PHYSICAL REVIEW B VOLUME 39, NUMBER 16 1 JUNE 1989

Zero modes and the quantized Hall conductance of the two-dimensional lattice in a magnetic field

Mahito Kohmoto*
*Laboratoire de Physique des Solides, Universite Paris-Sud, Bâtiment 510 Centre Universitaire, 91405 Orsay CEDEX, France
and Laboratoire Léon Brillouin, Centre d'Etudes Nucléaires de Saclay, 91191 Gif-sur-Yvette CEDEX, France*
(Received 2 November 1988)

Several properties of the tight-binding electrons in two dimensions with a uniform magnetic field perpendicular to the plane are discussed. The existence of zero-energy states is explicitly shown. The quantized value of the Hall conductance of each magnetic subband which is the Chern number of a certain fiber bundle on a two-torus is calculated. The holelike Hall coefficient can be explained by the quantum mechanics of negatively charged electrons.

I. INTRODUCTION

The tight-binding Hamiltonian on the square lattice is written as

$$H=-t_a \sum_{\langle i,j\rangle} {}_x c_j^\dagger c_i e^{i\theta_{ij}} - t_b \sum_{\langle i,j\rangle} {}_y c_j^\dagger c_i e^{i\theta_{ij}} , \tag{1.1}$$

where c_i is the usual fermion operator on the lattice. The first summation is taken over all the nearest-neighbor sites along the x direction and the second sum along the y direction. The phase factor $\theta_{ij}=-\theta_{ji}$ is defined on a link and represents the magnetic flux through the lattice, i.e., $2\pi\phi=\sum_{\text{plaquette}}\theta_{ij}$ is the magnetic flux through the plaquette in units of the magnetic flux quantum ch/e. We only consider a uniform magnetic field, so ϕ is constant throughout the lattice. The lattice spacing is taken to be unity. The generalization of the present work to a rectangular lattice is straightforward and will not be discussed here.

In traditional solid-state physics, one starts with the energy dispersion *without* a magnetic field,

$$E(k_x,k_y)=-2t_a\cos k_x-2t_b\cos k_y$$

and makes the Peierls-Onsager substitution $\mathbf{k}\rightarrow(\mathbf{p}+e/c\,\mathbf{A})/\hbar$ to have

$$H=-2t_a\cos[(p_x+e/c A_x)/\hbar] -2t_b\cos[(p_y+e/c A_y)/\hbar] . \tag{1.2}$$

Here $\mathbf{p}$ is the quantum-mechanical momentum operator and $\mathbf{A}$ is a vector potential. It can be shown that (1.1) and (1.2) have exactly the same energy spectrum and states if (1.2) is treated carefully.

The problem of two-dimensional electrons in a magnetic field is an old one (see, for example, Wannier,[1] Hofstadter[2] and references therein) and it shows extremely rich and interesting behavior. Also, the Hamiltonian (1.1) has some relevance to the mean-field theory of the resonating-valence-bond theory.[3]

The spectrum is symmetric with respect to $E=0$, and when ϕ is a rational number p/q (p and q are integers which are prime to each other), it consists of q bands. As ϕ is changed continuously, p and q change wildly. The spectrum has an extremely rich structure as shown by Hofstdater.[2] In fact, if ϕ is irrational, it is a Cantor set which consists of infinitely many "bands" with scaling properties.[4,5]

Thouless *et al.*[6] showed that each band carries an integral Hall conductance. This is discussed by Avron *et al.*[7] using the homotopy theory. The integer has a topological origin.[8] It is the Chern number of a fiber bundle which is defined by the wave functions on a two-torus, the reciprocal space of this problem.

II. HARPER EQUATION

A site i on the square lattice has a Cartesian coordinate (n,m) where n and m are integers. Let us choose a gauge in which $\theta_{ij}=0$ for the links along the x direction and $\theta_{ij}=2\pi\phi n$ for the link between $i=(n,m)$ and $j=(n,m+1)$ along the y direction. This gauge gives a uniform magnetic field whose flux through a plaquette is $2\pi\phi$. A rather straightforward calculation transforms (1.1) to

$$H=\frac{1}{(2\pi)^2}\int_{-\pi}^{\pi}dk_x\int_{-\pi}^{\pi}dk_y H(\mathbf{k}) , \tag{2.1}$$

with

$$\begin{aligned} H(\mathbf{k})=&-2t_a\cos k_x c^\dagger(\mathbf{k})c(\mathbf{k}) \\ &-t_b[e^{-ik_y}c^\dagger(k_x+2\pi\phi,k_y)c(k_x,k_y) \\ &+e^{ik_y}c^\dagger(k_x-2\pi\phi,k_y)c(k_x,k_y)] , \end{aligned} \tag{2.2}$$

where $c(\mathbf{k})$ is a fermion operator defined by

$$c_{nm}=\frac{1}{(2\pi)^2}\int_{-\pi}^{\pi}dk_x\int_{-\pi}^{\pi}dk_y \exp[i(k_x n+k_y m)] \times c(k_x,k_y) . \tag{2.3}$$

$c(\mathbf{k})$ is defined in the reciprocal space (the Brillouin zone): $-\pi\le k_x\le\pi$, $-\pi\le k_y\le\pi$. One has to identify $\mathbf{k}+2\pi(j,l)$ as $\mathbf{k}$, and we have a two-torus and its covering space. There is no coupling between different k_y's and k_x couples to $k_x+2\pi\phi$ and $k_x-2\pi\phi$ ($\phi=p/q$). Therefore the Schrödinger equation $H|\Psi\rangle=E|\Psi\rangle$ is reduced to

$$-t_b(e^{-ik_y}\psi_{j-1}+e^{ik_y}\psi_{j+1})-2t_a\cos(k_x^0+2\pi\phi j)\psi_j = E(k_x^0,k_y)\psi_j \ . \quad (2.4)$$

Here k_x is written as $k_x^0+2\pi\phi j$ and

$$\psi_{j+q}=\psi_j \ . \quad (2.5)$$

The state $|\Psi\rangle$ is given by

$$|\Psi\rangle = \sum_{j=1}^{q} \psi_j c^\dagger(k_x^0+2\pi\phi j, k_y)|\,\rangle \ , \quad (2.6)$$

where $|\,\rangle$ represents the vacuum state. When ϕ is a rational number p/q, (2.4) has q eigenvalues for fixed values of k_x^0 and k_y. Therefore, the original band for the tight-binding model is split into q bands due to the application of the magnetic field, and each band has a reduced Brillouin zone

$$-\pi/q \le k_x^0 \le \pi/q, \quad -\pi \le k_y \le \pi \ . \quad (2.7)$$

Equation (2.4) is known as the Harper equation and was derived from (1.2). Here it was derived from (1.1), so the two Hamiltonians have the same properties. The present derivation is certainly not new.

III. DUALITY AND GAUGE TRANSFORMATION

The Harper equation (2.4) has a duality due to Aubry and André,[9] who showed the existence of a transition between localized and extended states of ψ_j when ϕ is an irrational number. In the present case, we take a rational $\phi=p/q$. Let us write

$$\psi_j = \sum_{l=1}^{q} e^{i2\pi\phi j l} f_l \ , \quad (3.1)$$

and substitute it into (2.4). Then we get

$$-t_a(e^{-ik_x^0}f_{l-1}+e^{ik_x^0}f_{l+1})-2t_b\cos(k_y+2\pi\phi l)f_l = E(k_x^0,k_y)f_l \ . \quad (3.2)$$

Note that this is the same as (2.4) if one exchanges t_a and t_b, and k_x^0 and k_y.

This duality can be interpreted as a gauge transformation. To derive (2.4) we used the gauge in which $\theta_{ij}=0$ for the links along the x direction and $\theta_{ij}=2\pi\phi n$ for the link between $i=(n,m)$ and $j=(n,\ m+1)$ along the y direction. If one chooses instead a gauge in which $\theta_{ij}=0$ for the links along the y direction and $\theta_{ij}=2\pi\phi m$ for the link between $i=(n,m)$ and $j=(n+1,m)$ along the x direction, one would obtain

$$-t_a(e^{-ik_x}f_{l-1}+e^{ik_x}f_{l+1})-2t_b\cos(k_y^0+2\pi\phi l)f_l = E(k_x,k_y^0)f_l \ . \quad (3.3)$$

This equation is almost the same as the dual equation (3.2) except one has $-\pi \le k_x \le \pi$ and $-\pi/q \le k_y^0 \le \pi/q$ instead of k_x^0 and k_y. This implies that the dispersion $E(k_x^0,k_y)$ is q-fold degenerate, i.e., $E(k_x^0,k_y) = E(k_x^0,k_y+2\pi n/q)$. This will be shown explicitly in the next section.

IV. ZERO MODES

After a transformation $\psi_{j'}=\psi_j e^{ik_y j}$, the secular equation for (2.4) is written as

$$\det\begin{vmatrix} v_1-E & -t_b & & & -t_b e^{-iqk_y} \\ -t_b & v_2-E & -t_b & & \\ & & \ddots & 0 & \\ & 0 & & & \\ & & -t_b & v_{q-1}-E & -t_b \\ -t_b e^{iqk_y} & & & -t_b & v_q-E \end{vmatrix} = 0 \ , \quad (4.1)$$

where $v_j=-2t_a\cos(k_x^0+2\pi\phi j)$. It is easy to check that the dependence of k_y in the determinant comes only from a term $2t_b^q\cos(qk_y)$. Then the duality tells that the only k_x^0 dependence is a term $2t_a^q\cos(qk_x^0)$. So (4.1) is written as

$$F_{p/q}(E)=2t_a^q\cos(qk_x^0)+2t_b^q\cos(qk_y) \ , \quad (4.2)$$

where $F_{p/q}(E)$ is a qth order polynomial in E and has q real roots since the Hamiltonian is Hermitian. They can be degenerate and we shall show that a degeneracy could take place at $E=0$. If one notices that the argument in $v_j=-2t_a\cos(k_x^0+2\pi\phi j)$ is uniformly distributed on the circle ($\mathbf{R} \bmod 2\pi$), it is shown by inspecting (4.1) that $F_{p/q}(E)$ is an even function if q is even and an odd function if q is odd. So the spectrum is symmetric with respect to $E=0$ and if q is odd, $E=0$ is at the center of the middle band. For even q, it can be shown by inspecting (4.1) that $F_{p/q}(0)=2t_b^q$ when $t_a=0$. Then the duality tells that

$$F_{p/q}(0)=2t_a^q+2t_b^q \ . \quad (4.3)$$

Thus (4.2) has a solution $E=0$ when

$$(k_x^0,k_y)=(n\pi/q, m\pi/q) \ , \quad (4.4)$$

where n and m are integers. Near the zeroes (4.4), $E(k_x^0,k_y)$ has a linear dispersion where the pinnacles of two cones touch at $E=0$. Namely, near $k_x^0=k_y=0$ one has

$$E \approx \pm\gamma q(t_a^q k_x^{02}+t_b^q k_y^2)^{1/2} \ , \quad (4.5)$$

where $-1/\gamma^2$ is the coefficient of E^2 term in $F_{p/q}(E)$.

The fact that the secular equation depends on energy through E^2 only suggests that a unitary transformation makes the Hamiltonian in a form,

$$\begin{bmatrix} 0 & D \\ D^* & 0 \end{bmatrix}^2 ,$$

where D is a $q/2 \times q/2$ matrix.

V. ADIABATIC APPROXIMATION

In order to derive a formula for the Hall conductance, we use the adiabatic approximation (see, for example, Ref. 10). We assume that the system evolves slowly, so up to the first order in time derivative a state is given by

$$|\Psi_\alpha(t)\rangle = \exp\left[\frac{-i}{\hbar}\int_0^t E_\alpha(t')\,dt'\right]\left[|\alpha(t)\rangle + i\hbar\sum_{\beta\neq\alpha}\frac{|\beta(t)\rangle\langle\beta(t)|\partial/\partial t|\alpha(t)\rangle}{E_\beta(t)-E_\alpha(t)}\right] , \tag{5.1}$$

where $|\alpha(t)\rangle$ is a state of the *time-independent* Schrödinger equation with a *parameter t* and $E_\alpha(t)$ is its eigenvalue, i.e.,

$$H(t)|\alpha(t)\rangle = E_\alpha(t)|\alpha(t)\rangle . \tag{5.2}$$

The expectation value of an operator is given to this order as

$$\langle O\rangle_\alpha = \langle\Psi_\alpha(t)|O|\Psi_\alpha(t)\rangle = i\hbar\sum_{\beta\neq\alpha}\frac{\langle\alpha|O|\beta\rangle\langle\beta|\partial/\partial t|\alpha\rangle + \langle\alpha|\partial/\partial t|\beta\rangle\langle\beta|O|\alpha\rangle}{E_\beta(t)-E_\alpha(t)} . \tag{5.3}$$

In the next section we use this formula to calculate the Hall conductance.

VI. HALL CONDUCTANCE

Let us introduce an electric field in addition to the magnetic field and calculate a current. We use a time-dependent vector potential to represent the electric field. Note in the continuous case one has the relation $\mathbf{E} = -\partial\mathbf{A}/\partial t$. We introduce an electric field along the y direction. The phase factor in (1.1) becomes $\theta_{ij} = 2\pi\phi n - E_y t$ for the link between $i=(n,m)$ and $j=(n,m+1)$ along the y direction. Then the Hamiltonian (2.1) becomes

$$H = \frac{1}{(2\pi)^2}\int_{-\pi}^{\pi}dk_x\int_{-\pi}^{\pi}dk_y H(\mathbf{k},t) , \tag{6.1}$$

with

$$H(\mathbf{k},t) = -2t_a\cos k_x[c^\dagger(\mathbf{k})c(\mathbf{k})] - t_b[e^{-ik_y - iE_y t}c^\dagger(k_x+2\pi\phi,k_y+E_y t)c(k_x,k_y+E_y t) + e^{ik_y+iE_y t}c^\dagger(k_x-2\pi\phi,k_y+E_y t)c(k_x,k_y+E_y t)] , \tag{6.2}$$

Observe that the Hamiltonian (6.1) is still time independent due to the integration over $\mathbf{k}$. However, the system is indeed time dependent. This point has some similarity to the topological terms in field theory in the path-integral formulation.

An operator on a link ij which represents a particle flow from i to j is given by

$$v_{ij} = \frac{it_{ij}}{\hbar}(c_j^\dagger c_i e^{i\theta_{ij}} - c_i^\dagger c_j e^{-i\theta_{ij}}) . \tag{6.3}$$

From this, one can show that an electric current operator is written as

$$\mathbf{j} = \frac{e}{L^2\hbar(2\pi)^2}\int_{-\pi}^{\pi}dk_x\int_{-\pi}^{\pi}dk_y\nabla_k H(\mathbf{k},t) . \tag{6.4}$$

Here we consider an $L\times L$ lattice and let L go to infinity in the end. The electric current density along the x direction carried by a state α is

$$\langle j_x\rangle_\alpha = -i\hbar\sum_{\beta\neq\alpha}\frac{\langle\alpha|j_x|\beta\rangle\langle\beta|\partial/\partial t|\alpha\rangle + \langle\alpha|\partial/\partial t|\beta\rangle\langle\beta|j_x|\alpha\rangle}{E_\beta(t)-E_\alpha(t)} . \tag{6.5}$$

Here note the formula

$$\langle\alpha|\partial|\beta\rangle = \frac{\langle\alpha|\partial H(\mathbf{k},t)|\beta\rangle}{E_\beta(t)-E_\alpha(t)} , \tag{6.6}$$

where ∂ is a derivative operator with respect to a parameter in the Hamiltonian, e.g., $\partial/\partial t$ and $\partial/\partial k_y$. Using (6.6), one can obtain several equivalent formulas. First, we obtain

$$\langle j_x \rangle_\alpha = -ie\hbar \sum_{\beta\neq\alpha} \frac{\langle\alpha|j_x|\beta\rangle\langle\beta|\partial H(\mathbf{k},t)/\partial t|\alpha\rangle - \langle\alpha|\partial H(\mathbf{k},t)/\partial t|\beta\rangle\langle\beta|j_x|\alpha\rangle}{[E_\beta(t)-E_\alpha(t)]^2} . \tag{6.7}$$

Since $(\partial H/\partial t)=(e/\hbar)(\partial H/\partial k_y)E_y$ in the present gauge, the Hall conductance defined by

$$\langle j_x \rangle_\alpha = (\sigma_{xy})_\alpha E_y \tag{6.8}$$

is given by

$$(\sigma_{xy})_\alpha = -iL^2\hbar \sum_{\beta\neq\alpha} \frac{\langle\alpha|j_x|\beta\rangle\langle\beta|j_y|\alpha\rangle - \langle\alpha|j_y|\beta\rangle\langle\beta|j_x|\alpha\rangle}{[E_\beta(t)-E_\alpha(t)]^2} . \tag{6.9}$$

Also, one gets

$$(\sigma_{xy})_\alpha = -\frac{ie}{L^2}\left[\left[\frac{\partial}{\partial k_x}\langle\alpha|\right]\left[\frac{\partial}{\partial t}|\alpha\rangle\right] - \left[\frac{\partial}{\partial t}\langle\alpha|\right]\left[\frac{\partial}{\partial k_x}|\alpha\rangle\right]\right] \tag{6.10}$$

and

$$(\sigma_{xy})_\alpha = -\frac{ie^2}{\hbar L^2}\left[\left[\frac{\partial}{\partial k_x}\langle\alpha|\right]\left[\frac{\partial}{\partial k_y}|\alpha\rangle\right] - \left[\frac{\partial}{\partial k_y}\langle\alpha|\right]\left[\frac{\partial}{\partial k_x}|\alpha\rangle\right]\right] . \tag{6.11}$$

This can be written as

$$(\sigma_{xy})_\alpha = \frac{e^2}{h}\frac{2\pi}{iL^2}(\nabla_k \times \langle\alpha|\nabla_k|\alpha\rangle)_z , \tag{6.12}$$

where $(\ )_z$ represents the k_z component of a vector in k space.

Let us now consider the contribution to the Hall conductance from a single band,

$$\begin{aligned} \sigma_{xy} &= \frac{e^2}{h}\frac{1}{2\pi i}\int_0^{2\pi/q} dk_x^0 \int_0^{2\pi} dk_y \left[\left[\frac{\partial}{\partial k_x^0}\langle\alpha|\right]\left[\frac{\partial}{\partial k_y}|\alpha\rangle\right] - \left[\frac{\partial}{\partial k_y}\langle\alpha|\right]\left[\frac{\partial}{\partial k_x^0}|\alpha\rangle\right]\right] \\ &= \frac{e^2}{h}\frac{1}{2\pi i}\int_0^{2\pi/q} dk_x^0 \int_0^{2\pi} dk_y (\nabla_k \times \langle\alpha|\nabla_k|\alpha\rangle)_z . \end{aligned} \tag{6.13}$$

This formula has a subtle topological nature which is essential in the quantization of the Hall conductance. First, one may naively wish to apply the Stokes theorem to (6.13) to obtain

$$\sigma_{xy} = \frac{e^2}{h}\frac{1}{2\pi i}\int d\mathbf{k}\langle\alpha|\nabla|\alpha\rangle ,$$

where $\int d\mathbf{k}$ is a line integral around the reduced Brillouin zone. The integrand $\langle\alpha|\nabla|\alpha\rangle$ in fact represents the derivative of the phase of state α. So it appears that the quantization of σ_{xy} is already shown since the phase of the state must gain 2π times an integer after a revolution around the reduced Brillouin zone in order to have a single-valued wave function. However, this argument is rather incomplete. Assigning a phase to a state is a subtle problem. One example which drew large attention recently is Berry's phase in the adiabatic process.[11] (In the present formulation of the problem, k_y and time t evolve in an equivalent way. So we have a Berry's phase in this problem, but it does not play a role in the quantization.) The essential point here is that the reduced Brillouin zone is topologically a two-torus and in general it is not possible to define a global phase on it. Then the phase of the state defines a principal U(1) bundle over the two-torus and $\langle\alpha|\nabla|\alpha\rangle$ defines a connection. Now the expression in (6.13) represents a Chern number of the fiber bundle which is necessarily an integer. Therefore the Hall conductance due to a single band is an integer in unit of e^2/h, and the integer is the Chern number of the fiber bundle defined by the wave function. A detailed account of this point can be found in Ref. 8.

VII. QUANTIZED VALUES OF HALL CONDUCTANCE

Let us suppose the Fermi energy is in rth gap, namely there are r bands below the Fermi energy. Here we have three positive integers r, p, and q. The Darboux theorem implies that three positive integers always satisfy

$$r = qs_r + pt_r , \tag{7.1}$$

where s_r and t_r are integers, and $|t| \le q/2$ and $0 \le r \le q$. It was announced in Ref. 6 that t_r was the integral value of the Hall conductance. Then rth band, which is between $(r-1)$th and rth gaps, carries an integral Hall conductance I_r which satisfies

$$1 = qJ_r + pI_r , \tag{7.2}$$

where $I_r = t_r - t_{r-1}$ and $J_r = s_r - s_{r-1}$.

This relation was discussed by Dana *et al.*[12] in terms of the magnetic translation group. Here we show the details of the weak-coupling calculation in Sec. VII B. Also, a derivation from the symmetries of the Harper equation is given in Sec. VII C. First, it is perhaps useful to mention the Streda formula.[13]

A. Streda formula

The Streda formula for the Hall conductance is written as

$$\sigma_{xy}=\frac{e^2}{\hbar}\frac{\partial N}{\partial B} , \tag{7.3}$$

where N is the total density of states below the gap. Let us point out that this formula is not meaningful as it stands, unless N is shown to be a differentiable function of B. Nonetheless, this formula can give the correct value with certain manipulations. In the present case, (7.1) is divided by q to give

$$N=\frac{r}{q}=s_r+\frac{B}{2\pi}t_r . \tag{7.4}$$

So (7.3) would give

$$\sigma_{xy}=\frac{e^2}{h}t_r , \tag{7.5}$$

if s_r and t_r are assumed to be independent of B. This assumption can only be made plausible in the weak-coupling limit.

B. Weak-coupling limit

In order to calculate σ_{xy} explicitly, we consider the case of small t_b. When $t_b=0$, (2.4) gives a single band for a one-dimensional tight-binding model

$$E_0(k_x^m)=-2t_a\cos(k_x^m)=-2t_a\cos(k_x^0+2\pi\phi m) , \tag{7.6}$$

and the wave function is

$$\begin{aligned}\psi_j&=1 \quad \text{for } j=m\\ &=0 \quad \text{for } j\neq m .\end{aligned} \tag{7.7}$$

This spectrum is doubly degenerate and the term proportional to t_b in (2.4) gives the coupling between the two branches of the dispersion. The gaps open when $E_0(k_x)=E_0(k_x')$ (namely, $k_x'=-k_x+2\pi s$ for an integer s) and $k_x'=k_x-2\pi(p/q)t$, where t is an integer with $|t|\le q/2$. So k_x and k_x' couple by $|t|$th-order perturbation and the size of the gap is an order of $t_b^{|t|}$. If s is chosen appropriately, k_x is put between 0 and π and is written as $k_x=r/\pi$ with $0\le r\le q$. Thus we obtain (7.1) with an identification of t_r being the order of the rth gap size.

Let us now explicitly obtain the Hall conductance. In a band, except near the gaps, the wave functions are simply

$$\begin{aligned}|\psi_j(k_x^0,k_y)|&=1 \text{ for } j=m \text{ or } m'\\ &=0 \text{ otherwise} .\end{aligned} \tag{7.8}$$

Here $m'=m-t$ and we take $0\le k_x^0<2\pi/q$. The gaps are at $k_x^0=0$ and π/q. Near the edge, all the components of the wave function except ψ_m and $\psi_{m'}$ are negligible. Write $\psi_m=a$ and $\psi_m=b$, then an effective Schrödinger equation is

$$\begin{bmatrix}\varepsilon & \Delta e^{-ik_y t}\\ \Delta e^{ik_y t} & -\varepsilon\end{bmatrix}\begin{bmatrix}a\\ b\end{bmatrix}=E\begin{bmatrix}a\\ b\end{bmatrix} . \tag{7.9}$$

Here ε and E are measured from the middle of the gap, namely,

$$\varepsilon=2t_a\cos\left[k_x^0+2\pi\frac{p}{q}m\right] , \tag{7.10}$$

and $k_x^0=0$ for even r and $k_x^0=\pi/q$ for odd r. The parameter Δ is of the order of $t_b^{|t|}$. Write $b=b'e^{ik_y t}$, then one gets

$$\begin{bmatrix}\varepsilon & \Delta\\ \Delta & -\varepsilon\end{bmatrix}\begin{bmatrix}a\\ b'\end{bmatrix}=E\begin{bmatrix}a\\ b'\end{bmatrix} . \tag{7.11}$$

Here all the quantities are real and the solutions are

$$\begin{aligned}E_+&=(\varepsilon^2+\Delta^2)^{1/2} ,\\ (a,b')&=(\cos\theta,\sin\theta) ,\end{aligned} \tag{7.12}$$

with $\sin2\theta=\Delta/(\varepsilon^2+\Delta^2)^{1/2}$, and

$$\begin{aligned}E_-&=-(\varepsilon^2+\Delta^2)^{1/2} ,\\ (a,b')&=(-\sin\theta,\cos\theta) ,\end{aligned} \tag{7.13}$$

with $\sin2\theta=-\Delta/(\varepsilon^2+\Delta^2)^{1/2}$. For the upper edge of rth band one takes the solution E_-. As k_x^0 passes 0 (for even r) or π/q (for odd r) θ changes from $\pi/2$ to 0. Then (a,b) changes as $(-1,0)\to(0,e^{ik_y t_r})$. For the lower edge, we take E_+ and (a,b) changes as $(0,e^{ik_y t_{r-1}})\to(1,0)$.

Let I_r represent the contribution of rth band to the Hall conductance in unit of e^2/h. From (6.13) we have

$$I_r=\frac{h}{e^2}\sigma_{xy}=\frac{1}{2\pi i}\int_0^{2\pi/q}dk_x^0\int_0^{2\pi}dk_y\sum_{j=1}^{q}\left[\frac{\partial\psi_j^*}{\partial k_x^0}\frac{\partial\psi_j}{\partial k_y}-\frac{\partial\psi_j^*}{\partial k_y}\frac{\partial\psi_j}{\partial k_x^0}\right]=\frac{1}{2\pi i}\int_0^{2\pi/q}dk_x^0\int_0^{2\pi}dk_y\left[\nabla_k\times\sum_{j=1}^{q}\psi_j^*\nabla_k\psi_j\right]_z . \tag{7.14}$$

The energy dispersion does not depend on k_y in the weak-coupling limit and it is doubly degenerate for the two regions $k_x^0=[0,\pi/q]$ and $[\pi/q,2\pi/q]$. The wave function changes character across the boundaries of the two regions which contribute to I_r. Within each region the wave function is regular and we apply the Stokes theorem in the two regions separately. The result is

$$\frac{h}{e^2}\sigma_{xy}=I_r=t_r-t_{r-1} , \tag{7.15}$$

which was announced in the beginning of this section.

C. Derivation from symmetries

In order to understand the symmetry properties of (2.4), it is useful to make a transformation

$$\psi_j = \Psi_j e^{-ik_y j} . \tag{7.16}$$

Then one gets

$$-t_b(\Psi_{j-1}+\Psi_{j+1}) - 2t_a\cos(k_x+2\pi\phi j)\Psi_j = E(k_x,k_y)\Psi_j . \tag{7.17}$$

Note that we do not have explicit dependence on k_y, and k_x^0 is written as k_x. Now one can consider (7.17) as an infinite system. From (2.5) and (7.16) Ψ_j satisfies

$$\Psi_{j+q} = e^{iqk_y}\Psi_j , \tag{7.18}$$

which can be regarded as the Floquet (or Bloch) theorem.

The new equation (7.17) has several symmetries among which we use (a) $k_x \to k_x + 2\pi p/q$ and (b) $k_y \to k_y + 2\pi/q$. The wave function is transformed accordingly as (a) $\Psi_j \to \Psi_{j+1} e^{i\alpha(k_y)}$ and (b) $\Psi_j \to \Psi_j e^{i\beta(k_x)}$. Let us define

$$J = -\frac{q}{2\pi i L}\int_0^{2\pi p/q} dk_x \int_0^{2\pi/q} dk_y \sum_{j=1}^{L}\left[\frac{\partial \Psi_j^*}{\partial k_x}\frac{\partial \Psi_j}{\partial k_y} - \frac{\partial \Psi_j^*}{\partial k_y}\frac{\partial \Psi_j}{\partial k_x}\right] . \tag{7.19}$$

Here the wave function is normalized $\sum_{j=1}^{q}|\Psi_j|^2=1$. The above symmetries and a comparison of (7.14) and (7.19) with (7.16) give

$$J = -\frac{p}{q}I_r + \frac{1}{q} . \tag{7.20}$$

Thus the integral (7.19) is identified as the integer J_r in (7.2).

VIII. SUMMARY AND DISCUSSIONS

In Sec. IV, the existence of zero modes is shown explicitly. The duality plays an important role which is originated from the symmetry of the lattice. It is remarkable that the zero modes persist irrespective of the ratio of the couplings t_a/t_b. The zero modes disappear when the symmetry of the lattice is changed to, for example, a hexagonal lattice[14] and a spin-density-wave system.[15]

In Sec. VI, the formula for the Hall conductance is derived using the adiabatic approximation. The result is the same as the so-called Kubo formula which comes from the linear-response theory in dissipative systems. Note that in the case of the Hall conductance, electrons move perpendicular to the electric field. So one does not need a dissipation to have a current which is proportional to the electric field. In fact, one can eliminate the electric field in a moving frame with a constant velocity. In this frame, we have only a magnetic field and the moving lattice drags the electrons. So the electrons do not move faster than the velocity of the lattice classically. There is a possibility that the Hall current is exactly proportional to the electric field, namely all the higher-order corrections in the adiabatic approximation cancel each other. However, this point is not clear yet.

In Sec. VII, the integral values of the Hall conductance are calculated explicitly in the weak-coupling limit. Also, it is discussed in terms of the symmetries of the Harper equation. The relation (7.1) or (7.2) uniquely determines the Hall conductance except at a zero mode $E=0$ (q even and $r=q/2$). The contribution from a subband is symmetric with respect to $E=0$, and the sum of contributions from all the bands is zero. At a zero mode, we have two solutions $t_r=\pm q/2$. In fact, this point has a topological singularity. Since the two bands are degenerate, we do not have regular manifolds there.

The values of t_r change wildly as r is increased for general values of p and q. Also, the sign changes, so one has many changes between electronlike and holelike conductions as the Fermi energy is raised. This behavior might have some relevance to the Hall measurements of the quasi-two-dimensional organic conductors in which the Hall conductance is approximately quantized and it changes sign as a magnetic field is varied.[16]

In the simplest case of $p=1$, all the bands except ones at the center of the spectrum have $+1$ Hall conductance each. For odd q, the center band carries $-(q-1)$ Hall conductance and for even q the two bands at the center carry $-(q-2)$. Thus, if the Fermi energy is below the center of the spectrum we have an electronlike Hall coefficient, and if it is above the center we have a holelike Hall coefficient. Usually the holelike behavior is explained by assigning positively charged particles, namely holes, as the carrier of Hall current. This picture comes from a semiclassical treatment of the quantum-mechanical system. Here, however, it is explained by the full quantum-mechanical picture. The carrier is a negatively charged electron, and the holelike behavior is due to the wave nature of electrons, i.e., they get diffracted by the lattice.

The above behavior remains for $p\neq 1$ as well. Write $p/q = 1/(Q+p'/q')$, then the spectrum has $Q+1$ clusters. A sum of Hall conductance of each cluster is $+1$ and the middle cluster has $-Q$ integral Hall conductance. Note that in a realistic lattice spacing and a magnetic field, one has $\phi \ll 1$ and Q is a large number.

In three dimensions, if the magnetic field is parallel to the third axis, the coupling t_c in the third direction is decoupled with the first two couplings. The energy dispersion is simply $E = E_{2D} + (-2t_c \cos k_z)$. The Hall conductance is obtained simply by averaging the two-dimensional result over the new contribution to the density of states due to the third coupling. Therefore, the above discussion on the holelike Hall effects is applicable to three dimensions as well.

When the magnetic field is not parallel to the third axis, we have even richer phenomena including the fractional quantization of Hall conductivity.[17]

ACKNOWLEDGMENTS

Some parts of this work are based on discussions with D. J. Thouless. M. K. acknowledges support from the A. P. Sloan Foundation.

*Permanent address: Institute for Solid State Physics, University of Tokyo, Roppongi, Minato-ku, Tokyo 106, Japan.

[1]G. H. Wannier, Int. J. Quantum Chem. **13**, 413 (1979).
[2]D. R. Hofstadter, Phys. Rev. B **14**, 2239 (1976).
[3]G. Baskaran, Z. Zou, and P. W. Anderson, Solid State Commun. **63**, 973 (1987).
[4]M. Kohmoto, Phys. Rev. Lett. **51**, 1198 (1983); C. Tang and M. Kohmoto, Phys. Rev. B **34**, 2041 (1986).
[5]S. Ostlund and R. Pandit, Phys. Rev. B **29**, 1394 (1984).
[6]D. J. Thouless, M. Kohmoto, P. Nightingale, and M. den Nijs, Phys. Rev. Lett. **49**, 405 (1982).
[7]J. Avron, R. Seiler, and B. Simon, Phys. Rev. Lett. **51**, 51 (1983).
[8]M. Kohmoto, Ann. Phys. (N.Y.) **160**, 355 (1985).
[9]S. Aubry and G. André, Ann. Isr. Phys. Soc. **3**, 133 (1980).
[10]L. I. Schiff, *Quantum Mechanics* (McGraw-Hill, New York, 1955).
[11]M. Berry, Proc. R. Soc. London **392**, 45 (1984).
[12]I. Dana, Y. Avron, and J. Zak, J. Phys. C **18**, L679 (1985).
[13]P. Streda, J. Phys. C **15**, L717 (1982).
[14]F. H. Claro and G. H. Wannier, Phys. Rev. B **19**, 6068 (1979).
[15]A. M. Szpilka and M. Kohmoto (unpublished).
[16]B. Piveteau, L. Brossard, F. Creuwet, D. Jérome, R. C. Lacoe, A. Moradppour, and M. Ribault, J. Phys. C **19**, 4483 (1986).
[17]G. Montambaux and M. Kohmoto (unpublished).