Hydrodynamic theory of surface excitations of three-dimensional topological insulators

N. M. Vildanov

I. E. Tamm Department of Theoretical Physics, P. N. Lebedev Physics Institute, 119991 Moscow, Russia
(Received 24 November 2010; published 31 March 2011)

Edge excitations of a fractional quantum Hall system can be derived as surface excitations of an incompressible quantum droplet using one-dimensional chiral bosonization. Here we show that an analogous approach can be developed to characterize surface states of three-dimensional time-reversal invariant topological insulators. The key ingredient of our theory is the Luther's multidimensional bosonization construction.

Topological insulators (TIs) are materials with insulating bulk and topologically protected metallic edge or surface states.[1–3] Two-dimensional (2D) TIs are also called quantum spin Hall (QSH) systems. Edge states of QSH systems are analogous to quantum Hall edge states. However, there is an important distinction:[3,4] edge states of the QH systems with up and down spins propagate in the same direction, whereas QSH edge states with opposite directions of spin counterpropagate. This is the consequence of the time-reversal invariance of quantum spin Hall effect (QSHE), which is broken in conventional QH systems. Therefore the former are called chiral and the latter are called helical edge states.

There are two classes of nontrivial time-reversal invariant TIs in three dimensions which are called strong and weak.[5–7] While weak TIs are layered 2D QSH states (in the sense that these two systems can be adiabatically coupled to each other), strong TIs are purely three dimensional. Surface states of these TIs are massless 2D Dirac fermions. Weak and strong TIs are distinguished by the number of Dirac cones on the surface: Strong TIs have an odd number and weak TIs have an even number of Dirac cones on the surface. Gapless modes of strong TIs are robust and insensitive to weak interactions and disorder.

Surface states of three-dimensional time-reversal invariant TIs are spin filtered, such that $\langle \vec{s}(-\mathbf{k}) \rangle = -\langle \vec{s}(\mathbf{k}) \rangle$, which means that spin density and charge current are coupled (see, e.g., Ref. 8). This suggests that three-dimensional strong TIs realize quantum spin Hall effect in every radial direction. "Radial direction" here means that one considers tomographic projection of the surface states on a certain direction. This tomographically projected state in certain cases can be viewed as a two-dimensional quantum spin Hall system. In this Brief Report we explicitly show how to construct such mapping mathematically. In constructing this mapping we assume that a hydrodynamic theory of the QSHE edges can be developed in the similar way as for QHE edges.[9–11] Several interesting physical quantities characterizing three-dimensional TIs are introduced.

It should be made clear that many of the ideas presented here taken separately are not new. We have only combined them in a single picture. We also note that this Brief Report is illustrative rather than rigorous.

A good description of the hydrodynamic theory of the FQHE edge states can be found in Ref. 12. Here we briefly mention the main points. Then we modify this theory to describe surface states of strong topological insulators.

Suppose that the FQHE states are incompressible irrotational liquid without bulk excitations. Then the only low-lying excitations are the surface deformations of the quantum droplet. The droplet is confined by a smooth potential well. The electric field of the potential well will generate a persistent current along the edge fluctuating part of which is given by

$$\mathbf{j} = \sigma_{xy}[\mathbf{e}_z, \mathbf{E}]h(x), \quad \sigma_{xy} = \nu \frac{e^2}{h},$$

where ν is the filling fraction, $\mathbf{e}_z$ is the unit vector along the z axis, $h(x)$ is the displacement of the edge from its equilibrium value, x is the coordinate along the edge.

The one-dimensional density of the edge wave $\rho(x)$ is related to the displacement of the edge $h(x)$ through $\rho(x) = nh(x)$, where $n = \nu \frac{eH}{hc}$ is the two-dimensional electron density in the bulk. Then continuity equation reads

$$\partial_t h - \nu \partial_x h = 0, \tag{1}$$

where $v = c\frac{E}{H}$. This means that the electrons at the edge drift with the velocity v.

The Hamiltonian (energy) of the edge waves is given by

$$H = \frac{1}{2} e \int h\rho E dx = \frac{\pi v}{\nu} \int \rho^2 dx. \tag{2}$$

It is easy to quantize this Hamiltonian. Rewriting Eqs. (1) and (2) in the momentum representation and identifying ρ_k with the coordinate variable, one finds that the corresponding canonical momentum is given by $p_k = 2\pi i \rho_{-k}/\nu k$ (the zero mode with $k = 0$ is excluded from the theory due to the incompressibility of the liquid). From the commutation relations $[p_k, \rho_{k'}] = i\delta_{kk'}$ one obtains the Kac-Moody algebra for the currents,

$$[\rho_k, \rho_{k'}] = \frac{\nu}{2\pi} k \delta_{k+k'}. \tag{3}$$

This theory provides a complete description of low-lying excitations of the Laughlin state.[12]

Now we want to employ this theory to QSH systems. The easiest way to understand QSHE is to consider the case when spin $\hat{s}_z$ is conserved.[13] Then one can define two sectors with spin up and down. Such decomposition is possible even when spin is not conserved.[14] Then one can define the Chern numbers $n_{\uparrow,\downarrow}$ for the spin up and down sectors in the usual manner.[15,16] In a time-reversal invariant system the total Chern number $n_{\uparrow} + n_{\downarrow} = 0$ is zero. However, the difference of the two Chern numbers in general is not zero and one can define the spin Chern number according to $C_s = \frac{1}{2}(n_{\uparrow} - n_{\downarrow})$. Loosely

speaking, C_s determines the number of gapless edge modes in the system. It was shown that the edges of a QSH system with even C_s are localized by disorder, while edges of a QSH system with odd C_s are robust against small time-reversal invariant perturbations.[17,18] The case $n_\uparrow = -n_\downarrow = 1$ is the simplest and to describe edge states of such a system one needs two uncoupled incompressible liquids ρ_1 and ρ_2 with the edge state Hamiltonians (2), which are related by time-reversal symmetry. Edge excitations travel in opposite directions with the velocity $u = eE/hn$. Here the index 1 refers to spin-up and the index 2 refers to spin-down sectors.

In analogy with the above case, we assume that the low-lying excitations of a three-dimensional topological insulator are the surface excitations of two uncoupled incompressible quantum liquids confined by a potential well, with the following Hamiltonian:

$$H = \int \frac{1}{2} \rho_0 eE \left[h_1^2(x,y) + h_2^2(x,y) \right] dx\,dy, \qquad (4)$$

where ρ_0 is the density of the electronic liquid, $\mathbf{x} = (x,y)$, x and y are coordinates along the surface. The two-dimensional density of the surface waves is $\rho_i(x,y) = \rho_0 h_i(x,y)$, $i = 1,2$. Incompressibility of the liquid means that $\int h_i(x,y) dx\,dy = 0$. It will turn out that after making some assumptions about the dynamics of the fields $h_{i\theta}$ this case corresponds to a strong TI with a single Dirac cone on the surface.

The essence of Luther's approach is to consider a D-dimensional space as a set of one-dimensional "rays."[19–23] Following these ideas, we write the surface state Hamiltonian (4), as a sum over tomographic projections,

$$H = \int_{\mathcal{R}} d\theta \int \frac{1}{2} neE \left[h_{1\theta}^2(\xi) + h_{2\theta}^2(\xi) \right] d\xi, \qquad (5)$$

integrating over the right hemicircle $\mathcal{R} = \{-\pi/2 \leqslant \theta \leqslant \pi/2\}$. Here the surface displacement of the tomographically projected system $h_\theta(\xi)$ is defined as

$$h_{i\theta}(\xi) = \int_0^\infty \left(\frac{k\rho_0}{2\pi^3 n} \right)^{1/2} dk \int \cos k(\xi - \xi')$$
$$\times h_i(x',y') \, dx'\,dy'. \qquad (6)$$

$\mathbf{k} = |k|\hat{k}, \hat{k} = (\cos\theta, \sin\theta)$, $\xi = \hat{k} \cdot \mathbf{x}$, $\xi' = \hat{k} \cdot \mathbf{x}'$; since the z coordinate remains intact, we can consider the planes (ξ, z) which are labeled by θ. We call these planes tomographic planes or tomographic projections of the initial three-dimensional system. n is the two-dimensional density of the electronic liquid in the tomographic plane, which is determined from the consistency of the two descriptions. The electric field of the confining potential well is the same in both cases.[24] It is easily verified that tomographic projections are also incompressible liquids. Now we make the following assumption that these tomographic planes are QSH systems. From the topological band theory we can be sure that this is correct at least for three values of the parameter θ (see the discussion at the end of the Brief Report). The fact that this is correct for all θ will be justified below, because this is the only assumption that leads to the desired result: a single Dirac cone on the surface, which will be obtained after fermionization of the model. Then one finds that excitations have the linear spectrum $\omega_k = u|k|$, where $u = eE/hn$. Thus one can relate

the unknown parameter of the theory n to the parameters of the three-dimensional theory.

The choice of the range of θ is not unique and this reflects the fact that the splitting of the full Hilbert space induced by the time-reversal operation is not unique. When time-reversal symmetry is preserved such splitting is necessary to obtain a nontrivial Chern number of a 2D system, because the Chern number of the whole Hilbert space vanishes (in three dimensions one needs to consider certain 2D sections of the Brillouin zone and further split them using the time-reversal operation).

One can define Fourier components of $h_{i\theta}(\xi)$:

$$\tilde{h}_{i\theta}(k) = \int h_{i\theta}(\xi) e^{-ik\xi} d\xi. \qquad (7)$$

Since by our assumption $\tilde{h}_{i\theta}(k)$ are the edge modes of a QSH system, they must satisfy the equations of motion (continuity equations) $\partial_t \tilde{h}_{1\theta}(k) = iuk\tilde{h}_{1\theta}(k)$ and $\partial_t \tilde{h}_{2\theta}(k) = -iuk\tilde{h}_{2\theta}(k)$. From Eq. (6) one also has $\tilde{h}_{1\theta}(\pm|k|) = \sqrt{\frac{|k|\rho_0}{2\pi n}} h_1(\pm\mathbf{k})$. This finally leads to

$$\partial_t h_1(\pm\mathbf{k}) = \pm iu|k| h_1(\pm\mathbf{k}), \qquad (8)$$

$$\partial_t h_2(\pm\mathbf{k}) = \mp iu|k| h_2(\pm\mathbf{k}) \qquad (9)$$

when $\mathbf{k}$ is in the right hemicircle. In the second-quantized form, Hamiltonian (4) is

$$H_B = \sum_{\mathbf{k} \in \mathcal{R}} u|k|(\alpha_{\mathbf{k}}^\dagger \alpha_{\mathbf{k}} + \beta_{\mathbf{k}}^\dagger \beta_{\mathbf{k}}), \qquad (10)$$

where

$$[\alpha_{\mathbf{k}}, \alpha_{\mathbf{k}'}^\dagger] = [\beta_{\mathbf{k}}, \beta_{\mathbf{k}'}^\dagger] = \delta_{\mathbf{k},\mathbf{k}'}, \quad [\alpha_{\mathbf{k}}, \beta_{\mathbf{k}'}^\dagger] = 0,$$

e.g, $\alpha_{\mathbf{k}}$ is related to $h_1(\mathbf{k})$ through $\alpha_{\mathbf{k}} = \sqrt{\rho_0/n|k|} h_1(-\mathbf{k})$, $\alpha_{\mathbf{k}}^\dagger = \sqrt{\rho_0/n|k|} h_1(\mathbf{k})$ when $\mathbf{k} \in \mathcal{R}$. Hamiltonian (10) is half of the massless Klein-Gordon model, exactly what is needed to construct a massless two-component Dirac fermion.[19] This is similar to that chiral bosons in one dimension are half of the ordinary bosons. One can associate the following bosonic fields with this model:

$$\phi_1(\theta, \hat{k} \cdot \mathbf{x}) = \left(\frac{S}{2\pi^2} \right)^{1/2} \int_0^\infty e^{-\alpha k/2} dk$$
$$\times (\alpha_{\mathbf{k}} e^{ik(\hat{k}\cdot\mathbf{x})} + \text{H.c.}) \qquad (11)$$

$$\phi_2(\theta, \hat{k} \cdot \mathbf{x}) = -\left(\frac{S}{2\pi^2} \right)^{1/2} \int_0^\infty e^{-\alpha k/2} dk$$
$$\times (\beta_{\mathbf{k}}^\dagger e^{ik(\hat{k}\cdot\mathbf{x})} + \text{H.c.}). \qquad (12)$$

α is the cutoff which should be taken to zero at the end; S is the surface area. These fields do not correspond to any local observables. However, Luther showed that appropriate functions of these fields do.

Here we review some of the details of Luther's construction for completeness.[25] Suppose we have a fermionic Hamiltonian

$$H_F = u \sum_{\mathbf{k}} a_{\mathbf{k}}^\dagger (\vec{k} \cdot \vec{\sigma}) a_{\mathbf{k}}. \qquad (13)$$

This is a Hamiltonian of massless Dirac fermions. It can be diagonalized by a transformation

$$U = e^{iS}, \quad S = i \sum_{\mathbf{k}} \frac{\pi}{4} a_{\mathbf{k}}^{\dagger} \mathbf{V} \cdot \vec{\sigma} a_{\mathbf{k}}, \quad (14)$$

where $\mathbf{V} = (-\sin\theta, \cos\theta)$. Fermi operators transform according to

$$a_{\mathbf{k}}' = e^{-i\pi \mathbf{V}\cdot\vec{\sigma}/4} a_{\mathbf{k}}. \quad (15)$$

The diagonal Hamiltonian is

$$H_F' = \sum_{\mathbf{k}} u|k| a_{\mathbf{k}}'^{\dagger} \hat{\sigma}_z a_{\mathbf{k}}'. \quad (16)$$

The boson representation is given by

$$\psi(\theta, \hat{k} \cdot \mathbf{x}) = e^{-i\pi \mathbf{V}\cdot\vec{\sigma}/4} \psi'(\theta, \hat{k} \cdot \mathbf{x}), \quad (17)$$

$$\psi'(\theta, \hat{k} \cdot \mathbf{x}) = \frac{1}{2\pi\alpha} \begin{pmatrix} \exp[\phi_1(\theta, \hat{k} \cdot \mathbf{x})] \\ \exp[\phi_2(\theta, \hat{k} \cdot \mathbf{x})] \end{pmatrix} \quad (18)$$

(the Klein factors necessary to ensure anticommutation relations are omitted for simplicity; for details, see Refs. 19 and 20). The usual Fermi fields are given by

$$\psi(\mathbf{x}) = \int_{\mathcal{R}} d\theta \, \psi(\theta, \hat{k} \cdot \mathbf{x}). \quad (19)$$

Representation (11), (12), (17)–(19) is constructed in such a way that the correlation functions of free fermionic fields are correctly reproduced. Note a crucial point that in definition (19) the integration is only over half of the whole circle. Thus we see that Hamiltonian (4) together with equations of motion (8) and (9) [or alternatively Hamiltonian (5) where $h_{i\theta}(\xi)$ are edge states of a QSH insulator] is equivalent to the massless Dirac Hamiltonian. This corresponds to a strong TI with a single Dirac cone on the surface. These calculations confirm our initial intuition.

There is only one Z_2 invariant in two dimensions.[26] It can be defined as $C_s \bmod 2$.[27] In three dimensions there are four Z_2 invariants.[5–7] Three of them are equivalent to invariants of two-dimensional topological insulators and are defined as invariants of some sections of the Brillouin zone. The fourth topological invariant ν_0 is purely three dimensional. TIs with $\nu_0 = 0$ are called weak and TIs with $\nu_0 = 1$ are called strong. This invariant determines the number of Kramers degenerate Dirac points enclosed by the Fermi surface.

In any time-reversal invariant system with spin-orbit interactions there are two-dimensional Dirac points in the surface spectrum. Therefore the above considerations should

be clarified. A single Dirac fermion on the surface already means that the insulator is a strong topological insulator. We will show this directly and in parallel discuss the relation of our picture to the conventional theory of topological insulators. To make connection of this picture with the band topology, we use the simple argument for counting the topological invariants due to Roy,[28] which is quoted below. For us the important aspect of this work is how the Z_2 invariants of certain planes in the Brillouin zone, such as $p_x = p_y$, can be computed.

Let represent the Brillouin zone by a cube $\{-\pi \leqslant p_x, p_y, p_z \leqslant \pi\}$ and the Z_2 invariants associated with the planes $p_x = 0$, $p_x = \pi$, $p_y = 0$, $p_y = \pi$ by $\nu_1, \tilde{\nu}_1, \nu_2, \tilde{\nu}_2$, respectively. Then the Z_2 invariant of the plane $p_x = p_y$ equals $\nu_1 + \tilde{\nu}_2$ and the Z_2 invariant distinguishing strong topological insulators from weak topological insulators equals $\nu_0 = \nu_1 - \tilde{\nu}_1 = \nu_2 - \tilde{\nu}_2$. Any 3D topological insulator with time-reversal invariance can be characterized by four invariants, which may be chosen to be ν_1, ν_2, ν_3, and ν_0.

Above we assumed that every tomographic plane supports QSHE. In fact, it is sufficient to consider only three planes (then for the rest this would be satisfied automatically). Let these planes be $p_x = \pi$, $p_y = \pi$, $p_x = p_y$. We will consider for concreteness the Dirac point $(\pi, \pi, 0)$ and the surface states on the (x, y) plane having small 2D momentum $\mathbf{k}$ around $p_x = \pi, p_y = \pi$ (the direction of $\mathbf{k}$ is given by the angle θ). In this case $\theta = 0$ corresponds to the plane $p_y = \pi$, $\theta = \pi/2$ to the plane $p_x = \pi$, and $\theta = \pi/4$ to the plane $p_x = p_y$. Suppose that Z_2 invariant of each of these planes is odd, thus corresponding to a nontrivial insulator. Then we have $\tilde{\nu}_1 = \tilde{\nu}_2 = \nu_1 + \tilde{\nu}_2 = 1$, which gives $\nu_0 = 1$. This corresponds to a strong topological insulating phase. If there is no such point in the Brillouin zone, for which all three Z_2 invariants are 1, then the insulator is in the weak topological insulating phase with $\nu_0 = 0$ (see also Fig. 3 in Ref. 29).

In general, it is not possible to define a Z_2 invariant of a plane with arbitrary θ using topological band theory. It seems that this is possible only for planes such as $\tan\theta = m/n$, where m and n are two coprime integers. In the case we have considered, all such invariants are equal to 1.

Recently, bosonization approach was applied to topological insulators also in the Ref. 30, but in a different context. In passing we also note that surface excitations of certain 3D topological superconductors[31] can be viewed, in every radial direction, as edge states of a 2D topological superconductor in the same class. It would be interesting to explore this case as well.

In summary, we have shown that low-lying excitations of a strong TI with a single Dirac cone on the surface are the surface deformations of a droplet of incompressible quantum liquid. These excitations have very unusual form, however, they have a simple meaning when one considers tomographic projections of this liquid (as defined in the text): They are two chiral waves propagating in opposite directions. Thus, surface excitations of a strong TI with a single Dirac cone on the surface are the sum of QSH edge states. However, this is not true for the entire topological insulator, i.e., a strong TI can not be presented as a sum of 2D TIs. This Brief Report can be considered as an another physical illustration of the fact that 3D topological insulators can be characterized by 2D invariants.

---

[1] M. Z. Hasan and C. L. Kane, Rev. Mod. Phys. **82**, 3045 (2010).

[2] J. E. Moore, Nature (London) **464**, 194 (2010).

[3] X. L. Qi and S. C. Zhang, Phys. Today **63**, 33 (2010).

[4] M. Buttiker, Science **325**, 278 (2009).

[5] L. Fu, C. L. Kane, and E. J. Mele, Phys. Rev. Lett. **98**, 106803 (2007).

[6] J. E. Moore and L. Balents, Phys. Rev. B **75**, R121306 (2007).

[7] R. Roy, Phys. Rev. B **79**, 195322 (2009).

[8] S. Raghu, S. B. Chung, X.-L. Qi, and S.-C. Zhang, Phys. Rev. Lett. **104**, 116401 (2010).

[9] D. H. Lee and X. G. Wen, Phys. Rev. Lett. **66**, 1765 (1991).

[10] F. D. M. Haldane, Bull. Am. Phys. Soc. **35**, 254 (1990).

[11] M. Stone, Ann. Phys. (NY) **207**, 38 (1991); Int. J. Mod. Phys. B **5**, 509 (1991).

[12] X. G. Wen, Int. J. Mod. Phys. B **6**, 1711 (1992).

[13] C. L. Kane and E. J. Mele, Phys. Rev. Lett. **95**, 226801 (2005).

[14] E. Prodan, Phys. Rev. B **80**, 125327 (2009).

[15] D. J. Thouless, M. Kohmoto, M. P. Nightingale, and M. den Nijs, Phys. Rev. Lett. **49**, 405 (1982).

[16] J. E. Avron, R. Seiler, and B. Simon, Phys. Rev. Lett. **51**, 51 (1983).

[17] C. Wu, B. A. Bernevig, and S.-C. Zhang, Phys. Rev. Lett. **96**, 106401 (2006).

[18] C. Xu and J. E. Moore, Phys. Rev. B **73**, 045322 (2006).

[19] A. Luther, Phys. Rev. B **19**, 320 (1979).

[20] A. Luther, Phys. Rep. **49**, 261 (1979).

[21] H. Aratyn, Phys. Rev. D **28**, 2016 (1983); H. Aratyn, Nucl. Phys. B **227**, 172 (1983).

[22] P. W. Anderson, Phys. Rev. Lett. **67**, 2092 (1991).

[23] C. M. Sommerfeld, in *Strings and Symmetries*, edited by G. Aktaş, C. Saçlıoğlu, and M. Serdaroğlu, Lecture Notes in Physics Vol. 447 (Springer-Verlag, Berlin, 1995).

[24] We assume that E is constant on the surface or varies very slowly as a function of x and y.

[25] Although Luther considered three-dimensional fermions, it is easy to extend his results to two dimensions.

[26] C. L. Kane and E. J. Mele, Phys. Rev. Lett. **95**, 146802 (2005).

[27] D. N. Sheng, Z. Y. Weng, L. Sheng, and F. D. M. Haldane, Phys. Rev. Lett. **97**, 036808 (2006).

[28] R. Roy, New J. Phys. **12**, 065009 (2010).

[29] L. Fu and C. L. Kane, Phys. Rev. B **76**, 045302 (2007).

[30] G. Y. Cho and J. E. Moore, e-print arXiv:1011.3485 (to be published).

[31] X. L. Qi, T. L. Hughes, S. Raghu, and S. C. Zhang, Phys. Rev. Lett. **102**, 187001 (2009).