Quantum anomalous Hall effect and related topological electronic states

Hongming Weng, Rui Yu, Xiao Hu, Xi Dai & Zhong Fang

Published online: 07 Aug 2015.

Submit your article to this journal ⬚

View related articles ⬚

View Crossmark data ⬚

Taylor & Francis
Taylor & Francis Group

REVIEW ARTICLE

Quantum anomalous Hall effect and related topological electronic states

Hongming Weng[a,b], Rui Yu[c], Xiao Hu[c], Xi Dai[a,b] and Zhong Fang[a,b*]

[a]Beijing National Laboratory for Condensed Matter Physics and Institute of Physics, Chinese Academy of Sciences, Beijing 100190, People's Republic of China; [b]Collaborative Innovation Center of Quantum Matter, Beijing, People's Republic of China; [c]International Center for Materials Nanoarchitectonics (WPI-MANA), National Institute for Materials Science, Tsukuba 305-0044, Japan

Over a long period of exploration, the successful observation of quantized version of anomalous Hall effect (AHE) in thin film of magnetically doped topological insulator (TI) completed a quantum Hall trio—quantum Hall effect (QHE), quantum spin Hall effect (QSHE), and quantum anomalous Hall effect (QAHE). On the theoretical front, it was understood that the intrinsic AHE is related to Berry curvature and U(1) gauge field in momentum space. This understanding established connection between the QAHE and the topological properties of electronic structures characterized by the Chern number. With the time-reversal symmetry (TRS) broken by magnetization, a QAHE system carries dissipationless charge current at edges, similar to the QHE where an external magnetic field is necessary. The QAHE and corresponding Chern insulators are also closely related to other topological electronic states, such as TIs and topological semimetals, which have been extensively studied recently and have been known to exist in various compounds. First-principles electronic structure calculations play important roles not only for the understanding of fundamental physics in this field, but also towards the prediction and realization of realistic compounds. In this article, a theoretical review on the Berry phase mechanism and related topological electronic states in terms of various topological invariants will be given with focus on the QAHE and Chern insulators. We will introduce the Wilson loop method and the band inversion mechanism for the selection and design of topological materials, and discuss the predictive power of first-principles calculations. Finally, remaining issues, challenges and possible applications for future investigations in the field will be addressed.

	Contents	**PAGE**

*Corresponding author. Email: zfang@iphy.ac.cn

1. Introduction

1.1. *Hall Effect and anomalous Hall effect*

Edwin H. Hall discovered, in 1879, that when a conductor carrying longitudinal current was placed in a vertical magnetic field, the carriers would be pressed towards the transverse side of the conductor, which led to observed transverse voltage. This is called Hall effect (HE) [1], and it was a remarkable discovery, although it was difficult to understand at that time since the electron was not to be discovered until 18 years later. We now know that the HE is due to the Lorentz force experienced by the moving electrons in the magnetic field, which is balanced by a transverse voltage for a steady current in the longitudinal direction. The Hall resistivity ρ_{xy} under perpendicular external magnetic field $\mathbf{H}$ (along the z-direction) can be written as $\rho_{xy} = R_0 H$, where R_0 is the Hall coefficient, which can be related to the carrier density n as $R_0 = -(1/ne)$ (in the free electron gas approximation). The HE is a fundamental phenomenon in condensed matter physics, and it has been widely used as an experimental tool to identify the type of carrier and to measure the carrier density or the strength of magnetic fields.

In 1880, Edwin H. Hall further found that this "pressing electricity effect" in ferromagnetic (FM) conductors was larger than in non-magnetic (NM) conductors. This enhanced Hall effect was then called as the anomalous Hall effect (AHE) [2], in order to distinguish it from the ordinary HE. Later experiments on Fe, Co and Ni [3] suggested that the AHE was related to the sample magnetization $\mathbf{M}$ (along z), and an empirical relation for the total Hall effect in FM conductors was established as [4, 5]

$$\rho_{xy} = R_0 H + R_s M, \tag{1}$$

where the second term is the anomalous Hall resistivity, and its coefficient R_s is material-dependent (in contrast to R_0, which depends only on carrier density n).

As schematically shown in Figure 1, if we plot the Hall resistivity ρ_{xy} versus external magnetic field $\mathbf{H}$, we will generally expect a straight line (which crosses the origin of the coordinates) for NM conductors (see Figure 1(a)); however, non-linear behavior will be expected for FM conductors, wherein ρ_{xy} increases sharply at low field and crosses over into a linear region under high field (see Figure 1(b)). The initial sharp increase of ρ_{xy} is due to the saturation of magnetization of the sample under external field. After the saturation of magnetization, ρ_{xy} will depend on $\mathbf{H}$ linearly (for high-field region), which is dominated by the ordinary Hall contribution. Therefore, the slope of the linear part under high field gives us R_0. If we extrapolate this linear part to the zero-field limit ($\mathbf{H} = 0$), it will not go through the origin of coordinates, and its intercept on the y-axis gives us $R_s M$ as can be learned from Equation (1). This is a surprising result, which immediately implies two important facts: (1) A kind of Hall effect can be observed even in the
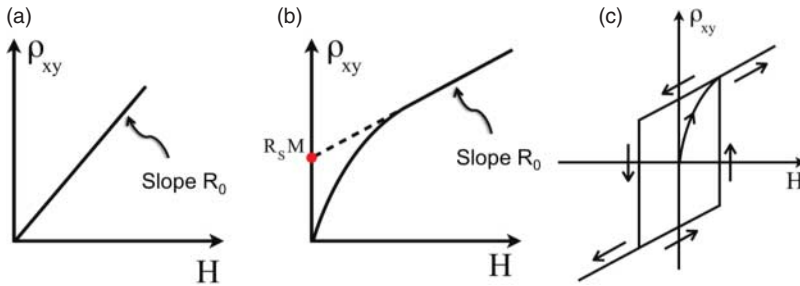
Figure 1. (Color online) Typical behaviors of the Hall and AHEs. The Hall resistivity ρ_{xy} is plotted versus external magnetic field **H**. (a) The Hall effect; (b) the AHE and (c) the hysteresis loop measured from the anomalous Hall effect.

absence of external magnetic field (i.e. no Lorentz force); (2) The anomalous Hall resistivity is sensitive to magnetic moment **M**, and it has been suggested that this property would be useful for the detection of magnetization of conducting carriers, particularly for the cases of weak itinerant magnetism, such as surface/interface magnetization [6] and dilute magnetic semiconductors [7, 8]. Experimentally, measurements are usually done in magnetization loops by scanning the magnetic field from positive to negative, and a hysteresis loop in the ρ_{xy} versus **H** relationship is to be expected. This is very similar to the familiar hysteresis loop observed in the **M** versus **H** curve, as usually found in FM conductors. The empirical relation Equation (1) is very simple and widely used. Unfortunately, its correctness is not well justified. As will be pointed out in a latter part of this article, the anomalous Hall resistivity may have a very complicated form, which is usually non-linear in **M**.

Although the HE and AHE look quite similar phenomenologically, their underlying physics are completely different. The HE is due to the Lorentz force's influence on the moving electrons under magnetic field; however, the AHE exists in the absence of external magnetic field, where there is no orbital effect of moving electrons. In other words, why do the electrons move towards the transverse direction in the absence of Lorentz force? The mechanism of AHE has been an enigmatic problem since its discovery, a problem that lasted almost a century. This problem involves concepts deeply related to topology and geometry that have been formulated only in recent years [9–15] after the Berry phase was recognized in 1984 [16].

Karplus and Luttinger provided a crucial step in unraveling this problem as early as 1954 [17]. They showed that moving electrons under external electric field can acquire an "anomalous velocity", which is perpendicular to the electric field and contributes to the transverse motion of electrons and therefore to the AHE. This "anomalous velocity" comes from the occupied electronic states in FM conductors with spin–orbit coupling (SOC). They suggested that the mechanism leads to an anomalous Hall resistivity proportional to the square of the longitudinal resistivity: $\rho_{xy} \propto \rho_{xx}^2$. Since this contribution depends only on the electronic band structures of perfect periodic crystal and is completely independent of scattering from impurities or defects, it is called *intrinsic* AHE, and it was not widely accepted until the concept of Berry phase was well established.

For a long time, two other *extrinsic* contributions had been considered to be the dominant mechanisms that give rise to the AHE. Smit [18, 19] suggested that there always exist defects or impurities in real materials, which will scatter the moving electrons. In the presence of SOC and ferromagnetism, this scattering is asymmetric and should lead to unbalanced transverse motion of electrons. This is called *skew scattering*, and Smit argued that it is the main source of the AHE. This mechanism predicted $\rho_{xy} \sim \rho_{xx}$, in contrast to the intrinsic contribution. Berger [20], on the other hand, argued that electrons should experience a difference in electric field when

approaching and leaving an impurity, and this leads to another asymmetric scattering process called *side jump*, which also contributes to the AHE. This mechanism marvelously predicts $\rho_{xy} \sim \rho_{xx}^2$, the same as the intrinsic contribution.

The debates about the origin of AHE lasted for a long time, and no conclusion could be drawn unambiguously. From the experimental point of view, defects or impurities in samples are unavoidable and are usually complicated with rich varieties. The contributions from both *intrinsic* and *extrinsic* mechanisms should generally coexist [21]. Experimentalists have tried hard to distinguish them [22–25]. The controversy also arises because of a lack of quantitative calculations that could be compared with experiments.

Still, the discovery of the quantum Hall effect (QHE) in 1980s [26] and the later studies on the geometric phase [27] and the topological properties of quantum Hall states [28–30] promoted the fruitful study of AHE significantly. Around the early years of this century, the *intrinsic* AHE mechanism proposed by Karplus and Luttinger was completely reformulated in the language of Berry phase and topology in electronic structures [9–13]. It was recognized for the first time that the so-called "anomalous velocity" originates from the Berry curvature of occupied eigen wave functions, which can be understood as effective magnetic field in momentum space. This effective magnetic field modifies the equation of motion of electrons and leads to the *intrinsic* AHE [9–15]. From this understanding, we now know that the *intrinsic* AHE can be evaluated quantitatively from the band structure calculations, thanks to the rapid development in the field of first-principles calculations. Detailed calculations for various compounds portray the contributions from the *intrinsic* AHE in a way that convincingly confirms the existing experimental results [11, 12, 23, 25, 31, 32], establishing the dominant role of the *intrinsic* contribution to the AHE. In the meantime, the *extrinsic* AHE was also analyzed carefully [14], and it was understood that its contribution dominates for the clean limit case (rather than the dirty limit), where the deviation of distribution of electronic states from the equilibrium distribution is significant. In this article, we will not discuss the *extrinsic* AHE. For readers who want to learn more details, refer to the review article by Nagaosa et al. [14] Introduction of the Berry phase mechanism for the understanding of the *intrinsic* AHE was a big step forward in the field, and it is also the fundamental base for understanding the quantum anomalous Hall effect (QAHE), which is the main subject of this article.

1.2. *Berry phase mechanism for intrinsic AHE*

In quantum mechanics, the Berry phase is the quantal phase acquired by the adiabatic evolution of wave function associated with the adiabatic change of the Hamiltonian in a parameter space $\mathcal{R}$, with $\mathcal{R} = (\mathcal{R}_1, \mathcal{R}_2, \ldots, \mathcal{R}_m)$ being the set of parameters (a vector). Let $|n(\mathcal{R})\rangle$ be the n-th eigenstate of the Hamiltonian $H(\mathcal{R})$. The overlap of two wave functions infinitesimally separated by $\Delta\mathcal{R}$ in the $\mathcal{R}$-space can be evaluated as

$$\langle n(\mathcal{R}) \mid n(\mathcal{R} + \Delta\mathcal{R})\rangle = 1 + \Delta\mathcal{R}\langle n(\mathcal{R}) \mid \nabla_{\mathcal{R}}|n(\mathcal{R})\rangle = \exp[-i\Delta\mathcal{R} \cdot \mathbf{A}_n(\mathcal{R})], \qquad (2)$$

where $\mathbf{A}_n(\mathcal{R}) = i\langle n(\mathcal{R})|\nabla_{\mathcal{R}}|n(\mathcal{R})\rangle$ is called the Berry connection. Here, $\mathbf{A}_n(\mathcal{R})$ is an important quantity, because it can be viewed as a vector potential and its curl $\mathbf{\Omega}_n(\mathcal{R}) = \nabla_{\mathcal{R}} \times \mathbf{A}_n(\mathcal{R})$, called the Berry curvature, gives an effective magnetic field in the parameter space $\mathcal{R}$ (as will be addressed below). The Berry phase γ_n can then be defined as the integral of the Berry connection along a closed loop $\mathcal{C}$ in the parameter space, or according to Stoke's theorem, equivalently as the integral of Berry curvature on the surface $\mathcal{S}$ enclosed by the adiabatic loop $\mathcal{C} \equiv \partial\mathcal{S}$ (i.e. the surface with the loop $\mathcal{C}$ as boundary):

$$\gamma_n = \oint_C \mathbf{A}_n(\mathcal{R}) \cdot d\mathcal{R} = \int_{\mathcal{S}} \mathbf{\Omega}_n(\mathcal{R}) \cdot d\mathcal{S}, \qquad (3)$$

where the second equality suggests that the Berry phase can also be regarded as the effective magnetic flux passing through a surface $\mathcal{S}$.

Although the concept of the Berry phase has broad applications in physics, its relevance to the band structure in solids has been recognized only in limited situations, such as the QHE under a strong magnetic field [28] and the calculation of electronic polarization in ferroelectrics [33, 34]. Here, we will show that the Berry phase concept is also important for understanding the intrinsic AHE. In this case, we treat the crystal momentum ($\mathbf{k}$) space as the parameter space. We consider a crystalline solid with discrete translational symmetry. The eigen equation of the system is given as $H(\mathbf{r})\psi_{n\mathbf{k}}(\mathbf{r}) = \varepsilon_{n\mathbf{k}}\psi_{n\mathbf{k}}(\mathbf{r})$, where $H(\mathbf{r})$ is the Hamiltonian, $\varepsilon_{n\mathbf{k}}$ and $\psi_{n\mathbf{k}}(\mathbf{r})$ are the eigen energy and the eigen wave function, respectively, and n is the band index. Because of the translational symmetry, the eigen wave function $\psi_{n\mathbf{k}}(\mathbf{r})$ of the system is $\mathbf{k}$-dependent, where $\mathbf{k}$ is the momentum defined in the first Brillouin zone (BZ) of momentum space. According to Bloch's theorem, the eigen wave function should take the form of Bloch state $\psi_{n\mathbf{k}}(\mathbf{r}) = \mathrm{e}^{\mathrm{i}\mathbf{k}\cdot\mathbf{r}}u_{n\mathbf{k}}(\mathbf{r})$, with $u_{n\mathbf{k}}(\mathbf{r})$ being the periodic part of the wave function. Then, the eigen equation of the system can be recast as $H_{\mathbf{k}}(\mathbf{r})u_{n\mathbf{k}}(\mathbf{r}) = \varepsilon_{n\mathbf{k}}u_{n\mathbf{k}}(\mathbf{r})$, where $H_{\mathbf{k}} = \mathrm{e}^{-\mathrm{i}\mathbf{k}\cdot\mathbf{r}}H\mathrm{e}^{\mathrm{i}\mathbf{k}\cdot\mathbf{r}}$ is the $\mathbf{k}$-dependent Hamiltonian. Following the above discussion, we can now define the Berry connection and Berry curvature in the parameter (momentum $\mathbf{k}$) space as

$$\mathbf{A}_n(\mathbf{k}) = i\langle u_{n\mathbf{k}}|\nabla_{\mathbf{k}}|u_{n\mathbf{k}}\rangle, \tag{4}$$

$$\mathbf{\Omega}_n(\mathbf{k}) = \nabla_{\mathbf{k}} \times \mathbf{A}_n(\mathbf{k}) = i\langle\nabla_{\mathbf{k}}u_{n\mathbf{k}}| \times |\nabla_{\mathbf{k}}u_{n\mathbf{k}}\rangle. \tag{5}$$

These two quantities are crucial in understanding the Berry phase mechanism for the intrinsic AHE. Before we can go further, we have to clarify several important properties of the Berry connection and Berry curvature.

- *Berry connection is gauge dependent*: As we have learned from the textbooks of solid-state physics, there exists an arbitrary phase factor for the eigen wave function $|u_{n\mathbf{k}}\rangle$ that is not uniquely determined by the eigen equation of the system. Under a U(1) gauge transformation, the eigen wave function $|u_{n\mathbf{k}}\rangle$ is transformed into $|u'_{n\mathbf{k}}\rangle$ as

$$|u'_{n\mathbf{k}}\rangle = \mathrm{e}^{\mathrm{i}\phi_n(\mathbf{k})}|u_{n\mathbf{k}}\rangle, \tag{6}$$

where $\phi_n(\mathbf{k})$ is a real and smooth scalar function of $\mathbf{k}$. It is easy to see that $|u'_{n\mathbf{k}}\rangle$ is still the eigen wave function of the system for the same eigen state, that is, $H_{\mathbf{k}}|u'_{n\mathbf{k}}\rangle = \varepsilon_{n\mathbf{k}}|u'_{n\mathbf{k}}\rangle$ is satisfied. However, the corresponding Berry connection will be changed by such a gauge transformation, becoming

$$\mathbf{A}'_n(\mathbf{k}) = i\langle u'_{n\mathbf{k}}|\nabla_{\mathbf{k}}|u'_{n\mathbf{k}}\rangle = \mathbf{A}_n(\mathbf{k}) - \nabla_{\mathbf{k}}\phi_n(\mathbf{k}). \tag{7}$$

If we regard the scalar function $\phi_n(\mathbf{k})$ as a kind of scalar potential in the $\mathbf{k}$-space, this form of transformation is the same as that of a vector potential of magnetic field in real space, and therefore the Berry connection $\mathbf{A}_n(\mathbf{k})$ can be viewed as an effective vector potential in momentum space. The gauge dependence of the Berry connection suggests that it is not physically observable. However, it becomes physical after integrating around a closed path (i.e. the Berry phase γ_n defined in Equation (3)). This is because the integration of the second term of Equation (7) around a closed path will only contribute an integer multiple of 2π, and the Berry phase γ_n is therefore invariant modulo 2π.

- *Berry curvature is gauge invariant*: This conclusion can be drawn directly from the factor that $\nabla_{\mathbf{k}} \times \nabla_{\mathbf{k}}\phi_n(\mathbf{k}) = 0$; therefore, $\mathbf{\Omega}_n(\mathbf{k}) = \nabla_{\mathbf{k}} \times \mathbf{A}'_n(\mathbf{k}) = \nabla_{\mathbf{k}} \times \mathbf{A}_n(\mathbf{k})$ is unchanged under the U(1) gauge transformation of Equation (6). This vector form of Berry curvature

suggests that it can be viewed as an effective magnetic field in momentum space. It is a gauge-invariant local manifestation of the geometric properties of the wave function in the parameter (**k**) space, and has been proven to be an important physical ingredient for the understanding of a variety of electronic properties [11, 15, 27, 33, 34].

- *Symmetry consideration*: Here, we consider two important symmetries in solid-state physics, namely TRS and inversion symmetry (IS). Following the above discussion, it is easy to learn that the Berry curvature has the following symmetry properties:

$$\boldsymbol{\Omega}_n(\mathbf{k}) = \boldsymbol{\Omega}_n(-\mathbf{k}) \quad \text{for IS,} \tag{8}$$

$$\boldsymbol{\Omega}_n(\mathbf{k}) = -\boldsymbol{\Omega}_n(-\mathbf{k}) \quad \text{for TRS,} \tag{9}$$

which suggests that $\boldsymbol{\Omega}_n(\mathbf{k}) \equiv 0$ for a system with both TRS and IS. They are strong symmetry constraints, which means that in order to study the possible physical effects related to the Berry curvature, a system with either broken TRS or broken IS is generally required. For example, the intrinsic AHE is observed in an FM system with SOC, where the TRS is broken. For a system with TRS only, in general $\boldsymbol{\Omega}_n(\mathbf{k}) \neq 0$, however, its integration over the whole BZ should be zero as implied by Equation (9). In such a case, the Berry curvature may take effect only through **k**-selective or band n-sensitive probes.

- *The choice of periodic gauge*: For the convenience of studies on solid states with discrete translational symmetry, we usually use the periodic gauge in practice. That is, we require the Bloch wave function to be periodic in momentum space, and use the condition $\psi_{\mathbf{k}}(\mathbf{r}) = \psi_{\mathbf{k}+\mathbf{G}}(\mathbf{r})$ (where **G** is a translational vector for the reciprocal lattice) to partly fix the phases of wave functions. This periodic gauge is used in most of the existing first-principles electronic structure calculation codes. However, we have to note that it is a very weak gauge-fixing condition, and abundant gauge degrees of freedom still remain.

- *Gauge field*: Spreading out each component of the vector form discussed above (Equation (5)), the Berry curvature can be written explicitly as an anti-symmetric second-rank tensor $\mathcal{F}_{\mu\nu}$,

$$\epsilon_{\mu\nu\xi}\Omega_{n,\xi} = \mathcal{F}_{n,\mu\nu} = \frac{\partial}{\partial k_\mu}A_{n,\nu}(\mathbf{k}) - \frac{\partial}{\partial k_\nu}A_{n,\mu}(\mathbf{k})$$

$$= i\left[\left\langle\frac{\partial u_{n\mathbf{k}}}{\partial k_\mu}\bigg|\frac{\partial u_{n\mathbf{k}}}{\partial k_\nu}\right\rangle - \left\langle\frac{\partial u_{n\mathbf{k}}}{\partial k_\nu}\bigg|\frac{\partial u_{n\mathbf{k}}}{\partial k_\mu}\right\rangle\right], \tag{10}$$

where $\epsilon_{\mu\nu\xi}$ is the Levi-Civita antisymmetric tensor, and μ, ν, ξ can be simply treated as the direction indices (i.e. x, y, z) for our purposes. For example, for many of our following discussions on the transport properties in a two-dimensional (2D) system, we will use the equality:

$$\Omega_{n,z} = \mathcal{F}_{n,xy}. \tag{11}$$

Writing it in such a way also helps us understand that the Berry curvature is simply the U(1) gauge field ($\mathcal{F}_{\mu\nu}$) in the language of gauge theory. In the presence of gauge freedom, in order to preserve the gauge invariance of Lagrangian under the gauge transformation, we need to use the gauge covariant derivative instead of the original derivative. For example, for the position operator $\hat{x}_\mu = i\partial_{k_\mu}$, its gauge covariant form is written as $\hat{x}_\mu = i\partial_{k_\mu} - A_\mu(k)$, where $A_\mu(k)$ is the vector potential (i.e. the Berry connection). In such a case, the gauge covariant position operators $\hat{x}_\mu$ and $\hat{x}_\nu$ are no longer commutable, and their commutation relation is given as $[\hat{x}_\mu, \hat{x}_\nu] = -i\mathcal{F}_{\mu\nu}$, which gives the gauge field $\mathcal{F}_{\mu\nu} = \partial_{k_\mu}A_\nu - \partial_{k_\nu}A_\mu$.

Keeping in mind all these properties discussed above for the Berry connection and Berry curvature, we are now ready to understand the intrinsic AHE from the viewpoint of gauge field and geometric Berry phase. For simplicity, let us first consider a 2D system with Hamiltonian $H_0(x, y)$ defined in the xy-plane. Under an external electric field E_y, the whole Hamiltonian of the system becomes $H(x, y) = H_0(x, y) + eE_y y$. By evaluating the velocity operator as $\dot{x} = -(i/\hbar)[x, H] = -(i/\hbar)[x, H_0] - (i/\hbar)[x, eE_y y]$, we find that the presence of electric field in the system leads to an additional velocity term, $v_x = -(i/\hbar)[x, eE_y y] = -(e/\hbar)\mathcal{F}_{xy}E_y = -(e/\hbar)\Omega_z E_y$, which is along the x-direction (being transverse to the external electric field E_y). This result comes from the non-vanishing commutation relation between the gauge covariant position operators. This additional velocity term is nothing but the anomalous velocity term proposed by Karplus and Luttinger [17], and is the origin of the intrinsic AHE.

Now, let us generalize the above discussion to three-dimensional (3D) space and considering both the electric and magnetic external fields, we find that the equation of motion of electrons, in the presence of gauge field, should be recast as [15]

$$\dot{\mathbf{r}} = \frac{1}{\hbar}\frac{\partial \epsilon_n(\mathbf{k})}{\partial \mathbf{k}} - \dot{\mathbf{k}} \times \Omega_n(\mathbf{k}), \tag{12}$$

$$\dot{\mathbf{k}} = -\frac{e}{\hbar}(\mathbf{E} + \dot{\mathbf{r}} \times \mathbf{B}), \tag{13}$$

where $\mathbf{E}$ and $\mathbf{B}$ are the external electric and magnetic fields, respectively. It is apparent that the group velocity (Equation (12)) is different from the textbook one by the inclusion of an additional contribution (i.e. the second term on the right-hand side). This term, $\dot{\mathbf{k}} \times \Omega_n$, is now known as the anomalous velocity term, and it is proportional to the $\mathbf{k}$-space Berry curvature $\Omega_n(\mathbf{k})$. This equation of motion can also be obtained from the semiclassical theory of Bloch electron dynamics. To do this, we have to consider a narrow wave-packet made out of the superposition of the Bloch state of a band. Here, we will not discuss the detailed derivation of this approach; readers refer to the review article by Xiao et al. for more details [15].

For an FM system in the presence of SOC, the TRS is broken; therefore, the non-vanishing Berry curvature contributes to the anomalous velocity term. Consider a system under external electric field $E_y \neq 0$ and without magnetic field (i.e. $B = 0$). The current density j_x can be then obtained from the equation of motion shown above as

$$j_x = -ev_x = -e\sum_n \int_{\text{BZ}} \frac{d^3\mathbf{k}}{(2\pi)^3} f_n(\mathbf{k})(-\dot{\mathbf{k}} \times \Omega_n(\mathbf{k}))_x$$

$$= -\frac{e^2}{\hbar}\sum_n \int_{\text{BZ}} \frac{d^3\mathbf{k}}{(2\pi)^3} f_n(\mathbf{k})(\mathbf{E} \times \Omega_n(\mathbf{k}))_x$$

$$= -\frac{e^2}{\hbar}\sum_n \int_{\text{BZ}} \frac{d^3\mathbf{k}}{(2\pi)^3} f_n(\mathbf{k})E_y\Omega_{n,z}(\mathbf{k}), \tag{14}$$

where $f_n(k)$ is the Fermi distribution function. The dc Hall conductivity is then given as

$$\sigma_{xy} = \frac{j_x}{E_y} = -\frac{e^2}{\hbar}\sum_n \int_{\text{BZ}} \frac{d^3\mathbf{k}}{(2\pi)^3} f_n(\mathbf{k})\Omega_{n,z}(\mathbf{k}) \tag{15}$$

$$= -\int_{-\pi}^{\pi}\left[\frac{e^2}{2\pi h}\sum_n \int_{-\pi}^{\pi}\int_{-\pi}^{\pi}(f_n(\mathbf{k})\Omega_{n,z}(\mathbf{k}))\, dk_x\, dk_y\right]\frac{dk_z}{2\pi}, \tag{16}$$

which is simply the BZ integral of the Berry curvature weighted by the occupation factor $f_n(\mathbf{k})$ of each state. For the convenience of our later discussions, in the last part of the equation, we

have explicitly written the 3D integral as a 2D integral in the $k_x k_y$-plane, supplemented with a line integral along k_z.

On the other hand, we know from textbooks that the DC Hall conductivity can also be derived from the Kubo formula as

$$\sigma_{xy} = \frac{e^2}{\hbar} \sum_{m \neq n} \int_{BZ} \frac{\mathrm{d}^3 \mathbf{k}}{(2\pi)^3} (f_n(\mathbf{k}) - f_m(\mathbf{k})) \times \mathrm{Im} \frac{\langle \psi_{n\mathbf{k}} | v_x | \psi_{m\mathbf{k}} \rangle \langle \psi_{m\mathbf{k}} | v_y | \psi_{n\mathbf{k}} \rangle}{(\varepsilon_{n\mathbf{k}} - \varepsilon_{m\mathbf{k}})^2}, \qquad (17)$$

which is basically equivalent to Equation (15), by expanding each term of the vector form of the Berry curvature explicitly with some algebra. However, we have to emphasize that the form of Equation (15) for Hall conductivity is not only compact, but also fundamentally important for the underlying physics. It demonstrates the Berry phase mechanism of the intrinsic AHE, and it is the key equation for our following discussions on the geometric meaning and topological nature of QAHE. For example, for a 2D system, the BZ integral of Berry curvature for a fully occupied band must give rise to an integer multiple of 2π. Therefore, the Hall conductivity σ_{xy} for a 2D insulator can be finite and be an integer multiple of e^2/h. This is the underlying physics for the quantization of the AHE.

In addition to the understanding of the Berry phase mechanism for the intrinsic AHE discussed above, another important step forward in this field is the quantitative and accurate evaluation of the anomalous Hall conductivity (AHC) for realistic materials, yielding results that can be compared with experimental observations. Due to the rapid increase of computational power and the development of first-principles calculation methods, such quantitative comparisons (though still difficult) become possible and contribute greatly to the development of the field. In a series of papers reporting quantitative first-principles calculations for $SrRuO_3$ [11], Fe [12] and $CuCr_2Se_{4-x}Br_x$ [31], it was demonstrated that the calculated AHC can be reasonably compared with experiments, suggesting the importance of the intrinsic AHE. In more recent years, accurate calculations of AHC have been achieved by using the Wannier function interpolation scheme [35, 36]. The presence of SOC and the breaking of TRS are crucially important for the intrinsic AHE in those systems.

In summary of this part, we have discussed the Berry phase mechanism for the intrinsic AHE. Conceptually, this is an important increment in our understanding because the intrinsic AHE is now directly linked to the geometric and topological properties of the Bloch states in momentum space. This understanding has also led to a great number of research projects on the exotic properties of electronic systems with SOC, such as QAHE and the topological insulators (TIs).

2. Quantum Hall family and related topological electronic states

One of the main subjects of condensed matter physics is the study and classification of different phases and various phase transitions of materials that possess rich varieties of physical properties. Landau developed a general symmetry-breaking theory to understand the phases and phase transitions of materials. He pointed out that different phases really correspond to different symmetries of compounds. An ordered phase, in general, can be described by a local order parameter, and the symmetry of the system changes as a material changes from one phase to another. Landau's symmetry-breaking theory is very successful in explaining a great many kinds of phases (or states) in materials, but it does not work for the topological quantum states. A topological quantum state cannot be described by a local order parameter, and it can change from one state to another without any symmetry breaking. Topology, a word mostly used in mathematics, is now used to describe and classify the electronic structures of materials. "Topological electronic state" means an electronic state that carries certain topological properties (in momentum space usually),

such as the states in TIs [37–54], topological semimetals [55–72] and topological superconductors [49, 50, 73–79]. One of the most important characteristics of topology is the robustness against local deformations, or in the language of physics, the insensitivity to environmental perturbations, which makes topological electronic states promising for future applications. To characterize the order of topological states, new parameters, namely the topological invariants, are needed. In this section, we will review some of the topological electronic states and their quantum physics, including the integer quantum Hall (IQH) state, the quantum anomalous Hall (QAH) state, the quantum spin Hall (QSH) state, and the topological semimetal (TSM) state. We will use the concept of Berry phase in momentum space to discuss the topological nature of topological states and the related topological invariants.

2.1. *Chern number, Chern insulator, and quantum AHE*

In 1980, Klaus von Klitzing discovered [26] that the Hall conductance, viewed as a function of strength of the magnetic field applied normal to the 2D electron gas plane, at very low temperature, was quantized and exhibited a staircase sequence of wide plateaus. The values of Hall conductance were integer multiples of a fundamental constant of nature: $e^2/h = 1/(25812.807572\Omega)$, with totally unanticipated precision, and independent of the geometry and microscopic details of the experiment. This is an important and fundamental effect in condensed matter physics, and it is called the integer quantum Hall effect (IQHE)—the quantum version of the Hall effect.

The IQHE is now well understood in terms of single particle orbitals of an electron in a magnetic field (i.e. the Landau levels), and the phenomenon of "exact quantization" has been shown to be a manifestation of gauge invariance. The robustness and the remarkable precision of Hall quantization can also be understood from the topological nature of the electronic states of 2D electron gas under magnetic field. The integer number, originally known as the TKNN number [28] in the Hall conductance derived from the Kubo formula, is now characterized as a topological invariant called "Chern number". This topological understanding of the IQHE is a remarkable leap of progress, opening up the field of topological electronic states in condensed matter physics. IQHE is therefore regarded as the first example of topologically non-trivial electronic states to be identified and understood.

This conceptual breakthrough with regard to the topological nature of IQHE, though important, was not widely generalized, and the relationship of IQHE to the rich variety of condensed materials was not revealed for a long time. This is because the IQHE is observed only in a very particular system, that is, 2D electron gas under a strong external magnetic field, and the formation of Landau levels (usually at very low temperature) is required. Under such extreme conditions, the material's details and its electronic band structure become irrelevant to the physics. In this sense, the lattice model proposed by Haldane in 1988 [37] is very stimulating for the study of topological electronic states. He proposed that a spinless fermion model on a periodic 2D honeycomb lattice without net magnetic flux can in principle support a similar IQHE. Although Haldane's model is very abstract and unrealistic (at least at that time), his result suggested that a kind of quantized Hall effect can exist even in the absence of magnetic field and the corresponding Landau levels. In other words, certain materials, other than the 2D electron gas under magnetic field, can have topologically non-trivial electronic band structures of their own, which can be characterized by a non-zero Chern number. Such materials were called Chern insulators later. In this way, the concept of topological electronic state was generalized and was connected to the electronic band structure of materials.

The progress in the study of IQHE was very inspiring in the 1980s; on the other hand, not much of it was related to the field of AHE. Around that time, the development of the AHE field

was rather independent, with almost no intersection with the study of IQHE. But the understanding of Berry phase mechanism for the intrinsic AHE [9–13], reformulated around the early years of twenty-first century, changed the situation significantly. This understanding established the connection between the AHE and the topological electronic states. By this connection, it is now understood [10, 11, 80] that the intrinsic AHE can have a quantum version—the QAHE, which is in principle similar to IQHE but without external magnetic field and the corresponding Landau levels. It turns out that this is exactly the same effect discussed by Haldane for a Chern insulator with a non-zero Chern number. Currently, the fields of IQHE and AHE are closely combined, and the realization of a Chern insulator becomes possible by finding the QAHE in suitable magnetic compounds with strong SOC.

The IQHE and the QAHE have their own characteristics; however, their underlying physics, in terms of the topological properties of their electronic structures, are basically the same. They are all related to the Berry connection, Berry curvature and Berry phase in momentum space, as discussed in the previous section. In this part, we will not discuss the particular details associated with each effect, but rather we will concentrate on their common features, namely the topological property and topological invariant (Chern number). To begin with, we will discuss several important concepts:

- *Hall conductivity as integral of Berry curvature:* Considering a 2D insulating system with broken TRS, the Hall conductivity of the system at low enough temperature can be written as

$$
\begin{aligned}
\sigma_{xy} &= -\frac{e^2}{\hbar} \sum_n \int_{\text{BZ}} \frac{\text{d}^2\mathbf{k}}{(2\pi)^2} f_n(\mathbf{k}) \Omega_{n,z}(\mathbf{k}) \\
&= -\frac{e^2}{2\pi h} \int_{\text{BZ}} \text{d}^2\mathbf{k} \sum_{n(occ)} \Omega_{n,z}(\mathbf{k}) \\
&= -\frac{e^2}{2\pi h} \int_{\text{BZ}} \text{d}^2\mathbf{k} \Omega_z(\mathbf{k}),
\end{aligned}
\tag{18}
$$

where the second equality follows because of the existence of an energy gap and $f_n(\mathbf{k}) = 1$ ($= 0$) for the occupied (unoccupied) state (see Equations (14)–(16) for a 3D version). This form of Hall conductivity can be derived either from the Berry phase formalism or from the Kubo formula, as discussed in the previous section. It is also important to note that this formula can be equally applied to the AHC of a 2D FM insulator without external field and to the Hall conductivity of a 2D electron gas under magnetic field. For the former case, it is straightforward; for the latter case, however, special considerations have to be taken. First, we assume that the external magnetic field is strong enough for the system to form the Landau levels. Second, the system must be insulating with the chemical potential located within the gap (i.e. between two neighboring Landau levels). Third, since the presence of external magnetic field will break the translational symmetry, we have to use the concept of magnetic translational symmetry to recover the periodicity of the system. As a result, magnetic Bloch states must be used for the evaluation of Berry curvature, and the magnetic BZ is used for the integration, correspondingly.
- *2D BZ as a torus:* By adopting the periodic gauge, the wave function $\psi_{n\mathbf{k}}(\mathbf{r})$ is periodic in momentum space. Under such a periodic boundary condition, the 2D BZ becomes a torus as schematically shown in Figure 2. Therefore, the integral of the Berry curvature Ω_z over
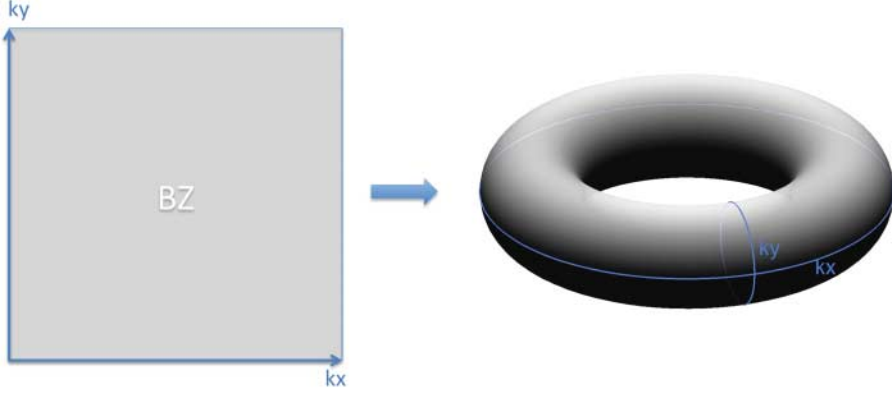
Figure 2. Under the periodic boundary condition for the Bloch states, the 2D BZ is equivalent to a torus.

the 2D BZ can be recast as a vector integral over the surface of the torus as

$$\int_{BZ} d^2\mathbf{k}\,\Omega_z(\mathbf{k}) = \int_{\mathcal{S}} \mathbf{\Omega}(\mathbf{k}) \cdot d\mathcal{S}. \tag{19}$$

Here, we have used an effective Berry curvature $\mathbf{\Omega}(\mathbf{k})$ defined over the torus surface, which satisfies the condition $\mathbf{\Omega}(\mathbf{k}) \cdot \mathbf{n} = \Omega_z(\mathbf{k})$, where $\mathbf{n}$ is the unit vector along the normal of the surface $\mathcal{S}$.

- *Quantization and Chern number:* The torus is a closed manifold without boundary, so according to Chern's theorem [81] in differential geometry, the vector integral of Berry curvature over a torus surface must be an integer multiple of 2π

$$\int_{\mathcal{S}} \mathbf{\Omega}(\mathbf{k}) \cdot d\mathcal{S} = 2\pi Z, \tag{20}$$

where Z is an integer number called Chern number. Having this property, it is now easy to see that the Hall conductivity for a 2D insulator must be quantized as

$$\sigma_{xy} = -\frac{e^2}{2\pi h} \times 2\pi Z = -\frac{e^2}{h}Z. \tag{21}$$

- *Absence of smooth gauge:* Based on Stokes' theorem, the surface integral of Berry curvature (Equation (20)) can also be evaluated as an loop integral of the Berry connection $\mathbf{A}(\mathbf{k})$ along the boundary of the BZ, that is, $\int_{\mathcal{S}} \mathbf{\Omega}(\mathbf{k}) \cdot d\mathcal{S} = \oint_C \mathbf{A}(\mathbf{k}) \cdot d\mathbf{k}$. Since the BZ is a torus, which has no boundary, the integral must be vanishing if $\mathbf{A}(\mathbf{k})$ is smoothly defined in the whole BZ. Therefore, a non-zero Chern number indicates that we cannot choose a smooth gauge transformation such that $\mathbf{A}(\mathbf{k})$ is continuous and single valued over the whole BZ. Thus, non-zero Chern number can be viewed as obstruction to continuously define the phase of the occupied wave functions on a 2D BZ, which is a torus [82–84].

For a system with multiple bands, the Berry curvature should be understood as the summation of contributions coming from all occupied bands. Having the properties discussed above, we can now define a Chern insulator as a 2D insulator whose electronic structure gives a non-zero Chern number. The above discussions about the Chern number and the topology of electronic states in momentum space may be still too abstract, in the following we will give a more explicit explanation, in which the non-zero Chern number manifests itself as the winding number of the

Wannier center evolution for the effective 1D systems with constant k_y. Here, we consider the simplest case: a 2D lattice with only one single occupied band (the band index n can therefore be neglected). Although it is impossible to choose a single smooth gauge over the whole BZ for a Chern insulator according to our above discussions, it is possible to choose a special gauge that is smooth and periodic along one direction (say k_x), but not necessarily along the other (say k_y) [82–85]. Thus, we can do the integration in Equation (19) explicitly for the k_x direction with fixed k_y and obtain (we set the lattice parameter $a = 1$)

$$
\begin{aligned}
2\pi Z &= -\int_{-\pi}^{\pi}\int_{-\pi}^{\pi} dk_x\,dk_y(\partial_{k_x}A_y - \partial_{k_y}A_x) \\
&= \int_{-\pi}^{\pi} dk_y\partial_{k_y}\left(\int_{-\pi}^{\pi} dk_x A_x(k_x, k_y)\right) \\
&= \int_{-\pi}^{\pi} d\theta(k_y).
\end{aligned}
\tag{22}
$$

Here, $\theta(k_y) = \int_{-\pi}^{\pi} dk_x A_x$ is an angle (i.e. Berry phase) calculated from the 1D integration of $A_x(k_x, k_y)$ along the k_x-axis (a closed loop) for each fixed k_y. We can then plot $\theta(k_y)$ as a function of k_y. As shown in Figure 3, over a cylinder surface (in the cylinder coordinates), k_y is along the longitudinal direction, and the azimuth is the angle $\theta(k_y)$ for each fixed k_y. Moving from $k_y = -\pi$ to $k_y = \pi$, we can see the difference between the trivial insulator ($Z = 0$) and the Chern insulator ($Z \neq 0$). The winding number of $\theta(k_y)$ over the cylinder surface is zero for the former (Figure 3(a)) and non-zero for the latter (Figure 3(b) and (c)). In this way, we relate the Chern number to some kind of winding number defined by the eigen functions, and see that the Chern insulator has a "twisted" energy band.

It is also interesting to note that the k_y-dependent Berry phase calculated as

$$
P_n(k_y) = \frac{\theta_n(k_y)}{2\pi} = \frac{1}{2\pi}\int_{-\pi}^{\pi} dk_x A_{n,x}(k_x, k_y)
\tag{23}
$$

can be related to the Wannier center of a 1D system (where we have recovered the band index n dependence). To see this, we consider a generic 2D system and denote the Wannier function in cell $\mathbf{R}$ associated with band n, given in terms of the Bloch states, as

$$
|n\mathbf{R}\rangle = \frac{1}{(2\pi)^2}\int_{BZ} d^2\mathbf{k}\,e^{-i\mathbf{k}\cdot\mathbf{R}}|\psi_{n\mathbf{k}}\rangle,
\tag{24}
$$

with

$$
|\psi_{n\mathbf{k}}\rangle = \sum_{\mathbf{R}} e^{i\mathbf{k}\cdot\mathbf{R}}|n\mathbf{R}\rangle,
\tag{25}
$$

where $\mathbf{k} = (k_x, k_y)$ is considered, and the periodic part of the Bloch function is defined as

$$
u_{n\mathbf{k}}(\mathbf{r}) = e^{-i\mathbf{k}\cdot\mathbf{r}}\psi_{n\mathbf{k}}(\mathbf{r}).
\tag{26}
$$

Then, the matrix elements of the position operator $\mathbf{r}$ between Wannier functions take the form

$$
\langle n\mathbf{R}|\hat{\mathbf{r}}|m\mathbf{0}\rangle = i\frac{1}{(2\pi)^2}\int_{BZ} d^2\mathbf{k}\,e^{i\mathbf{k}\cdot\mathbf{R}}\langle u_{n\mathbf{k}}|\nabla_{\mathbf{k}}|u_{m\mathbf{k}}\rangle,
\tag{27}
$$

from which we get the center of the Wannier function in the **0**-th cell as

$$
\bar{\mathbf{r}} = \langle n\mathbf{0}|\hat{\mathbf{r}}|n\mathbf{0}\rangle = i\frac{1}{(2\pi)^2}\int_{BZ} d^2\mathbf{k}\langle u_{n\mathbf{k}}|\nabla_{\mathbf{k}}|u_{n\mathbf{k}}\rangle.
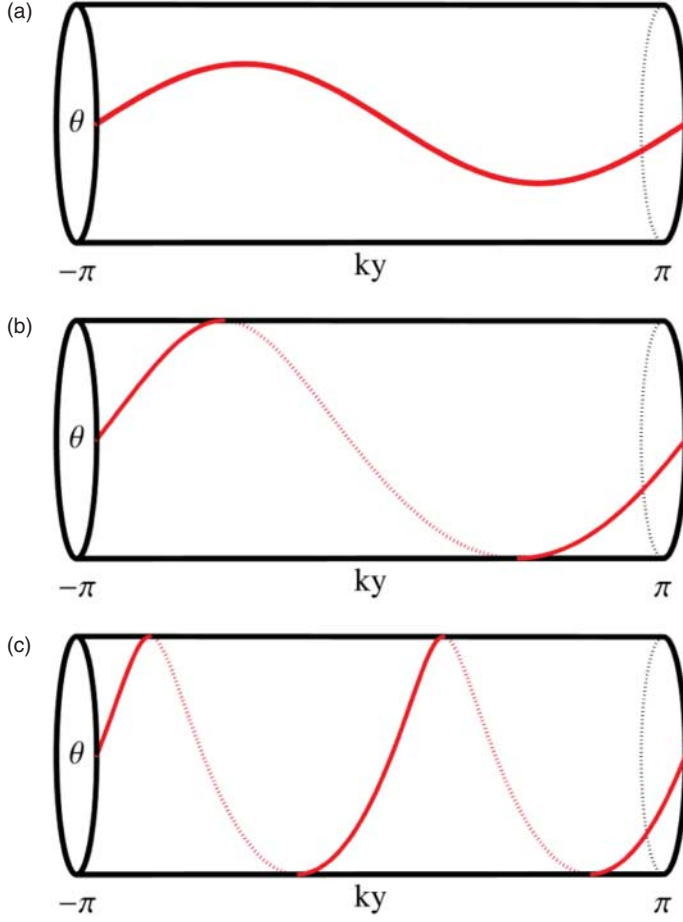\tag{28}
$$

Figure 3. The evolution of $\theta(k_y)$ (or 1D Wannier center) on cylinder surface for system with (a) $Z = 0$, (b) $Z = 1$ and (c) $Z = 2$.

Now, let us consider the k_x and k_y directions separately, and define the 1D Wannier function for each fixed-k_y as

$$|n, R_x, k_y\rangle = \frac{1}{(2\pi)} \int_{-\pi}^{\pi} dk_x \, e^{-ik_x R_x} |\psi_{n\mathbf{k}}\rangle. \tag{29}$$

Then, the 1D Wannier center (along the x-direction) as a function of k_y can be defined as the average value of position operator $\hat{x}$ as

$$\bar{x}(k_y) = \langle n, R_x = 0, k_y | \hat{x} | n, R_x = 0, k_y \rangle = \frac{i}{(2\pi)} \int_{-\pi}^{\pi} dk_x \langle u_{n\mathbf{k}} | \nabla_{k_x} | u_{n\mathbf{k}} \rangle = P_n(k_y), \tag{30}$$

from which we see that $P_n(k_y)$ is nothing but the 1D Wannier center of band n. The Chern number can therefore be understood as the winding number of 1D Wannier center when it evolves as a function of k_y (see Figure 3). We can also see that when the Chern number is not zero, we cannot choose a smooth gauge transformation such that $\mathbf{A}(\mathbf{k})$ is continuous and single valued over the whole BZ. Although the above discussion is only for the simplest systems with single occupied band, generalization to the multi-band situation is quite straightforward. For a general band insulator with multiple occupied bands, the Berry connection $\mathbf{A}_{nm}(\mathbf{k}) = i\langle u_{n\mathbf{k}} | \nabla_{\mathbf{k}} | u_{m\mathbf{k}} \rangle$ contains band
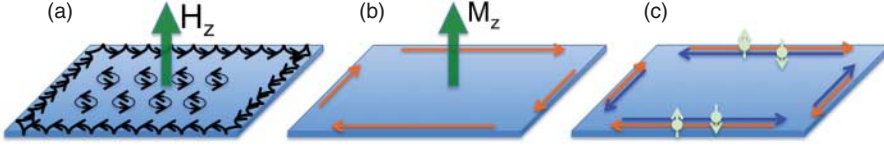
Figure 4. Schematic illustrations of the edge states in (a) QHE, (b) QAHE and (c) QSHE.

index n, m and becomes nonabelian. Then, the Berry curvature that defines the Chern number is obtained from the trace of the Berry connection $\mathbf{A}(\mathbf{k})$.

The Hall conductivity of a Chern insulator with a non-zero Chern number must be quantized as an integer multiple of e^2/h. Different Chern numbers give different states whose Hall conductivities are different, but their local symmetry can be the same. Therefore, to distinguish the states, we cannot use local order parameters (in the language of Landau's symmetry-breaking theory); instead, we need a topological invariant, the Chern number, as a global order parameter of the system. By rewriting Hall conductivity in terms of the Berry curvature and Berry phase, we can now unify QHE and QAHE. Readers will recall that the QAHE is nothing but the quantum version of the AHE realized in a Chern insulator without the presence of external magnetic field.

Similar to QHE, where the TKNN number is related to the number of edge states in a real 2D sample with boundary [28], the Chern number can also be physically related to the number of edge states for a 2D Chern insulator [37]. The existence of edge states is a direct result of the topological property of the bulk electronic structure, and is due to the phenomenon discussed in the literature as the bulk-boundary correspondence [86, 87]. In this case, due to the broken TRS, the edge state must be chiral (i.e. the electrons of the edge state can move only in one direction surrounding the sample boundary, either left- or right-handed, as shown in Figure 4(b)). As discussed with regard to IQHE, the charge transport of the edge state is in principle dissipationless, and back scattering is absent due to the lack of an edge state with opposite velocity [88, 89].

Although the topological properties of Chern insulators and the related QAHE are fundamentally the same as that discussed with regard to IQHE, they are conceptually broadly generalized to a wide field and to a rich variety of materials. This generalization is an important step forward, providing the building blocks for subsequent discussions of many possible topological electronic states. From the application point of view, the Chern insulator (or the related QAHE) is also important because the quantized Hall conductivity can be realized in the absence of magnetic field, greatly simplifying measurement conditions.

2.2. *Z_2 invariant, TI, and quantum spin Hall effect (QSHE)*

Although the Chern insulator (or the related QAHE) is the simplest topological electronic state, its realization occurred rather later and was much stimulated by the rapid development in the field of TIs—another interesting topological electronic state protected by TRS. Considering symmetry, it is easy to prove that the Chern number in Equation (20) must be vanishing for an insulator with TRS. However, this does not mean that the electronic state carries no topological property in this case. Kane and Mele [38] introduced a new topological invariant, the Z_2 number, to classify an insulating system with TRS. They proposed that a time-reversal invariant insulator can be further classified as a trivial insulator with $Z_2 = 0$ or a non-trivial TI with $Z_2 = 1$, which is a typical example of a symmetry protected topological state [90]. A Z_2 TI in 2D is also called a quantum spin Hall insulator (QSHI) because it can support the QSHE [39–42], which shares certain features with the IQHE and QAHE and can be understood from the viewpoint of "band-twisting" or winding number of Berry phase in momentum space. A QSHI is different from a

trivial insulator in the sense that it has gapped insulating states in the bulk but gapless states on the edge due to its topological property. It is also distinguished from the Chern insulator in the sense that QSHI has an even number of edge states, composed of pairs of counter-propagating chiral edge states, as shown in Figure 4(c)). The QSHE has been explicitly discussed for graphene lattice [39] and HgTe quantum-well structures [40–42]. In this part, however, we will take a simple task and discuss the Z_2 invariant and the topological state with TRS from the viewpoint used above for the Chern insulator.

For a 2D insulator with TRS, the total number of occupied electronic states must be even. Suppose that we can divide them into two subspaces, I and II, which are related by TRS and smoothly defined on the whole BZ. Evaluating the Chern number for each subspace independently, if one has Chern number Z, the other one must have Chern number $-Z$ (due to the TRS) and the total Chern number of the whole system is always zero. It seems that we can use Z as the topological index to classify the band insulators with TRS. While unfortunately, the smooth partition of the occupied states into two subspaces with one being the time reversal of another is only possible with extra good quantum numbers, such as S_z, where the Chern number obtained within the spin-up subspace can be used to describe the topology of the system and is called "spin Chern number". But generically such a smooth partition of the occupied states can only be made for half of the BZ (not the whole BZ) and we need at least two patches (A and B) to fully cover the whole BZ. The winding number of the U(2N) gauge transformation matrix $\hat{t}_{AB}(k)$ at half of the boundary between two patches can be used to define a new classification of the band insulators with TRS. As further proved by Fu and Kane, only the even and odd feature of the above winding number is unchanged under the U(2N) gauge transformation and all the band insulators with TRS can be classified into topological trivial and non-trivial classes, which is called Z_2 invariance accordingly. Fu and Kane further derived the expression for Z_2 invariance in terms of both Berry connection and Berry curvature as [91]

$$Z_2 = \frac{1}{2\pi} \left(\oint_{\partial(BZ/2)} \mathbf{A}(\mathbf{k}) \cdot d\mathbf{k} - \int_{BZ/2} d^2\mathbf{k}\Omega_z(\mathbf{k}) \right) \text{ mod } 2, \quad (31)$$

where the integral of Berry curvature is performed over the half BZ (i.e. BZ/2), and the integral of Berry connection is performed along the boundary of the half BZ (i.e. $\partial(BZ/2)$). To evaluate Z_2 invariance using the above equation, one needs to find a smooth gauge for the wave functions on half of the BZ, which is very difficult for the band structure calculations for realistic materials, and as a result, this formula is rarely used in practice. On the other hand, the Z_2 invariance is a gauge-invariant quantity, one should be able to compute it without any gauge-fixing condition. For that purpose, some of the authors of the present paper developed an alternative expression for the Z_2 invariance called the Wilson loop method, which will be introduced in detail in Section 3.2. Here, we just sketch its main idea briefly. Similar to the Wilson loop method for Chern number calculation, the purpose of the Wilson loop method is to calculate the "Wannier center" of each band for the effective 1D insulators with fixed k_y and determine the Z_2 invariance by looking at their evolution with k_y. With the presence of TRS, a generic system contains $2N$ occupied bands. As mentioned previously, now the Berry connection becomes $2N*2N$ matrix. The loop integral of such U(2N) Berry connection gives a $2N*2N$ unitary matrix $\hat{D}(k_y)$. The U(1) part of this matrix contributes to the Chern number as discussed before and the Z_2 invariance can be obtained from the remaining SU(2) part by taking the phases of its eigenvalues $\theta_n(k_y)$, where n denotes the band index. As proved in Ref. [92], the TRS only guarantees the double degeneracy of $\theta_n(k_y)$ at two time-reversal invariant loops $k_y = 0$ and $k_y = \pi$. The Z_2 invariance of the system is determined by whether the $\theta_n(k_y)$s switch parters or not as they evolves from $k_y = 0$ to π.
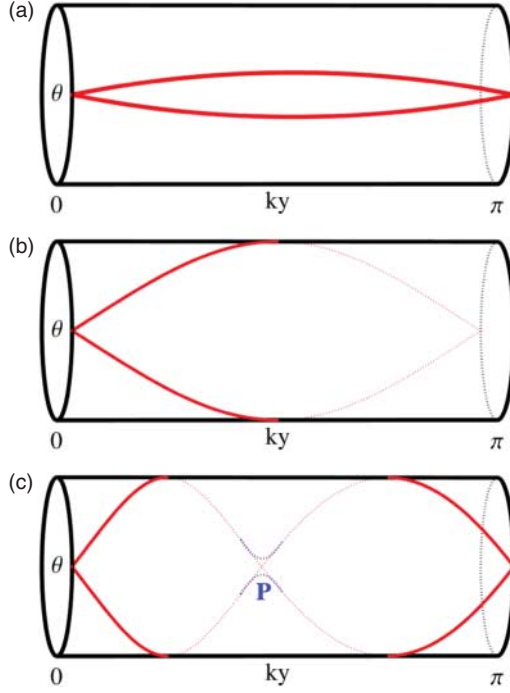
Figure 5. Evolution of Wannier centers ($P_n(k_y)$) for systems with (a) $Z_2 = 0$, (b) $Z_2 = 1$, (c) $|Z| = 2$ but $Z_2 = 0$. The Wannier centers wind the cylinder surface even (a, c) and odd (b) times respectively. Figure adapted from Ref. [92].

To be explicit, let us again consider the simplest example of 2D insulator, with only two occupied bands, $\psi_{I,\mathbf{k}}$ and $\psi_{II,\mathbf{k}}$, which are related by the TRS (say, for example, $\psi_{I,-\mathbf{k}} = \hat{T}\psi_{II,\mathbf{k}}$, where $\hat{T}$ is the time reversal operator). The two states form a Kramers pair and must be degenerate at the time-reversal invariant momentum (TRIM) of the BZ (i.e. with k_x or $k_y = 0, \pi$). We can evaluate $\theta(k_y)$ for all occupied bands (as we have done above for the Chern insulator). As mentioned previously, the two Wannier centers must be degenerate at $k_y = 0$ and π, which leads to three situations in general (as shown in Figure 5(a)–(c)). First, if the evolution path of the $\theta(k_y)$ does not enclose the half BZ, this is the trivial situation with Z_2 invariance $\nu = 0$ (Figure 5(a)). Second, if the evolution path of the $\theta(k_y)$ encloses the half BZ once, this is the non-trivial TI with Z_2 index $\nu = 1$, where the crossing of two $\theta(k_y)$ curves at $k_y = 0$ and π is protected by the TRS (Figure 5(b)). Third, if the evolution path of the $\theta(k_y)$ encloses the half BZ twice, the two $\theta(k_y)$ curves must cross at some k_y other than the TRIM. Such crossings are not protected by the TRS and are removed by small perturbations, which drive the crossings into anti-crossings. As a result, the system becomes equivalent to the trivial case with Z_2 index $\nu = 0$ (Figure 5(c)).

The edge states of 2D Z_2 TIs must appear in pairs due to the existence of TRS, and each pair of edge states is composed of two counter-propagating chiral edge states, which are related to each other by the TRS. In this case, if one state at the edge has velocity v, we can always find another state with opposite velocity $-v$. However, back scattering between the two states is again forbidden because the two states must have opposite spins. The existence of such pairs of edge states leads to the observable QSHE [38, 39, 41], where the spin Hall conductivity is quantized in units of $e/2\pi$ (if S_z is conserved). According to our above discussions of the Z_2 topology, if there

exist multiple pairs of edge states, two pairs can in principle couple together and open up a gap (because scattering between the pairs is not forbidden). As a result, the Z_2 number is physically related to the number of pairs of edge states mod 2, in contrast to the case of Chern insulators.

Using the 2D topological invariants (such as Z or Z_2) as building blocks, a 3D crystal can in principle be characterized by a triplet of 2D topological numbers for the three crystal orientations. In such a way, we can extend the topological classes from 2D to 3D. There are in general two situations by doing this: (1) one is the trivial extension leading the "weak" 3D topological class, which can be viewed as a simple stacking of 2D topological insulating layers along a certain crystal orientation; (2) another is a non-trivial extension leading to a "strong" topological class, which is a new and stable topological state in 3D. For example, by a simple stacking of 2D Chern insulating layers along the z-direction, if the band dispersion along k_z is weak enough, each 2D layer with fixed k_z will keep its original topological property (characterized by a non-zero Chern number). As a result, a 3D "weak" Chern insulator can be obtained. In a similar way, a 3D "weak" TI can be obtained too. However, it is important to note that, if we concentrate on insulators only, there will be no "strong" Chern insulator in 3D; on the other hand, a new 3D state with TRS, called "strong" TI, can be obtained [93–95] by extending the Z_2 number from 2D to 3D. (As will be discussed in the next section, the missing "strong" Chern class in 3D has to be a metal—the topological semimetal.) For a 3D insulator with TRS, there will be 8 TRIM points in the BZ (different from 2D BZ which has only 4 TRIM points). For any 4 TRIM points that lie in the same plane, we can define a 2D Z_2 number as discussed above. In total, there will be 16 invariant configurations and can be distinguished by four independent Z_2 indices, $\nu_0(\nu_x\nu_y\nu_z)$, as discussed in Ref. [93]. Here, ν_0 is the total Z_2 number, and the $\nu_{x,y,z}$ are the 2D Z_2 numbers along the x-, y-, z-directions, respectively. The three weak Z_2 indices $\nu_{x,y,z}$ can be evaluated for the $k_x = \pi$, $k_y = \pi$, and $k_z = \pi$ planes, as a convention, respectively. The total Z_2 index ν_0, however, is a strong index, and it has to be evaluated globally by considering the 3D structure. That is, we have to consider the change of 2D Z_2 number between two parallel planes (perpendicular to certain reciprocal lattice vector). For example, we can choose the $k_x = 0$ and π planes, or $k_y = 0$ and π planes, or $k_z = 0$ and π planes. If the Z_2 numbers of two planes are different, we have $\nu_0 = 1$, otherwise, if they are the same, we have $\nu_0 = 0$. A realistic example will be given in Section 3. For a 3D strong topological insulator with strong topological index $\nu_0 \neq 0$, on its boundary (i.e. surface), we will expect the Dirac-cone-type surface states, which are protected by the TRS [49,50].

2.3. *Magnetic monopole, Weyl node and topological semimetal*

As mentioned above, if we attempt to consider the possible "strong" topological state characterized by Chern numbers in 3D, we will find that it has to be a metal, and this leads to a very interesting new topological state of quantum matters—topological metals or semimetals. We will see that this is a much more general state, and the 2D Chern insulating state can be regarded as a special cut of the 3D topological semimetal state along a certain plane in momentum space. From another point of view, we can also raise a question like the following. Since we know from the above discussions that insulators can be further classified as topologically trivial and non-trivial insulators, can we do the same classification for metals? If possible, what will be the topological invariant for the proper description of the topological metallic state? This is a very interesting question, and we will show in this part that it is related to the magnetic monopoles [11, 96] and Weyl nodes in momentum space.

Berry phase is called quantum geometric phase since it has a very intuitive geometric picture. It is proportional to the solid angle subtended at the magnetic monopole by the adiabatic loop of Hamiltonian in momentum space. The magnetic monopole is the source or drain of the gauge

field $\boldsymbol{\Omega}_n(\mathbf{k})$. Since

$$\boldsymbol{\Omega}_n(\mathbf{k}) = -\mathrm{Im} \sum_{m \neq n} \frac{\langle u_m | \nabla_{\mathbf{k}} H(\mathbf{k}) | u_n \rangle \times \langle u_n | \nabla_{\mathbf{k}} H(\mathbf{k}) | u_m \rangle}{(E_m(\mathbf{k}) - E_n(\mathbf{k}))^2}, \tag{32}$$

it diverges at the point $\mathbf{k} = \mathbf{k}_0$ where $E_m(\mathbf{k}_0) = E_n(\mathbf{k}_0)$ (see Equations (16) and (17)). This means that the magnetic monopole is formed by energy level crossing, and a two-energy-level system is the simplest case. Consider the generic form of the Hamiltonian for any two-level system:

$$H = d_0(\mathbf{k}) I_{2\times2} + \mathbf{d}(\mathbf{k}) \cdot \sigma, \tag{33}$$

where $I_{2\times2}$ is the identity matrix, $\mathbf{d}$ is a 3D vector depending on momentum $\mathbf{k}$ and σ are Pauli matrices. There are two eigen states $\psi_{\pm}$ with eigen energies $E_{\pm} = d_0(\mathbf{k}) \pm \sqrt{\mathbf{d} \cdot \mathbf{d}}$. The term $d_0(\mathbf{k})$ is just a shift of zero energy level, and can be neglected. At the energy degeneracy or level crossing point ($\mathbf{k} = k_0$), $E_+ = E_-$ (i.e. $\mathbf{d}(\mathbf{k}_0) = 0$) is required. Obviously, $\mathbf{k}_0$ is not necessarily on the path of a adiabatic loop change of the Hamiltonian, which means that the band gap between E_+ and E_- can be well preserved if the Fermi level is away from the band crossing point. Around the neighborhood of the $\mathbf{k}_0$ point, $\mathbf{d}(\mathbf{k})$ can be expanded as $\mathbf{d}(\mathbf{k}) \approx \mathbf{d}(\mathbf{k}_0) + (\mathbf{k} - \mathbf{k}_0) \cdot \nabla \mathbf{d}(\mathbf{k})$. Taking the zero point of the parameter space as $\mathbf{k}_0 = 0$, we have $\mathbf{d}(\mathbf{k}) = (\mathbf{k} \cdot \nabla)\mathbf{d}(\mathbf{k})$. Now for the simplest case, where $\mathbf{d}(\mathbf{k}) = \pm \mathbf{k}$, we can have $H(\mathbf{k}) = \pm \mathbf{k} \cdot \sigma$, and the corresponding Berry curvature can be obtained as

$$\Omega_{\pm}(\mathbf{k}) = \mp \frac{\mathbf{k}}{2|k|^3}. \tag{34}$$

Obviously, such a magnetic field distribution in momentum space is similar to the electric field distribution of a point charge in real space, and can be understood as magnetic field around a "magnetic charge"–magnetic monopole [11]. In other words, the divergence of magnetic field $\boldsymbol{\Omega}(\mathbf{k})$ is no longer zero (i.e. $\nabla_{\mathbf{k}} \cdot \boldsymbol{\Omega}(\mathbf{k}) \neq 0$) but is related to the magnetic charge $\pm \frac{1}{2}$ at the source or drain. The degeneracy or level crossing point $\mathbf{k}_0$ is the place where the magnetic monopole with strength $\frac{1}{2}$ is located.

Zak [97] pointed out that such Berry phase and Berry curvature can also exist in periodic systems where the eigenstates are Bloch wave functions and the parameter space is the crystal momenta $\mathbf{k}$ which can vary in closed loops, such as BZ or Fermi surface. Then, the adiabatic evolution loop forms a compact manifold that has no boundary. For example, the BZ of a 2D system is a torus, and the Fermi surface of a 3D metal (in the simplest case) is an enclosed sphere. Obviously, Gauss's law ensures that the total flux penetrating the closed surface must be quantized and is equal to the magnetic charge of monopoles enclosed by the surface (either torus or Fermi sphere).

Now, we can see that the magnetic monopole in momentum space (or parameter space in general) plays a crucial role in determining the topology of electronic band structures. Besides the insulating case, in the following we concentrate on a 3D metal with a well-defined Fermi surface in momentum space. Taking the above two-level system as an example and assuming that $d_1(\mathbf{k}) = k_x$, $d_2(\mathbf{k}) = k_y$ and $d_3(\mathbf{k}) = k_z$, the magnetic monopole is located at the origin $k_x = k_y = k_z = 0$ and the low-energy Hamiltonian can be written as

$$H(\mathbf{k}) = \pm \mathbf{k} \cdot \sigma$$
$$= \pm \begin{bmatrix} k_z & k_x - ik_y \\ k_x + ik_y & -k_z \end{bmatrix}. \tag{35}$$

This Hamiltonian was first proposed by Weyl [98], who found that, for massless fermions, the 4×4 Dirac representation is reducible, and is composed of two 2×2 (irreducible) Weyl fermions

with positive ($+$) and negative ($-$) chirality (and opposite magnetic charge). Similar to the definition of Chern number for a torus (in a 2D insulator), the total flux of the gauge field passing through a Fermi surface of 3D metal must be quantized as a multiple of 2π. We can then define the Fermi surface Chern number C_{FS} as

$$C_{FS} = \frac{1}{2\pi} \int_{FS} \mathbf{\Omega}(\mathbf{k}) \cdot d\mathcal{S}, \tag{36}$$

and use it as a new topological invariant to describe topological metallic states. The Fermi surface Chern number C_{FS} is non-zero if a Weyl node (or a magnetic monopole) is enclosed by a Fermi surface, and this leads to a non-trivial topological metallic state (called Weyl metal). If the Fermi level happens to be exactly at the Weyl node, we will get a topological semimetal state (i.e. Weyl semimetal). Unfortunately, according to the "no-go theorem" [99, 100], for any lattice model, the Weyl nodes with opposite chirality have to appear in pairs (although they might be separated in momentum space), and the summation of C_{FS} for all pieces of Fermi surfaces must be vanishing [101]. This makes it difficult to give a proper definition for Weyl metal. Nevertheless, we should note that Weyl nodes (and magnetic monopoles) are stable topological objects, which can be well defined in 3D momentum space [102, 103]. Two Weyl nodes with opposite signs may be separated in momentum space and lead to two pieces of Fermi surfaces, each of which has a non-zero Fermi surface Chern number $C_{FS} \neq 0$. The two Weyl nodes may annihilate each other if and only if they overlap in $\mathbf{k}$-space [55, 57–59, 102].

Special attentions must be paid to the systems with both TRS and IS, where Kramers degeneracy exists for every momentum $\mathbf{k}$. In such cases, each pair of Weyl nodes, if any, must overlap exactly in the $\mathbf{k}$-space. In other words, the minimum effective Hamiltonian to describe such systems must be at least 4×4 and contain two Weyl nodes (with opposite signs) simultaneously, as shown below:

$$H(\mathbf{k}) = \begin{bmatrix} \mathbf{k} \cdot \boldsymbol{\sigma} & 0 \\ 0 & -\mathbf{k} \cdot \boldsymbol{\sigma} \end{bmatrix}. \tag{37}$$

This is called a 3D Dirac node, and is a straightforward extension of the 2D graphene to 3D space [104]. As we have discussed above, in such a case, a perturbative mass term can be introduced in principle, opening up a gap and leading to an insulating state. However, if we consider additional symmetries in the system, such as the crystalline symmetry, the mass term may again be forbidden, and this would stabilize the 3D Dirac cone and lead to a 3D Dirac metal or semimetal state [59, 60]. Therefore, the 3D Dirac node is not as stable as the Weyl node, however, it is a good starting point for us to reach the true Weyl semimetal state by breaking either TRS or IS in the 3D Dirac semimetals [58].

In general, the energy level crossing may happen in many materials and at any energy and momentum; however, it is quite rare to have band-crossing exactly at the Fermi level, particularly for the cases with broken TRS or IS. Up to now, only a few materials have been theoretically proposed to host such Weyl semimetal [55–57, 105] or Dirac semimetal states [59, 60, 62]. Among them, only Na_3Bi [59, 63] and Cd_3As_2 [62, 64] have been confirmed experimentally to be Dirac semimetals. The proposed Weyl semimetals breaking TRS, such as pyrochlore iridate [55] and $HgCr_2Se_4$ [56], have not been confirmed yet due to experimental difficulties. The proposals for Weyl semimetals keeping TRS but breaking IS are thought to be a way to overcome those difficulties. Presently, the following are typical proposals: One is a superlattice system formed by alternately stacking normal and topological insulators [106]. The second involves tellurium, selenium crystals or BiTeI under pressure [105, 107]. The third one is the solid solutions $LaBi_{1-x}Sb_xTe_3$ and $LuBi_{1-x}Sb_xTe_3$ [105] tuned around the topological transition points [103].

The fourth is TaAs-family compounds, including TaAs, TaP, NbAs and NbP, which are natural Weyl semimetals and each of them possesses a total of 12 pairs of Weyl points [108, 109].

The Weyl semimetal is a new state of quantum matters. It is of particular interest here that the Weyl semimetal with broken TRS is closely related to the Chern insulators and provides a unique way to realize QAHE. Let us first consider a Weyl semimetal with a single Weyl node in a continuous model defined as $H(\mathbf{k}) = v_F \mathbf{k} \cdot \boldsymbol{\sigma}$, which as mentioned previously generates a monopole in Berry curvature right at the origin. Gauss's theorem then requires that the total flux flowing through out any closed surface enclosing the Weyl point must be ± 1, that is, equal to the chirality of the Weyl point (Equation (36)). For a lattice system, there is an important "no-go theorem" indicating that Weyl points must appear in pairs with opposite chirality. Originally, this theorem has been proved by field theory [99, 100]. Here, we provide a much intuitive way to understand it in terms of the concept of Chern number. Given a Weyl semimetal in 3D lattice (assuming that all trivial states are far away from the Fermi level), first let us consider a 2D plane with fixed k_z, where band structures within the plane should be fully gapped unless the plane cut through a Weyl point exactly. The integral of Berry curvature over such a 2D insulating plane must be quantized and gives rise to a well-defined integer Chern number. Moving the plane (i.e. k_z) from $-\pi$ to π, we will then get the Chern number as a function of k_z, as shown in Figure 6. Now, it is important to note that the Chern number as a function of k_z should jump by $+1$ or -1 (depending on the chirality of the Weyl point), whenever the moving-plane goes across a Weyl point. This is because there exists a topological phase transition at the Weyl point, and the band gap of 2D plane is closed and re-opened when the plane goes across the Weyl point. This jump can also be understood by the following consideration. Selecting two planes at the opposite sides of a Weyl point, we can construct a closed manifold surrounding the Weyl point, that is, the cube formed by the two parallel planes and the four side-surfaces, shown in Figure 6 as the shaded area. The flux flowing through the four side-surfaces should exactly cancel each other, being net zero. Then, the total flux flowing through the closed manifold (i.e. the cube), which is
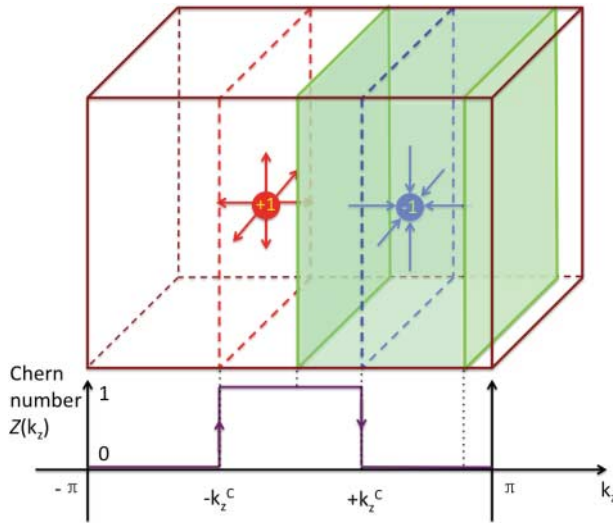

Figure 6. Two Weyl nodes with opposite chirality (± 1) distribute along k_z in 3D momentum space. The Chern number $Z(k_z)$, evaluated for each 2D plane with fixed k_z, is shown at the bottom as a function of k_z. Jumps of Chern number are seen when the plane moves across a Weyl point. A cube containing one Weyl node is shadowed.

now equivalent to the difference in Chern numbers for the two parallel planes, must be equal to the chirality of the Weyl point enclosed by the manifold (namely ± 1) as discussed above. Having understood the jump, let us look at the Chern number evolution as a function of k_z from $-\pi$ to π. In order to satisfy the periodic boundary condition in the BZ of lattice, the positive jump $(+1)$ and the negative jump (-1) must appear in pairs, if there is any. This leads to the important conclusion that Weyl points with opposite chirality must appear in pairs.

Finally, the total Hall conductivity of the system is given by the integral of k_z as

$$\sigma_{xy}^{\text{total}} = \frac{1}{2\pi} \int_{-\pi}^{\pi} \mathrm{d}k_z \sigma_{xy}(k_z). \tag{38}$$

In such a way, we see the relationship between 2D Chern insulators and 3D Weyl semimetals. Given a 3D Weyl semimetal as defined above, we can form a 2D thin film (or quantum-well structure) such that k_z is quantized by the sample thickness, and only particular values of k_z are allowed. If it happens that, for some particular thicknesses, the 2D Chern number are non-zero, we will expect quantized total Hall conductivity, and this leads to a 2D Chern insulator and the QAHE.

To explicitly show the Hall conductivity of an FM Weyl semimetal, we introduce a simple Hamiltonian based on a cubic lattice BHZ model [110]

$$H^{\text{WS}} = \sin(k_x)\sigma_x + \sin(k_y)\sigma_y + [2 + e_s - \cos(k_x) - \cos(k_y) - \cos(k_z)]\sigma_z. \tag{39}$$
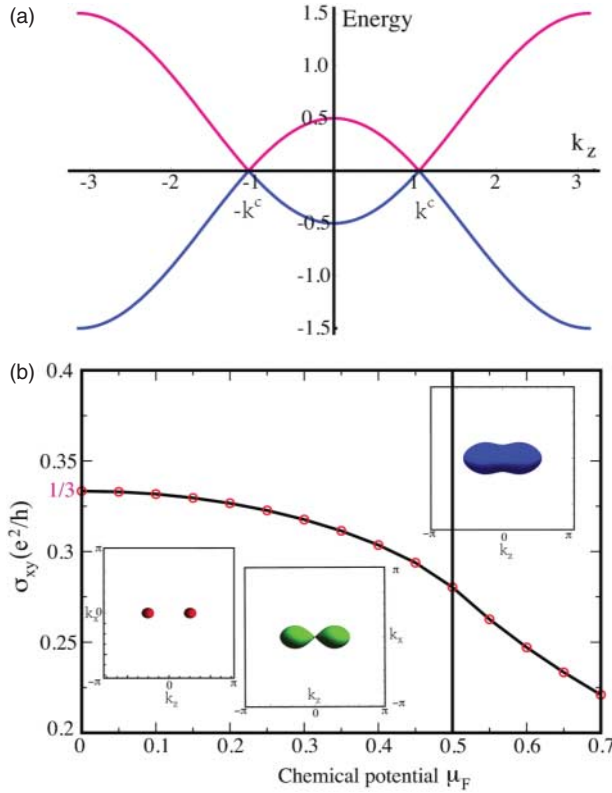

Figure 7. (a) The k_z band dispersion of FM Weyl semimetal. (b) The AHC dependence on the position of chemical potential μ_F. Inset: Fermi surface for different μ_F.

It has two Weyl nodes at $(0, 0, \pm k_z^c)$ with $\cos(k_z^c) = e_s$ when $0 < e_s < 1$. The energy dispersion along the k_z-axis is shown in Figure 7 for $e_s = \frac{1}{2}$. The Hall conductivity $\sigma_{xy}(k_z)$ is calculated by using Equations (18) and (32) for each plane with fixed k_z. When chemical potential μ_F is 0 and passes exactly through the Weyl nodes, $\sigma_{xy}(k_z)$ is quantized to be 1 for $|k_z| < k_z^c$ and 0 for $|k_z| > k_z^c$, consistent with the above discussions based on the continuum model. Integrating $\sigma_{xy}(k_z)$ over k_z, the total Hall conductivity σ_{xy} is given as $(e^2/h)(k_z^c/\pi)$ [111–113], which depends only on the separation of Weyl nodes in momentum space (i.e. k_z^c). Shifting the chemical potential μ_F away from 0 leads to partially occupied bands, and the corresponding Fermi surfaces change from two isolated spheres ($0 < \mu_F < 0.5$) to two touched ones ($\mu_F = 0.5$) and finally they merge into a peanut shape ($0.5 < \mu_F < 1.5$), as shown in Figure 7. The total Hall conductivity decreases with rising chemical potential μ_F, because the Berry curvature from the upper band tends to cancel out the contribution from the lower band. Finally, when both bands are fully occupied the sum of Berry curvature over all bands must vanish [114]. It is important to note that a fully occupied band (without a corresponding Fermi surface) may still contribute to the Hall conductivity, and this should be carefully considered in practical calculations. In such a case, the Hall conductivity is usually evaluated by the volume integrals of the Berry curvature of occupied bands over the whole BZ, as is done for the calculation of the intrinsic AHE [11, 12, 35]; this is also discussed by Chen et al. [111] and Vanderbilt et al. [36, 115–117].

3. First-principles calculations for topological electronic states

Theoretical predictions, particularly first-principles electronic structure calculations based on density functional theory [118], have played important roles in the exploration of topological electronic states and materials. This is not an accidental success, but rather due to several deep reasons: (1) To describe complicated electronic band structures of real materials, particularly with the SOC included, first-principles calculations are necessary; (2) First-principles calculations nowadays can reach the accuracy even up to 90% for most of the physical properties of "simple" materials (i.e. weakly correlated electronic materials), which makes prediction possible; (3) The topological electronic properties of materials are robust, non-perturbative, and not sensitive to small error bars, if any. Given those advantages of first-principles calculations and its great success in this field, however, we have to be aware that the numerical determination of topological invariants (such as integer Z or Z_2 numbers) is still technically demanding, because those numbers are related to the phases of eigen wave functions, which are gauge dependent and randomized in most of the calculations. In addition to that, we also have to note that (1) We still have a band-gap problem in either the local density approximation (LDA) or the generalized gradient approximation (GGA) for the exchange-correlation potential; (2) The present first-principles calculations based on LDA or GGA cannot treat strongly correlated systems properly; (3) The evaluation of Berry phase [27] and topological numbers may require some additional complicities, such as fine **k**-points meshes and gauge-fixing condition. In this section, we will discuss some important issues and techniques related to the first-principles studies of topological electronic states.

3.1. *Calculations of Berry connection and Berry curvature*

From the computational point of view, a self-consistent first-principles electronic structure calculation for a real material is typically performed based on the LDA or GGA for the exchange-correlation potential. Such a calculation will generate a set of single particle eigen states with wave functions $u_{n\mathbf{k}}(\mathbf{r})$, eigen values $\varepsilon_n(\mathbf{k})$ and occupation $f_n(\mathbf{k})$ for the ground state of the system. Having obtained those quantities, the remaining task is to evaluate the Berry connection

$\mathbf{A}_n(\mathbf{k}) = i\langle u_{n\mathbf{k}}|\nabla_{\mathbf{k}}|u_{n\mathbf{k}}\rangle$ and the Berry curvature $\mathbf{\Omega}_n(\mathbf{k}) = \nabla_{\mathbf{k}} \times \mathbf{A}_n(\mathbf{k})$. The Hall conductivity can be obtained either from the summation of Berry curvature over the BZ (Equation (16)) or from the Kubo formula directly (Equation (17)). In the latter case, the matrix element of velocity operator $\hat{\mathbf{v}}$ is given as

$$\langle \psi_{m\mathbf{k}}|\hat{\mathbf{v}}|\psi_{n\mathbf{k}}\rangle = \frac{1}{\hbar}\left\langle u_{m\mathbf{k}} \left|\left[\frac{\partial}{\partial \mathbf{k}}, H_{\mathbf{k}}\right]\right| u_{n\mathbf{k}}\right\rangle = \frac{1}{\hbar}(E_n(\mathbf{k}) - E_m(\mathbf{k}))\left\langle u_{m\mathbf{k}} \left|\frac{\partial}{\partial \mathbf{k}}\right| u_{n\mathbf{k}}\right\rangle. \qquad (40)$$

The strategy is simple, but the computational task is hard, because although the Berry curvature does not depend on the gauge choice, the Berry connection does. It is also true, according to our above discussions, that choosing a smooth gauge for a topologically non-trivial state is generally difficult or even impossible. In most calculations, the Berry connection is a rapidly varying function of momentum $\mathbf{k}$, which makes the convergence of the calculation difficult. To overcome this problem, a straightforward way is to increase the number of $\mathbf{k}$-points in momentum space. Another possible way is to use the Wannier representation, such as the maximally localized Wannier functions (MLWF) [119].

The first-principles calculations of Berry curvature and its integration over the BZ to get the intrinsic AHC have been performed for $SrRuO_3$ and body-centered cubic Fe by Fang et al. [11] and Yao et al. [12], respectively. These were the first quantitative demonstrations of strong and rapid variation of Berry curvature over the BZ. The sharp peaks and valleys make the calculation of AHC challenging. Millions of $\mathbf{k}$ sampling points are required to ensure convergence. Wang et al. [35] proposed an efficient approach by employing the Wannier interpolation technique. They first perform the normal *ab initio* self-consistent calculation by using relatively coarse k-grid—good enough to ensure the total energy convergence. Then, MLWF are constructed for the states around the Fermi level. The number of Wannier functions should be carefully chosen so that the isolated group of bands around the Fermi level can be well reproduced, especially those bands passing through the Fermi level (because the small energy splitting due to SOC or band anti-crossing can produce sharp peaks or valleys of Berry curvature). Once MLWFs are obtained, the interpolation technique is applied to obtain the necessary quantities, such as eigen energies, eigen states and velocity operator, on a much finer k-grid. Since the number of MLWFs is typically much smaller than the number of basis functions used in *ab initio* calculations, this approach saves a lot of computation cost, both in time and in storage.

In addition to the direct methods discussed above, some advanced techniques with well-controlled convergence with respect to the random phases can be developed. This is because Berry connection is not the physical quantity that we can measure, and our ultimate goal is to evaluate the Berry curvature which is gauge-invariant. For example, it was proposed that for calculation of electrical polarization within the modern theory of polarization [33, 34], the change of polarization is directly proportional to the sum of Berry phase terms of occupied states. Since Berry phase is the integral of the Berry connection along a closed loop in the parameter space (see Equation (3)), it is well defined and gauge independent. To calculate the change of polarization along the direction of z, a series of loops (called Wilson loops [120, 121]) in momentum space are constructed. Each loop has fixed $\mathbf{k}_\perp = (k_x, k_y)$, but is periodic along k_z. We can then discretize k_z as $k_z^j = (j/J)G_z$ ($j = 0, 1, \ldots, J-1$), where G_z is the reciprocal lattice vector along k_z. Taking the periodic gauge $u_{n,(\mathbf{k}_\perp,k_z^J)} = e^{-iG_z \cdot \mathbf{r}}u_{n,(\mathbf{k}_\perp,k_z^0)}$, the Berry phase along each loop can be calculated by

$$\gamma(\mathbf{k}_\perp) = \text{Im}\left[\ln \det \prod_{j=0}^{J-1}\langle u_{n,(\mathbf{k}_\perp,k_z^j)}|u_{m,(\mathbf{k}_\perp,k_z^{j+1})}\rangle\right] \qquad (41)$$

with band indices *n* and *m* for occupied states. By doing this, the random phase problem associated with eigen wave functions can be largely eliminated, and the main computational task is to evaluate the inner product of two neighboring wave functions along the loop.

For a metallic system, the intrinsic AHC can also be calculated through this method, as has been implemented by Wang et al. [36]. It was justified by Haldane [117] that the non-quantized part of the intrinsic AHC in an FM metal can be recast as a Fermi surface integration of the Berry curvature, so we can construct the Wilson loops over the Fermi surface rather than the BZ. Considering an FM metal with magnetization *M* along the *z*-direction, a 2D plane (in momentum space) perpendicular to k_z will cut through the Fermi surface. The intersection between the plane and the Fermi surface defines loops, which are the Wilson loops used for calculating the AHC. Following our above discussions, we can discretize each loop and evaluate the Berry phase along the loop efficiently by using Equation (41). To do this, several things should be borne in mind: (1) The periodic boundary condition has to be adapted as $u_{n,(\mathbf{k}_\perp, k_z^J)} = u_{n,(\mathbf{k}_\perp, k_z^0)}$ if the loop does not cross the BZ boundary; (2) The direction of Wilson loops should be defined consistently; (3) For the case with multiple branches, a continuity condition must be used to choose each branch carefully. Finally, in addition to the non-quantized part of AHC, the quantized contribution to the AHC (coming from the deep occupied states) can be determined by the Fermi-sea integration of the Berry curvature, which is gauge invariant and has no ambiguity.

This Wilson loop method is a convenient tool, and it can also be applied to investigate the topological invariant for systems with or without TRS as will be discussed below.

3.2. *Wilson loop method for evaluation of topological invariants*

In real materials, the band structures are usually very complicated, and the determination of topological invariants becomes essential and computationally demanding. The band degeneracy, either accidentally or due to symmetry, makes the numerical determination of phases of wave functions a tough task. Here, we will present the Wilson loop method for determining topological indices efficiently [85, 92, 122, 123]. This method is computationally easy and is equally applicable for Chern insulators, Z_2 TIs (2D and 3D) and the topological crystalline insulators [124, 125], which have attracted lots of interest recently.

For band insulators with IS, the Z_2 indices can easily be computed as the product of parity eigen values for half of the occupied states (Kramers pairs have identical parities) at the TRIM points [126]. In such a case, parity is a good quantum number for the TRIM points and can be computed from the eigen wave functions obtained from the first-principles calculations. This method has been used frequently and is very efficient. Unfortunately, for general cases where IS is absent, the evaluation of the Z_2 number becomes difficult. Three different methods have been considered for such a case: (i) Directly compute the Z_2 numbers from the integration of the Berry curvature $\mathbf{\Omega}(\mathbf{k})$ over half of the BZ [127]. In order to do so, one has to set up a fine mesh in the $\mathbf{k}$-space and calculate the corresponding quantity for each $\mathbf{k}$ point. Since the calculation involves the Berry connection $\mathbf{A}(\mathbf{k})$, one has to numerically fix the gauge on the half BZ, which is not easy for the realistic wave functions obtained by first-principles calculations. (ii) Start from an artificial system with IS, and then adiabatically deform the Hamiltonian towards the realistic one. If the energy gap never closes at any point in the BZ during the deformation process, the realistic system must share the same topological nature with the initial reference system, whose Z_2 number can easily be counted by the parity eigenvalue formula [126]. Unfortunately, making sure that the energy gap remains open on the whole BZ is very difficult numerically, especially in 3D. (iii) Calculate the boundary (edge or surface) states. Due to the open boundary condition, the first-principles calculation for the boundary states is numerically heavy. The Wilson loop method, which we will present here, has the following advantages: first, it uses only the periodic

bulk system; second, it does not require a gauge-fixing condition – thereby greatly simplifying the calculation; third, it can easily be applied to a general system with or without inversion symmetry.

As discussed in the previous section for 2D insulators, to determine the topological number (either Z or Z_2), the key is the $\theta(k_y)$ angle and its 1D evolution. In the single band model, the $\theta(k_y)$ angle is given simply as the 1D integration of $A_x(k_x, k_y)$ along the k_x-axis. However, for real compounds with multiple bands, its computation requires the Wilson loop method [92]. Once the θ angle is computed, all the topological numbers can be determined easily, following our previous discussions. For the mathematic details of this method, refer to our paper [92]. Here, we will focus on the practical steps for real calculations.

Considering an insulator with N occupied states, after the self-consistent electronic structure calculations, we get the converged charge density, the total energy, the band structure, etc. Then, without loss of generality, we consider a periodic plane of (k_x, k_y) with $-\pi < k_x < \pi$ and $-\pi < k_y < \pi$, and consider the k_x-axis as a closed loop. For each fixed k_y, we discretize the k_x line into an N_k mesh with mesh-index i ranging from 0 to $N_k - 1$. Due to the periodic condition, the $i = N_k$ point is equivalent to the $i = 0$ point. The Wilson loop method consists of the following steps:

(1) We need to calculate the eigen wave functions of occupied states for all mesh points. This can be done by continuing from the converged charge density and with no need for charge self-consistency furthermore.

(2) Using the obtained wave functions, we can calculate the inner products among them and evaluate the $N \times N$ matrix $F_{i,i+1}^{m,n} = \langle m, k_{x,i}, k_y \,|\, n, k_{x,i+1}, k_y \rangle$, where m and n are band indices of occupied states ranging from 1 to N. The F matrix [92] is nothing but the discretized version of Berry connection A_x.

(3) Now, we can construct the $N \times N$ matrix D as the product of F, $D(k_y) = F_{0,1} F_{1,2} \ldots F_{N_k-2,N_k-1} F_{N_k-1,0}$.

(4) Diagonalization of the $D(k_y)$ matrix gives us N eigenvalues $\lambda_n(k_y)$, whose phase angle can be obtained as $\theta_n(k_y) = \text{Im}[\log \lambda_n(k_y)]$.

This phase angle $\theta_n(k_y)$, which is now band dependent, is exactly what we need to determine the winding number and the corresponding topological numbers, as discussed in Section 2. In practice, for a system with many bands, we can draw an arbitrary reference line on the cylinder surface, parallel to the k_y-axis, and then count how many times the evolution lines of $\theta_n(k_y)$ cross the reference line [92]. For the calculation of Z_2 in time-reversal invariant systems, we can further reduce the computational task by considering only the $0 < k_y < \pi$ part. As an example, we show in Figure 8 the calculated θ_n and their evolutions for one 3D TI proposed recently [128]. All the six planes, each of them passing through four TRIM points, are shown. We see that the θ_n lines cross the reference line an odd number of times for all three planes containing Γ, but an even number of times for those planes containing (π, π, π). Therefore, we conclude that it is topologically non-trivial with Z_2 index (1;000).

Finally, through the whole calculation process, we can always take advantage of Wannier functions, particularly the MLWF as proposed by Vanderbilt [119, 129]. For example, after the self-consistent electronic structure calculations, we can construct the MLWF for the energy bands close to the Fermi level, which can usually be reproduced up to a very high accuracy. In this process, however, some of the unoccupied bands should be included [129]. Once the MLWF are obtained, we can easily calculate the eigen wave functions and their inner products for any **k**-points using the interpolation technique [35, 130]. This will greatly simplify our task for the computation of topological numbers. In addition, the MLWF are also useful for the calculations of boundary states, as discussed below.
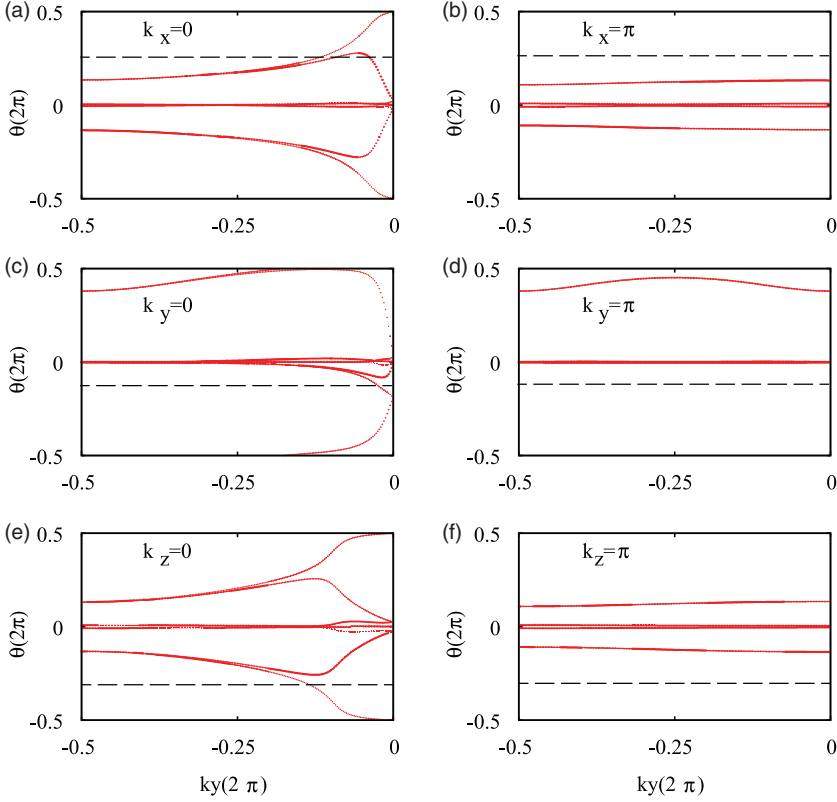
Figure 8. (Color online) Example of Wilson loop method for the 3D strong TI with Z_2 index (1;000). The calculated $\theta_m(k_y)$ for all occupied bands (m is the band index) are plotted as functions of k_y. The dashed lines are arbitrary reference lines. The $\theta_m(k_y)$ crosses the reference line an odd number of times in half of k_y direction ($k_y \in [0, \pi]$ or $[\pi, 2\pi]$) for (a) $k_x = 0$, (c) $k_y = 0$ and (e) $k_z = 0$ plane (left panel), but an even number of times for (b) $k_x = \pi$, (d) $k_y = \pi$ and (f) $k_z = \pi$ plane (right panel). Figures from Ref. [128].

3.3. *Boundary states calculations*

After the determination of bulk topological numbers by either of the methods presented above, the calculation of boundary states of real compounds becomes necessary. This is not just for the purpose of confirming topological properties – it is also very useful for experiments, such as angle-resolved photo-emission spectroscopy (ARPES) and scanning tunneling microscope (STM), because it provides necessary information for direct comparison.

In practice, two methods are usually used for the calculation of boundary states. First, the boundary states can be directly calculated from first-principles calculations by using a supercell of slab geometry, as is usually done for other surface or edge calculations [131]. In this way, the atomic details, such as the surface or edge terminations, the atomic relaxations or reconstructions, the possible defects or absorptions, can be carefully treated and studied. This method is straightforward, although it is computationally demanding. Due to the presence of boundary states, we have to use a very large supercell with a thick vacuum region to avoid possible mutual coupling between the boundaries. The convergence test as a function of cell size is usually necessary. This method can generate detailed surface eigen states, on the basis of which the spin texture of topological surface states (showing characteristic spin–momentum locking effect) can be obtained.

With SOC included, the eigen states are expressed as two-component spinors, $\psi_{n\mathbf{k}} = \begin{pmatrix} \phi_{n\mathbf{k}\uparrow} \\ \phi_{n\mathbf{k}\downarrow} \end{pmatrix}$, and the expectation value of the spin operator $\mathbf{S} = (\hbar/2)\boldsymbol{\sigma}$ can be calculated easily by $\psi_{n\mathbf{k}}^{\dagger}\mathbf{S}\psi_{n\mathbf{k}}$ [104, 128]. In addition, the layer-dependent distribution of surface states (i.e. the penetration depth) can also be studied carefully through these eigen states [104].

Second, we can take advantage of MLWF and calculate the boundary states from Green's functions of a semi-infinite system [132]. We can use the MLWF as a basis to construct an effective low-energy Hamiltonian H_{WF}, which can be regarded as a tight-binding model with its hopping parameters determined from the first-principles electronic structure calculations. The maximally localized property of MLWF guarantees the short-range hopping. Using this H_{WF} as building blocks, we can construct a large Hamiltonian H_{semi} for the semi-infinite system with only one boundary. This H_{semi} is a block tridiagonal matrix if we consider only the hopping terms between the nearest-neighbor building blocks. In practice, we find that the thickness of the building block should also be tested to ensure convergence since in some cases the hopping among MLWFs might extend to more than the nearest-neighbor. Then, we can use the iterative method [133–135] to solve the problem and get the projected Green's functions onto the boundary. From these boundary Green's functions, we can get necessary information such as the charge density of states and the spin density of states at the boundary [136, 137]. This procedure is an approximate way and it cannot treat the atomic details at the boundary precisely, however, it provides the most important information for the gapless nature of the boundary states of TIs.

4. Material predictions and realizations of QAHE

The topological electronic states discussed in previous sections are important and interesting, not only because they are conceptually new but also because they are realizable and testable in condensed materials. In recent years, this field has been developing very fast. Many candidate compounds with non-trivial topological electronic states are proposed, and some of them have been successfully synthesized and confirmed experimentally. The realizations of various topological electronic states have strongly promoted this field, and a lot of interesting experiments on their non-trivial topological properties are now becoming possible. In this section, we will focus on the topic of how to realize the QAHE in real materials.

In spite of great success in recent studies of TIs and QSHE, the realization of QAHE has taken a long time. The key to achieving this goal is to find a proper Chern insulator. Any realistic material system in which QAHE can be realized must satisfy the following conditions:

(1) It must be 2D.
(2) It must be insulating in the bulk.
(3) It must break the TRS with a certain magnetic ordering (the simplest case is an FM long-range order).
(4) The occupied bands must carry a non-zero Chern number (therefore strong SOC is usually required).

It is not difficult for a material to satisfy one or two conditions in the above list. But finding a realistic material that satisfies all of these conditions simultaneously turns out to be very challenging. This is the reason the Chern insulator is theoretically proposed long before the discovery of TIs, but its realization came much later. Nevertheless, the discovery of QSHE and TIs [38, 39, 41, 43, 49, 50, 93] have greatly stimulated the field of QAHE. It is natural to expect that starting from known 2D TIs and then breaking the TRS will be one of the simple ways to achieve the QAHE. The 2D TIs already satisfy three out of those four conditions, and these TIs can be effectively

viewed as two Chern insulator layers related by TRS. Breaking of TRS can, in principle, destroy one layer and keep the other layer active, which will lead to the QAHE. Based on this idea, we can go further: if the exchange-splitting is strong enough, we can even realize the QAHE starting from a 2D narrow gap semiconductor, where the strong exchange-splitting may cause band inversion for half of the states (i.e. for only one of the spin subspaces, as will be addressed below). Unfortunately, it is still very hard to induce exchange-splitting in 2D TIs or narrow gap semiconductors; two known procedures, the magnetic proximity effect and magnetic doping, are most commonly employed. Despite the issue of magnetic ordering temperature T_c (which can be very low), magnetic doping will obviously introduce impurities that may destroy the insulating state of parent compounds. On the other hand, using the magnetic proximity effect may have the advantage of avoiding the problem of impurity doping, while the possible exchange-splitting induced by this procedure is typically very weak. Therefore, in practice, we have to choose the candidate compounds and procedures carefully in order to balance various conditions and requirements. In this sense, the first-principles calculations play important, predictive roles for the final realization of the QAHE. As listed below, up to now, there have been several different types of proposals for the realization of the QAHE, and most of them start from TIs as the parent compounds.

- To sandwich a TI thin film with FM insulators. The FM proximity effect will open up a gap for the Dirac surface states on both surfaces of the TI, which gives rise to the QAHE [74].
- To dope the magnetic ions into the TI thin film and make it an FM semiconductor [138–140]. The QAHE can be obtained by proper control of the doping. This is the only one of these proposals that has succeeded so far [141, 142].
- To make a thin film (or quantum-well structure) of an FM Weyl semimetal (such as $HgCr_2Se_4$). The quantum size effect gives rise to the quantization of the crystal momentum along the growing axis, which at the same time quantizes the anomalous Hall coefficient at a certain film thickness [56, 143].
- To grow graphene, silicene or other honeycomb lattice on top of magnetic insulators or with magnetic adsorbates. The low energy bands (which are tunable by external electric field) acquire both the Zeeman exchange-splitting and the SOC, which drive the system towards QAHE [144–155].
- To mimic the honeycomb lattice by growing double-layer thin films of a transition metal oxide with perovskite structure along the (111) direction, or by forming a sheet of organic molecules that contain transition-metal ions (TMOs). The magnetic long-range order might be stabilized among the d electrons of the transition metal. The QAHE can be realized in such systems if a suitable exchange field is established [156–163].
- Other proposals, such as heavy metal on magnetic insulator substrate [164], strained epitaxial film with interface band inversion in EuO/GdN [165] or CdO/EuO [166].

Most of the proposals mentioned above are supported by first-principles calculations, and we will start this section with a brief introduction about how to explore new topological materials from the viewpoint of a band inversion mechanism. Then, we will discuss details of some of the proposals for the prediction and realization of the QAHE, with no attempting to cover them all.

4.1. *Band inversion mechanism*

Though we have found, in the previous sections, that the Wilson loop method can reduce computation effort in calculating various topological invariants, it is still quite tedious and inefficient for wide-range searching or design of topological materials. From this point of view, band inversion is the most intuitive picture and most practically useful guideline for the initial screening of possible topological materials. Band inversion involves the energy order switching of low energy
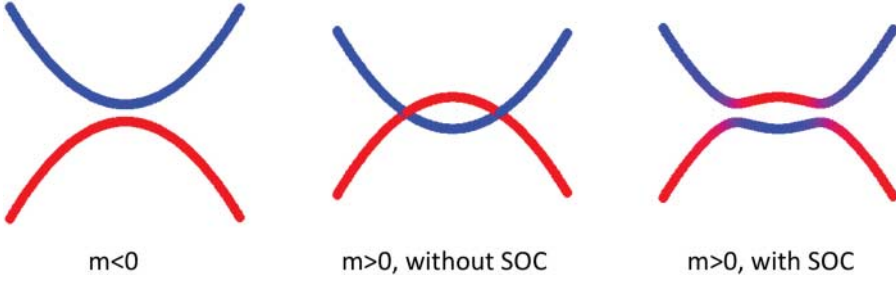
Figure 9. (Color online) Illustration of the band inversion mechanism. Schematic plot for the band structure with normal order ($m < 0$), inverted order ($m > 0$) without SOC, and inverted order with SOC, respectively. (See main text for detailed explanations).

electronic bands around a certain TRIM point in the BZ when compared with the energy ordering in the atomic limit (which is topologically trivial). Band inversion may happen when atoms form crystalline solids and acquire strong enough band dispersions. The mechanism was first demonstrated for a HgTe/CdTe quantum-well structure [41] and later was found to be applicable to many other materials, including 2D and 3D TIs [43, 53, 167–169], Chern insulators [139], topological Weyl semimetals [55, 56] and topological Dirac semimetals [59,62].

As schematically illustrated in Figure 9, we consider two bands close to the Fermi level around the Γ point. Far away from the Γ point, the red state is occupied and energetically lower than the blue state. However, close to the Γ point, they have an inverted energy ordering with the blue state lower and occupied at Γ. Let us assume that two bands do not couple when the SOC is absent, and they must cross at certain k-point away from the Γ point. For time-reversal invariant systems with SOC, the spin degree of freedom has to be taken into account, and at least four bands must be included in this simple model. The band crossing points are, in general, not stable and will open up a gap (in the presence of SOC). As a result, the system becomes an insulator if the chemical potential is located within the gap, and most importantly, this gapped state may be topologically non-trivial depending on the coupling terms. It should be noted that if additional crystal symmetries are considered, the fourfold-degenerated band crossings are possibly protected. This situation is demonstrated in topological Dirac semimetals [59,62].

To be more explicit, we consider two parabolic bands with energy dispersion $E_{\pm}(\mathbf{k}) = \pm(-m + \mathbf{k}^2)$. The $+$ and $-$ signs are for the blue and red bands, respectively. The band inversion (around $\mathbf{k} = 0$) happens for the parameter region of $m > 0$. We first focus on a 2D system with broken TRS. Using two bands as a basis, we can construct an effective 2×2 Hamiltonian,

$$H_{\text{eff}}(\mathbf{k}) = \begin{bmatrix} E_+(\mathbf{k}) & M(\mathbf{k})^* \\ M(\mathbf{k}) & E_-(\mathbf{k}) \end{bmatrix}, \tag{42}$$

where $M(\mathbf{k})$ is the coupling between the two bands that determines the topological properties of the resulting electronic structure. If $m < 0$, that is, without band inversion, we will in general expect nothing but a normal insulator. However, if $m > 0$ for the band-inverted system, we may find many interesting topological phases, depending on the choice of $M(\mathbf{k})$. For example, if $M(\mathbf{k}) = (k_x \pm ik_y)^n$ ($n \geq 1$), we will get a Chern insulator with Chern number $Z = n$ (for a 2D system). For a 2D system with TRS, the two bands should be replaced by four bands with each color (blue and red) representing a Kramers pair of bands, and the model Hamiltonian becomes 4×4 (block diagonal in its simplest form, such as the BHZ Hamiltonian [41]). If we now have a form of $M(\mathbf{k}) = (k_x \pm ik_y)^{2n-1}$, we will expect a 2D TI with $Z_2 = 1$. In the presence of SOC, these forms of coupling term $M(\mathbf{k})$ are in general allowed, and they depend on the material's

details. Similar discussions can be extended to 3D systems, and in this case, we can expect either 3D TIs [170] or 3D topological semimetals [56, 59, 62]. In real compounds, the band inversion may happen between $s - p$, $p - p$, $p - d$ and $d - f$ states, which leads to different material classes of TIs [167].

The band inversion mechanism discussed above is practically useful, yet a crude picture for real compounds. First, this picture comes from the low energy effective Hamiltonian of a continuous model. If we put the model onto a lattice, we have to be careful about the number of times band inversion occurs. Particularly for the Z_2 TIs, if band inversion happens an even number of times or it happens around an even number of TRIM points, we will return to a trivial insulator phase (in the sense of $Z_2 = 0$, although other topological phases, such as topological crystalline insulator [124, 125], may be defined). Second, real compounds may have many low energy bands that are strongly fixed and must be carefully analyzed. Third, the possible $M(\mathbf{k})$ terms can also be complicated and need to be treated carefully. In any case, a final conclusion must be carefully made based on the first-principles calculations as presented in previous sections.

4.2. *QAHE in magnetic TIs*

Several realistic systems, including Mn-doped HgTe quantum wells [138], magnetic-impurity doped Bi_2Se_3 thin films [139], Mn-doped InAs/GaSb quantum wells [140], and magnetically doped junction quantum wells [171], have been proposed to support the QAHE. The essence of these proposals is to break the TRS in a 2D slab of QSHI by magnetic doping. If the FM order can be established, the introduced exchange splitting may destroy the band inversion in half of the states (say, for one spin channel), while keeping the band inversion in the other half of the states, leading to the QAHE.

A HgTe quantum well with an appropriate thickness realized between two CdTe barriers can achieve the 2D QSH state [41, 42]. Based on this structure, Liu et al. proposed that the QAHE can be obtained by magnetic Mn doping in HgTe/CdTe quantum wells [138], if the doped Mn ions are ferromagnetically ordered with magnetic momentum polarized along the out-of-plane direction. Unfortunately, $(Hg_{1-x}Mn_x)Te$ is experimentally paramagnetic (rather than FM), which prohibits the realization of QAHE. Due to the paramagnetic nature, in practical measurement, a magnetic field is required to polarize the Mn magnetic moment. Therefore, it is hard to distinguish the Hall and the anomalous Hall contributions. With the same idea, Zhang et al. [171] proposed that a QSHI with bulk band gap around 0.1 eV can be realized in junction quantum wells comprising II–VI, III–V or IV semiconductors, and proper magnetic ordering can drive the system into the QAH state. Experimentally again, we have to wait for the synthesis and the well-controlled doping of a sample. In the following, we will take Bi_2Se_3 and Bi_2Te_3 (well-known TIs), as paradigms to show how the QAHE state is achieved in their thin film form with Cr doping. This is the only experimentally successful example so far [141].

The Bi_2Se_3 family of compounds have a rhombohedral crystal structure with space group D_{3d}^5 ($R\bar{3}m$). The system has a layered structure with five atomic layers as a basic unit, named a quintuple layer (QL), as shown in Figure 10. The materials Bi_2Se_3, Bi_2Te_3 and Sb_2Te_3 have been theoretically predicted and experimentally observed to be TIs with a large bulk band gap [43–52]. Recent experimental progress has shown that well-controlled layer-by-layer MBE thin film growth can be achieved [45, 172], and various magnetic transition metal elements (such as Ti, V, Cr and Fe) can be substituted into the Bi_2Se_3 family of parent compounds with observable ferromagnetism even above 100K [173–177].

We first discuss how to achieve an FM insulating phase in a semiconductor doped with dilute magnetic ions. To simplify the discussion, we divide the whole system into two subsystems describing the local moments and the other band electrons, respectively. We further assume that
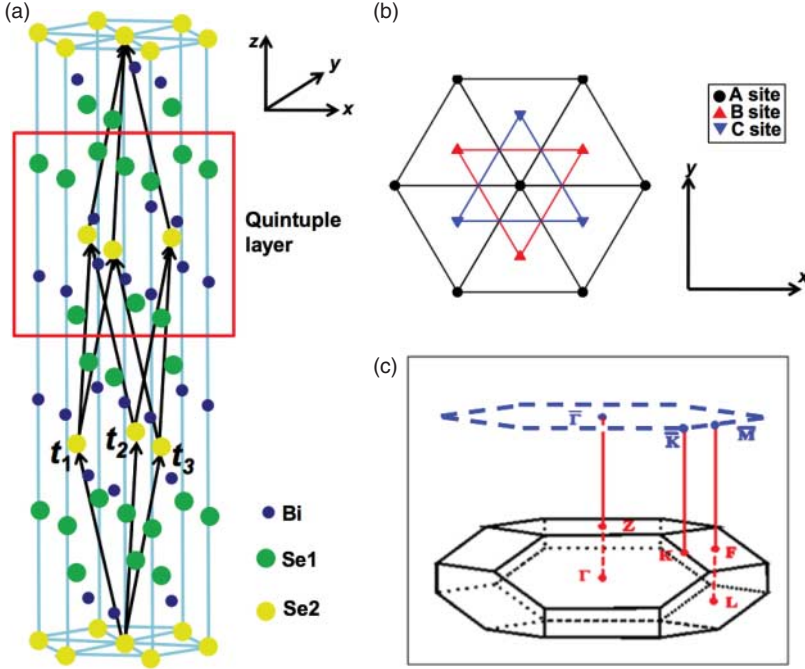
Figure 10. (Color online) Crystal structure of the Bi_2Se_3 family of compounds. (a) The hexagonal super-cell containing 15 atomic layers. (b) Top view of a QL in the triangular lattice. Three sets of different sites, labeled as A, B and C sublattices are presented. (c) The first BZ. Four non-equivalent TRIM points $\Gamma(0,0,0)$, $L(\pi,0,0)$, $F(\pi,\pi,0)$ and $Z(\pi,\pi,\pi)$ are denoted in the 3D BZ. The corresponding surface 2D BZ is represented by the dashed hexagon, and $\bar{\Gamma}$, $\bar{M}$ and $\bar{K}$ are the corresponding TRIM special k points in the surface BZ. Figure adapted from Ref. [104].

the magnetic exchange among local moments is mediated by the band electrons. If we consider only the spatially homogeneous phase, the total free energy of the system in an external magnetic field H can be written as

$$F_{\text{total}} = \tfrac{1}{2}\chi_L^{-1}M_L^2 + \tfrac{1}{2}\chi_e^{-1}M_e^2 - J_{\text{eff}}M_LM_e - (M_L + M_e)H, \tag{43}$$

where $\chi_{L/e}$ is the spin susceptibility of the local moments/electrons, $M_{L/e}$ denotes the magnetization for the local moment and electron subsystem and J_{eff} is the magnetic exchange coupling between them. The spontaneous FM phase can be realized when the minimization procedure of the free energy gives a non-zero magnetization in the zero magnetic field limit ($H = 0$). This leads to the requirement of $J_{\text{eff}}^2 - \chi_L^{-1}\chi_e^{-1} > 0$, or equivalently $\chi_e > 1/(J_{\text{eff}}^2\chi_L)$. This result means that in order to have a non-zero FM transition temperature (T_c), a sizable χ_e is needed. Within linear response theory, the spin susceptibility χ_e for a insulator is given by the van Vleck formula

$$\chi_e^{zz} = \sum_{v,c,k} 4\mu_0\mu_B^2 \frac{\langle vk|\hat{S}_z|ck\rangle\langle ck|\hat{S}_z|vk\rangle}{E_{ck} - E_{vk}}, \tag{44}$$

where μ_0 is the vacuum permeability, μ_B is the Bohr magneton, $|ck\rangle$ and $|vk\rangle$ are the Bloch functions in the conduction and valence bands respectively, E_{ck} and E_{vk} are the eigen energies of $|ck\rangle$ and $|vk\rangle$, and $\hat{S}_z$ is the spin operator of electrons and we consider only the z-direction, that is, perpendicular to the 2D plane, for simplicity.
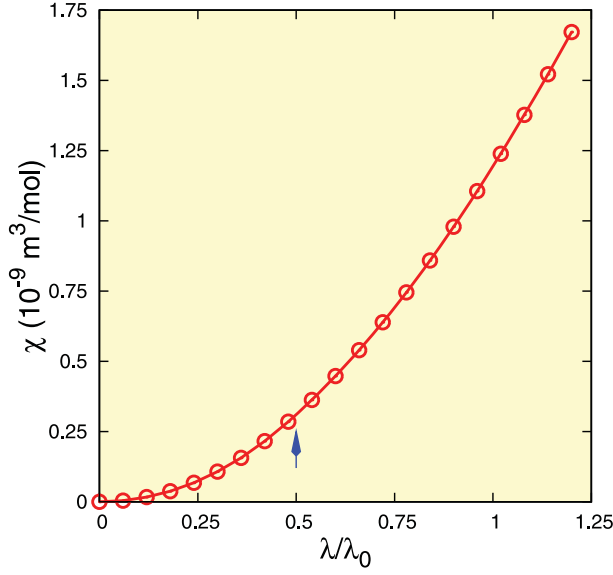
Figure 11. Van-Vleck-type spin susceptibility of Bi_2Se_3 bulk as a function of SOC (SOC) strength (λ_0 is the actual SOC strength). Figure from Ref. [139].

In most dilute magnetic semiconductors, such as $(Ga_{1-x}Mn_x)As$, which has *s*-like conduction and *p*-like valence bands, the matrix elements $\langle ck|\hat{S}_z|vk\rangle$ are very small. The electronic spin susceptibility is therefore negligible for the insulating phase, and a finite carrier density is required to mediate the magnetic coupling (for example, by the RKKY mechanism [178]). For the Bi_2Se_3 family of materials, however, the semiconductor gap is opened by the SOC between two *p*-orbital bands. The spin operator S_z has finite matrix elements between the valence and conduction bands, which are further enhanced due to the band inversion [139]. In Figure 11, we show the calculated spin susceptibility of bulk Bi_2Se_3. At the Γ point, the band inversion occurs when the relative SOC strength λ/λ_0 exceeds 0.5, and after that the spin susceptibility starts to increase rapidly. At the actual SOC strength, Bi_2Se_3-family materials have a considerable spin susceptibility χ_e^{zz}. Hence the local moments of the magnetic impurities in such materials can be ferromagnetically coupled through the van Vleck mechanism [179]. Recently the interplay between magnetism and the topological property in a material has been discussed in more details in Ref. [180, 181].

From first-principles calculations, we find that the insulating magnetic ground state discussed above can indeed be obtained by a proper choice of magnetic dopants. Experiments [173–175] and theoretical calculations [139, 182] show that the magnetic dopants will mostly substitute for Bi ions, which has a nominal 3+ valence state. In order to avoid introducing free carriers into the parent material when we dope the system, it is natural to choose those transition metal elements that also have a stable 3+ chemical state, such as Ti, V, Cr and Fe. Self-consistent first-principles calculations for Bi_2Se_3 doped with Ti, V, Cr and Fe have been performed. The calculated densities of states (DOS) shown in Figure 12 suggest that an insulating magnetic state is obtained only for Cr or Fe doping, while the states are metallic for Ti or V doping cases.

The above results can be understood as illustrated in Figure 13. We first point out that, for all the cases, the dopants are nearly in the 3+ valence state, and we always obtain the high-spin state due to the large Hund's coupling of 3*d* transition metal ions. For the Fe-doped case, the Fe^{3+} favors the $d^{5\uparrow}d^{0\downarrow}$ configuration in a high-spin state, resulting in a gap between the majority and minority spins, which directly leads to the insulating state in Fe-doped samples as shown in
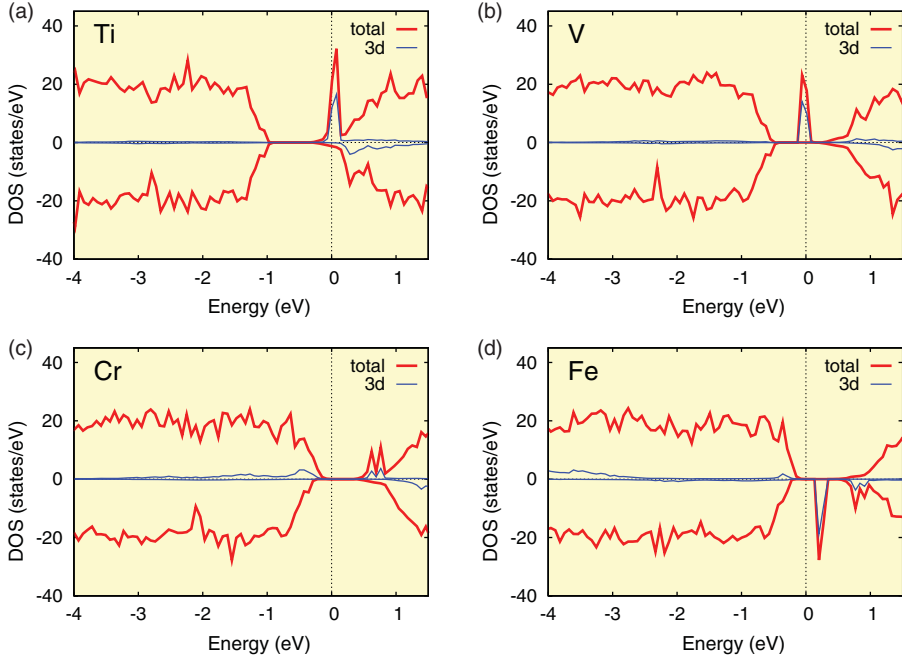
Figure 12. Calculated densities of states (DOS) for Bi_2Se_3 doped with different transition metal elements. The Fermi level is located at energy zero, and the positive and negative values of DOS are used for up and down spin, respectively. The lines are projected partial DOS of the $3d$ states of transition metal ions. It is shown that Cr or Fe doping will give rise to the insulating magnetic state. Figure adapted from Ref. [139].
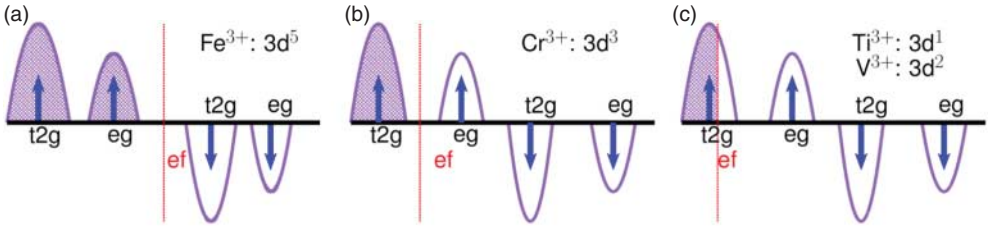

Figure 13. The configuration of the d electrons for the (a) Fe, (b) Cr, (c) Ti and V doping cases.

Figure 13(a). For the Cr-doped case, Cr ions substitute in the Bi sites and feel an octahedral crystal field formed by the six nearest neighboring Se^{2-} ions. Such a local crystal field splits the d-shell into t_{2g} and e_g manifolds. This splitting is large enough to stabilize the $t_{2g}^{3\uparrow}e_g^{0\uparrow}t_{2g}^{0\downarrow}e_g^{0\downarrow}$ configuration of a Cr^{3+} ion, resulting in a gap between the t_{2g} and e_g manifolds, as shown in Figure 13(b). For the case of Ti or V doping, the t_{2g} manifold is partially occupied, leading to the metallic state, as shown in Figure 13(c). We note that, although the LDA in density functional theory may underestimate the electron correlation effects, the inclusion of electron–electron interaction U (such as in the LDA $+ U$ method) should further enhance the gap (it may also reduce the $p - d$ hybridization and T_c [183]).

The exchange field Δ can be estimated as $\Delta = xJ_{\text{eff}}\langle S \rangle$ at the mean field level, where x is the doping concentration and $\langle S \rangle$ is the mean field expectation value of the local spin, J_{eff} is the effective exchange coupling between the local moments and the band electrons, which is

estimated to be around 2.7 eV for Cr doping and 2.8 eV for Fe doping in Bi_2Se_3 by first-principles calculations. With the concentration of the magnetic dopants at 10%, the FM Curie temperature can be estimated in the mean field approximation [184] to reach the order of tens of K.

Once the FM order is achieved in TI at a reasonable temperature, the QAHE can in principle be realized in 2D thin films of such systems by tuning the exchange splitting (i.e. the density of magnetic doping). Since the bulk states are always gapped, we focus on the simplest low energy effective Hamiltonian consisting of Dirac-type surface states only:

$$
\begin{aligned}
H &= H_{sf} + H_{\text{Zeeman}} \\
&= \begin{bmatrix} -v_F(k \times \sigma)_z & m_k^* \\ m_k & v_F(k \times \sigma)_z \end{bmatrix} + \begin{bmatrix} gM\sigma_z & 0 \\ 0 & gM\sigma_z \end{bmatrix},
\end{aligned}
\tag{45}
$$

with the basis of $|t \uparrow\rangle, |t \downarrow\rangle, |b \uparrow\rangle, |b \downarrow\rangle$, where "$t$" ("$b$") represents the surface states on the top (bottom) surface, and "$\uparrow$", "$\downarrow$" represent the spin up and down states, respectively. v_F is the Fermi velocity, m_k describes the tunneling effect between the top and bottom surface states, g is the effective g-factor, σ are the Pauli matrices, and M represents the exchange field along the z-direction, introduced by the FM ordering. For simplicity, spatial IS is assumed, which requires that v_F, g and M take the same values for top and bottom surfaces. In thick enough slab geometry ($m_k \approx 0$), the spatially separated pairs of surface states are well defined for the top and bottom surfaces. However, with the reduction of the film thickness, quantum tunneling between the top and bottom surfaces becomes more and more pronounced, giving rise to a finite mass term m_k, which can be expanded as $m_k = m_0 + B(k_x^2 + k_y^2)$ up to the second order of k. The Hamiltonian can be rewritten in terms of the symmetric and anti-symmetric combination of the surface states on top and bottom surfaces as

$$
H = \tilde{H}_{sf} + \tilde{H}_{\text{Zeeman}} = \begin{bmatrix} h_k + gM\sigma_z & 0 \\ 0 & h_k^* - gM\sigma_z \end{bmatrix},
\tag{46}
$$

with the following new basis: $|+ \uparrow\rangle, |- \downarrow\rangle, |+ \downarrow\rangle, |- \uparrow\rangle$, where $|\pm \uparrow\rangle = (|t \uparrow\rangle \pm |b \uparrow\rangle)/\sqrt{2}$, $|\pm \downarrow\rangle = (|t \downarrow\rangle \pm |b \downarrow\rangle)/\sqrt{2}$. Here $h(k) = m_k\sigma_z + v_F(k_y\sigma_x - k_x\sigma_y)$, which is similar to the BHZ model describing the low energy physics in HgTe/CdTe quantum wells [41]. When $m_0B < 0$, band inversion occurs, and the system will be in the QSH phase if this is the only band inversion between two subbands with opposite parity. Regardless of whether this condition is satisfied or not, a strong enough exchange field will induce the QAH effect in this system, thanks to the presence of the σ_z matrix in the exchange field ($gM\sigma_z$) and the opposite signs of the Zeeman coupling terms in the upper and lower blocks of the effective Hamiltonian in Equation (46). The exchange field breaks the TRS and increases the mass term of the upper block, while reducing the mass term for the lower block. More importantly, a sufficiently large exchange field can change the Chern number of one of the two blocks. There are two cases, as illustrated in Figure 14. In the first case, the four-band system is originally in the topologically trivial phase; the exchange field will induce band inversion in the lower block and push the two subbands in the upper block even farther away from each other. Therefore, the 2D model with a negative mass in the lower block contributes $-e^2/h$ for the Hall conductance (see Figure 14(a)). In the other case, the system is originally in the topologically non-trivial phase, and both blocks have inverted band structures. A sufficiently large exchange field can increase the band inversion in the lower block and release the band inversion in the upper block. Again the negative mass in the lower block contributes $-e^2/h$ for the Hall conductance (see Figure 14(b)). Such a mechanism is general for the thin films of TIs with FM ordering, it is guaranteed by the fact that the surface states on the top and bottom surfaces have the same g-factor.
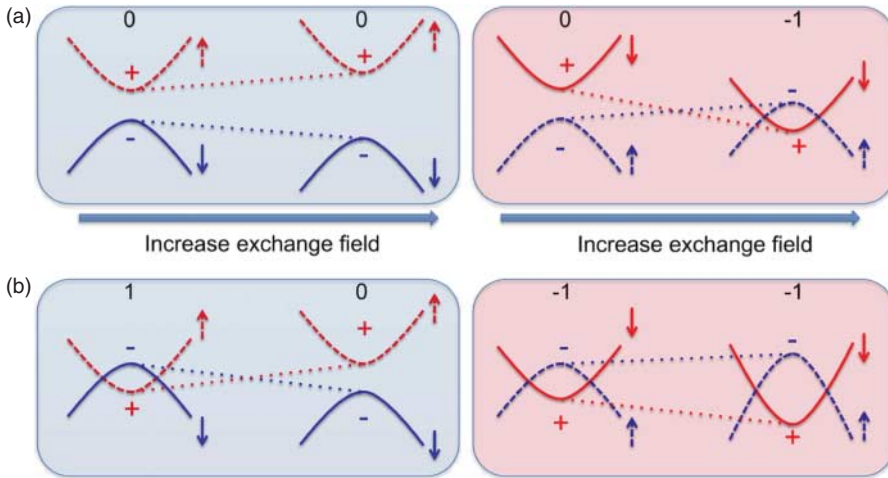
Figure 14. Evolution of the subbands structure upon increasing the exchange field. The red color denotes the subbands that have even parity at the Γ point, and blue color denotes subbands with odd parity at the Γ point. The dashed lines denote the spin up electrons; solid lines, spin down electrons. (a) The initial subbands are not inverted. When the exchange field is strong enough, a pair of inverted subbands appears (red solid line and blue dashed line). (b) The initial subbands are already inverted. The exchange field releases the band inversion in one pair of subbands (red dashed line and blue solid line) and increases the band inversion in the other pair (red solid line and blue dashed line).

In Figure 15(A)–(C), we plot the lowest four subband levels at the Γ point as functions of exchange field. Level crossings between the lowest conduction bands (blue lines) and valence bands (red lines) are found for all three values of layer thickness, signaling the quantum phase transition to the QAH state. In the insulator case, where the chemical potential is located inside the energy gap between the conduction and valence subbands, the Hall conductance is determined by the Chern number of the occupied bands and must be an exact integer multiple of e^2/h. The calculated Hall conductance for Bi_2Se_3 thin films of three typical thicknesses (Figure 15(D)) jumps from 0 to 1 at the corresponding critical exchange field for the level crossing. The QAH effect with higher Hall conductance plateaus in the Cr-doped $Bi_2(Se, Te)_3$ was further discussed in Ref. [185], where the stronger exchange field leads to more band crossings and a larger Chern number.

In Figure 16, we plot the subband dispersion and corresponding Hall conductance as functions of chemical potential for 3 QL Bi_2Se_3 thin films with different exchange fields. A quantized plateau in Hall conductance should be observed when the chemical potential is inside the gap. For the QAH phases (Figure 16(D)), a non-zero integer plateau is observed when the chemical potential falls into the energy gap, while for the non-QAH phases (Figure 16(B)), the corresponding Hall conductance is zero. In a recent experiment, Chang et al. [141] prepared the Cr-doped $(Bi,Sb)_2Te_3$ thin films with well-controlled chemical potential and long-range FM order. Careful control of the Bi and Sb concentrations is needed to separate the surface states from the bulk states. By tuning the gate voltage (and therefore, the chemical potential), they observed the quantization of the Hall resistance at h/e^2 at zero external magnetic field with temperature around 30 mK, accompanied by a considerable drop in the longitudinal resistance. A similar experiment was further performed by Checkelsky et al. [142], confirming the existence of QAHE in this system. For readers who want to know more experimental details, refer to their original
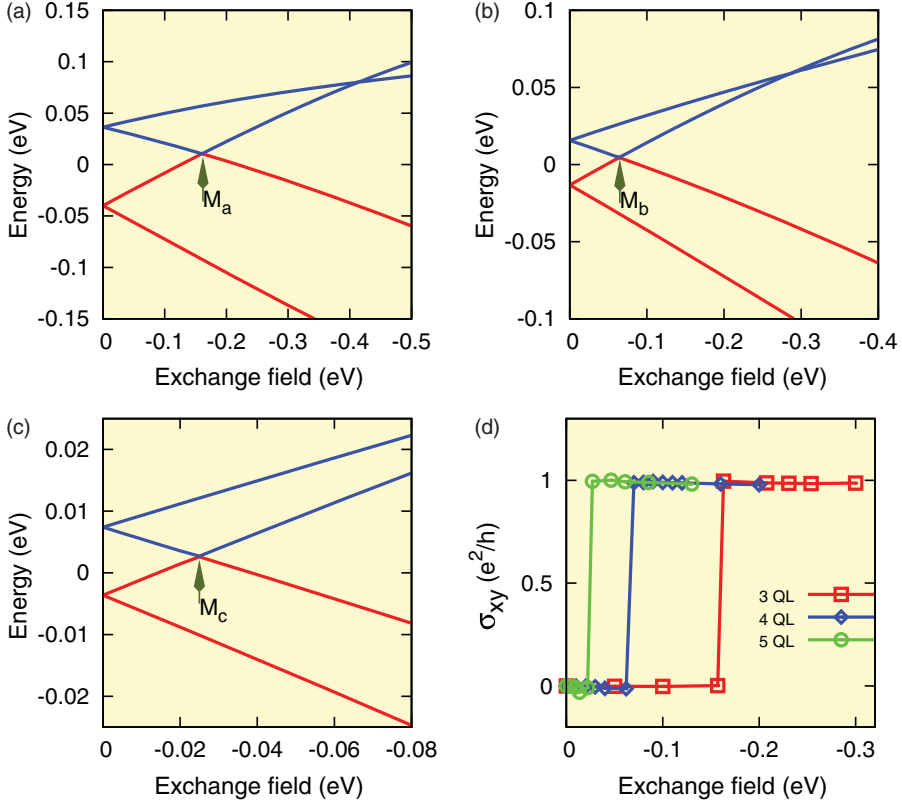
Figure 15. Quantized anomalous Hall (QAH) conductance. (a), (b) and (c) are the lowest subbands at the Γ point, plotted versus the exchange field, for Bi_2Se_3 films with thicknesses of 3, 4 and 5 quintuple layers (QL), respectively. The red lines denote the top occupied (bottom unoccupied) states. With an increase of the exchange field, a level crossing occurs (indicated by arrows), indicating a quantum phase transition to QAH state. (d) Calculated Hall conductance for 3, 4 and 5 QL Bi_2Se_3 films under ferromagnetic exchange field. As we expect, the Hall conductance is zero before the QAH transition, but $\sigma_{xy} = e^2/h$ afterward. Figure adapted from Ref. [139].

papers. To conclude this part, we finally point out that the QAHE can be induced even by in-plane magnetization, as discussed in Ref. [186].

4.3. *QAHE in thin film of Weyl semimetals*

As discussed in Section 2.3, Weyl semimetal is a new topological state of 3D quantum matter, unlike the 3D TIs. It can be characterized by Weyl nodes at the Fermi level in the bulk and Fermi arcs on surfaces. Around Weyl nodes, the low-energy physical behavior is like that of 3D two-component Weyl fermions [98] and can be described by the Hamiltonian $H = v_f \boldsymbol{\sigma} \cdot \mathbf{k}$, where v_f is the Fermi velocity, $\boldsymbol{\sigma}$ is the vector of the Pauli matrices, and $\mathbf{k}$ is the momentum as measured from the Weyl node. A Weyl fermion is half of a Dirac fermion, and Weyl fermions must appear in pairs. The Weyl Hamiltonian is robust against perturbations since it uses all three of the Pauli matrices. The perturbations only shift the position of the gap-closing point in momentum space and the Weyl nodes can be removed only when a pair of them meet in momentum space. To obtain a Weyl semimetal, either TRS or IS must be broken, and there have been several proposals
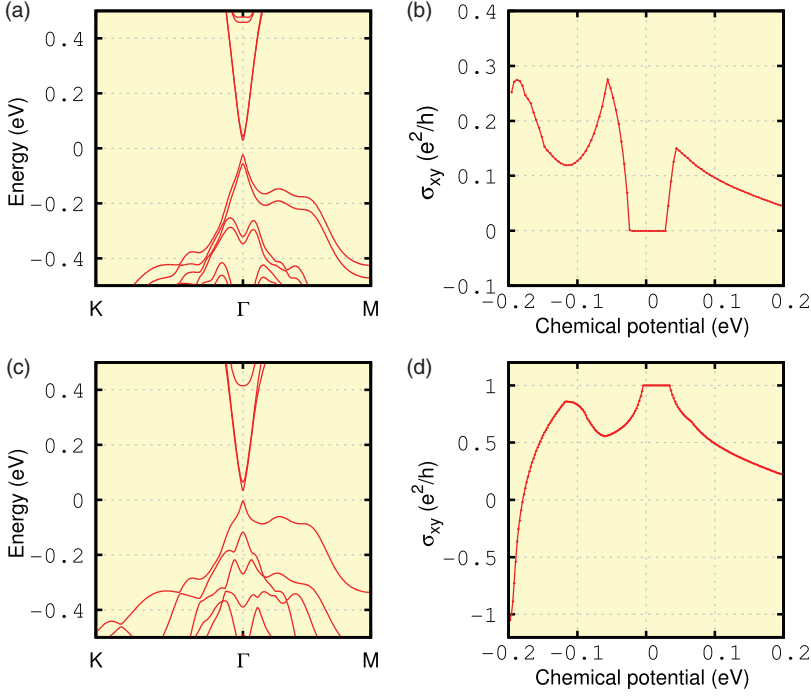
Figure 16. The subbands dispersion of 2D Bi_2Se_3 thin film are plotted for 3QL slab with exchange field (a) -0.05 eV, (c) -0.24 eV. The corresponding Hall conductance are plotted in the right panels as a function of the chemical potential. The zero chemical potential corresponds to the ideal stoichiometry case. The systems in (a) is normal insulator, which has zero platforms in Hall conductance when the chemical potential locates in the energy gap. The system in (c) is QAH insulators, in which the Hall conductance has quantized value e^2/h. Figure adapted from Ref. [139].

to realize this [55–57,105–109]. In particular, Weyl semimetals are closely related to the chiral anomaly [187], and the QAHE can naturally be expected in the quantum-well structure of Weyl semimetals with broken TRS. In this section, based on first-principles calculations, we will show that $HgCr_2Se_4$ is an FM Weyl semimetal with Fermi arcs on its surface, and its 2D thin film is a promising system for realizing the QAHE.

$HgCr_2Se_4$ is an FM spinel exhibiting large coupling effects between electronic and magnetic properties [189]. The spinel structure (space group $Fd\bar{3}m$) of $HgCr_2Se_4$ with Cr^{3+} octahedrally surrounded by Se ions, as shown in Figure 17, yields a half-filled t_{2g} shell and a spin number of $S = \frac{3}{2}$. First-principles calculations [56] confirm that FM solution is considerably (2.8 eV/f.u.) more stable than non-magnetic solution, and the calculated moment (6.0 μ_B/f.u.) is in good agreement with experiments [190, 191]. Without SOC, the electronic structures shown in Figure 18(a) and (b) suggest that it is nearly a "zero-gap half-metal". It is almost a half-metal because a gap exists only in the up-spin channel just above the Fermi level; it is nearly zero-gapped because of the band-touching around the Γ point just below the Fermi level in the down-spin channel. Quite large exchange splitting and strong octahedral crystal field together lead to a high spin state of Cr^{3+} ions in a $t_{2g}^{3\uparrow}e_g^{0\uparrow}t_{2g}^{0\downarrow}e_g^{0\downarrow}$ configuration with a gap between the $t_{2g}^{3\uparrow}$ and $e_g^{0\uparrow}$ manifolds as schematically shown in Figure 18(c). The Se-4p states (located from about -6 eV to 0 eV) are almost fully occupied and contribute to the top part of the valence band dominantly. The hybridization with Cr-3d states makes the Se-4p slightly spin-polarized with the opposite moment
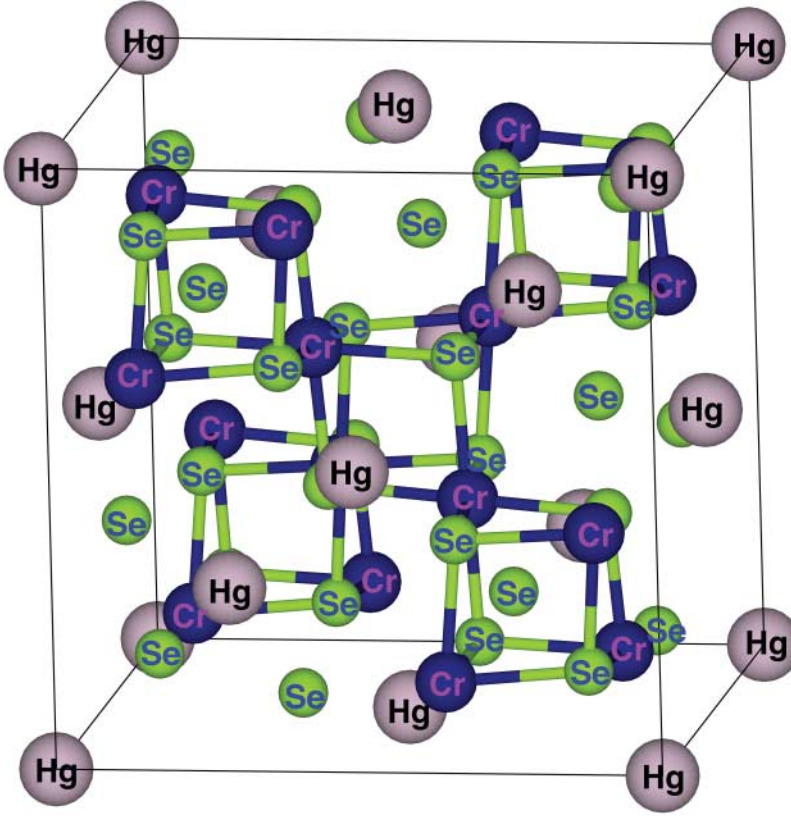
Figure 17. (Color online) Crystal structure of $HgCr_2Se_4$ spinels.

(about $-0.08\,\mu_B/Se$). The zero-gap behavior in the down spin channel is the most important characteristic here (Figure 18(b)), because it suggests the inverted band structure around Γ, similar to the case in HgSe or HgTe [192,193].

Taking the four (eight if spin is considered) low energy states at the Γ point as basis, which are identified as $|P_x\rangle$, $|P_y\rangle$, $|P_z\rangle$, and $|S\rangle$, the low energy physics becomes the same as in HgSe or HgTe, and the only difference is the presence of exchange splitting. The band inversion (see Figure 18(c); $|S, \downarrow\rangle$ is lower than $|P\rangle$) is due to the following two factors: (1) Hg-5d states around $-7.0\,eV$ below the Fermi level are very shallow and their hybridization with Se-4p states has pushed the anti-bonding Se-4p states quite high, similar to the situation in HgSe. (2) The hybridization between unoccupied Cr-3$d^{\downarrow}$ and Hg-6$s^{\downarrow}$ states will push the Hg-6$s^{\downarrow}$ state to lower in energy. These two key factors lead to band inversion, with the $|S, \downarrow\rangle$ lower than the $|P_{3/2}, \frac{3}{2}\rangle$ states by about 0.4 eV. This value is further enhanced to be 0.55 eV in the presence of SOC. Further, LDA $+ U$ calculations (with effective U around 3.0 eV [188, 194]) indicate that the band inversion remains even if the correlation effect is properly considered, unless U is unreasonably large (> 8.0 eV). Experimental observations of metallic behavior at low temperature for various samples [195–197] strongly support this conclusion of inverted band structure.

In the presence of SOC, the low energy eigenstates at Γ are given as $|\frac{3}{2}, \pm\frac{3}{2}\rangle, |\frac{3}{2}, \pm\frac{1}{2}\rangle, |\frac{1}{2}, \pm\frac{1}{2}\rangle$, and $|S, \pm\frac{1}{2}\rangle$, which can be constructed from the $|P\rangle$ and $|S\rangle$ states, similar to the case of HgSe and HgTe again. Considering the exchange splitting, the eight states at Γ are separated in energy. The $|\frac{3}{2}, \frac{3}{2}\rangle$ has the highest energy and the $|S, -\frac{1}{2}\rangle$ has the lowest. Due to the band inversion,
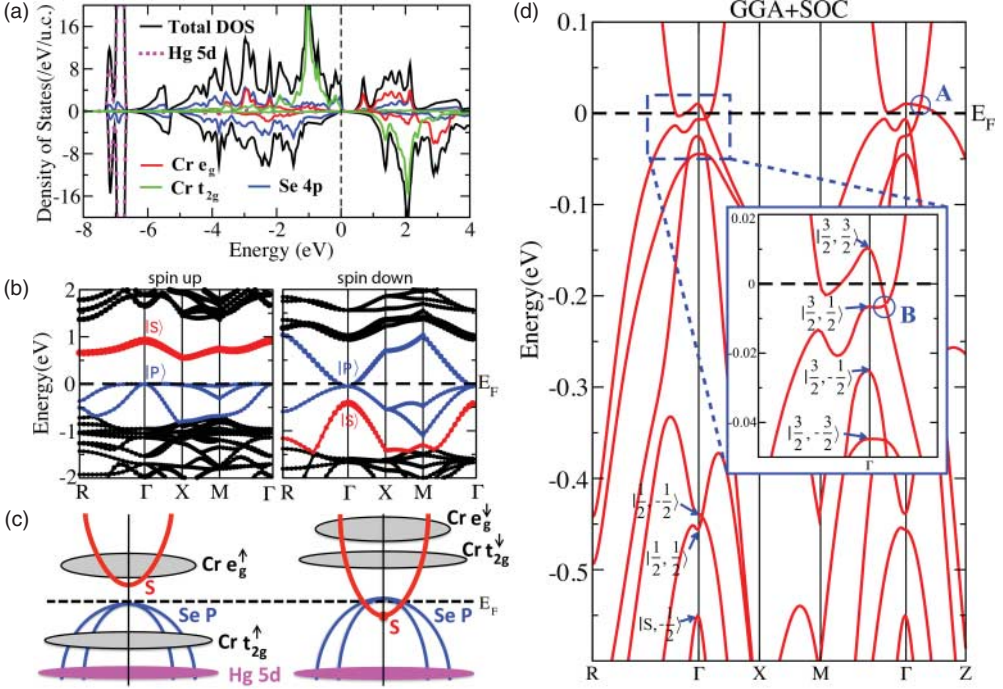
Figure 18. (Color online) Electronic structures of $HgCr_2Se_4$. (a) The total and projected partial density of states (DOS), (b) the band structures without SOC (showing the up and down spin parts separately), (c) schematic plot of band inversion, where the $|S, \downarrow\rangle$ state is lower than the $|P\rangle$ states and (d) the band structure including SOC (with majority spin along the (001) direction). The detail around Γ is enlarged in inset. Figure from Ref. [56].

several band-crossings are observed in the band structure. Among them, however, only two kinds of band-crossings (called A and B) are important for the states very close to the Fermi level, as shown in Figure 18(d). As schematically shown in Figure 19(a), the A-type crossing gives two points located at $k_z = \pm k_z^c$ along the $\Gamma - Z$ line. For the 2D planes with fixed-k_z ($k_z \neq 0, \pm k_z^c$), the in-plane band structures are all gapped in the sense that a curved Fermi level is defined. Therefore, the Chern number C for each k_z-fixed plane can be evaluated. It is found that the crossing A (i.e. $|k_z| = k_z^c$) is located at the phase boundary between the $C = 2$ and $C = 0$ planes, that is, $C = 0$ for the planes with $|k_z| > k_z^c$, while $C = 2$ for the planes with $|k_z| < k_z^c$ and $k_z \neq 0$. This change in Chern number indicates that these two crossing points (Weyl points) are topologically protected so long as the Weyl points remain separated in momentum space. The B-type crossing is a closed node line surrounding the Γ point in the $k_z = 0$ plane. It is just accidental and is due to the presence of crystal mirror symmetry with respect to the $k_z = 0$ plane. The B-type crossing is not as stable as the A-type crossing, in the sense that it can be avoided by changing of the crystal symmetry (such as the reorientation of magnetic momentum away from the z-direction).

Assuming that the magnetic momentum of the system is oriented to the z-axis, the low energy physics here can be caught by the following effective two-band model [56, 198]:

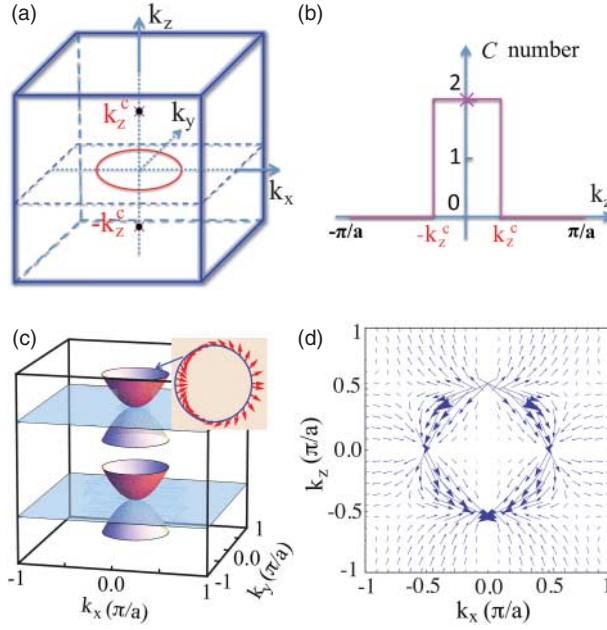$$H_{\text{eff}} = \begin{bmatrix} M & Dk_z k_-^2 \\ Dk_z k_+^2 & -M \end{bmatrix}, \tag{47}$$

Figure 19. (Color online) Weyl nodes and gauge flux in HgCr$_2$Se$_4$. (a) The band crossing points in $\vec{k}$-space, (b) Chern number as function of k_z, (c) schematic plot of the band dispersion around the Weyl nodes in the $k_z = \pm k_z^c$ plane, and the inset shows the chiral spin texture, (d) gauge flux evaluated as Berry curvature in the (k_x, k_z) plane. Figure adapted from Ref. [56].

in bases $|\frac{3}{2}, \frac{3}{2}\rangle$ and $|S, -\frac{1}{2}\rangle$. Here, $k_{\pm} = k_x \pm ik_y$ and $M = M_0 - \beta k^2$ is the mass term expanded up to the second order of k, with parameters $M_0 > 0$ and $\beta > 0$ to ensure band inversion. Since the two bases have opposite parity, the off-diagonal element has to be an odd function of k. In addition, $k_{\pm}^2$ has to appear to conserve the angular moment along the z-direction. Therefore, to the leading order, $k_z k_{\pm}^2$ is the only possible form for the off-diagonal element. The energy dispersion $E(k) = \pm\sqrt{M^2 + D^2 k_z^2 (k_x^2 + k_y^2)^2}$ has two gapless solutions: one is the degenerate points along $\Gamma - Z$ with $k_z = \pm k_z^c = \pm\sqrt{M_0/\beta}$; the other is a circle around the Γ point in the $k_z = 0$ plane determined by the equation $k_x^2 + k_y^2 = M_0/\beta$. They are exactly the band-crossings obtained from the first-principles calculations. Due to the presence of $k_{\pm}^2$ in the off-diagonal element [10], it is easy to find that Chern number C equals to 2 for the planes with $|k_z| < k_z^c$ and $k_z \neq 0$; otherwise, $C = 0$. The band dispersions near the Weyl nodes at $k_z = \pm k_z^c$ plane are thus quadratic rather than linear, with chiral in-plane spin texture, as shown in the inset of Figure 19(c). The two Weyl nodes located at $\pm k_z^c$ have opposite chirality and form a single pair of magnetic monopoles. The gauge flux starting from and ending at them in $\vec{k}$-space is shown in Figure 19(d). The band-crossing loop, that is, node line, in the $k_z = 0$ plane is not topologically unavoidable; however, its existence requires that the gauge flux in the $k_z = 0$ plane (except the loop itself) must be zero.

This Weyl semimetal (also called Chern semimetal) state realized in HgCr$_2$Se$_4$ will lead to novel physical consequences, which can be measured experimentally. First, each k_z-fixed plane with non-zero Chern number can be regarded as a 2D Chern insulator, and there must be chiral edge states for such planes if an edge is created. The number of edge states is two for the case of $C = 2$ (see Figure 20(a)) or zero for the case of $C = 0$. If the chemical potential is located within the gap, only the chiral edge states can contribute to the Fermi surface, which are isolated points
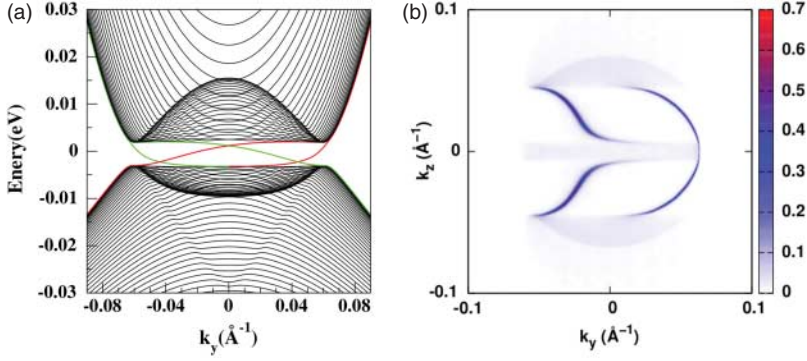
Figure 20. (Color online) Edge states and Fermi arcs of HgCr$_2$Se$_4$. (a) Calculated edge states for the plane with $k_z = 0.06\pi$. A ribbon with two edges is used, and there are two edge states for each edge since $C = 2$. The states located at different edges are indicated by different colors. (b) Calculated Fermi arcs for the side (k_y, k_z) surface. Figure adapted from Ref. [56].

for each plane with $C = 2$ but nonexistent for the plane with $C = 0$. Therefore, the trajectory of such points in the (k_x, k_z) surface or (k_y, k_z) surface form non-closed Fermi arcs, which can be measured by ARPES. As shown in Figure 20(b), the Fermi arcs connect the Weyl points at $k_z = \pm k_z^c$, and are interrupted by the $k_z = 0$ plane. This is very different from conventional metals, in which the Fermi surfaces must be either closed or interrupted by the BZ boundary. The possible Fermi arcs have also been discussed for pyrochlore iridates [55].

The QAHE, on the other hand, is a unique physical consequence characteristic of the Chern semimetal nature of HgCr$_2$Se$_4$ by considering its quantum-well structure. For 2D Chern insulators, the transverse Hall conductance should be quantized as $\sigma_{xy} = C(e^2/h)$, where C is the Chern number. Considering the k_z-fixed planes, the Chern number C is non-zero for limited regions of k_z, and this is due to the band inversion around Γ as discussed above. In the quantum-well structure, however, those low energy states around Γ should be further quantized into subbands (labeled as $|H_n\rangle$ and $|E_n\rangle$ for hole and electron subbands, respectively), whose energy levels change as functions of film thickness. As shown in Figure 21(a), when the film is thin enough, the band inversion in the bulk band structure will be removed entirely by the finite size effect. With the increment of film thickness, the finite size effect gets weaker and the band inversion among these subbands is restored subsequently, which leads to jumps in the Chern number or the Hall coefficient σ_{xy} [138]. As shown in Figure 21(b), if the film is thinner than 21 Å, σ_{xy} is zero; once the film thickness is larger than the critical thickness, we find subsequent jumps of σ_{xy} in units of $2e^2/h$. In fact, the AHE has been observed for the bulk samples of HgCr$_2$Se$_4$ [199], and more recently, n-type HgCr$_2$Se$_4$ was confirmed to be a single s-band half metal [200]. This is in sharp contrast to pyrochlore iridates, in which the AHE should be vanishing due to the AF ordering [55].

This proposal of realizing QAHE in thin films of FM Weyl semimetals has some advantages compared to the previous one for the magnetic-ion-doped TIs thin films. First, it is based on the stoichiometric material and avoids the difficulty of doping, which is in general good for raising the mobility of the sample. Second, FM ordering has been established in bulk HgCr$_2$Se$_4$ at temperature as high as 100 K [201], which makes QAHE plausible at high temperatures. Third, the proposed system can reach a QAHE with a high Chern number (i.e. Chern number 2 or even higher) and higher plateau [185]. Therefore, further experimental efforts in realizing QAHE in thin film of HgCr$_2$Se$_4$ are much anticipated and very promising.
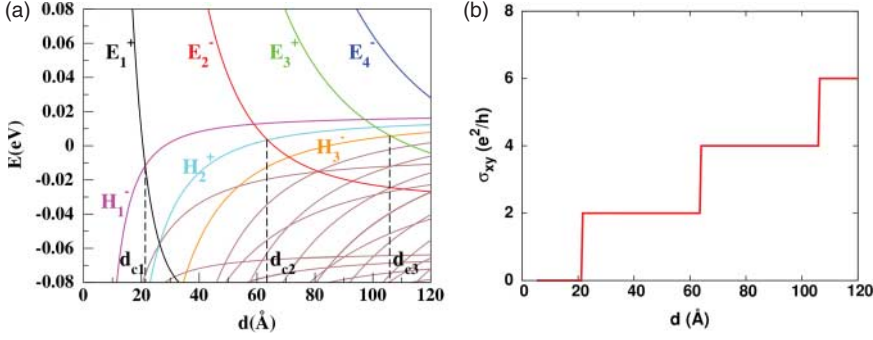
Figure 21. (Color online) QAHE in HgCr$_2$Se$_4$ thin film. (a) The subband energy levels at Γ point as a function of film thickness. (b) The Hall conductance as function of film thickness. Figure adapted from Ref. [56].

4.4. *QAHE on honeycomb lattice*

Since honeycomb lattice plays a unique role in exploiting non-trivial topology [37, 39], it is instructive to take a look at the geometric property of its electronic band structures, paying attention to possible configurations of Berry curvatures. There are two sites in a unit cell of honeycomb lattice, and with the nearest neighbor hoppings, the valance band and conduction band touch linearly and form the Dirac cones at the six corners of the BZ. Taking into account the C$_3$ symmetry, one can consider the two inequivalent Dirac cones at $K, K' = (4\pi/3d)(\pm 1, 0)$ with lattice constant d, which are called valleys and denoted by $\tau_z = \pm 1$. The Hamiltonian for low-energy physics is given by

$$H_{\tau_z} = \begin{pmatrix} 0 & \tau_z k_x + ik_y \\ \tau_z k_x - ik_y & 0 \end{pmatrix}, \tag{48}$$

where the momenta are measured from the valleys. This is the 2D version of Equation (35) with the basis of the 2×2 matrix referring to AB sublattices of honeycomb lattice. Therefore, the electrons are chiral at the valleys, which makes the honeycomb lattice special in generating non-trivial topology.

Let us introduce a staggered field U on the AB sublattices, and rewrite the Hamiltonian (48) as

$$H_{\tau_z} = \tau_z k_x \sigma_x + k_y \sigma_y + U\sigma_z, \tag{49}$$

with Pauli matrices $\sigma_{x,y,z}$. The dispersion is given by $E = \pm\sqrt{U^2 + k^2}$, with a gap opened by the staggered potential U. For this insulating state, one can define a fictitious magnetic field $\mathbf{d} = (\tau_z k_x, k_y, U)$ in Hamiltonian (49) on the pseudo-spin σ, which exhibits a topological structure around each valley known as *meron*. Its topological property can be characterized by counting the winding number of the *meron* structure (i.e. the integration of Berry curvature over a patch covering the region of the *meron*, or in other words, the Berry flux passing through the 2D BZ in the vicinity of the each valley, divided by 2π) as [15]

$$c = \frac{1}{4\pi} \int dk_x \, dk_y \mathbf{n} \cdot (\partial_{k_x}\mathbf{n} \times \partial_{k_y}\mathbf{n}) = \frac{\text{sgn}(\tau_z U)}{2}, \tag{50}$$

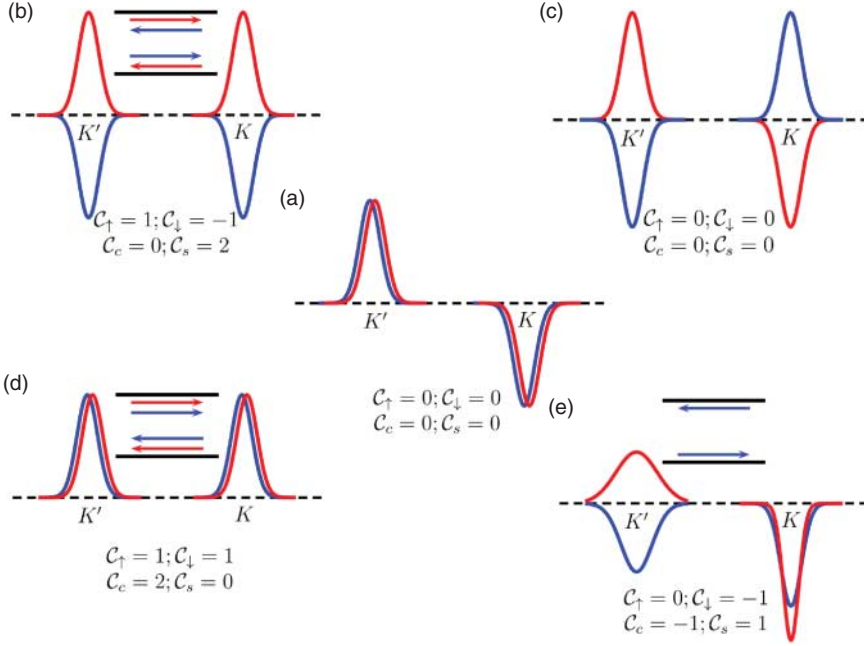with $\mathbf{n} = \mathbf{d}/|\mathbf{d}|$.

Figure 22. (Colour online) Possible configurations of Berry curvatures on honeycomb lattice: (a) CDW, (b) QSHE, (c) SDW, (d) QAHE, (e) spin-resolved QAHE characterized by simultaneous non-zero charge and spin Chern numbers [161]. Red/blue for spin up/down channels. Edge currents are displayed in insets in (b), (d) and (e) for topologically non-trivial states. $C_\uparrow$ and $C_\downarrow$ are Chern numbers in spin-up and -down channels, and $C_{c/s} = C_\uparrow \pm C_\downarrow$ are charge and spin Chern numbers respectively. Figure from Ref. [161].

4.4.1. *Combinations of Berry curvatures on honeycomb lattice*

For a staggered electric potential $H_{elec} = V\sigma_z$, where V is a constant, the signs of the Berry curvature around two valleys are opposite, as seen in Equation (50), which gives a topologically trivial insulator called a charge density wave (CDW) as displayed in Figure 22(a). On the other hand, if a staggered magnetic flux is introduced to spinless electrons on honeycomb lattice [37], the staggered potential term can be written as $H_{sf} = t\tau_z\sigma_z$ (around two valleys), which depends on the valley variable in contrast to the staggered electric field. Since the topological charges are now given by $c = \text{sgn}(t)/2$, the Berry curvatures around two valleys are aligned, which characterizes a QAHE state (see Figure 22(d) where spin is included in addition to the Haldane model). This staggered field can be generated by shining circularly polarized light on graphene, as revealed by the Floquet theorem [202]. Taking into account the true spin degree of freedom, the SOC term of $H_{so} = \lambda s_z\tau_z\sigma_z$ can also be included as suggested in the Kane–Mele model [39], and the SOC term again takes the form of a staggered field, and depends additionally on spin variable s_z. In this case, the Berry curvatures around two valleys in the spin-up and down channels are aligned individually and oppositely for different spin channels, resulting in the configuration in Figure 22(b), which characterizes the QSHE state [39]. Antiferromagnetic (AFM) exchange field $H_{afm} = Ms_z\sigma_z$ is another possible staggered field, which reverses the sign of the Berry curvature in the spin-down channel for both valleys (compared to the CDW state shown in Figure 22(a), assuming the same sign for M and V), resulting in the configuration in Figure 22(c) associated with a topologically trivial state called spin density wave (SDW). Coupling of the valley degree of freedom to the AFM order in manganese chalcogenophosphates was recently discussed [203].

We observe that the configuration of Berry curvature given in Figure 22(e) corresponds to an additional topological state [161] (see also [204]), by flipping the sign of the Berry curvature around K′ valley in the spin up channel from the CDW state. In contrast to QSHE with a non-zero spin Chern number $C_s = C_\uparrow - C_\downarrow$ [205, 206], which is equivalent to the Z_2 index, and QAHE with a non-zero charge Chern number $C_c = C_\uparrow + C_\downarrow$, this state is characterized by simultaneous non-zero charge and spin Chern numbers. We may call the state in Figure 22(e) *spin-resolved* QAHE, as opposed to other QAHE states where spin information is lost.

Note that configurations of the Berry curvature in Figure 22(a)–(d) are converted to each other by flipping the sign of two Berry curvatures either at the same valley or with the same spin. This reflects certain symmetries in these states: TRS in Figure 22(a), TRS and spacial IS in Figure 22(b), IS in Figure 22(d), and the combination of TRS and IS in Figure 22(c) (although both symmetries are broken individually). All symmetries are broken in the state portrayed in Figure 22(e). One can realize the state in Figure 22(e) by using three fields simultaneously – SOC, AFM exchange field and the staggered electric potential – which makes fine control on electronic states possible:

$$H = \tau_z k_x \sigma_x + k_y \sigma_y + (\lambda \tau_z s_z + M s_z + V)\sigma_z. \tag{51}$$

We reveal behaviors of the system upon tuning the electric potential V with the M field and λ fixed. For $\lambda > M > 0$, the system takes a QSHE state at $V = 0$ (see Figure 22(b)). The degeneracy in band dispersions is protected by the product of TRS and IS, a pseudo-TRS. The gaps are $2(\lambda + M)$ and $2(\lambda - M)$ at the two valleys. As illustrated in Figure 23(a), when V increases from zero, the gap in the spin-up channel shrinks, closes at $V = \lambda - M$ and reopens at the K′ valley; this quantum phase transition makes the spin-up channel topologically trivial, while the spin-down channel remains topologically non-trivial. This QAHE state is possible since the electric field V breaks the IS, therefore breaking the pseudo-TRS. When V increases further, the gap in the spin-down channel shrinks, closes at $V = \lambda + M$ and reopens at K valley (not shown explicitly). This second quantum phase transition makes the spin-down channel trivial as well, and the system transforms into the topologically trivial CDW state.

For $M > \lambda > 0$, the system is in a topologically trivial SDW state at $V = 0$ (see Figure 22(c)). As seen in Figure 23(b), when V increases, the gap in the spin-down channel shrinks, closes at $V = M - \lambda$ and reopens at the K′ valley, which brings the spin-down channel into a topological state. When V increases further, the gap in the spin-down channel shrinks, closes at $V = \lambda + M$ and reopens at K valley, which drives the system to the topologically trivial CDW state. It is clear that the spin-resolved QAHE state is realized when the sizes (absolute values) of the three fields satisfy the condition of being able to form a triangle, which indicates clearly that the interplay among the three fields is crucial. In this case, the sign of the total staggered field $U = \lambda \tau_z s_z + M s_z + V$ at one valley in one spin channel is opposite to the other three cases, which yields the configuration of Berry curvature in Figure 22(e).

The above discussions for the possible topological phases in honeycomb lattice by fine tuning of parameters can be generally applied to many different material systems, such as graphene [144–152], silicene [153–155] and perovskite bilayers [156–162]. Valleys, as an additional degree of freedom of electrons on honeycomb lattice, have attracted considerable interest, for which valleytronics has been coined as a new concept [207, 208]. We will leave the discussions of topological states in graphene and silicene to readers, and in the following, we concentrate on the perovskite bilayer systems, which are interesting and promising due to the presence of transition metal elements.

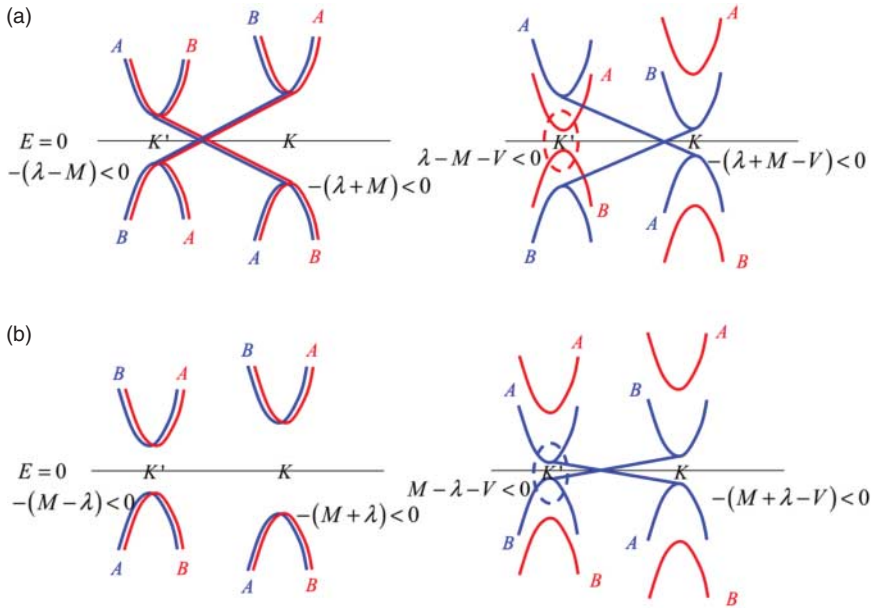Figure 23. (Color online) Band configurations for $V = 0$ (left) and $|\lambda - M| < V < \lambda + M$ (presuming $\lambda, M > 0$) for (a) $M < \lambda$ and (b) $M > \lambda$. Red (blue) curves are for spin-up (-down) channels, and A and B for the sublattices. The straight lines are for topological edge states in ribbon systems, and the dashed ellipses indicate the valleys where band inversions occur upon increasing V.

4.4.2. *Honeycomb lattices by perovskite structures*

Material realization of TIs has been mainly focused on narrow band-gap semiconductors with heavy elements such as Hg or Bi, in which the electronic properties are dominated by s and p orbitals. Recently, a completely different materials class – heterostructures of TMOs involving d electrons – was proposed [156]. Artificial heterostructures of TMOs are becoming available owing to recent developments in the fields of oxide superlattices and oxide electronics [209–211]. In particular, layered structures of TMOs can now be prepared with atomic precision [212–214], thus offering a high degree of control over important material properties, such as lattice constant, carrier concentration, SOC and correlation strength. These advantages can be readily exploited in the design of TIs. TMOs constitute a wide class of compounds that exhibit a variety of intriguing properties and electronic states associated with the electron–electron interactions, encompassing superconductivity, magnetism, ferroelectricity and Mott insulators. Combined with the TI phase, TMO heterostructures provide a very promising platform to explore various topological effects. In particular, possible 2D TIs have been discussed for bilayers of perovskite-type TMOs grown along the [111] crystallographic axis [156].

Perovskite compounds [215] are very common and have the chemical formula ABO_3, where A stands for alkaline or rare earth metal, B stands for transition metal (TM) and O stands for oxygen. The perovskite structure is composed of two simple cubic lattices of A and B elements penetrating each other, with B atoms caged by oxygen octahedra as shown in Figure 24(a) and (b). The A element usually contributes charge to the system, with remained electronic states lying deep in the valence band. The TM element B contributes to bands close to the Fermi level and produces rich physics. In an ideal perovskite structure, the octahedral crystalline field splits the TM d orbitals into twofold degenerate $e_g (d_{3z^2-r^2}, d_{x^2-y^2})$ and threefold degenerate $t_{2g} (d_{yz}, d_{zx}, d_{xy})$ levels, well separated by so-called 10Dq gap of $\sim 3\,\text{eV}$. Such a lattice geometry usually does not
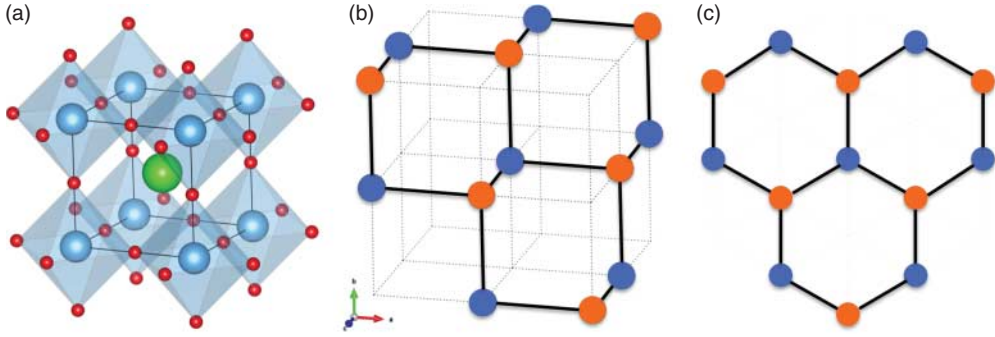
Figure 24. (Color online) Perovskite structure as a buckled honeycomb lattice grown along the [111] crystal direction. Adapted from Ref. [156].

support Dirac points. Instead, if bilayers of the perovskite structure grown along the [111] direction are considered [156], as shown in Figure 24(c), the TM ions will be located on an analogous honeycomb lattice consisting of two trigonal sublattices formed by two different layers. This lattice geometry has three consequences: first, it is well known from the study of graphene that electrons hopping on a honeycomb lattice generally give rise to Dirac points in the band structure; second, a layer potential difference can easily be created by applying a perpendicular electric field or by sandwiching the bilayer between two different substrates, which allows experimental control of the band topology; and, third, the honeycomb lattice further reduces the symmetry of the crystalline field from octahedral (O_h) to trigonal (C_{3v}), and introduces additional level splitting of the d orbitals. The key idea is to start with a band structure that possesses Dirac points in the BZ without the SOC, and then examine whether an energy gap can be opened at those points with SOC turned on. If an energy gap does open, combined with proper filling, the resulting state could be a TI. Based on tight-banding models and first-principles calculations, it has been demonstrated that both t_{2g} and e_g orbitals can generate topologically non-trivial states. In particular, it is predicted that $LaAuO_3$ bilayers sandwiched by two $LaAlO_3$ substrates have a topologically non-trivial energy gap of about 0.15 eV [156], which is sufficiently large to realize the QSHE at room temperature.

Various possible topological phases have been explored for the perovskite bilayer systems [158–160]. Except the QSH phase, the QAH phase can be readily expected by breaking the TRS, as discussed before. Thanks to the presence of transition metal elements, different magnetic orderings may be realized in such systems with no need of additional doping. If the strong electron interaction among the d electrons is properly taken into account, we can expect more interesting physics [157] for such honeycomb lattices formed by bilayer perovskite structures. In a recent study, it was predicted [162] by the LDA + Gutzwiller method [216], that (111) $LaCoO_3$ bilayer grown on a $SrTiO_3$ substrate may support the QAHE. Here, the interplay between SOC and Coulomb interaction stabilizes a robust FM insulating state, which is a Chern insulator with non-zero Chern number. It was pointed out in this study that the SOC splitting (which is usually not strong for $3d$ elements) can be greatly enhanced by the strong Coulomb interaction. It has also been revealed that the possibility of some intriguing phenomena, such as fractional QHE, associated with the nearly flat topologically non-trivial bands can be found in e_g systems [156].

The perovskite bilayer structure can also be used to generate the spin-resolved QAHE state. Choosing elements adequately, one can fill t_{2g} orbitals while leaving e_g orbitals partially occupied, and the 10Dq gap (crystal-field splitting) will prevent t_{2g} orbitals from participating in the low-energy physics. For an atomic sheet of buckled honeycomb lattice, the two B atoms in a unit cell
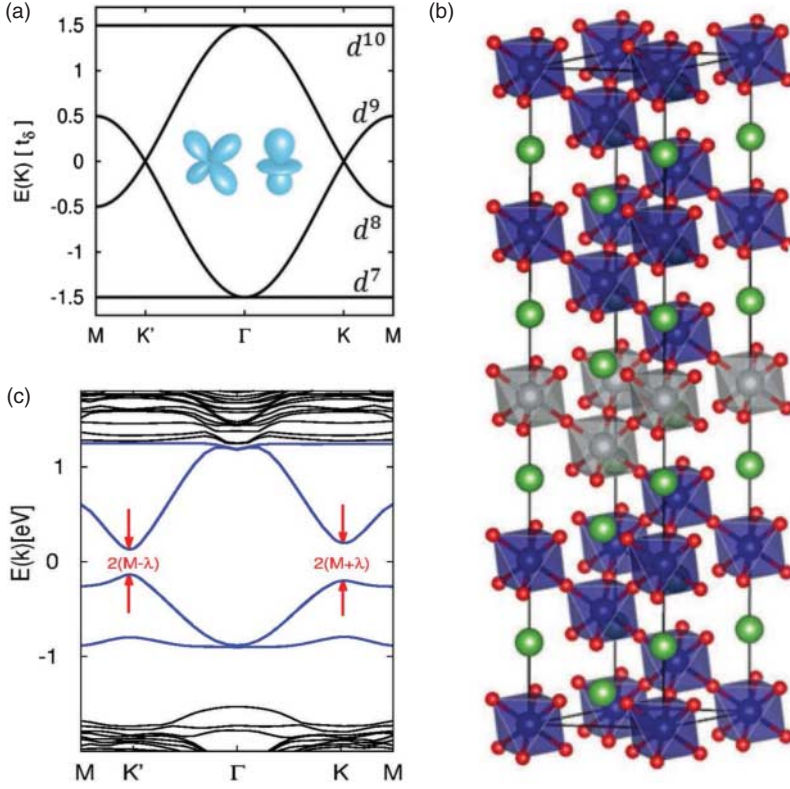
Figure 25. (Color online) (a) Noninteracting tight-binding band structure of the e_g system, where each band is labeled with the corresponding electron configuration if it is occupied. Inset shows the two e_g orbits. (b) Schematic of a unit cell $(La_2Au_2O_6)/(La_2Cr_2O_6)_5$. (c) Band structures based on first-principles calculation. Figures adapted from Ref. [161].

contribute four e_g orbits in total. The band structure for these four orbitals is shown in Figure 25(a) based on the tight-binding Hamiltonian with the hopping integrals given by the Slater–Koster formula [161]. Two dispersive bands are contributed from $d^{8,9}(t_{2g}^6 e_g^{2,3})$, which cross each other linearly and form the desired Dirac cones at K and K′ due to the honeycomb structure.

A candidate perovskite material that generates d^8 electrons is $LaAuO_3$, wherein Au takes the $+3$ charge state. In order to generate an AFM exchange field, one can sandwich the atomic sheet of buckled honeycomb lattice with d^8 electrons by two thick layers of insulating perovskite material grown in the [111] direction. A candidate material for the substrates is $LaCrO_3$, a Mott insulator with G-type AFM order, which induces opposite exchange fields on the two sublattices in the buckled honeycomb lattice. The large Mott gap of $LaCrO_3$ prevents undesired influences on the Dirac behavior of $d^{8,9}$ electrons except for the exchange field. Since La is chosen as the common A-site cation, the lattice mismatch between substrate and the targeting atomic sheet is minimal. First-principles calculations were performed in terms of VASP [217] code with a super-cell of $(La_2Au_2O_6)/(La_2Cr_2O_6)_5$ shown in Figure 25(b). As displayed in Figure 25(c), the black bands are mainly from the substrate $LaCrO_3$ with a large gap of 3eV, characteristic of a Mott insulator. The bands inside the Mott gap are contributed by the e_g electrons of Au^{+3} ions. There are two almost flat bands and two dispersive bands, as captured by the tight-binding Hamiltonian (see Figure 25(a)). Gaps are opened at K and K′ since the Dirac electrons now feel the

AFM exchange field from LaCrO$_3$ and SOC. The two gaps are different in magnitude and given explicitly by $2(M + \lambda)$ and $2(M - \lambda)$ respectively (see Figure 23(b) for $V = 0$), from which one obtains $M = 0.166$ eV and $\lambda = 32.91$ meV for this system. The SOC effect is quite large in the present case, due to the orbital hybridization between the two e$_g$ orbits in the new frame along the [111] crystalline direction. All the states in Figure 25(b) are doubly degenerate due to the product symmetry of TRS and IS. If a gate electric field is applied along the [111] direction, the Au ions in the buckled honeycomb lattice should experience staggered electric potentials. By increasing the electric potential, the system should be driven into the spin-resolved QAHE, as summarized in Figure 23(b). The required electric field is on the order of 0.1 V/Å. As the maximal non-trivial gap is given by $2\lambda \simeq 600$ K, this topological state in the LaCrO$_3$/LaAuO$_3$ superlattice may remain stable even around room temperature. Interaction-driven QAHE phase can also be found in the SrTiO$_3$/LaCoO$_3$/SrTiO$_3$ quantum well as discussed in detail in Ref. [218].

It is worth noticing that, although the edge current of the spin-resolved QAHE state is fully spin polarized, the total system does not show any net spin magnetization, since the numbers of occupied states in the two spin channels of the buckled honeycomb lattice are the same, and the substrates are AFM insulators. The spin polarization of the edge current can be reversed by the gating electric field, which is superior for future spintronic applications.

5. Discussions and future prospects

At this stage, the quantum Hall trio (namely, QHE, QSHE, and QAHE) has been completed. [219] Looking back to history, it was a long-term exploration towards the final realization of QAHE. First, the understanding of the Berry phase mechanism of the intrinsic AHE was an important step, because it established the link between the AHE and the topological properties of electronic structure. Second, the rapid progresses in the field of TIs strongly stimulated the studies on QAHE. Third, first-principles calculations play important roles in predicting and selecting candidate materials and systems. This also establishes a nice example that theoretical and computational studies can in some cases really go ahead of and greatly help experiments. Finally, the well-controlled growth and magnetic doping of high quality samples and transport measurement at ultra-low temperature make the experiments possible, thanks to the progress of modern experimental technology. All these aspects combined together finally gave birth to the QAHE.

On the other hand, the successful realization of QAHE is not the end of the story. Instead, we believe that this success will inspire more extensive research in this field. Two issues become immediately important for future studies:

(1) How can we realize a higher plateau with a Chern number larger than 1?
(2) How can we increase the temperature range of QAHE (it is now observed only in the tens mK range)?

From the theoretical point of view, the QAH states with higher Chern numbers can be derived in a way similar to that described in Section 4.2 if the magnetic exchange field can be made larger [185, 220]. Larger exchange field will, in general, generate more band inversions among subbands and thus induce higher Chern numbers. While in practice this idea is very difficult, because the system with large enough exchange field will usually become completely metallic before the second plateau state appears. An alternative way to realize the QAHE with a high Chern number is to start with materials that have multiple band inversion regimes in the BZ. An interesting proposal is the magnetically doped topological crystalline insulators (Pb,Sn)(Te,Se), which have band inversions around all the four T points in the BZ [221]. Similar to the situation in Bi$_2$Se$_3$ family compounds, here the band inversions appear between p-bands with opposite

parity, and the Van Vleck mechanism also works, which can lead to a stable FM semiconductor ground state after magnetic doping. Therefore, for thin films of this system with proper thickness, it is possible to realize the QAHE with a Chern number between ± 4 [221].

The experimental observation of QAHE in thin films of Cr-doped $(Bi, Sb)_2Te_3$ requires an extremely low temperature (approximately 30 mK), which hinders realistic applications [141]. This is presumably due to the small band gap opened by exchange coupling and the low carrier mobility of 760 cm^2/Vs in the sample. Therefore, searching for realistic systems with non-trivial band structures, strong exchange coupling and high sample quality is essential to realize the QAH effect at a higher temperature. Liu et al. proposed to realize QAHE in a more conventional diluted magnetic semiconductor, the Mn-doped InAs/GaSb type II quantum well, which has a higher carrier mobility of 6000 cm^2/Vs and a predicted FM transition temperature of 30 K [140]. Other theoretical purposes for high-temperature QAHE systems are based on the stanene and germanene materials. These two materials are recently predicted to be QSHI with a large insulating gap [222]. The QSH phase appears due to an energy inversion between the Sn(Ge)-s and Sn(Ge)-p_{xy} bands, which have opposite parities and lead to a sizable bulk gap of 0.3 eV. In Ref. [223], Wu et al. proposed that half-passivated stanene and germanene become Chern insulators with energy gaps of 0.34 and 0.06 eV, respectively. The proposed FM order in this system comes from the unpassivated sublattice that exhibits dangling bonds. Strong coupling between the spin-polarized dangling bond states and the inherent inverted bands opens a considerable energy gap. The estimated mean-field Curie temperature is as high as 243 K and 509 K for Sn and Ge lattices, respectively. In spite of those theoretical proposals, it is still challenging to reach QAHE at higher temperature experimentally. To achieve this goal, searching for new host materials with simpler electronic structure, easier synthesis and less extreme growth conditions is necessary.

The QAHE and Chern insulators can also be helpful for the possible realization of Majorana fermions (MF) [224], which are exotic charge-neutral particles that are their own antiparticles. In recent years, MF have been intensively studied in the field of condensed matter physics due to their importance to both fundamental science and potential applications in topological quantum computing. Theoretically, MF have been predicted to exist in several spin–orbit-coupled superconducting systems, such as superconductor–topological insulator interfaces [73], semiconductor–superconductor heterostructures [225–227], nanotube–superconductor devices [228], and half-metal–superconductor interface [79]. Experimentally, great efforts have been made in this direction, and some important evidence has been successfully observed [229, 230]. On the other hand, in Ref. [75], a different way to generate MF is proposed in a hybrid device that consists of a QH or QAH insulator layer in proximity to a fully gapped superconductor layer on top. In this case, it can generally be expected that a chiral topological superconductor phase with an odd number of chiral Majorana edge modes will exist. Compared to the QH system in the proposal, the QAH system has the advantage that it does not require external magnetic field, which makes the realization of superconductor proximity effect much easier. The Majorana zero mode in a model combined with a QAHE system and an s-wave superconductor is also discussed in Ref. [231].

In summary of the present paper, we have reviewed the history of achieving the QAHE. In 1880, E.H. Hall found the AHE, 17 years before the discovery of the electron. In the following 100 years, however, the developments in this field, particularly for the understanding of AHE, were slow. There were debates on whether intrinsic or extrinsic contribution is dominant. The breakthrough was the discovery of the QHE in 1980, from which people learned that there exist new quantum states characterized by the topology of electronic structures. The concept of the Berry phase, established in 1984 by M.V. Berry, is the solid base of various topological quantum states. We have reviewed how the Berry phase is related to the AHE, QHE, QAHE, QSHE, 3D TIs and 3D Weyl/Dirac semimetals. Such a unified view has led us to a unified scheme, called the

Wilson loop method, in calculating various topological invariants, including the Chern number, Z_2 index, mirror Chern number and Fermi surface Chern number. This method is efficient in rigorously determining the topology of real materials. However, in practical calculations, the band inversion mechanism is much more intuitive and efficient in predicting, selecting and designing candidate topological materials. Finally, we have listed the necessary and complete four conditions for realizing the QAHE. Following these conditions, we have demonstrated several ways to achieve the QAHE. The theoretical prediction for the QAHE in magnetic-ion Cr- and Fe-doped Bi_2Se_3 and Bi_2Te_3 family TIs has been recently confirmed by experimental observation. The key difference of this scheme from other proposals is the FM ordering of Cr and Fe dopants induced by the van Vleck paramagnetism, which makes this proposal successful.

Disclosure statement

No potential conflict of interest was reported by the authors.

References

 [1] E.H. Hall, Philos. Mag. Ser. 5 10(63) (1880), pp. 301–328.
 [2] E.H. Hall, Philos. Mag. Ser. 5 12(74) (1881), pp. 157–172.
 [3] A. Kundt, Wied. Ann. 49 (1893), p. 257.
 [4] E.M. Pugh, Phys. Rev. 36(9) (1930), pp. 1503–1511.
 [5] E.M. Pugh and T.W. Lippert, Phys. Rev. 42(5) (1932), pp. 709–713.
 [6] G. Bergmann, Phys. Rev. Lett. 41(4) (1978), pp. 264–267.
 [7] H. Ohno, H. Munekata, T. Penney, S. von Molnár, and L.L. Chang, Phys. Rev. Lett. 68(17) (1992), pp. 2664–2667.
 [8] H. Ohno, A. Shen, F. Matsukura, A. Oiwa, A. Endo, S. Katsumoto, and Y. Iye, Appl. Phys. Lett. 69(3) (1996), pp. 363–365.
 [9] T. Jungwirth, Q. Niu, and A. H. MacDonald, Phys. Rev. Lett. 88(20) (2002), p. 207208.
[10] M. Onoda and N. Nagaosa, J. Phys. Soc. J. 71(1) (2002), pp. 19–22.
[11] Z. Fang, N. Nagaosa, K.S. Takahashi, A. Asamitsu, R. Mathieu, T. Ogasawara, H. Yamada, M. Kawasaki, Y. Tokura, and K. Terakura, Science 302(5642) (2003), pp. 92–95.
[12] Y.G. Yao, L. Kleinman, A.H. MacDonald, J. Sinova, T. Jungwirth, D.-S. Wang, E.G. Wang, and Q. Niu, Phys. Rev. Lett. 92(3) (2004), p. 037204.
[13] J. Sinova, T. Jungwirth, and J. Cerne, Int. J. Mod. Phys. B 18(08) (2004), pp. 1083–1118.
[14] N. Nagaosa, J. Sinova, S. Onoda, A.H. MacDonald, and N.P. Ong, Rev. Mod. Phys. 82(2) (2010), pp. 1539–1592.
[15] D. Xiao, M.-C. Chang, and Q. Niu, Rev. Mod. Phys. 82(3) (2010), pp. 1959–2007.
[16] M.V. Berry, Pro. R. Soc. Lond. A 392(1802) (1984), pp. 45–57.
[17] R. Karplus and J.M. Luttinger, Phys. Rev. 95(5) (1954), pp. 1154–1160.
[18] J. Smit, Physica 21(6–10) (1955), pp. 877–887.
[19] J. Smit, Physica 24(1–5) (1958), pp. 39–51.
[20] L. Berger, Phys. Rev. B 2(11) (1970), pp. 4559–4566.

[21] S. Onoda, N. Sugimoto, and N. Nagaosa, Phys. Rev. Lett. 97 (2006), p. 126602.

[22] T. Miyasato, N. Abe, T. Fujii, A. Asamitsu, S. Onoda, Y. Onose, N. Nagaosa, and Y. Tokura, Phys. Rev. Lett. 99 (2007), p. 086602.

[23] Y. Tian, L. Ye, and X.F. Jin, Phys. Rev. Lett. 103 (2009), p. 087206.

[24] L. Ye, Y. Tian, X.F. Jin, and D. Xiao, Phys. Rev. B 85 (2012), p. 220403.

[25] C.G. Zeng, Y.G. Yao, Q. Niu, and H.H. Weitering, Phys. Rev. Lett. 96 (2006), p. 037204.

[26] K.V. Klitzing, G. Dorda, and M. Pepper, Phys. Rev. Lett. 45(6) (1980), pp. 494–497.

[27] F. Wilczek and A. Shapere, Geometric Phases in Physics, Vol. 5, World Scientific, 1989.

[28] D.J. Thouless, M. Kohmoto, M.P. Nightingale, and M. den Nijs, Phys. Rev. Lett. 49(6) (1982), pp. 405–408.

[29] R.E. Prange and S.M. Girvin, The Quantum Hall Effect, Springer, 1987.

[30] M. Stone, Quantum Hall Effect, World Scientific, 1992.

[31] Y.G. Yao, Y.C. Liang, D. Xiao, Q. Niu, S.Q. Shen, X. Dai, and Z. Fang, Phys. Rev. B 75 (2007), p. 020401.

[32] W.-L. Lee, S. Watauchi, V.L. Miller, R.J. Cava, and N.P. Ong, Science 303(5664) (2004), pp. 1647–1649.

[33] R.D. King-Smith and D. Vanderbilt, Phys. Rev. B 47 (1993), pp. 1651–1654.

[34] R. Resta, Rev. Mod. Phys. 66 (1994), pp. 899–915.

[35] X.J. Wang, J.R. Yates, I. Souza, and D. Vanderbilt, Phys. Rev. B 74(1) (2006), p. 195118.

[36] X.J. Wang, D. Vanderbilt, J.R. Yates, and I. Souza, Phys. Rev. B 76 (2007), p. 195109.

[37] F.D.M. Haldane, Phys. Rev. Lett. 61(18) (1988), pp. 2015–2018.

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

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

[40] B.A. Bernevig and S.-C. Zhang, Phys. Rev. Lett. 96 (2006), p. 106802.

[41] B.A. Bernevig, T.L. Hughes, and S.-C. Zhang, Science 314(5806) (2006), pp. 1757–1761.

[42] M. König, S. Wiedmann, C. Brüne, A. Roth, H. Buhmann, L.W. Molenkamp, X.-L. Qi, and S.-C. Zhang, Science 318(5851) (2007), pp. 766–770.

[43] H.J. Zhang, C.-X. Liu, X.-L. Qi, X. Dai, Z. Fang, and S.-C. Zhang, Nat. Phys. 5(6) (2009), pp. 438–442.

[44] P. Roushan, J. Seo, C.V. Parker, Y.S. Hor, D. Hsieh, D. Qian, A. Richardella, M.Z. Hasan, R.J. Cava, and A. Yazdani, Nature 460(7259) (2009), pp. 1106–1109.

[45] T. Zhang, P. Cheng, X. Chen, J.-F. Jia, X.C. Ma, K. He, L.L. Wang, H.J. Zhang, X. Dai, Z. Fang, X.C. Xie, and Q.-K. Xue, Phys. Rev. Lett. 103 (2009), p. 266803.

[46] Y. Xia, D. Qian, D. Hsieh, L. Wray, A. Pal, H. Lin, A. Bansil, D. Grauer, Y.S. Hor, R.J. Cava, and M.Z. Hasan, Nat. Phys. 5(6) (2009), pp. 398–402.

[47] Y.L. Chen, J.G. Analytis, J.-H. Chu, Z.K. Liu, S.-K. Mo, X.L. Qi, H.J. Zhang, D.H. Lu, X. Dai, Z. Fang, S.C. Zhang, I.R. Fisher, Z. Hussain, and Z.-X. Shen, Science 325(5937) (2009), pp. 178–181.

[48] Z. Alpichshev, J.G. Analytis, J.-H. Chu, I.R. Fisher, Y.L. Chen, Z.X. Shen, A. Fang, and A. Kapitulnik, Phys. Rev. Lett. 104 (2010), p. 016401.

[49] M.Z. Hasan and C.L. Kane, Rev. Mod. Phys. 82 (2010), pp. 3045–3067.

[50] X.-L. Qi and S.-C. Zhang, Rev. Mod. Phys. 83(4) (2011), pp. 1057–1110.

[51] I. Knez, R.-R. Du, and G. Sullivan, Phys. Rev. Lett. 107 (2011), p. 136603.

[52] H. Beidenkopf, P. Roushan, J. Seo, L. Gorman, I. Drozdov, Y.S. Hor, R.J. Cava, and A. Yazdani, Nat. Phys. 7(12) (2011), pp. 939–943.

[53] Y. Ando, J. Phys. Soc. Jpn. 82(10) (2013), p. 102001.

[54] A.P. Schnyder, S. Ryu, A. Furusaki, and A.W.W. Ludwig, Phys. Rev. B 78(1) (2008), p. 195125.

[55] X.G. Wan, A.M. Turner, A. Vishwanath, and S.Y. Savrasov, Phys. Rev. B 83(20) (2011), p. 205101.

[56] G. Xu, H.M. Weng, Z.J. Wang, X. Dai, and Z. Fang, Phys. Rev. Lett. 107(18) (2011), p. 186806.

[57] A.A. Burkov and L. Balents, Phys. Rev. Lett. 107(12) (2011), p. 127205.

[58] A.A. Zyuzin, S. Wu, and A.A. Burkov, Phys. Rev. B 85(16) (2012), p. 165110.

[59] Z.J. Wang, Y. Sun, X.-Q. Chen, C. Franchini, G. Xu, H.M. Weng, X. Dai, and Z. Fang, Phys. Rev. B 85(1) (2012), p. 195320.

[60] S.M. Young, S. Zaheer, J.C.Y. Teo, C.L. Kane, E.J. Mele, and A.M. Rappe, Phys. Rev. Lett. 108 (2012), p. 140405.

[61] A.M. Turner, Y. Zhang, R.S.K. Mong, and A. Vishwanath, Phys. Rev. B 85 (2012), p. 165120.

[62] Z.J. Wang, H.M. Weng, Q.S. Wu, X. Dai, and Z. Fang, Phys. Rev. B 88(1) (2013), p. 125427.

[63] Z.K. Liu, B. Zhou, Y. Zhang, Z.J. Wang, H.M. Weng, D. Prabhakaran, S.-K. Mo, Z.X. Shen, Z. Fang, X. Dai, Z. Hussain, and Y.L. Chen, Science 343(6) (2014), pp. 864–867.

[64] Z.K. Liu, J. Jiang, B. Zhou, Z.J. Wang, Y. Zhang, H.M. Weng, D. Prabhakaran, S.-K. Mo, H. Peng, P. Dudin, T. Kim, M. Hoesch, Z. Fang, X. Dai, Z.X. Shen, D.L. Feng, Z. Hussain, and Y.L. Chen, Nat. Mater. 13(7) (2014), pp. 677–681.

[65] H.M. Yi, Z.J. Wang, C.Y. Chen, Y.G. Shi, Y. Feng, A.J. Liang, Z.J. Xie, S.L. He, J.F. He, Y. Peng, X. Liu, Y. Liu, L. Zhao, G.D. Liu, X.L. Dong, J. Zhang, M. Nakatake, M. Arita, K. Shimada, H. Namatame, M. Taniguchi, Z.Y. Xu, C.T. Chen, X. Dai, Z. Fang, and X.J. Zhou, Sci. Rep. 4 (2014), pp. 6106.

[66] M. Neupane, S.Y. Xu, R. Sankar, N. Alidoust, G. Bian, C. Liu, I. Belopolski, T.R. Chang, H.T. Jeng, H. Lin, A. Bansil, F.C. Chou, and M.Z. Hasan, Nat. Commun. 5 (2014), p. 3786.

[67] B.-J. Yang and N. Nagaosa, Nat. Commun. 5 (2014), p. 4898.

[68] S.J. Jeon, B.B. Zhou, A. Gyenis, B.E. Feldman, I. Kimchi, A.C. Potter, Q.D. Gibson, R.J. Cava, A. Vishwanath, and A. Yazdani, Nat. Mater. 13(9) (2014), pp. 851–856.

[69] S.-Y. Xu, C. Liu, S.K. Kushwaha, R. Sankar, J.W. Krizan, I. Belopolski, M. Neupan, G. Bian, N. Alidoust, T.-R. Chang, H.-T. Jeng, C.-Y. Huang, W.-F. Tsai, H. Lin, P.P. Shibayev, F. Chou, R.J. Cava, and M.Z. Hasan, Science 347(6219) (2015), pp. 294–298.

[70] A.A. Burkov, M.D. Hook, and L. Balents, Phys. Rev. B 84(23) (2011), p. 235126.

[71] L. Lu, L. Fu, J.D. Joannopoulos, and M. Soljačić, Nat. Photonics 7(4) (2013), pp. 294–299.

[72] H. Weng, Y. Liang, Q. Xu, R. Yu, Z. Fang, X. Dai, and Y. Kawazoe, Phys. Rev. B 92 (2015), p. 045108. DOI: 10.1103/PhysRevB.92.045108.

[73] L. Fu and C.L. Kane, Phys. Rev. Lett. 100(9) (2008), p. 096407.

[74] X.-L. Qi, T.L. Hughes, and S.-C. Zhang, Phys. Rev. B 78(19) (2008), p. 195424.

[75] X.-L. Qi, T.L. Hughes, and S.-C. Zhang, Phys. Rev. B 82(18) (2010), p. 184516.

[76] L. Fu and E. Berg, Phys. Rev. Lett. 105(9) (2010), p. 097001.

[77] Y.S. Hor, A.J. Williams, J.G. Checkelsky, P. Roushan, J. Seo, Q. Xu, H.W. Zandbergen, A. Yazdani, N.P. Ong, and R.J. Cava, Phys. Rev. Lett. 104(5) (2010), p. 057001.

[78] J.L. Zhang, S.J. Zhang, H.M. Weng, W. Zhang, L.X. Yang, Q.Q. Liu, S.M. Feng, X.C. Wang, R.C. Yu,L.Z. Cao, L. Wang, W.G. Yang, H.Z. Liu, W.Y. Zhao, S.C. Zhang, X. Dai, Z. Fang, and C.Q. Jin, Proc. Natl. Acad. Sci. 108(1) (2011), pp. 24–28.

[79] H.M. Weng, G. Xu, H.J. Zhang, S.C. Zhang, X. Dai, and Z. Fang, Phys. Rev. B 84 (2011), p. 060408.

[80] M. Onoda and N. Nagaosa, Phys. Rev. Lett. 90(20) (2003), p. 206601.

[81] S.-S. Chern, Ann. Math. 46(4) (1945), p. 674.

[82] D.J. Thouless, J. Phys. C: Solid State Phys. 17(12) (1984), pp. L325–L327.

[83] T. Thonhauser and D. Vanderbilt, Phys. Rev. B 74(23) (2006), p. 235111.

[84] C. Brouder, G. Panati, M. Calandra, C. Mourougane, and N. Marzari, Phys. Rev. Lett. 98(4) (2007), p. 046402.

[85] A.A. Soluyanov and D. Vanderbilt, Phys. Rev. B 83(3) (2011), p. 035108.

[86] Y. Hatsugai, Phys. Rev. B 48(16) (1993), pp. 11851–11862.

[87] A.M. Essin and V. Gurarie, Phys. Rev. B 84(1) (2011), p. 125132.

[88] A.H. MacDonald and P. Středa, Phys. Rev. B 29(4) (1984), pp. 1616–1619.

[89] B.I. Halperin, Phys. Rev. B 25(4) (1982), pp. 2185–2190.

[90] Z.-C. Gu and X.-G. Wen, Phys. Rev. B 80(15) (2009), p. 155131.

[91] L. Fu and C.L. Kane, Phys. Rev. B 74(19) (2006), p. 195312.

[92] R. Yu, X.L. Qi, A. Bernevig, Z. Fang, and X. Dai, Phys. Rev. B 84(7) (2011), p. 75119.

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

[94] J.E. Moore and L. Balents, Phys. Rev. B 75(12) (2007), p. 121306.

[95] R. Roy, Phys. Rev. B 79(19) (2009), p. 195322.

[96] G.E. Volovik, The Universe in a Helium Droplet, Oxford, 2009.

[97] J. Zak, Phys. Rev. Lett. 62(23) (1989), pp. 2747–2750.

[98] H. Weyl, Zeitschrift für Physik 56(5–6) (1929), pp. 330–352.

[99] H.B. Nielsen and M. Ninomiya, Nucl. Phys. B 185(1) (1981), pp. 20–40.

[100] H.B. Nielsen and M. Ninomiya, Nucl. Phys. B 193(1) (1981), pp. 173–194.

[101] F.D.M. Haldane, *Attachment of surface "Fermi arcs" to the bulk Fermi surface: "Fermi-level plumbing" in topological metals*, preprint (2014). Available at arXiv:1401.0529 [cond-mat].

[102] L. Balents, Physics 4 (2011), p. 36.

[103] S. Murakami, New J. Phys. 9(9) (2007), p. 356.

[104] W. Zhang, R. Yu, H.-J. Zhang, X. Dai, and Z. Fang, New J. Phys. 12(6) (2010), p. 065013.

[105] J.P. Liu and D. Vanderbilt, Phys. Rev. B 90 (2014), p. 155316.

[106] G.B. Halász and L. Balents, Phys. Rev. B 85 (2012), p. 035103.

[107] M. Hirayama, R. Okugawa, S. Ishibashi, S. Murakami, and T. Miyake, Phys. Rev. Lett. 114 (2015), p. 206401. DOI: 10.1103/PhysRevLett.114.206401.

[108] H. Weng, C. Fang, Z. Fang, B.A. Bernevig, and X. Dai, Phys. Rev. X 5 (2015), p. 011029.

[109] S.-M. Huang, S.-Y. Xu, I. Belopolski, C.-C. Lee, G. Chang, B. Wang, N. Alidoust, G. Bian, M. Neupane, A. Bansil, H. Lin, and M. ZahidHasan, *An inversion breaking Weyl semimetal state in the TaAs material class*, Ref. 109 is formally published as Nature Commnunications 6, (2015), p. 7383.

[110] X.-L. Qi, Y.-S. Wu, and S.-C. Zhang, Phys. Rev. B 74(8) (2006), p. 085308.

[111] Y. Chen, D.L. Bergman, and A.A. Burkov, Phys. Rev. B 88 (2013), p. 125110.

[112] A.A. Burkov, Phys. Rev. B 89 (2014), p. 155104.

[113] A.A. Burkov, Phys. Rev. Lett. 113 (2014), p. 187202.

[114] B.A. Bernevig and T.L. Hughes, Topological Insulators and Topological Superconductors, Princeton University Press, 2013.

[115] D. Vanderbilt, I. Souza, and F.D.M. Haldane, Phys. Rev. B 89 (2014), p. 117101.

[116] Y. Chen, D.L. Bergman, and A.A. Burkov, *Reply to comment by Vanderbilt, Souza and Haldane: arxiv:1312.4200*, preprint (2013). Available at ArXiv:1312.5751.

[117] F.D.M. Haldane, Phys. Rev. Lett. 93 (2004), p. 206602.

[118] R.M. Martin, Electronic Structure: Basic Theory and Practical Methods, Cambridge University Press, 2004.

[119] N. Marzari, A.A. Mostofi, J.R. Yates, I. Souza, and D. Vanderbilt, Rev. Mod. Phys. 84(4) (2012), pp. 1419–1475.

[120] K.G. Wilson, Phys. Rev. D 10(8) (1974), pp. 2445–2459.

[121] R. Giles, Phys. Rev. D 24(8) (1981), pp. 2160–2168.

[122] M. Taherinejad, K.F. Garrity, and D. Vanderbilt, Phys. Rev. B 89(1) (2014), p. 115102.

[123] Z. Ringel and Y.E. Kraus, Phys. Rev. B 83(24) (2011), p. 245115.

[124] L. Fu, Phys. Rev. Lett. 106(10) (2011), p. 106802.

[125] T.H. Hsieh, H. Lin, J.W. Liu, W.H. Duan, A. Bansi, and L. Fu, Nat. Commun. 3 (2012), p. 982.

[126] L. Fu and C.L. Kane, Phys. Rev. B 76(4) (2007), p. 045302.

[127] T. Fukui and Y. Hatsugai, J. Phys. Soc. J. 76(5) (2007), p. 053702.

[128] X.-L. Sheng, Z.J. Wang, R. Yu, H.M. Weng, Z. Fang, and X. Dai, Phys. Rev. B 90 (2014), p. 245308.

[129] I. Souza, N. Marzari, and D. Vanderbilt, Phys. Rev. B 65(3) (2001), p. 035109.

[130] J.R. Yates, X.J. Wang, D. Vanderbilt, and I. Souza, Phys. Rev. B 75(19) (2007), p. 195121.

[131] S. Yip, Handbook of Materials Modeling, Springer, 2007.

[132] A.H. Walter, Electronic Structure and the Properties of Solids: The Physics of the Chemical Bond, Dover, 1989.

[133] E.M. Godfrin, J. Phys.: Cond. Matter 3(40) (1999), pp. 7843–7848.

[134] M.P.L. Sancho, J.M.L. Sancho, and J. Rubio, J. Phys. F: Metal Phys. 14(5) (2000), pp. 1205–1215.

[135] M.P.L. Sancho, J.M.L. Sancho, J.M.L. Sancho, and J. Rubio, J. Phys. F: Metal Phys. 15(4) (2000), pp. 851–858.

[136] H.-J. Zhang, C.-X. Liu, X.-L. Qi, X.-Y. Deng, X. Dai, S.-C. Zhang, and Z. Fang, Phys. Rev. B 80(8) (2009), p. 085307.

[137] X. Dai, T.L. Hughes, X.-L. Qi, Z. Fang, and S.-C. Zhang, Phys. Rev. B 77(12) (2008), p. 125319.

[138] C.-X. Liu, X.-L. Qi, X. Dai, Z. Fang, and S.-C. Zhang, Phys. Rev. Lett. 101(14) (2008), p. 146802.

[139] R. Yu, W. Zhang, H.-J. Zhang, S.-C. Zhang, X. Dai, and Z. Fang, Science 329(5987) (2010), pp. 61–64.

[140] Q.-Z. Wang, X. Liu, H.-J. Zhang, N. Samarth, S.-C. Zhang, and C.-X. Liu, Phys. Rev. Lett. 113 (2014), p. 147201.

[141] C.-Z. Chang, J.S. Zhang, X. Feng, J. Shen, Z.C. Zhang, M.H. Guo, K. Li, Y.B. Ou, P. Wei, L.-L. Wang, Z.-Q. Ji, Y. Feng, S.H. Ji, X. Chen, J.F. Jia, X. Dai, Z. Fang, S.-C. Zhang, K. He, Y.Y. Wang, L. Lu, X.C. Ma, and Q.-K. Xue, Science 340(6129) (2013), pp. 167–170.

[142] J.G. Checkelsky, R. Yoshimi, A. Tsukazaki, K.S. Takahashi, Y. Kozuka, J. Falson, M. Kawasaki, and Y. Tokura, Nature Physics 10(10) (2014), pp. 731–736.

[143] K.-Y. Yang, Y.-M. Lu, and Y. Ran, Phys. Rev. B 84(7) (2011), p. 075129.

[144] Z.H. Qiao, S.Y.A. Yang, W.X. Feng, W.-K. Tse, J. Ding, Y.G. Yao, J. Wang, and Q. Niu, Phys. Rev. B 82(16) (2010), p. 161414.

[145] R. Nandkishore and L. Levitov, Phys. Rev. B 82(11) (2010), p. 115124.

[146] W.-K. Tse, Z.H. Qiao, Y.G. Yao, A.H. MacDonald, and Q. Niu, Phys. Rev. B 83(1) (2011), p. 155447.

[147] F. Zhang, J. Jung, G.A. Fiete, Q. Niu, and A.H. MacDonald, Phys. Rev. Lett. 106 (2011), p. 156801.

[148] J. Ding, Z.H. Qiao, W.X. Feng, Y.G. Yao, and Q. Niu, Phys. Rev. B 84(19) (2011), p. 195444.

[149] Z.H. Qiao, W.-K. Tse, H. Jiang, Y.G. Yao, and Q. Niu, Phys. Rev. Lett. 107(25) (2011), p. 256801.

[150] H.B. Zhang, C. Lazo, S. Blgel, S. Heinze, and Y. Mokrousov, Phys. Rev. Lett. 108(5) (2012), p. 056802.

[151] Z.H. Qiao, H. Jiang, X. Li, Y. G. Yao, and Q. Niu, Phys. Rev. B 85 (2012), p. 115439.

[152] Z.H. Qiao, W. Ren, H. Chen, L. Bellaiche, Z.Y. Zhang, A.H. MacDonald, and Q. Niu, Phys. Rev. Lett. 112(1) (2014), p. 116404.

[153] M. Ezawa, Phys. Rev. Lett. 109(5) (2012), p. 055502.

[154] M. Ezawa, New J. Phys. 16(6) (2014), p. 065015.

[155] M. Ezawa, Phys. Rev. Lett. 110(2) (2013), p. 026603.

[156] D. Xiao, W.G. Zhu, Y. Ran, N. Nagaosa, and S. Okamoto, Nat. Commun. 2 (2011), p. 596.

[157] K.-Y. Yang, W.G. Zhu, D. Xiao, S. Okamoto, Z.Q. Wang, and Y. Ran, Phys. Rev. B 84(20) (2011), p. 201104.

[158] A. Rüegg and G.A. Fiete, Phys. Rev. B 84(20) (2011), p. 201103.

[159] X. Hu, A. Rüegg, and G.A. Fiete, Phys. Rev. B 86(23) (2012), p. 235141.

[160] A. Rüegg, C. Mitra, A.A. Demkov, and G.A. Fiete, Phys. Rev. B 85 (2012), p. 245131.

[161] Q.-F. Liang, L.-H. Wu, and X. Hu, New J. Phys. 15(6) (2013), p. 063031.

[162] Y.L. Wang, Z.J. Wang, Z. Fang, and X. Dai, Phys. Rev. B 91 (2015), p. 125139. DOI: 10.1103/PhysRevB.91.125139.

[163] Z.F. Wang, Z. Liu, and F. Liu, Phys. Rev. Lett. 110(19) (2013), p. 196801.

[164] K.F. Garrity and D. Vanderbilt, Phys. Rev. Lett. 110(1) (2013), p. 116802.

[165] K.F. Garrity and D. Vanderbilt, Phys. Rev. B 90 (2014), p. 121103.

[166] H.J. Zhang, J. Wang, G. Xu, Y. Xu, and S.-C. Zhang, Phys. Rev. Lett. 112 (2014), p. 096804.

[167] H.J. Zhang and S.-C. Zhang, Phys. Status Solide RRL. 7(1–2) (2013), pp. 72–81.

[168] H.M. Weng, X. Dai, and Z. Fang, Phys. Rev. X 4 (2014), p. 011002.

[169] H.M. Weng, J.Z. Zhao, Z.J. Wang, Z. Fang, and X. Dai, Phys. Rev. Lett. 112(1) (2014), p. 016403.

[170] C.-X. Liu, X.-L. Qi, H.J. Zhang, X. Dai, Z. Fang, and S.-C. Zhang, Phys. Rev. B 82(4) (2010), p. 045122.

[171] H.J. Zhang, Y. Xu, J. Wang, K. Chang, and S.-C. Zhang, Phys. Rev. Lett. 112 (2014), p. 216803.

[172] Y. Zhang, K. He, C.-Z. Chang, C.-L. Song, L.-L. Wang, X. Chen, J.-F. Jia, Z. Fang, X. Dai, W.-Y. Shan, S.-Q. Shen, Q. Niu, X.-L. Qi, S.-C. Zhang, X.-C. Ma, and Q.-K. Xue, Nat. Phys. 6(8) (2010), pp. 584–588.

[173] V.A. Kulbachinskii, A.Yu. Kaminskii, K. Kindo, Y. Narumi, K. Suga, P. Lostak, and P. Svanda, JETP Lett. 73(7) (2001), pp. 352–356.

[174] J.S. Dyck, P. Hajek, P. Lostak, and C. Uher, Phys. Rev. B 65(11) (2002), p. 115212.

[175] Y.J. Chien, *Transition metal-doped sbte and bite diluted magnetic semiconductors*, Ph.D. diss., University of Michigan, 2007.

[176] C.W. Niu, Y. Dai, L. Yu, M. Guo, Y.D. Ma, and B.B. Huang, Appl. Phys. Lett. 99(14) (2011), p. 142502.
[177] Z.H. Zhou, Y.-J. Chien, and C. Uher, Phys. Rev. B 74(22) (2006), p. 224418.
[178] M.A. Ruderman and C. Kittel, Phys. Rev. 96(1954), pp. 99–102.
[179] J.H. Van Vleck, The Theory of Electronic and Magnetic Susceptibilities, Oxford University Press, London, 1932.
[180] J.G. Checkelsky, J.T. Ye, Y. Onose, Y. Iwasa, and Y. Tokura, Nat. Phys. 8(10) (2012), pp. 729–733.
[181] J.S. Zhang, C.Z. Chang, P.Z. Tang, Zu.C. Zhang, X. Feng, K. Li, L.-L. Wang, X. Chen, C.X. Liu, W.H. Duan, K. He, Q.-K. Xue, X.C. Ma, and Y.Y. Wang, Science 339(6127) (2013), pp. 1582–1586.
[182] J.-M. Zhang, W.G. Zhu, Y. Zhang, D. Xiao, and Y.G. Yao, Phys. Rev. Lett. 109(26) (2012), p. 266405.
[183] K. Sato, P.H. Dederichs, H. Katayama-Yoshida, and J. Kudrnovsky, Phys. B: Cond. Matter 340 (2003), pp. 863–869.
[184] T. Dietl, Phys. Rev. B 63(19) (2001), p. 195205.
[185] J. Wang, B. Lian, H.J. Zhang, Y. Xu, and S.-C. Zhang, Phys. Rev. Lett. 111 (2013), p. 136801.
[186] X. Liu, H.-C. Hsu, and C.-X. Liu, Phys. Rev. Lett. 111 (2013), p. 086802.
[187] H.B. Nielsen and M. Ninomiya, Phys. Lett. B 130(6) (1983), pp. 389–396.
[188] A.N. Yaresko, Phys. Rev. B 77(11) (2008), p. 115106.
[189] P. Wojtowicz, IEEE Trans. Magn. 5(4) (1969), pp. 840–848.
[190] P.K. Baltzer, H.W. Lehmann, and M. Robbins, Phys. Rev. Lett. 15(11) (1965), pp. 493–495.
[191] P.K. Baltzer, P.J. Wojtowicz, M. Robbins, and E. Lopatin, Phys. Rev. 151(2) (1966), pp. 367–377.
[192] C.-Y. Moon and S.-H. Wei, Phys. Rev. B 74(4) (2006), p. 045205.
[193] A. Delin, Phys. Rev. B 65(15) (2002), p. 153205.
[194] C.J. Fennie and K.M. Rabe, Phys. Rev. B 72(21) (2005), p. 214123.
[195] N.I. Solin, V.V. Ustinov, and S.V. Naumov, Phys. Solid State 50(5) (2008), pp. 901–908.
[196] H.W. Lehmann and F.P. Emmenegger, Solid State Commun. 7(14) (1969), pp. 965–968.
[197] A. Selmi, A. Mauger, and M. Heritier, J. Magn. Magn. Mater. 66(3) (1987), pp. 295–316.
[198] L.C. Lew Yan Voon and M. Willatzen, The k·p Method: Electronic Properties of Semiconductors, 2009. book.
[199] N.I. Solin and N.M. Chebotaev, Phys. Solid State 39(5) (1997), pp. 754–758.
[200] T. Guan, C.J. Lin, C.L. Yang, Y.G. Shi, C. Ren, Y.Q. Li, H.M. Weng, X. Dai, Z. Fang, S.S. Yan, and P. Xiong, *Single s-band Half-Metallicity in n-type HgCr$_2$Se$_4$*, preprint (2015). Available at arXiv:1503.03190 [cond-mat].
[201] M.R. Chaves, J.L. Ribeiro, A. Selmi, and P. Gibart, Phys. Status Solidi (A) 92(1) (1985), pp. 263–271.
[202] N. Tsuji, T. Oka, and H. Aoki, Phys. Rev. Lett. 103(4) (2009), p. 047403.
[203] X. Li, T. Cao, Q. Niu, J.R. Shi, and J. Feng, Proc. Natl. Acad. Sci. 110(10) (2013), pp. 3738–3742.
[204] M. Ezawa, Phys. Rev. B 87(15) (2013), p. 155415.
[205] D.N. Sheng, Z.Y. Weng, L. Sheng, and F.D.M. Haldane, Phys. Rev. Lett. 97(3) (2006), p. 036808.
[206] Y.Y. Yang, Z. Xu, L. Sheng, B.G. Wang, D.Y. Xing, and D.N. Sheng, Phys. Rev. Lett. 107(6) (2011), p. 66602.
[207] D. Xiao, W. Yao, and Q. Niu, Phys. Rev. Lett. 99 (2007), p. 236809.
[208] X.D. Xu, W. Yao, D. Xiao, and T.F. Heinz, Nat. Phys. 10(5) (2014), pp. 343–350.
[209] J. Mannhart and D.G. Schlom, Science 327(5973) (2010), pp. 1607–1611.
[210] P. Zubko, S. Gariglio, M. Gabay, P. Ghosez, and J.-M. Triscone, Annu. Rev. Cond. Matter Phys. 2(1) (2011), pp. 141–165.
[211] H.Y. Hwang, Y. Iwasa, M. Kawasaki, B. Keimer, N. Nagaosa, and Y. Tokura, Nat. Mater. 11(2) (2012), pp. 103–113.
[212] K. Ueda, Science 280(5366) (1998), pp. 1064–1066.
[213] S. Chakraverty, A. Ohtomo, M. Okude, K. Ueno, and M. Kawasaki, Crys. Growth Des. 10(4) (2010), pp. 1725–1729.
[214] B. Gray, H.N. Lee, J. Liu, J. Chakhalian, and J.W. Freeland, Appl. Phys. Lett. 97(1) (2010), p. 3105.
[215] R.H.P. Mitchell, Perovskites: Modern and Ancient, Almaz Press, Ontario, Canada, 2002.
[216] X.Y. Deng, L. Wang, X. Dai, and Z. Fang, Phys. Rev. B 79(7) (2009), p. 075114.

[217] G. Kresse and J. Furthmüller, Phys. Rev. B 54 (1996), pp. 11169–11186.

[218] Y.L. Wang, Z.J. Wang, Z. Fang, and X. Dai, Phys. Rev. B 91 (2015), p. 125139.

[219] S. Oh, Science 340(6129) (2013), pp. 153–154.

[220] H. Jiang, Z.H. Qiao, H.W. Liu, and Q. Niu, Phys. Rev. B 85(4) (2012), p. 045445.

[221] C. Fang, M.J. Gilbert, and B.A. Bernevig, Phys. Rev. Lett. 112 (2014), p. 046801.

[222] Y. Xu, B.H. Yan, H.-J. Zhang, J. Wang, G. Xu, P.Z. Tang, W.H. Duan, and S.C. Zhang, Phys. Rev. Lett. 111(13) (2013), p. 136804.

[223] S.C. Wu, G.C. Shan, and B.H. Yan, Phys. Rev. Lett. 113 (2015), p. 256401. DOI: 10.1103/PhysRevLett.113.256401.

[224] S.R. Elliott and M. Franz, Rev. Mod. Phys. 87 (2015), pp. 137–163.

[225] R.M. Lutchyn, J.D. Sau, and S. Das Sarma, Phys. Rev. Lett. 105(7) (2010), p. 077001.

[226] J.D. Sau, R.M. Lutchyn, S. Tewari, and S. Das Sarma, Phys. Rev. Lett. 104(4) (2010), p. 040502.

[227] Y. Oreg, G. Refael, and F. von Oppen, Phys. Rev. Lett. 105(17) (2010), p. 177002.

[228] J. Klinovaja, S. Gangadharaiah, and D. Loss, Phys. Rev. Lett. 108(19) (2012), p. 196804.

[229] V. Mourik, K. Zuo, S.M. Frolov, S.R. Plissard, E.P.A.M. Bakkers, and L.P. Kouwenhoven, Science 336(6084) (2012), pp. 1003–1007.

[230] S. Nadj-Perge, I.K. Drozdov, J. Li, H. Chen, S. Jeon, J. Seo, A.H. MacDonald, B.A. Bernevig, and A. Yazdani, Science 346(6209) (2014), pp. 602–607.

[231] X.L. Liu, Z.Q. Wang, X.C. Xie, and Y. Yu, Phys. Rev. B 83(1) (2011), p. 125105.