Nuclear Physics B323 (1989) 631–659
North-Holland, Amsterdam

PARITY ANOMALY AND FERMION–BOSON TRANSMUTATION IN 3-DIMENSIONAL LATTICE QED

A. COSTE* and M. LÜSCHER

Deutsches Elektronen-Synchroton, DESY, Notkestrasse 85, 2000 Hamburg 52, FRG

Received 22 February 1989

We calculate the effective gauge field action due to a single two-component fermion of mass m and charge e, using a cubic lattice as an ultra-violet regulator and Wilson type fermion actions to avoid a doubling of fermion species. In the limit where the cutoff is removed, we find that the answer is finite but not unique: the various lattice fermion actions belong to different universality classes labelled by an integer n. In all cases, the parity invariance of the classical (continuum) action at $m = 0$ is violated by quantum effects, i.e. there is an anomaly. In addition, for $n \neq 0$, a Chern–Simons term is generated at large m, and this implies a non-zero photon mass, deconfinement and unusual spin and statistics of charged particles. In particular, for the class of actions with $n = -1$, we show by a straightforward semi-classical argument that the fundamental fermion is converted to a particle with spin 0.

1. Introduction

The interest in 3-dimensional gauge theories has recently been revived because they are apparently related, in various ways, to high-temperature superconductivity [1–5]. In particular, it has been suggested that abelian gauge fields $A_\mu(x)$ with an action including a term proportional to the Chern–Simons action★

$$I[A] = \frac{1}{8\pi^2} \int d^3x \, \epsilon_{\mu\nu\rho} A_\mu(x) \, \partial_\nu A_\rho(x) \tag{1.1}$$

can affect the spin and statistics of charged particles [6–14] and in this way lead to peculiar properties of the elementary excitations in models of superconducting materials. It is also well-known that the Chern–Simons term (1.1) makes the photon massive [15–17]. Independently of these developments, the possibility of a fermion–boson transformation in three space-time dimensions is interesting in itself and it is certainly very important to understand this phenomenon from different points of view.

* On leave from CPT, CNRS Marseille, France.

★ The space-time metric is euclidean and the totally anti-symmetric symbol $\epsilon_{\mu\nu\rho}$ is normalized such that $\epsilon_{123} = 1$. Repeated indices $\mu, \nu, \ldots$ are summed over from 1 to 3.

0550-3213/89/$03.50 © Elsevier Science Publishers B.V.
(North-Holland Physics Publishing Division)

An intriguing observation in this context is that the Chern–Simons action is radiatively generated by charged fermions coupled minimally to the gauge field $A_\mu(x)$ [18, 19]. Thus, consider a two-component fermion field $\psi(x)$ with (continuum) action

$$S_F = -\int d^3x\, \bar{\psi}(x)(D-m)\psi(x), \tag{1.2}$$

where the Dirac operator D is given by

$$D = \gamma_\mu\big(\partial_\mu + ieA_\mu(x)\big), \tag{1.3}$$

$$\{\gamma_\mu, \gamma_\nu\} = 2\delta_{\mu\nu}, \qquad \gamma_\mu^\dagger = \gamma_\mu, \tag{1.4}$$

and e denotes the electric charge of ψ. The effective gauge field action $\Gamma[A]$ generated by such a fermion field is, formally,

$$\Gamma[A] = -\ln\det[(D-m)/(D_0-m)], \tag{1.5}$$

where $D_0 = D|_{A_\mu=0}$.

Now there are two ways in which the Chern–Simons action can appear. First, there is the well-known parity anomaly [18–22], viz.

$$\lim_{m\to 0} \operatorname{Im}\Gamma[A] = c_0 e^2 I[A] + \pi h[A]. \tag{1.6}$$

Here, c_0 is a constant, to be discussed below, and $h[A]$ is an integer, which vanishes when A_μ is sufficiently close to zero ($h[A]$ changes by ± 1 when an eigenvalue of D crosses zero) [20–22]. Eq. (1.6) is referred to as an anomaly, because it implies that the reflection symmetry

$$\psi(x) \to \psi(-x), \qquad \bar{\psi}(x) \to -\bar{\psi}(-x), \qquad A_\mu(x) \to -A_\mu(-x) \tag{1.7}$$

of the classical fermion action (1.2) at $m = 0$ is broken through quantum effects.

The second instance where the Chern–Simons action appears is in the large mass limit, where it has been shown that [18, 19]

$$\lim_{m\to\infty} \Gamma[A] = ic_\infty e^2 I[A]. \tag{1.8}$$

This is a very interesting and, at first sight, rather surprising result. It implies that a heavy fermion does not decouple from the theory but leaves behind a computable local term which has long-range effects: it makes the photon massive and it changes the spin and statistics of charged particles. Thus, it is important to understand how exactly this can happen. In particular, the coefficient c_∞ should be calculated since the photon mass and the amount of spin change of charged particles are determined by this number (and the electric charges of the particles).

Although c_0 and c_∞ have already been computed elsewhere [18–22], we do not quote any numbers at this point, because these coefficients are actually not well determined through the formal expression (1.5) for the effective action. First there is a sign ambiguity which arises from the fact that there are two inequivalent irreducible representations of the gamma matrices in three dimensions and that the fermion mass parameter m can be given either sign. We fix these ambiguities once and for all by setting

$$\gamma_1\gamma_2\gamma_3 = i \quad \text{and} \quad m \geqslant 0 . \tag{1.9}$$

After this the numbers c_0 and c_∞ are still not uniquely determined, because the determinant (1.5) is ultra-violet divergent. Regularization is thus required and it turns out that the values (and not only the signs) of c_0 and c_∞ depend on the details of the regularization procedure even if we restrict ourselves to gauge invariant regularizations [16].

In this paper we completely determine the ambiguity which remains and in particular compute c_0 and c_∞ in a lattice formulation of the theory where a doubling of fermions is avoided by Wilson's method. In the context of high-temperature superconductivity, the introduction of a lattice as an ultraviolet regulator is natural. In addition we hope that our calculation will prove useful in attempts to find a lattice form of the fermion–boson transformation in $2+1$ dimensions analogous to the Jordan–Wigner transformation in $1+1$ dimensions. Finally, we mention that anomalies in even space-time dimensions have previously been worked out on the lattice (see refs. [23, 24] and references quoted there), the aim being to demonstrate that these subtle quantum effects are correctly reproduced by lattice fermions (which is not entirely obvious in view of the species doubling problem).

The outcome of our investigation is summarized by

$$c_0 = c_\infty + \pi , \qquad c_\infty = 2\pi n . \tag{1.10,1.11}$$

Here, n can be any integer, depending on the lattice action chosen. We shall provide explicit examples with $n = 0$ and $n = -1$. The quantization (1.11) of c_∞ is also found when one uses a Pauli–Villars cutoff with one or several ghost fields with variable mass signs. In fact we shall argue that it follows from two simple properties of the regularization procedure employed: gauge invariance and compatibility with respect to gauge group reductions. In the explicit lattice calculation the quantization is seen to arise from the fact that c_∞ is proportional to the winding number of the free fermion propagator viewed as a mapping from a 3-dimensional torus, the Brillouin zone, to the space of non-singular 2×2 matrices.

The non-universality of the effective action $\Gamma[A]$ can be absorbed by including a Chern–Simons term in the bare gauge field action with an adjustable coefficient c. Different regularizations can then be matched by performing a (finite) renormalization of c. This point of view has been taken by many authors, but the price to pay is that now the theory contains one more free parameter and hence becomes less predictive. Of course, one can always impose some condition on c, for example that the full photon propagator be parity invariant at large momenta, where mass effects can be neglected.

Although not unattractive, such a condition may, however, not be natural in an application of the model to a problem in solid state physics, where the lattice and the bare action may be given a priori, without Chern–Simons term. This is the point of view we adopt in this paper, and we shall therefore assume that the bare gauge field action contains no other term besides the usual Maxwell action. We shall then discuss the physical properties of the model in some detail for two lattice versions, corresponding to $n=0$ and $n=-1$. In the latter case the already famous fermion–boson transmutation takes place, i.e. the fundamental charged fermions turn into spin-zero bosons with only short-range interactions. As we shall see, this result is crucially dependent on the fact that c_∞ happens to be equal to -2π; any other value would imply a non-zero and in general fractional spin. Although it is hardly an accident that the radiatively induced Chern–Simons term comes with precisely this coefficient, we do not really understand (in terms of a physical picture) why this is so.

The organization of our paper is as follows. First we discuss the lattice calculation in detail (sect. 2). As already mentioned, we employ Wilson fermions so that there is no fermion doubling. However, we allow the Wilson regulating term to have either sign, which is perfectly admissible since in both cases reflection positivity holds and the classical continuum limit is as given by eq. (1.2). Thus, there is no apparent reason to exclude any one of them, but as we shall show we have the conventional result $c_\infty=0$ in the first case while in the other one gets $c_\infty=-2\pi$. In sect. 3 we argue that (1.10) and (1.11) must be true on general grounds and for any reasonable regularization. The implications of our results on the properties of lattice QED, including all degrees of freedom, are discussed in sect. 4. In particular, using a semi-classical argument, we show that for lattice actions in the universality class $n=-1$, the observable spin of the fundamental charged fermion vanishes. Finally, conclusions are drawn in sect. 5 and two appendices are included which contain the technical details of the semi-classical calculation mentioned above.

2. Calculation of the effective action $\Gamma[A]$ in lattice QED

In this section we compute the numbers c_0 and c_∞ using a lattice as a gauge invariant ultra-violet cutoff. Since the Chern–Simons action (1.1) is a polynomial in the gauge field A_μ, the computation can be done in perturbation theory, i.e. we only need to evaluate the effective action to second order in e and to study what happens in the continuum limit. This latter task will be greatly simplified by invoking the Reisz power counting theorem for lattice Feynman diagrams [25, 26].

2.1. BASIC LATTICE NOTATIONS

We consider a 3-dimensional cubic lattice with vertices

$$x = (an_1, an_2, an_2), \qquad n_\mu \in \mathbb{Z}, \tag{2.1}$$

where a denotes the lattice spacing. In the "non-compact" formulation of abelian lattice gauge theories (which we adopt here), a gauge field is an assignment of a real number $A_\mu(x)$ to every lattice point x and direction μ. The field tensor $F_{\mu\nu}(x)$ is then defined through

$$F_{\mu\nu}(x) = \partial_\mu A_\nu(x) - \partial_\nu A_\mu(x), \tag{2.2}$$

where the lattice derivative is given by

$$\partial_\mu f(x) = (f(x + a\hat{\mu}) - f(x))/a, \tag{2.3}$$

and $\hat{\mu}$ denotes the unit vector in the positive μ-direction. $F_{\mu\nu}(x)$ is invariant under the gauge transformation

$$A_\mu(x) \to A_\mu(x) + \partial_\mu \Lambda(x), \tag{2.4}$$

and so is the gauge field action

$$S_G = \tfrac{1}{4} a^3 \sum_x F_{\mu\nu}(x) F_{\mu\nu}(x). \tag{2.5}$$

Incidentally, the Chern–Simons action (1.1) can be readily translated to the lattice in the present formulation. For example, the expression

$$I[A] = \frac{a^3}{8\pi^2} \sum_x \epsilon_{\mu\nu\rho} A_\mu(x - a\hat{\mu}) \partial_\nu A_\rho(x) \tag{2.6}$$

is gauge invariant and it obviously reduces to (1.1) in the classical continuum limit.

Dirac fields $\psi(x)$ on the lattice are two-component spinor fields as in the continuum. Under a gauge transformation they transform as

$$\psi(x) \to \mathrm{e}^{-ie\Lambda(x)} \psi(x). \tag{2.7}$$

Accordingly, the lattice derivatives

$$\nabla_\mu \psi(x) = \{U(x,\mu)\psi(x+a\hat{\mu}) - \psi(x)\}/a\,,$$

$$\nabla_\mu^* \psi(x) = \left\{\psi(x) - U(x-a\hat{\mu},\mu)^{-1}\psi(x-a\hat{\mu})\right\}/a\,, \tag{2.8}$$

where we have introduced the link variables

$$U(x,\mu) = \exp\left(ieaA_\mu(x)\right), \tag{2.9}$$

are gauge covariant. Following Wilson, we thus define a lattice Dirac operator and fermion action through

$$D = \tfrac{1}{2}\gamma_\mu\left(\nabla_\mu^* + \nabla_\mu\right) + \tfrac{1}{2}sa\,\nabla_\mu^*\nabla_\mu\,, \tag{2.10}$$

$$S_{\mathrm{F}} = -a^3\sum_x \bar{\psi}(x)(D-m)\psi(x)\,. \tag{2.11}$$

The term proportional to the coefficient s in eq. (2.10) which is formally of order a is introduced to avoid the well-known doubling of fermions on the lattice. Originally, Wilson chose $s=1$, but there is actually no very deep reason to restrict s to this particular value. In particular, for $s=-1$ and $m<1/a$, the theory defined through the total action

$$S = S_{\mathrm{G}} + S_{\mathrm{F}}\,, \tag{2.12}$$

can be shown to share all the essential properties with the original Wilson model: there is no fermion doubling and the free propagator in momentum space is regular for all momenta $p \neq 0 \pmod{2\pi/a}$, the lattice Feynman rules reduce to the continuum rules in the limit $a \to 0$, and a self-adjoint positive transfer matrix exists which acts in a (positive metric) Hilbert space of physical states. It is therefore tempting to assume that the lattice theories with $s=\pm 1$ (to which we restrict ourselves in what follows) are in the same universality class, but as we shall see this is not true, not even in perturbation theory★.

The effective action $\Gamma[A]$ is now defined as in the continuum (eq. (1.5)). From the definition (2.10) of the Dirac operator it is easy to check that

$$\Gamma[A] = \Gamma[-A]\,, \tag{2.13}$$

$$\Gamma[A]^* = \Gamma[A^P]\,, \qquad A_\mu^P(x) = -A_\mu(-x-a\hat{\mu})\,. \tag{2.14}$$

★ A similar effect, related to the axial anomaly, occurs in even dimensional lattice gauge theories. See ref. [34] for a summary and further references.

Eq. (2.13) is nothing other than charge conjugation invariance (Furry's theorem). The second property (2.14) implies that c_∞ must be real.

2.2. PROPERTIES OF THE FREE FERMION PROPAGATOR

The free fermion propagator $S_0(x, y)_{\alpha\beta}$ is defined by

$$[(-D_0 + m)S_0(x, y)]_{\alpha\beta} = \frac{1}{a^3}\delta_{\alpha\beta}\delta_{xy}. \tag{2.15}$$

In momentum space we have

$$S_0(x, y) = \int_{-\pi/a}^{\pi/a} \frac{d^3p}{(2\pi)^3} e^{ip(x-y)}\tilde{S}_0(p), \tag{2.16}$$

$$\tilde{S}_0(p) = Q(p)^{-1}, \qquad Q(p) = m + \tfrac{1}{2}sa\hat{p}^2 - i\gamma\cdot\tilde{p}, \tag{2.17}$$

where $\hat{p}_\mu$ and $\tilde{p}_\mu$ are defined through

$$\hat{p}_\mu = \frac{2}{a}\sin\left(\tfrac{1}{2}ap_\mu\right), \qquad \tilde{p}_\mu = \frac{1}{a}\sin(ap_\mu). \tag{2.18, 2.19}$$

We also define a scalar function $R(p)$ through

$$R(p) = \left(m + \tfrac{1}{2}sa\hat{p}^2\right)^2 + \tilde{p}^2, \tag{2.20}$$

and it is then easy to show that

$$Q(-p)Q(p) = \det Q(p) = R(p), \qquad Q(p)^\dagger = Q(-p), \tag{2.21}$$

and hence

$$\tilde{S}_0(p) = Q(-p)/R(p). \tag{2.22}$$

It is obvious from these equations that the propagator $\tilde{S}_0(p)$ converges to the expected continuum expression in the limit $a \to 0$, p fixed. In particular, the continuum limit is independent of the parameter s.

From the definition (2.20) it follows that the denominator $R(p)$ of the propagator is non-negative and smooth. Furthermore, for $s = \pm 1$ we have

$$R(p) = m^2 + (1 + sam)\hat{p}^2 + \tfrac{1}{2}a^2 \sum_{\mu<\nu} \hat{p}_\mu^2\hat{p}_\nu^2, \tag{2.23}$$

and hence the bound

$$R(p) \geq C\left(m^2 + \hat{p}^2\right), \tag{2.24}$$

where $C=1$ for $s=1$ and $C=1-am$ for $s=-1$. In the latter case we shall always assume that $m<1/a$. The bound (2.24) then implies that there is no fermion doubling. In fact it is easy to show from eq. (2.23) that the one-particle energy $\omega(\boldsymbol{p})$, defined through

$$R(p)=0, \qquad p=(\pm i\omega(\boldsymbol{p}),\boldsymbol{p}), \tag{2.25}$$

is real, single valued and vanishes only when $m=0$ and $\boldsymbol{p}=\mathbf{0}$ (mod $2\pi/a$).

2.3. PERTURBATION EXPANSION OF THE EFFECTIVE ACTION

We here assume that $A_\mu(x)$ is a given external gauge potential and expand the effective action $\Gamma[A]$ in powers of e. Through the link variables (2.9), the Dirac operator D depends on e and it thus follows that

$$D=\sum_{k=0}^{\infty} e^k D_k, \tag{2.26}$$

where for $k \geqslant 1$ we have

$$D_k\psi(x)=\frac{(ia)^k}{2ak!}\sum_{\mu=1}^{3}\left\{A_\mu(x)^k(s+\gamma_\mu)\psi(x+a\hat{\mu})\right.$$
$$\left.+(-1)^k A_\mu(x-a\hat{\mu})^k(s-\gamma_\mu)\psi(x-a\hat{\mu})\right\}. \tag{2.27}$$

Inserting (2.26) into the definition (1.5) of the effective action, we obtain the expansion

$$\Gamma[A]=\sum_{k=1}^{\infty} e^k\Gamma_k[A]. \tag{2.28}$$

Note that because of charge conjugation symmetry, $\Gamma_k[A]$ vanishes for odd k.

As already mentioned above, to compute c_0 and c_∞ we only need to work out the 2-point function

$$\Gamma_2[A]=\mathrm{Tr}\left\{\tfrac{1}{2}\left[D_1(D_0-m)^{-1}\right]^2-D_2(D_0-m)^{-1}\right\}. \tag{2.29}$$

If we introduce the Fourier transform $\tilde{A}_\mu(q)$ of the gauge potential through

$$A_\mu(x)=\int_{-\pi/a}^{\pi/a}\frac{\mathrm{d}^3q}{(2\pi)^3}\mathrm{e}^{iq(x+\frac{1}{2}a\hat{\mu})}\tilde{A}_\mu(q), \tag{2.30}$$

eq. (2.29) may be written as

$$\Gamma_2[A]=\tfrac{1}{2}\int_{-\pi/a}^{\pi/a}\frac{\mathrm{d}^3q}{(2\pi)^3}\tilde{A}_\mu(-q)\hat{\Pi}_{\mu\nu}(q)\tilde{A}_\nu(q), \tag{2.31}$$

where the vacuum polarization tensor $\hat{\Pi}_{\mu\nu}(q)$ is given by

$$\hat{\Pi}_{\mu\nu}(q)=\int_{-\pi/a}^{\pi/a}\frac{d^3p}{(2\pi)^3}\left[1-T_0(q)\right]\mathrm{tr}\left\{\partial_\mu Q(p)Q(p+\tfrac{1}{2}q)^{-1}\partial_\nu Q(p)Q(p-\tfrac{1}{2}q)^{-1}\right\} \tag{2.32}$$

The symbol $T_0(q)$ in this equation implies a Taylor subtraction at zero momentum. For any $n \geqslant 0$ and any function $f(q)$ this operation is defined through

$$T_n(q)f(q)=\sum_{k=0}^{n}\frac{1}{k!}\frac{\partial^k}{\partial t^k}f(tq)\bigg|_{t=0}. \tag{2.33}$$

Thus, we have

$$\hat{\Pi}_{\mu\nu}(0)=0, \tag{2.34}$$

and it is also not difficult to verify that

$$\hat{q}_\mu\hat{\Pi}_{\mu\nu}(q)=0, \tag{2.35}$$

which is just expressing the gauge invariance of the effective action to second order in e.

2.4. SMALL MOMENTUM BEHAVIOUR OF $\hat{\Pi}_{\mu\nu}(q)$

In this subsection we restrict ourselves to the massive case ($m>0$). The propagator $\tilde{S}_0(p)$ then is a completely regular function for all momenta p in the Brillouin zone

$$\mathscr{B}=\left\{p\in\mathbb{R}^3 \,\middle|\, |p_\mu|\leqslant\pi/a\right\}. \tag{2.36}$$

The vacuum polarization tensor $\hat{\Pi}_{\mu\nu}(q)$ can therefore be expanded in a power series around $q=0$. Taking gauge invariance and the discrete lattice symmetries into account, the first two terms in this expansion can be shown to be of the form

$$\hat{\Pi}_{\mu\nu}(q)=a_0\,\epsilon_{\mu\nu\rho}q_\rho+b_0\left(q^2\delta_{\mu\nu}-q_\mu q_\nu\right)+\mathrm{O}(q^3), \tag{2.37}$$

where a_0 and b_0/m are some constants depending on s and am.

An interesting observation now is that the leading order coefficient a_0 has a topological significance. To see this, first note that by differentiating eq. (2.32) one obtains

$$a_0=\frac{1}{48\pi^3}\int_{\mathscr{B}}d^3p\,\epsilon_{\mu\nu\rho}\,\mathrm{tr}\left\{\left(Q^{-1}\partial_\mu Q\right)\left(Q^{-1}\partial_\nu Q\right)\left(Q^{-1}\partial_\rho Q\right)\right\}. \tag{2.38}$$

Without affecting the value of the integral, we can replace the matrix $Q(p)$ in this equation by

$$U(p) = Q(p)/\sqrt{R(p)}\,. \tag{2.39}$$

In view of eq. (2.21), this is a smooth periodic function with values in the group SU(2). In other words, $U(p)$ is a mapping from the Brillouin zone $\mathscr{B}$, viewed as a 3-dimensional torus, into SU(2) and the integral (2.38) is nothing other than the winding number of this mapping divided by 2π.

For $s = 1$ the winding number vanishes, because we can make m very large and $U(p)$ is then homotopically deformed to the trivial mapping $U(p) = 1$. For $s = -1$, we cannot do this since the mass must satisfy $m < 1/a$ to guarantee the regularity of $U(p)$ (cf. eq. (2.23)). In fact, a little thought shows that the winding number must be equal to ± 1: it is only at $p = 0$ that $U(p)$ assumes the value 1 and since the mapping is locally invertible there, it follows that a whole neighborhood of 1 in SU(2) is covered exactly once. Thus, we have

$$a_0 = \begin{cases} 0 & \text{if } s = 1\,; \\ -1/2\pi & \text{if } s = -1\,. \end{cases} \tag{2.40}$$

Since a_0 is independent of the lattice spacing, it is already clear at this point that the effective action at low momenta is not universal in the continuum limit, i.e. here is the origin of the lattice action dependence of the coefficients c_0 and c_∞ alluded to in sect. 1 (we come back to this issue later on).

2.5. CONTINUUM LIMIT OF $\hat{\Pi}_{\mu\nu}(q)$

We now proceed to evaluate

$$\Pi_{\mu\nu}(q) = \lim_{a\to 0} \hat{\Pi}_{\mu\nu}(q) \tag{2.41}$$

assuming $m > 0$. To this end, we first rewrite eq. (2.32) in the form

$$\hat{\Pi}_{\mu\nu}(q) = a_0 \epsilon_{\mu\nu\rho} q_\rho + \int_{-\pi/a}^{\pi/a} \frac{\mathrm{d}^3 p}{(2\pi)^3} [1 - T_1(q)]\, F_{\mu\nu}(p, q; m, a)\,, \tag{2.42}$$

where a_0 is the coefficient calculated in subsect. 2.4 and the function $F_{\mu\nu}$ is defined through

$$F_{\mu\nu}(p, q; m, a) = \mathrm{tr}\left\{ \partial_\mu Q(p) Q\left(p + \tfrac{1}{2}q\right)^{-1} \partial_\nu Q(p) Q\left(p - \tfrac{1}{2}q\right)^{-1} \right\}. \tag{2.43}$$

The reason for rewriting $\hat{\Pi}_{\mu\nu}(q)$ in this particular way is that the additional subtraction at $q = 0$ of the integrand in eq. (2.42) makes the integral better behaved in the continuum limit. In fact, this limit can now be controlled with the help of the power counting theorem of Reisz [25] (see ref. [26] for an introduction to the theorem). Making equal denominators, the integrand in eq. (2.42) assumes the standard form

$$[1 - T_1(q)]\, F_{\mu\nu}(p, q; m, a) = V(p, q; m, a)/C(p, q; m, a), \tag{2.44}$$

where V is a polynomial in m with the coefficients depending smoothly on the momenta, and

$$C(p, q; m, a) = R(p + \tfrac{1}{2}q)\, R(p - \tfrac{1}{2}q)\, R(p)^3. \tag{2.45}$$

The lattice degree of divergence (as introduced by Reisz) of these two functions is given by

$$\deg V \leqslant 6, \qquad \deg C = 10, \tag{2.46}$$

and the total degree of the integral is hence less than 0. Thus, the theorem applies and we conclude that

$$\Pi_{\mu\nu}(q) = a_0\, \epsilon_{\mu\nu\rho} q_\rho + \int_{-\infty}^{\infty} \frac{\mathrm{d}^3 p}{(2\pi)^3} \lim_{a \to 0} [1 - T_1(q)]\, F_{\mu\nu}(p, q; m, a). \tag{2.47}$$

Note that apart from the subtractions at $q = 0$, the resulting integral on the r.h.s. of this equation is just the integral which one would obtain using continuum Feynman rules. The subtractions are however essential, because they convert an initially linearly divergent integral to an absolutely convergent one, as one may easily verify by (ordinary) power counting.

Summarizing the discussion so far, we have shown that the lattice vacuum polarization $\hat{\Pi}_{\mu\nu}(q)$ converges to a well-defined tensor $\Pi_{\mu\nu}(q)$ in the continuum limit. In particular, there are no divergent terms and no renormalization of the effective action is hence needed. However, the continuum vacuum polarization (2.47) is not the same for all lattice actions, because it involves the coefficient a_0 which depends on the parameter s as we showed previously.

We finally evaluate the integral in eq. (2.47) using Feynman parameters and obtain

$$\Pi_{\mu\nu}(q) = A(q^2)\, \epsilon_{\mu\nu\rho} q_\rho + B(q^2)\big(q^2 \delta_{\mu\nu} - q_\mu q_\nu\big), \tag{2.48}$$

where the amplitudes $A(q^2)$ and $B(q^2)$ are given by

$$A(q^2) = a_0 + \frac{1}{4\pi} \int_0^1 \mathrm{d}t \left\{1 - m\left[m^2 + t(1-t)q^2\right]^{-1/2}\right\}, \tag{2.49}$$

$$B(q^2) = \frac{1}{2\pi} \int_0^1 \mathrm{d}t\, t(1-t)\left[m^2 + t(1-t)q^2\right]^{-1/2}. \tag{2.50}$$

2.6. RESULTS

We first discuss the parity anomaly (1.6). From the explicit expressions above, one finds

$$\lim_{m\to 0} \Pi_{\mu\nu}(q) = \left(a_0 + \frac{1}{4\pi}\right)\epsilon_{\mu\nu\rho}q_\rho + \frac{1}{16|q|}\left(q^2\delta_{\mu\nu} - q_\mu q_\nu\right). \tag{2.51}$$

In particular, we have

$$c_0 = 4\pi^2 a_0 + \pi, \tag{2.52}$$

and the parity anomaly thus depends on the lattice action chosen. That the difference between the two cases considered is merely a sign in the anomaly coefficient c_0 is an accident as we shall discuss below.

In the large mass limit, the vacuum polarization reduces to

$$\lim_{m\to\infty} \Pi_{\mu\nu}(q) = a_0\,\epsilon_{\mu\nu\rho}q_\rho \tag{2.53}$$

and it thus follows that

$$c_\infty = 4\pi^2 a_0. \tag{2.54}$$

Comparing with eqs. (1.10) and (1.11), we see that $s = 1$ corresponds to $n = 0$ and $s = -1$ to $n = -1$.

So far we have only discussed the second-order contribution $\Gamma_2[A]$ to the effective action. As already mentioned, $\Gamma_k[A]$ vanishes for odd k. For even $k \geqslant 4$, it is possible to show, using the Reisz power counting theorem once more, that the continuum limit of $\Gamma_k[A]$ exists and that it is independent of the details of the lattice action, i.e. it is universal. In other words, the only non-universality in the effective action is contained in the coefficient a_0 which appears in the small momentum expansion (2.37) of the lattice vacuum polarization.

The final formulae (2.48)–(2.50) for $\Pi_{\mu\nu}(q)$ are in complete agreement with the result which one obtains using Pauli–Villars regularization with a single ghost field. The cases $s = \pm 1$ here correspond to the possibility to switch the sign of the ghost field mass term. We mention at this point that the values for the coefficient c_∞ quoted in the literature disagree in many cases with our result (the value most often quoted is one-half of our minimal value $c_\infty = \pm 2\pi$). The reason for this is that in those papers a counter-term proportional to the Chern–Simons action is added to the effective action in such a way that the amplitude $A(q^2)$ in eq. (2.48) vanishes for large q^2 (at the same time, the parity anomaly at $m = 0$ is cancelled). After that the expression for the so renormalized effective action is unambiguously determined and one finds $c_\infty = -\pi$. However, as discussed in sect. 1, we decided to stick to the definition (1.5) of the effective action (with some regularization satisfying the weak requirements listed below), and eqs. (1.10) and (1.11) are then the correct result for the coefficients c_0 and c_∞.

By adding further formally irrelevant operators to the fermion action (2.11), it is easy to construct lattice theories for which the continuum limit of the effective action is exactly the same as in the Wilson model except that the coefficient a_0 is any integer multiple of $1/2\pi$. For example, an action with $a_0 = -1/\pi$ is

$$S_F = -a^3 \sum_x \bar{\psi}(x)(D-m)(1-2aD)\psi(x), \tag{2.55}$$

where D is the Dirac operator (2.10) with $s = -1$. Of course, this simple action is equivalent to adding a second fermion to the Wilson model with mass $M = 1/2a$. In more complicated situations, such a factorization of the action will not be possible in general, but it is anyway likely that in models with quantum numbers $n \neq 0, -1$ in eq. (1.11), there are effectively further heavy particles in the theory with masses of the order of the cutoff.

3. Regularization scheme dependence of the effective action

Although the effective action $\Gamma[A]$ is finite in the limit where the ultra-violet cutoff is removed, a divergence is only avoided, because cancellations between individually divergent contributions occur. This is clearly seen in the lattice calculation, where e.g. the term proportional to D_2 in eq. (2.29) is linearly divergent. Moreover, the integral (2.38) has a lattice degree of divergence equal to zero so that a logarithmic divergence is expected to show up. That it does not is due to the anti-symmetric character of the integrand. In view of these remarks it is not altogether surprising that the effective action, though finite, is regularization dependent to a certain extent.

We here explain why, for any reasonable regularization of the effective action, one is bound to obtain the result (1.10), (1.11) for the coefficients c_0 and c_∞. The qualification "reasonable" here means that we only admit regularization prescriptions which are gauge invariant and which are consistent with respect to gauge group reductions. This latter term refers to the following property.

Suppose G is a compact Lie group and let R be any unitary representation of G in a complex space of finite dimension N. The associated representation of the Lie algebra of G by hermitian $N \times N$ matrices will also be denoted by R. We then consider a $2N$ component fermion field $\psi(x)^A_\alpha$ with Dirac index $\alpha = 1,2$ and "color" index $A = 1,\ldots,N$. Under a gauge transformation $g(x) \in \mathrm{G}$, $\psi(x)$ transforms as

$$\psi(x) \to \mathrm{R}(g(x))^{-1}\psi(x), \tag{3.1}$$

and, for any given external gauge field $\mathscr{A}_\mu(x)$ taking values in the Lie algebra of G, we can define a gauge invariant fermion action through eq. (1.2), where the Dirac operator D is given by*

$$D = \gamma_\mu\left(\partial_\mu + i\mathrm{R}\left(\mathscr{A}_\mu(x)\right)\right). \tag{3.2}$$

The associated effective action will be denoted by $\Gamma[\mathscr{A}]_{\mathrm{G,R}}$.

A regularization procedure is called consistent with respect to gauge group reductions if the following two conditions are satisfied. First, if R is a reducible representation, the addition formula

$$\Gamma[\mathscr{A}]_{\mathrm{G,R_1\oplus R_2}} = \Gamma[\mathscr{A}]_{\mathrm{G,R_1}} + \Gamma[\mathscr{A}]_{\mathrm{G,R_2}} \tag{3.3}$$

should hold. Secondly, if the gauge field $\mathscr{A}_\mu(x)$ takes values in the Lie algebra of a compact Lie subgroup H of G (so that it can be alternatively regarded as a gauge field with respect to the gauge group H), we have

$$\Gamma[\mathscr{A}]_{\mathrm{G,R}} = \Gamma[\mathscr{A}]_{\mathrm{H,R}}. \tag{3.4}$$

It is obvious that the lattice regularization, the Pauli–Villars cutoff procedure and many other regularization prescriptions satisfy these requirements.

In what follows we choose $\mathrm{G} = \mathrm{SU}(2)$ and $\mathrm{H} = \mathrm{U}(1)$, the subgroup of diagonal matrices. Furthermore, the representation R is taken to be the fundamental 2-dimensional representation so that $\mathrm{R}(\mathscr{A}_\mu(x)) = \mathscr{A}_\mu(x)$ is a traceless, hermitian 2×2 matrix. If we now choose

$$\mathscr{A}_\mu(x) = \begin{pmatrix} eA_\mu(x) & 0 \\ 0 & -eA_\mu(x) \end{pmatrix}, \tag{3.5}$$

the properties listed above imply that for any fixed regularization prescription, we have

$$\Gamma[\mathscr{A}]_{\mathrm{G,R}} = \Gamma[A] + \Gamma[-A]. \tag{3.6}$$

As will become clear shortly, the interest in this relation in the present context is that it allows us to deduce certain restrictions on the possible changes of the abelian effective action $\Gamma[A]$ when we switch to a different regularization scheme.

It is well known that the ultra-violet sensitive parts of the effective action $\Gamma[\mathscr{A}]_{\mathrm{G,R}}$ can be isolated in perturbation theory, i.e. we only need to study the fermion loops with $k \geqslant 2$ external gauge field legs. For $k = 2$, the degree of divergence of the loop integral in momentum space is 1 so that a regularization ambiguity can at most be a polynomial of degree 1 in the external momentum.

* For notational simplicity, we do not introduce a gauge coupling constant here. To match with the notation used in sect. 2 in the abelian case, set $\mathrm{R}(\mathscr{A}_\mu(x)) = \mathscr{A}_\mu(x) = eA_\mu(x)$.

Similarly, for $k=3$, the possible ambiguity is a constant and for $k \geqslant 4$ the loop integral is power counting convergent and hence independent of the regularization employed. It follows from these considerations that the difference $\Delta\Gamma[\mathscr{A}]_{\mathrm{G,R}}$ between the effective actions calculated in two different regularization schemes is given by

$$\Delta\Gamma[\mathscr{A}]_{\mathrm{G,R}} = \int \mathrm{d}^3x\, P\big(\mathscr{A}_\mu(x), \partial_\nu \mathscr{A}_\rho(x)\big), \tag{3.7}$$

where P is a sum of monomials of dimension 2 or 3 ($\mathscr{A}_\mu$ has dimension 1 and $\partial_\mu \mathscr{A}_\nu$ dimension 2). The only expression of this type which is invariant under infinitesimal gauge transformations is the Chern–Simons action

$$I[\mathscr{A}]_{\mathrm{G}} = \frac{1}{8\pi^2}\int \mathrm{d}^3x\, \epsilon_{\mu\nu\rho}\, \mathrm{tr}\big\{\mathscr{A}_\mu(x)\,\partial_\nu \mathscr{A}_\rho(x) + \tfrac{2}{3} i \mathscr{A}_\mu(x)\mathscr{A}_\nu(x)\mathscr{A}_\rho(x)\big\} \tag{3.8}$$

and thus we have

$$\Delta\Gamma[\mathscr{A}]_{\mathrm{G,R}} = icI[\mathscr{A}]_{\mathrm{G}} \tag{3.9}$$

for some (possibly complex) coefficient c.

At this point, a crucial observation (which has been made by several authors in recent years [16, 17, 20–22, 32]) is that the Chern–Simons action (3.8) is not invariant under "large" gauge transformations, i.e. gauge transformations $g(x) \in \mathrm{SU}(2)$ which are homotopically non-trivial. Instead what happens is that $I[\mathscr{A}]_{\mathrm{G}}$ changes by an integer equal to the winding number of $g(x)$. Since we insisted that the admitted regularization procedures respect gauge invariance, the Boltzmann factor $\exp(-\Gamma[\mathscr{A}]_{\mathrm{G,R}})$ must be invariant under such gauge transformations too, and hence we conclude that

$$c = 0 (\mathrm{mod}\, 2\pi). \tag{3.10}$$

Combining eqs. (3.6), (3.9), and (3.10), the result of our argumentation is that under a change of regularization scheme, the abelian effective action changes by

$$\Delta\Gamma[A] = i2\pi k e^2 I[A], \tag{3.11}$$

where k can be any integer. It is easy to check, e.g. by working out the vacuum polarization with a Pauli–Villars cutoff with several ghost fields and variable mass signs, that k in fact does assume any arbitrary integer value.

The quantization of the coefficients c_0 and c_∞ according to eqs. (1.10) and (1.11) is now immediate, since it holds on the lattice with Wilson fermions (as shown previously) and hence (by eq. (3.11)) for any other regularization of the effective action.

4. Physical particle properties

The generation of a Chern–Simons term through virtual fermion pairs has a rather dramatic impact on the observable properties of the fundamental particles in lattice QED. We describe the most important of these effects here and discuss, in particular, how the spin of charged particles is affected. The fermion action is assumed throughout to be given by eqs. (2.10), (2.11) with $s = \pm 1$ and $m < 1/a$ if $s = -1$. Also, we will restrict ourselves to the limit where the dimensionless parameter e^2/m is small so that perturbation theory and semi-classical arguments are applicable.

4.1. PHOTON MASS GENERATION

For the original Wilson model $s = 1$, the coefficient c_∞ vanishes and no unusual effects are expected to happen. In particular, the fermions are confined by the Coulomb potential

$$V(r) \underset{r\to\infty}{=} \alpha \ln r + \mathrm{O}(1), \qquad \alpha = \frac{e^2}{2\pi} + \mathrm{O}(e^4). \tag{4.1}$$

Besides the photon (which remains massless), the observable particles are heavy fermion anti-fermion bound states. The size of these particles is proportional to $(e^2 m)^{-1/2}$ and since the bound fermions can annihilate into photons, they are actually narrow resonances.

The physical picture just described follows from ordinary perturbation theory and the non-relativistic approximation (which should be appropriate when m is large). It is also corroborated by a well-known lattice technique, the hopping parameter expansion. To derive it, one first rewrites the fermion action (2.11) in the familiar form

$$S_\mathrm{F} = \sum_x \Big\{ \bar\chi(x)\chi(x) - K \sum_\mu \big[\bar\chi(x)(1+\gamma_\mu)U(x,\mu)\chi(x+a\hat\mu) + \bar\chi(x+a\hat\mu)(1-\gamma_\mu)U(x,\mu)^{-1}\chi(x) \big] \Big\}, \tag{4.2}$$

$$K = 1/(6+2am), \qquad \chi(x) = a\psi(x)/\sqrt{2K}\,. \tag{4.3}$$

The correlation functions of gauge invariant operators such as $\bar\chi(x)\chi(x)$ can then be expanded in powers of the "hopping" parameter K. The terms which one generates in this way can be readily evaluated, because the functional integral one has to calculate in each case is of the gaussian type. Since K goes to zero for large masses m, it is evident from this expansion that heavy fermions decouple. Confinement is also quite obvious, because the terms in the expansion which correspond to widely separated fermion lines are strongly suppressed by the associated gauge factor (Wilson loop).

An entirely different physical picture emerges when we choose $s = -1$. In this case we have $c_\infty = -2\pi$ and the photon acquires a mass. To see this, first note that the total gauge field action in the continuum limit and at $m = \infty$ is exactly given by

$$(S_G + \Gamma)[A] = \int d^3x \left\{ \tfrac{1}{4} F_{\mu\nu}(x) F_{\mu\nu}(x) - i \frac{e^2}{4\pi} \epsilon_{\mu\nu\rho} A_\mu(x) \partial_\nu A_\rho(x) \right\}. \tag{4.4}$$

This is the action of a free field theory and it is easy to convince oneself that the model describes a single particle with mass $\mu = e^2/2\pi$ and spin-1 [15,16]. Since the particle comes with only one possible spin orientation, the breaking of parity (which is manifest in the action) is an observable effect.

A photon mass is also generated for any fixed value of the ultra-violet cutoff $\Lambda = 1/a$ and for any fermion mass m in the range $0 < m < \Lambda$. This is evident from eqs. (2.37), (2.40), which imply that the photon pole in the propagator of the gauge field tensor $F_{\mu\nu}(x)$ is shifted away from zero. Furthermore, the photon mass μ is given by the asymptotic formula

$$\mu = e^2/2\pi + O(e^4) \tag{4.5}$$

(the terms of order e^4 are calculable in perturbation theory and are unlikely to vanish). We note incidentally that the fermion action (2.11) with $s = -1$ is also equivalent to the action (4.2), but here the hopping parameter K is related to the fermion mass through $K = (6 - 2am)^{-1}$. These values of K are above the critical line $K = 1/6 + O(ae^2)$ in the (e^2, K) plane and are therefore not accessible to the hopping parameter expansion, i.e. there is no contradiction to what we have said above about the small K region.

An immediate consequence of the fact that the photon has become massive is that the Coulomb potential $V(r)$ is no longer confining, but instead falls off like $e^{-\mu r}$ for large distances r. One expects, therefore, that stable charged particles exist, which have a mass approximately equal to m and which behave as free particles when separated from each other by a distance much greater than μ^{-1}.

4.2. FERMION–BOSON TRANSMUTATION

We now concentrate on the interesting case, $s = -1$, and proceed to discuss the influence of the radiatively generated Chern–Simons action on the observable spin of the fundamental charged particles in the theory. For simplicity we assume from now on that the continuum limit has been taken and that the fermion mass parameter m is so large that the total gauge action is accurately given by eq. (4.4) (possible corrections of order e^2/m are discussed at the end of this section).

As explained above, the fundamental fermions are deconfined and can be studied in isolation. The observable spin of these particles is naively expected to be 1/2, but actually this may not be true, because a charged particle is always accompanied by a static Coulomb field which may contribute to the total angular momentum of the state. Due to parity invariance, such an effect cannot occur in ordinary QED. However, the Chern–Simons term in eq. (4.4) violates parity and the associated Coulomb field, though static, circulates around the charge in such a way that the field angular momentum is exactly $-1/2$. Thus, the observable spin of the fundamental charged particles vanishes, and, assuming that the usual spin statistics connection holds, one is led to conclude that these particles are in fact bosons (for a recent discussion of spin and statistics in three dimensions, see refs. [27, 28]).

When calculating the angular momentum of the Coulomb field, we have employed the expression (in euclidean notation)

$$T_{\mu\nu} = F_{\mu\rho}F_{\nu\rho} - \tfrac{1}{4}\delta_{\mu\nu}F_{\rho\lambda}F_{\rho\lambda} \tag{4.6}$$

for the energy momentum tensor, which one obtains by studying the variation of the action (4.4) under an infinitesimal change of the space-time metric. While this is a natural definition, the fact that other choices for the energy–momentum tensor have been proposed has cast some doubt on the validity of the calculation [15]. In addition, since the Coulomb field is singular at the origin, it is not entirely obvious that there is no short distance contribution to the angular momentum other than the bare spin of the particle.

For these reasons we think it is useful to establish the vanishing of the observable spin of the fundamental charged particles along a different line, using a semi-classical approach. The argument follows to some extent Polyakov's derivation [7, 8], but since we enforce a semi-classical situation, formal manipulations with path integrals can be avoided and the whole discussion is actually completely rigorous.

In outline, we wish to show in what follows that the spin phase factor which is associated to the (semi-classical) propagation of a Dirac particle with charge e along a closed path $\mathscr{C}$ is exactly cancelled by the Wilson loop when the average over all gauge fields is taken with the action (4.4). Up to a normalization factor, the semi-classical amplitude for the propagation along $\mathscr{C}$ then is the same as for a neutral spin-0 particle so that one is led to conclude that the observable spin of the Dirac particle must vanish.

As a preparation, it is useful to first consider the case of a scalar particle whose propagator $G(x, y)$ is determined by

$$\left\{-\hbar^2\Delta + U(x)\right\}G(x, y) = \hbar^2\delta(x-y), \quad \Delta = \partial_\mu\partial_\mu . \tag{4.7}$$

Here, we have introduced Planck's constant $\hbar$, to ease the study of the semi-classical limit, and an external potential $U(x) > 0$, which we assume is smooth and equal to

m^2 for large x. Physically, this potential should be thought to represent the effect of some apparatus which forces the particle to move along certain curved trajectories in the limit $\hbar \to 0$. The device is needed in the present discussion, because the spin factor alluded to above is trivial for straight paths, i.e. the spin of a particle can only be observed through its response to an external force.

It is well known that the scalar propagator $G(x, y)$ can be represented by a Feynman–Kac path integral with a positive weight (we are in euclidean space). In this representation, the semi-classical nature of the limit $\hbar \to 0$ is obvious, because only those paths contribute in this limit which are infinitesimally close to a certain classical trajectory $r_\mu(t)$ (t is an arbitrary curve parameter with $0 \leqslant t \leqslant 1$ and $r(1) = x$, $r(0) = y$). The classical equations of motion, which $r(t)$ has to satisfy, imply that this trajectory is a geodesic with respect to the riemannian metric

$$h_{\mu\nu}(x) = U(x)\,\delta_{\mu\nu}. \tag{4.8}$$

Furthermore, for the semi-classical expansion of the propagator $G(x, y)$ (which is just a saddle point expansion of the path integral about this curve), one obtains

$$G(x, y) = \mathrm{e}^{-\Omega(x, y)/\hbar}\{g_0(x, y) + \hbar g_1(x, y) + \cdots\}, \tag{4.9}$$

where $\Omega(x, y)$ denotes the geodesic distance between x and y and the leading order amplitude $g_0(x, y)$ is given by

$$g_0(x, y) = U(x)^{-1/2}\frac{|\det M(x, y)|^{1/2}}{4\pi\Omega(x, y)}U(y)^{-1/2}, \tag{4.10}$$

$$M(x, y)_{\mu\nu} = \frac{1}{2}\frac{\partial^2}{\partial x_\mu\,\partial y_\nu}[\Omega(x, y)]^2. \tag{4.11}$$

A derivation of this result, and also of all the other semi-classical formulae which follow, is sketched in appendix A.

We now proceed to discuss the propagation of a Dirac particle of charge e in a background gauge field $A_\mu(x)$. The associated propagator $S(x, y)_{\alpha\beta}$ is defined through

$$\left\{-\hbar D + \sqrt{U(x)}\right\}S(x, y) = \hbar^2\,\delta(x - y), \tag{4.12}$$

where D is the Dirac operator (1.3) and $U(x)$ the potential introduced above. The semi-classical expansion of $S(x, y)$ reads

$$S(x, y) = \mathrm{e}^{-\Omega(x, y)/\hbar}\{s_0(x, y) + \hbar s_1(x, y) + \cdots\} \tag{4.13}$$

with $\Omega(x, y)$ as before and

$$s_0(x, y)_{\alpha\beta} = 2U(x)^{1/4} g_0(x, y) U(y)^{1/4} z_\alpha(1) \exp\{-i(\Phi_g + \Phi_s)\} \bar{z}_\beta(0). \quad (4.14)$$

Besides a positive factor which is essentially the same as in the scalar case, this formula involves two phases Φ_g, Φ_s and a complex two-component spinor $z_\alpha(t)$ which are defined as follows. First, the gauge phase Φ_g is given by the familiar expression

$$\Phi_g = e\int_0^1 dt\, \dot{r}_\mu(t) A_\mu(r(t)), \quad (4.15)$$

where the dot indicates a derivative with respect to t. The spinor $z_\alpha(t)$ is best thought to be a CP^1 field on the classical trajectory $r(t)$. Up to an arbitrary t-dependent phase it is determined by

$$(\dot{r}_\mu(t)\gamma_\mu + |\dot{r}(t)|) z(t) = 0, \qquad |z(t)| = 1. \quad (4.16)$$

An associated CP^1 gauge field $\mathscr{A}(t)$ may then be introduced through

$$\mathscr{A}(t) = \frac{1}{i}\bar{z}(t) \cdot \dot{z}(t), \qquad \bar{z} = (z_1^*, z_2^*), \quad (4.17)$$

and the spin phase Φ_s is finally given by

$$\Phi_s = \int_0^1 dt\, \mathscr{A}(t). \quad (4.18)$$

Note that the combination (4.14) is invariant under the "gauge" transformation $z(t) \to e^{i\alpha(t)} z(t)$, and the phase ambiguity involved in the definition of $z_\alpha(t)$ is hence irrelevant.

We now consider the semi-classical propagation of a Dirac particle along a closed curve $\mathscr{C}$ parametrized by $r_\mu(t)$ ($0 \leqslant t \leqslant 1$, $r(0) = r(1)$; we are also free to choose $z(t)$ such that $z(0) = z(1)$). If we take the trace over Dirac indices, the total phase factor associated with the propagation is $\exp - i(\Phi_s + \Phi_g)$. After integrating over all gauge fields $A_\mu(x)$, using eq. (4.4) for the gauge field action, the amplitude becomes

$$W(\mathscr{C}) \exp\{-i\Phi_s\}, \quad (4.19)$$

where

$$W(\mathscr{C}) = \left\langle \exp\left(-ie\oint_{\mathscr{C}} dr_\mu A_\mu(r)\right)\right\rangle \quad (4.20)$$

denotes the Wilson loop.

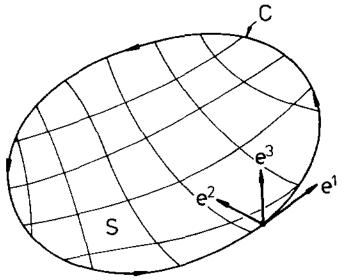

Fig. 1. Picture of a closed classical trajectory $\mathscr{C}$ of a Dirac particle, together with the associated surface $\mathscr{S}$ and the moving frame e^a.

$W(\mathscr{C})$ has recently been calculated by Polyakov [7, 8] in the limit where the curve $\mathscr{C}$ is large compared to the screening length μ^{-1}. In the derivation one also has to assume that $\mathscr{C}$ is at least twice continuously differentiable and that it is the boundary of a smooth surface $\mathscr{S}$ which is not self-intersecting. Since the shape of $\mathscr{C}$ is controlled by the potential $U(x)$, we are free to arrange an "experimental" set-up such that the classical particle trajectory has all these properties. To state the result of Polyakov's calculation, we introduce a moving frame $e^a(t)$, $a = 1, 2, 3$, along $\mathscr{C}$ with

$$e^a \cdot e^b = \delta^{ab}, \qquad e^1 \cdot (e^2 \times e^3) = 1, \tag{4.21}$$

and $e^1_\mu = \dot{r}_\mu / |\dot{r}|$ (cf. fig. 1). Furthermore, we choose e^2 such that it is tangential to the surface $\mathscr{S}$ and inward pointing. e^3 is then normal to $\mathscr{S}$.

As one runs along $\mathscr{C}$, the moving frame rotates in the plane orthogonal to e^1 with an angular velocity

$$\omega = \dot{e}^2 \cdot e^3. \tag{4.22}$$

Polyakov's formula for the Wilson loop now simply reads*

$$W(\mathscr{C}) = \exp\left\{ -cL + \frac{i}{2} \int_0^1 \mathrm{d}t\, \omega(t) \right\}, \tag{4.23}$$

where c is some constant and L the length of the loop. Thus, in the limit of a large loop $\mathscr{C}$, the phase of $W(\mathscr{C})$ is exactly one-half of the total (projected) angle of rotation of the basis vectors e^2, e^3. Note that the phase vanishes for a planar curve; it is only non-trivial when the torsion of $\mathscr{C}$ is non-zero.

* The coefficient multiplying the integral over $\omega(t)$ in eq. (4.23) is inversely proportional to the coefficient of the Chern–Simons term in the gauge field action. The value quoted is appropriate for the action (4.4). Unfortunately, our result for this number does not quite agree with ref. [7].

The phase factor in eq. (4.23) can be written in various equivalent forms. As we show in appendix B, one possibility is to express it in terms of the CP^1 field $z_\alpha(t)$ introduced earlier. In fact what we prove is that

$$\tfrac{1}{2}\int_0^1 \mathrm{d}t\, \omega(t) = \int_0^1 \mathrm{d}t\, \mathscr{A}(t) + \pi (\mathrm{mod}\, 2\pi)\,. \qquad (4.24)$$

Recalling eqs. (4.18) and (4.19), we now see that the spin phase associated with the propagation of the Dirac particle is exactly cancelled by the gauge factor. Thus, in the experimental situation considered, there is apparently no visible spin effect left (apart from normalization factors) and so we are led to conclude that the observable spin of the fundamental charge particles vanishes.

Although we have reached this conclusion neglecting higher order corrections to the total gauge field action (4.4), we believe that it actually remains unaffected to any order of e^2/m and thus, barring non-perturbative effects, is most likely an exact property of the model. To understand why this is so, recall that the semi-classical motion of the particle is completely controlled by the potential $U(x)$ and thus does not depend on e^2/m. The only place where this parameter enters the argumentation above is when we calculate the Wilson loop $W(\mathscr{C})$. Here we should really use the complete gauge field action and this leads to a number of additional contributions.

The crucial input at this point is the observation that the full vacuum polarization tensor $\Pi_{\mu\nu}(q)$ has the Lorentz structure (2.48) and $A(0) = A(0)|_{\text{1-loop}}$, i.e. all higher loop contributions to $A(0)$ vanish. This result has been established by explicit calculation to two loops [29, 30], and later, using Ward identities, to any number of loops [31]. Since the parity odd term in $\Pi_{\mu\nu}(q)$ dominates at small q, the full propagator at large distances in position space is equal to the one-loop propagator and it follows that the sum of all those contributions to the Wilson loop, which only involve the full propagator, gives exactly the phase (4.23) (for a large loop $\mathscr{C}$). Thus, it remains to be shown that all other contributions of $W(\mathscr{C})$ are negligible. We have not attempted to give a completely rigorous proof that this is indeed the case, but considering a few examples we convinced ourselves that for dimensional reasons, these terms must vanish with an inverse power of the loop size.

Thus, as far as we can see, there are no corrections to the phase of the Wilson loop and our semi-classical argument hence remains valid, so that the observable spin of the fermion is really zero to all orders of e^2/m.

5. Conclusions

The basic result obtained in this paper is that the effective gauge field action $\Gamma[A]$ due to a single, massive two-component fermion is, though finite, not uniquely determined by the classical fermion action (1.3), but depends to a certain extent on the ultra-violet regularization scheme employed. Making some weak assumptions on the properties of the admitted schemes, we have shown that in the limit where the cutoff is removed, the various regularizations fall into universality classes labelled by an integer n. The difference $\Delta\Gamma[A]$ between the effective actions corresponding to the classes n_1 and n_2 is given by

$$\Delta\Gamma[A] = i2\pi(n_1 - n_2)e^2 I[A], \tag{5.1}$$

where $I[A]$ is the Chern–Simons action (1.1). In particular, the Wilson lattice fermions studied in detail in sect. 2 correspond to $n=0$ and $n=-1$, respectively, depending on the sign of the (formally irrelevant) Wilson term in the action. We mention in passing that universality is restored, if one introduces several (massive) fermions in the model in such a way that they form parity doublets, and if the regularization employed respects parity. This is the case, for example, for staggered fermions [33].

For all values of n, the parity invariance of the classical fermion action (1.2) at $m=0$ is broken by quantum effects, i.e. there is always an anomaly. The anomaly coefficient c_0 depends on n and is given by eqs. (1.10), (1.11). The physical properties of lattice QED (with no Chern–Simons term in the bare gauge field action) are largely determined by the radiatively generated Chern–Simons action, which comes with a coefficient proportional to n. Thus, while there are no special effects for $n=0$, the radiative corrections to the gauge field action induce a photon mass for $n \neq 0$. Furthermore, the fundamental charged particles in the theory, which would naively be expected to be spin 1/2 fermions, are deconfined and assume fractional spin and statistics. We have studied this effect in detail for the special case $n=-1$, which is realized by one of the lattice models considered. Using a semi-classical argument, we found that the fermions acquire an additional spin of exactly $-1/2$ so that the total observable spin is equal to 0. This result is in quantitative agreement with the simple physical picture that the additional spin of a charged particle arises from the attached Coulomb field, which, due to the Chern–Simons term in the total gauge field action, carries angular momentum.

That the fermion in QED can turn into a boson through its own radiative effects, is a curious phenomenon for which we have no simple physical explanation. That is, what we do not really understand is why, for $n=-1$, the coefficient c_∞ assumes precisely the value needed to convert the fermions into spinless bosons. At present, it is just the result of a Feynman diagram calculation (with some topological flavour), which is apparently unrelated to the spin changing mechanism.

Although we have not studied the cases with $n \neq 0, -1$ in any detail, it is quite clear from the above that the total spin carried by the fundamental charged particles will be fractional in general, because the Coulomb field attached to the particle carries an angular momentum equal to $1/2n$ and the total angular momentum is hence $(n+1)/2n$. For the same reasons we have argued that the total spin vanishes exactly for $n = -1$, we do not expect that this result will receive any corrections of order e^2/m. However, we would like to stress at this point that it would certainly be desirable to clarify the question of such corrections on a more rigorous level, perhaps by performing partial summations of perturbation theory in a systematic way. Such improved techniques would also be required if one wants to calculate the scattering matrix of the fundamental charged particles.

We finally mention that all our calculations concerning the effective action can immediately be carried over to the case of non-abelian gauge theories. In particular, the regularization dependence, the associated universality classes and the parity anomaly are as in the abelian case, with the obvious changes. The physical implications of the radiatively generated Chern–Simons term at large fermion masses are, however, not entirely obvious, because the gauge field is itself charged in this case and the gauge particles thus may change their properties in an even more radical way than the photon in QED. Still, some interesting progress has recently been made in the theory with pure Chern–Simons action (and no fermions) [32].

A.C. is indebted to C.P. Korthals-Altes and O. Napoly for prior work on anomalies, and to A. Luther and P. Di Vecchia for kind hospitality in Copenhagen.

Note added in proof

We were kindly informed by S. Aoki that the parity anomaly in 3-dimensional lattice gauge theories has previously been studied by H. So [35].

Appendix A

We here derive the semi-classical expansion (4.13), (4.14) of the Dirac propagator $S(x, y)$. To this end, we first consider the analogous expansion of the scalar propagator $G(x, y)$. Inserting eq. (4.9) in the differential equation (4.7) and equating the terms of the same order of $\hbar$, one obtains

$$p_\mu p_\mu = U(x), \tag{A.1}$$

$$p_\mu \partial_\mu g_0(x, y) = -\tfrac{1}{2}\Delta\Omega(x, y) g_0(x, y), \tag{A.2}$$

where we have introduced the abbreviation

$$p_\mu = \partial_\mu \Omega(x, y). \tag{A.3}$$

In addition, $\Omega(x, y)$ and $g_0(x, y)$ must satisfy the boundary conditions

$$\Omega(x, x) = 0, \qquad \Omega(x, y) \geqslant 0, \tag{A.4}$$

$$g_0(x, y) \underset{x \to y}{\sim} \frac{1}{4\pi|x-y|}. \tag{A.5}$$

With the help of the space-time metric (4.8), eq. (A.1) can be written as

$$h^{\mu\nu}p_\mu p_\nu = 1\,. \tag{A.6}$$

This is a first-order partial differential equation for $\Omega(x, y)$. The associated characteristic curves are precisely the geodesics in the metric $h_{\mu\nu}$ and, taking into account the boundary condition (A.4), it thus follows that $\Omega(x, y)$ is the geodesic distance between x and y. Furthermore, p_μ is tangential to the geodesic $r(t)$, connecting x and y, at the end point $x = r(1)$.

To solve eq. (A.2), first note that it can be rewritten in the form

$$h^{\mu\nu}\,\partial_\mu\Omega\,\partial_\nu g_0' = -\frac{1}{2}\left(\frac{1}{\sqrt{h}}\,\partial_\mu\sqrt{h}\,h^{\mu\nu}\,\partial_\nu\Omega\right)g_0'\,, \tag{A.7}$$

where h denotes the determinant of $h_{\mu\nu}$ and

$$g_0'(x, y) = U(x)^{-1/4}g_0(x, y)\,. \tag{A.8}$$

Treating Ω and g_0' as scalar functions, eq. (A.7) is invariant under general coordinate transformations and we may take advantage of this fact by passing to Riemann normal coordinates $s_\mu(x)$, choosing y for the origin of the coordinate system. Explicitly, for any fixed y, $s_\mu(x)$ is defined by

$$s_\mu(x) = -\Omega(x, y)U(y)^{-1/2}\frac{\partial}{\partial y_\mu}\Omega(x, y)\,. \tag{A.9}$$

In these coordinates, we have $\Omega = |s|$, and the metric

$$\tilde{h}_{\mu\nu} = h_{\rho\sigma}\frac{\partial x_\rho}{\partial s_\mu}\frac{\partial x_\sigma}{\partial s_\nu} \tag{A.10}$$

satisfies

$$\tilde{h}_{\mu\nu}s_\nu = s_\mu\,. \tag{A.11}$$

Furthermore, eq. (A.7) reduces to

$$s_\mu\left(\partial_\mu + \tfrac{1}{4}\,\partial_\mu \ln \tilde{h}\right)g_0' = -g_0' \tag{A.12}$$

(the derivatives here are with respect to s_μ). Taking the boundary condition (A.5) for $s \to 0$ into account, this differential equation is easy to integrate and one obtains

$$g_0(x, y) = \tilde{h}(s)^{-1/4}\frac{(U(x)U(y))^{1/4}}{4\pi\Omega(x, y)}\,. \tag{A.13}$$

The final result (4.10), (4.11) now follows from eqs. (A.9), (A.10).

To derive the semi-classical expansion of the Dirac propagator $S(x, y)$, we proceed in exactly the same way as in the scalar case. From the differential equation (4.12) it follows that

$$\left\{ p_\mu \gamma_\mu + \sqrt{U(x)} \right\} s_0(x, y) = 0, \tag{A.14}$$

$$\left\{ p_\mu \gamma_\mu + \sqrt{U(x)} \right\} s_1(x, y) = D s_0(x, y). \tag{A.15}$$

In addition, to reproduce the δ-function singularity on the r.h.s. of eq. (4.12), the boundary condition

$$s_1(x, y) \underset{x \to y}{\sim} -\frac{\gamma_\mu (x_\mu - y_\mu)}{4\pi |x-y|^3} \tag{A.16}$$

must hold. For eq. (A.14) to admit non-vanishing solutions it is necessary that the determinant of the matrix in the curly bracket vanishes. This condition is equivalent to eq. (A.1) and $\Omega(x, y)$ is hence the same function as above.

We now introduce a CP^1 field $z_\alpha(x, y)$ through

$$(p_\mu \gamma_\mu + |p|) z = 0, \qquad |z| = 1. \tag{A.17}$$

If we restrict $z_\alpha(x, y)$ to the geodesic $x = r(t)$, the CP^1 field $z_\alpha(t)$, as defined in subsect. 4.2, is recovered. We may therefore use the same symbol z_α for the two fields without creating any confusion. Through eq. (A.17), z_α is determined up to a phase. The general solution of eq. (A.14) is thus given by

$$s_0(x, y)_{\alpha\beta} = z_\alpha(x, y) \bar{w}_\beta(x, y), \tag{A.18}$$

where $\bar{w}_\beta(x, y)$ is an arbitrary complex spinor.

To compute $\bar{w}_\beta(x, y)$, we must refer to eq. (A.15). Contracting with $\bar{z}_\alpha$, we have

$$\bar{z} \cdot \gamma_\mu (\partial_\mu + ieA_\mu(x)) s_0 = 0 \tag{A.19}$$

and hence

$$p_\mu (\partial_\mu + ieA_\mu(x)) \bar{w} = |p| (\bar{z} \cdot \gamma_\mu \partial_\mu z) \bar{w}. \tag{A.20}$$

Next, we make use of the identity $\epsilon_{\mu\nu\rho} \partial_\nu p_\rho = 0$ to show that

$$|p| (\bar{z} \cdot \gamma_\mu \partial_\mu z) = -\tfrac{1}{2} \Delta \Omega + p_u \left(\tfrac{1}{2} \partial_\mu \ln|p| - \bar{z} \cdot \partial_\mu z \right). \tag{A.21}$$

Inserting this result in eq. (A.20) and recalling the solution of eq. (A.2), one obtains

$$\bar{w}_\beta(x, y) = U(x)^{1/4} g_0(x, y) \exp\{ -i(\Phi_g + \Phi_s) \} \bar{u}_\beta(x, y), \tag{A.22}$$

where the phases Φ_g and Φ_s are given by eqs. (4.15) and (4.18). The spinor $\bar{u}_\beta$ is constant along the classical trajectory, viz.

$$\frac{d}{dt}\bar{u}_\beta(r(t), y) = 0, \tag{A.23}$$

but is otherwise arbitrary at this point. However, via eq. (A.15), the boundary condition (A.16) implies a certain behaviour of $s_0(x, y)$ in the limit $x \to y$ and this then is sufficient to fix $\bar{u}_\beta$ completely. As a result one obtains the desired expression (4.14) for the amplitude $s_0(x, y)$.

Appendix B

In this appendix we establish the relation (4.24). For notational convenience, the dependence on the curve parameter t of the quantities involved is not explicitly indicated in what follows. Starting from the defining equation

$$(e^1 \cdot \gamma + 1)z = 0, \tag{B.1}$$

a solution to the associated equation

$$(e^1 \cdot \gamma - 1)w = 0 \tag{B.2}$$

is given by

$$w_1 = z_2^*, \qquad w_2 = -z_1^*. \tag{B.3}$$

It is easy to check that z and w form an orthonormal basis in the space of all complex two-component spinors.

Consider now the (complex) vector $v_\mu = \bar{w} \cdot \gamma_\mu z$. Contracting with e_μ^1 and using eq. (B.1), it follows immediately that this vector is orthogonal to e^1 and hence given by

$$\boldsymbol{v} = a\boldsymbol{e}^2 + b\boldsymbol{e}^3, \tag{B.4}$$

where a and b are some complex numbers. Next, using the identity

$$\gamma_\mu \gamma_\nu = \delta_{\mu\nu} + i\epsilon_{\mu\nu\rho}\gamma_\rho, \tag{B.5}$$

one obtains

$$i\epsilon_{\mu\nu\rho}v_\rho = \bar{w} \cdot \gamma_\mu \gamma_\nu z, \tag{B.6}$$

which, when contracted with e_ν^1, yields $\boldsymbol{e}^1 \times \boldsymbol{v} = i\boldsymbol{v}$. Inserting eq. (B.4), we conclude that

$$\boldsymbol{v} = a(\boldsymbol{e}^2 - i\boldsymbol{e}^3). \tag{B.7}$$

Now we employ another identity for the γ-matrices,

$$(\gamma_\mu)_{\alpha\beta}(\gamma_\mu)_{\gamma\delta} = 2\delta_{\alpha\delta}\delta_{\beta\gamma} - \delta_{\alpha\beta}\delta_{\gamma\delta}, \tag{B.8}$$

to show that

$$v_\mu^* v_\mu = 2|a|^2 = 2. \tag{B.9}$$

In other words, a is a pure phase which we may parametrize by

$$a = e^{i\alpha}, \qquad \alpha(1) = \alpha(0) + 2\pi\nu, \qquad \nu \in \mathbb{Z}. \tag{B.10}$$

Finally, recalling the definition (4.22) of the angular velocity ω and using the identity (B.8) once more, it follows that

$$v_\mu^* \dot{v}_\mu = 2i(\dot{\alpha} + \omega) = 4\bar{z}\cdot\dot{z}, \tag{B.11}$$

and hence

$$\tfrac{1}{2}\int_0^1 \mathrm{d}t\,\omega(t) = \int_0^1 \mathrm{d}t\,\mathscr{A}(t) - \nu\pi. \tag{B.12}$$

That the winding number ν is odd follows from a topological consideration (deform $\mathscr{C}$ to a circle).

References

[1] G. Baskaran and P.W. Anderson, Phys. Rev. B37 (1988) 580

[2] E. Dagotto, On the Relation between High T_c Superconductors and Lattice Gauge Theories, talk given at the Symposium on Lattice Field Theory, Fermilab (1988), to be published in the proceedings

[3] J.B. Marston, U(1) Gauge Theory of the Heisenberg Antiferromagnet, Princeton preprint (1988)

[4] P.B. Wiegmann, Phys. Rev. Lett. 60 (1988) 821

[5] I. Dzyaloshinskii, A.M. Polyakov and P.B. Wiegmann, Phys. Lett. A127 (1988) 112

[6] F. Wilczek and A. Zee, Phys. Rev. Lett. 51 (1983) 2250

[7] A.M. Polyakov, Mod. Phys. Lett. A3 (1988) 325

[8] A.M. Polyakov, in "Fields, strings and critical phenomena," E. Brézin and J. Zinn-Justin eds., Les Houches session XLV (North-Holland, Amsterdam) (to appear)

[9] P.K. Panigrahi, S. Roy and W. Scherer, Rotational Anomaly and Fractional Spin in the Gauged CP^1 non-linear Sigma Model with Chern–Simons Term, Rochester preprint, UR-1062 (1988)

[10] S. Deser and A.N. Redlich, Phys. Rev. Lett. 61 (1988) 1541

[11] A.S. Goldhaber, R. MacKenzie and F. Wilczek, Field Corrections to Induced Statistics, Harvard preprint, HUTP-88/A044

[12] T.H. Hansson, M. Rocek, I. Zahed and S.C. Zhang, Phys. Lett. B214 (1988) 475

[13] L. Huerta and M. Ruiz-Altaba, Phys. Lett. B216 (1989) 371

[14] X.G. Wen and A. Zee, On the possibility of a Statistics-Changing Phase Transition, Santa Barbara preprint NSF-ITP-88-114

[15] J. Schonfeld, Nucl. Phys. B185 (1981) 157

[16] S. Deser, R. Jackiw and S. Templeton, Ann. Phys. (NY) 140 (1982) 372

[17] R. Jackiw, in Relativity, groups and topology II, B.S. DeWitt and R. Stora, eds., Les Houches session XL, (North-Holland, Amsterdam, 1984)
[18] A. Niemi and G.W. Semenoff, Phys. Rev. Lett. 23 (1983) 2077
[19] A.N. Redlich, Phys. Rev. D29 (1984) 2366
[20] L. Alvarez-Gaumé, S. Della Pietra and G. Moore, Ann. Phys. (NY) 163 (1985) 288
[21] S. Forte, Nucl. Phys. B288 (1987) 252
[22] S. Forte, Nucl. Phys. B301 (1988) 69
[23] A. Coste, C.P. Korthals-Altes and O. Napoly, Nucl. Phys. B289 (1987) 645;
S. Aoki, Phys. Rev. D35 (1987) 1435
[24] T. Jolicoeur, R. Lacaze and O. Napoly, Nucl. Phys. B293 (1987) 215
[25] T. Reisz, Commun. Math. Phys. 116 (1988) 81
[26] M. Lüscher, Selected Topics in Lattice Field Theory, in Fields, strings and critical phenomena, E. Brézin and J. Zinn-Justin, eds., Les Houches session XLV (North-Holland, Amsterdam) (to appear)
[27] J. Fröhlich, Statistics and monodromy in two- and three-dimensional quantum field theory, Proc. 1987 Como Conference, eds. K. Bleuler et al.
[28] J. Fröhlich and P.-A. Marchetti, Quantum field theory of anyons, Padua preprint, 88-0591
[29] M. Bernstein and T. Lee, Phys. Rev. D32 (1985) 1020
[30] Y. Kao and M. Suzuki, Phys. Rev. D31 (1985) 2137
[31] S. Coleman and B. Hill, Phys. Lett. B159 (1985) 184
[32] E. Witten, Commun. Math. Phys. 121 (1989) 351
[33] C. Burden and A.N. Burkitt, Europhys. Lett. 3 (1987) 545
[34] J. Smit, Acta Phys. Pol. B17 (1986) 531
[35] H. So, Prog. Theor. Phys. 73 (1985) 528; 74 (1985) 585