Quantum Spin Hall Effect

B. Andrei Bernevig and Shou-Cheng Zhang
Department of Physics, Stanford University, Stanford, California 94305, USA
(Received 18 April 2005; published 14 March 2006)

The quantum Hall liquid is a novel state of matter with profound emergent properties such as fractional charge and statistics. The existence of the quantum Hall effect requires breaking of the time reversal symmetry caused by an external magnetic field. In this work, we predict a quantized spin Hall effect in the absence of any magnetic field, where the intrinsic spin Hall conductance is quantized in units of $2\frac{e}{4\pi}$. The degenerate quantum Landau levels are created by the spin-orbit coupling in conventional semiconductors in the presence of a strain gradient. This new state of matter has many profound correlated properties described by a topological field theory.

DOI: 10.1103/PhysRevLett.96.106802 PACS numbers: 73.43.−f, 72.25.Dc, 72.25.Hg, 85.75.−d

Recently, the intrinsic spin Hall effect has been theoretically predicted for semiconductors with spin-orbit coupled band structures [1,2]. The spin Hall current is induced by the external electric field according to the equation

$$j_j^i = \sigma_s \epsilon_{ijk} E_k \quad (1)$$

where j_j^i is the spin current of the ith component of the spin along the direction j, E_k is the electric field, and ϵ_{ijk} is the totally antisymmetric tensor in three dimensions. The spin Hall effect has recently been detected in two different experiments [3,4], and there is strong indication that at least one of them is in the intrinsic regime [5]. Because both the electric field and the spin current are even under time reversal, the spin current could be dissipationless, and the value of σ_s could be independent of the scattering rates. This is in sharp contrast with the extrinsic spin Hall effect, where the effect arises only from the Mott scattering from the impurity atoms [6].

The independence of the intrinsic spin Hall conductance σ_s on the impurity scattering rate naturally raises the question whether it can be quantized under certain conditions, similar to the quantized charge Hall effect. We start off our analysis with a question: Can we have Landau Level (LL) -like behavior *in the absence* of a magnetic field [7]? The quantum Landau levels arise physically from a *velocity dependent force*, namely, the Lorentz force, which contributes a term proportional to $\vec{A} \cdot \vec{p}$ in the Hamiltonian. Here $\vec{p}$ is the particle momentum and $\vec{A}$ is the vector potential, which in the symmetric gauge is given by $\vec{A} = \frac{B}{2}(y, -x, 0)$. In this case, the velocity dependent term in the Hamiltonian is proportional to $B(xp_y - yp_x)$.

In condensed matter systems, the only other ubiquitous velocity dependent force besides the Lorentz force is the spin-orbit coupling force, which contributes a term proportional to $(\vec{p} \times \vec{E}) \cdot \vec{\sigma}$ in the Hamiltonian. Here $\vec{E}$ is the electric field, and $\vec{\sigma}$ is the Pauli spin matrix. Unlike the magnetic field, the presence of an electric field does not break the time reversal symmetry. If we consider the particle momentum confined in a two-dimensional geometry, say the xy plane, and the electric field direction confined in the xy plane as well, only the z component of the spin enters the Hamiltonian. Furthermore, if the electric field $\vec{E}$ is not constant but is proportional to the radial coordinate $\vec{r}$, as it would be, for example, in the interior of a uniformly charged cylinder $\vec{E} \sim E(x, y, 0)$, then the spin-orbit coupling term in the Hamiltonian takes the form $E\sigma_z(xp_y - yp_x)$. We see that this system behaves in such a way as if particles with opposite spins experience the opposite "effective" orbital magnetic fields, and a Landau level structure should appear for each spin orientations.

However, such an electric field configuration is not easy to realize. Fortunately, the scenario previously described is realizable in zinc-blende semiconductors such as GaAs, where the shear strain gradients can play a similar role. Zinc-blende semiconductors have the point-group symmetry T_d which is half of the cubic-symmetry group O_h, and does not contain inversion as one of its symmetries. Under the T_d point group, the cubic harmonics xyz transform like the identity, and off-diagonal symmetric tensors ($xy + yx$, etc.) transform in the same way as vectors on the other direction (z, etc.), and represent basis functions for the T_1 representation of the group. Specifically, strain is a symmetric tensor $\epsilon_{ij} = \epsilon_{ji}$, and its off-diagonal (shear) components are, for the purpose of writing down a spin-orbit coupling Hamiltonian, equivalent to an electric field in the remaining direction:

$$\epsilon_{xy} \leftrightarrow E_z; \qquad \epsilon_{xz} \leftrightarrow E_y; \qquad \epsilon_{yz} \leftrightarrow E_x. \quad (2)$$

The Hamiltonian for the conduction band of bulk zinc-blende semiconductors under strain is hence the analogous to the spin-orbit coupling term $(\vec{v} \times \vec{E}) \cdot \vec{\sigma}$. In addition, we have the usual kinetic p^2 term and a trace of the strain term $tr\epsilon = \epsilon_{xx} + \epsilon_{yy} + \epsilon_{zz}$, both of which transform as the identity under T_d:

$$H = \frac{p^2}{2m} + Btr\epsilon + \frac{1}{2}\frac{C_3}{\hbar}[(\epsilon_{xy}p_y - \epsilon_{xz}p_z)\sigma_x + (\epsilon_{zy}p_z - \epsilon_{xy}p_x)\sigma_y + (\epsilon_{zx}p_x - \epsilon_{yz}p_y)\sigma_z]. \quad (3)$$

The constant $\frac{C_3}{\hbar}$ is 8×10^5 m/s for GaAs [8] and 1.8×10^6 m/s for InSb [9]. This Hamiltonian is not new and was previously written down in Refs. [10–13] but the analogy with the electric field and its derivation from a Lorentz force is suggestive enough to warrant repetition. The mechanism for generation of spin-orbit coupling in the conduction band is hybridization between the p-valence band (where spin-orbit coupling is very large, comparable to the kinetic energy) and the n band. As such, the internal, $\vec{k}$-dependent magnetic field due to spin-orbit coupling is uniform and sizable in the bulk of the sample, as reflected in spin-drag experiments which measure the spin precession around the spin-orbit coupling field [14].

Let us now presume a strain configuration in which $\epsilon_{xy} = 0$ but ϵ_{xz} has a constant gradient along the y direction while ϵ_{yz} has a constant gradient along the x direction. This case then mimics the situation of the electric field inside a uniformly charged cylinder discussed above, as $\epsilon_{xz}(\leftrightarrow E_y) = gy$ and $\epsilon_{yz}(\leftrightarrow E_x) = gx$, g being the magnitude of the strain gradient. With this strain configuration and in a symmetric quantum well in the xy plane, which we approximate as being parabolic, the above Hamiltonian becomes

$$H = \frac{p_x^2 + p_y^2}{2m} + \frac{1}{2}\frac{C_3}{\hbar} g(yp_x - xp_y)\sigma_z + D(x^2 + y^2). \quad (4)$$

We first solve this Hamiltonian and come back to the experimental realization of the strain architecture in the later stages of the Letter. We make the coordinate change $x \to (2mD)^{-1/4}x$, $y \to (2mD)^{-1/4}y$, and $R = \frac{1}{2}\frac{C_3}{\hbar}\sqrt{\frac{2m}{D}}g$. $R = 2$ or $D = D_0 \equiv \frac{2mgC_3^2}{16\hbar^2}$ is a special point, where the Hamiltonian can be written as a complete square, namely $H = \frac{1}{2m}(\vec{p} - e\vec{A}\sigma_z)^2$ with $\vec{A} = \frac{mC_3 g}{2\hbar e}(y, -x, 0)$. At this point, our Hamiltonian is exactly equivalent to the usual Hamiltonian of a charged particle in an uniform magnetic field, where the two different spin directions experience the opposite directions of the effective magnetic field. Any generic confining potential $V(x, y)$ can be written as $D_0(x^2 + y^2) + \Delta V(x, y)$, where the first term completes the square for the Hamiltonian, and the second term $\Delta V(x, y) = V(x, y) - D_0(x^2 + y^2)$ describes the additional static potential within the Landau levels. Since $[H, \sigma_z] = 0$ we can use the spin on the z direction to characterize the states. In the new coordinates, the Hamiltonian takes the form:

$$H = \begin{pmatrix} H_\uparrow & 0 \\ 0 & H_\downarrow \end{pmatrix}, \quad H_{\downarrow\uparrow} = \sqrt{\frac{D}{2m}}[p_x^2 + p_y^2 + x^2 + y^2 \pm R(xp_y - yp_x)]. \quad (5)$$

The $H_{\downarrow\uparrow}$ is the Hamiltonian for the spin-$\downarrow$ and spin-$\uparrow$ σ_z respectively. Working in complex-coordinate formalism and choosing $z = x + iy$ we obtain two sets of raising and lowering operators:

$$a = \partial_{z^*} + \frac{z}{2}, \qquad a^\dagger = -\partial_z + \frac{z^*}{2}$$
$$b = \partial_z + \frac{z^*}{2}, \qquad b^\dagger = -\partial_{z^*} + \frac{z}{2} \quad (6)$$

in terms of which the Hamiltonian decouples:

$$H_{\downarrow\uparrow} = 2\sqrt{\frac{D}{2m}}\left[\left(1 \mp \frac{R}{2}\right)aa^\dagger + \left(1 \pm \frac{R}{2}\right)bb^\dagger + 1\right]. \quad (7)$$

The eigenstates of this system are harmonic oscillators $|m, n\rangle = (a^\dagger)^m(b^\dagger)^n|0, 0\rangle$ of energy $E_{m,n}^{\downarrow\uparrow} = \frac{1}{2}\sqrt{\frac{D}{2m}} \times [(1 \mp \frac{R}{2})m + (1 \pm \frac{R}{2})n + 1]$. We shall focus on the case of $R = 2$ where there is no additional static potential within the Landau level.

For the spin-$\uparrow$ electron, the vicinity of $R \approx 2$ is characterized by the Hamiltonian $H_\uparrow = \frac{1}{2}\frac{C_3}{\hbar}g(2aa^\dagger + 1)$ with the lowest Landau-level (LLL) wave function $\phi_n^\uparrow(z) = \frac{z^n}{\sqrt{\pi n!}} \times \exp(\frac{-zz^*}{2})$. a is the operator moving between different Landau levels, while b is the operator moving between different degenerate angular momentum states within the same LL: $L_z = bb^\dagger - aa^\dagger$, $L_z\phi_n^\uparrow(z) = n\phi_n^\uparrow(z)$. The wave function, besides the confining factor, is holomorphic in z, as expected. These spin-$\uparrow$ electrons are the chiral, and their charge conductance is quantized in units of e^2/h.

For the spin-$\downarrow$ electron, the situation is exactly the opposite. The vicinity of $R \approx 2$ is characterized by the Hamiltonian $H_\downarrow = \frac{1}{2}\frac{C_3}{\hbar}g(2bb^\dagger + 1)$ with the LLL wave function $\phi_m^\downarrow(z) = \frac{(z^*)^m}{\sqrt{\pi m!}}\exp(\frac{-zz^*}{2})$. b is the operator moving between different Landau levels, while a is the operator between different degenerate angular momentum states within the same LL: $L_z = bb^\dagger - aa^\dagger$, $L_z\phi_m^\downarrow(z) = -m\phi_m^\downarrow(z)$. The wave function, besides the confining factor, is antiholomorphic in z. These spin-$\downarrow$ electrons are antichiral, and their charge conductance is quantized in units of $-e^2/h$.

The picture that now emerges is the following: our system is equivalent to a bilayer system; in one of the layers we have spin-$\downarrow$ electrons in the presence of a down-magnetic field whereas in the other layer we have spin-$\uparrow$ electrons in the presence of an up-magnetic field. These two layers are placed together. The spin-$\uparrow$ electrons have positive charge Hall conductance while the spin-$\downarrow$ electrons have negative charge Hall conductance. As such, the charge Hall conductance of the whole system vanishes. The time reversal symmetry reverses the directions of the effective orbital magnetic fields, but interchanges the layers at the same time. However, the spin Hall conductance remains finite, as the chiral states are spin up while the antichiral states are spin down, as shown in Fig. 1. The spin Hall conductance is hence quantized in units of $2\frac{e^2}{h}\frac{\hbar}{2e} = 2\frac{e}{4\pi}$. Since an electron with charge e also

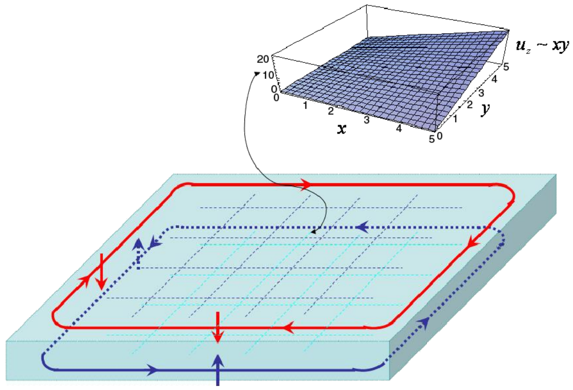

FIG. 1 (color online). Spin-↑ and spin-↓ electrons have opposite chirality as they feel the opposite spin-orbit coupling force. Total charge conductance vanishes but spin conductance is quantized. The inset shows the lattice displacement leading to the strain configuration.

carries spin $\hbar/2$, a factor of $\frac{\hbar}{2e}$ is used to convert charge conductance into the spin conductance.

We propose an entirely electric measurement to detect the quantum spin Hall effect. The edge states are *non chiral* and characterized by the number n of fermion pairs on one edge. There will be $4n$ fermion states crossing the bulk gap. In the ballistic regime, the longitudinal conductance will therefore be given by the Landauer-Buttiker formula $G_{xx} = 4ne^2/h$. Other experiments on the new state could involve the injection of spin-polarized edge states, which would acquire different chirality depending on the initial spin direction.

We now discuss the realization of a strain gradient of the specific form proposed in this Letter. The strain tensor is related to the displacement of lattice atoms from their equilibrium position u_i in the familiar way $\epsilon_{ij} = (\partial u_i/\partial x_j + \partial u_j/\partial x_i)/2$. Our strain configuration is $\epsilon_{xx} = \epsilon_{yy} = \epsilon_{zz} = \epsilon_{xy} = 0$ as well as the strain gradients $\epsilon_{zx} = gy$ and $\epsilon_{yx} = gx$. Having the diagonal strain components nonzero will not change the physics as they add only a chemical potential term to the Hamiltonian. The above strain configuration corresponds to a displacement of atoms from their equilibrium positions of the form $\vec{u} = (0, 0, 2gxy)$. This can be possibly realized by pulverizing GaAs on a substrate in molecular beam epitaxy at a rate which is a function of the position of the pulverizing beam on the substrate. The GaAs pulverization rate should vary as $xy \sim r^2 \sin(2\phi)$, where r is the distance from one of the corners of the sample where the GaAs depositing was started. Conversely, we can keep the pulverizing beam fixed at some r and rotate the sample with an angle-dependent angular velocity of the form $\sin(2\phi)$. We then move to the next incremental distance r, increase the beam rate as r^2 and again start rotating the substrate as the depositing procedure is underway.

The strain architecture we have proposed to realize the quantum spin Hall effect is by no means unique. In the present case, we have recreated the so-called symmetric gauge in magnetic-field language, but, with different strain architectures, one can create the Landau-gauge Hamiltonian and indeed many other gauges. The Landau-gauge Hamiltonian may be the easiest to realize in an experimental situation, by growing the quantum well in the [110] direction. This situation creates an off-diagonal strain $\epsilon_{xy} = \frac{1}{4}S_{44}T$, and $\epsilon_{xz} = \epsilon_{yz} = 0$ where T is the lattice mismatch (or impurity concentration), s_{44} is a material constant and x, y, z are the cubic axes. The spin-orbit part of the Hamiltonian is now $\frac{C_3}{\hbar}\epsilon_{xy}(p_x\sigma_y - p_y\sigma_x)$. However, since the growth direction of the well is [110] we must make a coordinate transformation to the x', y', z' coordinates of the quantum well (x', y' are the new coordinates in the plane of the well, whereas z' is the growth coordinate, perpendicular to the well and identical to the [110] direction in cubic axes). The coordinate transformation reads: $x' = \frac{1}{\sqrt{2}}(x - y)$, $y' = -z$, $z' = \frac{1}{\sqrt{2}}(x + y)$, and the momentum along z' is quantized.

$$H = \frac{p^2}{2m} + \frac{C_3}{\hbar} g y' p_{x'} \sigma_{z'} + Dy^2 \qquad (8)$$

where we have added a confining potential. At the suitable match between D and g, this is the Landau-gauge Hamiltonian. One can also replace the soft-wall condition (the Harmonic potential) by hard-wall boundary condition.

We now estimate the Landau level gap and the strain gradient needed for such an effect, as well as the strength of the confining potential. In the case $R \approx 2$ the energy difference between Landau levels is $\Delta E_{\text{Landau}} = 2 \times \hbar \frac{1}{2} \times \frac{C_3}{\hbar} g = C_3 g$. For a gap of 1 mK, we hence need a strain gradient or 1% over 60 μm. Such a strain gradient is easily realizable experimentally, but one would probably want to increase the gap to 10 mK or more, for which a strain gradient of 1% over 6 μm or larger is desirable. Such strain gradients have been realized experimentally, however, not exactly in the configuration proposed here [15,16]. The strength of the confining potential is in this case $D = 10^{-15}$ N/m, which corresponds roughly to an electric field of 1 V/m for a sample of 60 μm. In systems with higher spin-orbit coupling, C_3 would be larger, and the strain gradient field would create a larger gap between the Landau levels. The number of filled Landau levels varies as $n = \rho\hbar^2/(2mC_3 g)$ (the factor of 2 in the denominator comes from the fact that our Landau levels are Kramers doubled). For GaAs for which $C_3/\hbar = 8 \times 10^5$ m/s, a moderately low density of $\rho = 10^{13}$ m^{-2} would fill $n = 5$ Landau levels for $g = 1\%$ over 6 μm. For InSb, for which $C_3/\hbar = 1.8 \times 10^6$ m/s [9] a conduction band density of 1.4×10^{13} m^{-2} would fill $n = 3$ Landau levels for the same value of strain gradient as before. Attaining fractional fillings requires either lower density samples or higher gradient strains.

We now turn to the question of the many-body wave function in the presence of interactions. For our system this is very suggestive, as the wave function incorporates both holomorphic and antiholomorphic coordinates, by contrast to the pure holomorphic Laughlin states. Let the spin-↑ coordinates be z_i while the spin-↓ coordinates are the w_i. z_i enter in holomorphic form in the wave function whereas w_i enter antiholomorphically. While if the spin-↑ and spin-↓ electrons would lie in separate bi-layers the many-body wave function would be just $\prod_{i<j}(z_i - z_j)^m \prod_{k<l}(w_k^* - w_l^*)^m e^{-1/2(\sum_i z_i z_i^* + \sum_k w_k w_k^*)}$, where m is an odd integer. Since the particles in our state reside in the same quantum well and may possibly experience the additional interaction between the different spin states, a more appropriate wave function is

$$\psi(z_i, w_i) = \prod_{i<j}(z_i - z_j)^m \prod_{k<l}(w_k^* - w_l^*)^m \times \prod_{r,s}(z_r - w_s^*)^n e^{-1/2(\sum_i z_i z_i^* + \sum_k w_k w_k^*)}. \quad (9)$$

The above wave function is symmetric upon the interchange $z \leftrightarrow w^*$ reflecting the spin-↑-chiral–spin-↓-antichiral symmetry. This wave function is, of course, analogous to the Halperin's wave function of two different spin states [17]. The above many-body wave function is valid in the limit of strong Haldane pseudopotentials. As in the case of the classical Laughlin wave functions, the low exponent (m, n) functions tend to be variationally better in the limit of infinitely short-range repulsion. The physics and value of the Coulomb interactions are not modified by the spin-orbit coupling and we hence expect that the same repulsive interaction that stabilizes the Laughlin fractional states will stabilize our above state.

Many profound topological properties of the quantum Hall effect are captured by the Chern-Simons-Landau-Ginzburg theory [18]. While the usual spin-orbit coupling for spin-1/2 systems is T-invariant but P-breaking, our spin-orbit coupling is also P invariant due to the strain gradient. The low energy field theory of the spin Hall liquid is hence a double Chern-Simons theory with the action:

$$S = \frac{\nu}{4\pi} \int \epsilon^{\mu\nu\rho} a_\mu \partial_\nu a_\rho - \frac{\nu}{4\pi} \int \epsilon^{\mu\nu\rho} c_\mu \partial_\nu c_\rho \quad (10)$$

where the a_μ and c_μ fields are associated with the left and right movers of our theory while ν is the filling factor. Essentially, the two Chern-Simons terms have the same filling factor ν. Such special theories avoid the chiral anomaly [19] and their Berry phases have been recently proposed as preliminary examples of topological quantum computation [19].

A similar situation of Landau levels without magnetic field arises in rotating Bose-Einstein condensates (BECs) where the mean-field Hamiltonian is similar to either $H_\uparrow$ or $H_\downarrow$. In the limit of rapid rotation, the condensate expands and becomes effectively two-dimensional. The L_z term is induced by the rotation vector Ω [20]. The LLL behavior is achieved when the rotation frequency reaches a specific value analogous to the case $R \approx 2$ in our Hamiltonian. In the BEC literature this is the so-called mean-field quantum Hall limit and the ground state wave function is a Laughlin-type one. In contrast to our case, the theory is still T breaking.

We wish to thank S. Kivelson, E. Fradkin, J. Zaanen, D. Santiago and C. Wu for useful discussions. This work is supported by the NSF under Grant No. DMR-0342832 and the US Department of Energy, Office of Basic Energy Sciences under Contract No. DE-AC03-76SF00515.

Note added.—After the submission of this Letter, we became aware of the related work by Kane and Mele [21], in which an integer quantum spin Hall effect is predicted for a model of graphene. Three new papers [22–24] discussed various aspects of the quantum spin Hall.

[1] S. Murakami, N. Nagaosa, and S. Zhang, Science **301**, 1348 (2003).
[2] J. Sinova *et al.*, Phys. Rev. Lett. **92**, 126603 (2004).
[3] Y. Kato *et al.*, Science **306**, 1910 (2004).
[4] J. Wunderlich *et al.*, Phys. Rev. Lett. **94**, 047204 (2005).
[5] B. Bernevig and S. Zhang, Phys. Rev. Lett. **95**, 016801 (2005).
[6] N. Mott, Proc. R. Soc. A **124**, 425 (1929).
[7] A related question was posed by F. D. M. Haldane [Phys. Rev. Lett. **61**, 2015 (1988)] on whether the quantum Hall effect necessarily requires an external magnetic field or can arise as a consequence of magnetic ordering in two-dimensional systems. The state proposed breaks time reversal due to a local magnetic flux density of zero total flux through the unit cell.
[8] M. I. Dyakonov *et al.*, Sov. Phys. JETP **63**, 655 (1986).
[9] R. Ranvaud *et al.*, Phys. Rev. B **20**, 701 (1979).
[10] J. E. Pikus and A. Titkov, *Optical Orientation* (North Holland, Amsterdam, 1984), p. 73.
[11] W. Howlett and S. Zukotynski, Phys. Rev. B **16**, 3688 (1977).
[12] A. Khaetskii and Y. Nazarov, Phys. Rev. B **64**, 125316 (2001).
[13] T. B. Bahder, Phys. Rev. B **41**, 11 992 (1990).
[14] Y. Kato *et al.*, Nature (London) **427**, 50 (2004).
[15] Q. Shen *et al.*, Phys. Rev. B **54**, 16 381 (1996).
[16] Q. Shen and S. Kycia, Phys. Rev. B **55**, 15 791 (1997).
[17] B. Halperin, Helv. Phys. Acta **56**, 75 (1983).
[18] S. C. Zhang, H. Hansson, and S. Kivelson, Phys. Rev. Lett. **62**, 82 (1989).
[19] M. Freedman *et al.*, cond-mat/0307511.
[20] T. Ho, Phys. Rev. Lett. **87**, 060403 (2001).
[21] C. L. Kane and E. J. Mele, Phys. Rev. Lett. **95**, 226801 (2005).
[22] X. L. Qi, Y. S. Wuand, and S. C. Zhang, cond-mat/0505308.
[23] C. L. Kane and E. J. Mele, Phys. Rev. Lett. **95**, 146802 (2005).
[24] L. Sheng *et al.*, Phys. Rev. Lett. **95**, 136602 (2005).