RAPID COMMUNICATIONS

PHYSICAL REVIEW B VOLUME 30, NUMBER 2 15 JULY 1984

Gauge invariance and fractional quantum Hall effect

R. Tao and Yong-Shi Wu
Department of Physics FM-15, University of Washington, Seattle, Washington 98195
(Received 27 April 1984)

It is shown that gauge-invariance arguments imply the possibility of the fractional quantum Hall effect; the Hall conductance is accurately quantized to a rational value. The ground state of a system showing the fractional quantum Hall effect must be degenerate; the nondegenerate ground state can only produce the integral quantum Hall effect.

The discovery of the fractional quantum Hall effect in a two-dimensional electron gas in a strong magnetic field by Tsui, Störmer, and Gossard[1,2] has prompted a series of interesting theoretical investigations.[3–7] One important and controversial problem is whether the ground state of the fractional quantum Hall system is degenerate. In our many-body theory[4,5] for this effect, the degeneracy is explicit. Anderson[8] also suggests that Laughlin's wave function may have a broken symmetry. On the other hand, Laughlin[4] and Haldane[7] claim that the ground state is nondegenerate.

In a noteworthy paper of 1981, Laughlin[9] showed that integral quantization of the Hall conductance is a consequence of gauge invariance. Can gauge-invariance arguments also imply fractional quantization of Hall conductance? And why are the Hall plateaus at the filling factor $\nu=\frac{1}{3},\frac{2}{3},\frac{2}{5},\ldots$ accurately quantized to rational values,[2] e.g., at $\nu=\frac{1}{3}$ to better than 10^{-4}?

This Rapid Communication presents some answers to the above questions. Gauge-invariance arguments also imply the possibility of fractional quantization of Hall conductance; the Hall conductance is accurately quantized to a rational value. The ground state of a system showing the fractional quantum Hall effect must be degenerate; the nondegenerate ground state can only produce the integral quantum Hall effect. The presence of an energy gap is a necessary but, perhaps, not a sufficient condition for this effect. Our gauge-invariance arguments in this Rapid Communication do not tell which value of the filling factor is more stable and have not explained the odd denominator rule observed in the experiments,[2] but they suggest another possible explanation for this rule: the ground state of the Hall system at a filling factor with an even denominator has a special topological property in Hilbert space.

We consider the geometry proposed by Laughlin:[9] a ribbon of two-dimensional system bent into a loop of circumference L, and pierced everywhere by a strong magnetic field $\vec{B}$ normal to its surface (Fig. 1). We also put a small solenoid at the center of the loop. Initially, the solenoid is not turned on. The radius of the loop is big enough so that the surface of the loop can be considered as a plane. In order to make our system capable of producing Hall current, we assume that electrons can be fed in at one edge and taken away from the other, but we only consider the electrons on the surface. The ground state, Ψ_0, of the two-dimensional electron gas on the surface satisfies

$$H(\vec{p}_1-e\vec{A}_1,\ldots,\vec{p}_N-e\vec{A}_N)\Psi_0=E_0\Psi_0\ , \tag{1}$$

where $\vec{A}$ is the vector potential of the strong magnetic field $\vec{B}$ in a particular gauge; E_0 is the ground state energy; we set $c=1$. The dependence of the Hamiltonian, H, on $\vec{r}_1,\ldots,\vec{r}_N$ is not explicitly written in the formula. We also assume that an energy gap separates the ground state from the excited states. In the geometry considered here Ψ_0 is periodic in the y direction, with period L,

$$\Psi_0(\vec{r}_1,\ldots,\vec{r}_j+L\hat{y},\ldots)=\Psi_0(\vec{r}_1,\ldots,\vec{r}_j,\ldots)\ ,\quad j=1,2,\ldots,N\ . \tag{2}$$

Now we switch the solenoid on and adiabatically increase the magnetic flux of the solenoid from zero to an arbitrary value ϕ. During this process some electrons can be transferred from one edge to the other. Because of the energy gap, the system remains in a ground state which may be different from the original one. The new wave function Ψ is given by

$$H(\vec{p}_1-e(\vec{A}_1+\vec{a}),\ldots,\vec{p}_N-e(\vec{A}_N+\vec{a}))\Psi=E'\Psi\ , \tag{3}$$

where $\vec{a}=a\hat{y}$ is the vector potential of the solenoid at the surface of the loop; E' may depend on a. Let

$$\Psi(\vec{r}_1,\vec{r}_2,\ldots,\vec{r}_N)=\exp\left[i(ea/\hbar)\sum_{j=1}^{N}y_j\right]\times u(\vec{r}_1,\vec{r}_2,\ldots,\vec{r}_N)\ , \tag{4}$$

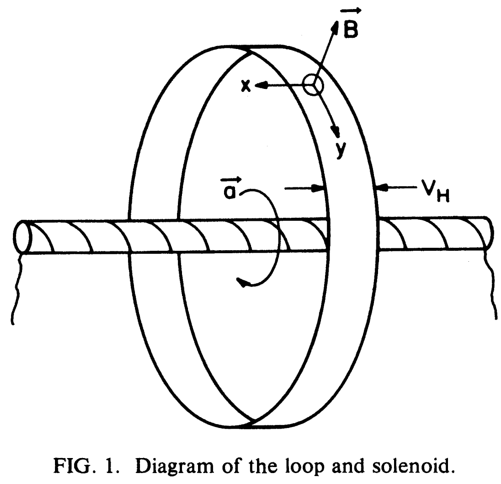

FIG. 1. Diagram of the loop and solenoid.

so that

$$H(\vec{p}_1 - e\vec{A}_1, \ldots, \vec{p}_N - e\vec{A}_N)u = E'u \quad . \tag{5}$$

Since the wave function Ψ is periodic in the y direction, we have, from Eq. (4),

$$u(\vec{r}_1, \ldots, \vec{r}_j + L\hat{y}, \ldots) = \exp(-ie\phi/\hbar)u(\vec{r}_1, \ldots, \vec{r}_j, \ldots) \ , \quad j = 1, \ldots, N \ , \tag{6}$$

where $\phi = aL$ is the magnetic flux of the solenoid. If $\phi \neq n\phi_0$ ($\phi_0 = h/e$, n is an integer), u cannot be periodic in the y direction, and u is different from Ψ_0.

Now let us consider the case $\phi = \phi_0$. From Eq. (6), u is also periodic in the y direction. By gauge invariance[10] we must also have $E' = E_0$. This can be proved directly. Because u is also an eigenstate of Hamiltonian $H(\vec{p}_1 - e\vec{A}_1, \ldots, \vec{p}_N - e\vec{A}_N)$ with the lowest eigenvalue and satisfies the same boundary conditions as Ψ_0, u and Ψ_0 must have the same eigenvalue, i.e., $E_0 = E'$. But this does not mean that u and Ψ_0 must be the same. We have two possibilities to consider: either the ground state of the system is nondegenerate or degenerate.

If the ground state is nondegenerate, u is the same as Ψ_0; by gauge invariance, the system simply maps back into the initial state. The net physical result is that N_0 electrons are transferred from one edge to the other. Then as Laughlin[9] showed, the energy increase due to this transfer is

$$\Delta U = N_0 e V_H \quad , \tag{7}$$

where V_H is the potential drop from one edge to another. The Hall current is

$$I_H = \partial U/\partial\phi = \Delta U/\phi_0 = V_H N_0 e^2/h \tag{8}$$

and the Hall conductance is

$$\sigma_H = I_H/V_H = N_0 e^2/h \quad . \tag{9}$$

Clearly only integral Hall conductance is produced.

If the ground state is degenerate, at $\phi = \phi_0$, u can be different from Ψ_0 though, by gauge invariance, they both are ground states of $H(\vec{p}_1 - e\vec{A}_1, \ldots, \vec{p}_N - e\vec{A}_N)$. Therefore, after the magnetic flux of the solenoid changes from zero to ϕ_0, the system may still not map back into the initial state. We then increase the magnetic flux of the solenoid ϕ to $2\phi_0, 3\phi_0, \ldots$. If the system maps back to Ψ_0 at a finite value of ϕ, say, $p\phi_0$, then we have fractional quantization of Hall conductance. Suppose q electrons in total are transferred from one edge to the other. The same argument as above yields

$$\sigma_H = (q/p)e^2/h \quad . \tag{10}$$

The fractional Hall conductance is now accurately quantized to a rational value as in the integral case. The above discussion clearly demonstrates that if a system shows the fractional quantum Hall effect, the ground state of that system must be degenerate. On the other hand, if the system can never come back to its initial state as the magnetic flux ϕ increases, there is no quantization of Hall conductance even though there is an energy gap. For example, if the ground state of a Hall system has an infinite degeneracy, the system may never map back to its initial state.

Gauge-invariance arguments can also relate σ_H to the filling factor. Their relationship can be obtained by considering the angular momentum.[11] The charge carriers here are electrons (or holes). Our argument does not tell which value of filling factor is more stable. This should be determined by calculation of energy gaps. The above discussion does not explain the odd denominator rule observed in the experiments. Usually the guess is that this odd denominator rule is due to absence of an energy gap at fillings with an even denominator. Our gauge-invariance arguments suggest another possibility: the ground state of the Hall system at these fillings has a special topological property in Hilbert space such that it can never come back to its initial state as the magnetic flux of the solenoid increases. We consider this possibility to be more interesting since some recent numerical calculations show the possible presence of energy gaps at filling factors with even denominator.[12] We speculate there is a more profound reason for this rule.

One of us (R.T.) wishes to thank Professor D. J. Thouless for helpful discussions. This work was supported by the National Science Foundation under Grant No. DMR-83-19301 and by the U.S. Department of Energy under Contract No. DE-AC06-81ER-40048.

[1]D. C. Tsui, H. L. Störmer, and A. C. Gossard, Phys. Rev. Lett. **48**, 1559 (1982).
[2]H. L. Störmer, A. Chang, D. C. Tsui, J. C. M. Hwang, and W. Wiegmann, Phys. Rev. Lett. **50**, 1953 (1983).
[3]R. B. Laughlin, Phys. Rev. Lett. **50**, 1395 (1983).
[4]R. Tao and D. J. Thouless, Phys. Rev. B **28**, 1142 (1983).
[5]R. Tao, Phys. Rev. B **29**, 636 (1984).
[6]B. I. Halperin, Phys. Rev. Lett. **52**, 1583 (1984).
[7]F. D. M. Haldane, Phys. Rev. Lett. **51**, 605 (1983).
[8]P. W. Anderson, Phys. Rev. B **28**, 2264 (1983).
[9]R. B. Laughlin, Phys. Rev. B **23**, 5632 (1981).
[10]T. T. Wu and C. N. Yang, Phys. Rev. D **12**, 3845 (1975).
[11]R. Tao (unpublished).
[12]See, for example, D. Yoshioka, Phys. Rev. B (to be published).