Title	Spherical topological insulator
Auther(s)	Imura, Ken-Ichiro; Yoshimura, Yukinori; Takane, Yositake; Fukui, Takahiro
Citation	Physical Review B , 86 (23) : 235119
Issue Date	2012
DOI	10.1103/PhysRevB.86.235119
Self DOI	
URL	http://ir.lib.hiroshima-u.ac.jp/00034062
Right	(c) 2012 American Physical Society
Relation	

Spherical topological insulator

Ken-Ichiro Imura,[1] Yukinori Yoshimura,[1] Yositake Takane,[1] and Takahiro Fukui[2]
[1]*Department of Quantum Matter, AdSM, Hiroshima University, Higashi-Hiroshima 739-8530, Japan*
[2]*Department of Physics, Ibaraki University, Mito 310-8512, Japan*
(Received 22 May 2012; published 13 December 2012)

The electronic spectrum on the spherical surface of a topological insulator reflects an *active* property of the helical surface state that stems from a constraint on its spin on a curved surface. The induced spin connection can be interpreted as an effective vector potential associated with a fictitious magnetic monopole induced at the center of the sphere. The strength of the induced magnetic monopole is found to be $g = \pm 2\pi$, being the smallest finite (absolute) value compatible with the Dirac quantization condition. We have established an explicit correspondence between the bulk Hamiltonian and the effective Dirac operator on the curved spherical surface. An explicit construction of the surface spinor wave functions implies a rich spin texture possibly realized on the surface of topological insulator nanoparticles. The electronic spectrum inferred by the obtained effective surface Dirac theory, confirmed also by the bulk tight-binding calculation, suggests a specific photoabsorption/emission spectrum of such nanoparticles.

DOI: 10.1103/PhysRevB.86.235119 PACS number(s): 73.20.−r, 73.63.Fg, 72.25.Mk, 61.72.Lk

I. INTRODUCTION

It was only several years ago that the idea of a topological insulator had been proposed as a possible candidate for the new state of matter in the field of condensed matter.[1] The original theoretical idea has already been extended in various aspects, made applicable to a broader range of phenomena, including superconductivity and superfluidity.[2,3] The related research areas are now reclassified and recognized as that of the topological quantum phenomena. Naturally the outbreak of this new research field owes much to a rapid success of experimental studies that have demonstrated that the new theoretical idea has much reality.[4]

The existence of a single gapless Dirac cone in its surface spectrum is a hallmark of strong topological insulators. Here we focus on a specific property of this robust and protected surface state on a curved surface,[5] the "spin-to-surface locking." It is indeed specific to the topological insulator surface state and distinguishes it from other realizations of gapless Dirac cones in condensed matter such as in graphene[6,7] and related carbon materials. The role of spin-to-surface locking may be most accentuated in the (pseudocylindrical) wire-shaped geometry in which an anomalous Aharonov-Bohm type of oscillation has been reported.[8] Motivated by the reality of such transport measurements which may allow for a direct observation of the spin Berry phase, theorists have extensively studied the role of this phenomenon in the transport characteristics of the surface state.[9–13]

A remarkable consequence of the spin-to-surface locking in the cylindrical geometry is the *half-integral* quantization of the *orbital* angular momentum. Clearly such half-integral quantization leads to an appearance of a finite-size energy gap in the hitherto gapless surface electronic spectrum. Interestingly, introduction of a physical magnetic flux of half of a unit flux quantum through (piercing) the cylinder compensates the Berry phase associated with the spin-to-surface locking, and closes the gap. The same mechanism applies to the classification of gapless electronic states bound to a crystal dislocation line penetrating an otherwise surfaceless sample of a three-dimensional topological insulator.[14] A more systematic consideration[15] on such gapless electronic states associated with a topological defect in a topological mother system has been developed from the viewpoint of classifying topological insulators and superconductors in a unified way solely from their symmetry class.[16–19]

The specificity of the cylindrical surface is that it is *flat* in the sense that it has everywhere a vanishing Gaussian curvature. On the surface of a topological insulator of more generic shape or geometry yielding a finite curvature, the effect of spin-to-surface locking mentioned earlier will be modified by that of a finite curvature. A spherical surface of a topological insulator[20] is a prototypical example in which such an interplay is expected. We show in this paper that the two effects are both expressed in terms of a Berry phase, but of contrasting nature (see Table I). The two types of Berry phase both contribute to the formation of a finite-size energy gap. The resulting surface electronic spectrum on the sphere is shown to have a substantial compatibility with the result of tight-binding calculations performed for a cubic system (for tight-binding calculation involving the bulk, cubic implementation is much more straightforward). A related but different scenario on the fate of such a (planar) gapless Dirac cone embedded on the curved spherical surface has been proposed in the study of the electronic states in fullerene.[21–25]

In addition to the spectrum, the structure of the surface spinor wave function is another highlight of the paper. On the curved spherical surface of a topological insulator the strong spin-orbit coupling in the bulk, combined with the twisting of the phase shift due to the two types of Berry curvature, leads to a nontrivial *spin texture*. By explicitly constructing the surface spinor wave function we reveal such a rich spin texture possibly realized on the surface of topological insulator nanoparticles.

The paper is organized as follows. In Sec. II the effective surface Dirac theory is derived from the gapped bulk Hamiltonian, in which two types of Berry phase appear. The nature of these two types of Berry phase is discussed and contrasted in Sec. III. The solution of the effective surface Dirac equation is given explicitly in Sec. IV. The surface wave function is shown to be expressed in terms of the Jacobi's polynomials.

TABLE I. Characterization of the two types of Berry phase manifesting on the surface of a spherical topological insulator.

Type	(A)	(B)
(Geometrical) origin	Curvature in the polar (θ) direction; effect of closing the surface at the north and south poles	Effect of rolling the surface in the azimuthal (ϕ) direction
Appearance (where, how)	In the covariant derivatives, or $\partial_\theta \to \partial_\theta + \frac{1}{2}\cot(\theta/2)$	$-i\partial_\phi \to -i\partial_\phi + 1/2$
Shifting the spectrum?	Yes	Yes
Relation to spin-to-surface locking	Breaks the locking	Expression of the (tendency to) spin-to-surface locking
Sensitivity to the choice of basis	No	Yes
Other examples?	Fullerene (buckyball)	Cylindrical TI

The obtained discrete energy spectrum is compared with the result of (bulk) tight-binding calculation in Sec. V. This leads us to our conclusions. Some details of the formulation are left to the Appendices.

II. DERIVATION OF THE SURFACE EFFECTIVE HAMILTONIAN

Let us first derive an effective "Dirac operator" on the spherical surface, starting with a bulk Hamiltonian. Our starting point is the following gapped bulk effective Hamiltonian[26,27] in the continuum limit:

$$H_{\text{bulk}} = \epsilon(\boldsymbol{p})\mathbf{1} + m(\boldsymbol{p})\tau_z + A\tau_x(p_x\sigma_x + p_y\sigma_y + p_z\sigma_z), \quad (1)$$

describing a three-dimensional (3D) $\mathbb{Z}_2$ topological insulator, where

$$m(\boldsymbol{p}) = m_0 + m_2\left(p_x^2 + p_y^2 + p_z^2\right) \quad (2)$$

is a (generalized) mass term containing both the constant and quadratic (Wilson) terms. For simplicity we have chosen the Wilson term to be isotropic. The two types of Pauli matrices $\boldsymbol{\sigma} = (\sigma_x,\sigma_y,\sigma_z)$ and $\boldsymbol{\tau} = (\tau_x,\tau_y,\tau_z)$ represent, respectively, the real and orbital spin degrees of freedom, and $\mathbf{1}$ is the 4×4 identity matrix. The Hamiltonian (1) is time-reversal invariant, that is, invariant under the operation of $\Theta = i\sigma_y K$, where K represents complex conjugation. The two types of Pauli matrices represent independent degrees of freedom acting on spinors living in a different space. To make this point explicit one may express Eq. (1) in the following 4×4 matricial form,

$$H_{\text{bulk}} = \epsilon(\boldsymbol{p})\mathbf{1} + \begin{bmatrix} m(\boldsymbol{p}) & Ap_z & 0 & Ap_- \\ Ap_z & -m(\boldsymbol{p}) & Ap_- & 0 \\ 0 & Ap_+ & m(\boldsymbol{p}) & -Ap_z \\ Ap_+ & 0 & -Ap_z & -m(\boldsymbol{p}) \end{bmatrix}, \quad (3)$$

where $p_\pm = p_x \pm ip_y$.

In the following demonstration we choose $\epsilon(\boldsymbol{p})$ to be null so that the spectrum can be symmetric with respect to $E = 0$ (particle-hole symmetric). The vanishing of the $\epsilon(\boldsymbol{p})\mathbf{1}$ term upgrades the symmetry of the model from class AII to DIII (see Appendix A for details), but leaves unchanged the distinction between topologically trivial ($m_0/m_2 > 0$) and nontrivial ($m_0/m_2 > 0$) phases. Note that the minimal model we consider contains only three control parameters, m_0, m_2, and A. We also consider the spherical geometry, assuming that the topological insulator described by Eqs. (1) and (2) occupies the interior of a sphere of radius R. We introduce standard 3D spherical coordinates: (r,θ,ϕ) related to the Cartesian coordinates as

$$x = r\sin\theta\cos\phi, \quad y = r\sin\theta\sin\phi, \quad z = r\cos\theta. \quad (4)$$

The momentum components (p_x,p_y,p_z) in Eq. (3) expressed in the Cartesian coordinates are rewritten in terms of the derivatives with respect to the spherical coordinates (r,θ,ϕ) by following standard procedure. We also introduce unit vectors $\hat{\boldsymbol{r}}$, $\hat{\boldsymbol{\theta}}$, and $\hat{\boldsymbol{\phi}}$ pointed, respectively, in the direction of the increase of (r,θ,ϕ). The momentum operator $\boldsymbol{p} = p_x\hat{\boldsymbol{x}} + p_y\hat{\boldsymbol{y}} + p_z\hat{\boldsymbol{z}}$ can be reprojected onto the directions of such unit vectors in the spherical coordinates as $\boldsymbol{p} = p_r\hat{\boldsymbol{r}} + p_\theta\hat{\boldsymbol{\theta}} + p_\phi\hat{\boldsymbol{\phi}}$, where $p_r = -i(\partial_r + 1/r)$, $p_\theta = -i\partial_\theta/r$, and $p_\phi = -i\partial_\phi/(r\sin\theta)$.

To derive the surface effective Hamiltonian in the spirit of $k\cdot p$ approximation, we divide H_{bulk} into two parts: $H_{\text{bulk}} = H_\perp + H_\parallel$, where $H_\perp = H|_{p_\theta=0,p_\phi=0}$, and first solve the radial eigenvalue problem

$$H_\perp|\psi\rangle = E_\perp|\psi\rangle, \quad (5)$$

instead of $H_{\text{bulk}}|\psi\rangle = E|\psi\rangle$. Let us consider a possible form of the surface solutions of Eq. (5). $|\psi\rangle$ may take the following form:

$$|\psi\rangle = |\psi(r,\theta,\phi)\rangle = e^{\kappa(r-R)}|u(\theta,\phi)\rangle, \quad (6)$$

where κ^{-1} measures the penetration of the surface wave function into the bulk. Here, taking a linear combination of the solutions of the form of Eq. (6), we construct a solution of Eq. (5) which is compatible with the boundary condition[13,26–28]

$$|\psi(r = R)\rangle = \mathbf{0}, \quad (7)$$

that is, the condition that all four components of the wave function ψ vanish on the surface of the sphere (at $r = R$). As shown in Appendix B, this can be matched by superposing two damped solutions of the form of Eq. (6). Importantly, the solutions of such a boundary value problem must satisfy the zero-energy condition (B14), that is, the (surface) Dirac point is at $E = 0$ in our model.

The zero-energy condition (B14) helps simplify the solution of the radial eigenvalue problem (5). Since $E_\perp = 0$, solving

Eq. (5) is equivalent to finding ψ that satisfies[29,30]

$$\tau_z H_\perp |\psi\rangle = 0, \tag{8}$$

where for $|\psi\rangle$ taking the form of Eq. (6),

$$\tau_z H_\perp |\psi\rangle = [m_\perp[\kappa] + \tau_y A\kappa \hat{\boldsymbol{r}}\cdot\boldsymbol{\sigma}]|\psi\rangle, \tag{9}$$

where $m_\perp[\kappa] = m_0 - m_2\kappa^2$. This implies that the orbital part of the eigenspinor $u(\theta,\phi)$ can be chosen as an eigenstate of τ_y,

$$\tau_y|\tau_y\pm\rangle = \pm|\tau_y\pm\rangle. \tag{10}$$

To fix the notation, let us express the explicit vectorial representation of $|\tau_y\pm\rangle$ as

$$|\tau_y+\rangle = \frac{1}{\sqrt{2}}\begin{bmatrix}1\\ i\end{bmatrix}, \quad |\tau_y-\rangle = \frac{1}{\sqrt{2}}\begin{bmatrix}1\\ -i\end{bmatrix}. \tag{11}$$

The real spin part of Eq. (9) can be also diagonalized by pointing the eigenstates of σ_z in the direction of $\hat{\boldsymbol{r}}$, that is, by

$$\begin{aligned}|\hat{\boldsymbol{r}}+\rangle &= \frac{1}{\sqrt{2}}\begin{bmatrix}e^{-i\phi/2}\cos\frac{\theta}{2}\\ e^{i\phi/2}\sin\frac{\theta}{2}\end{bmatrix},\\ |\hat{\boldsymbol{r}}-\rangle &= \frac{1}{\sqrt{2}}\begin{bmatrix}e^{-i\phi/2}\sin\frac{\theta}{2}\\ -e^{i\phi/2}\cos\frac{\theta}{2}\end{bmatrix}.\end{aligned} \tag{12}$$

Combining these two types of spinors, one can compose the spinorial part of ψ that can be matched with the condition (8), leading to

$$[m_\perp[\kappa] + \kappa A\alpha\beta]|\hat{\boldsymbol{r}}\alpha\rangle_\beta = 0, \tag{13}$$

where

$$|\hat{\boldsymbol{r}}\alpha\rangle_\beta = |\hat{\boldsymbol{r}}\alpha\rangle \otimes |\tau_y\beta\rangle. \tag{14}$$

Equation (13) implies

$$\kappa = \frac{-\alpha\beta A \pm \sqrt{A^2 + 4m_0 m_2}}{2m_2}. \tag{15}$$

The radial part of the wave function $\rho(r)$ that is compatible with the boundary condition (7) takes the form given in Eq. (B9), here, with $\kappa_\pm$ being the two solutions of Eq. (15). Clearly both κ_+ and κ_- must be positive for the wave function ψ to describe a surface state localized in the vicinity of the spherical surface ($r \simeq R$). Thus, in order to cope with the boundary condition, one must have both $\alpha\beta < 0$ and $m_0 m_2 < 0$ for the choice of model parameters such that $A/m_2 > 0$. Notice that the second condition $m_0 m_2 < 0$, which has appeared here automatically from the boundary condition, is a requirement for the system to be in the topologically nontrivial phase.

We have thus successfully found the two basis eigenstates of $H_\perp$ for constructing the effective surface Hamiltonian. For simplicity of the notation we denote them as $\psi = |\pm\rangle\rangle$, where

$$|+\rangle\rangle = \rho(r)|\hat{\boldsymbol{r}}+\rangle_-, \quad |-\rangle\rangle = \rho(r)|\hat{\boldsymbol{r}}-\rangle_+. \tag{16}$$

To avoid misunderstanding of the notations let us express explicitly the four-component vectorial form of $|\hat{\boldsymbol{r}}\pm\rangle_\mp$,

$$\begin{aligned}|\hat{\boldsymbol{r}}+\rangle_- &= \frac{1}{2}\begin{bmatrix}\begin{bmatrix}1\\ -i\end{bmatrix}e^{-i\phi/2}\cos\frac{\theta}{2}\\ \begin{bmatrix}1\\ -i\end{bmatrix}e^{i\phi/2}\sin\frac{\theta}{2}\end{bmatrix},\\ |\hat{\boldsymbol{r}}-\rangle_+ &= \frac{1}{2}\begin{bmatrix}\begin{bmatrix}1\\ i\end{bmatrix}e^{-i\phi/2}\sin\frac{\theta}{2}\\ -\begin{bmatrix}1\\ i\end{bmatrix}e^{i\phi/2}\cos\frac{\theta}{2}\end{bmatrix}.\end{aligned} \tag{17}$$

Here the arrangement of the basis is made in accordance with that of Eq. (3). Notice that the eigenvectors of Eq. (17) are *double valued* with respect to the azimuthal angle ϕ. This does not happen for the polar angle θ since the domain of definition for θ is restricted to a finite range $\theta \in [0,\pi]$ and not periodic, in contrast to ϕ. The double valuedness stems from our choice of the (arbitrary) phase factor in front of Eq. (12). This is, on the other hand, merely a choice, and one can equally formulate the same problem consistently using a pair of *single-valued* eigenvectors. We leave further arguments on this point to Sec. III and here take these *double-valued* eigenvectors as a basis for constructing the surface effective Hamiltonian.

The effective surface "Hamiltonian" $\mathcal{H}_{\rm dv}$ acts on a two-component spinor,

$$\boldsymbol{\alpha} = \begin{bmatrix}\alpha_+\\ \alpha_-\end{bmatrix}. \tag{18}$$

Within the $k\cdot p$ approximation any surface state $|\boldsymbol{\alpha}\rangle\rangle$ can be represented as a linear combination of $|+\rangle\rangle$ and $|-\rangle\rangle$ with the amplitude specified, respectively, by α_+ and α_-, that is,

$$|\boldsymbol{\alpha}\rangle\rangle = \alpha_+|+\rangle\rangle + \alpha_-|-\rangle\rangle \tag{19}$$

and

$$\begin{bmatrix}\langle\langle +|H_\parallel|\boldsymbol{\alpha}\rangle\rangle\\ \langle\langle -|H_\parallel|\boldsymbol{\alpha}\rangle\rangle\end{bmatrix} \equiv \mathcal{H}_{\rm dv}\boldsymbol{\alpha}. \tag{20}$$

The explicit form of $\mathcal{H}_{\rm dv}$ can be determined by evaluating each of the matrix elements $H_\parallel$ against the basis vectors $|\pm\rangle\rangle$, that is, $\langle\langle\pm|H_\parallel|\mp\rangle\rangle$. The procedure we follow here is precisely in parallel with that of the standard degenerate perturbation theory. $H_0 = H_\perp$ is an unperturbed Hamiltonian and $|\pm\rangle\rangle$ are its (twofold) degenerate eigenstates. To find the (degeneracy-lifted) spectrum of the perturbed Hamiltonian $H_{\rm tot} = H_0 + H'$, where $H' = H_\parallel$, $H_{\rm tot} = H_{\rm bulk}$, we first calculate the matrix elements $\langle\langle\alpha|H'|\beta\rangle\rangle \equiv (\mathcal{H}_{\rm dv})_{\alpha\beta}$ $(\alpha,\beta = \pm)$, then diagonalize this 2×2 coefficient matrix.

The explicit matrix form of $H_\parallel$ is

$$H_\parallel = \begin{bmatrix} m_\parallel & -iAD_z & 0 & -iAD_-\\ -iAD_z & -m_\parallel & -iAD_- & 0\\ 0 & -iAD_+ & m_\parallel & iAD_z\\ -iAD_+ & 0 & iAD_z & -m_\parallel\end{bmatrix}, \tag{21}$$

where $D_\pm$ and D_z are defined in Eqs. (B1) and (B2). Performing the r integral in $\langle\langle\alpha|H_\parallel|\beta\rangle\rangle$ one can safely replace the r dependence in these expressions with the radius R of the sphere, assuming that the surface wave function is well localized in the vicinity of the surface. Alternatively, one can

equally regard $D_\pm$ and D_z as

$$D_\pm \simeq \frac{e^{\pm i\phi}}{R}\left[\cos\theta\frac{\partial}{\partial\theta} \pm \frac{i}{\sin\theta}\frac{\partial}{\partial\phi}\right], \quad D_z \simeq -\frac{\sin\theta}{R}\frac{\partial}{\partial\theta}. \tag{22}$$

At leading order in the expansion with respect to $1/R$ the diagonal terms of $H_\parallel$ can be neglected since

$$m_\parallel = m_2\frac{\boldsymbol{L}^2}{r^2} \sim \frac{1}{R^2} \tag{23}$$

[see also Eqs. (B3) and (B5)]. Within this accuracy the coefficient matrix $\mathcal{H}_{\rm dv}$ is found, after some algebra, to be

$$\mathcal{H}_{\rm dv} = \frac{A}{R}\begin{bmatrix} 0 & -\partial_\theta + \frac{i\partial_\phi}{\sin\theta} - \frac{\cot\theta}{2} \\ \partial_\theta + \frac{i\partial_\phi}{\sin\theta} + \frac{\cot\theta}{2} & 0 \end{bmatrix}. \tag{24}$$

Apart from an overall constant in front of the expression, this can be identified as the "Dirac operator" for a free massless fermion on the sphere.[21–25,31–33]

The origin of the Berry phase term can be attributed to the covariance of the derivatives ∂_θ and ∂_ϕ on a curved spherical surface.[21–25] In this regard, the Berry phase term appears as a spin connection in $\mathcal{D}_\phi$ as

$$\mathcal{D}_\phi = \partial_\phi + \frac{i\sigma_z}{2}\cos\theta, \tag{25}$$

replacing ∂_ϕ in Eq. (24) as

$$\mathcal{H}_{\rm dv} = A\left[\left(\frac{-i\partial_\theta}{R}\right)\sigma_y - \left(\frac{-i\mathcal{D}_\phi}{R\sin\theta}\right)\sigma_x\right]. \tag{26}$$

Alternatively, the Berry phase term can be absorbed in ∂_θ by introducing

$$\tilde{\partial}_\theta = \partial_\theta + \tfrac{1}{2}\cot\theta. \tag{27}$$

In terms of $\tilde{\partial}_\theta$, the Dirac operator (24) can be also rewritten as (cf. Table I)

$$\begin{aligned}\mathcal{H}_{\rm dv} &= \frac{A}{R}\begin{bmatrix} 0 & -\tilde{\partial}_\theta + \frac{i\partial_\phi}{\sin\theta} \\ \tilde{\partial}_\theta + \frac{i\partial_\phi}{\sin\theta} & 0 \end{bmatrix} \\ &= A\left[\left(\frac{-i\tilde{\partial}_\theta}{R}\right)\sigma_y - \left(\frac{-i\partial_\phi}{R\sin\theta}\right)\sigma_x\right].\end{aligned} \tag{28}$$

III. NATURE OF THE TWO TYPES OF BERRY PHASE

The advantage of considering the spherical geometry is that the existence of two different types of Berry phase becomes apparent; each associated, respectively, with an electronic motion in the polar (θ) [type (A)] and azimuthal (ϕ) [type (B)] directions (see Table I). The type (A) Berry phase is intrinsic to the curvature of the spherical surface, while the type (B) is associated with the so-called spin-to-surface locking. For a cylindrical surface, on the contrary, only the latter [type (B)] manifests since the cylindrical surface has a vanishing Gaussian curvature. The contrasting behaviors of the two types of Berry phase are summarized in Table I.

To highlight the distinct behaviors of the two types of Berry phase, let us reconsider the Dirac operator (24) expressed against the double-valued basis vectors (17). As mentioned earlier, this was not a unique choice of the basis. One can equally choose them to be *single valued* as

$$\begin{aligned}|+\rangle\rangle &= \frac{1}{2}\begin{bmatrix} \begin{bmatrix}1\\-i\end{bmatrix}\cos\frac{\theta}{2} \\ \begin{bmatrix}1\\-i\end{bmatrix}e^{i\phi}\sin\frac{\theta}{2}\end{bmatrix}, \\ |-\rangle\rangle &= \frac{1}{2}\begin{bmatrix} \begin{bmatrix}1\\i\end{bmatrix}\sin\frac{\theta}{2} \\ -\begin{bmatrix}1\\i\end{bmatrix}e^{i\phi}\cos\frac{\theta}{2}\end{bmatrix}.\end{aligned} \tag{29}$$

This type of a single-valued choice of the basis is often employed in the $\boldsymbol{k}\cdot\boldsymbol{p}$ description of the electronic states in graphene.[7] Once this choice of basis vectors is adopted, one can repeat the same procedure as we described in the last section to find the surface effective Hamiltonian $\mathcal{H}_{\rm sv}$, or the Dirac operator in this basis, as

$$\mathcal{H}_{\rm sv} = \frac{A}{R}\begin{bmatrix} 0 & -\partial_\theta + \frac{i\partial_\phi}{\sin\theta} - \frac{1}{2}\cot\frac{\theta}{2} \\ \partial_\theta + \frac{i\partial_\phi}{\sin\theta} - \frac{1}{2}\tan\frac{\theta}{2} & 0 \end{bmatrix}. \tag{30}$$

In passing from Eq. (24) to (30), the matrix elements are replaced as

$$-i\partial_\phi \to -i\partial_\phi + \tfrac{1}{2}. \tag{31}$$

Here the additive factor $1/2$ is nothing but the Berry phase π associated with the spin-to-surface locking in cylindrical surfaces that has been extensively discussed in the literature.[9–13]

The appearance of the type (A) Berry phase (see Table I) is not restricted to the topological insulator surface state. It has already appeared in the study of the electronic spectrum of fullerene, typically the one called "buckyball" (or Buckminsterfullerene).[21–25] The type (B) Berry phase, on the contrary, is specific to the topological insulator surface state. By its nature whether this type of Berry phase appears explicitly in the effective Dirac operator depends on the choice of the basis. In $\mathcal{H}_{\rm dv}$ [Eq. (24)] the Berry phase is superficially hidden in the antiperiodicity of the basis spinor (17) with respect to the azimuthal angle ϕ. When one considers the orbital part of the wave function or spinor (18), this point must be carefully taken into account in its periodicity with respect to ϕ. This point will be clarified in the next section.

The Berry phase term, or more precisely the spin connection of the form of $\pm\cot\theta/2$ in Eq. (24), or equivalently either $\cot(\theta/2)/2$ or $\tan(\theta/2)/2$ in the two off-diagonals of Eq. (30), can be interpreted as a vector potential generated by an effective *magnetic monopole*. Indeed, the magnetic field associated with a magnetic monopole of strength g can be successfully encoded in a vector potential, for example,

$$\boldsymbol{A}_{\rm I} = \frac{g}{4\pi r}\tan\frac{\theta}{2}\,\hat{\boldsymbol{\phi}}, \tag{32}$$

$$\boldsymbol{A}_{\rm II} = \frac{-g}{4\pi r}\cot\frac{\theta}{2}\,\hat{\boldsymbol{\phi}}, \tag{33}$$

by introducing the concept of Dirac's string. Here $\boldsymbol{A}_{\rm I}$ and $\boldsymbol{A}_{\rm II}$ correspond to a choice of the gauge in which the Dirac's string

runs, respectively, on the $-z$ ($+z$) axis [direction of the south ($\theta=\pi$) vs north ($\theta=0$) poles]. Equations (24) and (30) imply that the strength of the "induced" monopole is, respectively, $g=-2\pi$ for α_+, and $g=2\pi$ for α_-. A fictitious magnetic monopole of an *opposite* charge is effectively induced at the center of the sphere for the two spin components of the surface spinor wave function (18) [see also Eq. (42)].

IV. SURFACE EIGENSTATES ON THE SPHERE

To find the electronic spectrum of the spherical topological insulator surface state, one needs to solve the surface Dirac equation

$$\mathcal{H}_{\rm dv}\boldsymbol{\alpha}=E_\parallel\boldsymbol{\alpha} \tag{34}$$

explicitly. In accordance with the decomposition of the Hamiltonian $H_{\rm bulk}=H_\perp+H_\parallel$, we have also separated the energy eigenvalue E of the original eigenvalue equation

$$H_{\rm bulk}|\boldsymbol{\alpha}\rangle\rangle=E|\boldsymbol{\alpha}\rangle\rangle \tag{35}$$

into two parts: $E=E_\perp+E_\parallel$. But, of course, since $E_\perp=0$ [Eq. (B14)], $E=E_\parallel$. As we have already emphasized, there exists some freedom in the choice of basis vectors, but here mainly to ease comparison with the literature we take them double valued as in Eq. (17). To solve the eigenvalue equation (34) explicitly we first separate the variables as $\boldsymbol{\alpha}(\theta,\phi)=e^{im\phi}\boldsymbol{\alpha}_m(\theta)$, or in terms of the components,

$$\begin{bmatrix}\alpha_+(\theta,\phi)\\ \alpha_-(\theta,\phi)\end{bmatrix}=e^{im\phi}\begin{bmatrix}\alpha_{m+}(\theta)\\ \alpha_{m-}(\theta)\end{bmatrix}. \tag{36}$$

Clearly the quantum number m signifies physically the z component of the orbital angular momentum. Here the important point is that this m takes *half-odd integral* values

$$m=\pm\tfrac{1}{2},\pm\tfrac{3}{2},\pm\tfrac{5}{2},\ldots \tag{37}$$

in contrast to the case of fullerene.[21,22,25] Any solution of the original Schrödinger equation $H_{\rm bulk}|\psi\rangle=E|\psi\rangle$, or its surface solution $|\boldsymbol{\alpha}\rangle\rangle$ [see Eq. (19)], obeys to a periodic boundary condition with respect to ϕ. To be explicit, any bulk (or surface) solution $|\psi\rangle=|\psi(r,\theta,\phi)\rangle$ satisfies $|\psi(r,\theta,\phi+2\pi)\rangle=|\psi(r,\theta,\phi)\rangle$, and in the same sense,

$$|\boldsymbol{\alpha}(\theta,\phi+2\pi)\rangle\rangle=|\boldsymbol{\alpha}(\theta,\phi)\rangle\rangle. \tag{38}$$

On the other hand, the basis spinor (17) spanning the space of the surface solutions $|\boldsymbol{\alpha}(\theta,\phi)\rangle\rangle$ [Eq. (19)] is antiperiodic with respect to 2π rotation of the azimuthal angle, that is, $|\pm\rangle\rangle\to-|\pm\rangle\rangle$ under $\phi\to\phi+2\pi$. In order to ensure the periodicity of the total wave function (38) this minus sign must be compensated or absorbed in the prefactor $\alpha_\pm=\alpha_\pm(\theta,\phi)$ of Eq. (19), that is, $\alpha_\pm(\theta,\phi+2\pi)=-\alpha_\pm(\theta,\phi)$.

In terms of $\alpha_{m+}(\theta)$ and $\alpha_{m-}(\theta)$, the eigenvalue equation (34) becomes a couple of first-order linear differential equations,

$$\begin{aligned}-\frac{A}{R}\left[\frac{d}{d\theta}+\frac{m}{\sin\theta}+\frac{\cot\theta}{2}\right]\alpha_{m-}(\theta)&=E_\parallel\alpha_{m+}(\theta),\\ \frac{A}{R}\left[\frac{d}{d\theta}-\frac{m}{\sin\theta}+\frac{\cot\theta}{2}\right]\alpha_{m+}(\theta)&=E_\parallel\alpha_{m-}(\theta).\end{aligned} \tag{39}$$

These two equations combine to give

$$\left[\frac{1}{\sin\theta}\frac{d}{d\theta}\sin\theta\frac{d}{d\theta}-\frac{1}{\sin^2\theta}\left(m-\frac{\sigma}{2}\cos\theta\right)^2+\lambda^2-\frac{1}{2}\right]\times\alpha_{m\sigma}(\theta)=0, \tag{40}$$

in which two important parameters σ and λ have been introduced. $\sigma=\pm$ specifies the spin index in the subscript of $\alpha_{m\sigma}$, whereas λ parametrizes the energy as

$$E_\parallel=\frac{A}{R}\lambda. \tag{41}$$

Let us remark here that Eq. (40) is formally equivalent to a differential equation defining the so-called monopole harmonics,[34] the latter describing (the angular part of) the electronic motion in the presence of a magnetic monopole. The role of the nontrivial Berry curvature imposed by the spherical geometry can be thus interpreted as an effective *magnetic monopole* induced at the center of the sphere. An electron in the surface state of a spherical topological insulator "sees" such an effective magnetic monopole and behaves accordingly. In the notation of Ref. 34, the quantum number q specifies physically the strength g of the effective magnetic monopole placed at the origin as $g=4\pi q$. The Dirac's quantization condition restricts the "allowed" value of this quantum number q to be half integral $q=0,\pm1/2,\pm1,\ldots$. Here, in Eq. (40), q is identified as

$$q=-\frac{\sigma}{2}=\pm\frac{1}{2}. \tag{42}$$

The eigenfunction $\alpha_{m\sigma}(\theta)$ of Eq. (40) is indeed related to the monopole harmonics specified by this value of q. Thus at the center of a spherical topological insulator, a magnetic monopole of strength $\pm2\pi$, the smallest finite value compatible with the Dirac's quantization condition is effectively induced. By its nature the induced monopole is automatically regularized by a Dirac's string.

Introducing a new independent variable $\zeta=\cos\theta$, one can rewrite Eq. (40) as[23]

$$\left[\frac{d}{d\zeta}(1-\zeta^2)\frac{d}{d\zeta}-\frac{m^2-\sigma m\zeta+1/4}{1-\zeta^2}+\lambda^2-\frac{1}{4}\right]\alpha_{m\sigma}=0. \tag{43}$$

This can be further rewritten in the form of a Jacobi-type differential equation. As shown below, its normalizable solutions are known to be expressed in terms of Jacobi polynomials $P_n^{\mu\nu}[\zeta]$ (see Appendix C for our conventions). Changing the dependent variables as

$$\alpha_{m\sigma}[\zeta]=(1-\zeta)^{\frac{1}{2}|m-\frac{\sigma}{2}|}(1+\zeta)^{\frac{1}{2}|m+\frac{\sigma}{2}|}\beta_{m\sigma}[\zeta], \tag{44}$$

and using the fact that m is half-integral [Eq. (37)], one can verify

$$\left[(1-\zeta^2)\frac{d^2}{d\zeta^2}+\left\{\sigma\frac{m}{|m|}-(2|m|+2)\zeta\right\}\frac{d}{d\zeta}-|m|(|m|+1)+\lambda^2-\frac{1}{4}\right]\beta_{m\sigma}[\zeta]=0. \tag{45}$$

Comparing this with the standard form of the Jacobi's differential equation [cf. Eq. (C3)],

$$\left[(1-\zeta^2)\frac{d^2}{d\zeta^2}+\{\nu-\mu-(\mu+\nu+2)\zeta\}\frac{d}{d\zeta} +n(n+\mu+\nu+1)\right]P_n^{\mu\nu}[\zeta]=0, \tag{46}$$

one can identify the parameters as

$$\mu=|m|-\frac{\sigma}{2}\frac{m}{|m|}=\left|m-\frac{\sigma}{2}\right|, \qquad \nu=|m|+\frac{\sigma}{2}\frac{m}{|m|}=\left|m+\frac{\sigma}{2}\right|, \tag{47}$$

and

$$\lambda^2=n(n+\mu+\nu+1)+|m|(|m|+1)+\tfrac{1}{4} =\left(n+|m|+\tfrac{1}{2}\right)^2, \tag{48}$$

where the normalizability of the wave function requires n to be non-negative integers,

$$n=0,\,1,\,2,\,3,\,\ldots. \tag{49}$$

In Eqs. (47) $\sigma=\pm$ refers to the subscript of $\beta_{m\sigma}[\zeta]$. The second equality of Eqs. (47) holds since m is half-integral. Thereby, apart from a normalization constant $c_{nm\sigma}$, which will be determined later, the surface wave function $\beta_{m\sigma}[\zeta]$ is expressed in terms of the nth order Jacobi polynomial as

$$\beta_{m\sigma}[\zeta]=\beta_{nm\sigma}[\zeta]=c_{nm\sigma}P_n^{|m-\frac{\sigma}{2}|,|m+\frac{\sigma}{2}|}[\zeta]. \tag{50}$$

Clearly n is common to $\beta_{m+}[\zeta]$ and $\beta_{m-}[\zeta]$ once λ is chosen to be an appropriate quantized value imposed by Eq. (48). Substituting Eq. (48) into (41), one finds the surface energy spectrum

$$E=E_{\parallel}=\pm\frac{A}{R}\left(n+|m|+\frac{1}{2}\right)\equiv E_\lambda, \tag{51}$$

where m and n take, respectively, half-integral and non-negative integral values [see Eqs. (37) and (49)]. The energy spectrum (51) has a couple of specific features.

(1) The discrete energy levels E_λ are placed with an equal distance A/R symmetrically on the positive and negative side of E, excluding the zero energy $E=0$.

(2) The degeneracy g_λ of each energy level increases linearly with $|E|$, taking every (positive) even numbers $g_\lambda=2,\,4,\,6,\ldots$ when one starts counting it at the first positive and negative energy level with $\lambda=\pm1$.

The discrete energy levels E_λ can be also cast from the view of finite-size energy gap. On the surface of an infinitely long cylinder the surface electronic spectrum shows an energy gap as a consequence of the spin-to-surface locking and the resulting half-integer quantization analogous to Eq. (37). The size of the obtained energy gap on a cylindrical surface is inversely proportional to the radius of the cylinder. Here in Eq. (51) the zero energy $E=0$ is indeed excluded, which can be regarded as a remnant of the energy gapped in the continuous spectrum. Note that this has nothing to do with the discreteness of the spectrum due to finite-size "quantization." What counts here is the absolute position of the entire (discrete) spectrum. The size of the "energy gap"

$$\Delta(R)=E_1-E_{-1}=\frac{2A}{R} \tag{52}$$

is again inversely proportional to R, the radius of the sphere.[20]

In order to determine the relative magnitude and phase of c_{nm+} and c_{nm-}, one needs to go back to Eqs. (39) and substitute $\beta_{m\sigma}[\zeta]$ obtained as in Eqs. (50) into this couple of equations. This is straightforward, but turns out to be rather tedious work. Leaving an explicit demonstration of this to Appendix D, here we refer only to its result

$$c_{nm-}=-\mathrm{sign}[\lambda m]c_{nm+}. \tag{53}$$

Choosing c_{nm+} to be real, and taking the normalization

$$2\pi\int d(\cos\theta)|\boldsymbol{\alpha}_{nm}(\theta)|^2=1 \tag{54}$$

also into account [cf. Eq. (C6) for the normalization of $P_n^{\mu\nu}[\zeta]$], one can give an explicit list of the coefficients $c_{nm\sigma}$ in which

$$c_{nm+}=\frac{1}{\sqrt{2\pi}}\frac{\sqrt{n!(n+2|m|)!}}{2^{|m|+1/2}(n+|m|-1/2)!}. \tag{55}$$

In Eq. (54) we introduced the notation $\boldsymbol{\alpha}_{nm}(\theta)$, that is, $\boldsymbol{\alpha}_m(\theta)=\boldsymbol{\alpha}_{nm}(\theta)$, in accordance with Eq. (50).

As a solution of the effective surface Dirac equation (34), Eqs. (55), (53), (50), and (44) specify an explicit form of the surface eigenspinor $\boldsymbol{\alpha}(\theta,\phi)=e^{im\phi}\boldsymbol{\alpha}_{nm}(\theta)$, with a definite *half-integral* angular momentum m. It may be suggestive to give a few concrete examples of $\boldsymbol{\alpha}_{nm}(\theta)$ for small values of n and $|m|$. For simplicity let us restrict ourselves to the case of $\lambda>0$ (positive energy). For $n=0$ with an arbitrary half-integral m, $\boldsymbol{\alpha}_{nm}(\theta)$ is found to be

$$\boldsymbol{\alpha}_{0m}(\theta)=\frac{1}{\sqrt{4\pi}}\frac{\sqrt{(2|m|)!}}{(|m|-\frac{1}{2})!}\begin{bmatrix}\left(\sin\frac{\theta}{2}\right)^{|m-\frac{1}{2}|}\left(\cos\frac{\theta}{2}\right)^{|m+\frac{1}{2}|}\\ -s_m\left(\sin\frac{\theta}{2}\right)^{|m+\frac{1}{2}|}\left(\cos\frac{\theta}{2}\right)^{|m-\frac{1}{2}|}\end{bmatrix}, \tag{56}$$

where s_m is an abbreviated notation for

$$s_m=\mathrm{sign}[m]=\frac{m}{|m|}. \tag{57}$$

More specifically,

$$\boldsymbol{\alpha}_{0\frac{1}{2}}=\begin{bmatrix}\alpha_{0\frac{1}{2}+}\\ \alpha_{0\frac{1}{2}-}\end{bmatrix}=\frac{1}{\sqrt{4\pi}}\begin{bmatrix}\cos\frac{\theta}{2}\\ -\sin\frac{\theta}{2}\end{bmatrix}, \tag{58}$$

$$\boldsymbol{\alpha}_{0,-\frac{1}{2}}=\frac{1}{\sqrt{4\pi}}\begin{bmatrix}\sin\frac{\theta}{2}\\ \cos\frac{\theta}{2}\end{bmatrix}, \tag{59}$$

$$\boldsymbol{\alpha}_{0\frac{3}{2}}=\sqrt{\frac{3}{2\pi}}\begin{bmatrix}\sin\frac{\theta}{2}\cos^2\frac{\theta}{2}\\ -\sin^2\frac{\theta}{2}\cos\frac{\theta}{2}\end{bmatrix}, \tag{60}$$

$$\boldsymbol{\alpha}_{0,-\frac{3}{2}}=\sqrt{\frac{3}{2\pi}}\begin{bmatrix}\sin^2\frac{\theta}{2}\cos\frac{\theta}{2}\\ \sin\frac{\theta}{2}\cos^2\frac{\theta}{2}\end{bmatrix}. \tag{61}$$

For $n = 1$ and $m = \pm 1/2$,

$$\boldsymbol{\alpha}_{1\,\frac{1}{2}} = -\frac{1}{2\sqrt{2\pi}} \begin{bmatrix} \cos\frac{\theta}{2}(1-3\cos\theta) \\ \sin\frac{\theta}{2}(1+3\cos\theta) \end{bmatrix}, \tag{62}$$

$$\boldsymbol{\alpha}_{1,-\frac{1}{2}} = \frac{1}{2\sqrt{2\pi}} \begin{bmatrix} \sin\frac{\theta}{2}(1+3\cos\theta) \\ -\cos\frac{\theta}{2}(1-3\cos\theta) \end{bmatrix}. \tag{63}$$

The two eigenstates (58) and (59) constitute the two lowest energy (degenerate) eigenstates on the $E > 0$ side, whereas Eqs. (60)–(63) correspond to the four second lowest (first excited) states.

Let us point out that the surface spinor wave functions thus constructed imply a rich *spin texture* that the surface eigenstates manifest. This can be already seen in the simplest example, the case of the eigenstate $\boldsymbol{\alpha}_{0\,\frac{1}{2}}$ given as in Eq. (58). At the north pole ($\theta = 0$) the spin in this eigenstate is pointed in the $+z$ direction, perpendicular to the tangential plane of the sphere at this point. As θ increases, the spin tends to lie closer to the tangential plane. At $\theta = \pi/2$, that is, on the equator, the spin is completely in-plane to the spherical surface. Now, staying at $\theta = \pi/2$, if one lets ϕ vary from 0 to 2π, that is, as the electron hypothetically travels around the equator, the spin also completes a 2π rotation in the (x, y) plane, manifesting the feature of "spin-to-surface locking." As one further increases θ, the spin-down component starts to dominate, before it dominates completely at the south pole.

Another example showing a more complicated spin texture is the case of $\boldsymbol{\alpha}_{1\,\frac{1}{2}}(\theta)$ given as in Eq. (62) [the two spin components of $\boldsymbol{\alpha}_{1\,\frac{1}{2}}(\theta)$ are depicted in Fig. 1]. At the north pole the spin in this eigenstate is again pointed in the $+z$ direction. As θ increases, it starts to lie, but tends to lie more strongly than the case of $\boldsymbol{\alpha}_{0\,\frac{1}{2}}$. At

$$\theta = 2\arccos\frac{3+\sqrt{5}}{6} \simeq 0.232\pi \equiv \theta_1 \tag{64}$$

the two spin components acquire an equal weight. At this angle $\theta = \theta_1$, if one lets ϕ vary from 0 to 2π, complete spin-to-surface locking occurs. As θ exceeds θ_1, the $|\hat{\boldsymbol{r}}-\rangle$ component starts to dominate. At

$$\theta = \arccos[1/3] \simeq 0.392\pi \equiv \theta_2 \tag{65}$$

the centrifugal spin component $\alpha_{1\,\frac{1}{2}+}(\theta)$ vanishes. Therefore, at this point the $|\hat{\boldsymbol{r}}-\rangle$ component dominates completely, and the spin is pointed to the center of the sphere. As θ further increases, the spin gradually tilts back to the tangential plane (at $\theta = \pi/2$, spin-to-surface locking is recovered), then it starts to be further tilted toward the outside of the sphere. At $\theta = \arccos[-1/3] = \pi - \theta_2$, the centripetal spin component $\alpha_{1\,\frac{1}{2}-}(\theta)$ vanishes, and the spin finds itself purely in the $|\hat{\boldsymbol{r}}+\rangle$ state. At $\theta > \pi - \theta_2$, $|\hat{\boldsymbol{r}}-\rangle$ starts to dominate again, and the spin is finally pointed in the $+z$ direction at the south pole. The behavior of the spin as it varies ϕ from 0 to 2π is the same as the case of $\boldsymbol{\alpha}_{0\,\frac{1}{2}}$ since the two states share the same quantum number $m = 1/2$.

The spin also rotates more drastically in the ϕ direction for the eigenstates with $|m| \geqslant 3/2$. Clearly the surface spinor wave functions with quantum numbers n and $|m|$ higher than the examples given in Eqs. (58)–(63) show a richer spin structure on the spherical surface as a consequence of the interplay between the two types of Berry phase; cf. the case of cylindrical geometry in which only a single type of Berry phase manifests.

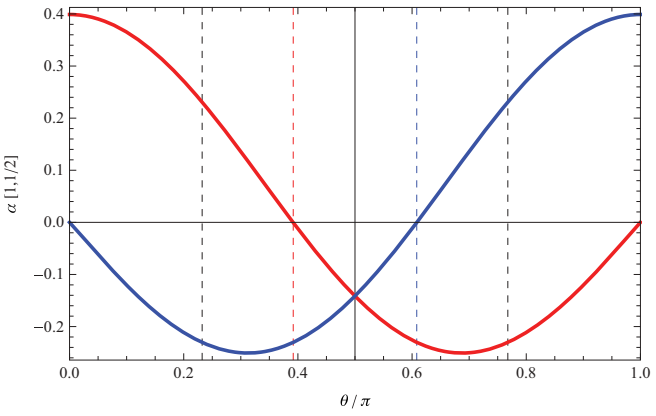

FIG. 1. (Color online) θ dependence of the two spin components of $\boldsymbol{\alpha}_{1\,\frac{1}{2}}(\theta)$. $\alpha_{1\,\frac{1}{2}+}(\theta)$ and $\alpha_{1\,\frac{1}{2}-}(\theta)$ shown, respectively, in red and in blue. The two spin components acquire an equal weight at $\theta = \theta_1 \simeq 0.232\pi$, $\theta = \pi/2$, and at $\theta = \pi - \theta_1$. $\alpha_{1\,\frac{1}{2}+}(\theta)$ vanishes at $\theta = \theta_2 \simeq 0.392\pi$ and at $\theta = \pi$, whereas $\alpha_{1\,\frac{1}{2}-}(\theta)$ vanishes at $\theta = 0$ and at $\theta = \pi - \theta_2$.

V. COMPARISON WITH THE TIGHT-BINDING CALCULATION

We have so far investigated specific features of the topological insulator surface state occupying a finite volume, taking as an example the case of spherical geometry. Starting with the gapped bulk effective Hamiltonian, we have derived and solved the surface Dirac equation, from which we have deduced the surface electronic spectrum (51) and the explicit form of the spinor wave functions [as given through Eqs. (44), (50), (53), and (55)] on a perfect spherical surface. The role of two distinct types of Berry phase has been revealed. Here we take another viewpoint; namely, we go back to the bulk effective Hamiltonian [Eq. (1)] and implement it as a nearest-neighbor tight-binding model, which allows for obtaining the spectrum of the surface solutions by exact diagonalization. We show that the basic features on the surface energy spectrum we have found so far in the idealized spherical geometry with exact rotational symmetry is *still valid* when that symmetry is *weakly broken*.

Let us employ the following lattice implementation of H_{bulk} [i.e., Eq. (1)] on a cubic lattice:

$$H_{\text{lattice}} = \tau_z m(\boldsymbol{k}) + A\tau_x \sum_{j=x,y,z} \sigma_j \sin k_j, \tag{66}$$

where

$$m(\boldsymbol{k}) = m_0 + 2\sum_{j=x,y,z} m_2(1-\cos k_j) \tag{67}$$

is a lattice version of Eq. (2). The model specified by Eqs. (66) and (67) can be regarded as a tight-binding model with only the nearest neighbor hopping. This couple of equations determine the structure of energy bands over the entire Brillouin zone, which also reproduces, in the vicinity of the Γ point, the

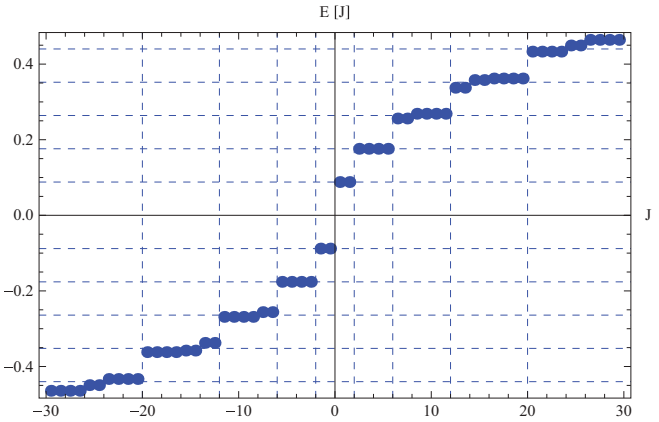

FIG. 2. (Color online) Pseudodegenerate spectrum $E[J]$ of the surface eigenstates, obtained by diagonalization of the tight-binding Hamiltonian on a cubic lattice. The system has a cubic shape of linear dimension $L = 16$. $m_0/m_2 = -1$, $A = 1$.

bulk effective Hamiltonian H_{bulk} [Eq. (1)] in the continuum limit. The system we consider here has a cubic shape of linear dimension L, in which the lattice points placed with a unit lattice spacing are restricted to

$$1 \leqslant x, y, z \leqslant L. \tag{68}$$

The obtained surface energy spectrum $E = E[J]$ is shown in Fig. 2. Here J is an index for numbering each surface eigenstate with an increasing order of E. For an aesthetic reason, and also to ease counting of the degree of (pseudo-) degeneracy, we have chosen this index J to be half-integral $J = \pm 1/2, \pm 3/2, \ldots$. Numbering of the eigenstate is done in such a way that starting with $E = 0$, $J = 1/2, 3/2, 5/2, \ldots$ is attributed to each eigenstate with increasing (decreasing) order of E on the positive (negative) E side [i.e., a positive (negative) J corresponds to a positive (negative) energy level]. Depending on the number of lattice sites contained in the system ($=L^3$) and on the value of model parameters, certain numbers of states appear in the bulk energy gap. Those surface eigenstates of relatively small $|E| < m_0$ are also characterized by the spatial profile of the corresponding wave function; their wave function is localized in the vicinity of the cubic surface. In this demonstration, the system's size is $L = 16$, and the model parameters are chosen as $m_0/m_2 = -1$, $A = 1$.

$E = E[J]$ plotted in Fig. 2 shows a "pseudodegenerate" feature very reminiscent of the quantization characteristic on the spherical surface, the spectrum (51), and the degeneracy rule [Eqs. (49) and (37)]. Notice that the zero-energy state $E = 0$ is clearly excluded. Horizontal gridlines are located at the positions of E that are an integer multiple of $E[1]$, the first (positive) energy level. These gridlines are shown for verifying that the energy levels are equally placed, one of the characteristic features of the spectrum on the spherical surface. Small deviation from the "expected" spectrum (51) can be seen for $|\lambda| > 3$, which can be interpreted as a consequence of the breaking of spherical symmetry. Vertical gridlines are drawn for highlighting the degree of the pseudodegeneracy of each level. If one recalls that a magnetic monopole is effectively induced at the center of the sphere, the (pseudo-) equally spaced spectrum illustrated in Fig. 2 can be interpreted as Landau levels. For $L = 16$ the value of $E[1]$ is found to be $E[1] \simeq 0.0880$ in units of $\hbar A/a$, where A is the group velocity of the surface state, and a is the lattice constant. Taking experimentally realistic values for A and a,[27] the characteristic energy scale $\hbar A/a$ is estimated to be on the order of 0.1–1 eV.

VI. CONCLUDING REMARKS

The protected surface state of a topological insulator has an "active" property that it reveals only when it is embedded onto a curved surface. On a cylindrical surface, it induces an effective magnetic solenoid of total flux $\pm\pi$. In the same sense, a magnetic monopole of strength $\pm 2\pi$ is induced when it is embedded onto a sphere. In this paper we have explicitly examined this active property of the topological insulator surface state, focusing on a most suggestive case of the spherical surface. As a result, the following unique profile of such surface states has been found. The two important features are

(1) A unique quantization rule; equally spaced spectrum with the exception of the "evaporated" zero-energy state, and the simple degeneracy rule.

(2) A rich spin texture resulting from the nature of complicated spinor wave function.

These characteristic features derived analytically using an idealized spherical geometry is then contrasted with a tight-binding calculation on a cubic lattice of cubic shape. Unexpectedly profound agreement of the two results suggests that those features which we have demonstrated on a perfect sphere capture the essential characteristic of the surface states occupying a finite volume of a more generic shape, inevitably involving a curved surface. In this sense it is natural to expect that the obtained spectrum be applicable to the spectroscopy of topological insulator *nanoparticles*. In particular, we predict a unique photoabsorption/emission spectrum resulting from the equally spaced energy levels of the low-energy surface eigenstates.

Most of the existing works characterizing the topological insulator surface states are based on Bloch states. Topological properties, however, do not depend on the translational symmetry. Here we have demonstrated this, focusing on the angular momentum[35–37] (instead of the linear momentum) as the good quantum number.

ACKNOWLEDGMENTS

K.I. acknowledges Akihiro Tanaka for valuable comments on the manuscript. The authors are supported by KAKENHI; K.I. and T.F. by the "Topological Quantum Phenomena" [No. 23103511 and No. 23103502], and Y.T. by a Grant-in-Aid for Scientific Research (C) [No. 24540375].

APPENDIX A: SYMMETRY CLASS AND TOPOLOGICAL INVARIANTS

In the classification of the Dirac Bogoliubov–de Gennes type Hamiltonian in terms of the time-reversal (Θ), particle-hole ($\mathcal{C}$), and chiral (Γ_5) symmetries,[15–19,38] our starting bulk effective Hamiltonian (1) falls on the class of AII, to which $\mathbb{Z}_2$ topological insulators in two (2D) and three spatial

(3D) dimensions are classified. This class of models has the symmetry $\Theta^2 = -1$, $\mathcal{C}^2 = 0$, and $\Gamma_5 = 0$ (here "0" indicates that the system does not possess that type of symmetry), and are characterized by $\mathbb{Z}_2$-type bulk topological invariants both in 2D[39] and 3D.[40–42] For the specific choice $\epsilon(\boldsymbol{p}) = 0$, this symmetry is upgraded to class DIII, yielding $\Theta^2 = -1$, $\mathcal{C}^2 = 1$, and $\Gamma_5 = 1$, where for the specific Hamiltonian (1) $\mathcal{C}$ and Γ_5 are given by $\mathcal{C} = \sigma_y\tau_y K$ and $\Gamma_5 = \tau_y$. This symmetry class obeys to a $\mathbb{Z}$-type bulk topological classification, characterized by a winding number $\mathcal{N}$ for the Wilson-Dirac-like operator in three dimensions.

In the following we explicitly construct and evaluate this $\mathbb{Z}$-type winding number $\mathcal{N}$. Here, to ease the notation, we rewrite Eq. (1) in the specific case of $\epsilon(\boldsymbol{p}) = 0$ as

$$H = A\tau_1\sigma_\mu p_\mu + \tau_3 m(\boldsymbol{p}), \tag{A1}$$

where $\mu = 1, 2, 3$ and $m(\boldsymbol{p}) \equiv m_0 + m_2 p_\mu^2$. Note that

$$H^2 = A^2 p_\mu^2 + m^2(\boldsymbol{p}) \equiv R^2(\boldsymbol{p})\mathbf{1}. \tag{A2}$$

Therefore, a deformed Hamiltonian

$$\widetilde{H} \equiv \frac{H}{R} \tag{A3}$$

has eigenvalues ± 1. To characterize the topological property of $\widetilde{H}$, let us make a rotation in the τ space such that $\tau_1 \to \tau_2$, $\tau_2 \to \tau_3$, $\tau_3 \to \tau_1$. Then the Hamiltonian is converted into

$$\begin{aligned}\widetilde{H} &\to \frac{A\tau_2\sigma_\mu p_\mu + \tau_1 m(\boldsymbol{p})}{R} \\ &= \frac{1}{R}\begin{pmatrix} & m(\boldsymbol{p}) - iA\sigma_\mu p_\mu \\ m(\boldsymbol{p}) + iA\sigma_\mu p_\mu & \end{pmatrix} \\ &\equiv \begin{pmatrix} & Q(\boldsymbol{p}) \\ Q^\dagger(\boldsymbol{p}) & \end{pmatrix}.\end{aligned} \tag{A4}$$

Here a 2×2 SU(2) matrix Q emerges,

$$Q \equiv \frac{m(\boldsymbol{p}) - iA\sigma_\mu p_\mu}{R}, \tag{A5}$$

satisfying $Q^\dagger Q = \mathbf{1}$ and $\det Q = 1$. Since $\pi_3[\mathrm{SU}(2)] = \mathbb{Z}$, Q should be characterized by an integer winding number generically.

The winding number $\mathcal{N}$ is defined by

$$\begin{aligned}\mathcal{N} &= \frac{1}{24\pi^2}\int \mathrm{tr}\,(Q^\dagger dQ)^3 \\ &= \frac{1}{24\pi^2}\int d^3p\,\epsilon_{\mu\nu\rho}\mathrm{tr}\,(Q^\dagger\partial_\mu Q)(Q^\dagger\partial_\nu Q)(Q^\dagger\partial_\rho Q),\end{aligned} \tag{A6}$$

where d stands for the exterior derivative with respect to p_μ, and ∂_μ is the derivative with respect to p_μ. For the normalization of the above equation, see, for example, Eq. (66) of Ref. 43.

Let us look into the nature of this winding number more precisely. Since Q is a function of p_μ, the winding number $\mathcal{N}$ characterizes the mapping from $\mathbb{R}^3$ to SU(2). However, at the infinity of $\mathbb{R}^3$, namely, $|p| \to \infty$, Q becomes a single element of SU(2), $Q \to \mathrm{sgn}(m_2)\mathbf{1}$. Therefore, in the present mapping, the infinity $|p| \to \infty$ can be regarded as a single point, which make it possible to regard $\mathbb{R}^3$ as $\mathbb{S}^3$. By the use of the fact that $\mathrm{SU}(2) \sim \mathbb{S}^3$, the winding number $\mathcal{N}$ characterizes the mapping from $\mathbb{S}^3$ to $\mathbb{S}^3$, which can be classified by $\pi_3(\mathbb{S}^3) = \mathbb{Z}$.

It is not difficult to guess the winding number in the following way. At the origin $|p| = 0$, $Q = \mathrm{sgn}(m_0)\mathbf{1}$. Therefore, if $\mathrm{sgn}(m_0 m_2) = 1$, Q can be deformed into $Q = \mathrm{sgn}(m_0)\mathbf{1}$ by taking the limit $|m_0| \to \infty$, giving rise to a trivial winding number $\mathcal{N} = 0$. On the other hand, if $\mathrm{sgn}(m_0 m_2) = -1$, Q at $|p| = 0$ cannot be deformed into Q at $|p| = \infty$, so this case gives $\mathcal{N} = \pm 1$.

Let us compute the winding number $\mathcal{N}$ concretely. After a tedious but straightforward calculation, one finds

$$\epsilon_{\mu\nu\rho}\mathrm{tr}\,(Q^\dagger\partial_\mu Q)(Q^\dagger\partial_\nu Q)(Q^\dagger\partial_\rho Q) = 12\frac{\boldsymbol{p}^2 - m(\boldsymbol{p})}{R^4}. \tag{A7}$$

We reach, therefore,

$$\begin{aligned}\mathcal{N} &= \frac{-12}{24\pi^2}\int_{\mathrm{R}^3} d^3p\,\frac{m_0 - m_2\boldsymbol{p}^2}{[\boldsymbol{p}^2 + (m_0 + m_2\boldsymbol{p}^2)^2]^2} \\ &= \frac{-12\times(4\pi)}{24\pi^2}\int_0^\infty dp\,\frac{p^2(m_0 - m_2 p^2)}{[p^2 + (m_0 + m_2 p^2)^2]^2} \\ &= -\frac{1}{2}[\mathrm{sgn}(m_0) - \mathrm{sgn}(m_2)].\end{aligned} \tag{A8}$$

The last formula indicates that the system is indeed in the topologically nontrivial ($\mathcal{N} = \pm 1$) phase when m_0 and m_2 have the *opposite* sign, albeit in the trivial ($\mathcal{N} = 0$) phase when they have the *same* sign.

APPENDIX B: THE ZERO-ENERGY CONDITION

The radial eigenvalue problem considered in Sec. II has two basic ingredients: The eigenvalue equation (5) and the boundary condition (7) at $r = R$ (on the surface of the sphere). Here we prove that the solutions of this radial boundary problem satisfies automatically the energy condition $E_\perp = 0$. This observation paves the way for constructing the basis eigenspinors given in Eqs. (17). In other "simpler" geometries, such as the case as a semi-infinite system with a flat boundary,[28] or a cylindrical system infinitely long in the axial direction,[13] the scenario applies, but the explicit use of the zero-energy condition may not be indispensable because of the simplicity of the problem.

We first need to find the explicit matrix form of $H_\perp = H_{\mathrm{bulk}}|_{p_\theta=0,\,p_\phi=0}$. Focus on the three-momentum operators $p_\pm = -i\partial_\pm$, $p_z = -i\partial_z$ that have appeared in Eq. (3), and in addition, $\boldsymbol{p}^2 = -\nabla^2$ in $m(\boldsymbol{p})$. In the spherical coordinates, these can be expressed as

$$\begin{aligned}\partial_\pm &= e^{\pm i\phi}\left[\sin\theta\frac{\partial}{\partial r} + \frac{1}{r}\left(\cos\theta\frac{\partial}{\partial\theta} \pm \frac{i}{\sin\theta}\frac{\partial}{\partial\phi}\right)\right] \\ &\equiv e^{\pm i\phi}\sin\theta\frac{\partial}{\partial r} + D_\pm,\end{aligned} \tag{B1}$$

$$\partial_z = \cos\theta\frac{\partial}{\partial r} - \frac{\sin\theta}{r}\frac{\partial}{\partial\theta} \equiv \cos\theta\frac{\partial}{\partial r} + D_z, \tag{B2}$$

and

$$\nabla^2 = \frac{\partial}{\partial r}^2 + \frac{2}{r}\frac{\partial}{\partial r} + \frac{\boldsymbol{L}^2}{r^2}, \tag{B3}$$

where $\boldsymbol{L}^2$ is a square of the orbital angular momentum operator. Since $D_\pm$, D_z, and $\boldsymbol{L}^2$ involve only angular derivatives, we put them in $H_\parallel$. Keeping only the first terms of Eqs. (B1), (B2),

and (B3), one finds

$$H_\perp[\kappa] = \begin{bmatrix} m_\perp[\kappa] & -i\kappa a & 0 & -i\kappa b e^{-i\phi} \\ -i\kappa a & -m_\perp[\kappa] & -i\kappa b e^{-i\phi} & 0 \\ 0 & -i\kappa b e^{i\phi} & m_\perp[\kappa] & i\kappa a \\ -i\kappa b e^{i\phi} & 0 & i\kappa a & -m_\perp[\kappa] \end{bmatrix}, \tag{B4}$$

where we have assumed a surface solution of the form of Eq. (6), and replaced the r derivatives with κ, the inverse of the penetration depth. Equation (B4) can be indeed regarded as a 4×4 c-number matrix specified by a parameter κ. We used the notation $H_\perp[\kappa]$ to make this point explicit. We have also introduced, for shortening the expression, the notations $a = A\cos\theta$, $b = A\sin\theta$, and

$$m_\perp[\kappa] = m_0 - m_2\left(\kappa^2 + \frac{2}{r}\kappa\right). \tag{B5}$$

Here in Eq. (B5) an r dependence in the last term looks cumbersome. But as long as the surface wave function is well localized in the vicinity of the surface at $r = R$, one can safely replace this coordinate r by a constant R. On the other hand, as long as the same assumption is applied, this last term itself becomes negligible since as long as $\kappa^{-1} \ll R$, the second term is much larger than the last term.

Thanks to the symmetric structure of the matrix form of Eq. (B4), the secular equation $\det|H_\perp[\kappa] - E_\perp| = 0$ for the radial eigenvalue problem (5) becomes as simple as

$$\det|H_\perp[\kappa] - E_\perp| = \{\kappa^2 A^2 - m_\perp[\kappa]^2 + E_\perp^2\}^2 = 0. \tag{B6}$$

This can be regarded as a quadratic equation for κ^2 under the approximation of $m_\perp[\kappa] = m_0 - m_2\kappa^2$, with its two solutions $\kappa_\pm^2$ satisfying

$$\kappa_+^2\kappa_-^2 = \frac{m_0^2 - E_\perp^2}{m_2^2}. \tag{B7}$$

Now, in order to cope with the boundary condition (7), two surface solutions of the form of Eq. (6), one with $\kappa = \kappa_+$ and the other with $\kappa = \kappa_-$ must be superposed, that is,

$$|\psi\rangle = c_+ e^{\kappa_+(r-R)}\boldsymbol{u}[\kappa_+] + c_- e^{\kappa_-(r-R)}\boldsymbol{u}[\kappa_-], \tag{B8}$$

where $\boldsymbol{u}[\kappa]$ is an eigenvector of $H_\perp[\kappa]$ given in Eq. (B4). The only way that this solution can be compatible with the boundary condition (7) is to have simultaneously $c_+ + c_- = 0$ and $\boldsymbol{u}[\kappa_+] = \boldsymbol{u}[\kappa_-]$, that is, $|\psi\rangle$ takes the following form:

$$|\psi\rangle = N[e^{\kappa_+(r-R)} - e^{\kappa_-(r-R)}]\boldsymbol{u}[\kappa_+] \equiv \rho(r)\boldsymbol{u}[\kappa_+], \tag{B9}$$

where N is a normalization constant.

In the reduction from Eq. (B8) to Eq. (B9), the second condition stating that two eigenvectors belonging to different κ's should coincide, was crucial. Indeed, this coincidence occurs only under a very specific condition. In order to clarify this point, let us consider the following quantity:

$$\Delta[\kappa] \equiv \frac{H_\perp[\kappa] - E_\perp}{\kappa} = \begin{bmatrix} M_1[\kappa] & -ia & 0 & -ibe^{-i\phi} \\ -ia & -M_2[\kappa] & -ibe^{-i\phi} & 0 \\ 0 & -ibe^{i\phi} & M_1[\kappa] & ia \\ -ibe^{i\phi} & 0 & ia & -M_2[\kappa] \end{bmatrix}, \tag{B10}$$

where

$$M_1[\kappa] = \frac{m_\perp[\kappa] - E_\perp}{\kappa}, \quad M_2[\kappa] = \frac{m_\perp[\kappa] + E_\perp}{\kappa}. \tag{B11}$$

Notice that $\boldsymbol{u}[\kappa]$ is a zero-eigenvalue eigenvector of this matrix, that is, $\Delta[\kappa]\boldsymbol{u}[\kappa] = \boldsymbol{0}$. In order that two of such eigenvectors $\boldsymbol{u}[\kappa]$ belonging to different κ's ($=\kappa_\pm$), and therefore to different $\Delta[\kappa]$'s ($\Delta[\kappa_+]$ and $\Delta[\kappa_-]$) coincide, both $M_1[\kappa]$ and $M_2[\kappa]$ must coincide. Namely, one must have, simultaneously, $M_1[\kappa_+] = M_1[\kappa_-]$ and $M_2[\kappa_+] = M_2[\kappa_-]$. Clearly, a solution of the form of Eq. (B9) is meaningful only when $\kappa_+ \neq \kappa_-$. Therefore, $M_1[\kappa_+] = M_1[\kappa_-]$ signifies

$$\kappa_+\kappa_- = -\frac{m_0 + E_\perp}{m_2}, \tag{B12}$$

whereas $M_2[\kappa_+] = M_2[\kappa_-]$ leads to

$$\kappa_+\kappa_- = -\frac{m_0 - E_\perp}{m_2}. \tag{B13}$$

Recalling that Eqs. (B12) and (B13) must follow independently, one can convince oneself that the surface solution must satisfy the zero-energy condition

$$E_\perp = 0. \tag{B14}$$

Note that Eqs. (B12) and (B13) are consistent with Eq. (B7), but impose a stronger constraint on the values of κ's and $E_\perp$.

APPENDIX C: A BRIEF REMINDER ON THE JACOBI'S POLYNOMIALS/DIFFERENTIAL EQUATION

(1) The explicit form of the Jacobi's polynomials is given by the following (differential) Rodrigues' formula:

$$P_n^{\mu\nu}[\zeta] = \frac{(-1)^n}{2^n n!}\frac{1}{\rho_{\mu\nu}[\zeta]}\frac{d^n}{dx^n}[(1-\zeta^2)^n\rho_{\mu\nu}[\zeta]], \tag{C1}$$

where

$$\rho_{\mu\nu}[\zeta] = (1-\zeta)^\mu(1-\zeta)^\nu. \tag{C2}$$

As is clear from its construction, Eq. (C1) can be also expressed in the form of a contour integral (the *integral* Rodrigues' formula).

(2) The Jacobi's differential equation (46) is a simple rewriting of the hypergeometric differential equation,

$$\left[\xi(1-\xi)\frac{d^2}{d\xi^2} + \{\mathcal{C} - (1+\mathcal{A}+\mathcal{B})\xi\}\frac{d}{d\xi} - \mathcal{A}\mathcal{B}\right] \times F(\mathcal{A},\mathcal{B},\mathcal{C},\xi) = 0, \tag{C3}$$

by the change of the independent variable,

$$\xi = \frac{1-\zeta}{2}, \tag{C4}$$

and choice of the parameters,

$$\mathcal{A} = -n, \quad \mathcal{B} = n + \mu + \nu + 1, \quad \mathcal{C} = \mu + 1. \tag{C5}$$

(3) The Jacobi's polynomials $P_n^{\mu\nu}[\zeta]$ satisfy the following orthonormal relation:

$$\int_{-1}^{1} d\zeta\rho_{\mu\nu}[\zeta]P_{n_1}^{\mu\nu}[\zeta]P_{n_2}^{\mu\nu}[\zeta] = \delta_{n_1n_2}\frac{2^{\mu+\nu+1}\Gamma(n_1+\mu+1)\Gamma(n_1+\nu+1)}{n_1!(2n_1+\mu+\nu+1)\Gamma(n_1+\mu+\nu+1)}. \tag{C6}$$

APPENDIX D: PROOF OF EQ. (53)

In order to determine the relative magnitude and phase of c_{nm+} and c_{nm-}, one needs to go back to Eqs. (39) and substitute $\beta_{m\sigma}[\zeta]$ given in Eqs. (50) into this couple of equations [naturally, the change of the *dependent* variables must be taken into account; see Eq. (44)]. Changing the *independent* variable from θ to $\zeta = \cos\theta$, let us rewrite Eqs. (39) as

$$\begin{aligned}\left[\sqrt{1-\zeta^2}\frac{d}{d\zeta} - \frac{m+\zeta/2}{\sqrt{1-\zeta^2}}\right]\alpha_{m-}[\zeta] = \lambda\alpha_{m+}[\zeta],\\ -\left[\sqrt{1-\zeta^2}\frac{d}{d\zeta} + \frac{m-\zeta/2}{\sqrt{1-\zeta^2}}\right]\alpha_{m+}[\zeta] = \lambda\alpha_{m-}[\zeta].\end{aligned} \tag{D1}$$

Performing the derivatives explicitly, and changing the variables from $\alpha_{m\sigma}$ to $\beta_{m\sigma}$ [using Eq. (44)], one finds

$$\begin{aligned}\left[(1-\zeta)\frac{d}{d\zeta} - \left(m+\frac{1}{2}\right)\right]\beta_{m-}[\zeta] = \lambda\beta_{m+}[\zeta],\\ \left[-(1+\zeta)\frac{d}{d\zeta} - \left(m+\frac{1}{2}\right)\right]\beta_{m+}[\zeta] = \lambda\beta_{m-}[\zeta]\end{aligned} \tag{D2}$$

for $m \geqslant 1/2$, and

$$\begin{aligned}\left[(1+\zeta)\frac{d}{d\zeta} - \left(m-\frac{1}{2}\right)\right]\beta_{m-}[\zeta] = \lambda\beta_{m+}[\zeta],\\ \left[-(1-\zeta)\frac{d}{d\zeta} - \left(m-\frac{1}{2}\right)\right]\beta_{m+}[\zeta] = \lambda\beta_{m-}[\zeta]\end{aligned} \tag{D3}$$

for $m \leqslant 1/2$. Recall that $\beta_{m\sigma}$'s as given in Eq. (50) are proportional to the nth order Jacobi's polynomial. Equations (D2) and (D3) can be further simplified on account of the following identities [cf. Eqs. (A.7a) and (A.7b) of Ref. 23], applicable to the derivative of the Jacobi's polynomials with a *specific* choice of parameters μ and ν that are implied in these relations through Eq. (50), that is,

$$(1-\zeta)\frac{d}{d\zeta}P_n^{m+\frac{1}{2},m-\frac{1}{2}} = \left(m+\frac{1}{2}\right)P_n^{m+\frac{1}{2},m-\frac{1}{2}} - \left(n+m+\frac{1}{2}\right)P_n^{m-\frac{1}{2},m+\frac{1}{2}}, \tag{D4}$$

$$-(1+\zeta)\frac{d}{d\zeta}P_n^{m-\frac{1}{2},m+\frac{1}{2}} = \left(m+\frac{1}{2}\right)P_n^{m-\frac{1}{2},m+\frac{1}{2}} - \left(n+m+\frac{1}{2}\right)P_n^{m+\frac{1}{2},m-\frac{1}{2}}, \tag{D5}$$

where $m \geqslant 1/2$. These identities can be explicitly verified, for example, by the use of the integral counterpart of Eq. (C1).

The final part of the proof of Eq. (53) lies in the comparison of Eqs. (D2), (D3) and (D4), (D5). For $m \geqslant 1/2$, one can safely take off the operation of absolute value to the superscripts of Jacobi's polynomial in Eq. (50), yielding

$$\beta_{m\sigma}[\zeta] = \beta_{nm\sigma}[\zeta] = c_{nm\sigma}P_n^{m+\frac{\sigma}{2},m-\frac{\sigma}{2}}[\zeta]. \tag{D6}$$

Then, by simply comparing Eqs. (D2) with (D4) and (D5), and recalling $\lambda = \pm(n+m+1/2)$, one can verify $|c_{nm+}| = |c_{nm-}|$ with a relative sign of

$$c_{nm-} = -\text{sign}[\lambda]c_{nm+}. \tag{D7}$$

For $m \leqslant -1/2$, notice that

$$\left|m-\frac{\sigma}{2}\right| = |m| + \frac{\sigma}{2}, \tag{D8}$$

that is, for such m, Eq. (50) becomes

$$\beta_{m\sigma}[\zeta] = \beta_{nm\sigma}[\zeta] = c_{nm\sigma}P_n^{|m|+\frac{\sigma}{2},|m|-\frac{\sigma}{2}}[\zeta]. \tag{D9}$$

This allows for the use of Eqs. (D4) and (D5) with m replaced by $|m|$ in Eqs. (D3). Taking note of $\lambda = \pm(n+|m|+1/2)$, one can again verify $|c_{nm+}| = |c_{nm-}|$, but this time with a relative sign of opposite value,

$$c_{nm-} = \text{sign}[\lambda]c_{nm+}. \tag{D10}$$

The relations (D7) and (D10), respectively, for the two possible regimes of m complete the proof of Eq. (53).

[1] J. E. Moore, Nature (London) **464**, 194 (2010).
[2] X.-L. Qi and S.-C. Zhang, Rev. Mod. Phys. **83**, 1057 (2011).
[3] Y. Tanaka, M. Sato, and N. Nagaosa, J. Phys. Soc. Jpn. **81**, 011013 (2012).
[4] M. Z. Hasan and C. L. Kane, Rev. Mod. Phys. **82**, 3045 (2010).
[5] D.-H. Lee, Phys. Rev. Lett. **103**, 196804 (2009).
[6] A. K. Geim and K. S. Novoselov, Nat. Mater. **6**, 183 (2007).
[7] T. Ando, J. Phys. Soc. Jpn. **74**, 777 (2005).
[8] H. Peng, K. Lai, D. Kong, S. Meister, Y. Chen, X.-L. Qi, S.-C. Zhang, Z.-X. Shen, and Y. Cui, Nat. Mater. **9**, 225 (2010).
[9] Y. Zhang and A. Vishwanath, Phys. Rev. Lett. **105**, 206601 (2010).
[10] P. M. Ostrovsky, I. V. Gornyi, and A. D. Mirlin, Phys. Rev. Lett. **105**, 036803 (2010).
[11] J. H. Bardarson, P. W. Brouwer, and J. E. Moore, Phys. Rev. Lett. **105**, 156803 (2010).
[12] K.-I. Imura, Y. Takane, and A. Tanaka, Phys. Rev. B **84**, 035443 (2011).
[13] K.-I. Imura, Y. Takane, and A. Tanaka, Phys. Rev. B **84**, 195406 (2011).
[14] Y. Ran, Y. Zhang, and A. Vishwanath, Nat. Phys. **5**, 298 (2009).
[15] J. C. Y. Teo and C. L. Kane, Phys. Rev. B **82**, 115120 (2010).
[16] A. P. Schnyder, S. Ryu, A. Furusaki, and A. W. W. Ludwig, Phys. Rev. B **78**, 195125 (2008).
[17] A. Kitaev, AIP Conf. Proc. **1134**, 22 (2009).
[18] A. P. Schnyder, S. Ryu, A. Furusaki, and A. W. W. Ludwig, AIP Conf. Proc. **1134**, 10 (2009).
[19] S. Ryu, A. P. Schnyder, A. Furusaki, and A. W. W. Ludwig, New J. Phys. **12**, 065010 (2010).
[20] V. Parente, P. Lucignano, P. Vitale, A. Tagliacozzo, and F. Guinea, Phys. Rev. B **83**, 075424 (2011).

[21] J. González, F. Guinea, and M. A. H. Vozmediano, Phys. Rev. Lett. **69**, 172 (1992).
[22] J. González, F. Guinea, and M. Vozmediano, Nucl. Phys. B **406**, 771 (1993).
[23] A. A. Abrikosov, Jr., arXiv:hep-th/0212134.
[24] A. A. Abrikosov, Int. J. Mod. Phys. A **17**, 885 (2002).
[25] D. V. Kolesnikov and V. A. Osipov, Eur. Phys. J. B **49**, 465 (2006).
[26] H. Zhang, C.-X. Liu, X.-L. Qi, X. Dai, Z. Fang, and S.-C. Zhang, Nat. Phys. **5**, 438 (2010).
[27] C.-X. Liu, X.-L. Qi, H. J. Zhang, X. Dai, Z. Fang, and S.-C. Zhang, Phys. Rev. B **82**, 045122 (2010).
[28] W.-Y. Shan, H.-Z. Lu, and S.-Q. Shen, New J. Phys. **12**, 043048 (2010).
[29] M. König, H. Buhmann, L. W. Molenkamp, T. Hughes, C.-X. Liu, X.-L. Qi, and S.-C. Zhang, J. Phys. Soc. Jpn. **77**, 031007 (2008).
[30] K.-I. Imura, A. Yamakage, S. Mao, A. Hotta, and Y. Kuramoto, Phys. Rev. B **82**, 085118 (2010).
[31] T. Fukui and T. Fujiwara, J. Phys. A **42**, 362003 (2009).
[32] T. Eguchi, P. B. Gilkey, and A. J. Hanson, Phys. Rep. **66**, 213 (1980).
[33] S. Deguchi and K. Kitsukawa, Prog. Theor. Phys. **115**, 1137 (2006).
[34] T. T. Wu and C. N. Yang, Nucl. Phys. B **107**, 365 (1976).
[35] Y. Li and C. Wu, arXiv:1103.5422.
[36] Y. Li, K. Intriligator, Y. Yu, and C. Wu, Phys. Rev. B **85**, 085132 (2012).
[37] Y. Li, X. Zhou, and C. Wu, Phys. Rev. B **85**, 125122 (2012).
[38] M. R. Zirnbauer, J. Math. Phys. **37**, 4986 (1996).
[39] C. L. Kane and E. J. Mele, Phys. Rev. Lett. **95**, 146802 (2005).
[40] L. Fu, C. L. Kane, and E. J. Mele, Phys. Rev. Lett. **98**, 106803 (2007).
[41] J. E. Moore and L. Balents, Phys. Rev. B **75**, 121306 (2007).
[42] R. Roy, Phys. Rev. B **79**, 195322 (2009).
[43] K.-I. Imura, T. Fukui, and T. Fujiwara, Nucl. Phys. B **854**, 306 (2012).