OPEN ACCESS

Phase transition between the quantum spin Hall and insulator phases in 3D: emergence of a topological gapless phase

To cite this article: Shuichi Murakami 2007 *New J. Phys.* **9** 356

New Journal of Physics
The open-access journal for physics

Phase transition between the quantum spin Hall and insulator phases in 3D: emergence of a topological gapless phase

Shuichi Murakami
Department of Applied Physics, University of Tokyo, Hongo, Bunkyo-ku, Tokyo 113-8656, Japan
E-mail: murakami@appi.t.u-tokyo.ac.jp

New Journal of Physics **9** (2007) 356
Received 4 June 2007
Published 28 September 2007
Online at http://www.njp.org/
doi:10.1088/1367-2630/9/9/356

Abstract. Phase transitions between the quantum spin Hall (QSH) and the insulator phases in three dimensions (3D) are studied. We find that in inversion-asymmetric systems there appears a gapless phase between the QSH and insulator phases in 3D which is in contrast with the 2D case. Existence of this gapless phase stems from a topological nature of gapless points (diabolical points) in 3D, but not in 2D.

Contents

PII: S1367-2630(07)51934-8

1. Introduction

In the intrinsic spin Hall effect (SHE) [1, 2], an external electric field applied to a doped semiconductor induces a transverse spin current. It has been attracting special interest from the following aspects. Firstly, it can produce spin current without breaking the time-reversal symmetry (T-symmetry), namely without magnetism or magnetic field, which may be potentially important for spintronics device applications. Secondly, the mechanism itself is dissipationless and thus it may open a way for spintronics devices with less power consumption. Thirdly, the SHE is driven by the spin–orbit coupling, which can be even larger at room temperature. Thus, it is expected to survive even at room temperature, as has been confirmed experimentally [3]–[5]. In particular, recent experiments on platinum shows the largest spin Hall conductivity so far [5], of about 240 Ω^{-1} cm^{-1} at room temperature. This might be attributed as an intrinsic SHE caused by near-degeneracy near the Fermi energy [6].

Henceforth, we restrict ourselves to time-reversal-symmetric systems. In relation to the SHE in conducting systems, there has been a growing interest in the SHE in insulating systems. The first proposal is spin Hall insulators [7]. It was shown that in insulators such as HgTe under uniaxial pressure or in PbTe the spin Hall conductivity is nonzero. In a sense, they are ordinary insulators with spin–orbit coupling. Another proposal is quantum spin Hall (QSH) systems, both in two dimensions (2D) [8]–[10] and in 3D [11, 12]. They are insulators in the bulk while the boundaries (i.e. edges in 2D or surfaces in 3D) are gapless and carry spin currents. They can be regarded as topological insulators; these gapless boundary states are topologically protected against T-symmetric perturbations [13, 14]. Experimental observations are yet to be made. We recently proposed that thin-film bismuth is a good candidate for the 2D QSH phase [15]. Another candidate for the 2D QSH phase is the CdTe/HgTe/CdTe quantum well [16]. The QSH phases are, however, yet to be realized experimentally. The distinction between the QSH and ordinary insulating (i.e. spin Hall insulator) phases is the absence or presence of (topologically protected) boundary states, which is characterized by the Z_2 topological number of the bulk states [9, 17, 18].

In the previous paper [19], we considered a phase transition between the QSH and the insulating phases in 2D. It was found that the phase transitions are classified into two cases, corresponding to the presence or absence of the inversion symmetry (I-symmetry) in the systems. It was also found that the effective theory describing the phase transition consists of two decoupled theories of two-component fermions.

In this paper, we study the phase transition between the two phases in 3D in the bulk. At first sight it might be similar to the 2D case studied in [19]. We find in this paper that it is not. Topological properties such as the present problem can be very different for different dimensions. In 3D, the inversion-symmetric (I-symmetric) case is similar to that in 2D. Nevertheless the inversion-asymmetric (I-asymmetric) case is different; when we vary an external parameter which drives the phase transition, a gapless phase appears between the QSH and the insulating phases. The existence of this gapless phase is enforced from a topological origin, and this phase does not exist in 2D. These discussions not only deepen our understanding toward the Z_2 topological numbers, but also imply some hints in the search for 3D QSH phase materials.

The paper is organized as follows. In section 2, we discuss generic phase transitions between the two phases in 3D. In section 3, we discuss the implications of the present theory

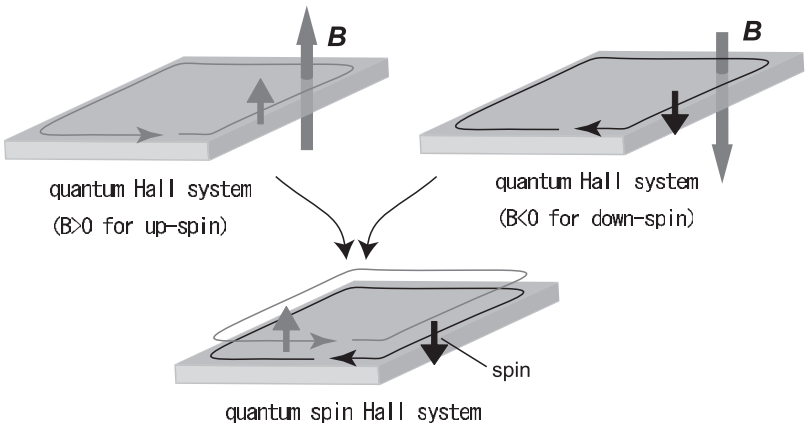

Figure 1. Schematic picture of the QSH system as a superposition of two QH systems.

for the search for QSH phase materials. Section 4 is devoted to conclusions and discussions. We neglect effects of interactions and impurities in the present paper.

2. Phase transitions between the QSH and insulating phases in 3D

2.1. Introduction on the QSH phase and Z_2 topological number

We first review the QSH phase and Z_2 topological number ν [8, 9]. The simplest example of the QSH phase can be realized by a superposition of two quantum Hall (QH) systems for the up- and down-spins having opposite (effective) magnetic field (see figure 1). Suppose for the up-spin (down-spin) subsystem the QH conductance is $\sigma_{xy}^{\uparrow} = e^2/h$ ($\sigma_{xy}^{\downarrow} = -e^2/h$). The whole system then has edge states with two spins propagating in the opposite direction. The whole system is T-symmetric, and the effective magnetic field can be realized by the spin–orbit coupling.

This is only the simplest special example. In general, the QSH allows a spin–orbit coupling term which mixes spins, without breaking T-symmetry. The QSH phase is defined as a T-symmetric system which is gapful in the bulk and gapless in the edge. Because of the T-symmetry, the edge states form Kramers pairs, consisting of two states with opposite spins counter-propagating from each other. The Z_2 topological number ν is used for distinguishing this QSH phase from the usual insulator phase, namely the spin Hall insulator phase. The Z_2 topological number can take only two different values for ν: even or odd, and this number simply means whether the number of Kramers pairs of edge states is even or odd. If $\nu =$ odd, the system is in the QSH phase, while if $\nu =$ even the system is in the insulating phase. This means that the system with an even number of Kramers pairs of edge states is equivalent to a system with no gapless edge state. It follows because general perturbations preserving T-symmetry can open a gap in the edge states when $\nu =$ even. In contrast, for the QSH phase ($\nu =$ odd), the edge states remain gapless even in the presence of T-symmetric perturbation, including nonmagnetic impurities and/or interaction, as long as they are not too strong [13, 14].

There are a number of equivalent expressions for the Z_2 topological number [11, 17, 18], and we only explain the ones relevant for the subsequent discussions, both in 2D and in 3D. We assume that the spectrum of the Hamiltonian has a gap, within which the Fermi energy E_{F} is located.

First we explain the 2D case. For I-asymmetric systems, the spectrum is doubly degenerate only at the four points $\boldsymbol{k} = \boldsymbol{k}_i \equiv \boldsymbol{G}/2$ ($i = 1, 2, 3, 4$), and non-degenerate at other points. In such systems, the Z_2 topological number ν is determined as

$$(-1)^{\nu} = \prod_{i=1}^{4} \delta_i, \tag{1}$$

where

$$\delta_i = \frac{\sqrt{\det[\mathrm{w}(\boldsymbol{k}_\mathrm{i})]}}{\mathrm{Pf}[w(\boldsymbol{k}_i)]} = \pm 1. \tag{2}$$

Here $w(\boldsymbol{k})$ is a unitary matrix with elements given by $w_{mn}(\boldsymbol{k}) = \langle u_{-\boldsymbol{k},m}|\Theta|u_{\boldsymbol{k},n}\rangle$, and $|u_{\boldsymbol{k},n}\rangle$ is the Bloch wavefunction of an nth band whose eigenenergy lies below E_F. Θ is the time-reversal operator, represented as $\Theta = \mathrm{i}\sigma_y K$ with K being complex conjugation. The branch of the square root of the determinant is so chosen that the wavefunctions (including their phases) are continuous in the whole Brillouin zone.

On the other hand, in I-symmetric systems, the formula simplifies drastically; there is no need to calculate the phases of the wavefunctions for the whole Brillouin zone, which is advantageous for numerical calculation. It is given by

$$(-1)^{\nu} = \prod_{i=1}^{4} \delta_i, \quad \delta_i = \prod_{m=1}^{N} \xi_{2m}(\boldsymbol{k}_i), \tag{3}$$

where $\xi_{2m}(\boldsymbol{k}_i)$ ($= \pm 1$) is the parity eigenvalue of the Kramers pairs at each of these points, and N is the number of Kramers pairs below E_F.

In 3D, there are four Z_2 topological numbers written as ν_0; ($\nu_1\nu_2\nu_3$) [11, 12], given by

$$(-1)^{\nu_0} = \prod_{i=1}^{8} \delta_i, \quad (-1)^{\nu_k} = \prod_{n_k=1; n_{j\neq k}=0,1} \delta_{i=(n_1 n_2 n_3)}, \tag{4}$$

where $\delta_{i=(n_1n_2n_3)}$ ($= \pm 1$) is defined for the wavevector $\boldsymbol{k}_i = \frac{1}{2}(n_1\boldsymbol{b}_1 + n_2\boldsymbol{b}_2 + n_3\boldsymbol{b}_3)$ ($n_i = 1, 2, 3$) and b_k ($k = 1, 2, 3$) are the primitive vectors of the reciprocal lattice. These eight wavevectors satisfy $\boldsymbol{k}_i = -\boldsymbol{k}_i \pmod{\boldsymbol{G}}$. These topological numbers in 3D determine the topology of the surface states for arbitrary crystal directions [11]. We note that among the four Z_2 topological numbers in 3D, only ν_0 is robust against nonmagnetic impurities, while the others (ν_k ($k = 1, 2, 3$)) are meaningful only for a relatively clean sample [11].

2.2. Phase transitions between the QSH and insulating phases

The problem of interest in this paper is how the Z_2 topological number changes with a change of an external parameter. For the I-symmetric systems it is easier to consider; because it is the product of the parity, it can change when the valence band and the conduction band with opposite parities touch and exchange their roles. On the other hand, in I-asymmetric systems it is not obvious from (2) how the Z_2 topological number changes. For 2D, this was studied in the previous paper [19], using the homotopy characterization of the Z_2 topological number in [12]. It is defined in the similar way as the Chern integer [12], but with some modification. From this definition, it follows that the Z_2 topological number in 2D can change when the valence and conduction bands touch each other at some $\boldsymbol{k} = \pm\boldsymbol{k}_0$. (The band touching occurs

simultaneously at $\boldsymbol{k} = \pm\boldsymbol{k}_0$ because of the T-symmetry.) In 2D, the effective Hamiltonian in the vicinity of the phase transition reduces to

$$\mathcal{H} = E_0(m, k_x, k_y) \pm (m - m_0)\sigma_z + (k_x - k_{x0})\sigma_x + (k_y - k_{y0})\sigma_y, \tag{5}$$

after unitary and scale transformations, where m is an external parameter which controls the phase transition [19]. Equation (5) describes the simplest and general case for the band crossing, which occurs at $(m, k_x, k_y) = (m_0, k_{x0}, k_{y0})$.

This story should be modified when the spatial dimension is three, especially for I-asymmetric systems. This is because (5) cannot accommodate the four parameters m, k_x, k_y and k_z, in contrast with the 2D counterpart (5). In the following, we answer this question, by finding that a gapless phase should lie between the two gapped phases. This gapless phase is a topological phase in the following sense. The phase transition is governed by monopoles, namely band crossing between two non-degenerate bands. Such monopoles are topological objects which can appear or disappear not by themselves, but by creation/annihilation of a pair of a monopole and an antimonopole.

We first explain the framework for describing the phase transition in 3D. As in the previous paper [19], for the purpose of describing generic phase transitions between the QSH and the insulating phases, we consider a *single* parameter m which controls the phase transition. This parameter m may be considered as externally controllable, and by changing this parameter the system undergoes the phase transition. At the phase transition, the Z_2 topological number must change, which necessitates the closing of the gap at some wavevector $\boldsymbol{k}$. There are various kinds of band crossing when we introduce a number of parameters; nevertheless, to study the phase transition, we restrict ourselves to 'generic' band crossing, and exclude band crossing achieved only by tuning more than one parameter.

To be specific, we consider a Hamiltonian matrix

$$H(\boldsymbol{k}) = \begin{pmatrix} h_{\uparrow\uparrow}(\boldsymbol{k}) & h_{\uparrow\downarrow}(\boldsymbol{k}) \\ h_{\downarrow\uparrow}(\boldsymbol{k}) & h_{\downarrow\downarrow}(\boldsymbol{k}) \end{pmatrix}, \tag{6}$$

where $\boldsymbol{k} = (k_x, k_y, k_z)$. We assume that the spectrum of the Hamiltonian has no extra degeneracies other than those imposed by symmetry. We also assume that the Fermi energy E_{F} lies within a gap of the Hamiltonian. The T-symmetry implies,

$$H(\boldsymbol{k}) = \sigma_y H^{\mathrm{T}}(-\boldsymbol{k})\sigma_y, \tag{7}$$

i.e. $h_{\uparrow\uparrow}(\boldsymbol{k}) = h^{\mathrm{T}}_{\downarrow\downarrow}(-\boldsymbol{k})$, $h_{\uparrow\downarrow}(\boldsymbol{k}) = -h^{\mathrm{T}}_{\uparrow\downarrow}(-\boldsymbol{k})$ and $h_{\downarrow\uparrow}(\boldsymbol{k}) = -h^{\mathrm{T}}_{\downarrow\uparrow}(-\boldsymbol{k})$. The Kramers theorem guarantees that the band structure of such a T-symmetric spin-1/2 system is symmetric with respect to $\boldsymbol{k} \leftrightarrow -\boldsymbol{k}$. For the respective cases considered, it suffices to choose the dimension of the Hamiltonian matrix to be the number of states involved in band crossing.

In 3D, as well as in 2D, there is another symmetry which is crucial for the nature of the phase transition: the I-symmetry. This is because it is the only symmetry beside the T-symmetry which transforms between $\boldsymbol{k}$ and $-\boldsymbol{k}$ for all $\boldsymbol{k}$. We only consider these two (T- and I-) symmetries in this paper. Various kinds of band crossings found in higher point-group symmetries may be considered as degenerate cases of the generic cases considered here. From the Kramers theorem for T-symmetric systems, the spectrum is doubly degenerate at the eight points $\boldsymbol{k} = \boldsymbol{k}_i = \boldsymbol{G}/2$ $(i = 1, \ldots, 8)$ in I-asymmetric systems, and is so for every $\boldsymbol{k}$ in I-symmetric systems.

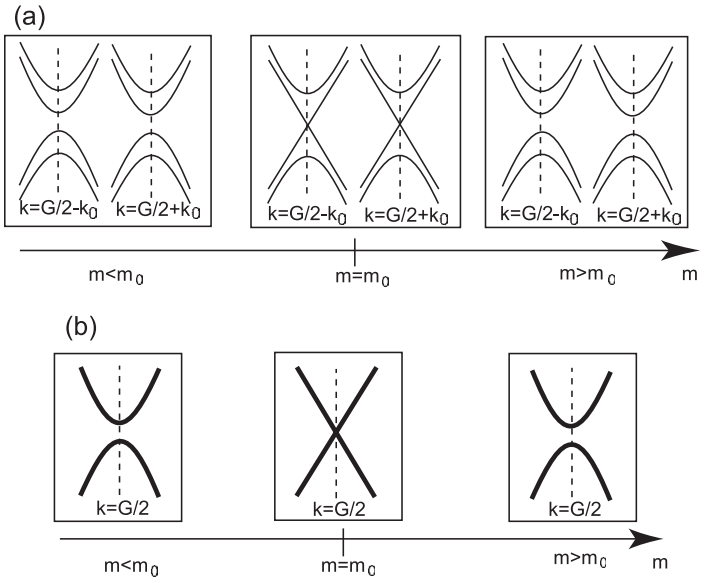

Figure 2. Phase transition in 2D between the QSH and insulating phases for (a) I-asymmetric and (b) I-symmetric cases. In the case (b) all the states are doubly-degenerate.

Let us explain the phase transitions in 2D as obtained in the previous paper [19]. In I-asymmetric systems, the band crossings occur at $\boldsymbol{k} = \pm\boldsymbol{k}_0 \neq \boldsymbol{G}/2$, between non-degenerate bands (figure 2). Because of the T-symmetry, band crossing occurs simultaneously at $\boldsymbol{k} = \pm\boldsymbol{k}_0$, at a single point $m = m_0$. On the other hand, in I-symmetric systems, the band crossings occur at $\boldsymbol{k} = \boldsymbol{k}_i = \boldsymbol{G}/2$ between two doubly degenerate bands. The two bands should have an opposite parity, and their parities are exchanged at the band crossing [19]. In the following we show that the phase transition in the I-symmetric systems are similar between 2D and 3D, whereas in the I-asymmetric systems they are quite different.

2.2.1. I-asymmetric systems. In 3D, in contrast with the 2D case, band crossing at $\boldsymbol{k} = \pm\boldsymbol{k}_0 \neq \boldsymbol{G}/2$ cannot lead to phase transition. The reason is the following. The energy bands for the I-asymmetric systems are non-degenerate for $\boldsymbol{k} \neq \boldsymbol{G}/2$. A crossing of two such energy bands has co-dimension three; namely, by tuning three parameters one can make two bands degenerate [20, 21]. To see this, let us consider a 2×2 Hamiltonian matrix

$$H = \begin{pmatrix} a & c \\ c^* & b \end{pmatrix}, \tag{8}$$

where a, b are real functions of $\boldsymbol{k}$ and m, and c is a complex function of $\boldsymbol{k}$ and m. A necessary condition for the two eigenvalues to be identical consists of three conditions $a = b$, $\mathrm{Re}\, c = 0$ and $\mathrm{Im}\, c = 0$, i.e. the co-dimension is three [20]. These three conditions determine a curve in the 4D space (m, k_x, k_y and k_z). Thus for generic m there will be in general a point (or points) $\boldsymbol{k}$ where the eigenvalues are degenerate. When m is changed continuously the $\boldsymbol{k}$ point moves in the $\boldsymbol{k}$ space, and the system remains gapless. It will be revealed later how the system can open a gap and run into either the QSH or the insulating phases.

On the other hand, for $\boldsymbol{k} = \boldsymbol{G}/2$, band crossing cannot occur in general. At the points $\boldsymbol{k} = \boldsymbol{G}/2$, the spectrum is doubly degenerate, and the co-dimension is five [22, 23]. It can be explicitly seen as follows. As the number of states involved is four, we consider a 4×4 Hamiltonian matrix with the constraint (7). This leads to a result

$$H(\boldsymbol{k} = \boldsymbol{k}_i) = E_0 + \sum_{i=1}^{5} a_i \Gamma_i, \tag{9}$$

where a_is and E_0 are real, and $\Gamma_1 = 1 \otimes \tau_x$, $\Gamma_2 = \sigma_z \otimes \tau_y$, $\Gamma_3 = 1 \otimes \tau_z$, $\Gamma_4 = \sigma_y \otimes \tau_y$ and $\Gamma_5 = \sigma_x \otimes \tau_y$. Its eigenenergies are given by $E_0 \pm \sqrt{\sum_{i=1}^{5} a_i^2}$. The two (doubly-degenerate) bands will touch when $a_i = 0$ for $i = 1, \ldots, 5$, which are not satisfied by tuning only one parameter m. (Note that the wavenumber $\boldsymbol{k}$ is fixed here and cannot be changed.) Thus, it is impossible to control the bands to touch at $\boldsymbol{k} = \boldsymbol{G}/2$ by tuning a single parameter m.

2.2.2. I-symmetric systems. In I-symmetric systems, the energies are doubly degenerate for every $\boldsymbol{k}$ by the Kramers theorem. The phase transition occurs when the gap between the two doubly-degenerate bands closes at some $\boldsymbol{k}$. Because there are four states involved, we consider the 4×4 Hamiltonian matrix $H(\boldsymbol{k})$. We impose the I-symmetry as

$$H(-\boldsymbol{k}) = P H(\boldsymbol{k}) P^{-1}, \quad u(-\boldsymbol{k}) = P u(\boldsymbol{k}), \tag{10}$$

where P is a unitary matrix independent of $\boldsymbol{k}$, and $u(\boldsymbol{k})$ is the periodic part of the Bloch wavefunction: $\varphi_{\boldsymbol{k}}(\boldsymbol{r}) = u(\boldsymbol{k}) \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}}$. After a judicious unitary transformation, all cases reduce to

$$P = \begin{pmatrix} P_\uparrow & \\ & P_\downarrow \end{pmatrix}, \quad P_\uparrow = P_\downarrow = \mathrm{diag}(\eta_a, \ \eta_b), \tag{11}$$

without losing generality. η_a and η_b represent the parity eigenvalues of the atomic orbitals involved.

The band crossings are different for $\eta_a = \eta_b$ and $\eta_a = -\eta_b$ in 3D, as is similar to 2D [19]. When $\eta_a = \eta_b = \pm 1$, the generic Hamiltonian becomes

$$H(\boldsymbol{k}) = E_0(\boldsymbol{k}) + \sum_{i=1}^{5} a_i(\boldsymbol{k}) \Gamma_i, \tag{12}$$

where a_is and E_0 are real even functions of $\boldsymbol{k}$. On the other hand, when $\eta_a = -\eta_b = \pm 1$, the Hamiltonian reads,

$$H(\boldsymbol{k}) = E_0(\boldsymbol{k}) + a_5(\boldsymbol{k}) \Gamma'_5 + \sum_{i=1}^{4} b^{(i)}(\boldsymbol{k}) \Gamma'_i, \tag{13}$$

where $E_0(\boldsymbol{k})$ and $a_5(\boldsymbol{k})$ are even functions of $\boldsymbol{k}$, $b^{(i)}(\boldsymbol{k})$ are odd functions of $\boldsymbol{k}$. The matrices $\Gamma'_1 = \sigma_z \otimes \tau_x$, $\Gamma'_2 = 1 \otimes \tau_y$, $\Gamma'_3 = \sigma_x \otimes \tau_x$, $\Gamma'_4 = \sigma_y \otimes \tau_x$ and $\Gamma'_5 = 1 \otimes \tau_z$ form the Clifford algebra. Therefore, for a generic point, $\boldsymbol{k} = \boldsymbol{k}_i \neq \boldsymbol{G}/2$. In both cases, $\eta_a = \eta_b$ and $\eta_a = -\eta_b$, the co-dimension is five, which exceeds the number of tunable parameters (m, k_x, k_y and k_z). Therefore, band crossing does not occur at a generic point $\boldsymbol{k}$ with $\boldsymbol{k} \neq \boldsymbol{k}_i \equiv \boldsymbol{G}/2$.

On the other hand, at the high-symmetry points $\boldsymbol{k} = \boldsymbol{k}_i = \boldsymbol{G}/2$. The number of parameters to achieve degeneracy is five for $\eta_a = \eta_b$ (in (12)) while it is one for $\eta_a = -\eta_b$ (in (13)). Because the wavenumber is fixed $\boldsymbol{k} = \boldsymbol{k}_i = \boldsymbol{G}/2$, there is only one changeable parameter m. Thus only

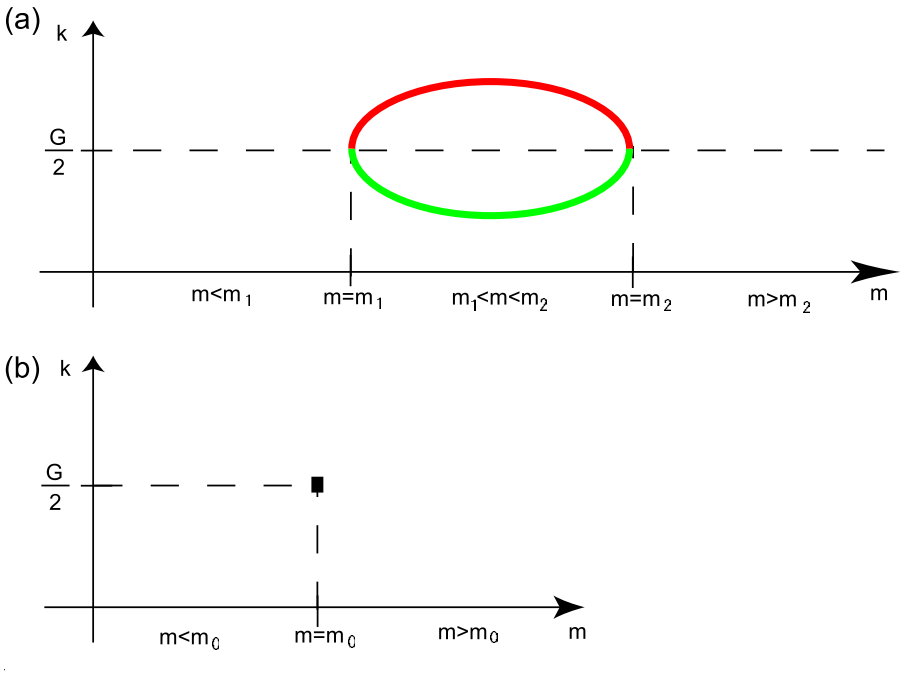

Figure 3. Location of the gapless points by changing the external parameter m in (a) I-asymmetric systems and (b) I-symmetric systems. In (a) the green and the red denotes trajectories of the monopole and antimonopole, respectively.

when $\eta_a = -\eta_b$, the two-doubly degenerate bands touch at $\boldsymbol{k} = \boldsymbol{k}_i = \boldsymbol{G}/2$. This situation is similar to the case in 2D [19].

Thus, we have seen that in I-symmetric 3D systems, the phase transition can occur at a single value of the parameter $m = m_0$. At one side (e.g. $m < m_0$), the system is in the QSH, while the other side (e.g. $m > m_0$) is in the insulating phase. Meanwhile, the phase transition at a single value of m is in general impossible in I-asymmetric cases. Here we encounter a question. How is the phase transition in the I-symmetric 3D system modified when some perturbation breaks I-symmetry. Because the two sides ($m \ll m_0$ and $m \gg m_0$) belong to the different phases, there should be a phase transition in between.

The answer is the following. Instead of a phase transition occurring at a single value of the parameter m, there appears a finite region of m where the system remains gapless. As was discussed previously, the band crossing in the I-asymmetric system cannot occur at a single value of m. The gapless points in the 4D m-$\boldsymbol{k}$ space form a 1D manifold (i.e. a curve). Thus, the only way to close the gap by changing m is to make a pair of gapless points, as shown in figure 3.

To see the behavior of such gapless points, we note that each gapless point carries a topological number. Such a gapless point, sometimes called a diabolical point, is regarded as a monopole in the $\boldsymbol{k}$ space [24]–[26]. Indeed, the Bloch wavefunctions with non-degenerate spectrum can be associated with a $U(1)$ gauge structure in the $\boldsymbol{k}$ space;

$$\boldsymbol{A}_n(\boldsymbol{k}) = -\mathrm{i}\langle \boldsymbol{k}n|\nabla_{\boldsymbol{k}}|\boldsymbol{k}n\rangle, \tag{14}$$

$$\boldsymbol{B}_n(\boldsymbol{k}) = \nabla_{\boldsymbol{k}} \times \boldsymbol{A}_n(\boldsymbol{k}), \tag{15}$$

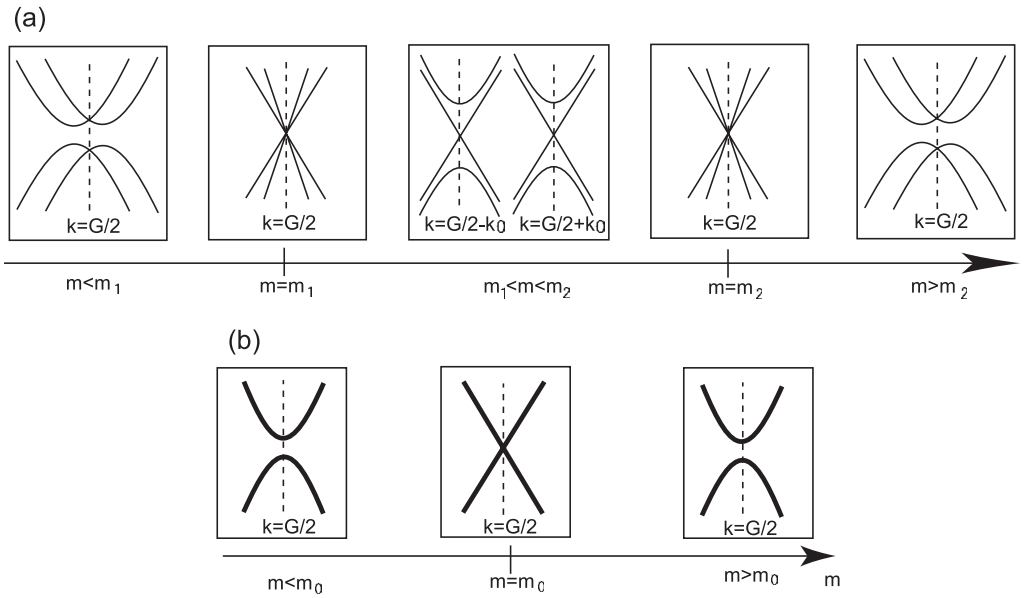

Figure 4. Phase transition in 3D between the QSH and insulating phases for (a) I-asymmetric and (b) I-symmetric cases. In the case (b) all the states are doubly degenerate.

$$\rho_n(\boldsymbol{k}) = \frac{1}{2\pi}\nabla_{\boldsymbol{k}} \cdot \boldsymbol{B}_n(\boldsymbol{k}). \tag{16}$$

The quantity $\rho(\boldsymbol{k})$ is called a monopole density. It vanishes when the nth band is not degenerate with other bands. At the band crossings between the nth band and other bands, $\rho(\boldsymbol{k})$ has a δ-function singularity; $\rho_n(\boldsymbol{k}) = \sum_l q_{ln}\delta(\boldsymbol{k} - \boldsymbol{k}_{ln})$, where q_{ln} is an integer called a monopole charge. For example, for the band crossing at $\boldsymbol{k} = \boldsymbol{k}_0$ with linear dispersion (Weyl fermion)

$$\mathcal{H} = E_0(\boldsymbol{k}) + \sum_{i=1}^{3} f_i(\boldsymbol{k})\sigma_i, \tag{17}$$

where $f_i(\boldsymbol{k} = \boldsymbol{k}_0) = 0$, the monopole charge for the lower band is $\mathrm{sgn}(\det(\frac{\partial f_i}{\partial k_j})_{ij})|_{\boldsymbol{k}=\boldsymbol{k}_0}$ $(= \pm 1)$.

In the present system, a pair of a monopole (charge $q = 1$) and an antimonopole (charge $q = -1$) is created at $m = m_1$, $\boldsymbol{k} = \boldsymbol{k}_i = \boldsymbol{G}/2$ when m is increased, and the system becomes gapless. When m is increased further, the monopole ($\boldsymbol{k} = \boldsymbol{k}_+$) and the antimonopole ($\boldsymbol{k} = \boldsymbol{k}_-$) moves in the $\boldsymbol{k}$ space, while the T-symmetry imposes that $\boldsymbol{k}_- = -\boldsymbol{k}_+$. This system can open a gap again only when the monopole and antimonopole annihilate together. The annihilation can occur only at the points with $\boldsymbol{k} = \boldsymbol{k}_i = \boldsymbol{G}/2$, again by the T-symmetry (figure 3). Thus, the overall feature of the phase transition is schematically expressed as in figure 4.

2.2.3. Summary for the phase transition in 3D. To summarize, the overall feature of the band crossing is schematically shown in figure 4. The overall phase diagram is schematically shown in figure 5, in a plane of the control parameter m and another parameter δ representing an I-symmetry breaking. As we have seen, when the I-symmetry is broken, the topological gapless

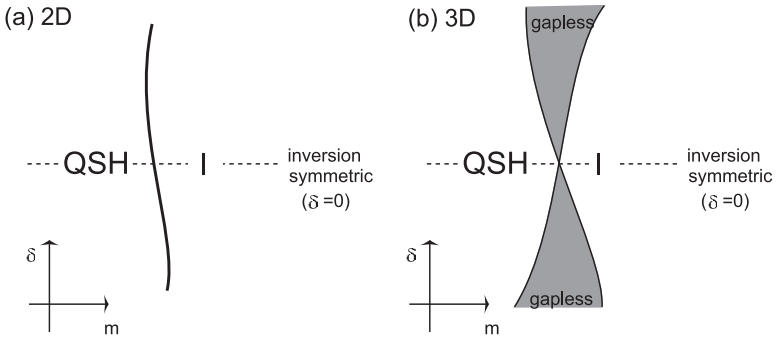

Figure 5. Phase diagram for the QSH and ordinary insulating (I) phases for (a) 2D and (b) in 3D. m is a control parameter which drives the phase transition, and δ represents a parameter which describes the breaking of I-symmetry. $\delta = 0$ is the case with I-symmetry.

phase appears between the two phases which are gapful in the bulk. The difference of the Z_2 topological numbers ν between the two sides of the phase transition can also be calculated as in the 2D case [19]. When the band crossing occurs at $\boldsymbol{k}_i = \frac{1}{2}(n_1\boldsymbol{b}_1 + n_2\boldsymbol{b}_2 + n_3\boldsymbol{b}_3)$, the factor $\delta_{i=(n_1n_2n_3)}$ in (4) changes sign, and some of the four Z_2 topological numbers, ν_0; $(\nu_1\nu_2\nu_3)$ in (4), change accordingly. This applies to the I-symmetric systems. The I-asymmetric cases are similarly treated because they can be associated with the I-symmetric cases with perturbation.

3. Towards materials search for QSH systems

3.1. Bismuth thin film and 2D QSH system

In [15], the bilayer bismuth is studied as a candidate for the 2D QSH phase. From the 2D bilayer tight-binding model, truncated from the 3D tight-binding model [27], the bilayer bismuth is proposed to be a 2D QSH system [15]. Bilayer antimony is studied in a similar way, and it is an ordinary insulator. It is predicted from the calculation of Z_2 topological number and from a band structure calculation for the geometry with edges (i.e. the strip geometry). The Z_2 topological number is calculated in [15] by the Pfaffian of the matrix for the time-reversal operator proposed in [8].

The calculation of the Pfaffian involves fixing of phases of the wavefunction as an analytic function of the wavenumber $\boldsymbol{k}$, which is numerically a challenging problem, even for a simplified model presented in [15]. It can be tackled by discretizing the $\boldsymbol{k}$ space and counting the vortex of the Pfaffian matrix [28, 29]. Instead, for I-symmetric systems, the method of calculating parity eigenvalues (3) proposed in [18] is much easier. For bilayer bismuth, we checked that this method leads us also to the same conclusion that the Z_2 topological number is odd and nontrivial, and it is in the QSH phase.

For the system to be in the QSH phase, it should have a gap in the bulk. The 3D bulk bismuth is semimetallic, and has a small band overlap between the conduction and the valence band, while the direct gap is finite for all wavenumbers. By making it into a thin film, the perpendicular motion is quantized and tends to open a band gap. Earlier theoretical estimates [30, 31] and experiments [32, 33] (see also [34]) show that thin-film

bismuth thinner than ∼30 nm becomes an insulator. However, in recent experiments by Hirahara *et al* [35] the angle-resolved photoemission spectra (ARPES) of ultrathin films of ten bilayers are measured. In their experimental data, as opposed to the earlier theories and experiments, the system remains semimetallic. This poses the question that the critical thickness might be much smaller than the earlier estimate [30]–[33].

3.2. Bismuth and 3D QSH system

In [18], it was suggested that 3D Bi has $\nu =$ even, while 3D Sb has $\nu =$ odd. This looks opposite to the case of 2D. To resolve this, numerical analysis was performed in [29] by artificially changing the interlayer hopping multiplied by a factor f $(0 < f < 1)$. It was found that for Bi, the regions $0 \leqslant f < 0.223$, $0.223 < f < 0.993$ and $0.993 < f \leqslant 1$ are the phases 0;(111), 1;(111), and 0;(000), respectively, whereas for Sb the regions $0 \leqslant f < 0.54$ and $0.54 < f \leqslant 1$ are the phases 0;(000) and 1;(111). This calculation was done by the calculation of the Pfaffian matrix, while this result was checked by the parity analysis proposed in [18]. The phase transitions by changing f are described by our theory developed in the previous section. The band crossing which accompanies the phase transition occurs for Bi at $(0, 0, 0)$ for $f = 0.223$, and at $(\pi, 0, 0)$, $(0, \pi, 0)$ and $(0, 0, \pi)$ for $f = 0.993$. For Sb, the band crossing occurs at (π, π, π) for $f = 0.54$. These systems are I-symmetric. According to our theory, the band crossing occurs only at a single value of the control parameter f, rather than having a gapless phase in between.

In reality, both Bi and Sb are semimetals, not insulators. The Z_2 topological numbers are defined in 3D, by assuming that the band overlap is lifted by some perturbation, thereby the bands below the gap are regarded as the 'valence band' which enters in the definition of the Z_2 topological numbers. This is possible because there is a direct gap in every $\boldsymbol{k}$. In this sense, although the 3D Z_2 topological number ν_0 defined as such is odd (nontrivial) in Sb, it is not the QSH phase, because there is no gap. If one can open a gap in the 3D Sb by some external perturbation, it becomes the QSH phase.

Nevertheless, because the Z_2 topological numbers can be defined in Bi and Sb in the above sense, they manifest themselves in the spectrum of surface states. The 3D bulk Bi and Sb has the topological numbers 0;(000) and 1;(111), from which the expected topology of the Fermi surface of the surface states can be easily sorted out for various directions of crystal surface [18]. The results can be compared with the experiments on ARPES for Bi (for example [36]) and for Sb [37]. To interpret these experiments to see whether it matches the prediction from the Z_2 topological number, we need to separate the Fermi surface of bulk states and that of surface states, which is not trivial experimentally.

To find a clear experimental manifestation for the nontrivial Z_2 topological number, we need to open a gap in the bulk. One example is $Bi_{1-x}Sb_x$ $(0.07 < x < 0.22)$, and ARPES experiments on this doping region are called for. To our knowledge [38] is the only experimental report in this doping region. Nonetheless, this experiment is limited only to the vicinity of the $\boldsymbol{k} = 0$, and is not sufficient to examine its topological phase from the ARPES data. Data on ARPES in $Bi_{1-x}Sb_x$ $(0.07 < x < 0.22)$ covering the whole surface Brillouin zone would be clear evidence for the QSH phase. Thus, there is no clear experimental evidence for the QSH phase as yet, whereas it looks well within our reach, and a good candidate is strongly called for.

3.3. Criterion for searching QSH systems

Apart from the T-symmetry, there are only two conditions for the QSH phases.

1. The bulk is a band insulator.
2. The Z_2 topological number is odd.

The first condition is clear, while the second requires calculation. At this stage, we need some strategy to search among the vast number of nonmagnetic insulators.

To find out the strategy, let us begin with a system without spin–orbit coupling, and switch on the spin–orbit coupling gradually. Insulators without the spin–orbit coupling have a trivial (i.e. even) Z_2 topological number. To reach the QSH phase, the gap should close in switching on the spin–orbit coupling, thereby the system should undergo a phase transition. This phase transition should be described within our theory developed in the previous section.

Therefore, the QSH phase near the phase transition may have the direct gap at $\boldsymbol{k} = \boldsymbol{k}_i = \boldsymbol{G}/2$, as a trace of the phase transition. When this phase transition occurs simultaneously at some points in the Brillouin zone, the changes of the Z_2 topological number add together. In this sense, when the material has an even number of equivalent points for the direct gaps which close at the phase transition, the Z_2 topological number ν_0 does not change at the transition, and the system remains an ordinary insulator. The example is the four equivalent L points in PbTe [18], and PbTe is indeed in the ordinary insulator phase.

Therefore, it is desirable to have direct gaps at an odd number of points in the Brillouin zone, such as the L points in $Bi_{1-x}Sb_x$ [18], or the Γ point in generic crystals. We note that this is not a necessary condition, but can be a reasonable guideline for searching candidate materials among nonmagnetic insulators.

It is also necessary to see whether the gap in the material considered is 'after' or 'before' the phase transition, when one turns on the spin–orbit coupling gradually. In other words, for the QSH phase, the gap should be originated from the spin–orbit coupling. This statement is somewhat vague; to make this more transparent, the author suggested that the susceptibility can be a measure to see whether the gap is of spin–orbit nature [15]. For example, bismuth is strongly diamagnetic, because of the inter-band matrix elements between the conduction and the valence bands due to the spin–orbit coupling. This class of materials, when gapped, are good candidates for the QSH phases. The criteria discussed so far will be useful for finding good candidates for the QSH phase.

4. Conclusions and discussions

In the present paper, we studied the phase transition between the QSH and the insulating phases in 3D. In contrast to the 2D systems, in the 3D I-asymmetric cases, there is a gapless phase between the two phases. This gapless phase originates from the topological nature of the monopoles (band crossing points) in 3D, which are responsible for the phase transition. Furthermore, the gap closing occurs only at the high-symmetry points $\boldsymbol{k} = \boldsymbol{G}/2$, which is also in contrast with 2D cases.

So far we have given topological argument in terms of the Bloch wavefunctions. One may suspect that the scenario may become invalid in the presence of interaction or disorder, thereby the topological order may be obscured. We give here qualitative argument that our

scenario is robust against small perturbations. Our argument is based on the fact that the Z_2 topological number (for gapped systems) can be defined even in the presence of interaction and disorder [18], by use of an analogue of the Laughlin's gedanken experiment [39]. In disordered systems, the Bloch wavenumber k_x becomes ill-defined. We then think of folding the system to a 'ring', periodic in the x-direction, and threading a flux Φ_x into its hole. The flux Φ_x plays the role of k_x. A similar procedure is taken for k_y and k_z. The Z_2 topological number is defined by using Φ_i $(i = x, y, z)$ instead of k_i. Therefore, as long as interaction and disorder are relatively weak, the bulk remains gapped and the Z_2 topological number remains well-defined and identical to that in the clean, noninteracting systems. As is related with this stability of the topological order, the gapless surface state in the 3D QSH phase remains gapless and shows antilocalization behavior in the presence of impurities. It is in the symplectic universality class, and is similar to the honeycomb lattice without intervalley scattering [40]. For this reason, the 3D QSH phase is robust against impurities.

Based on these observations, we can now see that the gapless topological phase ($m_1 < m < m_2$ in figure 4(a)) will survive in the presence of disorder and interaction. Since the two sides of the topological gapless phase are the gapped phases ($m < m_1$ and $m > m_2$ in figure 4(a)) with different Z_2 topological numbers, there should lie a gapless phase in between, even in the presence of interaction and disorder. (If the topological gapless phase becomes gapped, it can no longer accompany a phase transition.) A quantitative argument to see how this gapless phase is robust is involved and is beyond the scope of the present paper.

A following remark is in order. A calculation for the Z_2 topological number based on a simplified model requires some care, because the bands well below the Fermi energy might contribute to the Z_2 topological number. We present one example. In [7], we constructed a four-band tight-binding model for HgTe. It is a zero-gap system, but by opening a gap with uniaxial pressure, it becomes a band insulator. The question here is whether it is a simple insulator or a topological insulator (i.e. QSH). In fact, from the parities of the bands below the Fermi energy E_F from the four-band model in [7], the Z_2 topological number ν turns out to be even. This disagrees with the result in [18]. The reason for the disagreement is the following. This model only has four bands near the Fermi energy E_F, and the other bands well below E_F are discarded. In HgTe, one doubly-degenerate band just below the Fermi energy has a trivial Z_2 topological number, while the other discarded bands deep below the Fermi energy have a nontrivial Z_2 topological number; this leads to the seemingly contradicting results. Thus, in the calculation of the Z_2 topological number, one should be careful because it involves all the bands below the Fermi energy E_F.

The lesson we can learn from the examples of bismuth [15] and other materials [18] is that materials with odd Z_2 topological numbers are not rare. In nature, there may well exist systems with odd Z_2 topological number. The crucial difference between the two kinds of topological insulators, namely the QH and QSH systems, is that QH systems require a magnetic field as strong as several teslas, whereas the QSH systems do not require external fields to achieve topological phases. This is a crucial difference. To realize the QSH phase, the effective magnetic field produced from the spin–orbit coupling should be several tesla, as expected from the analogy with the QH system. We see from the examples of bismuth [15] and other materials [18] that the spin–orbit coupling is strong enough in some materials. Thus, it is natural to expect some materials in nature to be in the QSH phase, which would be of great interest both theoretically and experimentally.

Acknowledgments

The author would like to thank Y Avishai, T Hirahara, S Iso, N Nagaosa, M Onoda and S-C Zhang for fruitful discussions and helpful comments. This research is supported in part by Grant-in-Aid and NAREGI Nanoscience Project from the Ministry of Education, Culture, Sports, Science and Technology of Japan.

References

[1] Murakami S, Nagaosa N and Zhang S C 2003 *Science* **301** 1348
[2] Sinova J, Culcer D, Niu Q, Sinitsyn N A, Jungwirth T and MacDonald A H 2004 *Phys. Rev. Lett.* **92** 126603
[3] Saitoh E, Ueda M, Miyajima H and Tatara G 2006 *Appl. Phys. Lett.* **88** 185209
[4] Stern N P, Ghosh S, Xiang G, Zhu M, Samarth N and Awschalom D D 2006 *Phys. Rev. Lett.* **97** 126603
[5] Kimura T, Otani Y, Sato T, Takahashi S and Maekawa S 2007 *Phys. Rev. Lett.* **98** 156601
[6] Guo G Y, Murakami S, Chen T W and Nagaosa N 2007 Intrinsic spin Hall effect in platinum metal *Preprint* 0705.0409
[7] Murakami S, Nagaosa N and Zhang S C 2004 *Phys. Rev. Lett.* **93** 156804
[8] Kane C L and Mele E J 2005 *Phys. Rev. Lett.* **95** 146802
[9] Kane C L and Mele E J 2005 *Phys. Rev. Lett.* **95** 226801
[10] Bernevig B A and Zhang S C 2006 *Phys. Rev. Lett.* **96** 106802
[11] Fu L, Kane C L and Mele E J 2007 *Phys. Rev. Lett.* **98** 106803
[12] Moore J E and Balents 2007 *Phys. Rev.* B **75** 121306
[13] Wu C, Bernevig B A and Zhang S C 2006 *Phys. Rev. Lett.* **96** 106401
[14] Xu C and Moore J E 2006 *Phys. Rev.* B **73** 045322
[15] Murakami S 2006 *Phys. Rev. Lett.* **97** 236805
[16] Bernevig B A, Hughes T L and Zhang S C 2006 *Science* **314** 1757
[17] Fu L and Kane C L 2006 *Phys. Rev.* B **74** 195312
[18] Fu L and Kane C L 2006 *Phys. Rev.* B **76** 045302
[19] Murakami S, Iso S, Avishai Y, Onoda M and Nagaosa N 2007 Tuning phase transition between quantum spin Hall and ordinary insulating phases *Preprint* 0705.3696
[20] von Neumann V J and Wigner E 1929 *Phys. Z.* **30** 467
[21] Herring C 1937 *Phys. Rev.* **52** 365
[22] Avron J E, Sadun L, Segert J and Simon B 1988 *Phys. Rev. Lett.* **61** 1329
[23] Avron J E, Sadun L, Segert J and Simon B 1989 *Commun. Math. Phys.* **124** 595
[24] Volovik G E 1987 *JETP Lett.* **46** 98
[25] Volovik G E 2001 *Phys. Rep.* **351** 195
[26] Murakami S and Nagaosa N 2003 *Phys. Rev. Lett.* **90** 057002
[27] Liu Y and Allen R E 1995 *Phys. Rev.* B **52** 1566
[28] Fukui T and Hatsugai Y 2007 *Phys. Rev.* B **75** 121403
[29] Fukui T and Hatsugai Y 2006 Quantum spin Hall effect in three dimensional materials: lattice computation of Z_2 topological invariants and its application to Bi and Sb *Preprint* cond-mat/0611423
[30] Lutskii V N 1965 *Pis. Zh. Eksp. Teor. Fiz.* **2** 391
Lutskii V N 1965 *JETP Lett.* **2** 245 (Engl. Transl.)
[31] Sandomirskii V B 1967 *Zh. Eksp. Teor. Fiz.* **52** 158
Sandomirskii V B 1967 *Sov. Phys. —JETP* **25** 101 (Engl. Transl.)
[32] Hoffman C A, Meyer J R, Bartoli F J, Di Venere A, Yi X J, Hou C L, Wang H C, Ketterson J B and Wong G K 1993 *Phys. Rev.* B **48** 11431
[33] Hoffman C A, Meyer J R, Bartoli F J, Di Venere A, Yi X J, Hou C L, Wang H C, Ketterson J B and Wong G K 1995 *Phys. Rev.* B **51** 5535

[34] Chu H T 1995 *Phys. Rev.* B **51** 5532
[35] Hirahara T, Nagao T, Matsuda I, Bihlmayer G, Chulkov E V, Koroteev Yu M, Echenique P M, Saito M and Hasegawa S 2006 *Phys. Rev. Lett.* **97** 146803
[36] Koroteev Yu M, Bihlmayer G, Gayone J E, Chulkov E V, Blugel S, Echenique P M and Hofmann Ph 2004 *Phys. Rev. Lett.* **93** 046403
[37] Sugawara K, Sato T, Souma S, Takahashi T, Arai M and Sasaki T 2006 *Phys. Rev. Lett.* **96** 046411
[38] Höchst H and Gorovikov S A 2005 *J. Electron Spectrosc. Relat. Phenom.* **144–7** 351
[39] Laughlin R B 1981 *Phys. Rev.* B **23** 5632
[40] Suzuura H and Ando T 2002 *Phys. Rev. Lett.* **89** 266603

New Journal of Physics
The open-access journal for physics

Corrigendum

Phase transition between the quantum spin Hall and insulator phases in 3D: emergence of a topological gapless phase

Shuichi Murakami 2007 *New J. Phys.* **9** 356

New Journal of Physics **10** (2008) 029802
Online at http://www.njp.org/
doi:10.1088/1367-2630/10/2/029802

The discussion on page 9 including figures 3 and 4 needs to be corrected. The last paragraph ('In the present system...') of section 2.2.2 should be replaced by the following.

'When the system becomes gapless, a monopole (charge $q = 1$) and an antimonopole ($q = -1$) are created in pairs. Because of the T-symmetry, the distribution of monopole charges is symmetric with respect to $\boldsymbol{k} = \boldsymbol{G}/2$. Hence for the simplest case, two monopole–antimonopole pairs are created at $\boldsymbol{k} = \pm\boldsymbol{k}_0 + \boldsymbol{G}/2$ ($\boldsymbol{k}_0 \neq 0$) simultaneously when $m = m_1$, and the system becomes gapless. When m is increased further, the monopoles and antimonopoles move in the $\boldsymbol{k}$ space, while the distribution of the monopole charges remains symmetric with respect to $\boldsymbol{G}/2$. This system can open a gap again only when all the monopoles and antimonopoles annihilate in pairs. This occurs at $m = m_2$ as shown in figure 6. Thus the overall feature of the phase transition is schematically expressed as in figure 7.'

Correspondingly, figures 3 and 4 should be replaced by figure 6 and figure 7 respectively. These corrections do not affect the main conclusions of the article.

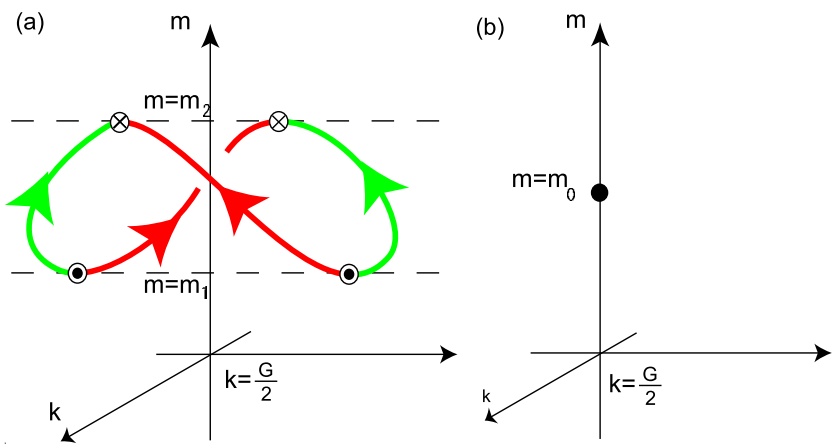

Figure 6. Location of the the gapless points by changing the external parameter m in (a) I-asymmetric systems and (b) I-symmetric systems. In (a) the green and the red denotes trajectories of the monopole and antimonopole, respectively.

1367-2630/08/029802+02$30.00

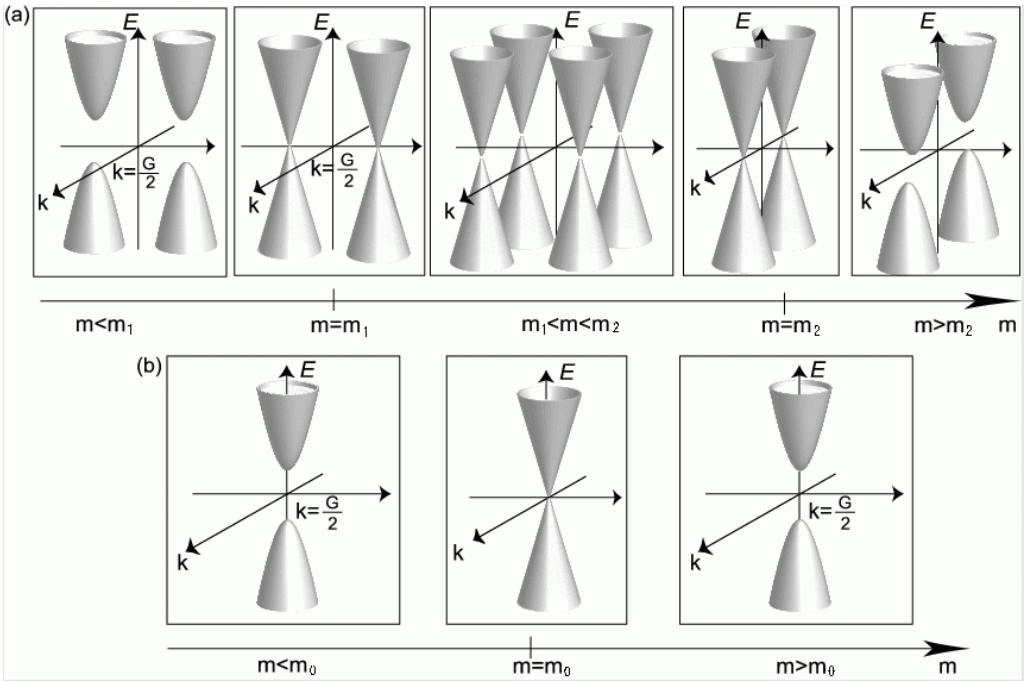

Figure 7. Phase transition in 3D between the quantum spin Hall (QSH) and insulating phases for (a) I-asymmetric and (b) I-symmetric cases. In the case (b) all the states are doubly degenerate.