A Pedagogic Review on Designing Model Topological Insulators

Tanmoy Das

Abstract | Following the centuries old concept of the quantization of flux through a Gaussian curvature (Euler characteristic) and its successive dispersal into various condensed matter properties such as quantum Hall effect, and topological invariants, we can establish a simple and fairly universal understanding of various modern topological insulators (TIs). Formation of a periodic lattice (which is a non-trivial Gaussian curvature) of ‘cyclotron orbits’ with applied magnetic field, or ‘chiral orbits’ with fictitious ‘momentum space magnetic field’ (Berry curvature) guarantees its flux quantization, and thus integer quantum Hall (IQH), and quantum spin-Hall (QSH) insulators, respectively, occur. The bulk-boundary correspondence associated with all classes of TIs dictates that some sort of pumping or polarization of a ‘quantity’ at the boundary must be associated with the flux quantization or topological invariant in the bulk. Unlike charge or spin accumulations at the edge for IQH and QSH states, the time-reversal (TR) invariant \( Z_2 \) TI class pumps a mathematical quantity called ‘TR polarization’ to the surface. This requires that the valence electron’s wavefunction (say, \( \psi^\uparrow(k) \)) switches to its TR conjugate \( \psi^\downarrow(-k) \) odd number of times in half of the Brillouin zone. These two universal features can be considered as ‘targets’ to design and predict various TIs. For example, we demonstrate that when two adjacent atomic chains or layers are assembled with opposite spin-orbit coupling (SOC), serving as the TR partner to each other, the system naturally becomes a \( Z_2 \) TI. This review delivers a holistic overview on various concepts, computational schemes, and engineering principles of TIs.
1 Introduction
Phase transition is distinguished by a change in symmetry, involving both a reduction or addition of symmetry in the ground state. A reduction of symmetry, which commonly involves translational, Time-Reversal (TR), rotational, gauge symmetries, among others, leads to a classical/quantum phase transition, and is defined by an order parameter within the Landau’s paradigm. On the contrary, a topological phase is defined by the emergence of a new quantum number (such as Landau level, Chern number, etc.), arising from the geometry or topology of the band structure. The topological invariant can be understood from a pure mathematical formalism of the Euler characteristic or Euler number. This implies that the net flux through a Gaussian curvature is always quantized. This precisely what happens, according to Laughlin’s argument, in two-dimensional (2D) lattices (which can be represented by a torus—a Gaussian curvature) when magnetic field is applied perpendicular to it. In this case, the magnetic flux through the 2D system or torus remains quantized, giving rise to Integer Quantum Hall (IQH) effect. This is the first realization of topological invariance in condensed matter science.

IQH is a well understood phenomenon, with different ways to quantify its topological invariants. For example, Thouless, Kohmoto, Nightingale, and Nijs (TKNN) have shown that the IQH effect can also be understood from the Berry phase paradigm in the momentum space. The corresponding topological invariant is thus known as TKNN number or Chern number. Haldane proposed a pioneering idea to obtain Quantum Hall (QH) effect without external magnetic field in a honeycomb lattice. Honeycomb lattice has two different sublattices. With the application of an external gauge field, the intra-sublattice hoppings possess chiral motion, and the chirality of two sublattices becomes opposite to each other. Each intra-sublattice hopping form counter-propagating triangular ‘cyclotron orbits’, each threading opposite ‘magnetic fields’. As we will go along in this review, we will identify that the formation of localized ‘cyclotron orbitals’ in 2D lattice without a magnetic field, which we call ‘chiral orbit’, is the key ingredient to obtain QH effect. Each such chiral orbit now encloses integer flux quanta, in the same language of the IQH effect, and due to TR symmetry breaking a net QH effect survives.

The next development to the TI field was put forward by Kane and Mele in 2005, for obtaining TR invariant TIs, as known by $Z_2$ TI. They realized that Haldane’s ‘gauge field’ can be achieved by SOC Spin-Orbit Coupling (SOC). Owing to the spin-momentum locking, the right- and left-moving electrons have opposite spin polarizations. Since the TR symmetry is intact here, the flux passing through different spin-resolved chiral orbits in a 2D lattice are exactly equal but opposite. This makes the net flux to be zero, leading to no charge pumping to the edge, but the difference between the two fluxes is finite, giving rise to a net spin-resolved QH effect, as known by Quantum Spin-Hall (QSH) effect. This is the foundation of the 2D TI.

The generalization of the $Z_2$ topological invariant to 3D cannot, however, be easily done in terms of chiral orbit formations, except in special cases of layered systems and heterostructures. There are several mathematical formulations of the $Z_2$ invariant, among which, the Kane-Mele method of TR polarization is widely used. They proposed a derivation of the bulk $Z_2$ topological invariant from the bulk-boundary correspondence, which is a necessary condition for all topological invariants. Recall that in the case of IQH and QSH insulators, charge and spin are accumulated at the edge. Based on this, they enquired that something similar must be pumped to the boundary in $Z_2$ TIs. Since QSH insulator also belongs to the $Z_2$ class, the quantity that is pumped to the edge must accommodate spin as a subset. Since the spin-up state at the +$\mathbf{k}$ and spin-down state at –$\mathbf{k}$ are TR conjugates to each other, Kane-Mele proposed that a more general mathematical quantity, called ‘TR polarization’, is accumulated at the edges in this case. Here, electrons possessing a particular wavefunction and their TR conjugate partners move to different edges. This incipiently requires that the electron exchanges its TR partner odd number of times in traversing half-of the Brillouin zone (BZ). In what follows, $Z_2$ topological invariant is nothing but counts of the number TR partner exchanges; odd number corresponds to $Z_2$ invariant $\nu = 1$, while even number implies $\nu = 0$. Since the TR symmetric topological invariant only takes two values here, it has $Z_2$ symmetry, but TR breaking IQH state can take arbitrarily large Chern number.

Chirality of electrons is an essential ingredient for TIs, which is obtained either by magnetic field, or in bipartite lattice (such as staggered hopping in Su-Schrieffer-Heeger (SSH) lattice, or graphene) or via SOC. Such chirality can also arise from the orbital texture inversion between even and odd parity orbitals at the TR invariant points, leading to a distinct class of spinless topological and Dirac materials. In simple terms, chirality means that the electron’s hopping matrix-element must be complex in the momentum space. As the electron’s hopping encloses a ‘chiral orbit’, in the
momentum space’, the complex phase manifests into a magnetic field at the center—either applied, or self-generated (Berry curvature). For \( Z_2 \) class, the TR partner switching discussed in the previous paragraph is nothing but the exchange of electron’s complex phase to its complex conjugate odd number of time in half-of the BZ.

Subsequently, Fu-Kane simplifies the calculation of \( Z_2 \) invariant from the parity analysis.\(^7\) They showed that if a system has both TR and inversion symmetries, the \( Z_2 \) topological invariant can be computed simply by counting how many times the electron exchanges it parity at the TR invariant momenta. If the valence band is not fully defined by a single parity, rather it exchanges parities odd number of times with the conduction band (as in the case of TR partner exchange), it gives a non-trivial topological invariant. This is also equivalent to chirality inversion in special cases as we will demonstrate in our engineering procedures. The band gap between the opposite parity conduction and valence bands at the TR invariant momenta also serves a ‘negative’ Dirac mass in the Dirac equation.\(^8,18\)

The consequence of this band inversion is very rich, allowing protected gapless surface or edge states with Dirac cone. This provides an alternative springboard to obtain numerous exotic properties originally proposed by solving Dirac equation in the high-energy theories.

This theory leads to the search of materials with band inversions at the TR invariant k-points as a simple tool to identify \( Z_2 \) TI. The mechanism for the parity inversion is not unique, and can, in-principle, vary to a wide range of tuning parameters as well as electron-electron interactions. Among them, SOC is responsible for the band inversion in most of the known TIs. Initially, the search for TI had been much ‘blind-folded’, seeking materials with odd number of band inversions triggered by SOC.\(^6,9,18,19\) Subsequently, more advanced methods of TI materials genome such as ‘adiabatic transformation’ method was developed,\(^20\) which can be applied to systems without inversion symmetry. In this method, one starts with a known TI, and continuously tunes the atomic number of the constituent elements, and arrives at a new material. In this process if the band gap has not closed and reopened at the TR invariant points, the new material must also be a non-trivial TI. These two methods have enabled the discoveries of a rich variety of TI materials.

Subsequently, various distinct classes of TI are predicted and discovered. For example, mirror symmetry and \( p \)-wave superconducting pairing symmetry can lead to two distinct classes of TIs, called topological crystalline insulator,\(^21\) and topological superconductor,\(^22,23\) respectively. Spontaneous TR symmetry breaking TIs, without magnetic field, are known as Quantum Anomalous Hall (QAH) insulator in 2D, or topological axion insulator in 3D with subtle differences between them.\(^24,25\) The axion insulator has a quantized magnetoelectric response identical to that of a (strong) TI, but lacks the protected surface states of the TI. Other methods of obtaining insulating state such as disorder, Kondo effect, or Hubbard interaction, is associated with odd number of band inversion, and are also proposed to give topological Anderson insulator,\(^26,27\) topological Kondo insulator,\(^28\) and topological Mott insulator,\(^29\) respectively. Finally, a new class of TI is proposed by the present author, which is called quantum spin-Hall density wave (QSHDW) insulator in quasi-2D.\(^30\) When two opposite chiral states are significantly nested in a given system, it renders a transitional symmetry breaking Landau order parameter (forming spin-orbit density wave),\(^31,32\) which can be associated with odd number of band inversion and \( Z_2 \) topological invariant for a special nesting vector. In this case, the parity or chirality inversion occurs in the real space between different lattice sites, breaking the transitional symmetry, but not the TR symmetry.

Density functional theory (DFT)\(^18,19\) calculations take a preceding role in predicting most of the TIs, many of which are followed by experimental realizations. Owing to weak SOC in graphene, this material has not been realized to be intrinsic TI despite its first prediction. HgTe/CdTe quantumwellstatewerereportedtobe2DTI,\(^7\) which was followed by its experimental realization.\(^33\) The 3D topological semimetal prediction\(^34\) and realized\(^35\) is Bi\(_2\)S\(_2\). The first 3D TI was discovered in Bi\(_2\)Se\(_3\), Bi\(_2\)Te\(_3\), and Sb\(_2\)Te\(_3\) families, both theoretically and experimentally.\(^18,19,35\) This is followed by a series of predictions of 3D TIs including gray tin,\(^7\) HgTe, InAs,\(^36\) ternary tetramytes Ge\(_{1-x}\)Bi\(_x\)Te\(_2\), series,\(^36\) half-Huesler compounds,\(^37,38\) TI-based III-V-VI\(_2\) chalcogenides,\(^39,40\) ternary I-III-VI\(_2\) and II-IV-VI\(_2\) chalcopryites, I\(_3\)-V-VI\(_4\) famatinites, and quaternary I\(_2\)-II-IV-VI\(_4\) chalcogenides,\(^41\) Li\(_2\)AgSb,\(^42\) LiAuSe honeycomb lattice,\(^42\) \(p\)-Ag\(_2\)Te,\(^43\) non-centrosymmetric Bi\(_2\)Te\(_3\) (\(X = Cl, I, Br\))\(^44,46\). Recently a number of materials were discovered to have a stable 2D structure, among which Si, Ge,\(^46\) Sn,\(^47\) As,\(^48\) Bi,\(^39,50\) P\(^51\) are predicted to exhibit QSH insulating state with SOC, and with other tuning. Pb, Sn, Se/Te\(_2\),\(^52–54\) and SnS\(^55\) are the only known topological crystalline insulators known to date, \(f\)-electron based compounds such as SmB\(_6\),\(^28,56\) and YbB\(_6\),\(^57\) are predicted to be topological Kondo

**Berry curvature:**
Berry curvature is a fictitious magnetic field that can be derived in the momentum space as to be associated with the Berry phase. The formula used for the derivation is analogous to the Pierels phase arising from applied magnetic field.

**Adiabatic transformation:**
Adiabatic transformation is referred to slow and continuous change of materials parameter such as lattice constant, atomic number etc and arriving at a different material or material phase.
insulators, while PbBi is considered a topological Mott insulator.\textsuperscript{58} U\textsubscript{3}Ru\textsubscript{2}Si\textsubscript{4} is considered as a candidate material for the spin-orbit density wave induced hidden topological order system.\textsuperscript{31,39,60} Ir, and Os-based oxides are proposed to be axion insulators.\textsuperscript{25} The list of QAH insulator is rather small,\textsuperscript{41} most of which requires external tuning with magnetic doping and thin films. La\textsubscript{X} (X = Br, I and Cl) family is predicted to be intrinsic QAH insulator with sizable band gap for applications.\textsuperscript{62} A complete materials repository and their individual properties can be found in Refs.\textsuperscript{63,64}

Despite this tremendous success as well as continuing research activities for harvesting more TI materials, the real struggle with the presently available materials is that the unwanted bulk conduction band lies below the Fermi level, and contributes to the transport phenomenon. There have been considerable efforts to eliminate the bulk band above the Fermi level by chemical doping,\textsuperscript{65,66} pressure,\textsuperscript{67,68} photo-doping,\textsuperscript{69} and heterostructure,\textsuperscript{70,71} among others. All these efforts have, however, little success, and the ultimate aim for obtaining pure edge current remained unachieved. In the 2D counterpart, the challenge extends to the lack of material diversity. Moreover, although HgTe/CdTe\textsuperscript{73} and InAs/GaSb\textsuperscript{72} are demonstrated by transport measurement to show QSH behavior, for other measurements, these samples are proven inap.

A jump-start to the field was offered by various engineering principles of TIs. There have also been considerable efforts to engineer TIs by proximity induced SOC in graphene through adatoms,\textsuperscript{75} or with interfacing with transition metal di-chalcogenides,\textsuperscript{74,75} and in artificial heterostructures.\textsuperscript{5,16,76–78} and in optical lattices.\textsuperscript{79–81} Artificially grown heterostructures can be very versatile as they offer higher materials flexibility and tunability. The present author has proposed a number of design principles for engineering quasi-1D,\textsuperscript{82} 2D,\textsuperscript{83} as well as 3D TIs\textsuperscript{5,16} by decorating atoms or layers. The basic idea lies in generating chiral orbits in a periodic fashion without breaking TR symmetry in 2D plane. Each chiral orbit thereby encloses a pseudo-magnetic field (Berry curvature) in the same sense as in the TKNN language and commence quantized Chern number.\textsuperscript{91} In real-space, electron-electron interaction induced translational symmetry breaking can lead to chirality inversion between different sublattices, leading to a new type of Z\textsubscript{2} quantum order, which is associated with topological invariant and edge states.\textsuperscript{90} The 3D generalization follows similarly in that the adjacent layers has opposite SOC, such that they serve as TR partners to each other. With a band inversion, a 3D TI naturally arises.\textsuperscript{5} A heterostructure of even and odd orbitals is also a fertile setup to generate Dirac materials.\textsuperscript{16} The second part of this article discusses various forms of these engineering principles.

TI has been reviewed extensively in various forms. Three review articles are published in the Review of Modern Physics.\textsuperscript{23,63,82} Topological band theory\textsuperscript{63,83} and topological field theories\textsuperscript{14} are also discussed extensively. A materials repository can be found in Refs.\textsuperscript{63,64} Topological superconductors are reviewed in Refs.\textsuperscript{23,84,85} QSH\textsuperscript{86} and QAH\textsuperscript{41} insulators are also reviewed separately. Reviews of Dirac and Weyl materials can be found in Refs.\textsuperscript{87,88} Two books are available to understand the basics of various TIs.\textsuperscript{89,90}

2 Theories of Topological Invariants

Topological band theory encompasses broadly defined computational schemes for the calculations of various topological invariants in systems with or without TR symmetry, inversion symmetry, particle-hole symmetry, and mirror symmetry. Depending on larger number of symmetries present in the Hamiltonian, the computation of the corresponding topological invariant is accordingly simplified. Both Chern number and Z\textsubscript{2} invariants are defined by single particle wavefunction, and can be calculated using the tight-binding or Wannier wavefunction within the DFT. For weakly interacting systems, such calculation can be extended to the quasiparticle spectrum within a Fermi-liquid or mean-field theory. For strongly correlated systems, the Chern number can be calculated by using the self-energy dressed Green’s function.\textsuperscript{91,92} In addition to rigorous calculations of Chern number or Z\textsubscript{2} invariants, there are also other simplified methods that can be used for preliminary diagnosis of a potential TI. For examples, one can determine the non-trivial TI by simply counting the odd number of band inversions at the TR symmetric k-points, or employing the adiabatic (band gap-) continuity between a known TI and an unknown insulator, and/or bulk-boundary correspondence.

We start this section with the historical development of topology, which forms the basis for modern topological invariant. Then we discuss various topological invariant calculations based on the number of symmetries present in a given case. Across most of the methods discussed below, some unified concepts can be excavated, which combine diverse formalisms of topological invariant. Among them, we have (i) Complex phase associated with electron hopping with quantized phase winding; (ii) Formation of cyclotron or chiral orbits in...
A Pedagogic Review on Designing Model Topological Insulators

Journal of the Indian Institute of Science
VOL 96:2 Apr.–Jun. 2016 journal.iisc.ernet.in

periodic lattice (Gaussian curvature); (iii) Odd number of exchange of electron’s complex phase, or chirality or TR conjugate in half of the BZ; (iv) Search for bulk property (s) which can lead to accumulation of charge (IQH), or spin, (QSH), or TR polarizibility ($\mathbb{Z}_2$ class). In the sections below, we discuss such corresponding basic principles used for the derivation of various topological invariants.

2.1 Gaussian curvature, Euler number, and genus

Inception of the concept of invariants can be traced back to the times of Leonhard Euler and Carl Friedrich Gauss in the 18th century. Gaussian curvature can be a good starting point to build up the discussion. It is defined by the product of two principle curvatures ($\kappa_1, \kappa_2$), along any two perpendicular directions at a given point (see Fig. 1) as $K = \kappa_1 \kappa_2$. If a principle curvature has a minimum (maximum) at the point (convex and concave curvatures), we assign its value to be $-1$ ($+1$). In this sense, if a point on the surface has a minima or maxima in both principle directions, then the corresponding Gaussian curvature is $+1$ [Figs 1(c–d)]. The outer and inner surfaces of a sphere provide the corresponding examples, both being topologically equivalent. On the other hand, if a point simultaneously possesses maximum and minimum in the two principle directions, the Gaussian curvature yields $-1$ [Fig. 1(e)]. The camel’s back or a torus is non-trivial Gaussian curvature with $K = -1$.

Euler characteristic (also known as Euler number) dictates that the flux through a Gaussian curvature is always quantized as

$$\oint_S K dS = 2\pi \nu, \quad \text{where } \nu = \text{integer.}$$

The above integral formula can also be expressed in terms of the Gauss-Bonnett formula, giving a topological invariant, called genus ($g$), relating the Euler characteristic as $\nu = 2 - 2g$. The Euler characteristic for a sphere is 2, giving $g = 0$ [Fig. 2(a)]. The same for a torus or Möbius strip is 0, with $g = 1$ [Fig. 2(b)]. Thus, the former geometry is attributed as topologically trivial, while the later (torus, Möbius strip) as non-trivial curvature. Double torus and a three–hole pretzel have Euler characteristic as $-2$, and $-4$, with $g = 2$, and 3, respectively [Fig. 2(c)]. From these examples, it is evident that the genus or the topological invariant is related to the number of holes present in a Gaussian geometry. Another
important observation can be made here that the number of holes or genus also dictates the number of distinct surface states. These results constitute the key mechanism for the emergence of topological invariants in the quantum and condensed matter world.

2.2 Laughlin’s argument and TKNN invariant

In simple Hall effect, as the magnetic field is applied perpendicular to the lattice, a potential gradient arises perpendicular to both applied magnetic field and applied current. This is because, due to magnetic field, the electron and holes feel opposite Lorentz forces and move towards different edges of the lattice (Fig. 3(b)). The transverse conductivity \( \sigma_{xy} \) initially increases linearly with the magnetic field strength (Fig. 3(a)). But with further increase of the field, the Hall conductivity becomes quantized and exhibits plateau with magnetic field, and the field, the Hall conductivity becomes quantized for the bulk topological invariant. We recall that here the Hall conductivity is itself a topological invariant: \( \sigma_{xy} = ve/h \), where \( v \) is called TKNN or Chern number. Therefore, from the general Kubo formula for conductivity, we can obtain our first definition of a topological invariant or the Chern number for the \( n \)th band \( (\nu_n) \) as

\[
\nu_n = i \sum_{k, q \neq 0} \left( f(E_{n,k}) - f(E_{n,k}) \right) \times \left[ \left( \psi_n^*(\mathbf{k}) \frac{\partial H}{\partial \mathbf{k}} \psi_n(\mathbf{k}) \right) \left( \psi_{n'}^*(\mathbf{k}) \frac{\partial H}{\partial \mathbf{k}} \psi_{n'}(\mathbf{k}) \right) \right] \left( E_{n,k} - E_{n,k} \right)^2
\]

where \( E_{n,k} \) is the eigenvalue of the Hamiltonian \( H \), \( \psi_n(\mathbf{k}) \) is the Wannier wavefunction, and \( f(E_{n,k}) \) is the Fermi-Dirac distribution function.

2.3 Quantum Hall calculation in arbitrary parameter space

The above formula is based on the variation of the Hamiltonian in two orthogonal momentum directions, and thus implicitly assumes a periodicity of the lattice. In some cases, as in disordered lattice, where a proper unit cell is...
difficult to define, the above formula apparently fails. However, Niu, Thouless and Wu generalized their TKNN invariant calculations to any two arbitrary parameter space, by implementing the so-called "twisted boundary condition". They introduced two fictitious parameters $\alpha, \beta$ (which does not require to have any physical relevance) and demanded that the lattice is periodic under them as

$$\psi(x_i + L_1) = e^{i\alpha L_1} e^{i(dh/\hbar)} y_i \psi(x_i),$$

$$\psi(y_i + L_2) = e^{i\beta L_2} \psi(y_i).$$

(4)

Here $x, y$ are the lattice site indices, and $L_1, L_2$ represent the system size. In this case, the velocity operators can be rewritten as $v_x = \frac{\partial H}{\partial k_x} \rightarrow \frac{\partial H}{\partial \alpha}$, and $v_y = \frac{\partial H}{\partial k_y} \rightarrow \frac{\partial H}{\partial \beta}$, and the QH invariant can be calculated in the $(\alpha, \beta)$ space by using Eq. (3). This generalization unravels an important insight that as the Hamiltonian is adiabatically varied in any closed parameter space, it gives the same topological invariant. This implies that as the wavefunction of the particle returns to its starting points, the expectation value of the velocity operators remains the same, but the wavefunction itself acquires an additional phase. This phase turns out to be the Berry phase, proposed independently.\(^{95}\)

The expectation value of the velocity operator in any parameter space is a most important factor for topological invariant, with the only requirement for a given parameter space being that it has to be periodic (Gaussian curvature), but not necessarily a physical parameter. This is crucial for the conceptualization of the 'chiral orbit' we use for the QSH effect which implies that 'chiral orbit' can be a mathematical object which can be 'created' in any periodic parameter space for the calculation of the topological invariant in 2D systems. For the same reason, when a system is driven periodically with time, the corresponding time evolution of the Hamiltonian (Floquet Hamiltonian) can also give rise to a 'Berry phase' in the time-domain and lead to QH or topological phase.

### 2.4 Berry connection and curvature

The above section discussed how an applied magnetic field’s flux quantization leads to the IQH as computed within the Kubo formula. Now we can reverse our derivation, and start with the Kubo formula version of the topological invariant in Eq. (3), and define a band dependent ‘magnetic field’ in the momentum space $F(k)$. We again demand that its flux in the reciprocal space is quantized:

$$\nu_n = \frac{1}{\Omega_{BZ}} \int_{BZ} F_n(k) \cdot N \,dk = \oint_{\partial BZ} A(k) \cdot dl,$$

(5)

where $\Omega_{BZ}$ is the BZ phase space area, and $N$ is the normal unit vector to the 2D BZ. In the last step, we have employed the Stokes’ theorem, which allowed us to define a momentum space ‘vector potential’ as $F_n(k) = V \times A_n(k)$. The formalism for $F(k)$ is simply the right-hand side of Eq. (3), and that for $A(k)$ can also be obtained subsequently. Another elegant formalism for $F$ and $A$ can be obtained by using the identity:

$$\partial_k \psi_n^\dagger (k) = \sum_{n' \neq n} \frac{\langle \psi_{n'}(k) | \frac{\partial H}{\partial k} | \psi_n^\dagger (k) \rangle}{E_n - E_{n'}},$$

(6)

which yields:

$$F_n (k) = i \langle \nabla_k \psi_n (k) | \times | \nabla_k \psi_n (k) \rangle,$$

(7)

$$A_n (k) = i \langle \psi_n (k) | \nabla_k \psi_n (k) \rangle.$$

(8)

To further elucidate the physical significance, we refer back to Eq. (5), which can be compared with the Peierls phase in real space, acquired by a charged particle moving in a magnetic field, $\varphi = \int A(\mathbf{r}) \cdot d\mathbf{l}$, where $A(\mathbf{r})$ is the vector potential, and $c_1$ and $c_2$ are starting and end points of the path. This implies that the topological invariant here is a momentum space ‘Peierls phase’ (equivalent to Aharonov-Bohm phase) acquired by the electron in traversing a closed path in the reciprocal space under an intrinsic gauge field $A(k)$. In this sense, $v_x$ is called the Berry phase,\(^{95}\) and $A(k)$ as the Berry connection, while $F(k)$ is the Berry curvature. Note that the Berry connection is gauge-dependent, and therefore topological invariant formulas [such as axion angle formalism in Eq. (19) below] involving $A(k)$ does not give unambiguous result. On the other hand, the Berry curvature is gauge invariant and observable. Therefore, for a band which possess a well-defined Berry phase in a close trajectory in the momentum space (translational symmetry is assumed as above), it possesses an intrinsic Chern number, and therefore, can give rise to a IQH effect without the application of an external magnetic field.

In the IQH effect, the applied magnetic field provides the ‘chirality’ for the electrons to form cyclotron orbit. Without magnetic field, the QH phenomenon can be thought of occurring in a reverse fashion. Here, a self-generated chirality of electrons creates a pseudo-magnetic field (Berry curvature) in the process of forming chiral orbits. In solid state systems, such intrinsic chirality can stem from a multiple origins, including SSH.

**Aharonov-Bohm**: The Aharonov–Bohm effect, sometimes called the Ehrenberg–Siday–Aharonov–Bohm effect, is a quantum mechanical phenomenon in which an electrically charged particle is affected by an electromagnetic potentials (scalar or vector potentials), despite being confined to a region in which both the magnetic field $\mathbf{B}$ and electric field $\mathbf{E}$ are zero.

--- Exerts from Wikipedia.
type staggered electron hopping,\textsuperscript{15} sublattice (as often referred to pseudospin) symmetry in the hexagonal lattice,\textsuperscript{96} SOC,\textsuperscript{3} or certain type of even-odd orbital texture mixing.\textsuperscript{16,17} In simpler terms, the chirality arises if the electron hopping is complex, because it naturally accompanies a phase associated with electron’s hopping. As the $k$-space magnetic field or the Berry curvature threads through a periodic lattice (Gaussian curvature), the Euler characteristic ensures a quantization of the flux [Eq. (1)], and bulk topological invariant arises. The intrinsic formation of ‘chiral orbit’ in a periodic lattice is the foundation of TR invariant TIs, which however, have different interpretations and mathematical expositions such as Pfaffian nodes, chiral vortex, momentum-space monopoles etc. as we uncover below.

### 2.5 Chern number as zero magnetic field topological invariants

If an 1D chain is made of two inequivalent sublattices, the hopping between the two sublattices becomes complex as used in the SSH model.\textsuperscript{15} According to the above prescription, the emergent chiral hopping can be associated with a winding number. Recall that in a QH state, since the cyclotron orbits cannot complete a full circle at the edges, it leads to the edge current, and thus, Hall effect arises. Something similar happens in 1D chain with chiral state. Since a localized chiral orbit cannot be assumed in 1D, the chirality of the electrons can be thought of as charge current following across the chain, allowing electrons and hole to be accumulated in opposite ends. This phenomenon naturally arises if we solve Eq. (5) with open boundary condition, which gives topologically protected polarizibility at the ends. This is called the Zak phase,\textsuperscript{97} which is a topological invariant (discussed further in Sec. 2.13 below).

E.D.H Haldane\textsuperscript{1} realized that such complex hopping can be easily obtained in 2D honeycomb lattice for the same reason, namely, due to the presence of two inequivalent sublattices. Each sublattice form triangular lattice, which are oppositely aligned, see Fig. 4. The low-energy Hamiltonian of a honeycomb lattice can be written in terms of the $2 \times 2$ Pauli matrices ($\sigma$) entangled with linear momentum, in which two sublattices provide the pseudospinor basis\textsuperscript{96}

$$H(k) = d(k) \cdot \sigma,$$

\text{where $d_{x,y}(k)$ contains linear-in-$k$ term, and $d_i$ gives the Dirac mass. Such Hamiltonian is analogous to the Dirac equation in 2D, and forms Dirac cone in the absence of the Dirac mass term. For such simplified Hamiltonian, the Chern number can be calculated from the d-vectors itself. Starting from Eq. (3) and substituting (9), one obtains}

$$\nu = \frac{1}{4\pi} \int_{BZ} d^2k \frac{d \cdot \left( \partial_{x_i} d \times \partial_{k_y} d \right)}{|d|^3}.$$

For the usual Dirac Hamiltonian, $d_i$ components are proportional to $k_i$ (where $i = x, y$), and therefore, it is easy to see that the Chern number is proportional to the $d_i$ term. $d_i$ term stems from the onsite energy difference between the two sublattices, and intrinsically remains zero.

A key ingredient is still missing here. Note that honeycomb lattice provides an imaginary hopping term between different sublattices, but the intra-sublattice hopping still remains real [Fig. 4]. Therefore, electrons hopping within each triangular sublattice do not have any chirality, and fail to form our desired chiral orbits. For a remedy, Haldane affixed an ‘extrinsic’ gauge field (but not a magnetic field) to the intra-sublattice hopping. Additionally, he imposed the condition that the ‘gauge field’ has different signs for different sublattices, such that the resulting chiral orbits for them are counter-propagating. Therefore, they treat opposite flux and the net magnetic field effect remains zero. However, the staggered ‘gauge field’ naturally induces different onsite energy to different sublattices, and therefore, $d_i$ term becomes finite, and Eq. (10) gives a finite Chern

\begin{figure}
\centering
\includegraphics[width=\textwidth]{figure4.png}
\caption{Haldane model. The net inter-sublattice hopping $t_\alpha e^{i\alpha}$ remains complex, since its complex conjugate hopping $t_\alpha e^{-i\alpha}$ is absent. On the other hand, the net intra-sublattice hopping $t_\alpha e^{i\alpha} + t_\alpha e^{-i\alpha} = 2 \cos (kb)$ becomes real. Haldane added an extrinsic phase ($\phi$) into the hopping as $t_\alpha e^{i\alpha+\phi}$ with opposite phase for different sublattices (blue and red). Therefore, the corresponding flux in the blue and red triangles are equal but opposite.}
\end{figure}
number. This signifies that two oppositely rotating triangular chiral orbits are split by a negative Dirac mass. Haldane’s proposal was important for the conceptual development of the QH effect without magnetic field, but for decades, it was assumed to be ‘unphysical’, since obtaining the required gauge field without magnetic field was not feasible. Very recently, researchers have successfully generated Haldane model by commencing time-dependent Hamiltonian with periodic pumping. The periodic time evolution naturally gives a Bloch Phase in the time-space, which provides Haldane’s gauge-field.\textsuperscript{96}

2.5.1 Spin Chern number: Kane and Mele turned on SOC to obtain Haldane’s ‘gauge-field’, which does not break TR symmetry.\textsuperscript{4} This gave birth to the TR invariant QH effect and eventually \( Z_2 \) TIs. They considered two copies of Haldane’s Hamiltonians for spin-up and spin-down states to form a Block diagonal Hamiltonian:

\[
H(k) = h_\uparrow (k) \oplus h_\downarrow (k).
\]

The Hamiltonian respects both TR and inversion symmetry with

\[
h_\uparrow (k) = -h_\downarrow (-k).
\]

The resulting Hamiltonian can be expressed in the \( 4 \times 4 \) Dirac matrix basis as

\[
H(k) = d(k), \quad \Gamma_i
\]

where \( d \)-vector has five components and the corresponding \( \Gamma \) (including identity matrix) are the usual Dirac matrices (Kane and Mele used few additional cross-terms, which we do not discuss here for simplicity). The \( k \)-dependence of each \( d_i \) component (\( d_0 \) and \( d_4 \) contain even power of \( k \), while others contain odd power) complements the symmetry of their corresponding \( \Gamma_i \) matrices to preserve the TR symmetry. In the case when two valence bands are fully spin-polarized, the Chern number for each band corresponds to different spin states \((\nu_\uparrow, \nu_\downarrow)\). Due to SU(2) symmetry \( \nu_\uparrow = -\nu_\downarrow \), which means two equal and opposite chiral orbits are stabilized due to the spin-momentum locking, as shown in Fig. 5. Therefore, the total Chern number \( \nu = \nu_\uparrow + \nu_\downarrow \) vanishes, while their difference, namely spin Chern number, \( \nu' = \nu_\uparrow - \nu_\downarrow \), becomes finite. Therefore, according to Kane and Mele, there is no charge pumping to the edge, but there is a net spin pumping, and hence, they refer the corresponding state as QSH effect.

Bernevig, Hughes, Zhang (BHZ) predicted that the Quantum Well (QW) state arising in the HgTe/CdTe heterostructure commences 2D QSH insulator above some critical thickness.\textsuperscript{97} They realized that the band structure for HgTe and CdTe are completely inverted across the Fermi level. In particular, the SOC split bands with \( \Gamma_8 \) and \( \Gamma_6 \) symmetries, respectively, constitute conduction and valence bands in HgTe, while they form valence and conduction bands in CdTe. Therefore, if we derive a 2 label Hamiltonian and expresses in terms of Pauli matrices as in Eq. (9), we immediately find that the Dirac mass \( d_i < 0 \) for HgTe and \( d_i > 0 \) for CdTe. Therefore, if one makes an HgTe/CdTe heterostructure, at their boundary \( d_i \) must vanish, which means gapless Dirac fermions emerge here. Based on this idea, they proposed a \( 4 \times 4 \) Hamiltonian using the Kramers pairs of \( \Gamma_8 \) and \( \Gamma_6 \) levels. The Hamiltonian is also block diagonal with each block representing different spin state as in Kane-Mele model. The Dirac mass inversion (which is same as band inversion or parity inversion as we will discuss in Sec. 2.7 below) guarantees that each block gives equal but opposite Chern numbers, and QSH insulator arises. Due to the block diagonal nature of the mode, it is popularly known as half-BHZ model.

When two spin states cannot be separated to assign individual Chern number, the present method does not work. Kane and Mele proposed more rigorous method to calculate the \( Z_2 \) invariant using TR ‘polarization’ as discussed in Sec. 2.6. As we will go along, we will learn more techniques and interpretations of various topological invariances.

2.5.2 Mirror Chern number: In the cases, where the band inversion occurs at non-TR symmetric points, a distinct topological invariant can be obtained if the system possess mirror symmetry.\textsuperscript{3,21,53} Let us consider a case where \( k_m \) represents a mirror plane in the BZ with the corresponding mirror operator \( (\mathcal{M}) \), defined by

\[
[H(k_m), \mathcal{M}] = 0.
\]

In such a case, the mirror plane can be decomposed into two subspaces, denoted by \( \pm \mathcal{M} \). Then, as in the case of half-BHZ model, the present Hamiltonian on the mirror plane can be split into two blocks, coming from two sub-space as

\[
H(k_m) = h_\uparrow (k_m) \oplus h_\downarrow (k_m).
\]

Each block (\( h_\uparrow \)) gives equal but opposite Chern number (due to TR symmetry). Therefore, their difference \( \nu' = (\nu_\uparrow - \nu_\downarrow)/2 \) leads to a finite value, called mirror Chern number. The corresponding TI family is refereed as topological crystalline insulator.\textsuperscript{21}
Given that the low-energy Hamiltonian here has mirror symmetry, the leading term in the edge state will have even power in momentum along this direction. Let us consider an example of a mirror plane at \( k_x = 0 \), which dictates \( H(k_x, k_y, k_z) = \{ -k_x, k_y, k_z \} \). Since a linear term in \( k_x \) violates this condition, it will drop out from the Hamiltonian. Therefore, the corresponding surface state will be quadratic. Along other directions lacking a mirror symmetry, the edge/surface state can be linear in momentum. In fact, the surface states of the topological crystalline insulator Sn_{0.5}PbTe^\text{83} contain both quadratic and linear bands.

### 2.6 \( Z_2 \) invariant and time-reversal polarization

In the case of TR breaking IQH insulator, Chern number can take any arbitrary value. However, this is not the case in TR invariant TIs. For such cases, spin or mirror Chern number can take only 0 or 1 \((\text{mod } 2)\) value, and thus, the topological invariant is represented by a more general \( Z_2 \) invariant.\(^4,6\) \( Z_2 \) invariant becomes equal to spin or mirror Chern number in cases where latter is defined, but there exist other methods of evaluating it. Although spin and mirror Chern numbers are observables via QH effect, \( Z_2 \) invariant is not a directly measurable quantity in the bulk, and is often diagnosed by the observation of topological surface state.

For TR symmetric cases, Kramers degeneracy at the TR invariant \( k \)-points dramatically reduces the full momentum space calculations into only TR invariant \( k \)-points. The antiunitary TR operator \( \Theta \) imposes the symmetry in the Hamiltonian as \( \Theta H(k) \Theta^{-1} = H(-k) \). Let us consider a system where the spin-rotational symmetry is broken, say due to SOC, without breaking the TR symmetry as \( \Theta|k, \uparrow\rangle = -|k, \downarrow\rangle \). The high-symmetric points \( k^* \), which are invariant under TR symmetry \([\text{e.g. } (0,0,0), (\pi,0,0), (0,\pi,0), (\pi,\pi,0), (\pi,\pi,\pi)\] etc. in a cubic lattice in Fig. 5\) are special. Here \( k^*_L \) and \( -k^*_L \) points are the same \([\text{up to a reciprocal lattice vector: } k^*_L = -k^*_L + G]\), so they must be spin degenerate. This is called the Kramers’ degeneracy.

Among many methods available for the evaluation of the \( Z_2 \) invariant, Fu-Kane-Mele method is often easier to implement, especially in the cases where both TR and inversion symmetries are present.\(^7\) To understand this method, we draw analogy with some of the properties of IQH, QSH states discussed above. In these insulators, it is the bulk Chern number that induces charge or spin polarizations, respectively, at the edge. Kane and Mele asked a similar question: What bulk property for TR invariant \( Z_2 \) class can pump a similar ‘polarization’ to the boundary. In QSH effect, opposite spins with opposite momentum, due to SOC, are pumped to the edge, requiring that the electron exchanges its spin in traversing half of the BZ odd number of times. Since opposite spins with opposite momentum are just the TR conjugate to each other, a more fundamental property to exchange in \( Z_2 \) TI is the TR partner of electrons. Based on this analogy, Kane, Mele proposed a mathematical concept, called ‘TR polarization’, in which they argued that electrons with one Bloch wavefunction and their complex conjugate partner are accumulated at the edge.\(^6,9\) This requires that the electron switches its TR partner odd number of times in traversing half of the BZ [green line in Fig. 4(a)]. [For systems with inversion symmetry, it is equivalent to the odd number of parity, or equivalently the Dirac mass or just simply band inversion in half of the BZ, as discussed in Sec. 2.7.]

They subsequently quantified this hypothesis\(^4,6,7,9\) by defining the matrix element of the TI operator between a Bloch state \( u(k) \) and its TR conjugate \( u^*_m(-k) \) in the Fermi sea, and constructed an antisymmetric, unitary matrix, with components \( w_{mn}(k) = \langle u^*_m(-k)|\Theta|u_n(k)\rangle \). We define \( P(k) = \text{Pf}[w(k)] \). For many TI Hamiltonians dealing with \( SU(2) \) spin, the filled state is two-fold degenerate, especially when inversion symmetry is present. Therefore, the above matrix-element is a \( 2 \times 2 \) matrix, in which the Pfaffian is just the off-diagonal term (the formula for topological invariant is, however, general to any number of filled bands).
Based on the value of $P(\mathbf{k})$, the BZ can be split into the ‘even’ and ‘odd’ subspaces. In the even subspace, $\Theta|\mu_+(\mathbf{k})\rangle$ is proportional to $|\mu_-(\mathbf{-k})\rangle$, making $|P_{\mathbf{k}}\rangle = 1$. In the odd subspace, $\Theta|\mu_-(\mathbf{k})\rangle$ is orthogonal to $|\mu_-(\mathbf{-k})\rangle$, implying $P(\mathbf{k}) = 0$, which is important for topological invariance. Let us assume that $\pm \mathbf{k}^*$ is a pair of points that are TR partners, where $P(\mathbf{k}) = 0$ [see Fig. 4(a)]. The phase of $P(\mathbf{k})$ about each of these points winds up in opposite directions (equivalent to having counterpropagating ‘chiral orbits’ in the momentum space). If $\pm \mathbf{k}^*$ coincides with any TR invariant point $k_0^*$, the two $k$-space ‘chiral orbits’ annihilate each other. Again, if there are even number of such points, say, $\pm k_1^*, \pm k_2^*$, then unless $k_1^*, k_2^*$ are protected by some additional symmetry, they can also annihilate each other by scattering or perturbation. But a single pair of $\pm \mathbf{k}^*$ does not have the option to scatter to another $k$-point, except to the corresponding $\mp \mathbf{k}$-points, which however, requires the corresponding spin to flip. Since spin flip is prohibited by the TR symmetry, such nodal points at $\pm \mathbf{k}^*$ remain protected from TR invariant perturbations.

Similarity, between the Berry connection formalism, and the Pfaffian $P(\mathbf{k})$ can be rigorously shown. Differentiating $w(\mathbf{k})$, and using the unitary property of the $w$ matrix, we obtain the Berry connection in terms of $w(\mathbf{k})$, and $P(\mathbf{k})$ as $\nu$ [using Eq. (8)]

$$A(\mathbf{k}) = -\frac{i}{2} \text{Tr}[w(\mathbf{k})^\dagger \nabla_k w(\mathbf{k})]$$
$$= -\frac{i}{2} \text{Tr}[
abla_k \log w(\mathbf{k})]$$
$$= -\frac{i}{2} \nabla_k \log \det w(\mathbf{k})$$
$$= -\frac{i}{2} \nabla_k \log |P(\mathbf{k})|$$

(11)

Based on this, $\nu$ invariant can be defined by calculating the winding number of the $P(\mathbf{k})$ over a single $k$-space ‘chiral orbit’, in a contour enclosing half of the BZ (so that only either $+\mathbf{k}$ or $-\mathbf{k}$ is included), as shown by green boundary in Fig. 4(a). So, from Eq. (5), $\nu$ invariant can be evaluated as:

$$\nu = \frac{1}{2\pi i} \oint_{C} d\mathbf{k} \cdot \nabla_k \log |P(\mathbf{k}) + i\delta|,$$

(12)

where a complex term $i\delta$ is introduced to evaluate the above integral in a complex contour plane. This simplifies the integral into a residue problem, with singularities occurring at the loci of $P(\mathbf{k}) = 0$.

Given that there should be an odd number of Pfaffian nodes in half of the BZ, it is expected that the corresponding nodes would occur on the high-symmetric $k$-directions. This simply means that $P(\mathbf{k})$ should change sign odd number of times on both sides of the nodes at the TR invariant $k$-points [see Fig. 4(b)]. Therefore, the calculation simply reduces to counting the sign of $P(\mathbf{k})$ at the TR invariant momenta in the first quadrant of the BZ only. If we take the product of the sign of the Pfaffian at all TR invariant points, and the result comes out to be negative, then there must be odd number of zeros in the first quadrant of the BZ. Since $\{\text{Pf}(w)\}^2 = \det (w)$, the sign of the Pfaffian can be defined in a formal way as

$$\delta_j = \frac{\sqrt{-\text{det} \left[ w(k_j^*) \right]}}{\text{Pf} \left[ w(k_j^*) \right]} = \pm 1.$$

(13)

Therefore, in a 1D system, the TR polarization can be defined as $(-1)^x = \delta_1 \delta_2$, where $\delta_1$ and $\delta_2$ are evaluated at the two TR invariant points. If $\delta_1$ and $\delta_2$ have opposite signs, we get the $\nu_1$ invariant $\nu = 1$, which signals the non-trivial topological phase. The formula generalizes to a higher dimensions as

$$(-1)^x = \prod_{j=1}^{N_1} \delta_j,$$

(14)

where $N_1$ is the total number of the TR invariant momenta in the first quadrant of the BZ. In a 2D square lattice, $N_1 = 4$, while in a 3D $C_{10}$ symmetric lattice $N_1 = 8$. If there are odd number of $\delta_1 = -1$ in this $k$-space, the right hand side of the above equation gives $-1$, which therefore yields $\nu = 1$, a non-trivial topological invariant. This is called the strong topological invariant (denoted by $\nu_j$). Note that for any arbitrarily large odd number of $\delta_1 = -1$, topological invariant remains $\nu_1 = 1$, otherwise 0. Therefore, unlike in IQH insulator, where arbitrarily large Chern number is possible, here one only gets two values of $\nu_1$, and the $Z_2$ symmetry emerges.

In some cases, there can be total even number of $k_j^*$-points with $\delta_j = -1$, but they lie in different planes (say on the $k_z = 0$, and $k_z = \pi$ planes) [see Fig. 5(c)]. Thus, those 2D planes contain odd number of $\delta_j = -1$, and constitute non-trivial 2D TIs, while the 3D system remains trivial TI. This is called the weak TI. Since in 3D, there are three orthogonal coordinate axes, there are three weak topological invariants ($\nu_1$, $\nu_2$, $\nu_3$). Fu, Kane, and Mele,\(^a\) thereby, introduced four $Z_2$ invariants ($\nu_{1,2,3,4}$) for 3D TIs. This part is explained with examples in Fig. 5 and discussed further in the following section.
2.7 $\mathbb{Z}_2$ calculation with inversion symmetry

If the system possesses inversion symmetry, in addition to TR symmetry, the calculations of topological invariants becomes exceptionally simpler. Suppose $\mathcal{P}$ is the parity operator defined by $\mathcal{P} (k, \uparrow) = - (k, \uparrow)$, under which the Hamiltonian transforms as $H(-k) = \mathcal{P} H(k) \mathcal{P}^{-1}$. Inserting $\mathcal{P}^2 = 1$ in the expression for $w_{\text{inv}}(k)$, and employing the identity that $[H, \mathcal{P} \Theta] = 0$, $\delta$ parameter at the TR invariant moment $k^*_s$ can be evaluated as

$$\delta_s = \prod_{m=1}^{N} \xi_m(k^*_s), \quad (15)$$

where $\xi_m(k^*_s) = \pm 1$ is the parity eigenvalue at $k^*_s$ defined as $\mathcal{P} u_m(k^*_s) = \xi_m(k^*_s) u_m(k^*_s)$. The product is computed for $N$ filled bands. In practice, one does not have to include all filled bands in the calculation, rather only those bands that undergo band inversion at the TR invariant points. For the two typical Dirac Hamiltonians that are expressed in terms of $2 \times 2$ Pauli matrices or $4 \times 4$ Dirac matrices, the parity term turns out to be $\sigma_z$, or $\Gamma_4 = \sigma_z \otimes 1_{2 \times 2}$, respectively. The corresponding $k$-dependent $d$-vector component can often be written as $d_s(k^*_s) = (\varepsilon_1(k^*_s) - \varepsilon_2(k^*_s))/2$, where $\varepsilon_1(k^*_s)$ and $\varepsilon_2(k^*_s)$ are the conduction and valence bands near the Fermi level. Therefore, the parity eigenvalue of the Hamiltonian is determined simply by $\xi_m(k^*_s) = \text{sgn} \left[ d_s(k^*_s) \right]$. For systems with both inversion and TR symmetries, the $\mathbb{Z}_2$ invariant is obtained by simply counting the number of band inversion at the high-symmetric momenta as

$$(-1)^{\nu} = \prod_{s=1}^{N} \text{sgn} \left[ \varepsilon_1(k^*_s) - \varepsilon_2(k^*_s) \right]. \quad (16)$$

In other words, if there are odd number of band inversions at all TR invariant momenta in an insulator, the system acquires non-trivial topological behavior with the same odd pair of edge or surface states. Interestingly, $d_s$ term is the mass term in the Dirac equation; that means parity inversion is equivalent to Dirac mass inversion from positive (trivial) to negative (non-trivial) value. For systems, where parity is not a good quantum number, the Dirac mass inversion serves as a good measure of the topological phase transition as often used in the literature (see Fig. 5).

The reason for the popularity of this method lies in the fact that identifying the parity of a given band within the tight binding model or Wannier method is quite straightforward. In most cases, it is nothing but knowing the orbital character of the valence and conduction bands. So band inversion simply refers to switching orbital character between these two classes. However, band inversion does not mean that an orbital entirely switches its position between conduction and valence bands at all $k$-points, rather it has to be done only at odd number of TR $k$-points, and not at other $k$-points. For example, in Fig. 7(a), if the odd parity conduction band drops fully below the Fermi level, the system still remains topologically trivial. This means the inter-orbital overlap matrix-element has to be strongly momentum dependent. Simple local inter-orbital hopping or crystal field splitting or onsite interactions such Hubbard $U$ or Hund’s coupling are often not adequate to commence such a $k$-dependent band inversion. SOC does this job in most of the known TIs.

Some caution has to be taken for the cases when a band at a given TR invariant $k$-point is not fully orbitally polarized, rather it contains a mixture of both even and odd orbitals. In such cases, the band inversion mechanism cannot be considered as conclusive. In this context, a term called band inversion strength is often used, which measures the amount of orbital weight is exchanged between the conduction and valence bands. Band inversion strength is also used as a measure for Dirac gap at the TR $k$-points [see, for example, Ref.]. Associated with the orbital weight transfer, the band topology also changes in this process. For example, if the top of the valence band has an upward curvature, it changes to a downward curvature around the TR $k$-point after the band inversion. This structure is sometimes referred as a ‘dent’ in the band structure, which is seen in the DFT band structure, as well in the experimental data.

Three representative examples for trivial, strong, and weak TIs are given in Fig. 7. Owing to TR symmetry, it is sufficient to consider only the first quadrant of the BZ to count the number of band inversions, since the other $k$-points are related to them by TR symmetry. As mentioned earlier, if the conduction and valence bands possess the same parity at all $k$-points, but different among them, Eq. (15) suggests that it is a trivial topological insulator, or not a topological insulator at all [Fig. 7(a)]. Fig. 7(b) depicts the case of a single band inversion at the $\Gamma$-point, indicating a strong TI ($\nu = 1$). In this case, all three surfaces of the lattice possess Dirac cones with the vertex of the cone lying at the same $k$-point where the band inversion has occurred. If the band inversions occur even number of times
in the first quadrant, a weak TI can be obtained if one or more BZ sides possess odd number of band inversions. For example, in Fig. 7(c), we consider the case of two band inversions at the \( \Gamma \) - and at \((\pi,0,0)\)-points, yielding \( \nu_0 = 0 \). But the \( k_x = 0 \) and \( k_x = \pi \)-planes contain only single band inversion, and the corresponding topological invariant becomes \( \nu_1 = 1 \). Each \( k \)-point where bands are inverted, a metallic Dirac cone arises at the corresponding edge or surface. (c) When even number of band inversions occurs, it does not render a strong 3D TI (\( \nu_0 = 0 \)). But those 2D planes accommodating odd number of band inversions become 2D TIs (such as \( k_y = 0 \) and \( k_y = \pi \) planes in this example): \( \nu_0 : \nu_1 : \nu_2 : \nu_3 = 0 : 1 : 0 : 0 \).

Figure 7: Band inversion and non-trivial TI. According to Eqs. (14 and 15), topological invariant is determined by the total odd number of inversions of the Pfaffian sign or parity (\( \delta \)) in the valence band. In this figure, the choice of signs of \( \delta \) for conducting and valence bands and the cubic lattice structure are completely arbitrary and for illustration purpose only. (a) There is no band inversion and thus it’s a trivial band insulator. Even when the conducting band drops below the Fermi level at all \( k \)-points, the system still remains topologically trivial. (b) When the band is inverted at odd number of TR invariant \( k \)-points, one obtains strong TI (\( \nu_0 = 1 \)). Each \( k \)-point where bands are inverted, a metallic Dirac cone arises at the corresponding edge or surface. (c) When even number of band inversions occurs, it does not render a strong 3D TI (\( \nu_0 = 0 \)). But those 2D planes accommodating odd number of band inversions become 2D TIs (such as \( k_y = 0 \) and \( k_y = \pi \) planes in this example): \( \nu_0 : \nu_1 : \nu_2 : \nu_3 = 0 : 1 : 0 : 0 \).

2.8 \( \mathbb{Z}_2 \) calculation without inversion symmetry

Subsequently, Fu and Kane have generalized the \( \mathbb{Z}_2 \) calculation for systems without inversion symmetry:\(^9\)

\[
\nu = \frac{1}{2\pi} \int_{\tau} d\tau \left[ \int_{\tau} A_n(k) dk - \int F_n(k) d^2k \right], \quad (17)
\]

\( \tau \) is half of the BZ, one in which the Berry curvature is to be computed, while \( d\tau \) is the boundary where the integral of the Berry connection is to be calculated (see Fig. 8). The difference between Eq. (17) and Eq. (5) is that here an additional surface integral over the Berry connection is present. This term appears in the process of gauge fixing as follows.

In systems with finite Chern number, the center of the cyclotron orbit or 'chiral orbit' poses an obstruction to smoothly affix a gauge to the wavefunction; because the phase of the wavefunction is supposed to acquire a discontinuity at the center of the orbit to commence finite phase winding or Chern number. If the wavefunction has a smooth gauge at all \( k \)-points, both \( A \) and \( F \) acquire the
same gauge, yielding \( \nu = 0 \) by Stokes’ theorem. For \( Z_2 \) TIs, although the Chern number is zero, but a similar obstruction arises at the nodes of the Pfaffian where the wavefunction switches to its TR conjugate [see Sec. 2.6]. This affirms that one cannot simultaneously obtain a node in the Pfaffian and a smooth gauge in the wavefunction.

However, Fu and Kane proposed \(^{99}\) that if the contour is restricted within half of the BZ, see Fig. 8, and also we choose a gauge that is periodic, i.e., \( \varphi\left(k\right) = \varphi\left(k + G\right) \), in addition to the TR symmetry, the problem can be solved, see discussions in Refs. \(^{100–102}\). In this case, Eq. (17) can be solved by converting the integral into summation over a uniform discretized \( k \)-mesh on the region \( \tau \). Imposing the TR symmetry on the boundary \( \partial \tau \), we can obtain a link matrix as \( W^m_n \left( k \right) = \left\langle u_m \left( k \right) | u_n \left( k + \delta k_i \right) \right\rangle \), and the unimodular link variable \( \Delta_\tau \left( k \right) = \det W^m_n / \det W^m_{mn} \), where \( \delta k_i \) is the step of the mesh in the direction of the reciprocal vectors \( G_i \left( G_j \right) \). Then we define the gauge potential as \( A_\tau \left( k \right) = \log \Delta_\tau \left( k \right) \), which gives \( F_\tau \left( k \right) = \partial \varphi / \partial k_i - \partial \varphi / \partial k_j \), then the \( Z_2 \) invariant can be calculated as

\[
\nu_\tau = \frac{1}{2 \pi} \sum_{k \in \partial \tau} A_\tau \left( k \right) - \sum_{k \in \partial \tau} F_\tau \left( k \right).
\]

For systems with inversion symmetry, this formula cannot be used since \( A \) and \( F \) terms cancel each other.

The 3D generalization of this calculation follows similarly.\(^{10}\) A 3D BZ has six inequivalent 2D planes at \( k_i = 0, \pi \), where \( i = x, y, z \). Using Eq. (18), one obtains six topological invariants on these six planes, namely \( \nu_{0,\tau} \). Among which, the strong topological invariant condition implies that \( \nu_{0,\tau}^2 - \nu_{0,\tau}^2 = \nu_{0,\tau}^2 = \nu_{0,\tau}^2 \). Therefore, as in Sec. 2.6, we get four independent topological invariants defined by \( \nu_0 = \nu_{0,\tau}^2, \nu_1 = \nu_{0,\tau}^2, \nu_2 = \nu_{0,\tau}^2, \nu_3 = \nu_{0,\tau}^2 \).

### 2.9 Axion angle as topological invariant

According to the axion electrodynamics,\(^{103}\) electric and magnetic fields can couple linearly, giving rise to an additional term in the Maxwell’s action

\[
S_\theta = \left( \frac{e}{4 \pi} \right) \int d^4 x dt \varepsilon_{\mu \nu \sigma \tau} F_{\mu \sigma} F_{\nu \tau}.
\]

The action becomes

\[
S_\theta = \left( \frac{e}{32 \pi} \right) \int d^4 x dt \varepsilon_{\mu \sigma \tau} F_{\mu \sigma} F_{\nu \tau}.
\]

As in the case of IQH effect, \( \theta \) becomes quantized to be 0 or \( \pi \) (mod \( 2\pi \)) if the system is TR invariant. For \( Z_2 \) TI, \( \theta \) reflects the same topological invariant as \( \theta = \nu \pi \). Therefore, \( \theta = 0 \) indicates a topologically trivial state, while \( \theta = \pi \) signifies a non-trivial topology. Interestingly, \( \nu \) which changes by integer values, \( \theta \) changes continuously from \( \pi \) to 0 as TR symmetry is lifted by spontaneously introducing magnetic moment. Again, unlike the \( Z_2 \) invariant, axion field \( \theta \) is an observable quantity, since it gives rise to a topological electromagnetic effect. Therefore, TI with small magnetic moment belongs to a new class, namely topological axion insulator. So far, this state is proposed in few materials,\(^{25,107,108}\) but not realized yet. As the magnetic moment is further increased, another class of TI, called quantum anomalous Hall (QAH) insulator, may arise if the system undergoes a similar band inversion. Here, a \( Z_2 \) classification is destroyed, and the system can, in principle, possess arbitrarily large value of Chern number. QAH state is proposed and realized in various engineered structures,\(^{61,109–111}\) and \( \text{LaX} \ (X = \text{Br, Cl, I}) \) is the only family predicted so far as intrinsic QAH insulator.\(^{25}\)
2.10 Topological invariant for interacting fermions

For non-interacting systems, the topological invariant can be extracted from the Kubo formula for the Hall conductivity in 2D [Eq. (3)]. For interacting systems, one can follow the same strategy. For such systems, a single-particle wavefunction cannot be defined, and thus we start from a different Kubo formula for interacting systems as \( \sigma_{xy} = \frac{e^2}{h} \text{Im} \frac{\partial}{\partial \omega} K(\omega + i\delta) \). The current-current correlation kernel \( K(\omega + i\delta) \) can be expressed in terms of the interacting Green's function \( G(k, i\omega) \) as

\[
K(i\omega) = -\frac{1}{\Omega_{BZ}} \sum_{k,\nu} \text{Tr} \left[ v_{\nu} G(k, i + iv) v_{\nu} G(k, i\omega) \right],
\]

where \( \Omega_{BZ} \) is the phase space volume, \( \beta = 1/k_B T \), and the velocity vertices are \( v_{\nu}(k) = \partial H(k)/\partial k_{\nu} \). For \((2 + 1)\)D systems, the vertex can be expressed within the Ward identity as \( v_{\nu}(p) = \partial G^{-1}(p)/\partial p_{\nu} \), where \( p = (k, i\omega) \). Integrating over the \((2 + 1)\) D phase space volume, we obtain a similar topological invariant (Chern number) in terms of generalized Green's function as

\[
v = \frac{\pi}{6} \int \frac{d^3 p}{(2\pi)^3} \text{Tr} \left[ e^{i\nu p} \frac{\partial G^{-1}}{\partial p_{\rho}} G \frac{\partial G^{-1}}{\partial p_{\rho}} \frac{\partial G^{-1}}{\partial p_{\sigma}} \right].
\]

(21)

Extension to higher dimensions follows the same procedure, in which the number of vertices is equal to the dimension of the system. Clearly, Eq. (21) also applies to non-interacting Green's function \( G_n(k, i\omega) = (i\omega - H(k))^{-1} \). For the multi-orbital systems, corresponding Green's function is a tensor. Electron-electron interaction or disorder effect can be incorporated within the Dyson, or the T - matrix formalism, among others, giving a generalized formalism \( G_n(k, i\omega)^{-1} = G_n(k, i\omega)^{-1} - \Sigma(k, i\omega) \), where \( \Sigma \) is the self-energy correction. Dynamical mean-field theory (DMFT), and momentum-resolved density fluctuation (MRDF) theory are two widely used methods to explore the dynamical correction effects, with the later method having an added advantage of incorporating the full momentum dependence of the correlation effects. For many systems such as transition metal oxides, di-chalcogenides, intermetallics, lanthanum and actinide compounds, the momentum dependence of the electron-correlation is significantly strong, which can lead to a characteristic change in the Berry curvature \( \mathbf{F}(k) \). These features can be captured within the MRDF method.

2.11 Topological invariants for superconductors

Superconductivity is a correlated phenomenon, which arises due to the condensation of electron-electron pair (Cooper pair) in the low-energy spectrum. Within the mean-field theory, the corresponding Hamiltonian can be casted into a single particle (quasiparticle) Hamiltonian in which the superconductivity opens a band gap at the Fermi level. Therefore, although in the two electrons picture, the system is superconducting (SC), in the single electron effective model it represents an insulator (assuming the SC gap opens everywhere on the Fermi surface). Interestingly, fully gapped superconductor, and topological insulator share an analogous Hamiltonian, and thus many of the topological concepts also apply in the former case. The single-band Hamiltonian for a superconductor can be expressed exactly by Eq. (9), with \( d_+ = \epsilon(k), d_- \) are the real and imaginary parts of the SC gap, \( \Delta(k) \). Furthermore, the owing to the criterion for the formation of 'chiral orbit' or 'chiral vortex'; the SC gap must be a chiral pairing symmetry, which is often obtained in \( p \)-wave superconductors (for spinful superconductors, this condition can be relaxed if SOC is present). The chiral \( p \)-wave superconductors have odd parity gap symmetry and breaks TR symmetry. In such a case, the topological invariant is obtained by the sum of the first Chern number \( \nu_n \) on each bands weighted by the sign of the gap as

\[
\mathcal{N} = \frac{1}{2} \sum_n \nu_n \text{sgn}(\Delta_{nk}).
\]

(22)

For spinful case, TR invariant topological superconductor can be obtained if the pairings \( \{ \psi_{k\uparrow}^\dagger \psi_{k\downarrow} \} \) and \( \{ \psi_{k\downarrow}^\dagger \psi_{k\uparrow} \} \) have opposite chirality, i.e., \( \Delta_{\uparrow\uparrow} = p_+ + ip_\sigma \) and \( \Delta_{\downarrow\downarrow} = p_- - ip_\sigma \). Here again, if the corresponding \( 4 \times 4 \) Hamiltonian can be split into the block diagonals, as in the case of half-BHZ model for QSH insulator, we can apply the same Chern number calculation to evaluate the topological invariant. Due to the associated particle-hole symmetry, the zero energy boundary modes must be a Majorana mode which means, its eigenstate must be real.

2.12 Adiabatic continuity

Adiabatic continuation is a simple and powerful tool to identify a non-trivial TI with reference to another known TI, if both these systems MRDF: MRDF method is introduced by the present author to calculate the correlated electron properties of materials residing in the weak to intermediate coupling regime.

Majorana: Majorana mode results from the solution of a special Dirac equation whose wavefunction is real. Therefore, these particles or fermions do not have any anti-particle component. Copper pairs of particle below the Fermi level always have hole-hole pair counterparts above the Fermi level. Thus those pairs which live at zero energy act as both particle and hole, and can be thought of Majorana excitations.
are adiabatically connected. Here ‘adiabatic connection’ simply means that as one transforms a non-trivial TI ‘A’ into another material ‘B’ by continuously changing the atomic number of the constituent elements, the bulk band gap of the ‘A’ system does not close and reopen in this whole process, then they are adiabatically connected or belong to the same non-trivial TI class.

The band evolution between a non-trivial TI to a trivial TI is shown in Fig. 9. Suppose with a given tuning, such as chemical potential, or atomic number, or pressure, the band gap simply decreases, without any band inversion, as in going from Fig. 9(a) to 9(b), then these two systems are adiabatically connected and belong to the same topological class. At the topological critical point, when the bulk band gap closes at the TR invariant point, it produces a bulk Dirac cone (also refereed as 3D Dirac cone), Fig. 9(c). Note that graphene is a non-trivial system lying at the topological critical point. Above the critical point, as the bands are inverted back, the system transforms into a trivial insulator. Such band evolution including the 3D Dirac cone formation is seen experimentally in BiTl(S$_{1-\delta}$Se$_\delta$)$_2$.

The difficulty with this method is that one requires to study the band topology by continuously changing the atomic number, i.e. by doping, which is not an easy calculation within the first principles methods. Yet, the method has been successful in predicting a number of materials, especially those that do not have inversion symmetry, by starting with a nearby known TI which has inversion symmetry. Lin et al. have shown\cite{20,37} that one can test a large class of materials by adiabatically changing the atomic number of the constituent elements, which is done by alloying or doping. For example, we can start with a hypothetical system with three elements MM$'$X, with nuclear charge $Z_{M'} = 47 + x$ and $Z_X = 51 - y$, respectively, where x and y are adjustable parameters. x and y are not necessarily integers, but the choice must maintain the charge neutrality. This mapping can start with $x = 0$ and $y = 0$, which corresponds to Li$_2$AgSb, and end with $x = 3$ and $y = 1$, which corresponds to the artificial compound He$_2$SnSn. Li$_2$AsSb is known to be a non-trivial TI. With increasing x, and y, we obtain Li$_2$AgBi, Li$_2$AuBi, and Li$_2$CdSn which are non-trivial topological metals. On the other hand, the end element Li$_2$CuSb is a trivial band insulator in which the bulk band gap has reopened above the critical point.

2.13 Bulk-boundary correspondence and surface states

Spontaneous (continuous) symmetry breaking leads to gapless Goldstone mode in the corresponding excitation spectrum. For example, the spin rotational symmetry breaking in quantum magnets leads to gapless magnons in the spin excitation spectrum, or the translational symmetry breaking in the formation of a lattice renders gapless phonon mode (acoustic modes). Similarly, as the cyclotron orbits or ‘chiral orbits’ become periodically arranged in a lattice (creating a Gaussian curvature), breaking the translational symmetry, which is associated with the emergence of non-trivial bulk topology, it manifests into gapless edge states at the boundary. Although a rigorous calculation to validate this premise is yet not explored, however, the application of Goldstone theory for the realization of bulk-boundary correspondence can be intriguing. For example, electromagnetic response of TIs stipulates two dynamical axion modes, one of them is gapless Goldstone-like mode, and another is Higgs-like gapped mode, as shown in earlier calculation.\cite{106}

According to the bulk-boundary correspondence of TI, the bulk topological
invariant dictates the number and characteristics of edge states at the boundary. For the case of IQH effect, the N Chern number prescribes N chiral edge states. For the spin or mirror Chern numbers, edge states form in pair; for example, N spin Chern number has 2N counter-propagating chiral edge states. The same principle also applies to $Z_2$ topological invariant with some modifications. Kane-Mele proposed that for TR invariant $Z_2$ TI, the TR partners are accumulated at different sides, leading to ‘TR polarization’. In the presence of additional inversion symmetry, TR polarization is equivalent to ‘parity polarization’, in which different parity states are pumped to different sides. This also implies that the edge or surface state must contain both TR partners, which means that the edge states must come in pairs. However, both states must also be degenerate at the TR invariant k-points, forming a Dirac cone, which is guaranteed by the $Z_2$ invariant in the bulk.

With the application of perpendicular magnetic field, electrons and hole are pushed towards opposite side of a plane, leading to Hall effect. The theory of bulk-boundary correspondence for TIs can be developed similarly, by studying the effect of the momentum space ‘magnetic field’ or Berry curvature in the boundary. This is precisely what is done by Zak, by calculating the Berry phase in a finite 1D chain to obtain the so-called Zak phase. Zak\textsuperscript{97} assumed a bulk periodic boundary condition in the momentum space, which yields \[ \gamma_n = \frac{2\pi i}{a} \int_{-a/2}^{a/2} dk \sum_k u_{nk}(x) \frac{\partial u_{nk}(x)}{\partial k} dx, \tag{23} \] where $u_{nk}(x)$ is the periodic part of the Bloch state, and a is the lattice constant. By using the corresponding Wannier function $a_n(x)$, he arrived at a spatial dependence of the Berry phase as
\[ \gamma_n = \frac{2\pi}{a} \int_{-\infty}^{\infty} x |a_n(x)|^2 dx. \tag{24} \]

We notice that the above equation is identical to that of the electric polarizability when multiplied with electric charge. As the inversion symmetry is imposed, $\gamma_n$ takes the values of either 0 or a$\ell$/2. At both ends, the integral range survives either $-\infty$ to 0, or 0 to $\infty$. Since the integrand is odd function of x, the integrals obtains $\pm \gamma_n$ at the $\pm \ell$ ends, respectively. By multiplying elementary charge e, we see that the polarization at the two ends have opposite signs, implying that the electrons and holes are separated into different ends of the 1D chain. This is referred as charge pumping in 1D TI [see Fig. 8(a)]. In the 2D IQH effect, similarly, electron and hole states are accumulated at the two edges, see Fig. 8(b). Since a single edge state has to smoothly connect both the electron (filled band) and hole (empty band) states, it has to pass through the neutral or zero energy mode. Therefore, the zero energy edge state is guaranteed by the bulk topology, and cannot be destroyed by any weak perturbation or disorder, as long as bulk topology remains intact.

For the case of QSH insulator, opposite spin Chern numbers dictate opposite charge pumping to a given edge. For example, if the spin-up state drives electrons toward the +x direction, then spin-down state will pump holes to the same side. This cancels the net charge in each edge but allows a net spin accumulation, see Fig. 8(c). Therefore, here always a pair of counter dispersive and spin-polarized edge states arises, which meet at the TR invariant momentum, $k_s^\ast$, owing to Kramers’ degeneracy. Let us assume an edge state along the $k_s$ direction. Since any band dispersion can be expanded via Taylor series around the $k_s$ as: $\epsilon(q,v_s) = \epsilon_0(q,v_s^\ast)$, with $q_s = k_y - k_{y_s} \ll 1$, and $v_s^\ast$ is the band velocity. Since $\epsilon$ is a constant term, we can neglect it. TR symmetry demands $\Theta|u_q(q,v_s)| = |u_q(-q,v_s)|$, implying that if we set $\epsilon_0 = q,v_s^\ast$, the band for the opposite spin must be $\epsilon_1 = -q,v_s^\ast$, or vice versa. For an electron below the Fermi level, its spin is interchanged as it traverses to the other side of $k_s^\ast$, or as $q \rightarrow -q$. Equal and opposite effect simultaneously occur for the hole state. Therefore, if we express our edge Hamiltonian in the spinor $(\psi_\uparrow, \psi_\downarrow)$, the minimal, low-energy Hamiltonian for an 1D edge state with eigenvalues $\pm q,v_s^\ast$ is:
\[ H_{edge} = v_s^\ast q_s \sigma_z. \tag{25} \]

The corresponding edge state is schematically shown in Fig. 10(c) (lower panel).

We can understand the emergence of 2D surface state in 3D TI in a similar way, with the idea of ‘TR polarization’.\textsuperscript{99} Let us start with a stack of 2D QSH insulators placed along the x-direction, each of which contains topological edge states along the y-direction. Now as we turn this into a ‘strong’ 3D TI, the spin pumping (or TR polarization, in general) must occur along both $y_s$ and $x$-directions. This constraints the surface Hamiltonian to have 2D spin-polarization, and can therefore be written in the same spinor $(\psi_\uparrow, \psi_\downarrow)$ as [see Fig. 10(d)]
\[ H_{surf} = v_s^\ast q \cdot \sigma. \tag{26} \]
Note that this surface Hamiltonian is analogous to the Rashba-type SOC 
\[ h_R = \frac{k^2}{2m} \sigma \times (q \times \sigma) \]
where \( \sigma \) is the SOC strength and the other symbols have the usual meanings. Indeed, owing to the loss of inversion symmetry at the surface, the surface states actually originate from the Rashba-type SOC, with a crucial difference. In the topological surface state, the kinetic energy of each spin \( \frac{k^2}{2m} \) is much smaller compared to the SOC strength, \( \frac{k^2}{2m} \alpha_R \). This is a manifestation of the bulk-boundary correspondence of the TI. For the Rashba-band, in Fig. 11(a), each spin band bends backward away from the TR momentum as the kinetic energy term dominates over SOC term at higher momenta. The Rashba-bands are not required to connect to the bulk states at all. On the other hand, topological surface states are required to adiabatically connect to both the bulk conduction and valence bands, and thus they cannot bend backward and rather disperse monotonically across the bulk insulating gap. This is illustrated in Fig. 11. It is obvious that all edge/surface states are not topological surface state. More importantly, even in a TI, all surface states are not necessarily topological surface state as well. A topological surface state must have a Dirac cone at the TR point where band inversion occurs in bulk state, and also has to connect...
both the bulk valence and conduction states. There might be situations where an apparently isolated topological state may arise, which is not connecting any other state. However, if one study the eigenvector or weight of the band, it may appear that the weight has changed from surface to bulk as one moves away from the Dirac node.

3 Engineering Topological Insulators

The above introduction highlights some key ingredients universally relating different TI families. The foremost ingredient is that electrons must contain a chirality, i.e., complex hopping term, which is obtained either via applied magnetic field (Pierels phase), or in bipartite lattice (such as SSH), or graphene lattice, or via SOC, or in orbital selective lattice, or even by artificial gauge fields as often done in optical lattice systems. These chiral states must also form cyclotron or chiral orbits, with or without magnetic field. The ‘chiral orbits’ naturally produces a k-space ‘magnetic fields’, or Berry curvature. As these ‘chiral orbits’ are arranged in a 2D lattice, the system can be represented by a Gaussian curvature (torus as in the case of IQH system), each threading counter propagating integer multiple of quantum flux of the Berry curvature. In Sec. 2.3, we also discussed that the ‘chiral orbits’ does not have to be formed in a physical parameter space, such as real or momentum space, but can also be formed in any generalized parameter space as long as they obey periodic boundary conditions. For such case, the Hall conductance can be precisely calculated. Therefore, for the cases where two chiral states can be fully separated in the Hamiltonian (as in the case of half-BHZ model, for example), each state can be assigned with spin or mirror Chern numbers, as applicable, with their net value vanishes, but the difference yields finite value. For the generalization to evaluating $Z_2$ invariant, odd pair of Pfaffian nodes is analogous to the center of odd pairs of ‘chiral orbits’ in the momentum space. Here one requires that the electron’s wavefunction must change into its TR conjugate odd number of times as it traverses through half of the BZ. These ingredients are used in the following methods as the targets to engineer TIs.

The idea would be to take a bottom-up approach to assemble TIs in desired dimensions. Here we start with an atomic chain or layer with SOC, and invert the SOC in its adjacent chain or layer, so that it acts as the TR conjugate partner to the former one. Therefore, the electron switches its TR partner in hopping from one chain/layer to another. The Pfaffian $P(k)$ thereby acquires a node in between them, and the $Z_2$ invariant becomes 1. This setup can be achieved by aligning the direction of SOC in each chain or layer by manipulating lasers (in optical lattice) or by reversing electron field directions (for RashbaSOC), or via interaction. As an insulating state occurs, the system is guaranteed to behave as a $Z_2$ TI.

In systems with inversion symmetry, $Z_2$ invariant can be evaluated by the odd number of band inversions between the two chiral states at the TR invariant momenta. This phenomena can be thought of as the 'chiral orbits' are split along the energy axis and/or in the momentum space with negative Dirac mass. At the end of this section, we will introduce another new concept for chiral band inversion in the real-space, due to interaction effect. We attribute the corresponding emergent topological phase as quantum spin-Hall density wave (QSHDW) insulator. Our research group has a major thrust in this research direction, among which we give below four representative examples.

3.1 Engineering topological ‘chiral orbits’ in 2D

In the first example, we start with a 1D chain of atoms with 1D SOC. Such state arises in quantum wires, optical lattices, as well as in bulk systems when Rashba- and Dresselhaus-type of SOCs have equal strength. We denote such a wire by ‘A’ SO wire. The corresponding Hamiltonian in the continuum limit is

\[ H_A = \frac{k^2}{2m} I_{2\times2} + i\alpha_k k_x \sigma_x, \]

where the first term gives the kinetic energy, and the second term gives the SOC with strength $\alpha_k$. (In optical lattice, in the process of creating SOC with lasers, a Zeeman-like terms also arise which has the form of $\Omega \sigma_y$ where $\Omega$ depends on the laser strength). The corresponding spin split state is shown in Fig. 12(b). So far researchers are only able to generate 1D SOC in optical lattice which is inadequate to create ‘chiral orbit’ in 2D plane, and thus poses a serious setback to obtain 2D or 3D TIs here. We propose to use the second SOC wire (called ‘B’ wire) with opposite SO such that its Hamiltonian can be written as $H_A(k) = H_A(-k)$. Without quantum tunneling between the two wires, the spin-polarization of bands for two wires are reversed, and the $\Gamma$-point has now four-fold degeneracy (see Fig. 10(c)). Therefore, bands at the $\Gamma$-point can now be gapped without breaking the TR symmetry (spin degeneracy is still

QSHDW: The word density wave is used here in analogy with the traditional density wave orders in which a well defined spin or charge density is modulated in the lattice and have a wavelength which is usually larger than the underlying lattice constant. Here the Hall density refers to the Hall conductivity for each spin which has a modulation vector larger than the lattice constant.
present), by switching on the inter-wire quantum tunneling, as shown in Fig. 12(d). Note that in this setup we have obtained chirality along the $k_x$ direction through SOC, but the same thing is still lacking for hopping between the wires. This can be obtained by allowing staggered inter-wire hopping, in which the distance between 'A' and 'B' wires and that between 'B' and next 'A' wires ('A' wire lying on opposite sides of 'B' wire) can be made different.

For this setup, the formation of localized 'chiral orbit' can be understood from the corresponding real-space view of the hoppings, see Fig. 13. Let us assume that a spin-up electron in the 'A' wire is right-moving, and that in the 'B' wire becomes left-moving, due to opposite SOC. Therefore, when a spin-up electron hops from 'A' wire to 'B' wire, its motion is reversed and as the electron hops back to the 'A' wire, it forms a 'chiral orbit' (reverse direction for the spin-down electron) as shown in Fig. 13(a). As the hopping amplitude is increased, these 'chiral orbits' become localized in the bulk, leading to a band insulator, as shown in Fig. 11(b). Due to the translational symmetry in the lattice, the TKNN invariants could have been assigned to each chiral state, with equal and opposite Chern numbers. However, due to SOC, the two chiral states switch TR partner across $\Gamma$-point. Therefore, we can calculate the 'TR polarization' as discussed in Sec. 2.6. Finally, with the calculation of Pfaffian, we show that the present engineered device becomes a non-trivial TI, with spin polarized edge state as shown in Fig. 12(e). We find that the resulting Pfaffian takes the form of $P(k) = 1 + e^{-i k_y}$, which therefore, depends on the Bloch phase associated with the inter-wire hopping $e^{i k_y}$. The Pfaffian has one pair of nodes inside the first BZ at $k_y = \pm \pi/2$. Therefore, irrespective of the parameters, the present setup is guaranteed to produce a non-trivial TI.

Figure 12: Band evolution from free fermion to TI in our first engineering principle in 1D system described in Sec. 2.1. (a) Free electron dispersion without SOC. (b) A typical Rashba-SOC split band structure in a single wire. (c) Two decoupled wires without any quantum tunneling between them. Since two wires have opposite SOC, they render opposite spin splitting. Therefore, the combination presents a spin degenerate band structure at all k-points. (d) As the quantum tunneling between the two wires is turned on, a band gap first opens at the $\Gamma$-point, without breaking the spin-degeneracy at all k-points. (e) As more pairs of TR coupled wires are joined in a lattice, a full band gap develops at k-points giving an insulating state with Dirac cone at the edge.

Figure 13: Real space view of 'chiral orbit' formation in 2D with two TR conjugate wires. (a) A atomic wire decorated with two orthogonal lasers (E1 and E2) to produce a 1D SOC $\alpha(k)$. In the adjacent 'B' wire, E1 laser is reversed to obtain TR conjugate SOC $\alpha^*(k)$. If the spin up state is left handed in 'A' wire, it becomes right handed in 'B' wire and vice versa. As a spin up atom/electron hops from 'A' to 'B' wire it will move backward, and then eventually hops back to 'A' wire. As the SOC strength and the hopping strength are tuned to their proper values, two counter-helical 'chiral orbits' form in the bulk and an insulating state commences. (b) As a 2D lattice (Gaussian curvature) is formed for the 'chiral orbits,' a non-trivial TI state is guaranteed to arise here with helical edge states.
3.2 Engineering 3D topological insulators with Rashba bilayers

Next we discuss our method of engineering 3D TIs by stacking 2D layers in a heterostructure, in which alternating layers have opposite SOCs. Here we start with a 2D electron gas (2DEG) with Rashba-type SOC $\alpha(\mathbf{k})$. The approach is based on growing bilayer of Rashba-type 2DEG with opposite SOC on adjacent planes of bilayers. We find that in the stack of bilayers grown along the (001)-direction, a topological phase transition occurs above a critical number of Rashba-bilayers, with the formation of a single spin-polarized Dirac cone at the $\Gamma$-point.

The band progression from trivial insulator to non-trivial TI is demonstrated in Fig. 14. The building block is a Rashba bilayer with opposite Rashba SOC in each layer denoted by $\alpha(\mathbf{k})$ and $\alpha^*(\mathbf{k})$. We tune the interlayer distance such that an anisotropic quantum tunneling, $D(\mathbf{k})$, couples them, as illustrated in Fig. 14(b). Here again, without breaking the TR symmetry, the Rashba-bilayer opens an insulating gap at the $\Gamma$-points, which is determined by $D(0)$. Then we add another Rashba-bilayer on top of the previous one with an inter-bilayer electron hopping, $t_z$, which is required to be different from $D(0)$ to eliminate the degeneracy related to the number of bilayers. For two Rashba bilayers, the bulk bands in the two interior single layers have reduced band gap, and a massive 'preformed' Dirac-like surface state appears, see Fig. 14(c). However, our parity analysis reveals that this setup still remains trivial insulator without any band inversion. However, as we add one more bilayer, a magic topological phase transition occurs, see Fig. 14(d). We see in the band structure that there exists a bulk band inversion between the valence and conduction bands at the $\Gamma$-point [indicated by arrows between Figs. 14(c), and 14(d)]. The band inversion can be easily visualized, confirmed by the parity analysis, from the change of curvature of the valence band near the $\Gamma$-point (or a 'dent' band structure). While the valence band is hole-like at all other momenta, it changes the topology to become electron-like at the $\Gamma$-point. Despite non-trivial band topology, the Dirac cone in the surface state obtains a tiny band gap due to finite size effect. As we grow a large size heterostructure with more Rashba bilayers, we see that band inversion strength in the bulk increases and the gap in the surface state gradually vanishes, see Fig. 14(e).

The low-energy effective model for a single Rashba-bilayer can be expressed as

$$H_{BL}(\mathbf{k}) = \begin{pmatrix} \varepsilon(\mathbf{k}) & \alpha(\mathbf{k}) & 0 & 0 \\ \alpha^*(\mathbf{k}) & \varepsilon(\mathbf{k}) & D(\mathbf{k}) & 0 \\ 0 & D(\mathbf{k}) & \varepsilon(\mathbf{k}) & \alpha^*(\mathbf{k}) \\ 0 & 0 & \alpha(\mathbf{k}) & \varepsilon(\mathbf{k}) \end{pmatrix}.$$  

"Figure 14: Engineering 3D TI. (a) Two decoupled TR conjugate Rashba-states with opposite SOC. (b) A coupled Rashba bilayer (1BL) with anisotropic inter-layer hopping [$D(\mathbf{k})$]. (c) Two Rashba bilayers setup, with inter-BL coupling $t \neq D(\mathbf{k})$ This is a trivial insulator. (d) 3 BL is guaranteed to give non-trivial TI with band inversion (indicated by the red dashed lines). Here due to finite size effect, two surfaces hybridize and open a gap at the Dirac cone. (e) Above about 6 BLs, a gapless Dirac cone arises."
with $\mathbf{D}(k) = (D_0 + D_1 k^2)$: The leading term in the
momentum dependence is quadratic which is easily
obtained due to ‘in-plane’ next-nearest-neighbor hopping, but
between the adjacent layers.

$\mathbf{v}$ The estimate of the
Fermi velocity does not
include correlation effect
which tends to reduce the velocity.

The estimate of the Fermi velocity does not include correlation effect which tends to reduce the velocity.

The estimate of the Fermi velocity does not include correlation effect which tends to reduce the velocity.

with $\mathcal{E}(k) = k^2/2m$, the intra-bilayer hopping
$\mathcal{D}(k) = (D_0 + D_1 k^2)$. Each such bilayer is now
stacked along the $z$-direction, which are connected by
the nearest-neighbor inter-bilayer hopping $T = t_{2D}^{\alpha\beta}$. Interestingly, the resulting insulating bulk gap depends on the two tunneling terms $D_0$ and $t_2$ which are readily tunable. The surface state is determined by $H_{\text{surf}} = \alpha_y (\sigma_k - \sigma_{k^*})$, with its surface Dirac fermion velocity fully controlled by the Rashba SOC strength. For values of the Rashba-coupling constant as large as $3.8$ eV A, achieved to date in bulk BiTeI, we get $v = 5.8 \times 10^5$ ms$^{-1}$, which is much larger than the highest speed achieved so far in 3D TIs as $v = 3 \times 10^5$ ms$^{-1.5}$.

Such Rashba-bilayer can be easily manufactured by creating a potential gradient between two 2DEGs with the help of gating, or by inserting oppositely polarized ferroelectric substrate between them, among others. This idea is applied in the GaAs/Ge/GaAs heterostructure with opposite semiconductor interfaces. The giant electric field generated by charge accumulation at the interfaces creates a Rashba-bilayer on both sides on Ge layer, and that allows a band inversion with an insulating gap of 15 meV or larger. Another example is the synthesis of bulk strong TI by stacking 2D weak TIs: Bi$_{14}$Rh$_2$I$_{16}$. Its Bi–Rh sheets are graphene analogues, but with a honeycomb net composed of RhBi$_2$ cubes rather than carbon atoms. The strong bismuth-related SOC renders each graphene-like layer a TI with a 2,400 K bandgap. The non-centrosymmetric BiTeCl is an experimentally realized 3D TI in which alternating layers of Bi and Te obtains opposite charge polarization, and thus opposite SOC. Therefore, our design principle is not only applicable in engineered heterostructure, but is also at play in bulk single crystals.

3.3 Spatial modulation of ‘chiral orbits’ and quantum spin Hall density wave insulator

Next we discuss an interesting situation where the energy level inversion occurs in real space between sublattices due to electron-electron interaction. To understand this phenomenon, we revisit the nearly free-electron model from the elementary condensed matter course. We know that when a weak periodic potential $V_{Q}$, where $G$ is the reciprocal wavevector, is applied to free electron gas, band gap opens at those $k$-points which obey $2k = G$ relation, see Fig. 15(a–b). In this case, those parts of the bands, $\epsilon(k)$, which lie above $2k > G$ are folded back into the reduced BZ by explicitly including it in the Hamiltonian as $\epsilon(k + G)$. The resulting system breaks translational symmetry, and thus opens a band gap at the zone boundary by $V_{Q}$. A similar translational symmetry can occur due to electron-electron or electron-phonon interactions generated potential at a preferential wavevector $V_{Q}$, where $Q$ is called the Fermi surface nesting vector, defined as $2k_p = Q$, where $k_p$ is the Fermi momenta. As a result, charge or spin density of the electrons at each lattice site becomes modulated and obtain a new periodicity which is different from the original periodic lattice. Similarly, at the edge of the reduced BZ [at $k = Q/2$], a quasiparticle gap opens. Here the gap is defined by the Landau-like order parameter, and the corresponding states are called charge/spin-density wave orders.

An analogous, but more exotic, situation arises when the Fermi surface nesting occurs between the two SOC split bands. As shown in Fig. 15(c–d), in such cases, the nesting simultaneously occurs for both spin states, giving rise to the spin-resolved potentials $V_{Q}^{\uparrow\downarrow}$ (unless TR symmetry is broken, the absolute values of the two potential can be

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{fig15.png}
\caption{Upper panel: BZ folding and band gap opening in the nearly free electron model. (a) Free electron dispersion in which $k$ and $k$ are connected by the reciprocal vector $G$ with a periodic potential $V_{G}$. (b) Band gap opens at the zone boundary and the bands above the zone is folded inside by doubling the unit cell. Lower panel: A similar situation arises with Rashba-SOC. (c) Two FS nestings simultaneously occur here between the two helical states ($V_{G}$, $V_{G}$). (d) Corresponding correlated band structure with SODW. Here the reduced BZ for the two helical states are different (red and blue shadings).}
\end{figure}
equal). For such a case, there will be two spin-density waves, with each spin density having opposite spin-polarization at a given site, due to SOC. In such a case, a distinct order parameter, namely Spin–Orbit Density Wave (SODW) arises, which breaks translational symmetry, but in most cases it preserves TR symmetry. A detailed discussion of this order parameter can be found in Refs. 31,32,59,130,136 Here we discuss how such a state can naturally give a $Z_2$ topological order parameter.

We illustrate this case for a Rashba-type SOC. Inside the reduced zone, the non-interacting band, and the Rashba SOC are $\epsilon(k)$, and $\alpha(k)$, and their folded counterparts are $\epsilon(k + Q)$, and $\alpha(k + Q)$. An interesting case arises when the nesting vector is exactly $Q = (\pi, 0)/(0, \pi)$. In this case, the folded Rashba SOC changes to the complex conjugate of the main SOC as $\alpha(k + Q) = \alpha^*(k)$. [This becomes obvious if we use the lattice form of the Rashba SOC where it takes the form of $\alpha(k) = \alpha_0 \sin k \sigma + \sin k \sigma$, which changes to $\alpha^*(k)$ as $k \rightarrow k + Q$]. This gives rise to a situation that is analogous to the Hamiltonian for TI written in Eq. (28). To see that we can express a mean-field Hamiltonian in the basis of $\Psi_k = \{\psi_{k+1\uparrow}, \psi_{k+1\downarrow}, \psi_{k+Q\uparrow}, \psi_{k+Q\downarrow}\}$ as

$$H = \begin{pmatrix}
\epsilon(k) & \alpha(k) & 0 & \Delta(k) \\
\alpha^*(k) & \epsilon(k) & \Delta^*(k) & 0 \\
0 & \Delta(k) & \epsilon(k + Q) & \alpha^*(k) \\
\Delta^*(k) & 0 & \alpha(k) & \epsilon(k + Q)
\end{pmatrix}.$$

(29)

By comparing Eq. (28) with Eq. (29), we see that the single-electron tunneling term $D(k)$ in Eq. (29), is replaced here by the interacting SODW gap $\Delta(k) = V\langle \psi_{k+1\uparrow}\psi_{k+1\downarrow}\rangle$, with $V$ being the interaction strength. The above Hamiltonian can also be expressed in terms of the Dirac matrices, in which the coefficient for the Parity operator $(\Gamma)$ is the Dirac mass term $M = (\epsilon(k) - \epsilon(k + Q))/2$. Therefore, the topological invariant can be evaluated by the parity analysis [Eq. (16)] by tracking the band inversion induced by the main and folded bands at the TR invariant points.

The physical interpretation of the SODW induced TI is that the electron-electron interaction induces a chirality inversion in the real-space between different sublattices. In 2D, the emergent QSH effect is, therefore, spatially modulated as demonstrated in Fig. 16, which we call Quantum Spin–Hall Density Wave (QSHDW) insulator. In typical topological classes, topological invariants arise from the non-trivial geometry of the band topology of non-interacting fermions. Electron-electron interaction does not directly drive a topological phase transition, except in few cases such as topological Mott, or Kondo or Anderson insulators, in which, however, no Landau-type order parameter develops. The proposed QSHDW phase is a new kind of Landau-type order parameter, which is associated with topological invariant. The realization of this order state requires SOC, strong coulomb interaction, as well as chemical potential tunability to obtain the desired nesting wavevector. Quantum wires of Pb, Bi and other SOC elements are ideal systems to study this problem. Because here due to quasi-1D nature, Fermi surface nesting is enhanced, and here all three parameters can be tuned both externally and internally. The QSHDW phase can also be explored in non-centrosymmetric heavy-fermion materials that allow SOC split band structure.

### 3.4 Spinless orbital texture inversion induced topological ‘chiral orbits’

So far we have discussed the formation of magnetic field free chiral electrons in two methods, namely via the staggered hopping in SSH model or in graphene lattice, and due to SOC. In some of the nodal superconductors, such as nodal $d$-wave copper-oxide superconductors, owing to the particle-hole symmetry, the quasiparticle dispersion around the discrete nodal point acts as Dirac excitations. Therefore, in the existing classes of Dirac materials, ‘massless’ Dirac fermions only appears in certain conditions. For example, Dirac cones only form in atomically thin layer of graphene, or on the surface states inside the bulk gap in TI. Therefore, the relevant materials choices are restricted to heavy elements which, by nature, have lower band velocity and higher correlation strength.

Recently, the present author has proposed a new theory for a distinct type of Dirac materials, called ‘Weyl/Dirac orbital semimetals’ and...
topological orbital insulator, which has Weyl/Dirac cone arising due to orbital texture inversion at discrete momenta between two orbitals with different symmetries. The general idea of this design principle lies in assembling different atoms with distinct conduction electrons in such a way that the inter-orbital electron hopping or tunneling term naturally obtains an odd function of energy-momentum dispersion. Finally, the intra-orbital terms also conspire in such a way that the corresponding low-energy Hamiltonian can be reduced to an effective $k.p$ - type Dirac Hamiltonian. The resulting Dirac/ Weyl cones at the orbital degenerate points are protected by lattice/ translational symmetry.

We discuss here a specific example for engineering Weyl cones. Weyl cones are split Dirac cones arising from the Dirac Hamiltonian when either inversion or TR symmetry is lifted, see Fig. 17. We take a layer-by-layer setup which includes even and odd parity orbitals in alternating layers - dubbed orbital selective superlattice. Such structure is odd under mirror symmetry along the superlattice growth axis. As shown in Fig. 18, we consider ‘s’ and ‘p$_z$’ orbitals, placed along the z-direction in different layers, such that inter-orbital or inter-layer hopping has same amplitude (unlike in the SSH model in which the amplitude itself is different), but acquires different sign. The net tunneling matrix-element then becomes purely imaginary as $i \varepsilon \sin(k_c)$, where $t_{sp}^{z}$ is the hopping amplitude and $c$ is the interlayer distance. Such complex hopping term can also arise in various other orbital combinations, such as a combination of bonding and antibonding states, or from the mixture of two even orbitals (such as $s$ and $d_{xy}$), or two odd orbitals (such as $p_x$ and $p_y$) orbitals when they are placed diagonally, as discussed in Ref. Let us fix $d_{k}(k) = \varepsilon \phi(k_c)$, hence we are left with obtaining imaginary in-plane hoppings corresponding to $d_s$ and $d_p$. Here we can seek the same principle as orbital selective in-plane lattice sites, or a hexagonal lattice, or $p_x + p_y$ - wave pairing or SOC, all yielding essentially similar Hamiltonians. We take a Rashba-type SOC which sets $d_{s} = \alpha_{s} \sin(k_a)$, and $d_{p} = -i \alpha_{p} \sin(k_a)$. This Rashba-SOC is the lattice generalization of the same Rashba-SOC discussed in Sec. 3.3. Now the Hamiltonian can be expressed in the usual $\Gamma$-matrix form in the orbital spinor $\Psi_{k} = \begin{pmatrix} \psi_{k_F}^{+} & \psi_{k_F}^{-} & \psi_{k_{F}}^{p} & \psi_{k_{F}}^{d} \end{pmatrix}$ as

$$H(k) = \varepsilon^{\pm}(k)\begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & \Gamma_{1} & \Gamma_{4} \\ 0 & 0 & \Gamma_{4} & \Gamma_{1} \end{pmatrix} + d(k)\Gamma_{5},$$

where $\varepsilon^{\pm}(k) = (\varepsilon^{+}(k) \pm \varepsilon^{-}(k))/2$, with $\varepsilon^{\pm}(k)$ being the intra-orbital dispersions. The energy spectrum is $\varepsilon^{\pm}(k) = \varepsilon^{0}(k) \pm \sqrt{(\varepsilon^{0}(k))^2 + |d(k)|^2}$. Therefore, the contour of $\varepsilon^{0}(k) = 0$ gives a gapless nodal ring, among which all the k-points become gapped by finite values of $d$-vectors except those discrete points at which all its components vanish simultaneously. At these discrete k-points, Dirac/Weyl cones arise. $\Gamma_{4}$ is the parity operator, and is even under TR, while $\Gamma_{1,2,3}$ are odd under TR. Since $\varepsilon^{0}(k)$ is even function of $k$, while $d_{1,2,3}$ are odd functions in momentum, the Hamiltonian is invariant under both inversion and TR symmetries. In such case, 3D Dirac cones can appear only at the TR invariant high-symmetric k-points, and are four-fold degenerate.
As the inversion symmetry is lifted, they split in the momentum space, while with TR symmetry breaking, the splitting can occur in both momentum and energy directions. The corresponding two-fold degenerate nodal cones are called the Weyl cones, each enclosing integer, but opposite Chern numbers. Therefore, Weyl cones always appear in pairs, each representing the center of counter helical ‘chiral orbits’, but remain protected by translational symmetry.137 As a negative band gap is opened at the 3D Dirac cone, the system becomes a $Z_2$ TI. Similarly, negative band gap at the Weyl nodes can also give rise to topological crystalline insulators if mirror symmetry is present. Such orbital texture inversion induced TIs is referred as topological orbital insulators. One can envision engineering topological orbital insulators in a similar fashion as discussed in Sec. 3.2, with layer-by-layer approach, in which stacking adjacent layers must host Weyl/Dirac cones with opposite chirality at the same momentum.

Recently, we have predicted that the ferromagnetic $V_S$ is an intrinsic Weyl orbital semimetal.17 The theory of Weyl orbital semimetals does not depend on the uncommon conditions such as sublattice symmetry in atomically thin graphene, or high value of SOC in TIs. Engineering Weyl orbital materials will expand the territory of the Dirac materials beyond the typical heavy elements’ Dirac systems or graphene to even lower atomic number system and thereby enhance the value of Fermi velocity. The nature of impurity scattering protection for different quantum orbitronics cases is characteristically different. In the present family, an electron can only scatter from one orbital state to another when the impurity vertex contains a corresponding anisotropic orbital-exchange matrix-element or if the electron dynamically passes through the momentum and energy of the Dirac cone. Another advantage of the Weyl orbital semimetal is that here the Dirac cone is even immune to TR symmetry breaking, and a bulk gap can be engineered by the lattice distortion. Therefore, the generation, transport and detection of orbitally protected electric current may lead to new opportunities for orbitronics. Chiral orbital current in the Weyl semimetals can be detected by Kerr effect.

4 Conclusion and Outlook
This article presented a thorough understanding of the topological invariant in the absence and presence of TR symmetry, as well as other symmetries such as inversion symmetry, mirror symmetry, or particle-hole symmetry. The key strategy that we followed here is that the historical development of the topological invariant, starting from the quantization of flux through a non-trivial Gaussian curvature to the IQH to QSH to $Z_2$ TI, provides a more transparent and step-by-step development of this field. A unique feature of this field is that it blends concepts from various fields including mathematics, condensed matter physics, chemistry, and particle physics. A detailed discussion of the materials chemistry of the TI systems is left out in this article and can be found elsewhere.63 We can highlight an interesting difference between the trivial insulator and TI from a chemist’s perspective. In a band insulator, the valence and conduction bands remain fully filled and empty, respectively. Therefore, such insulators are sought in elements or compounds whose outermost orbital contains even number of electrons. TI is also a ‘band insulator’ with a twist. Here the valence and conduction bands are also expected to be completely filled and empty, respectively, but at the same time, these bands are inverted at the TR invariant k-points, suggesting otherwise. Therefore, one can neither seek for elements with completely full outermost orbital, nor partially filled orbital (since partially filled state should give a metallic state). For the same reason, the material should neither be too covalent, nor too ionic.

Therefore, intermetallics are obvious elements to consider for non-interacting TI (interaction can change this simple explanation). Indeed, most of the TIs are made of intermetallics. Here we can consider at least two partially occupied elements which participate in orbital-overlap and/or coupled via SOC. Therefore, a band can remain fully occupied by accommodating two partially filled orbitals. If the exchange of the orbital character occurs at odd number of TR invariant points, it gives rise to ‘strong’ TI, otherwise even number of orbital weight switching can give rise to either a ‘weak’ TI or a trivial insulator.

From high-energy physicists’ perspective, TI field offers a plenty of new opportunities to predict new excitations as well as to realize some of the uncharted ‘particles’ predicted there. Weyl fermions138–144 and Majorana fermions142–144 which remained elusive for decades have only been realized recently in TI platforms. Axion,24,106 and anyons145,146 are two widely searched excitations which are predicted to be present in TIs. Recently, it is shown that supersymmetry (SUSY) can be found in TI and superconductor heterostructures.147,148

Cold atom physicists also find it interesting to contribute to the TI fields in various ways. Haldane, in his original paper,9 commented that “the particular model presented here is unlikely to be directly physically realizable”. Cold atom researchers have made it possible to create
the Haldane model with optically generated honeycomb lattice. They have engineered TR and inversion symmetry breakings by carefully enhancing next-nearest-neighbor hopping, and providing staggered onsite energies to different sublattices, respectively, as proposed by Haldane. More recently, protected edge state is generated in various optical lattice structure by employing synthetic gauge field. The apparent setback to realize the $Z_2$ TI due to the lack of 2D SOC can be overcome with our engineered structure (as discussed in Sec. 3.1), and the realization of TR invariant TI in optical lattice field is a matter of time. Due to tremendous controls over structural and quantum properties in these setups, an unprecedented tunability of TI properties for further exploration can be possible here.

Despite the predictions and discoveries of several TI materials classes, materials flexibility still remains a grand challenge. Bi$_x$Se$_{1-x}$ family is widely used for many experiments on TI, since both single crystal and thin films of this system can be easily grown. HgTe/CdTe and InAs/GaSb are the only two systems experimentally demonstrated as QSH insulators, which are however, not being used in other experiments. Magnetic doped thin film of Bi$_x$Se$_{1-x}$ is the only system synthesized so far to be QAH insulator. Pb$_p$Sn$_{1-p}$/Te$_2$ and SnS are the only two systems known to be topological crystalline insulator. SmB$_6$ is predicted, and subsequently realized to be topological Kondo insulator, although evidence against this conclusion is also present. Cd$_{As}$ and Na$_x$Bi are synthesized to be 3D TI, while TaAs and NbAs family are discovered recently to be Weyl semimetal. Many other families, such as topological Mott insulator, topological Anderson insulator, and topological axion insulators are yet to be discovered. Moreover, the inevitable presence of the bulk conductivity in most of the 3D TI samples poses a serious nuisance to experimentalists. On the other hand, engineering TI can be rather simpler. It also offers tremendous versatility in terms of materials growth, and obtaining quantum and topological properties. Therefore, the successful preparation of ‘home-made’ TI may cater to physics, chemistry and engineering fields seeking suitable materials with higher tunability and materials flexibility.

Acknowledgments

I have many people to thank for their generous collaborations, discussions, and lessons on various aspects of topological states of matter and other related concepts that led to the completion of this article. There is a possibility that I fail to acknowledge everybody here due to the rush in finishing this article; the people I can immediately recall to convey my gratitude are Arun Bansil, Alexander Balatsky, Hsin Lin, Le-Quy Duong, F.-C. Chuang, M. Zahid Hasan, Dipriman Sen, Anindya Das, Aavek Bid, Tirupathaiah Setti, Vidya Madhavan, Robert Markiewicz, Matthias Graf, Jian-Xin Zhou, Christoph Tegenkamp, J. Hugo Dil, Kapildeb Dolui, Gaurav Gupta, Sujay Ray, Ananya Ghatak, Sayonee Ray, Kallof Sen, Soumi Ghosh, and Kausik Ghosh.

Received 21 March 2016.

References

A Pedagogic Review on Designing Model Topological Insulators

Tanmoy Das is an Assistant Professor in the Department of Physics at Indian Institute of Science. His research field is theoretically condensed matter physics, with special emphasis on strongly correlated electrons, superconductivity and topological insulators. He received his PhD in Physics from Northeastern University in 2009, MSc at Indian Institute of Technology, Bombay in 2002 and BSc at University of Calcutta in 2002. He was a Senior research scientist at the Graphene Research center, National University of Singapore and a Postdoctoral fellow at Los Alamos National Laboratory, USA. He has published more than 70 scientific papers in peer-reviewed journals, and given more than 30 invited talks in various national and international conferences. He is also an editorial board member of the Scientific Reports Journal.

He was awarded “NERSC Award” for High Impact Scientific Achievement from Lawrence Berkeley National Laboratory, Berkeley, CA, USA in 2012 and Postdoc Distinguished Performance Award at Los Alamos National Laboratory, Los Alamos, USA in 2010. Presently, he is an Adjunct research Scientist Professor at Northeastern University, Boston, USA.