PHYSICAL REVIEW B VOLUME 27, NUMBER 2 15 JANUARY 1983

Theory of quantized Hall effect at low temperatures

Serge Luryi and Rudolf F. Kazarinov
Bell Laboratories, Murray Hill, New Jersey 07974
(Received 30 July 1982; revised manuscript received 28 October 1982)

At millikelvin temperatures the quantized Hall effect (QHE) is characterized by sharp steps connecting the quantized Hall resistance plateaus. We explain this behavior on the basis of the single-electron approximation and continuum percolation theory. It is shown that even when the magnetic field corresponds to partially filled Landau level on average, locally the sample breaks into patches having the occupation numbers 0 or 1. This represents a peculiar type of the metal-insulator transition, driven by disorder. Extended (global) electron states may not exist at equilibrium or arbitrarily small applied voltage V. Global states begin to appear at a certain critical voltage V_{cr}, which is of the order of characteristic magnitude of potential fluctuations on the scale of the sample size. For small $V > V_{cr}$ the fraction of global states is still small owing to the smallness of the average electric field compared with the fluctuating field in the inversion layer. Because of this, transitions between the QHE plateaus require only a minor change in the density of states and at $T=0$ they occur in small intervals of the magnetic field. Owing to the tail of the Fermi-Dirac distribution at nonzero T the deviation of the Hall conductivity from its nearest quantized value, $(e^2/h)\times$integer, is activated. This explains the plateau flatness and the high precision of the QHE measurements of the fine-structure constant.

The quantized Hall effect (QHE) is a microscopic quantum phenomenon recently discovered in silicon MOSFETs[1] (metal-oxide-semiconductor field-effect transistors) and then also observed in GaAs/GaAlAs heterojunctions.[2] Inversion layers formed at the interface in these structures represent, at low temperatures, a truly two-dimensional electronic system with the electron motion along the magnetic field (i.e., perpendicular to the plane of the inversion layer) frozen out. The QHE manifests itself in certain ranges of the applied magnetic field B and electron concentration N in the inversion layer. In these ranges, called "the plateaus," the Hall conductivity G_{xy} has quantized values given by

$$G_{xy} = \bar{n}e^2/h \tag{1}$$

with $\bar{n}$ integer, while the longitudinal conductivity G_{xx} and the longitudinal resistance $R_{xx} = G_{xx}/(G_{xx}^2 + G_{xy}^2)$ nearly vanish. Existence of these wide plateaus in which Eq. (1) holds, independent of sample geometry and unaffected by disorder, allowed remarkably accurate measurements of the quantity e^2/h and the fine-structure constant. The QHE is of great theoretical interest primarily because it reveals a long-range order in the fermionic system of electrons in the inversion layer.[3] The meaning of this order is that electrons delocalized along electrostatic equipotential lines remain self-coherent over macroscopic distances. Exactness of the quantization can also be viewed[4] as a topological property of equipotentials in the plane of the inversion layer. This is especially clear in the Corbino ring geometry where all equipotentials distinctly fall into two classes: *global* which encircle a central electrode and *local* which can be contracted to a point by a continuous deformation. It is the existence of electrons delocalized along global equipotentials which embodies the long-range order in the QHE.

An outstanding theoretical problem has been to explain the observed width of the plateaus, especially at low temperatures where it was found[5,6] that the quantized plateaus begin to occupy nearly 100% or the range of the parameter (B) variation. Regions between the plateaus (in which a dissipative current can flow) shrink with temperature gradually,[7] with no apparent discontinuity.

Several theories have been advanced to explain the plateau width. These include the Baraff-Tsui mechanism[8] (pinning of the Fermi level E_F by donor impurity bands), "exchange gap"[9] (increase of the effective gap between the Landau levels due to exchange interaction), and the Fukuyama-Platzman mechanism[10] (pinning and melting of charge-density waves, existence of which was earlier suggested by Fukuyama *et al.*[11]). It is of course generally understood[12] that the existence of real localized states in response to potential fluctuations in the inversion layer can also contribute to the plateau width. It is the purpose of the present work to develop the latter point of view on the basis of a microscopic model of both localized and delocalized states. Our model assumes that in the strong magnetic field all electron states are confined within narrow tubes or "fibers,"[4] extended along equipotential lines.[13] The potential is assumed to vary smoothly [on the scale of the magnetic length $a = (\hbar c/eB)^{1/2}$], so that at every point in the sample one has a well-defined Landau-level system. Therefore, we do not need to assume the existence of states with energies between the local Landau levels. We shall show that this picture alone is sufficient to explain the remarkable results of Refs. 5 and 6.

It is generally believed that the Hall plateaus correspond to the situations when the Fermi level is pinned between two adjacent Landau levels. This already implies the fundamental importance of disorder for the QHE. Indeed, in an ideal case with a uniform electron density σ, the Fermi level is pinned to a Landau level at all B except for discrete values, $B_n = n\sigma hc/e^2$, at which the Fermi level jumps between the nth and the $(n+1)$st level. This means that in an ideal system with no disorder, the plateaus are reduced to discrete points. The finite width of the plateaus is usually attributed to pinning of the Fermi level by localized states, whose energies are continuously distributed in the gaps between Landau levels. The integer $\bar{n}$ is identified as the total number of "filled" Landau levels, i.e., those below the Fermi level. In the presence of potential fluctuations this picture, however, must be qualified since in a disorderd system there are two-dimensional patches where a particular Landau level is filled and complementary regions where this level is empty, depending on the position of the local self-consistent potential with respect to the Fermi level. There also exists a one-dimensional manifold, in general multiply connected, where the Fermi level and a particular Landau level coincide.

The spatial inhomogeneity of the self-consistent potential is brought about by fluctuations in the fixed positive charge responsible for the creation of the inversion layer in a heterojunction system. The fact that the self-consistent potential in a strong magnetic field does indeed fluctuate, thus giving rise to patches with occupation numbers 0 or 1, is rather subtle and requires further elaboration. The crux of the matter is that the surface density Σ of the fixed positive charge much exceeds the electron density in the inversion layer (in GaAs QHE samples Σ is mainly compensated by a fixed negative charge), i.e., $\Sigma \gg \sigma$. Supppose we are dealing with the lowest Landau levels, so that $\sigma \approx \sigma_0 = e/2\pi a^2$, where $\sigma_0 \equiv e^2B/hc$ is the density of states per Landau level. Consider first an ideal situation in which we create a fluctuation $\delta\Sigma$. So long as $\delta\Sigma < \sigma_0$, the potential variation is perfectly screened by the inversion layer. On the other hand, when $\delta\Sigma > \sigma_0$, then screening does not occur and the self-consistent potential in the inversion layer fluctuates. The spatial scale of the fluctuation $\delta\Sigma = \sigma_0$ is given by

$$\lambda = \frac{(\Sigma/e)^{1/2}}{\sigma_0/e} = a\left(\frac{2\pi\Sigma}{\sigma_0}\right)^{1/2} . \tag{2}$$

It is seen from (2) that $\lambda \gg a$ provided $\Sigma \gg \sigma$, which makes our model self-consistent. Thus, even at a value of B for which the average σ in the inversion layer corresponds to a partially filled Landau level, the fixed-charge fluctuations induce a peculiar metal-insulator transition. As will be shown below, the remaining metallic regions (the above mentioned one-dimensional manifold, corresponding to intersecting Landau and Fermi levels) are, generally, disconnected, except for discrete values of the magnetic field.

Let us generalize the above identification of $\bar{n}$ to this situation: $\bar{n}$ is the number of the highest Landau level whose filled patches are globally connected forming a region we shall refer to as the *sea* (using the terminology adopted in the continuum percolation theory[14,15]). Within the sea there are both local and global fiber states but as shown in Ref. 4 the entire applied Hall voltage drops on the global fibers (similarly, in the case when current is injected into the sample, the entire Hall emf develops across the global fiber subsystem). The complement to the sea region of the inversion layer represents isolated *islands* within which the Landau level under consideration is not populated. When the magnetic field B is increased, the area of these islands grows. This occurs because the Landau-level degeneracy is proportional to the magnetic field, so that when B grows electrons fall onto the lower levels and the population of the top $\bar{n}$th level decreases. As a sufficiently large B one has a qualitatively different situation in which filled patches represent *lakes* in a globally connected *continent* corresponding to the $\bar{n}$th level unfilled. It is easy to see that in both of these limits the *coastline* corresponding to the intersecting Fermi and Landau levels is not percolating. The fundamental theorem of the continuum percolation theory[14,15] requires the existence of a transition point at which the islands-in-the-sea topology goes over into that corresponding to lakes in the continent. In two dimensions the transition occurs at such position of the Fermi level at which the areas of filled and empty patches are statistically equal.

An important question is whether the coastline percolates at the transition point, i.e., forms an infinite cluster. If the answer were positive, this would imply the existence of global equipotentials in equilibrium. We shall now argue, however, that this is not the case, at least in all practically studied QHE samples (both of the Hall-bridge and the Corbino geometry).

There are two different ways in which the coastline can, in principle, percolate at the transition point. One way corresponds to an isotropic cluster. In this case neither the sea nor the continent are present. An example of such a situation is provided by any periodic potential, a map of which at the percolation point breaks into lakes and islands, like a checkerboard, while the coastline does form a global cluster passing through an infinite set of saddle points of the potential surface. However, a random potential does not generate such an isoenergetic set of saddle points.

The other possibility corresponds to the coexistence at the percolation transition of *both* the sea and the continent. This situation is actually realized for random potentials, as follows from the well-known divergence of the correlation radius at the transition point.[16] The latter means that the area of certain lakes (islands) goes to infinity as we approach the percolation transition from the continent (sea) side. It is clear that the length of the largest shoreline must also diverge in this case. In a finite-size sample this implies the existence of equipotentials connecting opposite edges of the sample. However, in this case the coastline percolation in one direction (north-south) excludes the possibility of percolation in the other (east-west) direction. It is easy to see that for an uneven sample the percolation will be necessarily established in the shortest direction. This direction does not coincide, usually, with that of the Hall current.

The above arguments show that for a randomly disordered potential in the inversion layer at equilibrium there are no global equipotentials, either filled or empty, in the direction of the Hall current. It then follows that for an arbitrarily small applied voltage the QHE will not be observed. This brings us to the concept of a critical Hall voltage[17] $V_{\rm cr}$. At voltages greater than $V_{\rm cr}$ global equipotentials begin to appear. This is easy to visualize by thinking about a funnel with crimped surface. If the height of the funnel is greater than the characteristic height of the potential fluctuations on the scale of the distance between its edges, the water-level boundary (which, of course, coincides with an equipotential) will be globally connected. In this case there exists a whole range of global equipotential lines. For the QHE the value of $V_{\rm cr}$ must be of the order of the characteristic magnitude of potential fluctuations on the scale of the sample size.

When the Hall voltage is less than critical, $V < V_{\rm cr}$, there are no global states. These states do appear when $V \gtrsim V_{\rm cr}$ but even when $V >> V_{\rm cr}$ the number of global states is exceedingly small compared to the total number of electrons. This is due to the fact that the average applied field is negligible compared to the local field of the fluctuating potential (typically their ratio is less than 10^{-6}). Indeed, consider a fluctuation of geometric dimensions λ. For an uncorrelated distribution of the fixed charge one has $\delta\Sigma/\Sigma \approx 1/\sqrt{N}$ where $N = \lambda^2\Sigma/q$. This fluctuation of charge gives rise to a field $\delta F = \delta\Sigma/\epsilon$ or

$$\delta F = \frac{q}{\lambda\epsilon}\left(\frac{\Sigma}{q}\right)^{1/2} \quad . \tag{3}$$

The shorter the wavelength of a fluctuation the larger is the fluctuating field. However, the field due to fluctuations of wavelength less than d (where d is the distance between the fixed charge and the inversion layer) averages out without reaching the inversion layer. Thus, the largest fluctuating field is given by Eq. (3) with $\lambda = d$. For the typical QHE experiments this gives δF of order 10 kV/cm or greater, while the average applied field is only 10^{-2} V/cm. This means that the local topography of the potential surface is only slightly modified by the applied field so that the area occupied by the local states changes very little compared to 100% at equilibrium. The energy range of global states may be referred to as the percolation band. We see that this band emerges already having a finite width equal to the applied voltage, inasmuch as the entire Hall voltage develops across the global states.[4] At $V < V_{\rm cr}$ there is no quantum percolation band.[18]

We have shown that global states constitute a small fraction of the total number of states in the inversion layer. On the other hand, the variation of the chemical potential μ (the quasi-Fermi level) on the global states follows exactly the variation of the electrostatic potential ψ; in other words, all global states corresponding to the same Landau level are equally populated. This can be seen as follows. At voltages greater than critical, streams of the Hall current break the inversion layer into disjoint regions. Each of these regions (labeled i) is surrounded by a local (though quite extended) equipotential ψ_i and is, therefore, in equilibrium. These regions are macroscopic, and have a well-defined chemical potential μ_i and a fluctuating electron concentration. Since the average electron concentration is given by the surface charge density σ and is the same in each macroscopic region, the quantity $\mu_i - \langle\psi\rangle_i$ (where the average is taken over the isolated region i) has the same value throughout the sample. The gist of our argument is to note that $\psi_i = \langle\psi\rangle_i$. Indeed, the boundary of a macroscopic region is a very extended local equipotential. As such it must be very close in energy to the percolation threshold $\langle\psi\rangle_i$ of region i. (A similar argument was advanced recently by Trugman and Doniach in relation to vortices in inhomogeneous superconducting films.[19]) On the other hand, the edges of each Hall stream have the same energy ψ and chemical potential μ as the adjacent local equipotentials. We thus come to the conclusion that the difference $\mu - \psi$ is the same for all global states.

We can now give a simple explanation to the observed steplike behavior of the QHE plateaus (R_{xy} vs B curves) at low temperatures. Summing over all global states, as we did in Ref. 4 but with a Fermi-Dirac distribution function $f = f(\mu - \psi)$, we can take the function f out of the sum since it is constant on global states. The result is an expression of the form (1) with $\bar{n}$ given by

$$\bar{n} = \sum_{n=1}^{\infty} f(\Delta_n) \quad , \tag{4}$$

where $\Delta_n = \mu(x) - \psi(x) - n\hbar w_c$ is constant for each Landau level n. Transitions between the Hall plateaus are due to the variation of Δ_n with the magnetic field, mainly because of the dependence $\mu(B)$. For an ideal situation with no disorder, global states constitute 100% of all states and the Fermi level is pinned at every Landau level, jumping to the next level abruptly once a particular Landau level is filled. In this situation the Hall plateaus would reduce to a set of discrete points. The observed QHE plateaus are due to the fact that in a real system the opposite limit occurs as discussed above. The Fermi level is pinned to the global states only in a small interval of B proportional to the fractional area occupied by the global equipotentials. Outside this interval the value of $\bar{n}$ given by Eq. (4) is integer to within tails of the Fermi-Dirac distribution.

Thus, at zero temperature $T=0$ the width $\delta B(0)$ of the interplateau regions is finite but small as determined by the ratio of the number of global to local states in the sample. At a nonzero T these regions grow due to the washing-out of the Fermi step function, as expressed by Eq. (4). Varying the temperature at a fixed B outside $\delta B(0)$ we are simply tracing the tail of Fermi's distribution. Therefore, the dependence $G_{xy}(T)$ at fixed B must be exponential, characterized by an activation energy $\Delta_n = E_a(B)$. Such behavior was indeed observed in the low-temperature experiments.[7] When B is well outside $\delta B(0)$, i.e., far from the center of the interplateau region, then the activation energy is large and any deviation of $\bar{n}$ from an integer becomes intangible. This explains the high precision of the QHE in determining the fine-structure constant.

Finally, let us briefly discuss the longitudinal conductivity G_{xx} describing the dissipative current. In the plateau regions the conductivity G_{xx} (as well as the resistance R_{xx}) is exponentially small at low temperatures[20] as described by Mott's law for hopping conductivity. At $T=0$ this conductivity vanishes completely. No current can flow across global states due to the absence of scattering. Although, as discussed above, the area occupied by global states is small, streams of the quantum Hall current slice the sample into disjoint regions each of which remains at equilibrium. In the interplateau regions electrons in global states can scatter and the connectivity of the sample changes, giving rise to a possible metallic conductivity at $T=0$. This brings about the bursts of dissipative current observed in the low-temperature QHE experiments between the plateaus. Detailed analysis of the dissipative current components in the interplateau regions is beyond the reach of our qualitative approach.

ACKNOWLEDGMENTS

We are deeply grateful to Mikko Paalanen for a number of helpful discussions and especially for sharing with us some of his unpublished results. We are also grateful to W. F. Brinkman, D. S. Fisher, and T. M. Rice for helpful comments.

---

[1]K. von Klitzing, G. Dorda, and M. Pepper, Phys. Rev. Lett. 45, 494 (1980).
[2]D. C. Tsui and A. C. Gossard, Appl. Phys. Lett. 38, 550 (1981).
[3]R. B. Laughlin, Phys. Rev. B 23, 5632 (1981).
[4]R. F. Kazarinov and S. Luryi, Phys. Rev. B 25, 7626 (1982).
[5]M. A. Paalanen, D. C. Tsui, and A. C. Gossard, Phys. Rev. B 25, 5566 (1982).
[6]G. Ebert, K. von Klitzing, C. Probst, and K. Ploog (unpublished).
[7]M. A. Paalanen (private communication).
[8]G. Baraff and D. C. Tsui, Phys. Rev. B 24, 2274 (1982).
[9]P. W. Anderson (private communication).
[10]H. Fukuyama and P. M. Platzman, Phys. Rev. B 25, 2934 (1982).
[11]H. Fukuyama, P. M. Platzman, and P. W. Anderson, Phys. Rev. B 19, 5211 (1979).
[12]D. C. Tsui and S. J. Allen, Jr., Phys. Rev. B 24, 4082 (1981).
[13]Strictly speaking, these fibers are extended along the lines of constant classical energy. The latter includes also the kinetic energy $mv_H^2/2$ which is associated with the Hall velocity $v_H = cF/B$ and hence depends on the local electric field F, cf. Ref. 4. Throughout this work, speaking of the equipotentials, this comment is left understood.
[14]R. Zallen and H. Scher, Phys. Rev. B 4, 4471 (1971).
[15]B. I. Shklovsky and A. L. Efros, *Electronic Properties of Doped Semiconductors* (Nauka, Moscow, 1979), Chap. 5, in Russian.
[16]D. Stauffer, Phys. Rep. 54, 1 (1979).
[17]R. F. Kazarinov and S. Luryi, in Proceedings of the 1982 International Electron Devices Meeting, San Francisco, California, December 1982 (unpublished).
[18]This point of view clearly contradicts that of S. V. Iordansky (unpublished) who attempted to explain the QHE by postulating an infinitely narrow percolation band, which exists even at equilibrium.
[19]S. A. Trugman and S. Doniach (unpublished). We are indebted to D. S. Fisher for bringing this work to our attention.
[20]D. C. Tsui, H. L. Störmer, and A. C. Gossard, Phys. Rev. B 25, 1405 (1982).