WINDING NUMBER, FAMILY INDEX THEOREM, AND ELECTRON HOPPING IN A MAGNETIC FIELD

X.G. WEN and A. ZEE

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

In this paper we study the general electron hopping hamiltonian in a magnetic field $H = \sum t_{ij} c_i^\dagger c_j u_{ij}$ using topological methods. We find that the energy spectrum $E(k_x, k_y)$ of the hamiltonian always has at least p isolated zeros if the magnetic flux through each plaquette is $\Phi = 2\pi q/p$ with p even. In general each zero corresponds to a $(1 + 2)$ dimensional Dirac fermion in the continuum limit. Thus H has at least p families of Dirac fermions in the continuum limit. For the nearest neighbor hopping hamiltonian we are able to show (in a certain gauge) that the zeroes appear at $ak = (\pi/2 + 2\pi n_1/p, \pi/2 + 2\pi n_2/p)$. We also discussed the relation of our problem to Berry's phase and supersymmetry.

1. Introduction

In this paper, we study the electron hopping hamiltonian in a uniform magnetic field

$$H = t \sum_{\langle ij \rangle} c_i^\dagger c_j u_{ij} \tag{1.1}$$

on a two-dimensional square lattice, where u_{ij} represent complex numbers with unit modulus and $\langle ij \rangle$ are nearest neighbors. The product of the four u_{ij}'s around a plaquette $e^{i\Phi} = u_{12} u_{23} u_{34} u_{41}$ gives the flux Φ through the plaquette, where the sites $1, 2, 3, 4$ lie on the corners of the plaquette. We show that when the flux is equal to $2\pi q/p$ with p even, the energy dispersion relation of the electron $E(k_x, k_y)$ has (some number of) isolated zeroes around which the electron behaves like Dirac fermions. To show the existence of these zeroes, we find it useful to introduce some topological methods. Mathematically this problem is closely related to the index theorem and the family index theorem of Dirac operators [1]. In that case, one has a family of Dirac operators $D(a)$, where the parameter a spans a manifold M. For a general point in M, $D(a)$ has no zeroes. But sometimes, the Dirac operators as a

family have some topological structures which enable one to prove that the Dirac operators must have zeroes at certain points in M. A standard technique to determine whether a region $P \subset M$ contains zeroes or not is to construct a winding number on the boundary of P. The winding number is defined in such a way that a nonzero winding number on the boundary P implies the existence of the zeroes of $D(a)$ inside P. In this paper we will follow a similar approach. We will introduce some winding numbers relevant to the zeroes of H and show the existence of the zero energy states by showing that the relevant winding numbers are nonzero.

The hamiltonian (1.1) and its relatives has been studied by many people. Recently, the hamiltonian has attracted considerable attention because its relevance to the quantum Hall effect [2] and the mean field theory of the RVB (resonating valence bond) model [3]. Let us discuss the mean field theory of RVB theory in more detail. One may use fermion operator c_α, $\alpha = +, -$ to rewrite the Heisenberg model

$$H = \sum J_{ij} S_i \cdot S_j \qquad (1.2)$$

as

$$H = \sum J_{ij} c_i^\dagger \sigma c_i \cdot c_j^\dagger \sigma c_j \qquad (1.3)$$

with a constraint $\sum_\alpha c_\alpha^\dagger c_\alpha = 1$. In a mean field approach to the Heisenberg model one usually choses the order parameter to be

$$\langle S_i \rangle = \langle c_i^\dagger \sigma c_i \rangle = m_i$$

and finds that the mean field ground state is an anti-ferromagnetic state

$$m_i = (-)^i m$$

if $J_{ij} > 0$. RVB theory of the Heisenberg model intends to describe a spin liquid state without any long range order. To develop the mean field theory of this RVB state, one rewrites eq. (1.3) as

$$H = -\tfrac{1}{2} \sum_{\alpha, \beta} J_{ij} \left(c_{i\alpha}^\dagger c_{j\alpha} \right) \left(c_{j\beta}^\dagger c_{i\beta} \right) + \text{const}$$

and chooses

$$\sum_\alpha \langle c_{i\alpha}^\dagger c_{j\alpha} \rangle = -\eta u_{ij}$$

as the RBV mean field order parameter. The mean field effective hamiltonian reads

$$H_{\text{eff}} = \sum_\alpha \eta J_{ij} u_{ij} c_{j\alpha}^\dagger c_{i\alpha}.$$

If we only have the nearest neighbor coupling the above effective hamiltonian is equivalent to the hopping hamiltonian in eq. (1.1) with the spin index suppressed. The phases of $u_{ij} = \exp i\theta_{ij}$ are dynamical variables and correspond to an effective lattice U(1) gauge theory. In a more careful treatment of the Heisenberg model, one finds the RVB mean field theory actually contains a SU(2) gauge structure [3]. In the presence of an external magnetic field, the electron orbital coupling to the magnetic field can be absorbed in u_{ij}. Therefore the net effect of the magnetic field is only to shift the relative chemical potentials of spin-up electrons and spin-down electrons. As we will see, by introducing a "spinor" formalism we can describe the physics of our problem by an equation of the form

$$i\frac{\partial}{\partial t}\Psi(k,t) = \mathcal{H}\Psi(k,t), \qquad (1.4)$$

where $\mathcal{H}$ is a $p \times p$ matrix hamiltonian. In momentum space, the eigenvalues of $\mathcal{H}$ have isolated zeroes. Around these zeroes, we can expand eq. (1.4) as

$$i\frac{\partial}{\partial t}\Psi \simeq (q_x\alpha_x + q_y\alpha_y)\Psi,$$

where (q_x, q_y) represents some suitable momentum coordinates around the zero. From general arguments, we can show that the matrices α_x and α_y are Dirac matrices.

This paper is arranged as follows. In sect. 2, we Fourier transform the hamiltonian and reduce the problem to a $p \times p$ matrix problem. In sect. 3 we evaluate the energy eigenvalues explicitly for $p = 2$, 3 and 4. We find there are 2 and 4 families of Dirac fermions in the continuum limit for $p = 2$ and 4, respectively. In sect. 4 we summarize some general properties of the hopping hamiltonian. In sect. 5, we introduce some winding numbers and discuss some of their properties, especially their topological invariance and their relations to the zero modes. We argue that the zeroes of the hamiltonian with winding number ± 1, in general, correspond to Dirac fermions in the continuum limit. In sect. 6, we show that for even p, the hopping hamiltonian eq. (1.1) has at least p zeroes. Actually this is even true for more general hamiltonians

$$H = \sum t_{ij} c_i^\dagger c_j u_{ij}, \qquad (1.5)$$

where i, j may not be the nearest neighbor. We also introduce the concept of stable values of winding number; and demonstrate that for a generic hamiltonian the winding numbers only take their stable values. The winding numbers have zero probability of taking other unstable values. In our case, we show that for even p the hamiltonian always has a zero at $ak = (\pi/2 + 2\pi n_1/p, \pi/2 + 2\pi n_2/p)$ with the stable value of the winding number equal to ± 1. Therefore we conclude that

the hopping hamiltonian in general contains at least p families of Dirac fermions if p is even. In sect. 7 we evaluate the winding number explicitly for the following hamiltonian

$$H = \sum_i t_x c^\dagger_{i+\hat{x}} c_i u_{i+\hat{x},i} + t_y c^\dagger_{i+\hat{y}} c_i u_{i+\hat{y},i} + \text{h.c.}$$

where $t_y = t$, $t_x = \tau t$ and $\tau \ll 1$. When $\tau = 1$ the above hamiltonian reduces to our previous hopping hamiltonian. In sect. 8 we discuss the relation between the winding numbers and Berry's phase. In sect. 9 we collect some of our observations and conjectures.

2. The matrix hamiltonian

We work in a gauge in which $u_{i+\hat{y},i} = 1$. Here $\hat{x}$ and $\hat{y}$ are unit vectors in the x and y directions. To maintain translation invariance, we are then forced to have $u_{i+\hat{x},i} = (\alpha^{i_y})$ (and thus $u_{i-\hat{x},i} = (\alpha^{i_y})^*$ as required by the hermiticity of H) so that the flux through each plaquette is given by $e^{i\Phi} = \alpha$. The integer i_y is defined by writing the site vector as $i \equiv (i_x, i_y)a$. For u_{ij} to be periodic, we require $\alpha^p = 1$ for some integer p and thus we can have $\alpha = e^{-i2\pi q/p}$ with q and p incommensurate integers.

Transforming to momentum space we find that

$$H = t \sum_k \left(c^\dagger_{k+w} c_k e^{ik_x a} + c^\dagger_{k-w} c_k e^{-ik_x a} + 2 c^\dagger_k c_k \cos k_y a \right), \qquad (2.1)$$

where $w \equiv (0, 2\pi q/pa)$. The appearance of w reflects the periodic structure in coordinate space. Here k ranges over the Brillouin zone $-\pi/a \leqslant k_x$, $k_y \leqslant \pi/a$.

We can introduce a p-component spinor

$$\Psi_k = \begin{pmatrix} c_{k+w} \\ c_{k+2w} \\ c_{k+3w} \\ \vdots \\ c_{k+pw} \end{pmatrix}. \qquad (2.2)$$

The terminology "spinor" is of course purely formal and notational. In terms of the spinor Ψ we have the hamiltonian

$$H = t \sum_k{}' \Psi^\dagger_k \mathcal{H}_k \Psi_k, \qquad (2.3)$$
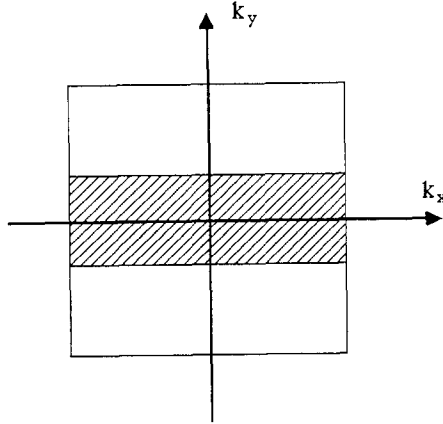
Fig. 1. The square is the first Brillouin zone, $-\pi/a \leqslant k_x$, $k_y \leqslant \pi/a$. The shaded area is the reduced zone, $-\pi/a \leqslant k_x \leqslant \pi/a$ and $-\pi/pa \leqslant k_y \leqslant \pi/pa$.

where the $p \times p$ matrix $\mathscr{H}$ is given by

$$
\mathscr{H} = 2 \begin{pmatrix} \cos(k_y + w)a & & & 0 \\ & \cos(k + 2w)_y a & & \\ & & \ddots & \\ 0 & & & \cos(k + pw)_y a \end{pmatrix}
$$

$$
+ \begin{pmatrix} 0 & z^* & & & z \\ z & 0 & z^* & 0 & \\ & z & \ddots & \ddots & \\ & & \ddots & 0 & z^* \\ z^* & 0 & & z & 0 \end{pmatrix}, \tag{2.4}
$$

and Σ'_k is the summation over the reduced zone $-\pi/a < k_x < \pi/a$, $-\pi/pa < k_y < \pi/pa$ (fig. 1). We have $z = e^{ik_x a}$. More specifically we have

$$
\mathscr{H}_{jl} = 2\delta_{jl}^{(p)} \cos(k_y + 2\pi jq/p)a + \delta_{j,l+1}^{(p)} z^* + \delta_{j,l-1}^{(p)} z, \tag{2.5}
$$

with $\delta_{1,p+1}^{(p)} = 1$. In what follows, we will set the lattice spacing a to unity. Our task is to study the properties of the matrix hamiltonian $\mathscr{H}$ as a function of k. The corresponding eigenvalue equation is given by

$$
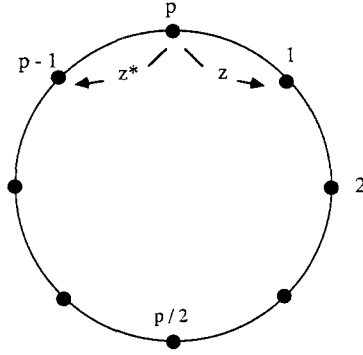E\psi_j = 2\cos(k_y + 2\pi jq/p)\psi_j + z^*\psi_{j-1} + z\psi_{j+1}. \tag{2.6}
$$

Fig. 2. The hamiltonian $\mathcal{H}$ (for fixed k) can be thought of as describing a particle hopping among p sites around a ring.

The problem of an electron moving on a lattice in a magnetic field has been studied extensively of course [4]. Eq. (2.6) is known as Harper's equation and has been studied most notably by Hofstadter [4] who found the range of E for which solutions exists as k_x and k_y range over the Brillouin zone. We expect that some of the results to be given here would be known, but hopefully, our topological approach and emphasis on Dirac modes may be new. We are interested here in the zeroes of $E(k_x, k_y)$. As we will see, around an isolated zero, the theory can be described in terms of Dirac fermions. The hamiltonian $\mathcal{H}$ can be thought of as describing a particle hopping among p sites arranged on a ring. (See fig. 2.) The hopping amplitude is z clockwise and z^* anticlockwise. There is a site energy given by $2\cos(k_y + 2\pi jq/p)$ at the jth site. The problem has an interesting duality property. Were the site energies absent, then the hopping matrix can be readily diagonalized. But this diagonalization would turn the diagonal site energy matrix into a matrix of hopping form.

3. Evaluation of energy eigenvalues

To orient ourselves, we consider the problem for some small values of p. For $p = 2$ $(q = 1)$ we have

$$\mathcal{H} = 2 \begin{pmatrix} -\cos k_y & \cos k_x \\ \cos k_x & \cos k_y \end{pmatrix}. \tag{3.1}$$

(The $p = 2$ case is actually special: in eq. (2.6) z and z^* "collide." However, we can readily show that $\mathcal{H}$ has the form given here.) Thus,

$$E(k_x, k_y) = \pm 2\sqrt{\cos^2 k_x + \cos^2 k_y} \tag{3.2}$$

and there are two isolated zeroes at (k_x, k_y) equal to $(\frac{1}{2}\pi, \frac{1}{2}\pi)$ and $(\frac{1}{2}\pi, -\frac{1}{2}\pi)$. At each of these two nodes, there is a doubly degenerate set of states. Expanding around these isolated points $k_x = \frac{1}{2}\pi + \frac{1}{2}q_x$, $k_y = \pm \frac{1}{2}\pi + \frac{1}{2}q_y$, we find $\mathcal{H} \simeq -(q_x\sigma_1 \pm q_y\sigma_3)$. This has the Dirac form

$$E\gamma_0 - q_x\gamma_x - q_y\gamma_y = 0 \tag{3.3}$$

with the identification $\gamma_0 = \sigma_2$, $\gamma_x = i\sigma_3$, $\gamma_y = \pm i\sigma_1$. The two sets of γ-matrices are related by the transformation $\gamma_\mu \leftrightarrow \gamma_x^\dagger \gamma_\mu \gamma_x$. Thus, the Dirac fermions around the two isolated zeroes are parity reflections of each other. For $p = 3$ and $q = 1$, we find the remarkably simple cubic equation for the eigenvalues of $\mathcal{H}$

$$E^3 - 6E - 2(\cos 3k_x + \cos 3k_y) = 0. \tag{3.4}$$

As k ranges over the Brillouin zone this equation has 3 real roots given by $E_1 = 2\sqrt{2} \cos\theta$, $E_2 = 2\sqrt{2} \cos(\theta + \frac{2}{3}\pi)$, and $E_3 = 2\sqrt{2} \cos(\theta - \frac{2}{3}\pi)$ with $\theta = \frac{1}{3}\arccos[(\cos 3k_x + \cos 3k_y)/2\sqrt{2}]$. Thus, there are three bands. There are four zero-energy lines defined by $k_x = \pm(k_y \pm \frac{1}{3}\pi)$ (with uncorrelated $\pm$ signs). For $p = 4$ and $q = 1$ we can write

$$\mathcal{H} = (\cos k_y - \sin k_y)\sigma_3 \otimes 1 \times (\cos k_y + \sin k_y)\sigma_3 \otimes \sigma_3$$
$$+ \cos k_x(1 + \sigma_1) \otimes \sigma_1 + \sin k_x(1 - \sigma_1) \otimes \sigma_2. \tag{3.5}$$

In general $\mathcal{H}(k)$ with different k will generate an SU(p) algebra. Here we note that if k satisfies $\sin k_x = 0$, $\mathcal{H}(k)$ in fact only generates the subalgebra SO(4) = SU(2) × SU(2) of SU(4). (To see this in a somewhat more symmetrical basis, we can rotate the Pauli matrices to the left of the direct product $\otimes$ by $\sigma_1 \rightarrow \sigma_3$, $\sigma_3 \rightarrow -\sigma_1$, and note that $(\sigma_1 \otimes 1, \sigma_3 \otimes \sigma_1, \sigma_3 \otimes \sigma_1)$ and $(1 \otimes \sigma_1, \sigma_1 \otimes \sigma_3, \sigma_1 \otimes \sigma_3)$ generate two commuting SU(2) algebras.) This implies that the energy has the "factorized" form

$$E(k_x = 0, k_y) = \pm\sqrt{2 + \sin 2k_y} \pm \sqrt{2 - \sin 2k_y} \tag{3.6}$$

($\pm$ signs not correlated). The solutions E_{++}, E_{+-}, E_{-+}, and E_{--}, define four branches. Notice that the branches E_{+-} and E_{-+} $(= -E_{+-})$ vanish and cross over when $\sin 2k_y = 0$. Under the shift $k_y \rightarrow k_y + \frac{1}{2}\pi$, E_{+-} and E_{-+} go over into each other. For general k_x and k_y we have to work out the eigenvalue equation. We find the remarkably simple form

$$E^4 - 8E^2 + 4 - 2(\cos 4k_x + \cos 4k_y) = 0, \tag{3.7}$$

with the solution

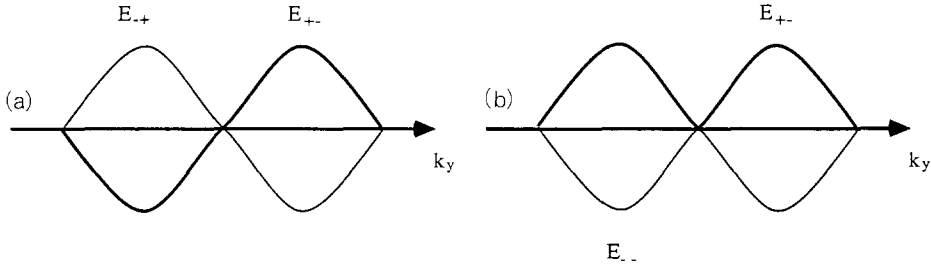$$E = \pm\left\{4 \pm 2\left[3 + \frac{1}{2}(\cos 4k_x + \cos 4k_y)\right]^{1/2}\right\}^{1/2}. \tag{3.8}$$

Fig. 3. The electron energy $E(k_x = 0, k_y)$ as a function of k_y for $p = 4$. (a) The two branches E_{--} and E_{+-} given by eq. (3.6). (b) The two branches E_{+-} and E_{--} given by eq. (3.8).

(Again, the $\pm$ signs are not correlated.) Thus, there are zeroes at $(k_x, k_y) = (0,0)$, $(0, \pm \frac{1}{2}\pi)$, $(\pm \frac{1}{2}\pi, 0)$, and $(\pm \frac{1}{2}\pi, \frac{1}{2}\pi)$. Again, four branches are defined by $E_{++}, E_{+-}, E_{-+}, E_{--}$ according to the four choices of sign in eq. (3.8). (Thus $E_{+-} = + \{4 - 2[3 + \frac{1}{2}(\cos 4k_x + \cos 4k_y)]^{1/2}\}^{1/2}$.) We can verify that for $\sin k_x = 0$ the four branches defined here indeed coincide with the four branches defined in eq. (3.6). However, they are not labeled in the same way. Evidently, as defined here, E_{+-} is always positive while E_{--} is always negative. They never cross zero. But when $\sin k_x = 0$, E_{+-} and E_{--} touches zero (at $k_y = \frac{1}{2}\pi$) with a cusp. (See fig. 3.)

4. General properties

In this section, we summarize some general properties of $\mathcal{H}$, most of which are already known in the literature [5]. Inspection of the eigenvalue equation (2.8) suggests that we define $\psi_j = z^{-j}\psi'_j$. The factor of z apparently disappears from the equation, but it doesn't quite and reappears in the boundary conditions. In eq. (2.8), we had to define $\psi_0 = \psi_p$ and $\psi_{p+1} = \psi_1$, these now become $\psi'_0 = z^{*p}\psi'_p$ and $\psi'_{p+1} = z^p\psi'_1$. In other words, when the eigenvalue equation is written in terms of ψ', the two equations "at the end" depend on z^p. This shows that E is invariant under the shift $k_x \to k_x + 2\pi/p$. By reflection, E is also invariant under the shift $k_y \to k_y + 2\pi/p$. Alternatively, we can write E as a function of e^{ipk_x} and e^{ipk_y} so $E(e^{ipk_x}, e^{ipk_y})$. Thus, we can restrict our attention to the square $0 \leqslant k_x \leqslant 2\pi/p$, $0 \leqslant k_y \leqslant 2\pi/p$. We see that this shift invariance is manifest in the explicit solutions given in sect. 3. Indeed, the energy spectrum depends only on $\cos pk_x + \cos pk_y$ for the cases we have examined ($p = 2$, 3 and 4). It is tempting to think that this may be true in general. Note that one must take care in verifying that this property is satisfied particularly when $\mathcal{H}$ is diagonalized only for special values of k_x or k_y, as is evidenced by the potential confusion in our discussion of the $p = 4$ case. The forms given in eq. (3.6) are evidently not shift invariant. But one can easily check the energy spectrum, i.e. the whole set of eigenvalues is indeed shift invariant. To have shift invariance for each eigenvalue, we must follow the branches with the

cusps (eq. (3.8)). We see that for p even, the cases $p = 4k$ and $p = 4k + 2$ are different. For $p = 4k$, $e^{ip\pi/2} = 1$, thus if E has a zero at $(\frac{1}{2}\pi, \frac{1}{2}\pi)$, then it has a zero at $(0, 0)$, $(0, 2\pi/p)$, $(2\pi/p, 0)$ and $(2\pi/p, 2\pi/p)$. For $p = 4k + 2$, however, if E has a zero at $(\frac{1}{2}\pi, \frac{1}{2}\pi)$, then it has a zero at $(\pi/p, \pi/p)$ but not necessarily at $(0, 0)$.

These shifts can be explicitly formulated. Define the matrix A and B by $(A\psi)_j = e^{-i2\pi q/p}\psi_j$ and $(B\psi)_j = \psi_{j+1}$. Then

$$A\mathcal{H}(k_x, k_y)A^{-1} = \mathcal{H}(k_x + 2\pi q/p, k_y), \tag{4.1}$$

$$B\mathcal{H}(k_x, k_y)B^{-1} = \mathcal{H}(k_x, k_y + 2\pi q/p). \tag{4.2}$$

Furthermore, we find that

$$AB = BA\, e^{i2\pi q/p} \tag{4.3}$$

and that

$$\mathcal{H} = e^{-ik_y}A + e^{+ik_x}B + \text{h.c.}. \tag{4.4}$$

Eqs. (4.3) and (4.4) evidently imply eqs. (4.1) and (4.2). Since q and p are incommensurate, there exists an integer n such that

$$nq/p = 1/p + \text{integer}, \tag{4.5}$$

n and p are incommensurate. Introducing $\tilde{A} = A^n$ and $\tilde{B} = B^n$, one readily shows that

$$\tilde{A}\mathcal{H}(k_x, k_y)\tilde{A}^{-1} = \mathcal{H}(k_x + 2\pi/p, k_y), \tag{4.6}$$

$$\tilde{B}\mathcal{H}(k_x, k_y)\tilde{B}^{-1} = \mathcal{H}(k_x, k_y + 2\pi/p). \tag{4.7}$$

Therefore $\tilde{A}$ and $\tilde{B}$ generate the shift transformation. They also satisfy an algebra

$$\tilde{A}\tilde{B} = \tilde{B}\tilde{A}\, e^{i2\pi q n^2/p}. \tag{4.8}$$

Next, we note that, regardless of whether p is even or odd, under $k_x \to k_x + \pi$ and $k_y \to k_y + \pi$, $\mathcal{H} \to -\mathcal{H}$ and $E \to -E$. For p even, we know in addition that because of shift invariance (k_x, k_y) and $(k_x + \pi, k_y + \pi)$ are the same point in the sense that the spectrum at (k_x, k_y) is the same as the spectrum at $(k_x + \pi, k_y + \pi)$. This implies that for p even the spectrum is such that for each state with energy E there is a state with energy $-E$. This reflection property may be expressed algebraically by saying that for p even, there exists a matrix Γ such that

$$\{\Gamma, \mathcal{H}\} = 0, \qquad \Gamma^2 = 1. \tag{4.9}$$

Indeed, it is easy to verify that explicitly Γ is given by

$$(\Gamma\psi)_j = (-)^j (i)^{p/2} \psi_{j+p/2}. \tag{4.10}$$

Γ plays the role of the γ_5 matrix in field theory discussions of the anomaly and index theorems. We note that the existence of Γ insures that the eigenvalue equation for p even is effectively only of order $p/2$. In fact, we also have

$$\{\Gamma, A\} = 0, \qquad \{\Gamma, B\} = 0; \qquad \{\Gamma, \tilde{A}\} = 0, \qquad \{\Gamma, \tilde{B}\} = 0, \quad (4.11, 4.12)$$

since eq. (4.5) holds for arbitrary k_x and k_y and n is an odd integer. Finally, we mention the properties of $\mathcal{H}$ under reflections of k_x and k_y. From eqs. (2.4) we see that if ψ is a solution of (k_x, k_y), then ψ^* is a solution for $(-k_x, k_y)$ with the same E. Also, ψ_{p-j} is a solution for $(k_x, -k_y)$ with the same E. We also know that for each state with energy E there is a state with energy $-E$. Thus, for p even, we conclude that $E(k_x, k_y)$ is either even or odd under $k_x \to -k_x$ or $k_y \to -k_y$.

We see from eq. (2.1) that under spatial parity $x \to -x$, $\alpha \to \alpha^*$, that is, the flux flips sign. Similarly, under $y \to -y$, we have $\alpha \to \alpha^*$.

5. Topological approaches

We now come to the main results of this paper, namely that topological methods may be applied to determining the zeroes of $\mathcal{H}$. From the explicit examples for $p = 2$ and 4, we suspect that the energy spectrum has zeroes at $k^* = (k_y^*, k_y^*) \equiv (\frac{1}{2}\pi, \frac{1}{2}\pi)$ and at the points connected to it by shift invariance. We have verified by brute force that this is indeed the case for $p = 6$ and 8. However, we would like to demonstrate the existence of zeroes for arbitrary even p. Our approach is based on the existence of a Γ such that $\{\Gamma, \mathcal{H}\} = 0$. Since Γ is hermitian and $\Gamma^2 = 1$, we can go to a basis in which

$$\Gamma = \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix}. \tag{5.1}$$

The matrix blocks here are $p/2$ by $p/2$. $\{\Gamma, \mathcal{H}\} = 0$ implies that $\mathcal{H}$ has the form

$$\mathcal{H}(k) = \begin{pmatrix} 0 & h^\dagger \\ h & 0 \end{pmatrix} \tag{5.2}$$

in this basis. h is not necessarily hermitian. We have $\det H = -\det h^\dagger h$. In a region where $\mathcal{H}$ has no zero mode, $\mathcal{H}^{-1}$ exists and we can define the line integral

$$\nu = \frac{1}{4\pi i} \oint_C \mathrm{Tr}\, \Gamma \mathcal{H}^{-1} \, \mathrm{d}\mathcal{H} \tag{5.3}$$

over a closed loop C in parameter space, i.e. k-space. The 1-form

$$\mathscr{A} \equiv \mathrm{Tr}\; \Gamma\mathscr{H}^{-1}\,\mathrm{d}\mathscr{H} = \mathrm{tr}\big(h^{-1}\,\mathrm{d}h - h^{\dagger-1}\,\mathrm{d}h^{\dagger}\big) \tag{5.4}$$

is closed (but not in general exact), since $\mathrm{d\,tr}(h^{-1}\,\mathrm{d}h) = -\mathrm{tr}(h^{-1}\,\mathrm{d}h)^2 = 0$. Thus, ν is a homotopic invariant: within the region where $\mathscr{A}$ is defined, the loop C may be distorted and ν remains unchanged. It also follows that ν is invariant under variations in $\mathscr{H}$ as long as ν continues to be defined. We can parameterize the change in $\mathscr{H}$ by a parameter τ, for example, we can take the specific form $\mathscr{H}(k, \tau) = \mathscr{H}(k) + \tau\delta\mathscr{H}(k)$ and extend the definition of $\mathscr{A}$ to be in terms of $\mathscr{H}(k, \tau)$. Since $\mathrm{d}\mathscr{A}$ still vanishes with the exterior derivative extended over the parameter τ, ν is invariant under distortion in $\mathscr{H}$. Therefore ν is a topological invariant. To see how ν counts the zero modes of $\mathscr{H}$, let us first show that a nonzero winding number ν along the loop C implies that $\mathscr{H}$ must have zeroes at certain points of the region M enclosed by the loop C. Otherwise, $\mathscr{A}$ would be well defined over M. Shrinking C to zero, we find that ν vanishes. Therefore a nonzero ν implies the existence of zero modes. For simplicity we may assume that $\mathscr{H}$ has zero mode only at a single point k_0 in M. In general, if there is no accidental degeneracy, the zero energy states of $\mathscr{H}(k_0)$ are doubly degenerate. Near k_0, $\mathscr{H}(k)$ takes the following form

$$\mathscr{H}_{\mathrm{eff}} = \begin{pmatrix} 0 & f^*(k) \\ f(k) & 0 \end{pmatrix}$$

in the subspace of nearly zero energies, where $f(k)$ satisfies $f(k_0) = 0$. One finds that the winding number ν is determined by $f(k)$

$$\nu = \frac{1}{2\pi i}\oint \frac{\mathrm{d}f}{f}$$

in this case. We will be mainly interested in the property of f when $\nu = 1$. Assuming $f(k)$ an analytic function, we can write $f(k)$ for $\nu = 1$ as

$$f(k) = \mathbf{n}_1 \cdot (k - k_0) + i\mathbf{n}_2 \cdot (k - k_0) + \mathrm{O}(k - k_0)$$

near k_0, where $\mathbf{n}_i$ are two component real vectors and are linearly independent. Choosing a proper coordinate $q_i = a_{ij}(k - k_0)_j$, we may rewrite

$$f(k(\mathbf{q})) = q_x + iq_y + \mathrm{O}(q^2).$$

One can easily check $f(k)$ indeed gives $\nu = 1$. In terms of $\mathbf{q}$, $\mathscr{H}$ in the nearly zero

energy subspace near k_0 takes the standard Dirac form

$$\mathscr{H}_{\text{eff}} = \begin{pmatrix} & f^*(k) \\ f(k) & \end{pmatrix} \cong q_x \sigma^x + q_y \sigma^y.$$

Therefore a zero of $\mathscr{H}$ with winding number $\nu = 1$ in general corresponds to a $(2+1)$-dimensional Dirac fermion in the continuum limit. (One may object that for $f(k) = r^3 e^{i\phi}$, where (r, ϕ) represents polar coordinates centered at k_0, ν is also 1 and such $f(k)$ does not correspond to a Dirac fermion. However, in general $|f(k)| = a|k - k_0| + b|k - k_0|^3 + \dots$. The term $|k - k_0|^2$ does not appear because $\nu = 1$ and $f(k)$ is analytic. The situation with $a = 0$ is accidental and for real systems has zero probability of happening. For a generic hamiltonian a Dirac fermion will appear for each zero of $\mathscr{H}$ with $\nu = 1$.) Since in the preceding discussion only the phase of $f(k)$ enters, this indicates that it is more direct to consider an alternative winding number defined by

$$u = \frac{1}{2\pi i} \oint_C D^{-1} \, dD \qquad (5.6)$$

where $D \equiv \det h$. We can show the equivalence of the two definitions (5.3) and (5.6) by defining $h = D\tilde{h}$ with $\det \tilde{h} = 1$. Then $\operatorname{tr} \tilde{h}^{-1} \, d\tilde{h} = d \operatorname{tr} \log \tilde{h} = d \log \det \tilde{h} = 0$ and thus $\operatorname{tr} h^{-1} \, dh \propto D^{-1} \, dD$. In fact we have $\mu = \nu$. The discussion above is in a specific basis. In general, we may wish to transform to another basis $\Gamma \to U\Gamma U^{-1}$ and $\mathscr{H} \to U\mathscr{H}U^{-1}$ via a unitary transformation U. In the new basis

$$\mathscr{A} = \operatorname{Tr} \Gamma \mathscr{H}^{-1} D\mathscr{H} \qquad (5.7)$$

is defined in terms of the covariant derivative $D\mathscr{H} = d\mathscr{H} + [U^{-1} dU, \mathscr{H}]$ if U depends on the parameters. In many cases, however, U does not depend on the parameters, for instance, the transformation from the basis given by (4.10) to the one given by (5.1) is independent of k. Incidentally we note that in general we can define

$$\mathscr{A}_m \equiv \operatorname{Tr} \Gamma (\mathscr{H}^{-1} \, d\mathscr{H})^m \qquad (5.8)$$

for any odd m. $\mathscr{A}_m$ is closed since $d \operatorname{tr}(h^{-1} \, dh)^m \propto \operatorname{tr}(h^{-1} \, dh)^{m+1} = 0$ by the cyclicity of the trace. In our specific problem, however, the parameter space (k_x, k_y) is only two dimensional and the $\mathscr{A}_m$'s all vanish identically for $m > 1$.

6. Existence of zero modes

We are now finally ready to show that $\mathscr{H}$ has isolated zero modes as suggested by the explicit examples for $p = 2$ and 4. In fact, using the topological approaches,

we only need to invoke the general properties given in eqs. (4.6)–(4.9) and (4.12) rather than the specific form of $\mathcal{H}$.

From eqs. (4.12) we learned that in the basis where $\Gamma = \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix}$, $\tilde{A}$ and $\tilde{B}$ have the form $\begin{pmatrix} 0 & \tilde{A}_1 \\ \tilde{A}_2 & 0 \end{pmatrix}$ and $\begin{pmatrix} 0 & \tilde{B}_1 \\ \tilde{B}_2 & 0 \end{pmatrix}$, respectively. Eq. (4.8) may then be restated as

$$\tilde{A}_1 \tilde{B}_2 = \tilde{B}_1 \tilde{A}_2 e^{i2\pi q n^2/p}, \tag{6.1}$$

$$\tilde{A}_2 \tilde{B}_1 = \tilde{B}_2 \tilde{A}_1 e^{i2\pi q n^2/p} \tag{6.2}$$

thus implying that

$$\det \tilde{A}_1^{-1} \tilde{A}_2 \det \tilde{B}_1 \tilde{B}_2^{-1} = e^{i\pi q n^2} = -1. \tag{6.3}$$

since both n and q are odd integers. Eqs. (4.6) and (4.7) may be rewritten as

$$\tilde{A}_1 h(k_x, k_y) \tilde{A}_2^{-1} = h^\dagger(k_x + 2\pi/p, k_y), \tag{6.4}$$

$$\tilde{B}_1 h(k_x, k_y) \tilde{B}_2^{-1} = h^\dagger(k_x, k_y + 2\pi/p). \tag{6.5}$$

It is convenient to perform a unitary transformation on all the operators here $\tilde{A} \to e^{i\theta\Gamma} \tilde{A} e^{-i\theta\Gamma}$, $\tilde{B} \to e^{i\theta\Gamma} \tilde{B} e^{-i\theta\Gamma}$, etc. so that $\det \tilde{A}_1^{-1} \tilde{A}_2 = 1$. Then $\det \tilde{B}_1 \tilde{B}_2^{-1} = -1$. We then have

$$D(k_x, k_y) = D^*(k_x + 2\pi/p, k_y), \tag{6.6}$$

$$D(k_x, k_y) = -D^*(k_x, k_y + 2\pi/p). \tag{6.7}$$

Consider then the square with corners on (k_x, k_y), $(k_x + 2\pi/p, k_y)$, $(k_x + 2\pi/p, k_y + 2\pi/p)$ and $(k_x, k_y + 2\pi/p)$. (See fig. 4.) Let us call the path by which we go around the square anticlockwise passing through these four corners C. As we follow the path C, D traces out a closed path in the complex plane. The index μ measures the number of times the complex number D winds around the origin. In fact, eqs. (6.6) and (6.7) imply that the index is not zero. This is most easily seen pictorially. Since we have $D(k_x, k_y) = -D^*(k_x + 2\pi/p, k_y) = -D(k_x + 2\pi/p, k_y + 2\pi/p) = D^*(k_x, k_y + 2\pi/p)$, the path traced out by D has a certain symmetry, namely that opposite sides are related by reflections in the real and imaginary axis, respectively. (See fig. 5.) In fig. 5a we depict a situation in which $\mu = +1$, while in fig. 5b we have $\mu = -1$. In general, as indicated in fig. 5c, the index μ can be any odd integer. In other words, $\mu(\mathrm{mod}\ 2) = 1$. It follows that D has to vanish somewhere inside the loop and $\mathcal{H}$ has zero modes at certain k. The advantage of this topological argument is of course that it shows that the existence of zeroes is
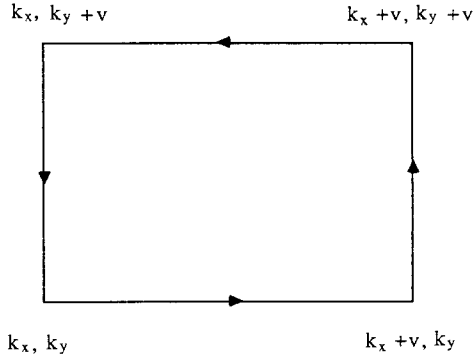
Fig. 4. The contour C in k-space along which the winding number μ (eq. (5.6)) is evaluated (where $v = 2\pi/p$).
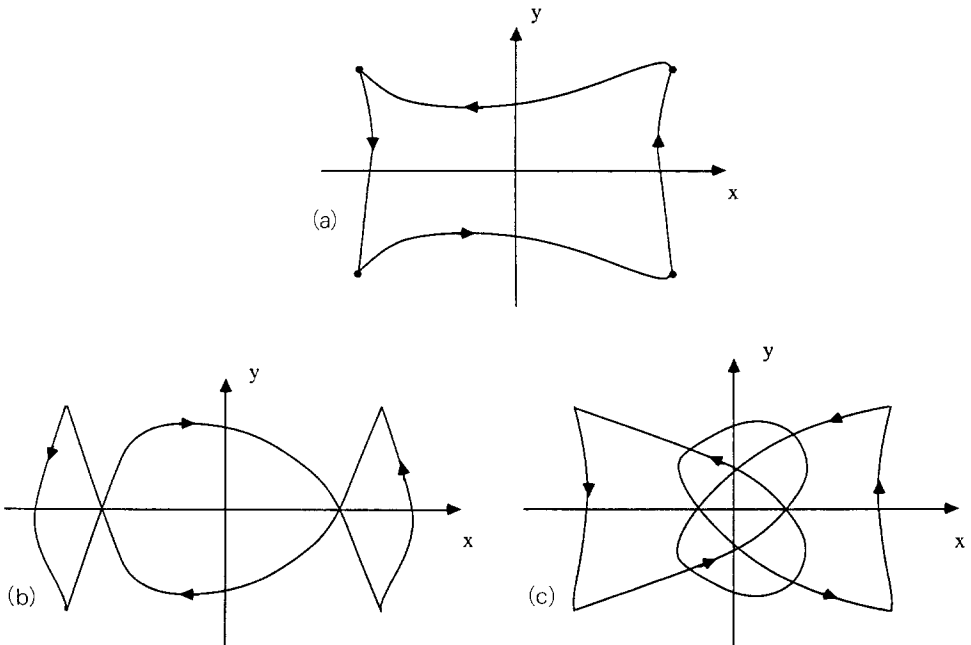

Fig. 5. The closed path traced out by $D(k_x, k_y)$ in complex plane when k goes around the square in fig. 4. (a) The winding number $\mu = 1$, (b) $\mu = -1$, (c) $\mu = 3$.
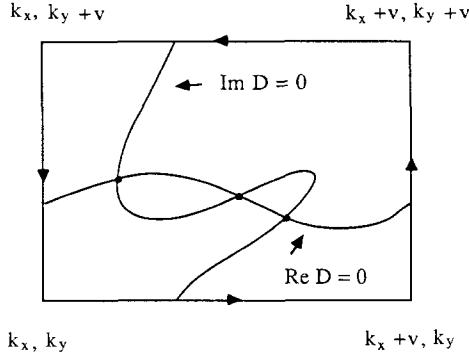
k$_x$, k$_y$ +v   k$_x$ +v, k$_y$ +v

k$_x$, k$_y$ k$_x$ +v, k$_y$

Fig. 6. The lines on which Re D or Im D vanish. The intersection of the two lines gives the locations of zeroes of $\mathcal{H}(k)$.

robust. Thus, we may consider an arbitrary hamiltonian on the square lattice. The zeroes and the resulting Dirac modes will persist as long as the general properties (eqs. (4.6)–(4.9), (4.12)) needed for this discussion hold. An alternative argument can be constructed by looking at the real and imaginary parts of D. Eq. (6.6) tells us that the imaginary part of D reverses sign going from the k_x side of the square to the $k_x + 2\pi/p$ side of the square. Thus, we know that there is, running across the square from the k_y side to the $k_y + 2\pi/p$ side, a curve on which Im D vanishes. Similarly, from eq. (6.7), we conclude that there is, running across the square from the k_x side to the $k_x + 2\pi/p$ side, a curve on which Re D vanishes. The intersections of these two curves give the location of the zeroes of $\mathcal{H}$ at which both Re D and Im D vanish. There must be an odd number of such zeroes. (See fig. 6.)

Since $\mathcal{H}$ has many additional symmetries, we can give another proof of the existence of the zeroes by using those symmetries. (Our previous proof only requires the shift invariance of $\mathcal{H}(p)$.) This proof is closely related to the index theorem [1], and may provide information about the location of some zeroes.

It is convenient to introduce $q = k - (\frac{1}{2}\pi, \frac{1}{2}\pi)$ and rewrite $\mathcal{H}$ in eq. (4.4) as

$$\mathcal{H}(k) = \tilde{\mathcal{H}}(q) = -i\,e^{iq_y}A + i\,e^{iq_x}B + \text{h.c.}.\qquad(6.8)$$

Defining an operator P by $(P\psi)_j = \psi_{p-j}$ we find

$$P^2 = 1,\qquad PAP = A^\dagger,\qquad PBP = B^\dagger,\qquad [P, \Gamma] = 0.\qquad(6.9)$$

Therefore

$$P\tilde{\mathcal{H}}(q)P = -\tilde{H}(-q).\qquad(6.10)$$

In particular for $q = 0$, $P\tilde{\mathcal{H}}(0)P = -\tilde{\mathcal{H}}(0)$. Now, zero-energy eigenstates of $\tilde{\mathcal{H}}(0)$

can be chosen to be the eigenstates of P. Assume n_+ (n_-) are the numbers of zero-energy eigenstates of $\tilde{\mathcal{H}}(0)$ with $P = +1$ $(P = -1)$, we may define an index I by

$$I = n_+ - n_- = \mathrm{Tr}_{E=0} P , \tag{6.11}$$

where $\mathrm{Tr}_{E=0}$ is the trace in zero-energy subspace. Due to eq. (6.10), P maps an eigenstate of $\tilde{\mathcal{H}}(0)$ with a nonzero energy E to a different state with energy $-E$. Therefore P has no diagonal terms in the subspace of nonzero energy states, hence $\mathrm{Tr}_{E \neq 0} P = 0$. We obtain

$$I = \mathrm{Tr}_{E=0} P + \mathrm{Tr}_{E \neq 0} P = \mathrm{Tr} P = 2 . \tag{6.12}$$

As long as $\tilde{\mathcal{H}}(0)$ satisfies eq. (6.10), its index is two and it has at least two zero-eigenstates. Thus we show that the hamiltonian must have zeroes at $q = 0$, or equivalently at $k^* = (\frac{1}{2}\pi, \frac{1}{2}\pi)$.

Using the properties of $\mathcal{H}$ under reflections and/or P given at end of sect. 4 and the above paragraph, we may introduce local winding numbers and reach the same results as we obtained before. First we define a local winding number at $k = 0$ by

$$\nu_0 = \frac{1}{4\pi i} \oint_{C_0} \mathrm{Tr}\, \Gamma \mathcal{H}^{-1} \mathrm{d}\mathcal{H} , \tag{6.13}$$

where C_0 is a small circle centered at $k = 0$. To show ν_0 is a topological invariant, we notice that if $\mathcal{H}$ has a zero at $k_1 \neq 0$, it also has a zero at $-k_1$, with the same winding number. This is because $\mathcal{H}$ satisfies

$$P\mathcal{H}(k)P = +\mathcal{H}(-k) \tag{6.14}$$

Therefore when we deform $\mathcal{H}$, zeroes of $\mathcal{H}$ always enter or leave the circle C_0 in pairs, as long as the deformation does not violate eq. (6.14). This can only change ν_0 by an even number. Therefore $\nu_0 \bmod Z_2$, or equivalently, $e^{i\pi\nu_0}$ is topologically invariant. Similarly one may define a winding number at k^* (i.e. at $q = 0$)

$$\nu_1 = \frac{1}{4\pi i} \oint_{C_1} \mathrm{Tr}\, \Gamma \mathcal{H}^{-1} \mathrm{d}\mathcal{H} . \tag{6.15}$$

where C_1 is a small circle centered at k^*. Using

$$P\mathcal{H}(k - k^*)P = -\mathcal{H}(-k + k^*) \tag{6.16}$$

one can easily show $\nu_1 \bmod Z_2$ or $e^{i\pi\nu_1}$ is a topological invariant for the hamiltonians satisfying eq. (6.10). In sect. 7 we will show $e^{i\pi\nu_1} = -1$, $e^{i\pi\nu_0} = 1$ for $p = 4k + 2$, and $e^{i\pi\nu_1} = -1$, $e^{i\pi\nu_0} = -1$ for $p = 4k$. Before ending this section, we would like to

introduce a concept of stable value of the winding number. n is called a stable value of the winding number of $\mathcal{H}$, if all the slightly deformed hamiltonians $\mathcal{H} + \delta\mathcal{H}$ (the deformations may be required to respect some symmetries) have the same value of the winding number. In our case $\nu_1 = \pm 1$, $\nu_0 = 0$ for $p = 4k + 2$, and $\nu_1 = \pm 1$, $\nu_0 = \pm 1$ for $p = 4k$ are stable values of the winding numbers. Let us illustrate this point by considering some examples. Assume $\tilde{\mathcal{H}}(q = 0)$ has two zero-energy states with $\nu_1 = 1$. Near $q = 0$ and in the subspace of the two nearly zero-energy states, we have

$$\tilde{\mathcal{H}}(q) = \begin{pmatrix} 0 & z^\dagger \\ z & 0 \end{pmatrix}, \qquad P = \begin{pmatrix} 1 & \\ & 1 \end{pmatrix}, \qquad \Gamma = \begin{pmatrix} 1 & \\ & -1 \end{pmatrix}, \qquad (6.17)$$

where $z = q_x + iq_y$. We verify that eq. (6.10) is indeed satisfied. Now let us deform $\tilde{\mathcal{H}}(q)$ to $\tilde{\mathcal{H}}(q) + \delta\tilde{\mathcal{H}}(q)$, where

$$\delta\tilde{\mathcal{H}}(q) = \epsilon \begin{pmatrix} f_{11}(q) & f_{12}(q) \\ f_{21}(q) & f_{22}(q) \end{pmatrix}.$$

Due to eq. (6.10) $f_{ij}(q) = -f_{ij}(-q)$. From eq. (4.9) we find $f_{11} = f_{22} = 0$. If $f_{12}(q) = f_{21}^*(q)$ is analytic, we have $f_{12}(q) \simeq aq_x + bq_y + O(q^3)$ near $q = 0$. For small ϵ, it is obvious that the winding number is not changed

$$\nu_1 = \frac{1}{2\pi i} \oint_{C_1} \frac{d(z + \epsilon f_{12})}{z + \epsilon f_{12}} = \frac{1}{2\pi i} \oint_{C_1} \frac{dz}{z} = 1.$$

Therefore $\nu_1 = 1$ is a stable value of the winding number. If $\nu_1 = 3$, the situation is very different. Now $\tilde{\mathcal{H}}$ in the small energy subspace (again assuming the zero-energy states of $\tilde{\mathcal{H}}(q = 0)$ is two fold degenerate) may take the form

$$\tilde{\mathcal{H}}(q) = \begin{pmatrix} 0 & z^{\dagger 3} \\ z^3 & 0 \end{pmatrix}.$$

After an arbitrary small deformation

$$\tilde{\mathcal{H}} \to \tilde{\mathcal{H}} + \epsilon \begin{pmatrix} 0 & z^\dagger \\ z & 0 \end{pmatrix}$$

the winding number becomes

$$\nu_1 = \frac{1}{2\pi i} \oint_{C_1} \frac{d(z^3 + \epsilon z)}{z^3 + \epsilon z} = 1,$$

where C_1 is an infinitesimal loop around $q = 0$. However, the winding number

before the deformation is

$$\nu_1 = \frac{1}{2\pi i} \int_{C_1} \frac{dz^3}{z^3} = 3.$$

This demonstrates that $\nu_1 = 3$ is an unstable value of the winding number. In fact only $\nu_1 = \pm 1$ are the stable values of the winding number. Similarly one can demonstrate that for $p = 4k + 2$, $\nu_0 = 0$ is the only stable value of winding number at $k = 0$.

For a generic hamiltonian, a winding number only takes its stable values. In our case, if $\mathcal{H}(k)$ satisfies eq. (6.14), in general we have $\nu_1 = \pm 1$, $\nu_0 = 0$ for $p = 4k + 2$ or $\nu_1 = \pm 1$, $\nu_0 = \pm 1$ for $p = 4k$. The winding numbers have a zero probability of taking other unstable values. From the argument in sect. 5 we find that the zeroes at $k = (\frac{1}{2}\pi, \frac{1}{2}\pi)$ and its equivalent points under the shift transformation, in general, give rise to a total of p families of Dirac fermions. The zeroes of $\mathcal{H}$ at other points may also give rise to Dirac fermions. The Hamiltonian $\mathcal{H}$ in general contains at least p families of Dirac fermions arising from $k = (\frac{1}{2}\pi, \frac{1}{2}\pi)$ and its shift equivalences.

7. Evaluation of the winding number

One advantage of introducing topological invariant is that we may be able to exploit our freedom to distort the hamiltonian $\mathcal{H}$. As given in eq. (2.4), $\mathcal{H}$ has the form $\mathcal{H} = \mathcal{H}_y + \mathcal{H}_x$ where $\mathcal{H}_y$ and $\mathcal{H}_x$ depends only on k_y and k_x respectively. A "natural" distortion is to consider $\mathcal{H}(k, \tau) = \mathcal{H}_y + \tau \mathcal{H}_x$. (This distortion can be realized physically by allowing for different hopping constants t_x and t_y along the x and y directions.) $\mathcal{H}(k, \tau)$ obviously satisfies eq. (6.14) for arbitrary τ. We can now treat $\tau \mathcal{H}_x$ as a perturbation on $\mathcal{H}_y$ by taking $\tau \to 0$. Notice that we cannot simply set $\tau = 0$: since $\mathcal{H}_y$ does not depend on k_x, it has lines of zeroes.

For $k_y \simeq \frac{1}{2}\pi$, $\mathcal{H}_y$ has two nearly zero eigenvalues, $\cos k_y$ and $\cos(k + \frac{1}{2}pw)_y$ and thus we should do perturbation theory in this nearly-degenerate two-dimensional subspace. (We are considering p even as always.) Referring to fig. 2, we see that these two nearly degenerate states are connected by $p/2$ hops through ($p/2 - 1$) intermediate states with energies not equal to zero, all clockwise or all anticlockwise. Since the amplitude for each hop is τz or τz^* according to whether the hop is clockwise or anticlockwise, the net amplitude linking the two states is proportional to either $\tau^{p/2}(z^{p/2} + z^{*p/2})$ or $\tau^{p/2}(z^{p/2} - z^{*p/2})$. It turns out that the former corresponds to $p = 4k + 2$, the latter to $p = 4k$. Thus, in the subspace we have the effective two-dimensional hamiltonians as $\tau \to 0$

$$\mathcal{H}_{\mathrm{eff}} = \cos k_y \sigma_3 + K_p \cos \tfrac{1}{2} p k_x \sigma_1, \qquad p = 4k + 2$$

or

$$\mathscr{H}_{\text{eff}} = \cos k_y \sigma_3 + K_p \sin \tfrac{1}{2} p k_x \sigma_1, \qquad p = 4k,$$

where K_p are real numbers.

Now we can evaluate the winding number at $k^* = (\tfrac{1}{2}\pi, \tfrac{1}{2}\pi)$ by considering a small loop C_1 around k^*. One finds immediately

$$\nu_1 = \frac{1}{2\pi i} \oint_{C_1} \frac{\mathrm{d}\left(\cos k_y - iK_p \cos \tfrac{1}{2} p k_x\right)}{\cos k_y - iK_p \cos \tfrac{1}{2} p k_x} = 1 \bmod Z_2. \tag{7.1}$$

For k near k $k = 0$, $\mathscr{H}(k, \tau)$ has no zero for small τ if $p = 4k + 2$. If $p = 4k$, $k = (0,0)$ and $k = (\tfrac{1}{2}\pi, \tfrac{1}{2}\pi)$ are equivalent under the shift transformation. Therefore

$$\nu_0 = \begin{cases} 0, & p = 4k + 2 \\ 1 \bmod Z_2, & p = 4k. \end{cases} \tag{7.2}$$

Although (7.1) and (7.2) are derived for small τ, they are valid for arbitrary τ since ν_0 and ν_1 are topological invariant.

8. Berry's phase

It may also be useful to analyze the situation here in terms of Berry's phase. Since $\{\Gamma, \mathscr{H}\} = 0$, when $\mathscr{H}$ has a zero eigenvalue, it is also doubly degenerate. Now Berry has taught us that when a hamiltonian has a double degeneracy, there is a gauge 1-form A such that the phase integral $e^{\oint A}$ around a closed loop encircling that degeneracy may have a nonzero phase. Thus, we can define $A = (\psi, \mathrm{d}\psi)$ where ψ is an eigenfunction of $\mathscr{H}$: $\mathscr{H}(k_x, k_y)\psi(k_x, k_y) = E(k_x, k_y)\psi(k_x, k_y)$. All we have to do is to integrate $\oint A$ around a closed loop in (k_x, k_y) space encircling the point in (k_x, k_y) space where we suspect the zero of energy to be. If we find that $e^{\oint A} = (-1)$, then there is indeed a zero. Otherwise there is no zero or at most an accidental degeneracy at zero energy. Unfortunately, to use this Berry's phase approach, we have to find the eigenfunctions ψ of $\mathscr{H}$, which would amount to solving the problem. (In the context of Berry's original problem, this step corresponds to determining an eigenfunction of the "instantaneous" hamiltonian.) Incidentally, this underlines the difficulty in the standard formalism that to calculate Berry's phase one has to essentially solve the hamiltonian.

9. Conclusions

It may be useful to record the explicit forms of the reduced "hamiltonian" h. For $p = 2$ we have

$$h = 2\left(\cos k_x + i \cos k_y\right), \tag{9.1}$$

for $p = 4$

$$h = 2\begin{pmatrix} \sin k_y & \cos k_x \\ -i \sin k_x & -\cos k_y \end{pmatrix}, \tag{9.2}$$

and for $p = 6$

$$h = \begin{pmatrix} ia_1 & z^* & z \\ z & ia_2 & z^* \\ z^* & z & ia_3 \end{pmatrix}. \tag{9.3}$$

We have defined $a_j = 2[\cos k_y + (j-1)\frac{1}{3}\pi]$.

As the discussion in sect. 5 suggests, $D = \det h$ should have a relatively simple form. Indeed, we have for $p = 2$

$$D = 2(\cos k_x + i \cos k_y), \tag{9.4}$$

for $p = 4$

$$D = -2(\sin 2k_y + i \sin 2k_x), \tag{9.5}$$

and for $p = 6$

$$D = 2(\cos 3k_x + i \cos 3k_y). \tag{9.6}$$

It is tempting to conjecture that this pattern holds for general p. We can now check the topological argument given in sect. 5 explicitly. Since we know the explicit forms of Γ and $\mathcal{H}$, it is straightforward though somewhat tedious to work out what h is. We find that for $p = 4k + 2$

$$h_{\alpha\beta} = ia_{2\alpha-1}\delta_{\alpha\beta}^{(p/2)} + \delta_{\alpha-k,\beta}^{(p/2)} z + \delta_{\alpha-k-1,\beta}^{(p/2)} z^*. \tag{9.7}$$

Here the indices α, β run from 1 to $p/2 = 2k + 1$. We have defined

$$a_j = 2\cos(k_y + 2\pi j/p). \tag{9.8}$$

Remarkably, the reduced "hamiltonian" h for $p = 4k + 2$ has essentially the same form as the original hamiltonian for $p/2 = 2k + 1$. (This is reminiscent of certain random systems in which a problem in d dimension can be related to a problem in $d/2$ dimension.) For $p = 4k$ we find that

$$h_{\alpha\beta} = -2\cos(k_y + 2\pi\alpha/p)\delta_{\alpha,\beta}^{(p/2)} - (-)^\alpha \delta_{\alpha+1,\beta}^{(p/2)} z - (-)^k(-)^\alpha \delta_{\alpha-1,\beta}^{(p/2)} z^*$$

$$+ \left[1 + (-)^k\right]\delta_{1,\beta}^{(p/2)}\delta_{p/2,\alpha}^{(p/2)} z + \left[1 - (-)^k\right]\delta_{1,\alpha}^{(p/2)}\delta_{p/2,\beta}^{(p/2)} z^*. \tag{9.9}$$

One may check that for $p = 4$, H given by the above formula is equivalent to eq. (9.2).

For $p = 8$

$$h = \begin{pmatrix} -2\cos\left(k_y + \frac{1}{4}\pi\right) & -z^* & 0 & z \\ z & 2\sin k_y & z^* & 0 \\ 0 & -z & 2\sin\left(k_y + \frac{1}{4}\pi\right) & -z^* \\ z^* & 0 & z & 2\cos k_y \end{pmatrix} \tag{9.10}$$

and

$$D = 2\left(\sin 4k_y + i\sin 4k_x\right). \tag{9.11}$$

We also note that the structure of the problem admits a hidden supersymmetry. We can define the supersymmetry generators $Q = \begin{pmatrix} 0 & 0 \\ h & 0 \end{pmatrix}$ and $Q^\dagger = \begin{pmatrix} 0 & h^\dagger \\ 0 & 0 \end{pmatrix}$. Then $Q^2 = Q^{\dagger 2} = 0$. We have $\mathscr{H} = Q + Q^\dagger$ and thus

$$\mathscr{H}^2 = \{Q, Q^\dagger\} = \begin{pmatrix} h^\dagger h & 0 \\ 0 & hh^\dagger \end{pmatrix}$$

is a supersymmetric hamiltonian.

Since $\mathscr{H}^2 Q = Q\mathscr{H}^2 = QQ^\dagger Q$, we have $[Q, \mathscr{H}^2] = 0$. Thus, if $\mathscr{H}^2\psi = E^2\psi$, $Q\psi$ is also an eigenvector with eigenvalue E^2. We note that if

$$\psi = \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix}$$

is a zero mode of $\mathscr{H}^2$, then $h^\dagger h\psi_1 = 0$ which implies $h\psi_1 = 0$ since $\psi_1^\dagger h^\dagger h\psi_1 = 0$. Similarly, $h^\dagger\psi_2 = 0$. Thus, $Q\psi = Q^\dagger\psi = 0$. The vacuum states (lowest energy states) of $\mathscr{H}^2$ respect the supersymmetry if and only if the vacuum energy is zero. Our previous result can be restated as in the family of the supersymmetric hamiltonian $\mathscr{H}^2(\mathbf{k})$ there must exist a Hamiltonian $\mathscr{H}^2(\mathbf{k}_0)$ whose vacuum states are invariant under the supersymmetry (the supersymmetry is not spontaneously broken).

References

[1] T. Eguchi, P.B. Gilkey and A. Hanson, Phys. Rev. 66, 215 (1980);
    P. Shanahan, An introduction to the Atiyah–Singer index theorem, Springer lecture notes in mathematics, No. 638 (Springer-Verlag, 1977);
    L. Alvarez-Gaumé and P. Ginsparg, Nucl. Phys. B243 (1984) 449

[2] D.J. Thouless et al., Phys. Rev. Lett. 49 (1982) 405;
P. Streda, J. Phys. C15 (1982) L717, L1299;
A.H. MacDonald, Phys. Rev. B28 (1983) 6713; B29 (1984) 3057;
I. Dana, Y. Avron and J. Zak, J. Phys. C18 (1985) L679

[3] G. Baskaran, Z. Zou and P.W. Anderson, Sol. St. Comm. 63 (1987) 973;
G. Kotliar, Phys. Rev. B37 (1988) 3664;
I. Affleck and J.B. Marston, Phys. Rev. B37 (1988) 3774;
I. Affleck, Z. Zou, T. Hsu and P.W. Anderson, Phys. Rev. B38 (1988) 745

[4] D.R. Hofstadter, Phys. Rev. B14 (1976) 2239;
G.H. Wannier, Phys. Status Solidi B88 (1978) 757;
M. Ya. Azbel, Zh. Eksp. Teor. Fiz. 46 (1964) 939; [Sov. Phys.-JETP 19, 634 (1964)]

[5] J. Zak, Phys. Rev. A134 (1964) 1602, 1607;
G. Weinreich, Solids: Elementary theory for advanced students (New York, Wiley, 1965) §§8.3, 8.4

[6] M.V. Berry, Proc. Roy. Soc. London Sez. A392 (1984) 45; Nature 326 (1987) 277