

## A Proof of the Nielsen–Ninomiya Theorem

D. Friedan\*

Service de Physique Theorique, CEN-Saclay, F-91190 Gif-sur-Yvette, France

**Abstract.** The Nielsen–Ninomiya theorem asserts the impossibility of constructing lattice models of non-selfinteracting chiral fermions. A new proof is given here. This proof fills a technical gap in the two proofs presented by the authors of the theorem. It also serves as prelude to an investigation of the chiral properties of the general lattice model.

### 1. Introduction

Nielsen and Ninomiya [1, 2] have demonstrated that there can be no net chirality in a lattice model of fermions in which the Hamiltonian satisfies the following conditions:

- (1) it is quadratic in the fields;
- (2) it is invariant under change of the phase of the fields;
- (3) it is invariant under translations of the (cubic) lattice; and
- (4) it is local, specifically in the sense that it is continuous in momentum space.

One way of coping with this result is to suppose that chiral fermions on the lattice must be selfinteracting. The Nielsen–Ninomiya theorem then might have a generalization equating the number of chiral fermions to some quantity associated with their interaction. This quantity would have to vanish with the selfinteraction. As a first step towards such a generalization, we reprove the original theorem here, in language potentially adaptable to models with selfinteraction. The structure of the proof suggests that the number of chiral fermions in the general lattice model might be determined entirely from examination of its high energy spectrum.

In [1], the theorem was proved using arguments from homotopy theory (algebraic topology). A second proof [2] called on intersection theory (differential topology). The present proof uses calculus (differential geometry). It is technically more complete than the original proofs. First, it shows the mathematical content to be a part of the theory of characteristic classes. Second, it is easily carried

---

\* address after 1 October 1981: Enrico Fermi Institute, University of Chicago, Chicago, IL 60637

through for a lattice of any odd dimension. Third, it deals with a somewhat esoteric technical possibility ignored in [1] and [2].

The theorem states, precisely, that in each irreducible representation of the internal symmetry group there are no chiral fermions. The internal symmetries are assumed to satisfy a locality condition analogous to (4) above. Nielsen and Ninomiya proved the absence of chiral fermions in a model containing  $N$  single component fermion fields. They did not distinguish among different representations of the internal symmetries. They then asserted that the fields transforming under a given representation  $Q$  of the internal symmetries (i.e., the fields of charge  $Q$ ) formed a sub-model of the same type as the original: consisting of  $N'$  single component fermion fields. With this assertion no more remains to be proved. But in fact it is completely consistent with locality that the fields of charge  $Q$  might form a nontrivial vector bundle over momentum space, in which case they would not consist of a set of single component fields. The proofs of [1] and [2] would have to be made rather more elaborate to account for this possibility.

The paper is organized as follows. The theorem is stated in Sect. 2 and proved in Sect. 3. Comments are provided in Sect. 4.

## 2. The Theorem

This section duplicates, for the most part, material presented in [1]. The principal differences are: (1) the dimension of space is an arbitrary odd positive integer; (2) a modified locality condition on the internal symmetries is used and (3) no mention is made of generic intersection of energy levels.

We consider free fermions on a cubic lattice of odd dimension  $d$ . They are described by an  $N$ -component fermion field  $\tilde{\psi}^a(x)$  defined on the lattice  $Z^d$ . The lattice momentum  $\mathbf{p}$  dual to the translation group  $Z^d$  ranges over the torus  $T^d$  (i.e., the first Brouillon zone with periodic boundary conditions). The Fourier transform of the field satisfies the canonical anti-commutation relations

$$[\psi^a(\mathbf{p}), \psi^b(\mathbf{q})]_+ = 0, \quad (1)$$

$$[\psi_b^*(\mathbf{p}), \psi^a(\mathbf{q})]_+ = \delta_b^a \delta(\mathbf{p} - \mathbf{q}). \quad (2)$$

The Hamiltonian, according to conditions (1–4) above, takes the form

$$H = \int d\mathbf{p} \psi_a^*(\mathbf{p}) K_b^a(\mathbf{p}) \psi^b(\mathbf{p}), \quad (3)$$

where  $K(\mathbf{p})$  is an  $N \times N$  hermitian matrix which depends smoothly on  $\mathbf{p}$ . The degree of smoothness has only temporary technical significance. In the end, the relevant property will be continuity in momentum space.

We suppose that the internal symmetries of the model act linearly on the fundamental fields and that they are translation invariant and local. The action of an internal symmetry is thus of the form  $\psi^a(\mathbf{p}) \rightarrow T_b^a(\mathbf{p}) \psi^b(\mathbf{p})$ , where  $T(\mathbf{p})$  is unitary and depends smoothly on  $p$ . For each irreducible representation  $Q$  of the internal symmetries there exists a projection  $P_Q(\mathbf{p})$ , depending on  $\mathbf{p}$ , such that the fields of charge  $Q$  are exactly those satisfying  $P_Q(\mathbf{p}) \psi_Q(\mathbf{p}) = \psi_Q(\mathbf{p})$ . The internal symmetries

commute with the Hamiltonian, so

$$[K(\mathbf{p}), P_Q(\mathbf{p})] = 0. \tag{4}$$

It will be evident below that the appropriate locality condition on the internal symmetries is that the projections  $P_Q(\mathbf{p})$  should depend smoothly on  $\mathbf{p}$ . This is stronger than the condition that the symmetry transformations  $T(\mathbf{p})$  be smooth in  $\mathbf{p}$ . The discrete chiral symmetries of [3], for example, satisfy the latter but not the former. The smoothness condition on  $P_Q(\mathbf{p})$  also expresses the possibility of coupling the charge  $Q$  locally to an external field, which is to say, the possibility of observing it.

In [1] and [2] the implicit assumption is made that  $P_Q(\mathbf{p})$  does not depend at all on  $\mathbf{p}$ . In such a case the charge  $Q$  portion of the field, i.e. the range of  $P_Q$ , can be represented as an  $N'$ -component field  $\psi_Q^a(\mathbf{p})$ . The same is true given the weaker assumption that there is some  $\mathbf{p}$ -dependent unitary transformation  $U_Q^a(\mathbf{p})$ , depending smoothly on  $\mathbf{p}$ , such that  $UP_QU^{-1}(\mathbf{p})$  is constant in  $\mathbf{p}$ . This amounts to the assumption that the charge  $Q$  portion of the field forms a trivial vector bundle over momentum space. But there do in fact exist nontrivial complex vector bundles over the torus  $T^d$ ; that is, projections  $P_Q(\mathbf{p})$  which cannot be made constant. The proof must take account of this possibility. It happens to be a somewhat nontrivial mathematical fact that a sufficiently local  $P_Q(\mathbf{p})$  can always be made constant.  $P_Q(\mathbf{p})$  must be such that its Fourier transform  $\tilde{P}_Q(\mathbf{x})$  is nonzero only on a finite set of points  $\mathbf{x}$ . [4] In the present circumstances such a locality condition is artificially strong.

The theorem now states that in each irreducible representation  $Q$  of the internal symmetries there must be equally as many left- as right-handed low energy elementary excitations (i.e. massless fermions). If we take the difference between the numbers of left and right handed fermions to be the number of chiral fermions, then the statement is that the number of chiral fermions of charge  $Q$  is zero, for every  $Q$ .

The low energy spectrum is to be found in a neighbourhood of the Fermi surface, which consists of those momenta  $\mathbf{p}_\alpha$  at which one or more of the eigenvalues of  $K(\mathbf{p})$  is zero. We require, for the existence of an acceptable continuum limit, that the low energy spectrum be interpretable as that of collection of relativistic massless fermions. This means, first of all, that the Fermi surface cannot be a surface; it must contain only a finite number of points  $\mathbf{p}_\alpha$ . Otherwise an accumulation point would exist at which the spectrum could not possibly appear relativistic.

Next, the elementary low energy excitations at momenta  $\mathbf{p}$  near  $\mathbf{p}_\alpha$  must resemble massless fermions. These excitations correspond to the eigenvectors of  $K(\mathbf{p})$  with eigenvalue near zero. We can always find unitaries  $U(\mathbf{p})$ , smooth in  $\mathbf{p}$  near  $\mathbf{p}_\alpha$ , which rotate the space of low energy eigenvectors to the space spanned by the first  $N_\alpha$  components. The requirement of a relativistic spectrum then becomes that, for  $\mathbf{p} \simeq \mathbf{p}_\alpha$ ,

$$UKU^{-1}(\mathbf{p}) \sim \left( \begin{array}{c|c} \frac{(\mathbf{p} - \mathbf{p}_\alpha)^i \Gamma_i^\alpha}{0} & 0 \\ \hline 0 & K_\alpha(\mathbf{p} - \mathbf{p}_\alpha) \end{array} \right), \tag{5}$$

where  $K_x(0)$  has only nonzero eigenvalues, and where  $\Gamma_i^\alpha$  is an  $N_x$ -dimensional representation of the Clifford algebra on  $d$  generators:

$$\Gamma_i^\alpha \Gamma_j^\alpha + \Gamma_j^\alpha \Gamma_i^\alpha = 2\delta_{ij}. \tag{6}$$

In the physically interesting case  $d = 3$ , these representations of the Clifford algebra would be direct sums of the irreducible representations by Dirac matrices:  $\frac{1}{2}(1 \pm \gamma_5)\gamma_0\gamma_i$ .

The excitations corresponding to the eigenvectors of  $K_x(\mathbf{p})$  are not seen in the low energy limit. By (4), the representation  $\Gamma_i^\alpha$  of the Clifford algebra splits up into the direct sum of representations  $\Gamma_i^{Q,\alpha}$  corresponding to the possible charges  $Q$ .

To describe the handedness of these excitations, define,

$$\Gamma_5^{Q,\alpha} = \frac{(-i)^{(d-1)/2}}{d!} \varepsilon^{i_1 \dots i_d} \Gamma_{i_1}^{Q,\alpha} \dots \Gamma_{i_d}^{Q,\alpha}, \tag{7}$$

which satisfies

$$(\Gamma_5^{Q,\alpha})^2 = 1, \tag{8}$$

$$[\Gamma_5^{Q,\alpha}, \Gamma_i^{Q,\alpha}] = 0. \tag{9}$$

The sub-representation of the  $\Gamma_i^{Q,\alpha}$  on which  $\Gamma_5^{Q,\alpha} = 1$  can be called right-handed; it consists of one irreducible representation for each right-handed massless fermion in the spectrum. The sub-representation with  $\Gamma_5^{Q,\alpha} = -1$  is called left-handed.

Define the chiral index at  $\mathbf{p}_\alpha$  to be

$$I_{Q,\alpha} = 2^{-(d-1)/2} \text{tr}(\Gamma_5^{Q,\alpha}). \tag{10}$$

The coefficient  $2^{(d-1)/2}$  is the dimension of the irreducible representations of the Clifford algebra. Define the total chiral index to be

$$I_Q = \sum_\alpha I_{Q,\alpha}. \tag{11}$$

$I_Q$  is the total number of right-handed massless fermions minus the number of left-handed massless fermions occurring in the low energy, charge  $Q$  spectrum of the model.

The Nielsen–Ninomiya theorem states that the chiral index of a model of free lattice fermions is zero; i.e.,  $I_Q = 0$  for each charge  $Q$ .

### 3. The Proof

We will use a construction on energy-momentum space. Points in this space will be denoted  $p = (p^\mu) = (p^0, \mathbf{p}^i)$ , where  $p^0$  is an arbitrary real number, the energy, and  $\mathbf{p} = (p^i)$  is a point in  $d$ -dimensional momentum space. We will write  $\partial_\mu$  for  $\partial/\partial p^\mu$ . The spectrum of  $K$  consists of the points  $p = (p^0, \mathbf{p})$  for which  $p^0$  is an eigenvalue of  $K(\mathbf{p})$ .

For each charge  $Q$  we will construct a  $(d+1)$ -current  $j_Q^\mu(p)$  on energy-momentum space.  $j_Q^\mu(p)$  will be called the chiral  $(d+1)$ -current;  $j_Q^0(p)$ , the chiral

charge density; and  $\mathbf{j}_Q(p)$ , the chiral  $d$ -current. It will have the following properties:

- (P1)  $\partial_\mu j_Q^\mu(p) = 0$ , i.e. chiral charge is conserved.
- (P2)  $j_Q^\mu(p)$  vanishes away from the spectrum of  $K$ ;
- (P3)  $j_Q^\mu(p)$  depends only on the spectral projection of  $K$  near  $p$ ; and
- (P4)  $j_Q^0(0, \mathbf{p}) = \sum_\alpha \delta(\mathbf{p} - \mathbf{p}_\alpha) I_{Q, \alpha}$ .

The theorem will then follow immediately. By property (P3), the chiral index is the total chiral charge at  $p^0 = 0$ :

$$I_Q = \int d\mathbf{p} j_Q^0(0, \mathbf{p}). \tag{12}$$

By property (P1) the conservation of chiral charge, and by the compactness of momentum space, the total chiral charge is the same at all energies:

$$\int d\mathbf{p} j_Q^0(0, \mathbf{p}) = \int d\mathbf{p} j_Q^0(p^0, \mathbf{p}) \tag{13}$$

for any  $p^0$ . But, because  $K(\mathbf{p})$  is a matrix of finite dimension and because momentum space is compact, the spectrum of  $K$  is bounded in the  $p^0$ -direction. Therefore, by (P2),

$$\int d\mathbf{p} j_Q^0(p^0, \mathbf{p}) = 0 \tag{14}$$

for  $p^0$  large enough. Combining (12–14) gives  $I_Q = 0$ .

To construct  $j_Q^\mu(p)$ , first define a projection valued distribution  $S(p)$  by

$$S(p) = P_Q(\mathbf{p})\theta(K(\mathbf{p}) - p^0). \tag{15}$$

$S(p)$  projects the eigenvectors of  $K(\mathbf{p})$  which have charge  $Q$  and eigenvalue  $\geq p^0$ .  $S(p)$  is a projection because  $[P_Q(\mathbf{p}), K(\mathbf{p})] = 0$ , and the product of two commuting projections is also one. The chiral current is given in terms of  $S(p)$  by:

$$j_Q^\mu(p) = \left[ d \left( \frac{d-1}{2} \right)! (2\pi i)^{(d-1)/2} \right]^{-1} \varepsilon^{\mu\nu_1 \dots \nu_d} \text{tr} (\partial_{\nu_1} S \dots \partial_{\nu_d} S)(p). \tag{16}$$

The definition would appear to be ambiguous because it requires multiplication of the distributions  $\partial_\mu S$ , but in fact the combination of the antisymmetrization and the trace removes all ambiguity.

Property (P1), that  $\partial_\mu j_Q^\mu = 0$ , is obvious from (16). To see that  $j_Q^\mu$  has property (P2), note that when  $S(p)$  is smooth in  $p$ , we have

$$\begin{aligned} \partial_\mu S &= \partial_\mu(S^2) = S(\partial_\mu S) + (\partial_\mu S)S, \\ \text{so } (1-S)\partial_\mu S &= (\partial_\mu S)S \text{ and } (\partial_\mu S)(1-S) = S(\partial_\mu S), \text{ so} \\ \partial_\nu S &= S(\partial_\nu S)(1-S) + (1-S)(\partial_\nu S)S. \end{aligned} \tag{17}$$

When the expression (16) for  $j_Q^\mu$  is expanded using (17), each of the resulting  $2^d$  terms contains at least once the product  $S(1-S)$ , because of the fact that  $d$  is odd and the trace is cyclic. Thus wherever  $S(p)$  is smooth in  $p$ ,  $j_Q^\mu(p)$  vanishes. From the definition (15) of  $S(p)$  it is clear that  $S(p)$  is smooth for  $p$  not in the spectrum of  $K$ . Therefore (P2) holds.

Property (P3) requires some calculation. We calculate  $j_Q^\mu(p)$  at a generic point  $p$  in the spectrum of  $K$ , namely a point  $p = (p^0, \mathbf{p})$  such that for all  $\mathbf{q}$  near  $\mathbf{p}$ ,  $K(\mathbf{q})$

has exactly one, possibly degenerate, eigenvalue  $k(\mathbf{q})$  near  $p^0$ . Let  $P_0(\mathbf{q})$  be the projection on the eigenspace (of charge  $Q$ ) corresponding to that eigenvalue.  $P_0(\mathbf{q})$  is smooth in  $\mathbf{q}$  near  $\mathbf{p}$  because  $K(\mathbf{q})$  is smooth in  $\mathbf{q}$  there.  $S(q)$  is given by

$$S(q) = \theta(k(\mathbf{q}) - q^0)P_0(\mathbf{q}) + P_s(\mathbf{q}), \tag{18}$$

where  $P_s(\mathbf{q})$  is the projection on the eigenspaces corresponding to the eigenvalues greater than  $k(\mathbf{q})$ .  $P_s(\mathbf{q})$  is smooth for  $\mathbf{q}$  near  $\mathbf{p}$ . Now substitute, using (18), in (16). Analogues to Eq. (17) hold also for the smooth projections  $P_0(\mathbf{q})$  and  $P_s(\mathbf{q})$ . After repeated use of them, we arrive at the formula

$$j_Q^\mu(p) = \left[ \left( \frac{d-1}{2} \right)! (2\pi i)^{(d-1)/2} \right]^{-1} \delta(k(\mathbf{p}) - p^0) \varepsilon^{\mu\nu_1 \dots \nu_d} \partial_{\nu_1} (k(\mathbf{p}) - p^0) \cdot \text{tr}(P_0 \partial_{\nu_2} P_0 \dots \partial_{\nu_d} P_0)(\mathbf{p}), \tag{19}$$

valid at any generic point  $p$ . Thus (P3) holds.

To obtain (P4) we must calculate  $j_Q^\mu(p)$  at  $p^0 = 0$ . The spectrum of  $K$  at  $p^0 = 0$  consists of the isolated points  $\mathbf{p}_\alpha$ , so it suffices to calculate  $j_Q^\mu(p)$  for  $p$  near  $(0, \mathbf{p}_\alpha)$ . We would like to use the canonical form (5) for  $K(\mathbf{p})$  near  $\mathbf{p}_\alpha$  in conjunction with formula (19) for  $j_Q^\mu$ . But the form (5) for  $K(\mathbf{p})$  results from performing a  $\mathbf{p}$ -dependent unitary transformation  $U(\mathbf{p})$  near  $\mathbf{p}_\alpha$ . We need to know how to take into account the effect on  $j_Q^\mu$ .

Examine first the effect of an infinitesimal transformation. Its generator is a skew hermitian matrix  $A(p)$ , depending smoothly on  $P$ . The change in  $S(p)$  is

$$\delta S(p) = [A(p), S(p)]. \tag{20}$$

The change in  $j_Q^\mu(p)$  is calculated from (16) to be

$$\delta j_Q^\mu(p) = \partial_\nu f^{\mu\nu}(p), \tag{21}$$

where

$$f^{\mu\nu}(p) = \left[ \left( \frac{d-1}{2} \right)! (2\pi i)^{(d-1)/2} \right]^{-1} \varepsilon^{\mu\nu\nu_2 \dots \nu_d} \text{tr}(A[S, \partial_{\nu_2} S \dots \partial_{\nu_d} S]). \tag{22}$$

Because of (17),  $f^{\mu\nu}(p)$  is, like  $j_Q^\mu(p)$ , supported on the spectrum of  $K$ .

It follows by integration that if  $U(p)$  is a unitary transformation, smooth in  $p$ , which can be connected smoothly to the identity transformation and if  $S(p)$  is replaced by  $USU^{-1}(p)$  then  $j_Q^\mu$  is replaced by  $j_Q^\mu + \partial_\nu f^{\mu\nu}$  for some  $f^{\mu\nu}$  supported on the spectrum of  $K$ . Any  $U(p)$  given only locally in  $\mathbf{p}$ , but for all  $p^0$ , can certainly be connected smoothly to the identity transformation. Therefore if we use expression (5) for  $K(\mathbf{p})$  near  $\mathbf{p}_\alpha$  we do no more than change  $j_Q^\mu(p)$  near  $(0, \mathbf{p}_\alpha)$  by a total divergence. We will determine  $j_Q^\mu(0, \mathbf{p}_\alpha)$  by calculating  $j_Q^\mu(p^0, \mathbf{p})$  for  $(p^0, \mathbf{p}) \sim (0, \mathbf{p}_\alpha)$  and then appealing to conservation of chiral charge. Adding a total divergence to  $j_Q^\mu$  will not affect the result.

We proceed with the calculation, using (9) in (19). Write  $p = (p^0, \mathbf{p}_\alpha + \mathbf{q})$  with  $(p^0, \mathbf{q}) \sim 0$ . From (5),

$$P_0(\mathbf{p}_\alpha + \mathbf{q}) \sim \frac{1}{2}(1 + \hat{q}^i \Gamma_i^{Q,\alpha}) \tag{23}$$

and

$$k(\mathbf{p}_\alpha + \mathbf{q}) \sim |\mathbf{q}|. \tag{24}$$

Substituting in (19) gives

$$j_Q^\mu(p^0, \mathbf{p}_\alpha + \mathbf{q}) \sim I_{Q,\alpha} \delta(p^0 - |\mathbf{q}|) [V(S^{d-1})|\mathbf{q}|^{d-1}]^{-1}(1, \hat{q}), \tag{25}$$

where  $V(S^{d-1})$  is the volume (area) of the unit sphere in Euclidean  $d$ -space. The limit  $p^0 \rightarrow 0$  yields

$$j_Q^0(0, \mathbf{p}) = I_{Q,\alpha} \delta(\mathbf{p} - \mathbf{p}_\alpha) \tag{26}$$

for  $\mathbf{p}$  near  $\mathbf{p}_\alpha$ .

#### 4. Comments

The above proof proceeded by identifying  $I_Q = \int d\mathbf{p} j_Q^0(0, \mathbf{p})$ , an integer associated with the low energy spectrum, with a quantity derived from the high energy spectrum:  $\int d\mathbf{p} j^0(\infty, \mathbf{p})$ . The proofs of [1] and [2] are based on a similar reasoning. We might remark that there is a simpler although less suggestive argument. Define

$$F_Q^{\mu\nu}(p) = \left[ d \left( \frac{d-1}{2} \right)! (2\pi i)^{(d-1)/2} \right]^{-1} \varepsilon^{\mu\nu\nu_2 \dots \nu_d} \text{tr}(S \partial_{\nu_2} S \dots \partial_{\nu_d} S)(p) \tag{27}$$

for  $p$  not in the spectrum of  $K$ . Clearly,

$$j_Q^\mu(p) = \partial_\nu F_Q^{\mu\nu}(p). \tag{28}$$

In particular,

$$j_Q^0(0, \mathbf{p}) = \partial_j F_Q^{0j}(0, \mathbf{p}), \tag{29}$$

so

$$\begin{aligned} I_Q &= \int d\mathbf{p} j_Q^0(0, \mathbf{p}) \\ &= \int d\mathbf{p} \partial_j F_Q^{0j}(0, \mathbf{p}) \\ &= 0. \end{aligned} \tag{30}$$

We see that to have a nonzero chiral index, we must have nontrivial high energy spectrum and also a failure in the definition (27) of  $F_Q^{\mu\nu}$ . For either of these to hold  $K(\mathbf{p})$  must be an infinite dimensional matrix.

Before continuing this comment, we digress to remark that the relativistic character of the low energy spectrum played a role only in the interpretation of the theorem, allowing the chiral index to be equated with the number of chiral fermions. The integer  $I_Q = \int d\mathbf{p} j_Q^0(0, \mathbf{p})$  is a characteristic of the low energy spectrum whether or not that is relativistic in appearance. This chiral index vanishes for a system of free lattice fermions, but it is nonzero, for example, for a gas of massless chiral fermions in the continuum at nonzero chemical potential  $\mu$ , since this gas is described by  $K(\mathbf{p}) = \mathbf{p}^i \sigma_i + \mu$ , which has the spectrum of the free relativistic Hamiltonian shifted by  $\mu$  in the  $p^0$ -direction.

Return now to the possibility of a nonzero chiral index for an infinite

dimensional  $K(\mathbf{p})$ . We will consider a trivial example which, although of no physical interest, does give a first indication of what might occur. The example is simply the continuum right-handed fermion made into a lattice object by blocking degrees of freedom. In momentum space this amounts to folding the continuum momentum space  $\mathbb{R}^d$  over the torus  $T^d$ . The field has an infinite number of components  $\psi^{\mathbf{a}}(\mathbf{p})$ , where the index  $\mathbf{a}$  is a  $d$ -plet of integers  $\mathbf{a} = (a^1, \dots, a^d)$ , and is subject to the folding boundary conditions

$$\psi^{\mathbf{a}}(\mathbf{p} + 2\pi\mathbf{b}) = \psi^{\mathbf{a}+\mathbf{b}}(\mathbf{p}). \tag{31}$$

The matrix  $K(\mathbf{p})$  is given, for momentum components  $p^i$  in the range  $-\pi \leq p^i \leq \pi$ , by

$$K_{\mathbf{b}}^{\mathbf{a}}(\mathbf{p}) = \delta(\mathbf{a}, \mathbf{b})(\mathbf{p} + 2\pi\mathbf{a})^i \sigma_i, \tag{32}$$

where  $\sigma_i$  is the irreducible right-handed representation of the Clifford algebra.

The only zero eigenvalue of  $K(\mathbf{p})$  occurs at  $\mathbf{p} = 0$  and the spectrum near there is exactly that of the continuum chiral fermion, so  $I_Q = 1$ . More comprehensively, (25) gives

$$j^0(p) = \sum_{\mathbf{a}} \delta(p^0 - |\mathbf{p} + 2\pi\mathbf{a}|) [V(S^{d-1})|\mathbf{p} + 2\pi\mathbf{a}|^{d-1}]^{-1}, \tag{33}$$

so that  $j^0(0, \mathbf{p}) = \delta(\mathbf{p})$  as in the continuum.  $F^{0i}(0, \mathbf{p})$ , defined by (27), is infinite, so Eq. (29), exhibiting  $j^0(0, \mathbf{p})$  as a divergence, does not obtain.

To examine  $j^0(p^0, \mathbf{p})$  at high energy  $p^0$  it is convenient to Fourier transform:

$$\tilde{j}(p^0, \mathbf{x}) = \int_{T^d} d\mathbf{p} e^{i\mathbf{x}\cdot\mathbf{p}} j^0(p^0, \mathbf{p}) \tag{34}$$

$$= \int_{\mathbb{R}^d} d\mathbf{p} e^{i\mathbf{x}\cdot\mathbf{p}} \frac{\delta(p^0 - |\mathbf{p}|)}{V(S^{d-1})|\mathbf{p}|^{d-1}}. \tag{35}$$

From (35) it is clear that, at least for  $d > 1$ ,

$$\lim_{p^0 \rightarrow \infty} j^0(p^0, \mathbf{x}) = \delta(\mathbf{x}). \tag{36}$$

Thus

$$\lim_{p^0 \rightarrow \infty} j^0(p^0, \mathbf{p}) = 1, \tag{37}$$

where the limit of distributions is meant. The equation between the low and high energy expressions for the chiral index reads

$$\int d\mathbf{p} \delta(\mathbf{p}) = \int d\mathbf{p} 1. \tag{38}$$

The spectrum of  $K$  at  $p^0 > 0$  is an immersion of the Euclidean  $(d - 1)$ -sphere of radius  $p^0$  by folding into the torus as described by (31–33). The chiral charge density  $j^0(p^0, \mathbf{p})$  is the unit measure concentrated on this immersed sphere. As  $p^0 \rightarrow \infty$  the immersed sphere becomes dense in momentum space in such a way that the chiral charge density becomes constant. Both  $j^0(0, \mathbf{p})$  and  $j^0(\infty, \mathbf{p})$  are simple in form, even though  $j^0(p^0, \mathbf{p})$  for  $0 < p^0 < \infty$  can be exceedingly complicated.

The mathematical content of the theorem can be regarded as a standard piece



of the theory of characteristic classes. Consider the direct argument of (27–30). The projection  $S(0, \mathbf{p})$  is smooth in  $\mathbf{p}$  except at the fermi surface (i.e., the  $\mathbf{p}_\alpha$  where  $K(\mathbf{p})$  has zero as an eigenvalue). Remove from momentum space a small open neighborhood of the fermi surface (i.e., a small open ball around each  $\mathbf{p}_\alpha$ ). Left behind is a  $d$ -manifold  $M$  with boundary. For the relativistic fermi surface, in which the  $\mathbf{p}_\alpha$  are isolated, the boundary is  $\partial M = \sum S_\alpha$ , where the  $S_\alpha$  are small  $(d-1)$ -spheres enclosing the  $\mathbf{p}_\alpha$ . The vector spaces range  $(S(0, \mathbf{p}))$  form a vector bundle  $W$  over  $M$ . The natural covariant derivative in  $W$  is

$$D_i = \partial_i - \partial_i S(0, \mathbf{p}). \quad (39)$$

Its curvature is

$$\begin{aligned} R_{ij}(\mathbf{p}) &= [D_i D_j] \\ &= [\partial_i S, \partial_j S](0, \mathbf{p}). \end{aligned} \quad (40)$$

We can now recognize  $F^{0i}$ , given by (27), as the dual of the highest component of the Chern character of  $W$ . [5] Since

$$I_Q = \int_{\partial M} \hat{n}_i F^{0i}, \quad (41)$$

$\hat{n}_i$  being the unit normal vector to  $\partial M$ ,  $I_Q$  should be called the Chern number of the bundle  $W$  restricted to the boundary of  $M$ . The theorem  $I_Q = 0$  simply states the fact that the Chern number is a cobordism invariant: i.e., that the Chern number of a boundary is necessarily zero.

Once we have the interpretation of  $I_Q$  as the characteristic number of a vector bundle we can drop the requirement of smoothness for  $K(\mathbf{p})$  and simply demand continuity. We then also know that still another proof of the theorem will exist, in which  $I_Q$  is expressed as the index of an elliptic differential operator on  $\partial M$ . This fourth, analytic proof is the proof of the cobordism invariance of the elliptic index [6]. There is also an evident generalization of the theorem to a setting in which the fermion field of the model itself lies in a nontrivial bundle  $V$  over momentum space. The bundle  $W$  is defined as a subbundle of  $V$ , as before, but its covariant derivative is

$$D_i = D_i^0 - (D_i^0 S), \quad (42)$$

where  $D_i^0$  is some arbitrarily chosen covariant derivative in  $V$ . And of course the theorem works whenever the momentum space is a compact manifold of odd dimension. Neither of these generalization seem of immediate physical interest.

The example of the folded fermion gives a bundle  $W$  of infinite rank. Its Chern number  $I_Q$  is well defined but not the density  $\hat{n}_i F_Q^{0i}$  which it is the integral of in (48). So there is at least a certain category of infinite dimensional bundles for which the Chern number, or more generally the elliptic index, need not be a cobordism invariant. The chiral index measures the amount by which cobordism invariance fails.

We conclude with a brief indication of how the constructions used in this proof

might be applicable to the general lattice model. Define first a Hilbert space by means of the following inner product on the fermionic operators of the model:

$$(B, A) = \langle \Omega, [B^*, A]_+ \Omega \rangle, \quad (43)$$

where  $\Omega$  is the ground state. The resulting Hilbert space is the fermionic subspace of the physical Hilbert space plus its complex conjugate. The usefulness of this expanded Hilbert space is that the inner product (43) depends smoothly on the momentum. Define an operator  $K$  on this Hilbert space by

$$(B, KA) = \langle \Omega, [B^*, [H, A]]_+ \Omega \rangle, \quad (44)$$

where  $H$  is the Hamiltonian. The spectrum of  $K$  consists of the fermionic part of the spectrum of  $H$  along with its reflection through the origin. Thus  $\theta(K)$  projects on the physical vectors.  $K(\mathbf{p})$  will describe the action of  $K$  on vectors of momentum  $\mathbf{p}$ . The use of the anti-commutator in (44) means that the locality properties of  $H$  express themselves as smoothness properties of  $K(\mathbf{p})$ . The goal from this point is to construct the chiral  $(d+1)$ -current  $j_Q^\mu(\mathbf{p})$  as in (15–16) and to show that it has the properties (P1–4). The key steps are: (1) to implement the derivative  $\partial_i$  with respect to  $\mathbf{p}^i$ , and (2) to make sense of the trace in the definition (16) of  $j_Q^\mu$ . It would then be possible to make the chiral index of a model visible in the high energy spectrum. This might lead to an improved understanding of the role of anomalies in lattice models. It might also make possible calculation of the chiral index, and so the number of chiral fermions, without a solution of the dynamics.

Work in these directions will be reported subsequently.

*Acknowledgements.* I acknowledge early conversations with H. B. Nielsen and with D. Foerster. I thank S. Fursteberg and H. Bass for providing reference 4.

After completion of this paper, I was informed by I. M. Singer that what here is called the chiral index was treated by different techniques in [7] under the name spectral flow.

## References

1. Nielsen, H. B., Ninomiya, M.: Nucl. Phys. **B185**, 20 (1981)
2. Nielsen, H. B., Ninomiya, M.: Niels Bohr Institute preprint, January 1981
3. Kogut, J., Susskind, L.: Phys. Rev. **D11**, 395 (1975)
4. Swann, R. G.: Trans. Am. Math. Soc. **237**, 111 (1978)
5. See for example Kobayashi, S., Nomizu, K.: Foundations of Differential Geometry Vol. II, New York: Interscience 1969
6. See for example Palais, R. S.: (ed.), Seminar on the Atiyah–Singer Index Theorem, Annals of Mathematics Studies N° 57, Princeton: Princeton University Press 1965
7. Atiyah, M. F., Patodi, W. K., Singer, I. M.: Math. Proc. Camb. Philos. Soc. **79**, 71 (1976)

Communicated by R. Stora

Received August 21, 1981