FS1) (1980), 60.
21. F. D. M. HALDANE, *Phys. Rev. Lett.* **51** (1983), 605; QIAN NIU, D. J. THOULESS AND Y. S. WU, Univ. Washington preprint, 1984.
22. A quantum mechanical particle moving on a circle can be regarded as a model for a pure gauge theory. See, for example, M. ASOREY, J. G. ESTEVE, AND A. F. PACHECO, *Phys. Rev. D* **27** (1983), 1852.
23. S. M. ROY AND V. SINGH, *Phys. Lett. B* **143** (1984), 179. See also R. RAJARAMAN AND J. S. BELL, *Phys. Lett. B* **116** (1982), 151.
24. This result originated in the mathematical literature. See Ref. [5], and P. B. GILKEY AND L. SMITH, *Comm. Pure Appl. Math.* **36** (1983), 85.
25. Y. FRISHMAN, D. GEPNER, AND S. YANKIELOWICZ, *Phys. Lett. B* **130** (1983), 66.