Quantized Hall conductance as a topological invariant

Qian Niu, D. J. Thouless,* and Yong-Shi Wu†
Department of Physics FM-15, University of Washington, Seattle, Washington 98195
(Received 21 September 1984)

Whenever the Fermi level lies in a gap (or mobility gap) the bulk Hall conductance can be expressed in a topologically invariant form showing the quantization explicitly. The new formulation generalizes the earlier result by Thouless, Kohmoto, Nightingale, and den Nijs to the situation where many-body interaction and substrate disorder are also present. When applying to the fractional quantized Hall effect, we draw the conclusion that there must be a symmetry breaking in the many-body ground state. The possibility of writing the fractionally quantized Hall conductance as a topological invariant is also discussed.

I. INTRODUCTION

In the experiments on both integral[1] and fractional[2] quantized Hall effect it is found that the appearance of a plateau in the Hall conductance is always accompanied by a dip in the longitudinal conductance. This well-observed fact suggests that the existence of the Fermi gap (the energy gap or mobility gap in which the Fermi energy of the system lies) is a necessary condition for the quantization of the Hall conductors. On the other hand, since the phenomenon is quite independent of the details of the devices used in the experiments, this condition must also be sufficient (of course, at zero temperature and in weak electric field).

By now, in the integral case, this relationship has been quite established by perturbation theory[3,4] or by gauge-invariance argument.[5] The latter is more profound for it only uses global properties of the electron system in the external fields. But the solenoid device typically employed in this theory seem to be artificial to most of the known experiments.

Another nonperturbative approach was proposed by Thouless *et al.*[6] (henceforth referred to as TKNdN), who considered an infinite two-dimensional electron gas in a periodic substrate potential commensurate to the perpendicular magnetic field. The Hall conductance calculated from the Kubo formula was rewritten into an integral which shows quantization explicitly. This expression has the advantage that it is independent of the detailed structure of the periodic potential. Later this integral was recognized as the first Chern class of a U(1) principal fiber bundle on a torus.[7,8] The fibers are the magnetic Bloch waves and the torus corresponds to the magnetic Brillouin zone.

Unfortunately, this theory cannot allow either impurity disorder or many-body interactions, because the use of Bloch waves is quite essential to their derivations. In this paper we generalize TKNdN's idea so that an invariant expression can still be constructed in the general case.

The method we are going to use is quite parallel to the generalized formulation of Niu and Thouless[9] for the quantization of particle transport induced by a potential varying slowly and periodically. We use the same

geometry as used by TKNdN, consequently we share with them the same deficiency of ignoring the edge effect.[10] The many-body wave functions are required to satisfy a particular boundary condition described by two fixed phase parameters. We then prove that the Hall conductance becomes independent of the phase parameters in the thermodynamic limit, so that it can be averaged over all the phases that prescribe different boundary conditions. The averaged quantity which equals the Hall conductance possesses an expression whose value is quantized explicitly. This expression is of the same form as TKNdN's integral, except that the roles played by the Bloch wave numbers are now played by the phase parameters in the boundary conditions. Thus the same topological identification can be made in the new theory.

In our generalized formulation it is found that the Hall conductance is quantized in an integer times e^2/h as long as the Fermi gap is finite and the many-body ground-state energy is nondegenerate. This quantized value is topological in the sense that it is unchanged under a variation of the potentials so long as the Fermi gap is kept open. Also, the result does not depend on how this gap is generated. It could be generated by the action of the magnetic field alone (Landau gap), together with a periodic substrate potential (gaps between the subbands), or with the many-body interactions.

To obtain a fractional quantization, we have to require, in addition to a finite Fermi gap, that the ground-state energy is degenerate and the ground states have a discrete symmetry breaking. In this case, the Fermi gap must be generated by the many-body interactions, since otherwise the degeneracy cannot be obtained, nor can the symmetry breaking. Recently, Tao and Wu[11] generalized Laughlin's gauge-invariance argument; our result agrees with theirs.

The degeneracy in the ground-state energy at fractional fillings has been clearly demonstrated by the numerical calculation of Su[12] for a few small systems with torus geometry. On the other hand, such degeneracy was not found in Haldane's[13] numerical calculations with spherical geometry. At the present time we cannot conclude (although we suspect) whether the Haldane system will eventually present a degeneracy in the thermodynamic limit, because it is not clear how the spherical geometry could

be fitted into our formulation.

Recently, the topological nature of the quantized Hall effect has been revealed in a different approach by Levine, Libby, and Pruisken.[14] They showed that the effective Lagrangian in a replica treatment of a two-dimensional (2D) disordered electron system in a transverse magnetic field contains a nonperturbative topological term which gives rise to the quantization of the Hall conductance. Unfortunately, this theory seems unable to deal with the many-body interaction and therefore the fractionally quantized Hall effect.

In Sec. II we present the formulation generalizing TKNdN's theory in the integral quantized Hall effect. In Sec. III we illustrate the idea of how the fractional quantized Hall conductance could be expressed as a topological invariant. Finally, in Sec. IV we give a brief discussion of the relation between our formulation and Laughlin's gauge-invariance argument.

II. THE INTEGER CASE

The method we will use relies on the fact that for an energy in the Fermi gap the Green function falls off exponentially with distance, and so the current induced by an electric field is a local function of the field. Therefore the response to a field in the interior (away from the edge) of a two-dimensional system can be calculated with only exponentially small errors by replacing the realistic boundary conditions with convenient artificial ones.

Edge currents are a separate problem and cannot be calculated in this way, since they depend on a delicate balance between the diamagnetic currents at the two edges. However, it is possible to devise geometrical conditions under which there are no edge currents. For example, we could consider a Corbino disc or a cylinder with the emf applied in the azimuthal direction.

We now consider a two-dimensional interacting electron system in both a magnetic field $B\hat{z}$ perpendicular to the plane and an electric field $E\hat{x}$ in the plane. The substrate potential may or may not be periodic in space. The Hall current which flows in the y direction can be calculated by the Kubo formula[15] derived from a linear-response theory as

$$\sigma = \frac{ie^2\hbar}{A} \sum_{n\,(>0)} \frac{(v_1)_{0n}(v_2)_{n0} - (v_2)_{0n}(v_1)_{n0}}{(E_0 - E_n)^2} , \qquad (2.1)$$

where $A = L_1 L_2$ is the area of the system; the subscripts 0 and n label the ground state and the excited states of the N-body Hamiltonian in the absence of the external electric field:

$$H = \sum_{i=1}^{N} \left[\frac{1}{2m_i} \left(-i\hbar\frac{\partial}{\partial x_i} \right)^2 + \frac{1}{2m_i} \left(-i\hbar\frac{\partial}{\partial y_i} - eBx_i \right)^2 \right]$$
$$+ \sum_{i=1}^{N} U(x_i,y_i) + \sum_{j=1}^{N} \sum_{\substack{i=1 \\ i<j}}^{N} V(|\mathbf{r}_i - \mathbf{r}_j|) . \qquad (2.2)$$

Correspondingly, E_0 and E_n are the eigenenergies. The velocity operators appearing in the Kubo formula are given by

$$v_1 = \sum_{i=1}^{N} \frac{1}{m_i} \left[-i\hbar\frac{\partial}{\partial x_i} \right] ,$$

$$v_2 = \sum_{i=1}^{N} \frac{1}{m_i} \left[-i\hbar\frac{\partial}{\partial y_i} - eBx_i \right] . \qquad (2.3)$$

A realistic boundary condition is

$$\psi(x_i = L_1) = \psi(x_i = 0) = 0 , \qquad (2.4a)$$

$$\psi(y_i + L_2) = e^{i\beta L_2}\psi(y_i) , \qquad (2.4b)$$

$$i = 1, 2, \ldots, N$$

where the phase parameter β is independent of the particle indices, as is required by the total antisymmetry. But we are only interested in the bulk contribution to the Hall conductance; the condition (2.4a) can then be relaxed to the following form:

$$\psi(x_i + L_1) = e^{i\alpha L_1}e^{i(eB/\hbar)y_i L_1}\psi(x_i) , \qquad (2.4c)$$

where the y-dependent phase factor is necessary for the Hamiltonian to be Hermitian. At this moment one should note that the boundary conditions (2.4b) and (2.4c) are appropriate only when the particular Landau gauge (2.2) is chosen, but in general we can use instead the gauge-covariant boundary conditions of the form

$$\mathcal{T}_i(L_1,\hat{x})\psi(x_i) = e^{i\alpha L_1}\psi(x_i) , \qquad (2.4d)$$

$$\mathcal{T}_i(L_2,\hat{y})\psi(y_i) = e^{i\beta L_2}\psi(y_i) , \qquad (2.4e)$$

$$i = 1, 2, \ldots, N$$

where $\mathcal{T}_i(L_1,\hat{x})$ and $\mathcal{T}_i(L_2,\hat{y})$ are the single-particle magnetic translation operators in the x and y directions.[16] With this generalization the argument will follow the same line, so we would rather stay with the special case (2.4b) and (2.4c).

Now we make the unitary transformation

$$\phi_n = \exp[-i\alpha(x_1 + \cdots x_N)]$$
$$\times \exp[-i\beta(y_1 + \cdots + y_N)]\psi_n . \qquad (2.5)$$

Then (2.1) becomes

$$\sigma = \frac{ie^2}{A\hbar} \sum_{n\,(>0)} \frac{\left\langle \phi_0 \left| \frac{\partial\tilde{H}}{\partial\alpha} \right| \phi_n \right\rangle \left\langle \phi_n \left| \frac{\partial\tilde{H}}{\partial\beta} \right| \phi_0 \right\rangle - \left\langle \phi_0 \left| \frac{\partial\tilde{H}}{\partial\beta} \right| \phi_n \right\rangle \left\langle \phi_n \left| \frac{\partial\tilde{H}}{\partial\alpha} \right| \phi_0 \right\rangle}{(E_0 - E_n)^2} , \qquad (2.6)$$

where $\tilde{H}$ is the transformed Hamiltonian. Equivalently, $\tilde{H}$ can be obtained from H by the following replacement:

$$-i\frac{\partial}{\partial x_i} \rightarrow -i\frac{\partial}{\partial x_i}+\alpha, \quad -i\frac{\partial}{\partial y_i} \rightarrow -i\frac{\partial}{\partial y_i}+\beta. \quad (2.7)$$

It is clear that $(1/\hbar)\partial\tilde{H}/\partial\alpha$ and $(1/\hbar)\partial\tilde{H}/\partial\beta$ are just the transformed velocity operators. By a simple manipulation we can express (2.6) in terms of the partial derivatives of the transformed wave function for the ground state of the many-body system:

$$\sigma = \frac{ie^2}{\hbar A}\left[\left\langle\frac{\partial\phi_0}{\partial\alpha}\bigg|\frac{\partial\phi_0}{\partial\beta}\right\rangle - \left\langle\frac{\partial\phi_0}{\partial\beta}\bigg|\frac{\partial\phi_0}{\partial\alpha}\right\rangle\right]$$

$$= \frac{ie^2}{\hbar}\left[\left\langle\frac{\partial\phi_0}{\partial\theta}\bigg|\frac{\partial\phi_0}{\partial\varphi}\right\rangle - \left\langle\frac{\partial\phi_0}{\partial\varphi}\bigg|\frac{\partial\phi_0}{\partial\alpha}\right\rangle\right], \quad (2.8)$$

where $\theta = \alpha L_1$ and $\varphi = \beta L_2$.

So far the derivatives are formal and we still cannot see why the Hall conductance should be quantized. To proceed further, we need to assume that there is always a finite energy gap between the ground state and the excitations under any given boundary conditions of the form in (2.4b) and (2.4c). Also, it is plausible to say that the bulk conductance as given by the Kubo formula should be insensitive to the boundary conditions if the particles do not have long-range correlations in the ground state. We leave the justification of this point to the Appendix. In fact, in the special case of zero interaction and flat substrate potential, one can explicitly show that the above expression is indeed independent of the parameters, even without taking the thermodynamic limit.

Consequently we can equate σ with its average over all the phases $(0 \leq \theta < 2\pi, \ 0 < \varphi \leq 2\pi)$ that specify different boundary conditions, i.e.,

$$\sigma = \bar{\sigma} = \frac{e^2}{h}\int_0^{2\pi}\int_0^{2\pi}d\theta\,d\varphi\frac{1}{2\pi i}\left[\left\langle\frac{\partial\phi_0}{\partial\varphi}\bigg|\frac{\partial\phi_0}{\partial\theta}\right\rangle\right.$$

$$\left.-\left\langle\frac{\partial\phi_0}{\partial\theta}\bigg|\frac{\partial\phi_0}{\partial\varphi}\right\rangle\right]. \quad (2.9)$$

This is of the same form as the integral that appeared in TKNdN's original theory, except that positions of the Bloch wave numbers are now taken by the phase parameters θ and φ. Because of the energy gap, the ground state must go back to itself (up to an overall phase factor) as θ or φ changes by 2π, unless the ground state is not uniquely determined by the boundary condition. Thus the Hall conductance is quantized into an integer times e^2/h whenever the ground state is nondegenerate and is separated from the excited states by a finite energy gap.

The integral in (2.9) is actually a topological invariant. It is the first Chern class of a U(1) principal fiber bundle of the ground-state wave functions on the base manifold of a torus T^2 parametrized by the phases θ and φ. Originally such a recognition was made by Avron, Seiler, and Simon[7] in the context of TKNdN's original theory. The base manifold was the magnetic Brillouin zone, and the fibers were the single-particle Bloch waves. Recently, Kohmoto[8] finished a detailed analysis showing how the

abstract topological idea is applied to their expression for the quantized Hall conductance. But since the use of single-particle Bloch waves is essential in TKNdN's theory, this topological idea loses sense as soon as the many-body interaction and the substrate disorder are taken into account.

Fortunately, all these can be recovered by manipulating the phases describing the boundary conditions. Apparently, the Hall conductance should be calculated under a fixed boundary condition as given in (2.8). But the insensitivity of the physical quantity to the boundary conditions allows us to make an average over the phases. In this way an expression for the Hall conductance similar to that of TKNdN's is obtained, so the same topological words can apply to the problem in the rather general situation.

In fact, the recognition of the Hall conductance as a topological invariant is not only of mathematical formality, but also of physical content, because we can explain the stability of the quantization of the Hall conductance against various kinds of perturbations. The reason lies in the fact that the existence of a finite Fermi gap above the ground state is a discrete property which does not depend upon the potentials continuously. Also, in the presence of slight disorder in the substrate, the Fermi level can be locked into the impurity spectra and allows the mobility gap to open in a finite range of the magnetic field. This explains the plateaus of the Hall conductance at the quantized values.

III. THE FRACTIONAL CASE

As mentioned in the introduction, the observation of fractional quantized Hall effect[2] (FQHE) is also accompanied with a vanishing of the longitudinal conductance in the zero-temperature limit. This suggests that a Fermi gap must also exist at the fractional fillings near which FQHE is observed. Several theoretical calculations[17-19] have already justified this point. Furthermore, the ground state is shown to be liquidlike, so correlation between electrons decays rapidly as their separation becomes large. Thus we can continue to use the method employed in the preceding section to equate the Hall conductance to its average over all different boundary conditions as in (2.9). Since a nondegenerate ground state always leads to an integral quantization, we must require a degeneracy in order to explain the fractional quantization. In this case, Eq. (2.9) should be written as

$$\sigma = \bar{\sigma}$$

$$= \frac{e^2}{hd}\sum_{K=1}^{d}\int_0^{2\pi}\int_0^{2\pi}d\theta\,d\varphi\frac{1}{2\pi i}\left[\left\langle\frac{\partial\phi_K}{\partial\varphi}\bigg|\frac{\partial\phi_K}{\partial\theta}\right\rangle\right.$$

$$\left.-\left\langle\frac{\partial\phi_K}{\partial\theta}\bigg|\frac{\partial\phi_K}{\partial\varphi}\right\rangle\right], \quad (3.1)$$

where d is the degree of the degeneracy, and $\{\phi_K\}$ is an orthogonal basis spanning the ground-state Hilbert space. In the above expression we have also used the fact that

there is no coupling between different ground states, because they are macroscopically separated in the sense that they cannot be obtained from one another by a few number of single-particle excitations.

Unlike the nondegenerate case, the integral over θ and φ is no longer a topological invariant, since the variation of θ or φ by 2π does not necessarily lead each ground state back to itself. But the summation over the integrals may still be a topological invariant. We consider this possibility in the following.

Consider the $1/p$ fillings first. We start from the parent states similar to those proposed by Tao and Thouless.[17] The single-particle states in the ground Landau level are

$$e^{i\lambda y}u(x-\lambda b^2), \quad b^2=\frac{\hbar}{eB} . \tag{3.2}$$

In order to satisfy the boundary conditions (2.4b) and (2.4c), we make the following linear combinations:

$$W_m(\alpha,\beta)=\sum_{n=-\infty}^{\infty} e^{i\delta_m\lambda_n b^2}e^{i\lambda_n y}u(x-\lambda_n b^2) , \tag{3.3}$$

where

$$\lambda_n=\beta+\frac{2\pi}{L_2}pn, \quad \delta_n=\alpha+\frac{2\pi}{L_1}m \tag{3.4}$$

and the linear dimensions of the system is chosen to give $1/p$ filling of the ground Landau level:

$$\frac{N}{L_1L_2}=\frac{1}{2\pi pb^2} . \tag{3.5}$$

Then one can show that these satisfy the following boundary conditions:

$$W_m(x+L_1)=e^{i\alpha L_1}e^{iL_1y/b^2}W_m(x) , \tag{3.6}$$

$$W_m(y+L_2)=e^{i\beta L_2}W_m(y) . \tag{3.7}$$

Also one should notice that there are N different states of the form (3.3) (for fixed α and β), since W_{m+N} differs from W_m by only an overall phase factor. This is not surprising because we only used those states in (3.2) whose centers (in the x direction) are separated by a multiple of $(2\pi/L_2)pb^2$ to construct the linear combinations. The parent state of the many-body system is then made by taking the determinant of these single-particle states. We claim that this parent state is equivalent to that proposed by Tao and Thouless,[17] because different linear combinations of the single-particle states give the same determinantal wave function. One can check explicitly that the same boundary conditions as in (3.6) and (3.7) are satisfied by the parent state just constructed, i.e.,

$$\psi(x_l+L_1)=e^{i\alpha L_1}e^{iL_1y_l/b^2}\psi(x_l) , \tag{3.8}$$

$$\psi(y_l+L_2)=e^{i\beta L_2}\psi(y_l) , \tag{3.9}$$

$$l=1,2,\ldots,N .$$

Now we come to the interesting point. Clearly the above boundary condition is specified by $\theta=\alpha L_1$ (mod 2π) and $\varphi=\beta L_2$ (mod 2π). But the parent state can-

not be uniquely determined in the same way, since a variation of αL_1 by 2π does not lead the parent state back to itself. In other words, under a given boundary condition we can construct many different parent states from which the true ground states would be generated by the adiabatic turning on of the many-body interactions. Since a variation of βL_2 by 2π and/or of αL_1 by $p 2\pi$ do lead the state back to itself, p different parent states are obtained:

$$\psi(\theta,\varphi),\psi(\theta+2\pi,\varphi),\ldots,\psi(\theta+(p-1)2\pi,\varphi) . \tag{3.10}$$

These states are orthogonal since they are constructed from different sets of single-particle states. In fact, the true ground states generated from them are also orthogonal to one another. The reason is that these parent states belong to different eigenstates (with different eigenvalues) of the magnetic translation operator $\mathscr{T}((2\pi/L_1)b^2,\hat{y})$ which translates all the particles by $(2\pi/L_1)b^2$ along the y direction, while the operator commutes with the total Hamiltonian and its interaction part from which the S matrix is constructed.[18] In other words the true ground states obtained by acting the S matrix on the parent states must also belong to the eigenstates of $\mathscr{T}((2\pi/L_1)b^2,\hat{y})$ with different eigenvalues, therefore they must be orthogonal to one another.

On the other hand, different parent states are connected by the total magnetic translation $\mathscr{T}((2\pi/L_2)b^2,\hat{x})$ and its powers. Since this operator also commutes with the S matrix, the states generated from different parent states must have the same energy.

Let us now look back to formula (3.1), where ϕ_K are just the true ground states transformed by relation (2.5). If there are no other ground states other than those generated from our parent states, then (3.1) can be written as

$$\sigma=\bar{\sigma}$$

$$=\frac{e^2}{hp}\int_0^{2\pi p}d\theta\int_0^{2\pi}d\varphi\frac{1}{2\pi i}\left[\left\langle\frac{\partial\phi_1}{\partial\varphi}\bigg|\frac{\partial\phi_1}{\partial\theta}\right\rangle\right.$$

$$\left.-\left\langle\frac{\partial\phi_1}{\partial\theta}\bigg|\frac{\partial\phi_1}{\partial\varphi}\right\rangle\right] , \tag{3.11}$$

where we have absorbed the summation into the integration over the extended range $[0\leq\theta\leq p 2\pi]$, because different ground states can be obtained from ϕ_1 by continuous variation of θ by 2π, 4π, etc. We can regard the extended zone $[0\leq\theta<p 2\pi, 0\leq\varphi<2\pi]$ as a torus, because ϕ_1 is led back to itself (up to an overall factor) as θ change by $2\pi p$ or φ by 2π. Thus

$$\sigma=\frac{e^2}{h}\frac{c}{p} , \tag{3.12}$$

where c is the integer given by the integral. As in the nondegenerate case we can continue to attach a topological meaning to this integer if the extended zone is regarded as the base manifold.

To determine the integer c we may turn off the varying part of the substrate potential; the integer is unchanged assuming the Fermi gap is unclosed by this process. We then transform to the moving frame in which the external

electric field becomes zero, we will find no electron current in this frame. Consequently the Hall conductance is directly related to the filling factor $1/p$, so c is unity.

In the above argument we used the concept of parent states to illustrate the idea how the fractional quantization can be obtained within our frame. We do not mean to give a rigorous proof, but we suggest that the following scenario might correspond to the physical reality:

(1) At a $1/p$ filling with p odd and small, the ground-state energy of the system has a p-fold degeneracy. As the phase parameter θ changes by 2π, 4π, etc., the system goes from one ground state to another, and it comes back to the original ground state after θ varies by $2\pi p$.

(2) At a q/p filling, again with p odd and small, the degeneracy is jp-fold with $j \leq q$. But these states fall into j groups, and the states in each group transform in the same way as in (1).

(3) As the density of electrons deviates a small but finite amount from one of the fillings considered above, the ground states may seek a similar structure as in the neighboring case in order to gain the commensurate energy (which is negative) achieved at the neighboring small denominator fillings. If this is right, then the finite plateaus of the Hall conductance observed near the small and odd denominator fillings can be understood.

(4) Since p is even, conjectures (1) and (2) might still be true, but (3) may break down by the fractional statistics recently suggested by a number of people.[20]

Before closing this section we would like to add one more comment. The argument about the degeneracies of the ground state has been presented in terms of the wave function proposed by Tao and Thouless,[17] but it can also be presented in terms of Laughlin's wave function.[19] With the boundary condition (2.4b) (periodicity in the y direction) a wave function confined to the region $0 < x_1 < hNp/eBL_2$ can be written in the form

$$\phi_1 = \exp\left[-\frac{1}{2b^2}\sum_{i=1}^{N}x_i^2 + \beta\sum_{i=1}^{N}z_i\right]$$

$$\times \prod_{\substack{j=1 \\ i<j}}^{N}\prod_{i=1}^{N}(e^{2\pi z_i/L_2}e^{2\pi z_j/L_2})^p \qquad (3.13)$$

where $z_j = x_j + iy_j$. This has the same local properties as the wave function Laughlin writes in cylindrical geometry or Haldane[21] in spherical geometry. This state can be modified to satisfy the condition (2.4c) by expanding it in terms of the single-particle Landau states (3.2), and replacing each of these by its periodic continuation in the x direction (3.3). This state is also degenerate with the $p-1$ orthogonal states which can be constructed by the action of the magnetic translation operator on it.

IV. RELATION TO LAUGHLIN'S GAUGE ARGUMENT

According to Laughlin's theory,[5] the Hall conductance can be expressed in terms of the number of electrons

transported from one edge of the sample to the other after an adiabatic change of the gauge flux by one quantum. We now try to relate this charge transport to the Kubo formula we have been using. The Hamiltonian with a time-dependent gauge parameter is given by (2.2) with the replacement

$$-i\frac{\partial}{\partial y_i} \rightarrow -i\frac{\partial}{\partial y_i} + \beta(t) . \qquad (4.1)$$

The current in the x direction induced by the adiabatic variation of $\beta(t)$ can be calculated from the formula[9,22]

$$J_x = \frac{i\hbar^2\dot{\beta}(t)}{L_1}\sum_{n\,(>0)}\frac{(v_2)_{0n}(v_1)_{n0}-(v_1)_{0n}(v_2)_{n0}}{(E_0-E_n)^2} , \qquad (4.2)$$

where the indices 0 and n indicate the instantaneous eigenstates of the Hamiltonian. The charge transport in a period T during which $\beta(t)$ changes by $2\pi/L_2$ (corresponding to a flux quantum) is thus

$$C = \int_0^T dt\,J_x$$

$$= \int_0^{2\pi/L_2}d\beta\frac{i\hbar^2}{L_1}\sum_{n\,(>0)}\frac{(v_2)_{0n}(v_1)_{n0}-(v_1)_{0n}(v_2)_{n0}}{(E_0-E_n)^2}$$

$$= \frac{h}{e^2}\int_0^{2\pi}\frac{d\varphi}{2\pi}\sigma(\varphi) , \qquad (4.3)$$

or

$$\bar{\sigma} = \frac{e^2}{h}C , \qquad (4.4)$$

where $\sigma(\varphi)$ is the Hall conductance calculated from the Kubo formula under a fixed gauge $\beta \equiv \varphi L_2$.

Note that it is the averaged Hall conductance that corresponds to the charge transport, so Laughlin's argument really involved an approximation. In fact, as he pointed out in his paper, when he made use of the Faraday's theorem, he actually replaced the adiabatic derivative $dU/d\Phi$ by the fraction $\Delta U/\Phi_0 \propto C$. Here U is the total energy of the electron system, Φ the flux, and Φ_0 the flux quantum. In other words, his approximation is of the same nature as ours.

One final comment. When Laughlin[5] tried to establish the quantization of the Hall conductance, he actually based his argument upon the belief that as the electronic states in the bulk go back to themselves, the particle transport from one edge to the other must be an integral number. Although this idea is physically intuitive, it is still not obvious because of the wave nature of the electrons. Our formulation presents a rigorous proof of this idea.

APPENDIX

The insensitivity of the Hall conductance to the boundary conditions is most easily understood in the noninteracting case. Then the Kubo formula (2.6) can be written in terms of the single-particle quantities as

$$\sigma = \frac{ie^2}{A\hbar} \sum_{\epsilon_m\,(<\epsilon_f)} \sum_{\epsilon_n\,(>\epsilon_f)} \frac{\left\langle \varphi_m \left| \frac{\partial \tilde{h}}{\partial \alpha} \right| \varphi_n \right\rangle \left\langle \varphi_n \left| \frac{\partial \tilde{h}}{\partial \beta} \right| \varphi_m \right\rangle - \left\langle \varphi_m \left| \frac{\partial \tilde{h}}{\partial \beta} \right| \varphi_n \right\rangle \left\langle \varphi_n \left| \frac{\partial \tilde{h}}{\partial \alpha} \right| \varphi_m \right\rangle}{(\epsilon_m - \epsilon_n)^2} \tag{A1}$$

where $\tilde{h}$ is the transformed single-particle Hamiltonian and φ_n are the single-particle wave functions. With some manipulations, this can in turn be written in terms of the Green functions as

$$\sigma = \frac{ie^2}{A\hbar} \oint \frac{dz}{2\pi i} \mathrm{Tr} \left[g \frac{\partial \tilde{h}}{\partial \alpha} g \frac{\partial \tilde{h}}{\partial \beta} g \right] , \tag{A2}$$

where the integral contour surrounds the filled state energies. The derivative of the Hall conductance σ with respect to θ is

$$\frac{\partial \sigma}{\partial \theta} = \frac{ie^2}{A\hbar L_1} \oint \frac{dz}{2\pi i} \mathrm{Tr} \left[2g \frac{\partial \tilde{h}}{\partial \alpha} g \frac{\partial \tilde{h}}{\partial \alpha} g \frac{\partial \tilde{h}}{\partial \beta} g + g \frac{\partial \tilde{h}}{\partial \alpha} g \frac{\partial \tilde{h}}{\partial \beta} g \frac{\partial \tilde{h}}{\partial \alpha} g + g \frac{\partial^2 \tilde{h}}{\partial \alpha^2} g \frac{\partial \tilde{h}}{\partial \beta} g \right] . \tag{A3}$$

Now, because of the existence of the Fermi gap the energy parameter z can be chosen away from the spectrum of the extended states, hence the Green function $g(\mathbf{r},\mathbf{r}')$ is exponentially bounded as $|\mathbf{r}-\mathbf{r}'|$ becomes large.[4] Thus $\partial\sigma/\partial\theta$ is of the order of $(l/L_1)e^2/h$, with l being the localization length of the Green function. Similarly we have

$$\frac{\partial \sigma}{\partial \varphi} \sim \frac{l}{L_2} \frac{e^2}{h} . \tag{A4}$$

When the many-body interactions are taken into account, we can use the method used in Ref. 9 to estimate $\partial\sigma/\partial\theta$ and $\partial\sigma/\partial\varphi$. Although the manipulations are complicated, the same result can be obtained under the assumption that the electrons do not have long-range correlations in the ground state.

---

*Also at Cavendish Laboratory, University of Cambridge, Madingley Road, Cambridge CB3 0HE, United Kingdom.

†On leave from Department of Physics, University of Utah, Salt Lake City, UT 84112 (address after April 1, 1985).

[1]K. V. Klitzing, G. Dorda, and M. Pepper, Phys. Rev. Lett. **45**, 494 (1980).

[2]D. C. Tsui, H. L. Störmer, and A. C. Gossard, Phys. Rev. Lett. **48**, 1559 (1982); H. L. Störmer, D. C. Tsui, A. C. Gossard, and J. C. M. Hwang, in Proceedings of the 16th International Conference on Physics of Semiconductors (unpublished).

[3]H. Aoki and T. Ando, Solid State Commun. **38**, 1079 (1981); R. E. Prange, Phys. Rev. B **23**, 4802 (1981); R. E. Prange and R. Joynt, *ibid.* **25**, 2943 (1982).

[4]D. J. Thouless, J. Phys. C **14**, 3475 (1981).

[5]R. B. Laughlin, Phys. Rev. B **23**, 5632 (1981).

[6]D. J. Thouless, M. Kohmoto, M. 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); B. Simon, Phys. Rev. Lett. **51**, 2167 (1983).

[8]M. Kohmoto (unpublished).

[9]Q. Niu and D. J. Thouless, J. Phys. A **17**, 2453 (1984).

[10]B. I. Halperin, Phys. Rev. B **25**, 2185 (1982).

[11]R. Tao and Y. S. Wu, Phys. Rev. B **30**, 1097 (1984).

[12]W. P. Su, Phys. Rev. B **30**, 1069 (1984).

[13]F. D. M. Haldane (private communication).

[14]H. Levine, S. B. Libby, and A. M. M. Pruisken, Nucl. Phys. B **240**, 30 (1984); **240**, 49 (1984); **240**, 71 (1984).

[15]At zero temperature this form of the Kubo formula is most easily derived by first-order static perturbation expansion.

[16]J. Zak, Phys. Rev. **134**, A1607 (1964).

[17]R. Tao and D. J. Thouless, Phys. Rev. B **28**, 1142 (1983).

[18]R. Tao, Phys. Rev. B **29**, 636 (1984).

[19]R. B. Laughlin, Phys. Rev. Lett. **50**, 1395 (1983); D. Yoshioka, B. I. Halperin, and P. A. Lee, Surf. Sci. **142**, 155 (1984).

[20]B. I. Halperin, Phys. Rev. Lett. **52**, 1583 (1984); D. Arovas, J. R. Schrieffer, and F. Wilczek, Phys. Rev. Lett. **53**, 722 (1984); R. Tao and Y. S. Wu (unpublished).

[21]F. D. M. Haldane, Phys. Rev. Lett. **51**, 605 (1983).

[22]D. J. Thouless, Phys. Rev. B **27**, 6083 (1983).