Quantized Hall conductivity in two dimensions

R. B. Laughlin
Bell Laboratories, Murray Hill, New Jersey 07974
(Received 20 January 1981)

It is shown that the quantization of the Hall conductivity of two-dimensional metals which has been observed recently by Klitzing, Dorda, and Pepper and by Tsui and Gossard is a consequence of gauge invariance and the existence of a mobility gap. Edge effects are shown to have no influence on the accuracy of quantization. An estimate of the error based on thermal activation of carriers to the mobility edge is suggested.

There has been considerable interest in the remarkable observation made recently by von Klitzing, Dorda, and Pepper[1,2] and by Tsui and Gossard[2] that, under suitable conditions, the Hall conductivity of an inversion layer is quantized to better than one part in 10^5 to integral multiples of e^2/h. The singularity of the result lies in the apparent total absence of the usual dependence of this quantity on the density of mobile electrons, a sample-dependent parameter. As it has been proposed[1] to use this effect to define a new resistance standard or to refine the known value of the fine-structure constant, an important issue at present is to what accuracy the quantization is exact, particularly in the regime of high impurity density. Some light has been shed on this question by the renormalized weak-scattering calculations of Ando,[3] who has shown that the presence of an isolated impurity does not affect the Hall current. A similar result has been obtained recently by Prange,[4] who has shown that an isolated δ-function impurity does not affect the Hall conductivity to lowest order in the drift velocity $v = cE/H$, even though it binds a localized state, because the remaining delocalized states carry exactly enough extra current to compensate for its loss. The exactness of these results and their apparent insensitivity to the type or location of the impurity suggest that the effect is due, ultimately, to a fundamental principle. In this communication, we point out that it is, in fact, due to the long-range phase rigidity characteristic of a supercurrent, and that quantization can be derived from gauge invariance and the existence of a mobility gap.

We consider the situation illustrated in Fig. 1, of a ribbon of two-dimensional metal bent into a loop of circumference L, and pierced everywhere by a magnetic field H_0 normal to its surface. The density of states of this system, also illustrated in Fig. 1, consists, in the absence of disorder, of a sequence of δ functions, one for each Landau level. These broaden, in the presence of disorder, into bands of extended states separated by tails of localized ones. We consider the disordered case with the Fermi level in a mobility gap, as shown.

We wish to relate the total current I carried around the loop to the potential drop V from one edge to another. This current is equal to the adiabatic derivative of the total electronic energy U of the system with respect to the magnetic flux ϕ through the loop. This may be obtained by differentiating with respect to a uniform vector potential A pointing around the loop, in the manner

$$I = c\frac{\partial U}{\partial \phi} = \frac{c}{L}\frac{\partial U}{\partial A} \quad . \tag{1}$$

This derivative is nonzero only by virtue of the phase coherence of the wave functions around the loop. If, for example, all the states are localized then the only effect of A is to multiply each wave function by $\exp(ieAx/\hbar c)$, where x is the coordinate around the loop, and the energy change and current are zero. If a state is extended, on the other hand, such a gauge transformation is illegal unless

$$A = n\frac{hc}{eL} \quad . \tag{2}$$

In the case on noninteracting electrons, phase coherence enables a vector potential increment to

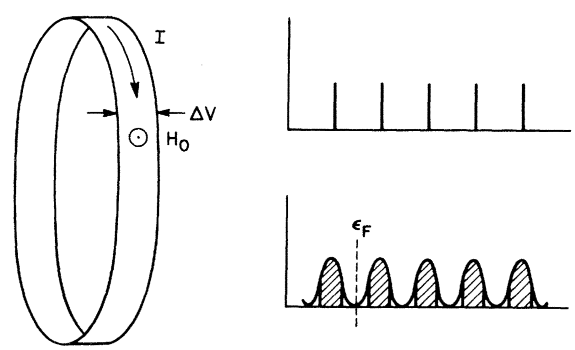

FIG. 1. Left: Diagram of metallic loop. Right: Density of states without (top) and with (bottom) disorder. Regions of delocalized states are shaded. The dashed line indicates the Fermi level.

change the total energy by forcing the filled states toward one edge of the ribbon. Specifically, if one adopts the usual isotropic effective-mass Hamiltonian,

$$H=\frac{1}{2m^*}\left(\vec{p}-\frac{e}{c}\vec{A}\right)^2+eE_0y \quad , \tag{3}$$

where E_0 is the electric field across the ribbon, and adopts Landau gauge

$$\vec{A}=H_0y\hat{x} \quad , \tag{4}$$

then the wave functions, given by

$$\psi_{k,n}=e^{ikx}\phi_n(y-y_0) \quad , \tag{5}$$

where ϕ_n is the solution to the harmonic-oscillator equation

$$\left[\frac{1}{2m^*}p_y^2+\frac{1}{2m^*}\left(\frac{eH_0}{c}\right)^2y^2\right]\phi_n=(n+\tfrac{1}{2})\hbar\omega_c\phi_n \quad , \tag{6}$$

and y_0 is related to k by

$$y_0=\frac{1}{\omega_c}\left(\frac{\hbar k}{m^*}-\frac{cE_0}{H_0}\right) \quad , \tag{7}$$

are affected by a vector potential increment $\Delta A\hat{x}$ only through the location of their centers, in the manner

$$y_0\rightarrow y_0-\frac{\Delta A}{H_0} \quad . \tag{8}$$

The energy of the state, still given by

$$\epsilon_{n,k}=(n+\tfrac{1}{2})\hbar\omega_c+eE_0y_0+\tfrac{1}{2}m^*(cE_0/H_0)^2 \tag{9}$$

thus changes linearly with ΔA. This gives rise to the derivative in Eq. (1), which may be conveniently evaluated via the substitution

$$\frac{\partial U}{\partial\phi}\rightarrow\frac{\Delta U}{\Delta\phi} \tag{10}$$

with $\Delta\phi=hc/e$ a flux quantum. Since, by gauge invariance (2), adding $\Delta\phi$ maps the system back into itself, the energy increase due to it results from the net transfer of n electrons (no spin degeneracy) from one edge to the other. The current is thus

$$I=c\frac{neV}{\Delta\phi}=\frac{ne^2V}{h}. \tag{11}$$

We now consider the dirty interacting system. As in the ideal case, gauge invariance is an exact symmetry forcing the addition of a flux quantum to result only in excitation or deexcitation of the original system. Also as in the ideal case, there is a gap, although the gap now exists between the electrons and holes affected by the perturbation, those contiguous about the loop, rather than in the density of states. Since adiabatic change of the many-body Hamiltonian cannot excite quasiparticles across this gap, it can only produce an excitation of the charge-transfer variety discussed in the ideal case, although the number of electrons transferred need not be the ideal number, and can be zero, as is the case for most systems with gaps. Therefore, Eq. (11) is always true, as a bulk property, for some integer n whenever the local Fermi level lies everywhere in a gap in the extended-state spectrum.

At the edges of the ribbon, the effective gap collapses and communication between the extended states and the local Fermi level is reestablished. Particles injected into this region rapidly thermalize to the Fermi level, in the process losing all memory of having been mapped adiabatically. This would be a significant source of error in Eq. (11) were it not for the fact that *isothermal* differentiation with respect to ϕ, the thermodynamically correct procedure for obtaining I, is equivalent to adiabatic differentiation in the sample interior and is reversible. Thus, slow addition of $\Delta\phi$ physically removes a particle from the local Fermi level at one edge of the ribbon and injects it at the local Fermi level of the other, acting as a pump. Since the Fermi energy is defined as the change in V resulting from the injection of a particle, and since eV is defined to be the Fermi-level difference, edge effects are not a source of error in Eq. (11).

Several other sources remain to be investigated, including possible ϕ dependence, the effect of substituting the ring geometry of Fig. 1 for the usual strip geometry, and effects of tunneling. However, we find it intuitively appealing that the quantum effect should go hand in hand with the persistence of currents, and thus that the physically significant source of error should be thermal activation of carriers to the mobility edge. These carriers produce a large, but finite, normal resistance per square R, which in the steady-state strip geometry, results in a Hall resistance too small in the amount

$$\left|\frac{\Delta R_H}{R_H}\right|=\left(\frac{R_H}{R}\right)^2 \quad . \tag{12}$$

In summary, we have shown that the quantum Hall effect is intimately related to the extended nature of the states near the center of the disorder-broadened Landau level, and that edge effects do not influence the accuracy of the quantization. We speculate that the only significant source of error is thermal activation of carriers to the mobility edge.

ACKNOWLEDGMENTS

I am grateful to P. A. Lee, D. C. Tsui, R. E. Prange, and H. Störmer for helpful discussions.

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

[2]Identical behavior has been seen in GaAs-$Al_xGa_{1-x}As$ heterostructures. D. C. Tsui and A. C. Gossard (unpublished).

[3]T. Ando, J. Phys. Soc. Jpn. 37, 622 (1974).

[4]R. E. Prange (unpublished).