Week 4
Lecture Notes:
Topological Condensed Matter Physics

Sebastian Huber and Titus Neupert

Department of Physics, ETH Zürich
Department of Physics, University of Zürich

If you notice mistakes or typos, please report them to michael.denner@physik.uzh.ch or periv@phys.ethz.ch

Chapter 4

From Chern insulators to 3D topological insulators

Learning goals

- We know Dirac fermions.
- We know what a Chern insulator is.
- We know the BHZ model.
- We can explain the idea of "pair-switching".

- G. Jotzu, M. Messer, R. Desbuquois, M. Lebrat, T. Uehlinger, D. Greif, and T. Esslinger, Nature **515**, 237 (2014)

Initially, we dealt with systems subject to a magnetic field $\boldsymbol{B}$. We could show how their ground state can be described with a topological invariant, the Chern number. Moreover, we have seen how topological phenomena can be observed in one dimensional systems without a magnetic field. In the present chapter, we try to extend these ideas. The main question we are trying to answer is the following: Can there be lattice systems with Bloch bands that are characterized by a non-zero Chern number even in the absence of a net magnetic field? Such an insulator would be termed a *Chern insulator*. Before we embark on this question, we need to understand a simple continuum problem called the Dirac model. Moreover, can we get rid of time-reversal symmetry breaking all-together and find topological insulators in time reversal invariant (TRI) systems?

4.1 Dirac fermions

Dirac fermions in two dimensions are described by the Hamiltonian

$$\mathcal{H}(\boldsymbol{k}) = \sum_i d_i(\boldsymbol{k})\sigma_i \qquad \text{with} \qquad d_1(\boldsymbol{k}) = k_x,\ \ d_2(\boldsymbol{k}) = k_y,\ \ d_3(\boldsymbol{k}) = m. \tag{4.1.1}$$

The energies and eigenstates are given by

$$\epsilon(\boldsymbol{k})_\pm = \pm d(\boldsymbol{k}) = \pm\sqrt{k^2+m^2} \qquad \text{and} \qquad \psi_\pm(\boldsymbol{k}) = \frac{1}{\sqrt{2d(\boldsymbol{k})[d(\boldsymbol{k})\pm d_3(\boldsymbol{k})]}}\begin{pmatrix} d_3(\boldsymbol{k})\pm d(\boldsymbol{k}) \\ d_1(\boldsymbol{k}) + \mathrm{i}d_2(\boldsymbol{k})\end{pmatrix}.$$

It is straight forward to show (exercise!) that the Berry connection of the lower band can be written as

$$\mathcal{A}_\mu(\boldsymbol{k}) = \mathrm{i}\langle\psi_-(\boldsymbol{k})|\partial_{k_\mu}\psi_-(\boldsymbol{k})\rangle = \frac{1}{2d(\boldsymbol{k})[d(\boldsymbol{k})-d_3(\boldsymbol{k})]}\left[d_2(\boldsymbol{k})\partial_{k_\mu}d_1(\boldsymbol{k}) - d_1(\boldsymbol{k})\partial_{k_\mu}d_2(\boldsymbol{k})\right] \tag{4.1.2}$$

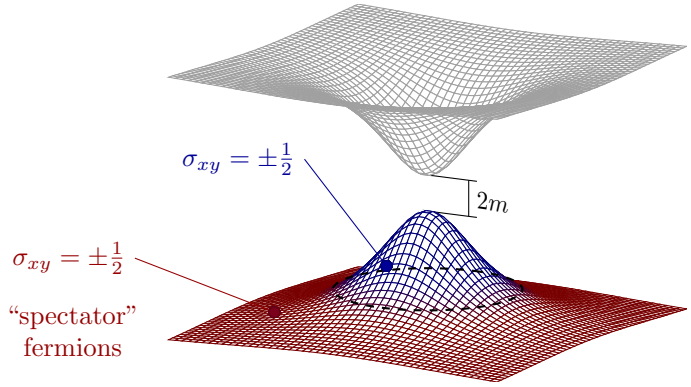


Figure 4.1: Regularization of the Dirac spectrum due to a lattice.

And the corresponding Berry curvature is given by

$$\mathcal{F}_{\mu\nu}(\boldsymbol{k}) = \frac{1}{2}\epsilon_{\alpha\beta\gamma}\hat{d}_\alpha(\boldsymbol{k})\partial_{k_\mu}\hat{d}_\beta(\boldsymbol{k})\partial_{k_\nu}\hat{d}_\gamma(\boldsymbol{k}) \qquad \text{with} \qquad \hat{\boldsymbol{d}}(\boldsymbol{k}) = \frac{\boldsymbol{d}(\boldsymbol{k})}{d(\boldsymbol{k})}. \tag{4.1.3}$$

Using our concrete d-vector we find

$$\mathcal{A}_x = \frac{k_y}{2\sqrt{k^2+m^2}(\sqrt{k^2+m^2}-m)} \qquad \text{and} \qquad \mathcal{A}_y = \frac{-k_x}{2\sqrt{k^2+m^2}(\sqrt{k^2+m^2}-m)}, \tag{4.1.4}$$

and therefore

$$\mathcal{F}_{xy} = -\frac{m}{2(k^2+m^2)^{3/2}}. \tag{4.1.5}$$

Let us plug that into the formula for the Hall conductance

$$\sigma_{xy} = \frac{e^2}{h}\frac{1}{2\pi}\int d\boldsymbol{k}\,\mathcal{F}_{xy} = -\frac{e^2}{h}\int_0^\infty dk k\frac{1}{2}\frac{m}{(k^2+m^2)^{3/2}} = -\frac{e^2}{h}\frac{\text{sign}(m)}{2}. \tag{4.1.6}$$

We can draw several important insights from this results:

1. $\sigma_{xy} \neq 0 \;\Rightarrow$ we must have broken time-reversal invariance. How did this happen?
2. $\sigma_{xy} \neq \frac{e^2}{h}\nu$ with $\nu \in \mathbb{Z}$. How can this be?

Let us start with the first question. We have to make the distinction between two cases. (i) If the σ-matrices encode a real spin-1/2 degree of freedom the time reversal operator is given by

$$\mathcal{T} = \mathrm{i}\sigma_y K,$$

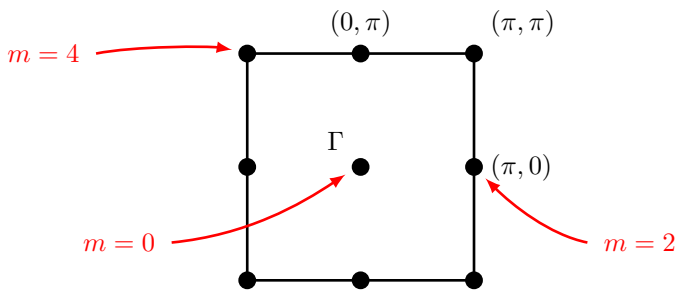


Figure 4.2: Band touching for a simple Chern insulator.

where K denotes complex conjugation. Therefore

$$\mathcal{T}\mathcal{H}(\boldsymbol{k})\mathcal{T}^{-1} = \sum_i -d_i(\boldsymbol{k})\sigma_i = -k_x\sigma_x - k_y\sigma_y - m\sigma_z.$$

If we want to above Hamiltonian to be time reversal invariant we need this to be

$$\mathcal{T}\mathcal{H}(\boldsymbol{k})\mathcal{T}^{-1} \stackrel{!}{=} \mathcal{H}(-\boldsymbol{k}) = -k_x\sigma_x - k_y\sigma_y + m\sigma_z.$$

From this we conclude that the Dirac fermions are only time reversal invariant for $d_3(\boldsymbol{k}) = m = 0$. However, in this case, there is no gap in the spectrum at $\boldsymbol{k} = 0$ and the calculation of σ_{xy} does not make sense. (ii) For the case that the Pauli matrices describe some iso-spin where $\mathcal{T} = K$, we need to have $H^*(\boldsymbol{k}) = H(-\boldsymbol{k})$. Or in other words

$$d_1(\boldsymbol{k}) = d_1(-\boldsymbol{k}), \;\; d_2(\boldsymbol{k}) = -d_2(-\boldsymbol{k}), \;\; d_3(\boldsymbol{k}) = d_3(-\boldsymbol{k}).$$

From these considerations we conclude that our Hamiltonian breaks time reversal invariance in either case and we can indeed expect a non-vanishing Hall conductance.
Let us now address the non-quantized nature of σ_{xy}. The quantization of σ_{xy} arises from the quantized value of the Chern number. We have seen in our derivation, however, that it was crucial that the domain over which we integrated the Berry curvature was closed and orientable. Here we are in a continuum model where the integral over all momenta extends over the whole $\mathbb{R}^2$. We have therefore no reason to expect σ_{xy} to be quantized.
There is value to formula (4.1.6), however. Imagine that the Dirac Hamiltonian arises from some low-energy expansion ($\boldsymbol{k}\cdot\boldsymbol{p}$) around a special point in the Brillouin zone of a lattice model. For the full lattice, the $k \to \infty$ integral would be regularized due to the Brillouin zone boundary. The whole system has a quantized Hall conductivity. However, the region close to the "Dirac-point" contributes $\pm 1/2$ to the Chern number, see Fig. 4.1. Moreover, imagine a gap closing and re-opening transition described by the Dirac Hamiltonian where m changes its value. In such a situation the change in Chern number $\Delta\mathsf{C}^{(1)} = \pm 2\pi$. Therefore, the Dirac model is an excellent way to study *changes in the Chern number*.
Before we continue to the simplest possible Chern insulator we state the following formula without proof (exercise!)

$$\mathcal{H}(\boldsymbol{k}) = \sum_{i,j=1}^{2} A_{ij}k_i\sigma_j + m\sigma_z \qquad \Rightarrow \qquad \sigma_{xy} = -\frac{e^2}{h}\frac{\mathrm{sign}(m)}{2}\mathrm{sign}(\det A). \tag{4.1.7}$$

4.2 The simplest Chern insulator

We obtain the simplest conceivable Chern insulator by elevating the Dirac model to a lattice problem

$$d_1 = k_x \to \sin(k_x), \;\; d_2 = k_y \to \sin(k_y). \tag{4.2.1}$$

The σ matrices now act in a space of orbitals. The fact that the coupling between them is odd in $\boldsymbol{k}$ means that they need to differ by one quantum of angular momentum, e.g., an s-type and a p-type orbital. By symmetry, there can be an even in $\boldsymbol{k}$ term within each orbital, so we add it to our model

$$d_3 = m \to -2 + m + \cos(k_x) + \cos(k_y).$$

The Hamiltonian is gapped ($d(\boldsymbol{k}) \neq 0 \;\forall\, \boldsymbol{k}$) $\forall\, m$ except at the special points in the Brillouin zone shown in Fig. 4.2.
We begin analyzing the Hamiltonian for $m \ll 0$ and $m \gg 4$. For $m = \pm\infty$, the eigenstates of the Hamiltonian are fully localized to single sites and the system certainly shows no Hall conductance. Another way to see this is to observe that

$$\mathsf{C}^{(1)} = \frac{1}{2\pi}\int_{\mathrm{BZ}} d\boldsymbol{k}\, \epsilon_{\alpha\beta\gamma}\hat{d}_\alpha\partial_{k_x}\hat{d}_\beta\partial_{k_y}\hat{d}_\gamma \tag{4.2.2}$$

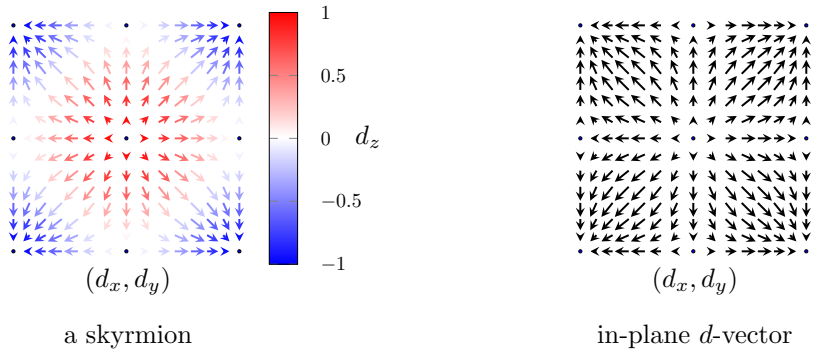


Figure 4.3: Left: spin-configuration of a skyrmion. Right: in-plane d-vector of H.

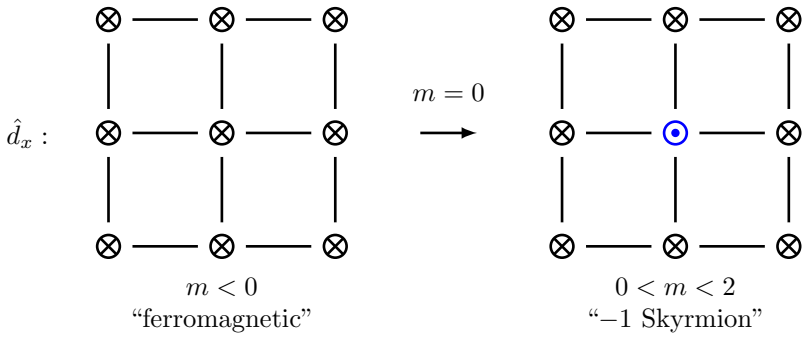


Figure 4.4: Change of the d_3 component at the first critical point.

counts the winding of $\hat{\boldsymbol{d}}(\boldsymbol{k})$ throughout the Brillouin zone, i.e., it provides us what we know as the *skyrmion number*. In Fig. 4.3(a) we show a spin-configuration corresponding to a skyrmion. When we now look at the planar part of the d-vector, we see that we have all laid out for a skyrmion. The only addition we need is a sign change of d_3 at the right places in the Brillouin zone. This does not happen for $m < 0$ or $m > 4$. Note that exactly this sign change closes the gap in a fashion describable by Dirac fermions. Hence we appreciate the importance of the above discussion. It is now trivial to draw the phase diagram.

The case $0 \leq m < 2$: We start from $m = -\infty$ where $\sigma_{xy} = 0$ and go through the gap-closing at $\boldsymbol{k} = 0$ for $m = 0$. Around $\boldsymbol{k} = 0$ we find

$$\mathcal{H} = k_x\sigma_x + k_y\sigma_y + m\sigma_x.$$

Therefore

$$\Delta\sigma_{xy} = -\frac{e^2}{h}\left[\frac{1}{2}\text{sign}(m)\Big|_{m>0} - \frac{1}{2}\text{sign}(m)\Big|_{m<0}\right] = -\frac{e^2}{h} = \sigma_{xy}.$$

The correspoding change in $d_3(\boldsymbol{k})$ is shown in Fig. 4.4.

The case $2 \leq m < 4$: At $m = 2$ the gap closes at $(\pi, 0)$ and $(0, \pi)$. Let us expand the Hamiltonian around these points

$$\mathcal{H}_{(\pi,0)} = \quad k_x\sigma_x - k_y\sigma_y + (-2+m)\sigma_z, \tag{4.2.3}$$
$$\mathcal{H}_{(0,\pi)} = -k_x\sigma_x + k_y\sigma_y + (-2+m)\sigma_z. \tag{4.2.4}$$

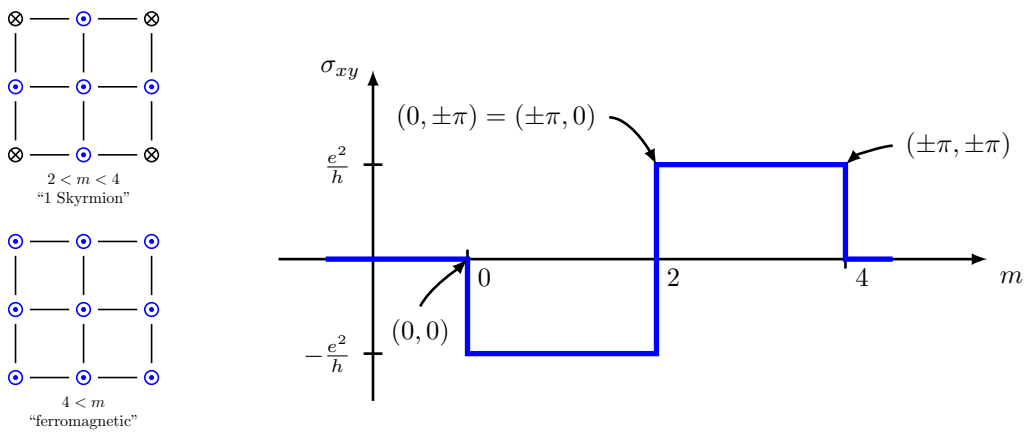


Figure 4.5: Left: d_3 component after the second and third gap closing. Right: Evolution of the topological index as a function of m.

From this we read out the change in σ_{xy}:

$$\Delta\sigma_{xy} = 2\frac{e^2}{h}\left[\frac{1}{2}\text{sign}(-2+m)\Big|_{m>2} - \frac{1}{2}\text{sign}(-2+m)\Big|_{m<2}\right] = 2\frac{e^2}{h}. \tag{4.2.5}$$

Note that the 2 in front stems from the two gap closings, and an additional $-$ sign arises from the odd sign of the determinant A, cf. Eq. (4.1.7). Together with the value of σ_{xy} for $0 < m < 2$ we obtain

$$\sigma_{xy} = +\frac{e^2}{h}.$$

The corresponding $d_3(\boldsymbol{k})$ is shown in Fig. 4.5.

The case $4 \leq m$: The last gap-closing happens at (π, π) for $m = 4$. At this point

$$\mathcal{H}_{(\pi,\pi)} = -k_x\sigma_x - k_y\sigma_y + (-4+m)\sigma_z.$$

As before the change in σ_{xy} is given by

$$\Delta\sigma_{xy} = -\frac{e^2}{h}\left[\frac{1}{2}\text{sign}(-4+m)\Big|_{m>0} - \frac{1}{2}\text{sign}(-4+m)\Big|_{m<0}\right] = -\frac{e^2}{h}.$$

And we arrive again at $\sigma_{xy} = 0$ as expected for a phase connected to the $m = \infty$ limit. Again, $d_3(\boldsymbol{k})$ is shown in Fig. 4.5 together with an overview of the whole analysis.
Before we move on, we also show the energy spectrum of the lattice Dirac Hamiltonian analyzed here. Fig. 4.6 shows such a spectrum for $m = 0.8$ on a half-plane: We calculate the spectrum on a cylinder of length $L = 60$. Around the circumference we use the translation symmetry to label states with respect to $k_\parallel$. If we also apply periodic boundary conditions in the direction where the cylinder has length L (turning the geometry into torus), we find the bulk spectrum shown in the left panel. By opening the boundaries and only selecting states on one side of the open cylinder we obtain the right panel showing the expected chiral edge mode traversing the gap.

4.3 Time reversal invariant topological insulators

In this chapter we try to understand what topological properties can arise for free fermion systems subject to some symmetry constraints. The exposition starts from a the simplest extension

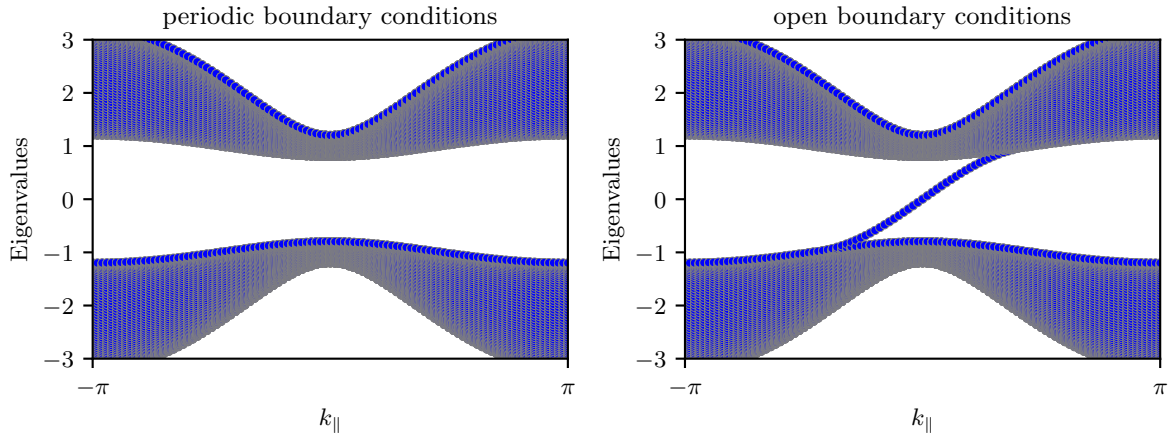


Figure 4.6: Spectrum of the lattice Dirac model for $m = 0.8$. The left panel shows the bulk spectrum as a function of the momentum along one direction ($k_\parallel$). The right panel shows the same setup for a half plane including one edge.

of the lattice Dirac model into a time-reversal invariant version. We then motivate on physical grounds how one can construct a topological index characterizing this new type of band insulator. The so derived topological index for two-dimensional systems readily generalizes to three dimensions. Note, that our path of going from the lattice Dirac model to the Bernevig-Hughes-Zhang model [1] below is not following the historical route. For an account of the two milestone papers by Haldane [2] and Kane and Mele [3], you can consult the App. B. Finally, note that we build in an essential way on Kramers-pairs which require the time reversal operator $\mathcal{T}^2 = -1$ to square to minus one, i.e., we are dealing with spinful electrons.

4.3.1 The BHZ model

We construct a time reversal invariant Hamiltonian from the lattice Dirac model

$$\mathcal{H}_{\mathrm{D}}(\boldsymbol{k}) = [m - 2 + \cos(k_x) + \cos(k_y)]\sigma_z + \sin(k_x)\sigma_x + \sin(k_y)\sigma_y, \tag{4.3.1}$$

by explicitly adding the time-reversed partner

$$\begin{aligned}\mathcal{H}_{\mathrm{BHZ}} &= \begin{pmatrix} \mathcal{H}_{\mathrm{D}}(\boldsymbol{k}) & 0 \\ 0 & \mathcal{H}^*_{\mathrm{D}}(-\boldsymbol{k}) \end{pmatrix} \\ &= \tau_0 \otimes \{[m - 2 + \cos(k_x) + \cos(k_y)]\sigma_z + \sin(k_y)\sigma_y\} + \sin(k_x)\tau_z \otimes \sigma_x. \end{aligned} \tag{4.3.2}$$

Note, that if we use $\boldsymbol{\tau}$ to denote a spin-1/2 degree of freedom, the time reversal operator reads now $\mathcal{T} = \mathrm{i}\tau_y \otimes \sigma_0 K$, where K denotes complex conjugation and $\mathcal{T}^2 = -\tau_0 \otimes \sigma_0$. In other words, this Hamiltonian simply describes two Chern insulators glued together, each with an opposite Chern number. As long as the z-component $\tau_z \otimes \sigma_0$ of the spin is conserved, one can immediately write down a topological index, the spin-Chern number [4]

$$\mathsf{C}_s^{(1)} := \frac{\mathsf{C}_\downarrow^{(1)} - \mathsf{C}_\uparrow^{(1)}}{2} \mod 2 \quad \in \mathbb{Z}_2. \tag{4.3.3}$$

When we look at the edge spectrum of the above Hamiltonian in Fig 4.7, it can be understood why this index is only in $\mathbb{Z}_2$ and not in $\mathbb{Z}$ as the Chern number. The crossing of the two edge states at $k_\parallel = 0$ is protected by the Kramer's degeneracy. This degeneracy arises, as $\boldsymbol{k} = 0$ is itself a time-reversal invariant momentum (TRIM). If we now would have two such Kramers pairs

at $\boldsymbol{k} = 0$, one could unlink the edge states while preserving the double-degeneracy. However, each time one has an odd number of such crossings, one is bound to remain, hence $\mathsf{C}_s^{(1)} \in \mathbb{Z}_2$. We can now add a perturbation that breaks the conservation of S_z, such as

$$\mathcal{H} = \mathcal{H}_{\mathrm{BHZ}} + \lambda \tau_x \otimes \sigma_y. \tag{4.3.4}$$

As this Hamiltonian still commutes with $\mathcal{T}$, we expect the Kramers argument of above to still hold. Indeed, as we see in the right panel of Fig. 4.7, the edge states persists also for $\lambda = 0.1$. Clearly we need a better index than the spin-Chern number defined above. While there is an extension of the spin-Chern number for weakly broken spin-conservation[1], we want to make progress to an index only based on time-reversal symmetry.

4.3.2 $\mathbb{Z}_2$ index

Charge polarization

We revisit Laughlin's pumping argument to make progress towards a $\mathbb{Z}_2$ index for TRI topological insulators. We have seen in the last chapter, that in the simple case of only one filled band in one spatial dimension, the charge polarization can be written as

$$\mathsf{P} := -\frac{\mathrm{i}}{2\pi} \log W = -\frac{1}{2\pi} \int_0^{2\pi} \mathcal{A}(k) \mathrm{d}k. \tag{4.3.5}$$

Two comments are in order:

1. If we re-gauge $|\psi_k\rangle \to e^{\mathrm{i}\varphi(k)} |\psi_k\rangle$ with a $\varphi(k)$ that is winding by $2\pi m$ throughout the Brillouin zone, the corresponding polarization changes to

   $$\mathsf{P} \to \mathsf{P} + m.$$

   This is ok, as charge polarization is anyway only defined up to a lattice constant.

[1] The two filled bands of H_{BHZ} together have a vanishing Chern number. The S_z quantum number, however, allowed us to label the two filled bands individually and calculate a Chern number per spin. Here we show how one can extend this to weak violation of S_z conservation. First, take the two eigenstates of the filled bands

$$P_{\mathrm{filled}}(\boldsymbol{k}) = \begin{bmatrix} \boldsymbol{u}_1(\boldsymbol{k}) & \boldsymbol{u}_2(\boldsymbol{k}) \end{bmatrix},$$

with the two column vectors $\boldsymbol{u}_\alpha(\boldsymbol{k})$, $\alpha = 1, 2$ of the two lower eigenstates of H_{BHZ} per $\boldsymbol{k}$. In case $\tau_z \otimes \sigma_0$ is a symmetry, the two labels are also labelling the eigenstates of S_z with $(1, 2) \to (\uparrow, \downarrow)$. How do we smoothly (as a function of $\boldsymbol{k}$) assign the labels α in case this symmetry is broken? We can project the symmetry into the space of filled states

$$S_z^{\mathrm{filled}}(\boldsymbol{k}) = P_{\mathrm{filled}}(\boldsymbol{k}) \tau_z \otimes \sigma_0 P_{\mathrm{filled}}^\dagger(\boldsymbol{k}).$$

As S_z is not commuting with H_{BHZ}, $S_z^{\mathrm{filled}}(\boldsymbol{k})$ is not diagonal anymore. However, we can diagonalize it

$$\begin{pmatrix} \chi_+ & 0 \\ 0 & \chi_- \end{pmatrix} = M(\boldsymbol{k}) S_z^{\mathrm{filled}}(\boldsymbol{k}) M^\dagger(\boldsymbol{k}).$$

As long as there is a spin gap $\Delta_S = |\chi_+ - \chi_-| > 0$ for all $\boldsymbol{k}$, we can now use the states

$$\begin{bmatrix} \boldsymbol{u}_+(\boldsymbol{k}) & \boldsymbol{u}_-(\boldsymbol{k}) \end{bmatrix} = M(\boldsymbol{k}) P_{\mathrm{filled}}(\boldsymbol{k})$$

to calculate the spin-Chern number [5] via

$$\boldsymbol{\mathcal{A}}_\pm = \mathrm{i} \langle \boldsymbol{u}_\pm(\boldsymbol{k}) | \nabla \boldsymbol{u}_\pm(\boldsymbol{k}) \rangle.$$

It is important to note that there is no well-established bulk-boundary relation for the spin-Chern number defined this way.

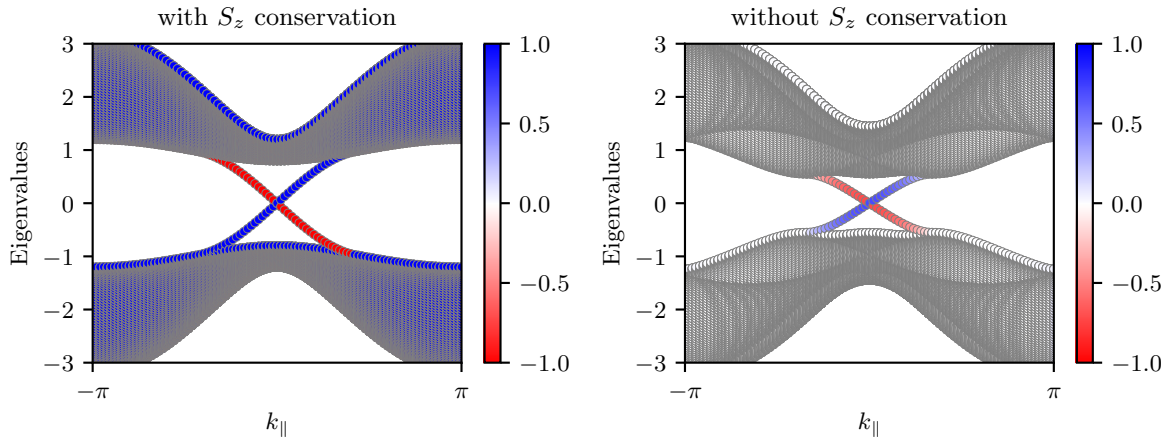


Figure 4.7: Spectrum of the BHZ model for $m = 0.8$. The left panel shows the spectrum for a half plane including only one edge for $\lambda = 0$, where S_z is a good quantum number. The color code indicates the spin S_z (Note, that the bulk has the same number of both red and blue dots). The right panel shows the same spectrum for $\lambda = 0.8$, where the two spin sectors are mixed leading to an un-polarized bulk. Due to the type of coupling, the spin-polarization is largely preserved along the edge.

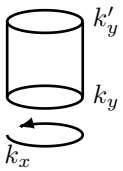
2. P depends on the chosen gauge. But *changes in* P *by a smooth change in system parameters are gauge independent.* So let us imagine a tuning parameter k_y with

$$\mathcal{H}(k_y) \to \mathcal{H}(k_y')$$

which is slow in time. The change in charge polarization is given by

$$\Delta\mathsf{P} = -\frac{1}{2\pi}\left[\int_{-\pi}^{\pi} dk \mathcal{A}(k, k_y) - \int_{-\pi}^{\pi} dk\, \mathcal{A}(k, k_y')\right] \tag{4.3.6}$$

If we use Stokes' theorem we arrive at



$$\Delta\mathsf{P} = \int_{k_y}^{k_y'} dk_y \int_{-\pi}^{\pi} dk\, \mathcal{F}(k, k_y). \tag{4.3.7}$$

By chosing $k_y' = k_y + 2\pi$, we find for the change in charge polarization $\Delta\mathsf{P} = \mathsf{C}^{(1)}$ where $\mathsf{C}^{(1)}$ is the Chern number. We known, however that $\mathsf{C}^{(1)} \sim \sigma_{xy}$. Indeed, this is nothing but Laughlin's pumping argument for the quantum Hall effect and hence is equal to zero for TRI systems.

Building on the above insight we try to refine the charge pumping of Laughlin to be able to characterize a TRI system.

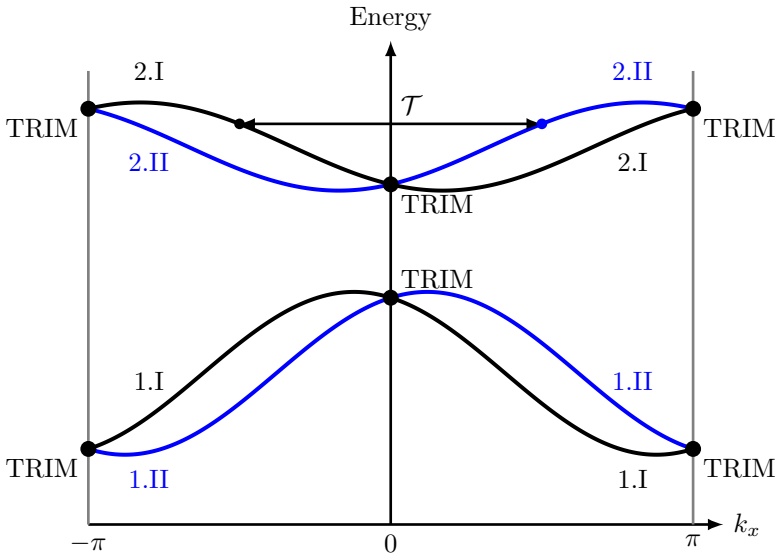


Figure 4.8: Energy levels for a time revesal invariant system.

Time reversal polarization

Let us now try to generalize the charge pumping approach to the TRI setup. For this it is beneficial to look at the structure of a generic energy diagram as shown in Fig 4.8. Under time-reversal, momenta k are mapped to $-k$. Moreover, there are special points in the Brillouin zone which are mapped onto themselves. This is true for all momenta which fulfill $k = -k + G$, where G is a reciprocal lattice vector. This is trivially true for $k = 0$, but also for special points on the borders of the Brillouin zone. On such time reversal invariant momenta (TRIM's), the spectrum has to be doubly degenerate due to Kramer's theorem.
Owing to the symmetry between k and $-k$ we can constrain ourselves to only *half the Brillouin zone*. In this half, we label all bands by 1.I, 1.II, 2.I, 2.II, ... The arabic number simply label pairs of bands. Due to the double degeneracy at the TRIMs, we need an additional (roman number) to label the two (sub)-bands emerging form the TRIMs. One can also say that the roman index labels Kramers pairs

$$\mathcal{T}|\varphi_{n.\mathrm{I}}(k)\rangle = e^{\mathrm{i}\chi_{n,k}}|\varphi_{n.\mathrm{II}}(-k)\rangle. \tag{4.3.8}$$

We now try to construct the polarization for only one of the two labels $s = \mathrm{I}$ or II

$$\mathsf{P}^s = -\frac{1}{2\pi}\int_{-\pi}^{\pi} dk\, \mathcal{A}^s(k) \qquad \text{with} \qquad \mathcal{A}^s(k) = \mathrm{i}\sum_{n \text{ filled}} \langle\varphi_{n.s}(k)|\partial_k|\varphi_{n.s}(k)\rangle. \tag{4.3.9}$$

It is clear that $\mathsf{P} = \mathsf{P}^{\mathrm{I}} + \mathsf{P}^{\mathrm{II}}$ will vanish. However, the same must not hold for the *time reversal polarization*

$$\mathsf{P}^{\mathcal{T}} = \mathsf{P}^{\mathrm{I}} - \mathsf{P}^{\mathrm{II}}. \tag{4.3.10}$$

The problem is, that we assigned the labels I and II. It is not a priori clear if this can be done in a gauge invariant fashion. In particular, the Slater determinant of a band insulator with $2n$ filled bands has a $SU(2n)$ symmetry, as basis changes of filled states do not affect the total wave function. With our procedure we explicitly broke this $SU(2n)$ symmetry. There is a way however, to formulate the same $\mathcal{T}$-polarization $\mathsf{P}^{\mathcal{T}}$ in a way that does not rely on a specific labeling of the Kramers pairs. This can be achieved by the use of the so-called *sewing matrix* [6]

$$B_{mn}(k) = \langle\varphi_m(-k)|\mathcal{T}|\varphi_n(k)\rangle. \tag{4.3.11}$$

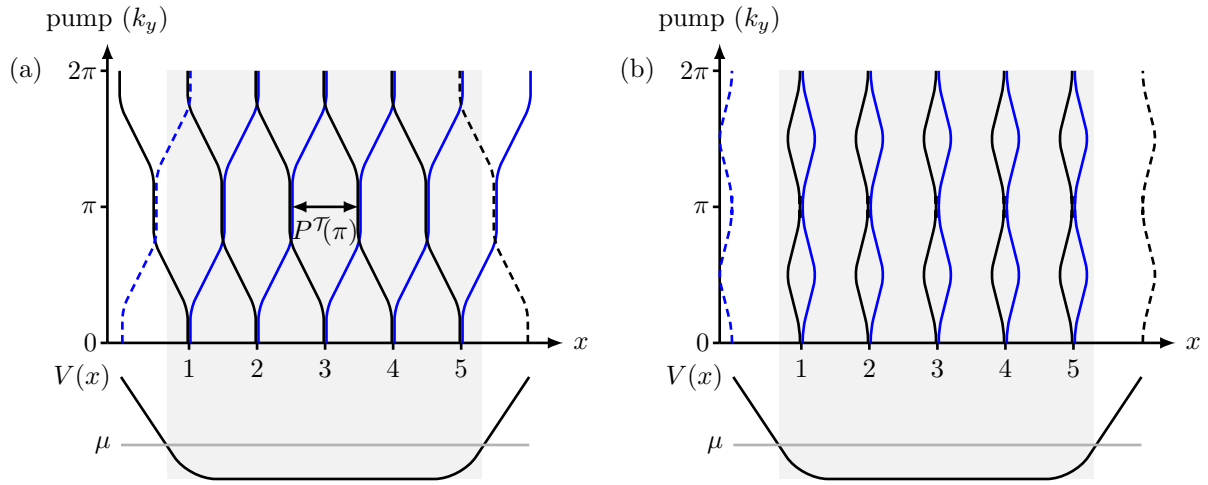


Figure 4.9: (a) Pumping of time reversal polarization in a topologically non-trivial state. (b) Pumping of time reversal polarization in a trivial state.

$B(k)$ has the following properties: (i) it is unitary, and (ii) it is anti-symmetric, i.e., $B^{\mathrm{T}}(k) = -B(k)$, only if k is a TRIM. Using this matrix one can show that

$$\mathsf{P}^{\mathcal{T}} = \frac{1}{\mathrm{i}\pi} \log \left[\frac{\sqrt{\det B(\pi)}}{\mathrm{Pf}\, B(\pi)} \frac{\mathrm{Pf}\, B(0)}{\sqrt{\det B(0)}} \right]. \tag{4.3.12}$$

This expression is manifestly invariant under $SU(2n)$ rotations within the filled bands. Moreover, it only depends on the two TRIMs $k = 0, \pi$, and it is defined modulo two.

The Pfaffian $\mathrm{Pf}\, B(k)$ of a $2n \times 2n$ anti-symmetric matrix matrix B is defined as

$$\mathrm{Pf}\, B = \frac{1}{2^n n!} \sum_{\sigma \in S_{2n}} \mathrm{sign}(\sigma) \prod_{i=1}^{n} b_{\sigma(2i-1),\sigma(1i)} \tag{4.3.13}$$

with the property

$$\mathrm{Pf}^2\, B = \det B. \tag{4.3.14}$$

Let us now see how we can describe changes in the time-reversal polarization under the influence of an additional parameter k_y. Written as in (4.3.12), it is only defined for $k_y = 0, \pi, 2\pi$, i.e, at TRIMs. In Fig. 4.9 we illustrate what we can expect from such a smooth change. We start at $k_y = 0$. If we now change k_y slowly, we know that due to TRI, we cannot build up a charge polarization. However, the Wannier centers of two Kramers pairs will evolve in opposite direction. At $k_y = \pi$, we can check how far these centers evolved away from each other. As $\mathsf{P}^{\mathcal{T}}$ is well defined and equal to 0 or 1 we have two options: (i) Each Wannier center meets up with one coming from a neighboring site [Fig. 4.9(a)]. This gives rise to $\mathsf{P}^{\mathcal{T}}(k_y = \pi) = 1$ and this effect is called *pair switching*. (ii) The centers fall back onto each other again [Fig. 4.9(b)], resulting in $\mathsf{P}^{\mathcal{T}}(k_y = \pi) = 0$.

Let us further assume that we have a smooth confining potential $V(x)$ in x-direction. As in the case of the quantum Hall effect, we see how states can be pushed up-hill or pulled down-hill as a function of k_y. However, as opposed to the quantum Hall effect, we have here the situation that on each edge we have both a state coming down in energy as well as one climbing up! From that we conclude that if we have pair-switching, we expect two counter-propagating edge states on *both sides of the sample*. The same observation can be made by looking at the Wilson loop spectrum as described in the last chapter.

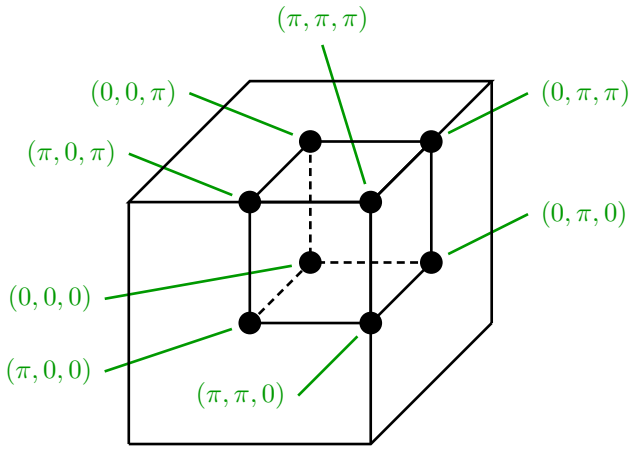


Figure 4.10: TRIMs of the three-dimensional Brillouin zone.

We can now construct a topological index for the two-dimensional system: If the $\mathcal{T}$-polarization at $k_y = 0$ and $k_y = \pi$ differ by one, we expect an odd number of pairs of edge states. Hence, we define

$$\nu = \prod_{l=1}^{4} \frac{\sqrt{\det B(\mathbf{\Lambda}_l)}}{\mathrm{Pf}\, B(\mathbf{\Lambda}_l)} \in \mathbb{Z}_2 \qquad \text{with} \qquad \mathbf{\Lambda}_l \ : \ \text{TRIM}. \tag{4.3.15}$$

4.3.3 Three dimensional topological insulators

The above formulation immediately suggests a three-dimensional generalization of the $\mathbb{Z}_2$ index

$$\nu_s = \prod_{l=1}^{8} \frac{\sqrt{\det B(\mathbf{\Lambda}_l)}}{\mathrm{Pf}\, B(\mathbf{\Lambda}_l)} \in \mathbb{Z}_2 \qquad \text{with} \qquad \mathbf{\Lambda}_l \ : \ \text{TRIM}, \tag{4.3.16}$$

where now the product runs over all eight TRIMs of the three-dimensional Brillouin zone shown in Fig. 4.10. This index is called the *strong* topological index. Additionally, one can think of a three-dimensional system to be made out of planes of two-dimensional topological insulators. In Fig. 4.11 we show how one can attribute a *weak* topological index (ν_x, ν_y, ν_z) corresponding to the stacking directions.
According to our reasoning above, when we cut the system perpendicular to the direction defined by the weak index, we expect *two Dirac cones* on the resulting surface (why?). However, if we have a strong topological index, there is a single Dirac cone irrespective of the way we terminate the bulk system. To wrap up, we mention that one usually gathers the indices to

$$\boldsymbol{\nu} = (\nu_s; \nu_x, \nu_y, \nu_z). \tag{4.3.17}$$

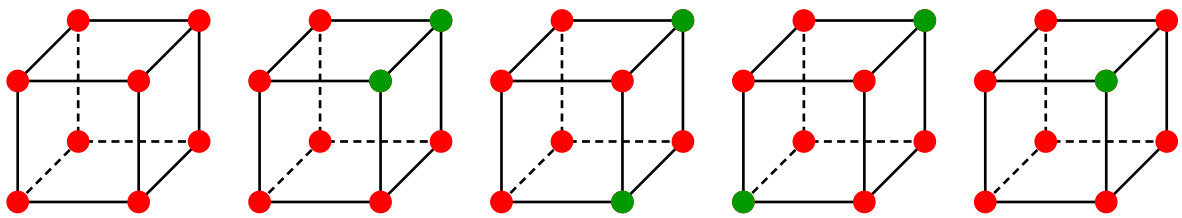


Figure 4.11: Stacking directions of 2D topological insulators.

References

1. Bernevig, B. A., Hughes, T. L. & Zhang, S.-C. "Quantum Spin Hall Effect and Topological Phase Transition in HgTe Quantum Wells". *Science* **15,** 1757. http://dx.doi.org/10.1126/science.1133734 (2006).
2. Haldane, F. D. M. "Model for a Quantum Hall Effect without Landau Levels: Condensed-Matter Realization of the "Parity Anomaly"". *Phys. Rev. Lett.* **61,** 2015. http://link.aps.org/doi/10.1103/PhysRevLett.61.2015 (1988).
3. Kane, C. L. & Mele, E. J. "Quantum Spin Hall Effect in Graphene". *Phys. Rev. Lett.* **95,** 226801. http://link.aps.org/abstract/PRL/v95/e226801 (2005).
4. Sheng, D. N., Weng, Z. Y., Scheng, L. & Haldane, F. D. M. "Quantum Spin-Hall Effect and Topologically Invariant Chern Numbers". *Phys. Rev. Lett.* **97,** 036808. http://dx.doi.org/10.1103/PhysRevLett.97.036808 (2006).
5. Prodan, E. "Robustness of the spin-Chern number". *Phys. Rev. B* **80,** 125327. http://dx.doi.org/10.1103/PhysRevB.80.125327 (2009).
6. Bernevig, B. A. & Hughes, T. L. *Topological insulators and superconductors* (Princeton University Press, 2013).

Appendix B

A more historical route to time reversal invariant topological insulators*

B.1 The Haldane Chern insulator*

In his seminal paper [1], Haldane considered a honeycomb model with no net magnetic flux but with complex phases $e^{\pm i\varphi}$ on the next-to-nearest neighbor hoppings. A possible *staggered* flux pattern giving rise to such a situation is shown in Fig. B.1. In Fig. B.1 we also indicate the sign structure of the phases. The model can be written as

$$H = \sum_{\langle i,j\rangle} c_i^\dagger c_j + t \sum_{\langle\langle i,j\rangle\rangle} e^{\pm \mathrm{i}\varphi} c_i^\dagger c_j + m \sum_i \epsilon_i c_i^\dagger c_i, \tag{B.1.1}$$

where $\epsilon_i = \pm 1$ for the two sub-lattices of the honeycomb lattice. Written in $\boldsymbol{k}$-space we find $\mathcal{H} = \epsilon(\boldsymbol{k}) + \sum_i d_i(\boldsymbol{k})\sigma_i$ with

$$d_1(\boldsymbol{k}) = \cos(\boldsymbol{k}\cdot\boldsymbol{a}_1) + \cos(\boldsymbol{k}\cdot\boldsymbol{a}_2) + 1, \tag{B.1.2}$$

$$d_2(\boldsymbol{k}) = \sin(\boldsymbol{k}\cdot\boldsymbol{a}_1) + \sin(\boldsymbol{k}\cdot\boldsymbol{a}_2), \tag{B.1.3}$$

$$d_3(\boldsymbol{k}) = m + 2t\sin(\varphi)\left[\sin(\boldsymbol{k}\cdot\boldsymbol{a}_1) - \sin(\boldsymbol{k}\cdot\boldsymbol{a}_2) - \sin(\boldsymbol{k}\cdot(\boldsymbol{a}_1 - \boldsymbol{a}_2))\right], \tag{B.1.4}$$

with $\boldsymbol{a}_1 = a(1,0)$ and $\boldsymbol{a}_2 = a(1/2, \sqrt{3}/2)$. We ignore the shift $\epsilon(\boldsymbol{k})$ in the following. What are the symmetries of this Hamiltonian? First, d_1 and d_2 are compatible with the time-reversal $\mathcal{T}$. However, $d_3(\boldsymbol{k}) = d_3(-\boldsymbol{k})$ holds only for $\varphi = 0, \pi$. We can therefore expect a non-vanishing Chern number for a general φ. The Hamiltonian has C_3 symmetry. Hence, the gap closings have to happen at the K or K' point, see Fig. B.2 (Prove!),

$$K = \frac{2\pi}{a}\left(1, \frac{1}{\sqrt{3}}\right), \qquad K' = \frac{2\pi}{a}\left(1, -\frac{1}{\sqrt{3}}\right),$$

where a denotes the lattice constant. To calculate the Chern number we follow the same logic as in the last chapter. We start from the limit $m \to \infty$ and track the gap-closings at the Dirac points at K and K'. The low energy expansion at these two points read

$$\mathcal{H}_K = \frac{3}{2}(k_y\sigma_x - k_x\sigma_y) + \left(m - 3\sqrt{3}t\sin(\varphi)\right)\sigma_z, \tag{B.1.5}$$

$$\mathcal{H}_{K'} = -\frac{3}{2}(k_y\sigma_x + k_x\sigma_y) + \left(m + 3\sqrt{3}t\sin(\varphi)\right)\sigma_z. \tag{B.1.6}$$

Note that the gap-closings at K and K' happen at different values of m (for $\varphi \neq 0, \pi$). Moreover, the two Dirac points give rise to a change in σ_{xy} of opposite sign has $\det(A)$ as a different sign. We can now construct the phase diagram

- $m > 3\sqrt{2}t\sin(\varphi)$: $\sigma_{xy} = 0$
- $-3\sqrt{2}t\sin(\varphi) < m \geq 3\sqrt{2}t\sin(\varphi)$ for $\varphi > 0$: At $m = 3\sqrt{2}t\sin(\varphi)$ the gap closes at K and we have a $\Delta\sigma_{xy} = -\frac{e^2}{h}$. The gap at K' stays open.
- $m \leq -3\sqrt{2}t\sin(\varphi)$ for $\varphi > 0$: The gap at K' closes at $m - 3\sqrt{2}t\sin(\varphi)$ and hence the Chern number changes back to 0.

For $\varphi < 0$ the signs of the Chern numbers are inverted. The resulting phase diagram is summarized in Fig. B.3.
The model of Haldane breaks time-reversal invariance $\mathcal{T}$. How can we build a model which is $\mathcal{T}$-symmetric? The easiest way is by doubling the degrees of freedom:

$$\mathcal{T}H\mathcal{T}^{-1} = H' \neq H \Rightarrow H_{\text{doubled}} = \begin{pmatrix} H & \\ & H' \end{pmatrix}.$$

We will see in the next section how Kane and Mele [2] took this step.

B.2 The Kane-Mele model*

In the last chapter we have seen that we can construct lattice models where the Bloch bands have a non-vanishing Chern number despite the absence of a net magnetic field. Here we try to build a time-reversal invariant version based on Haldane's honeycomb model for a Chern insulator.
We start from the low energy version of graphene

$$\mathcal{H}_0 = -\mathrm{i}\hbar v_{\mathrm{F}} \left[\sigma_x\tau_z\partial_x + \sigma_y\partial_y\right], \tag{B.2.1}$$

where $\boldsymbol{\sigma}$ acts on the sub-lattice index and $\boldsymbol{\tau}$ on the valley (K, K') space.
Let us add spin $\boldsymbol{s}$ to the game. With this we arrive at an 8×8 problem. The question is what kind of terms can we add in order to open a "non-trivial" gap. We have seen that $m\sigma_z + \tau_z\sigma_z 3\sqrt{3}t\sin(\varphi)$ does the job. However, this is not time-reversal symmetric for $\varphi \neq 0, \pi$ and m alone opens trivial gaps with $\mathsf{C}^{(1)} = 0$.
We construct a "non-trivial" time-reversal invariant gap step by step. First, in the sub-lattice space we need a σ_z term, otherwise we just move around the K and K' Dirac points in $\boldsymbol{k}$-space. Next, we need a spin dependent ($\boldsymbol{s}$) part to couple the two copies of the Haldane model. Let us try for the K point

$$\sigma_z \otimes s_z = \begin{pmatrix} \sigma_z & 0 \\ 0 & -\sigma_z \end{pmatrix}, \tag{B.2.2}$$

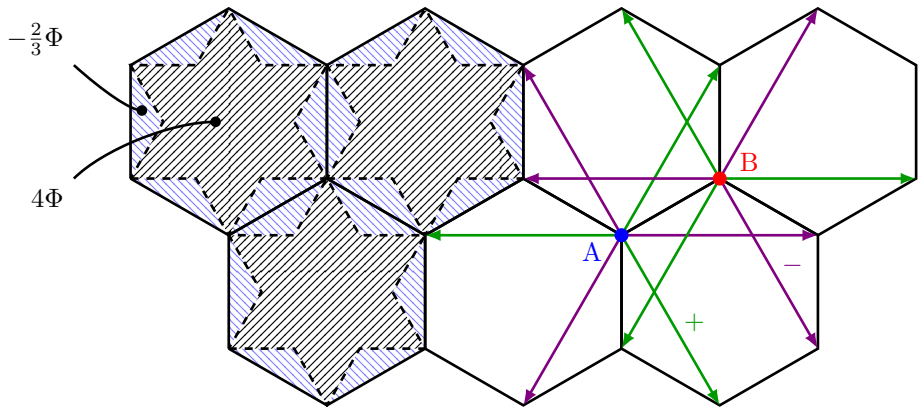


Figure B.1: The Haldane Chern insulator model.

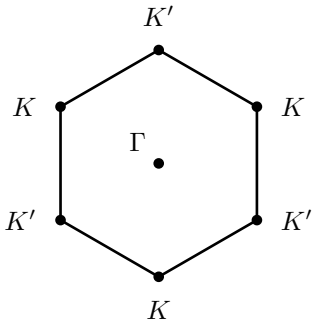


Figure B.2: Gap closings for the Haldane Chern insulator.

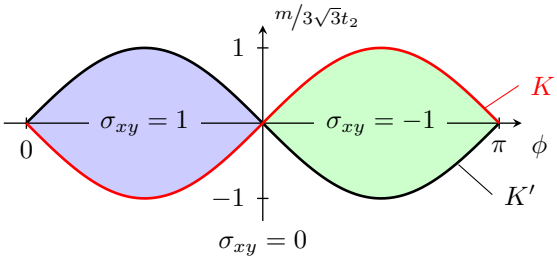


Figure B.3: Phase diagram of the Haldane model.

which gives us different gaps [with different "sign(m)"] for the two spins. How do we now add the valley degree of freedom ($\boldsymbol{\tau}$) in order to make it time-reversal invariant? The $\mathcal{T}$-operator acts in sub-lattice and spin space as

$$\mathcal{T} = \mathbb{1}_\sigma \otimes \mathrm{i}s_y K = \begin{pmatrix} 0 & -\mathbb{1}_\sigma \\ \mathbb{1}_\sigma & 0 \end{pmatrix}. \tag{B.2.3}$$

Therefore, the term $\propto \sigma_z \otimes s_z$ transforms as

$$\begin{aligned} \mathcal{T}\sigma_z \otimes s_z \mathcal{T}^{-1} &= \begin{pmatrix} 0 & -\mathbb{1}_\sigma \\ \mathbb{1}_\sigma & 0 \end{pmatrix} \begin{pmatrix} \sigma_z & 0 \\ 0 & -\sigma_z \end{pmatrix} \begin{pmatrix} 0 & \mathbb{1}_\sigma \\ -\mathbb{1}_\sigma & 0 \end{pmatrix} \\ &= \begin{pmatrix} 0 & -\mathbb{1}_\sigma \\ \mathbb{1}_\sigma & 0 \end{pmatrix} \begin{pmatrix} 0 & \sigma_z \\ \sigma_z & 0 \end{pmatrix} = \begin{pmatrix} -\sigma_z & 0 \\ 0 & \sigma_z \end{pmatrix} = -\sigma_z \otimes s_z. \end{aligned} \tag{B.2.4}$$

Under time reversal, $K \to K'$. Hence, we need the gap opening term in K' to be $\mathcal{T}\sigma_z \otimes s_z \mathcal{T}^{-1} = -\sigma_z \otimes s_z$ to have $\mathcal{T}\mathcal{H}(\boldsymbol{k})\mathcal{T}^{-1} = \mathcal{H}(-\boldsymbol{k})$. From this we conclude that the full gap opening term should be of the form

$$\mathcal{H}_{\mathrm{KM}} = \lambda_{\mathrm{SO}} \sigma_z \otimes \tau_z \otimes s_z. \tag{B.2.5}$$

We labelled the interaction with spin-orbit (λ_{SO}) to stress that $\mathcal{H}_{\mathrm{KM}}$ couples spin (s_z) and orbital (τ_z) degrees of freedom. Moreover, $\mathcal{H}_{\mathrm{KM}}$ is time-reversal invariant (TRI) by construction. Reverse engineering to a full lattice model we find

$$H_{\mathrm{KM}} = \sum_{\langle i,j\rangle,\alpha} c^\dagger_{i\alpha} c_{j\alpha} + \mathrm{i}\lambda_{\mathrm{SO}} \sum_{\langle\langle i,j\rangle\rangle,\alpha\beta} \nu_{ij} c^\dagger_{i\alpha} s^z_{\alpha\beta} c_{j\beta} + \lambda_v \sum_{i\alpha} \epsilon_i c^\dagger_{i,\alpha} c_{i,\alpha}, \tag{B.2.6}$$

where ϵ_i and the sign strucutre of ν_{ij} are the same as in Haldane's '88 model [1]. The above model was the first TRI topological insulator proposed by Kane and Mele in 2005 [2]. As it is TRI, the total Chern number cannot be non-zero. However, in the form (B.2.6), the spin

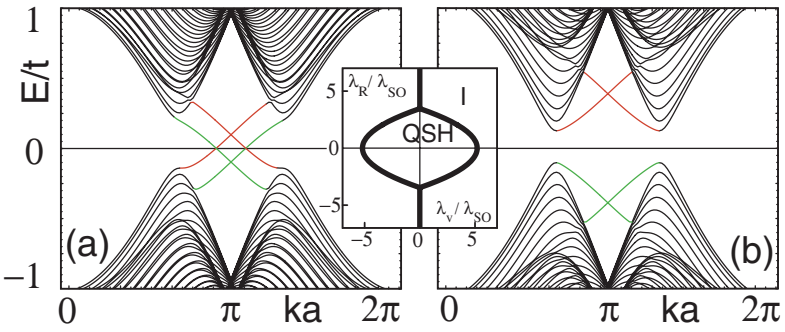


Figure B.4: Edge spectrum of the Kane Mele model for two different values of λ_R. On the left, two edge states cross the gap (colors label the edge). On the right, no edge states cross the gap. The inset shows the phase diagram as a function of λ_v and λ_R. Figure take from Ref. [2] (Copyright (2005) by The American Physical Society).

projections $|\uparrow\rangle$, $|\downarrow\rangle$ are good eigenstates. Therefore, we can use the Chern number $\mathsf{C}_\sigma^{(1)}$ in each spin-sector to characterize the phases. Indeed

$$\nu = \frac{\mathsf{C}_\uparrow^{(1)} - \mathsf{C}_\downarrow^{(1)}}{2} \bmod 2 \quad \in \mathbb{Z}_2 \tag{B.2.7}$$

defines a good topological index as we will see below [3]. The addition of a Rashba term

$$\mathcal{H}_R = \lambda_R \left[\sigma_x \tau_z s_x - \sigma_y s_x\right] \tag{B.2.8}$$

removes this conserved quantity. While $\mathcal{H}_R$ does not open a gap by itself (why?), it can influence the λ_{SO} induced gap, see Fig. B.4. However, the above topological index ν is not well defined anymore. In the following section we aim at deriving a $\mathbb{Z}_2$ index which does not rely on spin-Chern numbers.

References

1. Haldane, F. D. M. "Model for a Quantum Hall Effect without Landau Levels: Condensed-Matter Realization of the "Parity Anomaly"". *Phys. Rev. Lett.* **61,** 2015. http://link.aps.org/doi/10.1103/PhysRevLett.61.2015 (1988).
2. Kane, C. L. & Mele, E. J. "Quantum Spin Hall Effect in Graphene". *Phys. Rev. Lett.* **95,** 226801. http://link.aps.org/abstract/PRL/v95/e226801 (2005).
3. Sheng, D. N., Weng, Z. Y., Scheng, L. & Haldane, F. D. M. "Quantum Spin-Hall Effect and Topologically Invariant Chern Numbers". *Phys. Rev. Lett.* **97,** 036808. http://dx.doi.org/10.1103/PhysRevLett.97.036808 (2006).