Lecture notes on topological insulators

Ming-Che Chang

Department of Physics, National Taiwan Normal University, Taipei, Taiwan

(Dated: October 26, 2022)

CONTENTS

I. 2D TOPOLOGICAL INSULATOR

A. Z_2 topological number in topological insulator

1. Spinor Bloch state with time-reversal symmetry

The spin of an electron in a solid is often coupled with the electron's orbital motion via the **spin-orbit interaction** (SOI),

$$H_{so} = \lambda_{so}\boldsymbol{\sigma} \times \mathbf{p} \cdot \nabla V_L,\ \lambda_{so} = \frac{e\hbar}{4m^2c^2}, \tag{1.1}$$

where V_L is the lattice potential. Such an interaction is invariant under time-reversal symmetry (TRS), and invariant under space-inversion symmetry (SIS) *if* $V_L(-\mathbf{r}) = V_L(\mathbf{r})$. Because of the SOI, Bloch states $\psi_{n\mathbf{k}\pm}$, which are energy eigenstates, are in general not spin eigenstates $\psi_{n\mathbf{k}\uparrow/\downarrow}$.

Recall that the TR operator for fermion is,

$$\Theta = i\sigma_y K,\ \Theta^2 = -1. \tag{1.2}$$

In the presence of TRS, *if* the Bloch states are topologically trivial, then one can choose

$$\begin{cases} \Theta\psi_{n\mathbf{k}+} = -\psi_{n-\mathbf{k}-}, \\ \Theta\psi_{n\mathbf{k}-} = +\psi_{n-\mathbf{k}+}. \end{cases} \tag{1.3}$$

The $-$ sign in front of $\psi_{-\mathbf{k},-}$ is necessary because $\Theta^2 = -1$.

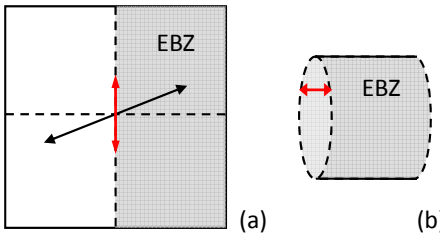

FIG. 1 (a) Time reversal conjugate pairs and effective Brillouin zone. (b) Folding the EBZ into a cylinder, with open edges.

The Bloch states, which are spinors now, are of the form,

$$\psi_{n\mathbf{k}+}(\mathbf{r}) = e^{i\mathbf{k}\cdot\mathbf{r}}\left[+a_{n\mathbf{k}}(\mathbf{r})\chi_\uparrow + b_{n\mathbf{k}}(\mathbf{r})\chi_\downarrow\right], \tag{1.4}$$
$$\psi_{n\mathbf{k}-}(\mathbf{r}) = e^{i\mathbf{k}\cdot\mathbf{r}}\left[-b^*_{n\mathbf{k}}(\mathbf{r})\chi_\uparrow + a^*_{n\mathbf{k}}(\mathbf{r})\chi_\downarrow\right], \tag{1.5}$$

where $a_{n\mathbf{k}}, b_{n\mathbf{k}}$ are cell-periodic functions, and $\chi_\uparrow = (1,0)^T, \chi_\downarrow = (0,1)^T$. If the SOI is weak, then $|b_{n\mathbf{k}}(\mathbf{r})| \ll 1$, so that $(+,-) \simeq (\uparrow,\downarrow)$. It is not uncommon to refer to $\pm$ simply as spin up/down.

If the Bloch states are topologically non-trivial, then one needs to write

$$\begin{cases} \Theta\psi_{n\mathbf{k}+} = -e^{i\chi_{n-k}}\psi_{n-\mathbf{k}-}, \\ \Theta\psi_{n\mathbf{k}-} = +e^{i\chi_{nk}}\psi_{n-\mathbf{k}+}. \end{cases} \tag{1.6}$$

It's possible *not* to have such a phase (in the so-called **TR-smooth gauge**). However, this would result in points of gauge singularity within the BZ.

2. Chern number

To understand the topology in topological insulator (TI), we follow Moore and Balent's argument for 2D TI (Moore and Balents, 2007). Because of time-reversal symmetry, the degenerate Bloch states for $\mathbf{k}$ and $-\mathbf{k}$ in a Brillouin zone are time-reversal conjugate (see Fig. 1(a)). As their Berry curvatures cancel with each other, the first Chern number for a filled band vanishes. Since the domain of *independent* Bloch states cover only half of the BZ (called **effective Brillouin zone**, or EBZ), one may wonder if the integral of the Berry curvature over the EBZ could be quantized.

Unfortunately, since the EBZ does not form a closed surface (see Fig. 1(b)), no quantization is guaranteed. To fix this, one can put two caps with TR conjugation to close the EBZ. This closed surface should have an integer C_1, but its value depends on the caps of choice. Nevertheless, Moore and Balents proved that, because of the TR conjugation, the caps can only change C_1 by an even integer. That is, C_1 mod 2 is independent of the caps of choice. Therefore, C_1 mod 2 should be an intrinsic property of the EBZ itself. We thus have two topological classes: 0 being the usual insulator, and 1 being the topological insulator. Hence, a 2D TI is characterized by a Z_2 topological number.

3. Winding number

Fu and Kane showed that the Z_2 topological number can be related to the winding number between two patches of gauge (Fu and Kane, 2006), which we now explain.

First, instead of Eq. (1.6), we will adopt the TR-smooth gauge,

$$\begin{cases} \Theta\psi_{n\mathbf{k}+} = -\psi_{n-\mathbf{k}-}, \\ \Theta\psi_{n\mathbf{k}-} = +\psi_{n-\mathbf{k}+}. \end{cases} \tag{1.7}$$

As a result, the phases of Bloch states cannot be uniquely defined over the whole BZ. As in the case of the magnetic monopole in Chap **??**, we need to use more than one patch of gauge to get rid of singularity. This is the *topological obstruction* mentioned earlier.

Fig. 2 shows the EBZ covered by two patches of gauge. Along their boundary, the wave functions are connected by gauge transformation,

$$|u_{\mathbf{k}\alpha}\rangle_B = U_{\alpha\beta}|u_{\mathbf{k}\beta}\rangle_A, \tag{1.8}$$

where U is a U(2) matrix. Again we consider only one Kramer pair for simplicity, so that the Berry connection is a 2×2 matrix.

Recall that the Berry connection and Berry curvature for band n are,

$$\mathbf{A}^n_{\alpha\beta}(\mathbf{k}) = i\langle u_{n\mathbf{k}\alpha}|\frac{\partial}{\partial\mathbf{k}}|u_{n\mathbf{k}\beta}\rangle, \tag{1.9}$$

$$\mathsf{F}^n_{k\ell} = \partial_k \mathsf{A}^n_\ell - \partial_\ell \mathsf{A}^n_k - i[\mathsf{A}^n_k, \mathsf{A}^n_\ell], \tag{1.10}$$

where α,β are "spin" indices $\pm$, and k,ℓ are space indices 1,2,3. Under a gauge transformation,

$$\mathsf{A}^B_\ell = \mathsf{U}^\dagger \mathsf{A}^A_\ell \mathsf{U} + i\mathsf{U}^\dagger\frac{\partial}{\partial k_\ell}\mathsf{U}. \tag{1.11}$$

The topology of the Bloch states can be characterized by the winding number w of the U(1) phase of U around the closed loop ∂A in Fig. 2,

$$w = \frac{1}{2\pi i}\oint_{\partial A} d\mathbf{k}\cdot \mathrm{tr}\left(\mathsf{U}^\dagger\frac{\partial}{\partial\mathbf{k}}\mathsf{U}\right). \tag{1.12}$$

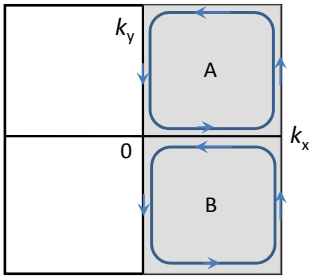

FIG. 2 Two patches of gauge in the EBZ.

Taking the trace of Eq. (1.11), we have

$$w = \frac{1}{2\pi}\oint_{\partial A} d\mathbf{k}\cdot\left(\mathbf{A}^A - \mathbf{A}^B\right), \tag{1.13}$$

in which $\mathbf{A}^{A/B} \equiv \mathrm{tr}\,\vec{\mathsf{A}}^{A/B}$.

Since $|u^A_{\mathbf{k}\alpha}\rangle$ is smoothly defined inside A, one has

$$\oint_{\partial A} d\mathbf{k}\cdot\mathbf{A}^A = \int_A d^2k\ F^A_z. \tag{1.14}$$

The same cannot be done for $|u^B_{\mathbf{k}\alpha}\rangle$, since it is *not* smoothly defined in A. Instead, we write

$$\begin{aligned}\oint_{\partial A} d\mathbf{k}\cdot\mathbf{A}^B &= \oint_{\partial EBZ} d\mathbf{k}\cdot\mathbf{A}^B - \oint_{\partial B} d\mathbf{k}\cdot\mathbf{A}^B \\ &= \oint_{\partial EBZ} d\mathbf{k}\cdot\mathbf{A}^B - \int_B d^2k\ F^B_z\end{aligned} \tag{1.15}$$

Finally, combine Eqs. (1.14) with (1.15), we have (Fu and Kane, 2006),

$$w = \frac{1}{2\pi}\left(\int_{EBZ} d^2k\ F_z - \oint_{\partial EBZ} d\mathbf{k}\cdot\mathbf{A}\right) \text{ mod } 2. \tag{1.16}$$

A modulo operation is imposed, since the second term is only gauge invariant modulo 2. This expression of w is different from the Chern number in systems with quantum Hall conductance, which only has the first term, and is an integral over a closed surface (the whole BZ).

Eq. (1.16) looks like the generalized **Gauss-Bonnet formula** for an open 2D surface M, in which the Berry curvature is replaced by the *Gaussian curvature* G, and the Berry connection is replace by the *geodesic curvature* k_g of the boundary,

$$\chi = \frac{1}{2\pi}\left(\int_M d^2r\ G - \oint_{\partial M} dr\ k_g\right). \tag{1.17}$$

For example, for a torus, $\chi = 0$, for a disk-like surface (which has a boundary), $\chi = 1$, and for a sphere, $\chi = 2$.

4. Z_2 topological number again

Fu and Kane found yet another way of calculating the Z_2 index (Fu and Kane, 2006). Its deduction is less

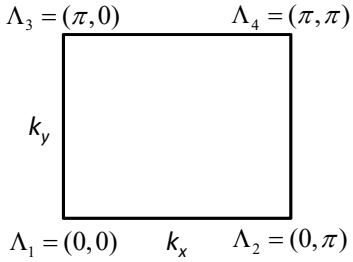

FIG. 3 The first quadrant of the Brillouin zone in a 2D lattice model. The four corners are the TRIM.

straightforward compared to the ones above, so here we will not explain how it is derived. Detailed explanation can be found in another set of my note.

Let's consider N filled Kramer pairs of Bloch bands in an insulator, and introduce the following quantity,

$$w_{m\alpha,n\beta}(\mathbf{k}) \equiv \langle u_{m-\mathbf{k}\alpha}|\Theta|u_{n\mathbf{k}\beta}\rangle, \quad n=1,\cdots,N; \alpha,\beta=\pm \tag{1.18}$$

They are the matrix elements of a $2N\times 2N$ matrix, often called the **sewing matrix**. Since the Bloch states from different bands are orthogonal to each other, one has

$$w_{m\alpha,n\beta}(\mathbf{k}) = \delta_{mn} w_{n\alpha\beta}(\mathbf{k}). \tag{1.19}$$

For example, for one Kramer pair, the sewing matrix is

$$\mathsf{w}_n = \begin{pmatrix} 0 & e^{i\chi_{n\mathbf{k}}} \\ -e^{i\chi_{n-\mathbf{k}}} & 0 \end{pmatrix}. \tag{1.20}$$

At a TRIM, it becomes an antisymmetric matrix,

$$\mathsf{w}_n = w_n(\mathbf{\Lambda}) \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, w_n(\mathbf{\Lambda}) \equiv e^{i\chi_{n\Lambda}}. \tag{1.21}$$

For *filled* bands at *each* TRIM, one can define

$$\delta_i = \prod_{n \text{ filled}} \frac{w_n(\Lambda_i)}{\sqrt{w_n^2(\Lambda_i)}}. \tag{1.22}$$

Note that if the argument of a complex number $z=re^{i\theta}$ is restricted to $[0,2\pi)$, then $\sqrt{z^2}$ has two possible values: If $\theta\in[0,\pi)$, then

$$\frac{z}{\sqrt{z^2}} = \frac{re^{i\theta}}{r\left(e^{2i\theta}\right)^{1/2}} = 1. \tag{1.23}$$

However, if $\theta\in[\pi,2\pi)$, then

$$\frac{z}{\sqrt{z^2}} = \frac{re^{i(\pi+\tilde{\theta})}}{r\left(e^{2i\tilde{\theta}}\right)^{1/2}}, \ \tilde{\theta}\in[0,\pi] \tag{1.24}$$

$$= -1. \tag{1.25}$$

Thus, $z/\sqrt{z^2}$ can be $+1$ or -1. That is, the δ_i above is product of $+1$ and -1.

Finally, the Z_2 topological index ν of a topological insulator is related to δ_i's,

$$(-1)^\nu = \delta_1\delta_2\delta_3\delta_4. \tag{1.26}$$

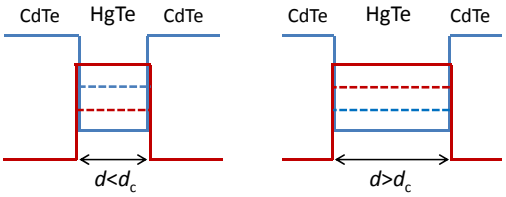

FIG. 4 HgTe is sandwitched between CdTe, forming a quantum well. The positions of the discrete energy levels in the QW depend on the width of the QW.

5. Lattice with inversion symmetry

Even though ν is known to have two possible values, 0 and 1, it is not easy to get explicit values of $\chi_{n\Lambda_i}$. Fortunately, *if the lattice has space inversion symmetry* (SIS), then we can determine ν from the parity $\zeta_n(\Lambda_i)$ of the Bloch state $\psi_{n\Lambda_i\pm}$ at TRIM.

If the lattice has SIS, then $\psi_{n\mathbf{k}\alpha}$ are parity eigenstates at $\mathbf{k}=\Lambda_i$,

$$\Pi\psi_{n\Lambda_i\alpha}(\mathbf{r}) = \zeta_{n\Lambda_i}\psi_{n\Lambda_i\alpha}(\mathbf{r}). \tag{1.27}$$

The parity eigenvalue $\zeta_{n\Lambda_i}=1$ or -1 is the same for the two Bloch states (with $\alpha=\pm$) in a Kramer pair. Fu *et al.*, 2007 showed that

$$w_n(\Lambda_i) = \zeta_n(\Lambda_i), \tag{1.28}$$

hence

$$\delta_i = \prod_{n\in\text{filled}} \zeta_n(\Lambda_i). \tag{1.29}$$

It is the cumulative parity of filled Bloch states (pick only one $\zeta_n(\Lambda_i)$ for each Kramer pair) at a TRIM. Finally,

$$(-1)^\nu = \prod_{i=1}^{4} \delta_i. \tag{1.30}$$

Band inversion often results in a change of the parity ζ_n, and this results in topological phase transition. For a crystal without inversion symmetry, one can deform it to one that has SIS and determine its ν by parities. This is a valid shortcut only *if* the energy gap remains open during the process of deformation.

B. Bernevig-Hughes-Zhang model

There is a close relation between the topology of energy band and the *band inversion* (*negative* energy gap) near the chemical potential. The first experimentally confirmed (2D) TI is made from semiconductor quantum well. Bulk HgTe has inverted band structure (due to spin-orbit coupling) near the Fermi energy. Unfortunately, it's a metal, not an insulator. Nevertheless, one can sandwitch it between CdTe (an ordinary insulator),

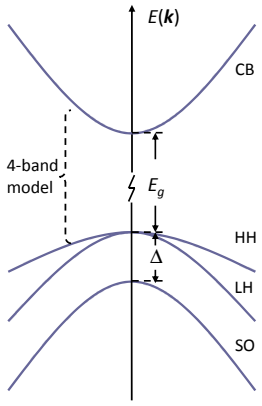

FIG. 5 Typical energy bands near the energy gap of a semiconductor, in which each line is two-fold degenerate. The conduction band is originated from the s-orbital. The valence band from the p-orbitals are composed of heavy-hole (HH) band, light-hole (LH) band, and spin-orbit split-off (SO) band.

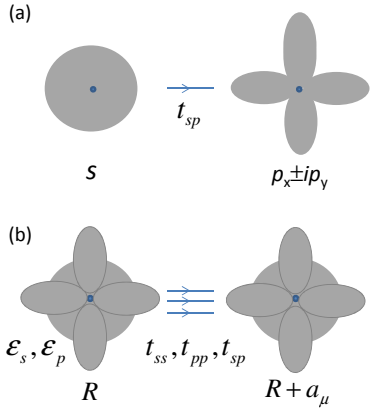

FIG. 6 (a) The hopping amplitude of an electron hopping from a s-orbital to the $p_x \pm ip_y$ orbitals of a nearest-neighbor atom is designated as t_{sp}. (b) Each atom has s-orbital and p-orbital, with energies ε_s and ε_p. An electron can hop between s-orbitals, between p-orbitals, or between an s-orbital and a p-orbital.

forming a quantum well (QW) and opening an energy gap near the Fermi energy (see Fig. 4). When the HgTe layer is thick, the discrete QW energy levels remain inverted, similar to the bulk states. However, if the HgTe layer is thinner than a critical width d_c, then the electron (E1) and hole (H1) levels in the QW would switch positions. Therefore, one can check if the topology of the QW states (signified by the emergence of helical edge states) depends on the width of the QW (König *et al.*, 2007).

Typical semiconductor band structure near $k = 0$ is shown in Fig. 5. In a QW, the LH bands split off the HH bands, so in the simplified Bernevig-Hughes-Zhang (BHZ) model, only conduction band and one HH valence band are considered (each are two-fold degenerate). In order to investigate the parities of Bloch states at TRIM, we follow the lattice version of the BHZ model proposed by Fu and Kane, 2007 (also, see Nomura, 2016)

For an atom at site $\mathbf{R}$, four states are considered,

$$|s\uparrow\rangle, |s\downarrow\rangle, |p_x + ip_y\uparrow\rangle, |p_x - ip_y\downarrow\rangle. \tag{1.31}$$

Without ambiguity, one can write $|p_x + ip_y\uparrow\rangle$ and $|p_x - ip_y\downarrow\rangle$ simply as $|p\uparrow\rangle$ and $|p\downarrow\rangle$. They are the states with quantum numbers $m_j = 3/2$ and $m_j = -3/2$.

We consider only the electron hopping between nearest neighbors. The relevant parameters are on-site energies $\varepsilon_s, \varepsilon_p$, and hopping amplitudes t_{ss}, t_{pp}, and t_{sp} (and its complex conjugate t_{ps}), see Fig. 6. In a 2D square lattice, the vectors $\mathbf{R} + \mathbf{a}_\mu$ ($\mu = \pm x, \pm y$) point to the four nearest neighbors of $\mathbf{R}$. The tight-binding Hamiltonian is therefore given as,

$$\begin{aligned} H &= H_0 + H_1, \\ H_0 &= \sum_{\mathbf{R}\sigma=\pm} \left(\varepsilon_s |\mathbf{R}s\sigma\rangle\langle\mathbf{R}s\sigma| + \varepsilon_p |\mathbf{R}p\sigma\rangle\langle\mathbf{R}p\sigma|\right), \end{aligned} \tag{1.32}$$

and (given $t_{ps} = t_{sp}$)

$$\begin{aligned} H_1 = -\sum_{\mathbf{R}\sigma} \sum_{\mu=\pm x,\pm y} & \left(t_{ss} |\mathbf{R} + \mathbf{a}_\mu s\sigma\rangle\langle\mathbf{R}s\sigma| \right. \\ & - t_{pp} |\mathbf{R} + \mathbf{a}_\mu p\sigma\rangle\langle\mathbf{R}p\sigma| \\ & + e^{i\theta_\mu \sigma} t_{sp} |\mathbf{R} + \mathbf{a}_\mu s\sigma\rangle\langle\mathbf{R}p\sigma| \\ & \left. + e^{-i\theta_\mu \sigma} t_{sp} |\mathbf{R}p\sigma\rangle\langle\mathbf{R} + \mathbf{a}_\mu s\sigma| \right), \end{aligned} \tag{1.33}$$

where $\theta_\mu = \angle(\hat{x}, \mathbf{a}_\mu)$. That is, $\theta_x = 0, \theta_y = \pi/2, \theta_{-x} = \pi, \theta_{-y} = 3\pi/2$. Such a system has both TRS and SIS (details later).

Because of the lattice translation symmetry, the Hamiltonian can be diagonalized using the momentum basis. We therefore introduce the Fourier transformation (N is the total number of lattice sites),

$$|\mathbf{R}s\sigma\rangle = \frac{1}{\sqrt{N}} \sum_{\mathbf{k}} e^{i\mathbf{k}\cdot\mathbf{R}} |\mathbf{k}s\sigma\rangle, \tag{1.34}$$

$$|\mathbf{R}p\sigma\rangle = \frac{1}{\sqrt{N}} \sum_{\mathbf{k}} e^{i\mathbf{k}\cdot\mathbf{R}} |\mathbf{k}p\sigma\rangle, \tag{1.35}$$

and get (the lattice constant is set as one)

$$H = \sum_{\mathbf{k}} \left(\langle\mathbf{k}s\uparrow|, \langle\mathbf{k}s\downarrow|, \langle\mathbf{k}p\uparrow|, \langle\mathbf{k}p\downarrow|\right) \mathsf{H}(\mathbf{k}) \begin{pmatrix} |\mathbf{k}s\uparrow\rangle \\ |\mathbf{k}s\downarrow\rangle \\ |\mathbf{k}p\uparrow\rangle \\ |\mathbf{k}p\downarrow\rangle \end{pmatrix},$$

where

$$\begin{aligned} &\mathsf{H}(\mathbf{k}) \\ &= \begin{pmatrix} \varepsilon_s - 2t_{ss}(\cos k_x + \cos k_y) & 2t_{sp}(\sigma_z \sin k_y - i \sin k_x) \\ 2t_{sp}(\sigma_z \sin k_y + i \sin k_x) & \varepsilon_p + 2t_{pp}(\cos k_x + \cos k_y) \end{pmatrix}. \end{aligned} \tag{1.36}$$

Or,

$$\begin{aligned}\mathsf{H}(\mathbf{k}) &= \left[\frac{\varepsilon_s+\varepsilon_p}{2}-(t_{ss}-t_{pp})(\cos k_x+\cos k_y)\right]1\otimes 1\\ &+\left[\frac{\varepsilon_s-\varepsilon_p}{2}-(t_{ss}+t_{pp})(\cos k_x+\cos k_y)\right]\tau_z\otimes 1\\ &+\ 2t_{sp}\sin k_x\tau_y\otimes 1+2t_{sp}\sin k_y\tau_x\otimes\sigma_z. \qquad (1.37)\end{aligned}$$

The Pauli matrices $\tau_{x,y,z}$ are for the orbital degree of freedom, and $\sigma_{x,y,z}$ are for the spin degree of freedom.

1. Time reversal and space inversion

Time reversal flips spin, but not orbital, therefore it acts only on the spin degree of freedom. The TR operator is thus,

$$\Theta=\begin{pmatrix} i\sigma_y K & 0\\ 0 & i\sigma_y K\end{pmatrix}=1\otimes i\sigma_y K. \qquad (1.38)$$

The s-orbital is even under space inversion, while the p-orbital is odd under SI. That is,

$$\begin{aligned}\Pi|\mathbf{k}s\sigma\rangle &= |\mathbf{k}s\sigma\rangle, \qquad (1.39)\\ \Pi|\mathbf{k}p\sigma\rangle &= -|\mathbf{k}p\sigma\rangle. \qquad (1.40)\end{aligned}$$

Therefore,

$$\Pi=\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}=\tau_z\otimes 1. \qquad (1.41)$$

One can check that the Hamiltonian $\mathsf{H}(\mathbf{k})$ is indeed invariant under these two transformations,

$$\begin{aligned}\Theta\mathsf{H}(\mathbf{k})\Theta^{-1} &= \mathsf{H}(-\mathbf{k}), \qquad (1.42)\\ \Pi\mathsf{H}(\mathbf{k})\Pi^{-1} &= \mathsf{H}(-\mathbf{k}). \qquad (1.43)\end{aligned}$$

2. Z_2 topological number

Since the BHZ model has SI symmetry, one can calculate the Z_2 topological number from the parities of the Bloch states at TRIM (see Fig. 3(b)). At a TRIM, the first line of Eq. 1.37 contributes a constant energy shift and can be ignored, and the third line is zero. Therefore,

$$\begin{aligned}\mathsf{H}(\Lambda_1) &= \left[\frac{\varepsilon_s-\varepsilon_p}{2}-2(t_{ss}+t_{pp})\right]\tau_z\otimes 1, \qquad (1.44)\\ \mathsf{H}(\Lambda_2) &= \left(\frac{\varepsilon_s-\varepsilon_p}{2}\right)\tau_z\otimes 1, \qquad (1.45)\\ \mathsf{H}(\Lambda_3) &= \left(\frac{\varepsilon_s-\varepsilon_p}{2}\right)\tau_z\otimes 1, \qquad (1.46)\\ \mathsf{H}(\Lambda_4) &= \left[\frac{\varepsilon_s-\varepsilon_p}{2}+2(t_{ss}+t_{pp})\right]\tau_z\otimes 1. \qquad (1.47)\end{aligned}$$

Note that they are proportional to the parity operator,

$$\mathsf{H}(\Lambda_i)=g_i\Pi,\quad \Pi=\tau_z\otimes 1. \qquad (1.48)$$

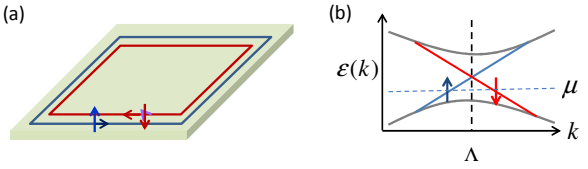

FIG. 7 (a) In real space, there is a helical edge state along the boundary of a 2D TI. (b) In momentum space, the energy dispersion curves of the edge states cross each other at a TRIM.

Hence, an energy eigenstate at Λ_i is also a parity eigenstate,

$$\Pi\psi=\zeta\psi \ \leftrightarrow\ \mathsf{H}(\Lambda_i)\psi=(g_i\zeta)\psi. \qquad (1.49)$$

Let's assume $\varepsilon_s>\varepsilon_p$ in the following discussion. If $\varepsilon_s-\varepsilon_p>4(t_{ss}+t_{pp})>0$, then the energies at Λ_1 are

$$\begin{aligned}\varepsilon_{1+} &= +\left[\frac{\varepsilon_s-\varepsilon_p}{2}-2(t_{ss}+t_{pp})\right], \qquad (1.50)\\ \varepsilon_{1-} &= -\left[\frac{\varepsilon_s-\varepsilon_p}{2}-2(t_{ss}+t_{pp})\right]. \qquad (1.51)\end{aligned}$$

The degenerate eigenstates ψ_{1+} have even parity, while ψ_{1-} have odd parity. Only the state ψ_{1-} is filled, so $\delta(\Lambda_1)=-1$. Similarly, one can also get $\delta(\Lambda_2)=\delta(\Lambda_3)=\delta(\Lambda_4)=-1$. Therefore,

$$(-1)^\nu=1,\ \text{or}\ \nu=0. \qquad (1.52)$$

This is a trivial insulator.

On the other hand, if $\varepsilon_s-\varepsilon_p<4(t_{ss}+t_{pp})$, then $\varepsilon_{1-}>\varepsilon_{1+}$. Due to the band inversion, now the states ψ_{1+} are filled instead. Therefore, $\delta(\Lambda_1)=1$. The other three parities are not changed. As a result,

$$(-1)^\nu=-1,\ \text{or}\ \nu=1. \qquad (1.53)$$

This is a topological insulator.

At the critical point, $\varepsilon_s-\varepsilon_p=4(t_{ss}+t_{pp})$, the energy gap closes. The transition from the trivial phase to the topological phase is accompanied by an *inversion of energy bands*, when ε_+ and ε_- switch positions.

3. Orbital basis versus spin basis

Note that in Eq. 1.36, every 2×2 block of the Hamiltonian matrix is diagonal. In this case, it is possible to block-diagonal the 4×4 matrix by re-arranging the order of the basis,

$$\begin{aligned}&|s\uparrow\rangle,|s\downarrow\rangle;|p\uparrow\rangle,|p\downarrow\rangle \qquad (1.54)\\ \rightarrow\ &|s\uparrow\rangle,|p\uparrow\rangle;|s\downarrow\rangle,|p\downarrow\rangle. \qquad (1.55)\end{aligned}$$

For convenience, we call the first choice the orbital basis, and the second the spin basis.

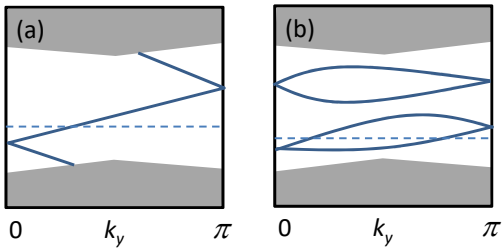

FIG. 8 Comparison of the edge states in (a) 2D topological insulator, and (b) 2D trivial insulator.

Under the spin basis, the Hamiltonian becomes (first switch the 2nd and the 3rd rows, then switch the 2nd and the 3rd columns of the matrix)

$$\mathsf{H}(\mathbf{k}) = \begin{pmatrix} \mathsf{h}(\mathbf{k}) & 0 \\ 0 & \mathsf{h}^*(-\mathbf{k}) \end{pmatrix}, \tag{1.56}$$

where

$$\begin{aligned} &\mathsf{h}(\mathbf{k}) \\ =& \begin{pmatrix} \varepsilon_s - 2t_{ss}(\cos k_x + \cos k_y) & 2t_{sp}(-i\sin k_x + \sin k_y) \\ 2t_{sp}(i\sin k_x + \sin k_y) & \varepsilon_p + 2t_{pp}(\cos k_x + \cos k_y) \end{pmatrix} \\ =& \ \varepsilon_0(\mathbf{k}) + 2t_{sp}\sin k_x \tau_y + 2t_{sp}\sin k_y \tau_x \\ +& \left[\frac{\varepsilon_s - \varepsilon_p}{2} - (t_{ss} + t_{pp})(\cos k_x + \cos k_y)\right]\tau_z. \end{aligned} \tag{1.57}$$

Because of the block diagonalization, the up and down spins are explicitly decoupled.

Note that the TR and SI operators also are altered under the new basis. They now become

$$\Theta = i\sigma_y K \otimes 1, \ \Pi = 1 \otimes \tau_z. \tag{1.58}$$

4. QWZ model versus BHZ model

When written in the block-diagonal form, the Hamiltonian $\mathsf{h}(\mathbf{k})$ is similar to the QWZ model Hamiltonian for the QAHE (see **??**). That is, the BHZ model is composed of two independent QWZ subsystems, $\mathsf{h}(\mathbf{k})$ and $\mathsf{h}^*(-\mathbf{k})$.

One can write

$$\mathsf{h}(\mathbf{k}) = \varepsilon_0(\mathbf{k}) + \mathbf{d}(\mathbf{k})\cdot\boldsymbol{\tau}, \tag{1.59}$$

then the Hall conductivity of the subsystem is,

$$\sigma_{xy} = -\frac{e^2}{h}\frac{1}{4\pi}\int_{BZ} d^2k \frac{1}{d^3}\mathbf{d}\cdot\frac{\partial \mathbf{d}}{\partial k_x}\times\frac{\partial \mathbf{d}}{\partial k_y}. \tag{1.60}$$

One can check that if $\varepsilon_s - \varepsilon_p > 4(t_{ss} + t_{pp})$, then the signs of $d_z(\mathbf{k})$ are all positive at the TRIM (analogous to Fig. **??**(a)). If $\varepsilon_s - \varepsilon_p < 4(t_{ss} + t_{pp})$, then $d_z(\Lambda_1)$ becomes negative, while the other three signs remain the same (see Fig. **??**(b)). According to the analysis of the QAHE in **??**, the first case has $\sigma_{xy} = 0$, while the second case has $\sigma_{xy} = e^2/h$.

When the subsystem is in the QAH phase, according to the discussion in **??**, it has chiral edge states. Since this subsystem consists only of spin-up electrons (see Eq. 1.55), the edge-state electrons are spin-up. On the other hand, the conjugate subsystem $h^*(-\mathbf{k})$ has $\sigma_{xy} = -e^2/h$. Its edge electrons transport along the *opposite* direction and the spins are *down* (see Fig. 7(a)).

In momentum space, the energy dispersion of the edge state is linear in the small k-limit. One has positive slope (positive velocity), and the other has negative slope (negative velocity), see Fig. 7(b). Because of the Kramer degeneracy, these two dispersion curves have to cross each other at a TRIM. This point degeneracy can be lifted *only if* the TRS is broken.

The topological phase of the BHZ model is called a 2D TI phase, aka a **quantum spin Hall** (QSH) phase. Its edge state, with one spin moving along one direction, and the opposite spin moving along the opposite direction, is called a **helical edge state**. It is robust in the sense that, even if there is a *non-magnetic* impurity $V_{imp}(\mathbf{r})$ blocking the way, the electron will not be back scattered since that requires a spin flip.

In general, if one plots a horizontal line (chemical potential) inside the energy gap, then it would cut the edge states of a TI odd number of times (Fig. 8(a)). For a trivial insulator, the chemical potential would cut its edge states even number of times (Fig. 8(b)). The former cannot be avoided by shifting or distorting the energy levels of edge states, while the latter can be avoided. Thus, the edge states in TI are robust, while those in trivial insulator are not.

Some proposed materials for 2D topological insulator are 2D transition metal dichalcogenides (such as the 1T' form of WTe_2) (Cazalilla *et al.*, 2014, Qian *et al.*, 2014), and single-layer $ZrTe_5$ (Weng *et al.*, 2014, Li *et al.*, 2016). Several experimental reports can be found in, e.g., Fei *et al.*, 2017, Wu *et al.*, 2018, and Ugeda *et al.*, 2018.

Exercise

1. Start from the tight-binding Hamiltonian in Eqs. (1.32) and (1.33), switch to the momentum basis, and verify Eq. (1.36).

REFERENCES

Cazalilla, M A, H. Ochoa, and F. Guinea (2014), "Quantum spin hall effect in two-dimensional crystals of transition-metal dichalcogenides," Phys. Rev. Lett. **113**, 077201.

Fei, Zaiyao, Tauno Palomaki, Sanfeng Wu, Wenjin Zhao, Xinghan Cai, Bosong Sun, Paul Nguyen, Joseph Finney, Xiaodong Xu, and David H. Cobden (2017), "Edge conduction in monolayer wte2," Nature Physics **13**, 677–682.

Fu, Liang, and C. L. Kane (2006), "Time reversal polarization and a Z_2 adiabatic spin pump," Phys. Rev. B **74**, 195312.

Fu, Liang, and C. L. Kane (2007), "Topological insulators with inversion symmetry," Phys. Rev. B **76**, 045302.

Fu, Liang, C. L. Kane, and E. J. Mele (2007), "Topological insulators in three dimensions," Phys. Rev. Lett. **98**, 106803.

König, Markus, Steffen Wiedmann, Christoph Brüne, Andreas Roth, Hartmut Buhmann, Laurens W. Molenkamp, Xiao-Liang Qi, and Shou-Cheng Zhang (2007), "Quantum spin hall insulator state in hgte quantum wells," Science **318** (5851), 766–770.

Li, Xiang-Bing, Wen-Kai Huang, Yang-Yang Lv, Kai-Wen Zhang, Chao-Long Yang, Bin-Bin Zhang, Y. B. Chen, Shu-Hua Yao, Jian Zhou, Ming-Hui Lu, Li Sheng, Shao-Chun Li, Jin-Feng Jia, Qi-Kun Xue, Yan-Feng Chen, and Ding-Yu Xing (2016), "Experimental observation of topological edge states at the surface step edge of the topological insulator zrte$_5$," Phys. Rev. Lett. **116**, 176803.

Moore, J E, and L. Balents (2007), "Topological invariants of time-reversal-invariant band structures," Phys. Rev. B **75**, 121306.

Nomura, K (2016), *Topological insulator and superconductor* (Maruzen Publishing Co.) in Japanese.

Qian, Xiaofeng, Junwei Liu, Liang Fu, and Ju Li (2014), "Quantum spin hall effect in two-dimensional transition metal dichalcogenides," Science **346** (6215), 1344–1347.

Ugeda, Miguel M, Artem Pulkin, Shujie Tang, Hyejin Ryu, Quansheng Wu, Yi Zhang, Dillon Wong, Zahra Pedramrazi, Ana Martín-Recio, Yi Chen, Feng Wang, Zhi-Xun Shen, Sung-Kwan Mo, Oleg V. Yazyev, and Michael F. Crommie (2018), "Observation of topologically protected states at crystalline phase boundaries in single-layer wse2," Nature Communications **9**, 3401.

Weng, Hongming, Xi Dai, and Zhong Fang (2014), "Transition-metal pentatelluride $ZrTe_5$ and $HfTe_5$: A paradigm for large-gap quantum spin hall insulators," Phys. Rev. X **4**, 011002.

Wu, Sanfeng, Valla Fatemi, Quinn D. Gibson, Kenji Watanabe, Takashi Taniguchi, Robert J. Cava, and Pablo Jarillo-Herrero (2018), "Observation of the quantum spin hall effect up to 100 kelvin in a monolayer crystal," Science **359** (6371), 76–79.