Theory of the breakdown of the quantum Hall effect

V. Tsemekhman, K. Tsemekhman, C. Wexler, J. H. Han, and D. J. Thouless

Department of Physics, University of Washington, Box 351560, Seattle, Washington 98195

(Received 22 November 1996; revised manuscript received 3 February 1997)

Breakdown of the quantum Hall effect at high values of injected current is explained as a consequence of an abrupt formation of a metallic conduction path from one edge of the sample to the other. Such a path is formed when regions of compressible liquid, where the long-range disorder potential is screened, get connected due to the strong electric field. Our theory explains various features of the breakdown and numerical simulations based on the theory yield critical currents consistent with experiment. [S0163-1829(97)50420-6]

In this paper we discuss the breakdown of the quantum Hall effect (QHE) at high injected currents. Observed as a sudden onset of dissipation when the injected current exceeds some critical value, the breakdown of the integer QHE has been extensively investigated since then. However, the phenomenon has not been given a satisfactory explanation. Existing theories of this effect include production of hot electrons, inter-Landau level transitions in the high local electric field (due to tunneling or emission of phonons), increase in the number of the delocalized states in the Landau level. These theories either produce critical current orders of magnitude higher than observed experimentally or need some additional assumptions to be consistent with experiment. Furthermore, none of them can explain a rich variety of phenomena associated with the breakdown.

Experiments on the breakdown of the QHE are performed on GaAs heterostructures or on high-quality metal-oxide-semiconductor field-effect transistor (MOSFET) devices, both characterized by the presence of long-range fluctuations of the disorder potential. In GaAs systems the disorder potential due to the remote dopants has predominantly long wavelength fluctuations (λ > d > l_H), where λ is the typical wavelength of the fluctuations, d is the spacer thickness, and l_H is the magnetic length. Long-range potential fluctuations are also present in high-mobility MOSFET devices though their origin is not evident. In such systems, the QHE for noninteracting electrons can be understood in terms of the percolation theory. The plane of the two-dimensional electron gas (2DEG) is divided into areas of completely filled and empty LL’s. The number of completely filled LL’s in the percolating region determines the value of the Hall conductance.

Coulomb interaction between electrons leads to the screening of the long-range potential fluctuations, changing this picture. As the kinetic energy is frozen out in the strong magnetic field, screening is perfect in some areas of the system where the LL is partially occupied and is absent in the rest of the system where LL’s are either completely filled or empty. According to Refs. 11 and 12, only in the narrow range of the filling factors around integers does incompressible region percolate through the sample, leading to the QHE. While this statement contradicts experimental observations of the wide QH plateaux at low temperatures, it is justified for temperatures T ~ 1 K, which are high compared to the energy scales of the short wavelength fluctuations of the disorder potential (λ < l_H) that are left unscreened and of the residual interelectron interaction. As most experiments on the breakdown of the QHE are performed at temperatures of this range, one should expect that when the magnetic field and the density of the 2DEG correspond to a filling factor close to an integer number, an incompressible region percolates through the sample, leading to the QHE. In the percolating incompressible region electric fields are about the same as those of the bare disorder potential

E_inc = \sqrt{E_{bare}} = \sqrt{n_0/32\pi(e/\varepsilon_0 d)},

where n_0 is the density of the 2DEG. For a typical clean sample this leads to E_inc = 0.1 \hbar \omega_c / e l_H, where \omega_c is the cyclotron frequency. There are, however, isolated compressible areas of perfectly screened disorder potential that behave like metallic liquids. In each of the regions, all states of the highest available Landau level are partially occupied, and the screened potential fluctuates around the Fermi level \varepsilon_F with an amplitude of the order of k_B T, which is much smaller than the amplitude of the potential fluctuations in the incompressible region. Below we consider the potential in the compressible regions to be flat. The main idea of our theory is that at high enough currents the insulating region separating two adjacent metallic regions suddenly breaks down due to the high electric field leading to the connection of the regions. When such connected regions form a metallic path percolating from one edge of the sample to the other, abrupt onset of the dissipative regime is observed. According to our model, breakdown of the QHE is totally different from the percolation transition. In complete analogy with the dielectric breakdown of the insulating media, breakdown of the QHE should occur at any concentration of the compressible liquid in the system, however small it is. Therefore the predictions of the percolation theory cannot be applied to this phenomenon.

In general, incompressible region separating two adjacent regions can be wide, and the potential in it may have a complicated structure. It is clear from physical intuition, and confirmed by our numerical simulations, that the regions closest to each other get connected first. Let us, therefore, consider a simple model in which two metallic regions are separated by a narrow insulating region with one parabolic potential fluctuation characterized by the root-mean-square electric field E = E_inc = 0.1 \hbar \omega_c / e l_H. We will now estimate the value of the external electric field that brings these lakes together. In equilibrium, all electrons in the regions have energies equal to \varepsilon_F (up to T), and all...
electrons in the incompressible region separating them have energies below the Fermi energy. The electric field is pushed out into the incompressible region whose width is much smaller than the average electric field. Therefore, an average electric field much smaller than $E_0$ is needed to connect the regions. When the two regions merge, the electric field is expelled from the former incompressible region now covered with compressible liquid. Therefore, the electric field becomes even stronger outside the newly formed larger region, facilitating further connections as the injected current is increased. After the metallic region acquires some critical length in the direction of the Hall electric field upon the increase of the injected current, the process of further connection must take an avalanche form. Finally, when the compressible liquid forms a path flowing from one edge to the other, dissipation drastically increases. It is important to notice that just before the formation of the metallic path, the system is still in the regime of exponentially small activated longitudinal conductivity; there are many regions on the way of the future path that are not yet connected, and the distance between them is as large as in equilibrium. This is the reason why the longitudinal conductivity jumps several orders of magnitude at the critical current: the mechanism of the conductivity abruptly switches from the activated to the metallic one.

To see how the breakdown occurs in real systems we performed numerical simulations (details will be published elsewhere\cite{14}). Positively charged ions and negatively charged electrons were assumed to occupy parallel plates of a capacitor separated by the spacer with thickness $d$. To avoid edge effects these plates, each $2.5\,\mu m \times 2.5\,\mu m$ in size, were modeled to be cut from the equally and uniformly charged infinite planes. To simulate the disorder, positive ions were placed at random on one of these plates with the same average density as in the rest of the plane. Charge distribution on the negatively charged plate was calculated self-consistently. Each plate was divided into a grid of elementary squares with the side of 250 Å. In equilibrium, the free energy of the interacting 2DEG in the strong magnetic field and in the smooth disorder potential (up to an overall factor)

$$F\left(\{n_i\}\right) = \sum_{N=0,1} \sum_i \left\{ k_B T \left( (1-n_i^{(N)}) \log(1-n_i^{(N)}) + n_i^{(N)} \log(n_i^{(N)}) \right) + \frac{s}{\pi \hbar e} \sum_{i,j \in \epsilon} \frac{n_i n_j}{|\mathbf{r}_i - \mathbf{r}_j|} \right\}$$

was minimized with respect to the occupation numbers in each square at finite temperature. In Eq. (1) $n_i^{(N)}$ is the occupation number of the $N$th Landau level in the $i$th elementary square; $n_i = n_i^{(0)} + n_i^{(1)}$, $r_i$ is a radius vector for the $i$th square; $s$ is the area of the elementary square, and $\epsilon$ is the dielectric constant. The external potential $U_j$ is a sum of the random disorder potential and the constant potential of an infinite capacitor fixing the chemical potential of the system. Coulomb interaction between electrons was taken into account in the mean-field approximation. The first two Landau levels of the spin-degenerate system were included in our calculations. Figure 2(a) shows the charge distribution for a clean sample at $\nu = 1.93$. Regions of compressible liquid ($\nu < 2$, gray areas) are immersed into the percolating incompressible region ($\nu = 2$, white areas). Injection of the Hall current leads to the appearance of the external electric field. If the relaxation processes are fast enough, the external field...
averaged over a distance much larger that the typical size of the lakes is constant across the sample and is equal to the average Hall electric field $E_{av}$. On a smaller length scale the electric field may, of course, fluctuate. Therefore, in our calculations the external electric field is taken to be uniform before the response of the charges in the system. Relaxation current is allowed to flow only in the compressible regions and through the boundary into the lakes, i.e., only into the areas where there is empty space for the electrons. The charge density of the same system with external field of 200 V/cm is shown on Fig. 2(b), showing no sign of breakdown. At 500 V/cm [Fig. 2(c)], several regions have merged into compressible channels parallel to the external electric field. Evidently, this field is already higher than the critical one and dissipation becomes important. These values are in good agreement with experiments. 1–4

Tables I and II show results of numerical simulations for the dependence of the critical field on the filling factor and temperature. The wide range of values for the same size sample corresponds to the fluctuations of the breakdown field in small systems due to different realizations of disorder. When the filling factor deviates from an integer, the size of the regions grows and the separation between them decreases. This causes stronger enhancement of the electric field in the incompressible region. Therefore, the critical electric field is largest at the center of the plateau and decreases when the filling factor is moved away from the integer. 1 The same growth of the metallic regions with increasing temperature leads to lower critical fields at higher temperatures. This dependence becomes noticeable, however, only at rather high temperatures, and the critical fields at temperatures in the range 0.5–1 K are basically the same.

Observed localization in space 2 of the region where the dissipation sets in naturally follows from our theory. Only if the probes measuring the longitudinal voltage appear on different banks of the path will it be possible to observe the breakdown. As the metallic path is very narrow at the breakdown, situations when the probes are on the same side of the path and when they are separated by the metallic region are both likely. Connection of the regions at fields below the critical one causes charge redistribution and energy relaxation. The amount of dissipated energy and the time scale of the relaxation process vary with the size of the connecting regions and the distance between them. This energy relaxation shows up as a broadband noise. 2

The observed transient switching 2 just before the breakdown and switching between different voltage levels after the breakdown is the result of the formation and immediate disconnection of the path at the fields just below critical or switching between metastable states with different number of metallic paths, respectively. This happens, since after the path is formed, relaxation is available all over the sample width. The system can then relax to a new state with incompressible region still percolating, or to the state with other paths formed. Our numerical simulations show such connection/disconnection processes.

As in any dielectric breakdown, hysteresis 3 is the result of the irreversible change of the system properties (namely the pattern of paths and regions) after the current was allowed to flow through the system. Finally, the peculiar steps observed in the magnetic field dependence of the longitudinal voltage in the critical current regime 5 show the evidence of the opening of new metallic channels when the system is moved away from the center of the plateau. As mentioned above, in equilibrium, when the filling factor is shifted from integer, the size of the regions of compressible liquid increases and their separation becomes smaller. Therefore, in the critical regime the number of percolating narrow metallic channels also grows, each path giving its own discrete contribution in the voltage drop. Further study is needed to see if the individual contributions in the longitudinal voltage from these paths are quantized.

We would like to emphasize that the nature of the phenomenon does not depend on the system size. However, the value of the breakdown electric field is not a universal quantity and is determined by the system size and by the disorder configuration. Any large system can be divided into smaller pieces such that each of them still breaks down through the avalanche process. Once an avalanche starts in any one of them it spreads across the system. Thus the breakdown field

<table>
<thead>
<tr>
<th>$\nu$</th>
<th>$E_{cr}$ (V/cm)</th>
<th>LL#</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.08</td>
<td>330 – 360</td>
<td>2</td>
</tr>
<tr>
<td>2.04</td>
<td>400 – 500</td>
<td>1.2</td>
</tr>
<tr>
<td>2.00</td>
<td>350 – 400</td>
<td>1</td>
</tr>
<tr>
<td>1.96</td>
<td>350 – 400</td>
<td>1</td>
</tr>
<tr>
<td>1.92</td>
<td>200 – 300</td>
<td>1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$T$ (K)</th>
<th>$E_{cr}$ (V/cm)</th>
<th>LL#</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.0</td>
<td>350 – 400</td>
<td>1</td>
</tr>
<tr>
<td>2.0</td>
<td>200 – 250</td>
<td>1</td>
</tr>
<tr>
<td>3.0</td>
<td>100 – 200</td>
<td>1</td>
</tr>
<tr>
<td>3.8</td>
<td>~ 120</td>
<td>1</td>
</tr>
</tbody>
</table>
in the larger system is determined by the weakest piece, and not by the average over all pieces comprising the sample. The larger the system the higher is the probability to find a piece with the lower critical electric field and, hence, the smaller is, on average, the breakdown field for the whole sample.\(^1\) Therefore, the numbers from Table I give the upper bounds for the critical field in the macroscopic samples. This, as well as possible deviations of the external electric field from the uniform one explain the fact that our values of the critical field are about two times higher than the experimental results. Quantitative conclusions about the breakdown electric field dependence on the size, filling factor, temperature, etc., as well as about its fluctuations require averaging over a large number of samples. This work is currently under way.

In conclusion, we propose a theory of the breakdown of the QHE based on the existence of the compressible regions in the inhomogeneous 2DEG. Calculations based on our theory produce values of the critical fields that are in good agreement with the results of measurements. We are also able to explain the observed dependence of the critical fields on the filling factor and temperature, as well as qualitatively explain some of the rich features observed in the experiments.

This work was supported by the NSF Grant No. DMR-9528345.

---

4 \textit{The Quantum Hall Effect}, edited by R. E. Prange and S. M. Girvin (Springer-Verlag, New York, 1990), and references therein.