Quantization of the Hall Conductance for General, Multiparticle Schrödinger Hamiltonians

Joseph E. Avron
Department of Physics, Technion–Israel Insitute of Technology, Haifa 32000, Israel

and

Ruedi Seiler
Fachbereich 3, Mathematik Technische Universität Berlin, D-1000 Berlin, Germany
(Received 15 October 1984)

We describe a precise mathematical theory of the Laughlin argument for the quantization of the Hall conductance for general multiparticle Schrödinger operators with general background potentials. The quantization is a consequence of the geometric content of the conductance, namely, that it can be identified with an integral over the first Chern class. This generalizes ideas of Thouless *et al.*, for noninteracting Bloch Hamiltonians to general (interacting and nonperiodic) ones.

PACS numbers: 03.65.Bz, 02.40.+m, 72.20.My

The integer quantization of the Hall conductance has been explained by Laughlin[1] making clever use of a nontrivial geometry: a ring threaded by a flux tube, combined with a gauge argument. The impact of this work on the development of the subject cannot be overestimated.

Our purpose here is to describe a precise mathematical theory of this argument. The two key issues are, first, Laughlin's identification of a physical quantity as the Hall conductance averaged over one unit of quantum flux (of the flux tube that threads the ring). Following Laughlin we shall slightly abuse the terminology and call it the Hall conductance.

The second main theme will be the identification of the geometric content of the Hall conductance: Roughly speaking, there is a natural notion of curvature describing how the state of the system is parallel transported in the Hilbert space of states. The Hall conductance is a suitable integral of the corresponding curvature. More precisely, it is an integral over the first Chern class.[2] This was first recognized by Thouless *et al.* in the special case of noninteracting Bloch Hamiltonians.[3] What is shown here is that this holds generally, with electron-electron interaction and general background potential. (No flux averaging is necessary for Bloch Hamiltonians.) Bellissard generalized the result of Thouless *et al.* from rational to real magnetic flux.[4]

We shall replace Laughlin's condition that the Fermi energy lies in the region of the localized states (which is not an appropriate condition for multiparticle Hamiltonians) by a condition of nondegeneracy of the multiparticle ground states (for all fluxes).

Our work has been independent of, but is nevertheless closely related to, a recent published paper of Niu and Thouless.[5] The general framework is similar, although there are some differences in the details and in the approach. In both works one needs the ground state to be separated from the rest of the spectrum by a finite gap (the nondegeneracy condition). In both approaches one considers time-periodic Hamiltonians (in Niu-Thouless only up to unitary equivalence). In Niu and Thouless the time dependence resides in the substrate potential and it comes from a Galilean transformation that removes the electric field. In our case, the time dependence comes from generating the electromotive force by a flux tube, and so resides in the minimal coupling term in the Hamiltonian. In both, strict quantization is obtained only after a suitable averaging: In Niu-Thouless the averaging is over boundary conditions and here, averaging is over the fluxes in Laughlin's flux tube.

We feel that avoiding the Galilean transformation is a distinct advantage of the present approach. Also, essentially all the structure we shall use of the Hamiltonian is minimal coupling. This makes it clear, for example, that there are no relativistic corrections, nor spin effects, nor finite-volume corrections, not even exponentially small ones, something which is less obvious in Ref. 5. We also note that our approach is more geometric and in our opinion, simpler.

We stress that strict quantization is proven for the Hall conductance averaged over the fluxes in Laughlin flux tubes. It is interesting to investigate under what conditions the Hall conductance for most flux values is close to its average. This problem has been investigated by Thouless[6] and Niu-Thouless,[5] who showed that under suitable additional conditions this is indeed the case in the thermodynamic limit of infinitely large systems. This is reasonable as it is easy to see that variations in the conductance are the same as those caused by changing the lengths of the connecting leads. The results in Ref. 5 are essentially perturbative, and are based on the assumption that correlations decay fast.

The nondegeneracy condition turns out to be a sufficient condition for integer quantization at zero temperature and finite volume (we shall study the infinite-volume limit elsewhere). From this one can also learn about fractions: Tao and Wu[7] and Tao[8] have extended Laughlin's argument to degenerate ground state and have shown that the ground-state degeneracy is related to the denominator of the fractional conductance. The present analysis extends to this case as well and agrees with these results. However, since degeneracies are nongeneric it is not clear how one can get plateaus at fractions. We speculate that this occurs via diamagnetism as the relation between the external and internal fields is singular near degeneracies because the ground-state energy is not smooth there. We have nothing to say about the odd-denominator rule.

We find it convenient to formulate the problem in the following setting. Consider a domain Λ in the plane, with two holes. The holes are threaded with two flux tubes with fluxes $\phi=(\phi_1,\phi_2)$. On the boundary of Λ we impose Dirichlet boundary conditions. Now consider the multielectron Schrödinger Hamiltonian for this system, $H(\phi)$, depending on ϕ through minimal coupling. The geometry, shown in Fig. 1(a), is motivated by that of the physical Hall effect shown in Fig. 1(b). The motivation is clear once the leads that connect to the sample in Fig. 1(b) are considered as part of the system. [One may remove the leads in Fig. 1(b) at the expense of imposing periodic boundary conditions.] We take one of the fluxes, say ϕ_1, linearly increasing with time thereby replacing the battery in Fig. 1(b). The second flux is the flux tube introduced by Laughlin in his original argument and is shown also in Fig. 1(b). The translation between the two geometries will always be evident and henceforth we stick with that of Fig. 1(a).

Adding one unit of quantum flux to ϕ_1 or ϕ_2 is equivalent to a gauge transformation. $H(\phi)$ is therefore a continuous function of two variables with a natural period of 2π in each (in units where $\hbar=e=1$). There is now a formal analogy with Bloch Hamiltonians in two dimensions where $H(k)$ is a function of the

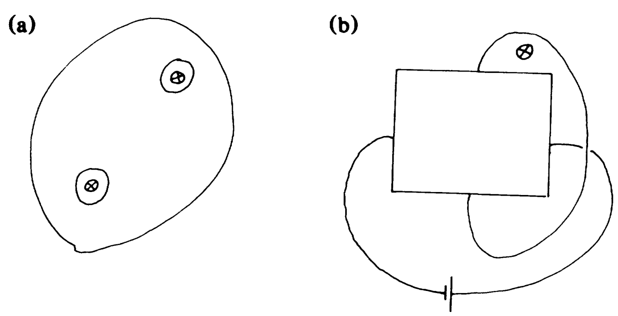

FIG. 1. (a) The domain Λ with the two holes threaded by the two flux tubes; (b) the standard Hall settings with the connecting leads considered as part of the system. Both geometries have two independent closed cycles. The battery in (b) is replaced by a time-dependent flux tube in (a). Both have an extra flux tube which is put there "by hand" in the Laughlin argument.

two Bloch momenta with natural period given by the Brillouin zone. This suggests that the ideas developed by Thouless and co-workers[3,5] and by Thouless[6] for Bloch Hamiltonians could be applied in the present circumstances as well.

For our purposes it is convenient to consider Hamiltonians that are smooth and periodic functions of ϕ. To achieve this, introduce two cuts in Λ so that the resulting set, $\tilde{\Lambda}$, is simply connected. The vector potential associated with ϕ is the gradient of the function $F(\phi)$ which is regular in $\tilde{\Lambda}$ and with discontinuities ϕ_1 and ϕ_2 across the cuts. $\tilde{H}(\phi)=\exp[-iF(\phi)]H(\phi)\times\exp[iF(\phi)]$ is formally ϕ independent. The ϕ dependence enters through the $\exp(i\phi_{1,2})$ boundary conditions that the wave function has to satisfy across the cuts. $\tilde{H}(\phi)$ is manifestly periodic in ϕ. $H(\phi)$ is periodic up to unitary equivalence.

All we shall need in order to prove the quantization is to assume that $H(\phi)$ has a nondegenerate ground state for all values of ϕ. Let us examine this condition. Recall that according to the Wigner–von Neumann no-crossing theorem, eigenvalue crossing has codimension three.[9] This says that Hamiltonians that depend on two parameters will not have any degeneracy (generically) while those that depend on three parameters have points of degeneracy. Thus, with ϕ as parameters there are no crossings, but if one varies the magnetic field as well, there will be crossings at special values of the magnetic field. These are the special values of B where the Hall conductance jumps. The proof of integer conductance we give holds for generic Hamiltonians at zero temperature. Had there been a way to prepare the system at an excited and isolated eigenstate, the Hall conductance would be quantized then too. It follows that at finite temperature it is, generically, a thermal average of integers.[10]

Let $|\Omega\rangle$ be the ground state of $H(\phi)$ with energy $E(\phi)$. By the assumption of nondegeneracy, it can be chosen normalized and smooth in ϕ (for all ϕ). The same holds for $|\tilde{\Omega}\rangle$. However, it is in general impossible to require periodicity in ϕ as well. In contrast, the spectral projection, $\tilde{P}(\phi)=|\tilde{\Omega}\rangle\langle\tilde{\Omega}|$, is both smooth and periodic in ϕ.

We shall establish the following set of formulas for the (flux-averaged) Hall conductance $\langle\sigma_{\rm H}\rangle$:

$$2\pi\langle\sigma_{\rm H}\rangle=\frac{i}{2\pi}\int_{\partial T}\langle\Omega|d\Omega\rangle \tag{1a}$$

$$=\frac{i}{2\pi}\int_T\langle d\tilde{\Omega}|d\tilde{\Omega}\rangle \tag{1b}$$

$$=\frac{i}{2\pi}\int_T \mathrm{Tr}(dPPdP) \tag{1c}$$

$$=\frac{i}{2\pi}\int_T \mathrm{Tr}(d\tilde{P}\tilde{P}d\tilde{P}) \tag{1d}$$

$$=\frac{i}{2\pi}\int_T\langle\Omega|dH\hat{R}^2dH|\Omega\rangle. \tag{1e}$$

T is the square of length 2π and ∂T is its boundary.

The physics lie in showing that (1a) follows from a reasonable definition of the Hall conductance. The equivalence of the various formulas turns out to be simple. Equations (1a) and (1b) are analogs of the basic equations in Thouless *et al.* and therefore express an integer by their argument. This is a standard mathematical fact and we shall sketch a simple proof of it towards the end of this note. Equations (1c) and (1d) were used before in Ref. 11. Equation (1e) is the standard Kubo formula (this is explained below). Here $\hat{R}=(2\pi i)^{-1}\oint R(z)/(z-E)dz=R(1-P)$ is the reduced resolvent.

If $\phi_1=-Vt$, $\phi_2=$const, there is an electromotive force V around hole 1 and "Hall current" $I_2=\sigma_{21}V$ around hole 2. We shall establish a formula for σ_{21}, in the limit $V\rightarrow 0$ assuming that at time $t=0$ the system is in a state Ψ which is an eigenstate of $H(\phi_1=0,\phi_2)$. We may choose $\Psi(t=0,\phi_2)=\Omega(\phi_1=0,\phi_2)$. Ψ evolves in time according to the Schrödinger equation $-iV\partial_1\Psi=H(\phi)\Psi$. In the limit $V\rightarrow 0$, Ψ evolves adiabatically in the ϕ_1 variable. Therefore we shall have to use in the following the adiabatic wave function

$$\Psi^{\mathrm{ad}}(\phi)=\lim_{V\rightarrow 0}[\exp(i/V)\int_0^{\phi_1}d\phi_1'\,E(\phi_1',\phi_2)]\Psi(\phi)$$

which satisfies the evolution equation[12]

$$\partial_1\Psi^{\mathrm{ad}}=[\partial_1 P,P]\Psi^{\mathrm{ad}}. \tag{2}$$

As a result of the fundamental expression for the current[1] $I_2=\langle\Psi|\partial_2 H|\Psi\rangle$, one finds for the conductivity

$$\sigma_{21}=-i(\partial_2\langle\Psi,\partial_1\Psi\rangle+\langle\partial_1\Psi,\partial_2\Psi\rangle-\langle\partial_2\Psi,\partial_1\Psi\rangle). \tag{3}$$

The adiabatic wave function does not represent $\Psi(\phi)$ well enough to compute σ_{21}, but it is sufficient to compute $\langle\sigma_{21}\rangle$.[13] The reason for that is that $\langle\sigma_{21}\rangle$ can be written as a line integral (rather than a surface integral):

$$2\pi\langle\sigma_{21}\rangle=-\frac{i}{2\pi}\{\int_0^{2\pi}d\phi_1\langle\Psi,\partial_1\Psi\rangle|_{\phi_2=0}^{2\pi}+\int_T\langle d\Psi,d\Psi\rangle\} \tag{4}$$

$$=-\frac{i}{2\pi}\int_0^{2\pi}d\phi_2\langle\Psi,\partial_2\Psi\rangle|_{\phi_1=0}^{2\pi} \tag{5}$$

$$=-\frac{i}{2\pi}\int_0^{2\pi}d\phi_2\langle\Psi^{\mathrm{ad}},\partial_2\Psi^{\mathrm{ad}}\rangle|_{\phi_1=0}^{2\pi}. \tag{6}$$

In going from (4) to (5) we used Stokes theorem; (6) follows from (5) due to periodicity of the ground-state energy $E(\phi)$. Since $\Psi^{\mathrm{ad}}(\phi)$ is a ground-state wave function it follows from our spectral assumption that $\Psi^{\mathrm{ad}}(\phi)$ equals $\Omega(\phi)$ up to a phase $\exp i\gamma(\phi)$ (Berry's phase[14,15]). Hence we get for the Hall conductivity

$$2\pi\langle\sigma_{21}\rangle=-\frac{i}{2\pi}\int_0^{2\pi}d\phi_2\{i\partial_2\gamma+\langle\Omega,\partial_2\Omega\rangle\}|_{\phi_1=0}^{2\pi}=-\frac{i}{2\pi}[\Gamma(2\pi)-\Gamma(0)]-\frac{i}{2\pi}\int_0^{2\pi}d\phi_2\langle\Omega,\partial_2\Omega\rangle|_{\phi_1=0}^{2\pi}, \tag{7}$$

where we used the notation $\Gamma(\phi_2)=\gamma(\phi_1=2\pi,\ \phi_2)$ and the fact $\gamma(\phi_1=0,\phi_2)=0$. To compute the Γ's we use Eq. (2)

$$0=\langle\Psi^{\mathrm{ad}},\partial_1\Psi^{\mathrm{ad}}\rangle=i\partial_1\gamma+\langle\Omega,\partial_1\Omega\rangle.$$

Integrating yields

$$\Gamma(2\pi)-\Gamma(0)=i\int_0^{2\pi}d\phi_1\langle\Omega,\partial_1\Omega\rangle|_{\phi_2=0}^{2\pi}. \tag{8}$$

Inserting (8) into (7) gives (1a).

A simple computation gives $d\langle\Omega|d\Omega\rangle=\langle d\Omega|d\Omega\rangle=\mathrm{Tr}(dPPdP)$. This establishes $a=c$ and $b=d$. The equivalence $c=d$ follows from

$$\mathrm{Tr}(dPPdP)=\mathrm{Tr}(d\tilde{P}\tilde{P}d\tilde{P})+id\,\mathrm{Tr}(\tilde{P}dF). \tag{9}$$

$\tilde{P}dF$ is continuous in $\boldsymbol{\phi}$ on T so the boundary term vanishes upon integration. $c=e$ follows from the operator identity:

$$dPPdP=\hat{R}(dH)P(dH)\hat{R}, \tag{10}$$

and the cyclicity of the trace.

Equation (1e) is Kubo's formula. To see that write $\mathbf{A}_i=\phi_i\mathbf{a}_i$ for the two vector potentials with $\mathbf{a}_i$ normalized by $\int_{c_j}\mathbf{a}_i\cdot\mathbf{d}_1=2\pi\delta_{ij}$; c_j is a loop around hole j. $dH=-\frac{1}{2}\{\mathbf{v}\cdot\mathbf{a}_i\}d\phi_i$ with $\mathbf{v}=\sum_{i=1}^{N}\mathbf{v}_i$ the cannonical velocity and N is the number of particles. Since in (1e) only the antisymmetric part enters, one of the terms in dH must be identified with the perturbation and the second with the observable. If we take $\phi_1(t)$ and ϕ_2 fixed it follows from minimal coupling that the $d\phi_1$ term in dH is indeed the perturbation. To see that the $d\phi_2$ term is the response we note that without loss of generality we can choose $\mathbf{a}_2=(2\pi)^{-1}(\mathrm{grad}\theta)$ where the polar coordinates (r,θ) have their origin in hole 2. The $d\phi_2$ term in dH is now naturally interpreted as the quantum mechanical observable associated with the rotation frequency $\theta/2\pi$ around hole 2. After some algebra this is seen to reduce (1e) to standard forms of the Kubo formula (with a $\boldsymbol{\phi}$ average). It is interesting that this version is free from the uncontrolled interchange of the $V\rightarrow 0$, $t\rightarrow\infty$ limits, or the adiabatic switching which makes the derivation of the dissipative part of Kubo formulas mathematically formal.

We conclude with a proof that (1b) is an integer when integrated on any two-dimensional manifold without a boundary (in our case a 2-torus).

Take $|\tilde{\Omega}(\boldsymbol{\phi})\rangle$ as our fixed basis and $|\tilde{\Psi}_1(\boldsymbol{\phi})\rangle$ satisfying the adiabatic dynamics along the loop 1. $|\tilde{\Omega}\rangle$ and $|\tilde{\Psi}_1\rangle$ are related by the holonomy[14] (Berry's) phase: $|\tilde{\Psi}_1\rangle = (\exp i\gamma_1)|\tilde{\Omega}(\phi)\rangle$. Now if 1 is the boundary of S, $\gamma_1 = i\int_1 \langle \tilde{\Omega}|d\tilde{\Omega}\rangle = i\int_S \mathrm{Tr}(d\tilde{P}\tilde{P}d\tilde{P})$.

On a closed orientable manifold such a 1 is in fact the boundary of both $S_{\rm out}$ and $S_{\rm in}$. Therefore, from the uniqueness $|\tilde{\Psi}\rangle$:

$$\left(\int_{S_{\rm out}} - \int_{S_{\rm in}}\right)\mathrm{Tr}(d\tilde{P}\tilde{P}d\tilde{P}) = 2\pi(\text{integer}). \qquad (11)$$

The integrand is smooth. Hence in the limit that $S_{\rm in}$ shrinks to a point and $S_{\rm out}$ is the entire manifold we obtain the requisite result $\exp i\int_T \mathrm{Tr}(d\tilde{P}\tilde{P}d\tilde{P}) = 1$.

In conclusion, we have shown that Laughlin's argument and its extension to fractions by Tao and Wu can be understood in terms of the first Chern class of the line bundle over the torus associated to the ground state of $\tilde{H}(\boldsymbol{\phi})$.

We thank H. Cycon, C. Erdmann, M. Klein, A. Raphaelian, and in particular Boris Shapiro, Barry Simon, and Rainer Wüst for useful discussions. One of us (J.A.) thanks the Technische Universität Berlin for its kind hospitality.

[1]R. B. Laughlin, Phys. Rev. B **23**, 5632–5633 (1981).

[2]J. W. Milnor and J. D. Stasheff, *Characteristic Classes*, Annals of Mathematics Studies, Vol. 76 (Princeton Univ. Press, Princeton, 1974).

[3]D. J. Thouless, M. Kohmoto, M. P. Nightingale, and M. den Nijs, Phys. Rev. Lett. **49**, 405–408 (1982).

[4]J. Bellissard, private communication. This most elegant version makes use of K theory.

[5]Q. Niu and D. J. Thouless, J. Phys. A **17**, 2453–2462 (1984).

[6]D. J. Thouless, Phys. Rev. B **27**, 6083–6087 (1983).

[7]R. Tao and Y. S. Wu, Phys. Rev. B **30**, 1097–1098 (1984).

[8]R. Tao, to be published.

[9]J. von Neumann and E. P. Wigner, Phys. Z **30**, 467–470 (1929).

[10]In the nongeneric case, where bands can cross, the ideas developed here together with a generalization of the adiabatic theorem can be used to show that the conductance is a fraction with a denominator that is a divisor of the number of bands that cross. This will be discussed in greater detail elsewhere: J. E. Avron, R. Seiler, and B. Simon, to be published; also see H. L. Stormer, A. Chang, D. C. Tsui, J. C. M. Hwang, and W. Wiegmann, Phys. Rev. Lett. **50**, 1953 (1983); R. Laughlin, Phys. Rev. Lett. **50**, 1395 (1983); F. D. M. Haldane, Phys. Rev. Lett. **51**, 1142 (1983); B. Halperin, Phys. Rev. Lett. **52**, 1583 (1984).

[11]J. E. Avron, R. Seiler, and B. Simon, Phys. Rev. Lett. **51**, 51–53 (1983).

[12]T. Kato, J. Phys. Soc. Jpn. **5**, 435–439 (1950).

[13]Ψ can be expanded in powers of V (adiabatic expansion). The zero-order term is up to a phase $\Psi^{\rm ad}$; to get σ_{21} one has to put the first-order term $\hat{R}^2\partial_1 H\Psi^{\rm ad}$ into the formula for the Hall current I_2. This gives σ_{21} in the form of a Kubo formula.

[14]M. V. Berry, Proc. Roy. Soc. London, Ser. A **392**, 45–57 (1984).

[15]B. Simon, Phys. Rev. Lett. **51**, 2167–2170 (1983).