Time-Reversal Symmetric Two-Dimensional Topological Insulators: The Bernevig-Hughes-Zhang Model
Contents

- Time Reversal Symmetry
- Kramers Degeneracy
- The Bernevig-Hughes-Zhang model
- Edge States
- The $\mathbb{Z}_2$ invariant
- Absence of Back Scattering
We define a time reversal operator which acts as

$$|n\rangle \rightarrow \mathcal{T}|n\rangle$$

Where $\mathcal{T}|n\rangle$ is the time reversed state.

For example: if $|n\rangle = |k\rangle$ we expect $\mathcal{T}|k\rangle = | - k\rangle$; if $|n\rangle = |x\rangle$ we expect $\mathcal{T}|x\rangle = |x\rangle$.

We also say a Hamiltonian $H$ has time reversal symmetry if

$$\mathcal{T}HT^{-1} = H$$

How do we define this in a rigorous way?
Time Reversal Symmetry

The infinitesimal time evolution of a state $|n\rangle$ is generated by the Hamiltonian, and is given by

$$|n, t_0 = 0; t = \delta t\rangle = \left(1 - \frac{iH}{\hbar}\delta t\right)|n\rangle$$

If we consider applying the several operator at $t = 0$, then time evolving the state, then at $t = \delta t$ we have $(1 - \frac{iH}{\hbar}\delta t)\mathcal{T}|n\rangle$. Intuitively we expect that this should equal $\mathcal{T}|n, t_0 = 0; t = -\delta t\rangle$. Together this gives

$$\left(1 - \frac{iH}{\hbar}\delta t\right)\mathcal{T}|n\rangle = \mathcal{T}\left(1 - \frac{iH}{\hbar}(-\delta t)\right)|n\rangle$$

If this is to be true for any state $|n\rangle$ we find

$$-iH\mathcal{T}|n\rangle = \mathcal{T}iH|n\rangle$$
Time Reversal Symmetry

So we found:

\[-iHT|n\rangle = \mathcal{T}iH|n\rangle\]

If \( \mathcal{T} \) was a unitary operator then we may directly cancel the \( i \)'s and find

\[-HT = \mathcal{T}H.\]

This does not work: as it implies

\[HT|n\rangle = -\mathcal{T}H|n\rangle = -\epsilon_n\mathcal{T}|n\rangle\]

Therefore we find that the time reversal operator must be anti unitary

\([\mathcal{T}c|n\rangle = c^\ast \mathcal{T}|n\rangle, \text{ where } c \in \mathbb{C}\), this means

\[\mathcal{T}H = HT\]
What is the effect of time reversal on the wave function?

\[
\mathcal{T}\left| n \right\rangle = \int d^3 x' \mathcal{T}(\left| x' \right\rangle \left\langle x' \right| n) \quad \mathcal{T}\left| n \right\rangle = \int d^3 k' \mathcal{T}(\left| k' \right\rangle \left\langle k' \right| n)
\]

\[
= \int d^3 x' \mathcal{T} \left| n \right\rangle \langle x' | x \rangle * \\
= \int d^3 x' \left| n \right\rangle \langle x' | x \rangle *
\]

So this shows that \( \mathcal{T}\psi(x') \rightarrow \psi(x')^* \) and \( \mathcal{T}\psi(p') \rightarrow \psi(-p')^* \).

The representation we choose crucially matters! We also see that two successive applications gives \( \mathcal{T}^2 = 1 \).

We will now show that this story differs for spin full particles.
Time Reversal Symmetry

The spin eigenstates of a spin half particle can be written as

\[ |\hat{n}, +\rangle = e^{-iS_z \alpha/\hbar} e^{-iS_y \beta/\hbar} |+\rangle \text{ and } |\hat{n}, -\rangle = e^{-iS_z \alpha/\hbar} e^{-iS_y (\pi + \beta)/\hbar} |-\rangle. \]

Now, noting that \( \mathcal{T} |\hat{n}, +\rangle = |\hat{n}, -\rangle \) we deduce

\[ \mathcal{T} = \eta e^{-i\pi J_y/\hbar} K \]

\[ \mathcal{T} \left( \mathcal{T} \sum |j, m\rangle \langle j, m |n\rangle \right) = \mathcal{T} \left( \eta \sum e^{-i\pi J_y/\hbar} |j, m\rangle \langle j, m |n\rangle^* \right) \]

\[ = |\eta|^2 \sum e^{-2i\pi J_y/\hbar} |j, m\rangle \langle j, m |n\rangle \]

Using the properties of the angular momentum eigenstates under rotation we note \( e^{-2i\pi J_y/\hbar} |j, m\rangle = (-1)^{2j} |j, m\rangle \), therefore we find

\[ \mathcal{T}^2 |j \text{ half-integer}\rangle = - |j \text{ half-integer}\rangle \]

\[ \mathcal{T}^2 |j \text{ integer}\rangle = + |j \text{ integer}\rangle \]
Kramers Degeneracy

If we consider a system of charged particles in a static electric field, with $V(x) = e\phi(x)$ then $[\mathcal{T}, H] = 0$ will still hold as the electrostatic potential is a real function of the time-reversal operator $x$.

We consider a state $|n\rangle$ and its time reversed parter $\mathcal{T}|n\rangle$, these must have the same energy - following $H\mathcal{T}|n\rangle = \mathcal{T}H|n\rangle = \epsilon_n \mathcal{T}|n\rangle$.

So are these the same state or different states?
Kramers Degeneracy

The states $|n\rangle$ and $\mathcal{T}|n\rangle$ can only differ by a phase $\eta$ with $\eta = e^{i\delta}$.

$$\mathcal{T}|n\rangle = e^{i\delta}|n\rangle$$

Now applying $\mathcal{T}$ once more

$$\mathcal{T}^2|n\rangle = \mathcal{T}e^{i\delta}|n\rangle = e^{-i\delta}\mathcal{T}|n\rangle = e^{-i\delta}e^{i\delta}|n\rangle = +|n\rangle$$

Which cannot hold for spin one-half particle. So we must conclude that these are different degenerate states.

This holds as long at the system obeys time-reversal symmetry!
Earlier we noted that $\mathcal{T}\psi(p') \rightarrow \psi(-p')^*$. It is also useful to realise we can split the time reversal operator into two parts $\mathcal{T} = \tau K$ where $K$ is an operator which complex conjugates and $\tau$ is a unitary operator which acts on internal degrees of freedom.

$$\mathcal{T} H_{\text{Bulk}} \mathcal{T}^{-1} = \sum_k |\mathbf{-k}\rangle\langle\mathbf{-k}| \otimes \mathcal{T} H (\mathbf{k}) \mathcal{T}^{-1} = \sum_k |\mathbf{k}\rangle\langle\mathbf{k}| \otimes \tau H^* (\mathbf{-k}) \tau^{-1}$$

So $\tau H^* (\mathbf{-k}) \tau^{-1} = H (\mathbf{k})$ in the TRS bulk, and therefore the Hamiltonian must be symmetric to inversion in the Brillouin zone.
The Bernevig-Hughes-Zhang model

How do we construct time reversal invariant Hamiltonians from a lattice model?
We take two copies of the system, where ones $H^* = KHK$ and couple in the same manner as we did with the Chern insulators.

$$H_{\text{TRI}} = \begin{pmatrix} H & C \\ C^\dagger & H^* \end{pmatrix}$$
We look for a coupling with $\mathcal{T}^2 = 1$. We have two options: we choose $\mathcal{T} = s_x K$.

This gives

$$s_x K H_{\text{TRI}} (s_x K)^{-1} = H_{\text{TRI}}$$

If we chose the coupling such that $C = C^T$.
Additionally we may choose an even simpler solution with $\mathcal{T} = K$ and not even double the Hilbert space

$$H_{\text{TRI}} = \frac{H + H^*}{2} = K H_{\text{TRI}} K$$
The Bernevig-Hughes-Zhang model

We look for a coupling with $T^2 = -1$, therefore we choose $T = i s_y K$. This gives

$$i s_y K H_{\text{TRI}} (i s_y K)^{-1} = H_{\text{TRI}}$$

If we chose the coupling such that $C = -C^T$.

So we see different types of TRS must be satisfied by different couplings.
The Bernevig-Hughes-Zhang model

We now introduce a toy model for a time-reversal invariant topological insulator.

\[ H_{\text{BHZ}}(k) = s_0 \otimes \left[ (u + \cos k_x + \cos k_y)\sigma_z + \sin k_y \sigma_y \right] + s_z \otimes \sin k_x \sigma_x + s_x \otimes \hat{C} \]

For \( C = 0 \) this reduces to a 4 band model of HgTe.

---

**Fig. 8.1** Stripe dispersion relations of the BHZ model, with sublattice potential parameter \( u = -1.2 \). Right/left edge states (more than 60\% weight on the last/first two columns of unit cells) marked in dark red/light blue. (a): uncoupled layers, \( \hat{C} = 0 \). (b): Symmetric coupling \( \hat{C} = 0.3\sigma_x \) gaps the edge states out. (c): Antisymmetric coupling \( \hat{C} = 0.3\sigma_y \) cannot open a gap in the edge spectrum.
Edge States
The $\mathbb{Z}_2$ invariant

$\mathcal{T}^2 = -1$ demands that we have an equal number of left movers and right moves at any given energy in the bulk gap.

$$N_+(E) = N_-(E)$$

Due to Kramers degeneracy adiabatic deformations of the edge states can only change the number of edge states by 4, a pair of Kramers pairs. Therefore we define the $\mathbb{Z}_2$ as the parity of the number of Kramers pairs, this is a topological invariant.

$$\mathbb{Z}_2 = 0(1)$$ for even(odd) number of Kramers pairs.
Absence of Back Scattering

How does edge state transport behave under the influence of local disorder?

Fig. 8.3 Disordered region (gray) obstructing electrons in a two-dimensional phase-coherent conductor. The scattering matrix $S$ relates the amplitudes $a_L^{(in)}$ and $a_R^{(in)}$ of incoming waves to the amplitudes $a_L^{(out)}$ and $a_R^{(out)}$ of outgoing waves.
Absence of Back Scattering

Consider a lattice model where the unit cells form a square lattice of size $N_x \times N_y$. If the system is translational invariant in the $x$-direction we may write the Bloch states as

$$|l\pm\rangle = |k_{l,\pm}\rangle \otimes |\Phi_{l,\pm}\rangle$$

where the momentum eigenstates are given by

$$|k\rangle = \frac{1}{N_x} \sum_{m_x=1}^{N_x} e^{ikm_x} |m_x\rangle$$

and $|\Phi_{l,\pm}\rangle$ incorporates the transverse modes and the sublattice degree of freedom.

We will use these current-normalised states defining

$$|l\pm\rangle_c = \frac{1}{\sqrt{|v_{l,\pm}|}} |l\pm\rangle$$

with $v_{l,\pm}$ being the group velocity.
Absence of Back Scattering

A wave incident on the scattering region is characterized by a vector of coefficients, in the lead index $L, R$

$$a^{(in)} = \left( a^{(in)}_{L,1}, a^{(in)}_{L,2}, \ldots, a^{(in)}_{L,N}, a^{(in)}_{R,1}, a^{(in)}_{R,2}, \ldots, a^{(in)}_{R,N} \right)$$

$$a^{(out)} = \left( a^{(out)}_{L,1}, a^{(out)}_{L,2}, \ldots, a^{(out)}_{L,N}, a^{(out)}_{R,1}, a^{(out)}_{R,2}, \ldots, a^{(out)}_{R,N} \right)$$

The corresponding energy eigenstate reads

$$|\psi\rangle = \sum_{l=1}^{N} a^{(in)}_{L,l} |l, +, L\rangle_c + a^{(out)}_{L,l} |l, -, L\rangle_c + a^{(in)}_{R,l} |l, -, R\rangle_c + a^{(out)}_{L,l} |l, +, R\rangle_c$$

The scattering matrix $S$ relates the two vectors

$$a^{(out)} = Sa^{(in)} \quad S = \begin{pmatrix} r & t' \\ t & r' \end{pmatrix}$$
Absence of Back Scattering

Kramers degeneracy gives the connection between states.

\[
|l, -, L\rangle_c = T|l, +, L\rangle_c \\
|l, +, R\rangle_c = T|l, -, R\rangle_c
\]

We may write the eigenstates outside the scattering region as

\[
|\psi\rangle = \sum_{l=1}^{N} a_{L,l}^{(in)} |l, +, L\rangle_c + (Sa_{L,l}^{(in)}) |l, -, L\rangle_c \\
+ a_{R,l}^{(in)} |l, -, R\rangle_c + (Sa_{R,l}^{(in)}) |l, +, R\rangle_c
\]
Absence of Back Scattering

We now look for the time reversed state

\[- \mathcal{T} \ket{\psi} = \sum_{l=1}^{N} -a^{(\text{in})\ast}_{L,l} \ket{l, -, L}_c + (S^* a^{(\text{in})\ast})_{L,l} \ket{l, +, L}_c \]

\[- a^{(\text{in})\ast}_{R,l} \ket{l, +, R}_c + (S^* a^{(\text{in})\ast})_{L,l} \ket{l, -, R}_c \]

Comparing \ket{\psi} and \(- \mathcal{T} \ket{\psi}\) we must conclude

\[S = \begin{pmatrix} r & t' \\ t & r' \end{pmatrix} = -S^{T} = \begin{pmatrix} -r & -t \\ -t' & -r' \end{pmatrix} \]

We gives \(r = -r\) therefore \(r = 0\).