6

Tight-binding models

6.1. tight-binding models

Tight-binding models are effective tools to describe the motion of electrons in solids. Here, we assume that the system is a discrete lattice and electrons can only stay on the lattice site. The kinetic energy is included by allowing electrons to hop from one site to another.

6.1.1. Example 1: a one-band model

Lets consider a 1D lattice with one atom per unit cell. For each atom, we consider only one quantum state. The creation (annihilation) operator $c_i^\dagger$ (c_i) creates and annihilates one particle on site i. The non-interacting Hamiltonian can be written as

$$H = -\sum_{ij}\left(t_{ij}\, c_i^\dagger\, c_j + t_{ji}\, c_j^\dagger\, c_i\right) + \sum_i V_i\, c_i^\dagger\, c_i \tag{6.1}$$

The first term $\sum_{ij} t_{ij}\, c_i^\dagger\, c_j$ describes hoppings from site j to i. The second term $\sum_{ij} t_{ji}\, c_j^\dagger\, c_i$ describes the hopping from i to j. The last term is the potential energy, which tells us how much energy we need to put an electron on each site V_i. Because the lattice contains only one type of atoms, V_i = constant (i.e. translational symmetry). Because H is Hermitian, we find that $t_{ij} = t_{ji}^*$ and V_i is a real number.

Proof:

$$H^\dagger = H \tag{6.2}$$

$$\left[-\sum_{ij}\left(t_{ij}\, c_i^\dagger\, c_j + t_{ji}\, c_j^\dagger\, c_i\right) + \sum_i V_i\, c_i^\dagger\, c_i\right]^\dagger = -\sum_{ij}\left(t_{ij}\, c_i^\dagger\, c_j + t_{ji}\, c_j^\dagger\, c_i\right) + \sum_i V_i\, c_i^\dagger\, c_i \tag{6.3}$$

$$-\sum_{ij}\left(t_{ij}^*\, c_j^\dagger\, c_i + t_{ji}\, c_i^\dagger\, c_j\right) + \sum_i V_i^*\, c_i^\dagger\, c_i = -\sum_{ij}\left(t_{ij}\, c_i^\dagger\, c_j + t_{ji}\, c_j^\dagger\, c_i\right) + \sum_i V_i\, c_i^\dagger\, c_i \tag{6.4}$$

By comparing the two sides, we find that $t_{ij}^* = t_{ji}$ and $V_i = V_i^*$.

Therefore, we can simplify the Hamiltonian

$$H = -\sum_{ij}\left(t_{ij}\, c_i^\dagger\, c_j + t_{ij}^*\, c_j^\dagger\, c_i\right) + V\sum_i c_i^\dagger\, c_i = -\sum_{ij}\left(t_{ij}\, c_i^\dagger\, c_j + t_{ij}^*\, c_j^\dagger\, c_i\right) + V\ N \tag{6.5}$$

In the last term, $N = \sum_i c_i^\dagger\, c_i$ is the total number of electrons in the system. Because $V\ N$ is a constant, this term just shifts the total energy by a constant, and thus has no other physical contribution (can be ignored if we are not interested in the total energy).

For tight-binding models, a typically approximation is to assume that electrons can only hop to its nearest-neighbor sites. In reality, long-range hopping is allowed, but their amplitudes are small (decay exponentially as distance increases). Therefore, in many cases, we just need the nearest-neighbor hopping terms to describe the systems. Due to the translational symmetry, all nearest-neighbor hoppings shall have the same hopping strength. If we assume that this hopping strength is a real number t, the Hamiltonian is

$$H = -t\sum_{\langle ij\rangle}\left(c_i^\dagger\, c_j + c_j^\dagger\, c_i\right) = -t\sum_i c_i^\dagger\, c_{i+1} + h.c. \tag{6.6}$$

Here, ⟨ij⟩ implies that i and j must be neighbors. On the right hand side, *h.c.* means Hermitian conjugate. Notice that the second term is in fact the Hermitian conjugate of the first term, so we just use *h.c.* to represent it.

Fourier series: typically, Fourier series is used to describe a periodic function in the real space, which will have a describe set of wave vectors in the k space. Here, it is the opposite. We have a continuous k-space and it is periodic (the Brillouin zone), but the real space is discrete.

$$c_k = \frac{\sqrt{a}}{\sqrt{2\pi}} \sum_i c_i \, e^{-i k x} \tag{6.7}$$

$$c_i = \frac{\sqrt{a}}{\sqrt{2\pi}} \int_{BZ} dk \, c_k \, e^{i k x} \tag{6.8}$$

Because $\{c_i, \, c_j^\dagger\} = \delta_{ij}$, it is easy to check that

$$\begin{aligned} &\{c_k, \, c_{k'}^\dagger\} = \Big\{ \frac{1}{\sqrt{2\pi/a}} \sum_i c_i \, e^{-i k x_i}, \; \frac{1}{\sqrt{2\pi/a}} \sum_j c_j^\dagger \, e^{i k' x_j} \Big\} = \\ &\sum_{i,j} \{c_i, \, c_j^\dagger\} \frac{1}{2\pi/a} e^{-i k x_i} \, e^{i k' x_j} = \sum_{i,j} \delta_{i,j} \frac{1}{2\pi/a} e^{-i (k-k') x_i} = \sum_i \frac{1}{2\pi/a} e^{-i (k-k') i a} = a \, \delta((k-k') a) = \delta(k-k') \end{aligned} \tag{6.9}$$

If we know $\{c_k, \, c_{k'}^\dagger\} = \delta(k-k')$, we can also show that $\{c_i, \, c_j^\dagger\} = \delta_{ij}$.

$$\begin{aligned} &\{c_i, \, c_j^\dagger\} = \Big\{ \frac{1}{\sqrt{2\pi/a}} \int_{BZ} dk \, c_k \, e^{i k x_i}, \; \frac{1}{\sqrt{2\pi/a}} \int_{BZ} dk' \, c_{k'}^\dagger \, e^{-i k' x_j} \Big\} = \\ &\frac{1}{2\pi/a} \int_{BZ} dk \int_{BZ} dk' \{c_k, \, c_{k'}^\dagger\} \, e^{i k x_i} \, e^{-i k' x_j} = \frac{1}{2\pi/a} \int_{BZ} dk \int_{BZ} dk' \, \delta(k-k') \, e^{i k x_i} \, e^{-i k' x_j} = \frac{1}{2\pi/a} \int_{BZ} dk \, e^{i k (x_i - x_j)} = \delta_{ij} \end{aligned} \tag{6.10}$$

Hamiltonian in k-space We can transfer the Hamiltonian into the k-space. In real space, it looks like

$$H = -t \sum_i c_i^\dagger \, c_{i+1} + h.c. \tag{6.11}$$

The first term is

$$\begin{aligned} &\sum_j c_j^\dagger \, c_{j+1} = \sum_j \frac{1}{\sqrt{2\pi/a}} \int_{BZ} dk \, c_k^\dagger \, e^{-i k a j} \frac{1}{\sqrt{2\pi/a}} \int_{BZ} dk' \, c_{k'} \, e^{i k' (j+1) a} = \\ &\sum_j \frac{1}{2\pi/a} \int_{BZ} dk \int_{BZ} dk' \, c_k^\dagger \, c_{k'} \, e^{i k' a} \, e^{-i (k'-k) a j} = \int_{BZ} dk \int_{BZ} dk' \, c_k^\dagger \, c_{k'} \, e^{i k' a} \, \delta(k-k') = \int_{BZ} dk \, c_k^\dagger \, c_k \, e^{i k a} \end{aligned} \tag{6.12}$$

The second term is the Hermitian conjugate of the first term, so

$$h.c. = \Big(\sum_j c_j^\dagger \, c_{j+1} \Big)^\dagger = \Big(\int_{BZ} dk \, c_k^\dagger \, c_k \, e^{i k a} \Big)^\dagger = \int_{BZ} dk \, c_k^\dagger \, c_k \, e^{-i k a} \tag{6.13}$$

$$H = -t \sum_i c_i^\dagger \, c_{i+1} + h.c. = -t \int_{BZ} dk \, c_k^\dagger \, c_k \, e^{i k a} - t \int_{BZ} dk \, c_k^\dagger \, c_k \, e^{-i k a} = -2 t \int_{BZ} dk \, c_k^\dagger \, c_k \cos k a = \int_{BZ} dk \, (-2 t \cos k a) \, c_k^\dagger \, c_k \tag{6.14}$$

For a solid, we know that the total energy of electrons (ignore interactions) is

$$E = \sum_n \int_{BZ} dk \; \epsilon_n(k) \, n_n(k) \tag{6.15}$$

where $\sum_n$ sums over all bands, $\epsilon_n(k)$ is the dispersion relation for band n and $n_n(k)$ is the occupation number for the Bloch wave state in band n with momentum k. In second quantization, this formula implies that the Hamiltonian is

$$H = \sum_n \int_{BZ} dk \; \epsilon_n(k) \, \gamma_{n,k}^\dagger \, \gamma_{n,k} \tag{6.16}$$

Here, $\gamma_{n,k}^\dagger$ is the creation operator for a Bloch wave $\psi_{n\,k}(r) = u_{n\,k}(r) \, e^{i k r}$. If we compare this formula with the tight-binding Hamiltonian (in k-space), we find immediately that the tight-binding model we considered here has only one energy band. And our $c_k^\dagger$ operator is in fact the creation operator for Bloch waves. And the dispersion relation for this band is $\epsilon_k = -2 t \cos k a$. Notice that ϵ_k is a periodic function of k with periodicity $2\pi/a$, which is exactly what we expect for Bloch waves.

6.1.2. Discrete Fourier transformation and Fourier series

In the formula above, we treat k as a continuous variable. In many cases, it is useful to treat k as discrete values (where $\int dk$ turns into $\sum_k$). Here, we briefly discuss about the discrete Fourier transform. Consider a discrete function f_i, where $i = 1, 2, 3 \ldots N$ marks different lattice site. In addition, we assume the periodic boundary condition $f_{N+i} = f_i$. We can define

$$\tilde{f}_k = \frac{1}{\sqrt{N}} \sum_j f_j e^{-i k a j} \tag{6.17}$$

where a is the lattice constant and j marks different lattice sites.

And it is easy to prove that

$$f_j = \frac{1}{\sqrt{N}} \sum_k \tilde{f}_k e^{i k a j} \tag{6.18}$$

Same as *j*, the wave vector *k* here is also a discrete variable. This is because $f_{N+i} = f_N$ (our system has a finite size)

$$f_{j+N} = \frac{1}{\sqrt{N}} \sum_k \tilde{f}_k e^{i k (a j + a N)} = \frac{1}{\sqrt{N}} \sum_k \tilde{f}_k e^{i k a j} e^{i k a N} \tag{6.19}$$

$$f_j = \frac{1}{\sqrt{N}} \sum_k \tilde{f}_k e^{i k a j} \tag{6.20}$$

If we compare the two equations, we find that $e^{i k a N}$=1, which implies

$$k = \frac{2\pi m}{N a} = \frac{2\pi m}{L} \tag{6.21}$$

Here $L = N a$ is the size of the system. For infinite systems, $L \to \infty$, the discrete sum $\sum_k$ turns into an integral $\int dk$.

Same as j, k also have a periodicity, and the periodicity is also *N*. This periodicity is just the Brillouin zone.

Define $k_m = 2\pi m / L$ and by definition,

$$\tilde{f}_{k_m} = \frac{1}{\sqrt{N}} \sum_j f_j e^{-i k_m a j} = \frac{1}{\sqrt{N}} \sum_j f_j e^{-i \left(\frac{2\pi}{N a} m\right) a j} = \frac{1}{\sqrt{N}} \sum_j f_j e^{-i \frac{2\pi}{N} m j} \tag{6.22}$$

For k_{m+N}, by definition

$$\tilde{f}_{k_{m+N}} = \frac{1}{\sqrt{N}} \sum_j f_j e^{-i k_{m+N} a j} = \frac{1}{\sqrt{N}} \sum_j f_j e^{-i \left[\frac{2\pi}{N a} (m+N)\right] a j} = \frac{1}{\sqrt{N}} \sum_j f_j e^{-i \frac{2\pi}{N} m j} e^{-i 2\pi j} \tag{6.23}$$

Because j is an integer, the factor $e^{-i 2\pi j} = 1$. Therefore

$$\tilde{f}_{k_{m+N}} = \frac{1}{\sqrt{N}} \sum_j f_j e^{-i \frac{2\pi}{N} m j} = \tilde{f}_{k_m} \tag{6.24}$$

As a result, we can limit the value of m to be $-N/2 \le m < N/2$. If m is not in this range, the value of $\tilde{f}_{k_m}$ can be obtained using the periodic condition $\tilde{f}_{k_{m+N}} = \tilde{f}_{k_m}$

$$m = -\frac{N}{2}, -\frac{N}{2} + 1, \ \ldots \frac{N}{2} - 1, \tag{6.25}$$

For the wave vector k, this means that

$$\tilde{f}_{k + 2\pi/a} = \tilde{f}_k \tag{6.26}$$

So, we can confine the value of k into the range of $-\pi/a \le k < \pi/a$, which is the first Brillouin zone. For k outside the first Brillouin zone, we can find the corresponding $\tilde{f}_k$ using the periodicity $\tilde{f}_{k+2\pi/a} = \tilde{f}_k$

$$k = -\frac{\pi}{a}, -\frac{\pi}{a} + 1 \times \frac{2\pi}{L}, -\frac{\pi}{a} + 2 \times \frac{2\pi}{L}, \ \ldots \frac{\pi}{a} - \frac{2\pi}{L}, \tag{6.27}$$

Two useful Identities:

$$\frac{1}{N}\sum_k e^{i k a j} e^{-i k a j'} = \delta_{j,j'} \tag{6.28}$$

$$\frac{1}{N}\sum_j e^{i k a j} e^{-i k' a j} = \delta_{k,k'} \tag{6.29}$$

6.1.3. Example 2: a two-band model in 1D

Now, let us consider a slightly more complicated situation, i.e. a 1d chain formed by two different types of atoms (*a* and *b*).

$$H = -t\sum_i \left(a_i^\dagger b_i + b_i^\dagger a_{i+1} + h.c.\right) + V_a\sum_i a_i^\dagger a_i + V_b\sum_i b_i^\dagger b_i \tag{6.30}$$

Here, the position of *a* sites in the jth unit cells is

$$r = a\times j + r_a \tag{6.31}$$

and the position of *b* sites in the jth unit cells is

$$r = a\times j + r_b \tag{6.32}$$

where, *a* is the lattice constant (the size of a unit cell)

Discrete Fourier transform (as discussed in the previous section)

$$a_k = \frac{1}{\sqrt{N}}\sum_i a_i e^{i k x} \tag{6.33}$$

$$a_i = \frac{1}{\sqrt{N}}\sum_k a_k e^{-i k x} \tag{6.34}$$

$$b_k = \frac{1}{\sqrt{N}}\sum_i a_i e^{i k x} \tag{6.35}$$

$$b_i = \frac{1}{\sqrt{N}}\sum_k a_k e^{-i k x} \tag{6.36}$$

Therefore, we know that

$$\begin{aligned}\sum_j a_j^\dagger b_j = \sum_j \frac{1}{\sqrt{N}}\sum_k a_k^\dagger e^{-i k (a j+r_a)} \frac{1}{\sqrt{N}}\sum_{k'} b_{k'} e^{i k' (a j+r_b)} = \frac{1}{N}\sum_k\sum_{k'} a_k^\dagger b_{k'}\sum_j e^{-i k (a j+r_a)} e^{i k' (a j+r_b)} = \\ \sum_k\sum_{k'} a_k^\dagger b_{k'}\frac{1}{N}\sum_j e^{-i (k-k') a j} e^{-i k r_a} e^{i k' r_b} = \sum_k\sum_{k'} a_k^\dagger b_{k'}\delta_{k,k'} e^{-i k r_a} e^{i k' r_b} = \sum_k a_k^\dagger b_k e^{i k (r_b-r_a)}\end{aligned} \tag{6.37}$$

This conclusion is generic. In k-space, hoppings from b to a result in a term in Hamiltonian $\sim\sum_k a_k^\dagger b_k e^{i k (r_b-r_a)}$. The phase factor here is determined by how far the electron hop, i.e. $r_b - r_a$. If we apply the same conclusion to the hopping from a to b, we find immediately that

$$\sum_j b_j^\dagger a_{j+1} = \sum_k b_k^\dagger a_k e^{i k(r_a-r_b)} \tag{6.38}$$

If $r_b - r_a = \frac{a}{2}$,

$$\sum_j a_j^\dagger b_j = \sum_k a_k^\dagger b_k e^{i k a/2} \tag{6.39}$$

$$\sum_j b_j^\dagger a_{j+1} = \sum_k b_k^\dagger a_k e^{i k a/2} \tag{6.40}$$

For the potential term, we also know that

$$\sum_i a_i^\dagger a_i = \sum_k a_k^\dagger a_k \tag{6.41}$$

and

$$\sum_i b_i^\dagger b_i = \sum_k b_k^\dagger b_k \tag{6.42}$$

Therefore, in the k-space, the Hamiltonian looks like

$$H = -t \sum_k \left(a_k^\dagger b_k e^{i k a/2} + b_k^\dagger a_k e^{i k a/2} + b_k^\dagger a_k e^{-i k a/2} + a_k^\dagger b_k e^{-i k a/2} \right) + V_a \sum_k a_k^\dagger a_k + V_b \sum_k b_k^\dagger b_k = \\ -2\, t \sum_k \left[\left(a_k^\dagger b_k \cos\left(\frac{k a}{2}\right) + b_k^\dagger a_k \cos\left(\frac{k a}{2}\right) \right) + V_a a_k^\dagger a_k + V_b b_k^\dagger b_k \right] \tag{6.43}$$

We can write it in a matrix form:

$$H = \sum_k \left(a_k^\dagger,\ b_k^\dagger \right) \begin{pmatrix} V_a & -2\, t\, \cos\left(\frac{k a}{2}\right) \\ -2\, t \cos\left(\frac{k a}{2}\right) & V_b \end{pmatrix} \begin{pmatrix} a_k \\ b_k \end{pmatrix} \tag{6.44}$$

For any tight-binding models with two quantum states per unit cell, the Hamiltonian can be written in terms of a two-by-two matrix in the k space:

$$H = \sum_k \left(a_k^\dagger,\ b_k^\dagger \right) \mathcal{H}(k) \begin{pmatrix} a_k \\ b_k \end{pmatrix} \tag{6.45}$$

where $\mathcal{H}$ is a 2×2 Hermitian matrix as a function of k. It is called the *kernel of the Hamiltonian*. Because $\mathcal{H}(k)$ contains all the information of the Hamiltonian H, it is often called the Hamiltonian in literature. However, it is important to keep in mind that $\mathcal{H}(k)$ is only part of the Hamiltonian. The Hamiltonian H must be gauge invariant, but $\mathcal{H}(k)$ is not. For example, if we change $a^\dagger \to a^\dagger e^{i\phi}$, the Hamiltonian H is invariant, but the kernel $\mathcal{H}(k)$ is NOT.

For the model considered here,

$$\mathcal{H}(k) = \begin{pmatrix} V_a & -2\, t\, \cos\left(\frac{k a}{2}\right) \\ -2\, t \cos\left(\frac{k a}{2}\right) & V_b \end{pmatrix} \tag{6.46}$$

For more generic cases, if one have *m* quantum states per unit cell, $\mathcal{H}$(k) will be a *m*×*m* Hermitian matrix.

When we have a matrix, we know what to do. We need to find the eigenvalue and eigenfunctions of the Hamiltonian, i.e. diagonalizing the matrix. The next two sections discuss the physical meanings of the eigenvalues and eigenvectors of $\mathcal{H}(k)$.

6.1.4. Eigenvalues of $\mathcal{H}(k)$

The eigenvalues of $\mathcal{H}(k)$ give us the dispersion relations. At each k point, one can define a unitary transformation

$$\begin{pmatrix} c_k \\ d_k \end{pmatrix} = U_k^{-1} \begin{pmatrix} a_k \\ b_k \end{pmatrix} \tag{6.47}$$

$$\begin{pmatrix} c_k^\dagger & d_k^\dagger \end{pmatrix} = \begin{pmatrix} a_k^\dagger & b_k^\dagger \end{pmatrix} U_k \tag{6.48}$$

where $U_k^{-1} = U_k^\dagger$

$$H = \sum_k \left(c_k^\dagger,\ d_k^\dagger \right) U_k^+ \mathcal{H}\, U_k \begin{pmatrix} c_k \\ d_k \end{pmatrix} \tag{6.49}$$

If we choose U_k such that $\mathsf{U_k}^+ \mathcal{H}\, \mathsf{U_k}$ is a diagonal matrix,

$$U_k \mathcal{H}\, U_k^{-1} = \begin{pmatrix} \epsilon_c & 0 \\ 0 & \epsilon_d \end{pmatrix} \tag{6.50}$$

Then, the Hamiltonian becomes

$$H = \sum_k \left(c_k^\dagger,\ d_k^\dagger \right) \begin{pmatrix} \epsilon_c & 0 \\ 0 & \epsilon_d \end{pmatrix} \begin{pmatrix} c_k \\ d_k \end{pmatrix} = \sum_k \epsilon_c(k)\, c_k^\dagger c_k + \sum_k \epsilon_d(k)\, d_k^\dagger d_k \tag{6.51}$$

We know that using Bloch waves (which are eigenstates of the Hamiltonian), the total energy is

$$E = \sum_{k,m} \epsilon_m(k)\, n_m(k) = \sum_k \epsilon_1\, n_1(k) + \sum_k \epsilon_2\, n_2(k) + \sum_k \epsilon_3\, n_3(k) + \ldots \tag{6.52}$$

where $\epsilon_m(k)$ is the dispersion relation for the mth band and $n_m(k)$ is the occupation number for the quantum state in the mth band with momentum k. Therefore, in terms of Bloch waves, the Hamiltonian should take the form:

$$H = \sum_k \epsilon_1(k)\, {\gamma_{1,k}}^\dagger \gamma_{1,k} + \sum_k \epsilon_2(k)\, {\gamma_{2,k}}^\dagger \gamma_{2,k} + \sum_k \epsilon_3(k)\, {\gamma_{3,k}}^\dagger \gamma_{3,k} + \ldots \tag{6.53}$$

where ${\gamma_{n,k}}^\dagger$ is the creation operator for a Bloch wave in band n at momentum k. If we compare this formula with Eq. 6.51, we find that in this model we get two energy bands. ${c_k}^\dagger$ creates a Bloch wave with momentum k in one of the bands. ${d_k}^\dagger$ creates a Bloch wave with momentum k in the other band. For the bands created by ${c_k}^\dagger$, its dispersion relation is $\epsilon_c(k)$, while the other hand has a dispersion relation $\epsilon_d(k)$.

For the problem we considered here, the dispersions are

$$\epsilon_c = \frac{V_a + V_b}{2} - \sqrt{\left[2\, t \cos\left(\frac{k\, a}{2}\right)\right]^2 + \left(\frac{V_a - V_b}{2}\right)^2} \tag{6.54}$$

$$\epsilon_d = \frac{V_a + V_b}{2} + \sqrt{\left[2\, t \cos\left(\frac{k\, a}{2}\right)\right]^2 + \left(\frac{V_a - V_b}{2}\right)^2} \tag{6.55}$$

which are the eigenvalues of the matrix $\mathcal{H}(k)$.

If we have a model with m orbitals per unit cell, $\mathcal{H}(k)$ would be a $m \times m$ matrix. This matrix will have m eigenvalues, which results in a m-band model and each eigenvalue $\epsilon_n(k)$ is the dispersion relation for one of the bands.

Bottom line: eigenvalues of $\mathcal{H}$(k) tell us the dispersion relation.

6.1.5. Eigenvectors of $\mathcal{H}(k)$

The eigenvectors $\mathcal{H}(k)$ gives us the Bloch waves.

Let's start from the two band model discussed above as an example. Here, the two-by-two matrix $\mathcal{H}(k)$ has two eigenvalues, $\epsilon_c(k)$ and $\epsilon_d(k)$. For each eigenvalue, the corresponding eignevector is a 2-component vector:

$$\mathcal{H}\begin{pmatrix} {v^{(-)}}_a(k) \\ {v^{(-)}}_b(k) \end{pmatrix} = \epsilon_c \begin{pmatrix} {v^{(-)}}_a(k) \\ {v^{(-)}}_b(k) \end{pmatrix} \tag{6.56}$$

$$\mathcal{H}\begin{pmatrix} {v^{(+)}}_a(k) \\ {v^{(+)}}_b(k) \end{pmatrix} = \epsilon_d \begin{pmatrix} {v^{(+)}}_a(k) \\ {v^{(+)}}_b(k) \end{pmatrix} \tag{6.57}$$

It can be easily checked that

$${c_k}^\dagger = {v^{(-)}}_a(k)\, {a_k}^\dagger + {v^{(-)}}_b(k)\, {b_k}^\dagger \tag{6.58}$$

$${d_k}^\dagger = {v^{(+)}}_a(k)\, {a_k}^\dagger + {v^{(+)}}_b(k)\, {b_k}^\dagger \tag{6.59}$$

We know that $a^\dagger$ and $b^\dagger$ are the creation operators for plane waves (because they come from Fourier transformations) and we also know that ${c_k}^\dagger$ and ${d_k}^\dagger$ create Bloch waves. The relation between plane waves and Bloch waves is

$$\psi_{n,k}(x) = u_{n,k}(x)\, \frac{e^{i\, k\, x}}{\sqrt{N}} \tag{6.60}$$

So, the coefficients ${v^{(\pm)}}_a$ and ${v^{(\pm)}}_b(k)$ here are in fact the discrete version of $u_{n,k}(x)$.

Bottom line: eigenvectors of $\mathcal{H}$(k) tell us $u_{n,k}$.

We have learned that the Berry connection $\mathcal{A}_n(k)$ is defined as

$$\vec{\mathcal{A}}_n = -i \left\langle u_{n,k} \middle| \nabla_{\vec{k}} \middle| u_{n,k} \right\rangle \tag{6.61}$$

For continuous models, this formula means

$$\vec{\mathcal{A}}_n = -i \int d\, x\, \left[u_{n,k}(x)\right]^* \nabla_{\vec{k}}\, u_{n,k}(x) \tag{6.62}$$

For lattice models (tight-binding models)

$$\vec{\mathcal{A}}_- = -i \begin{pmatrix} v^{(-)}{}_a(k)^* & v^{(-)}{}_b(k)^* \end{pmatrix} \nabla_{\vec{k}} \begin{pmatrix} v^{(-)}{}_a(k) \\ v^{(-)}{}_b(k) \end{pmatrix} \tag{6.63}$$

$$\vec{\mathcal{A}}_+ = -i \begin{pmatrix} v^{(+)}{}_a(k)^* & v^{(+)}{}_b(k)^* \end{pmatrix} \nabla_{\vec{k}} \begin{pmatrix} v^{(+)}{}_a(k) \\ v^{(+)}{}_b(k) \end{pmatrix} \tag{6.64}$$

In general, for a m band model, where $\mathcal{H}(k)$ is a $m \times m$ matrix, each of the eigenvector gives us the Bloch wavefunction for one of the m bands. We can use each eigenvector to compute the Berry connection, the Berry curvature and the Chern number for that band

Bottom line: the eigenvectors are Bloch wave functions.

6.1.6. Summary

- The key properties of a tight-binding model is coded in $\mathcal{H}_k$, which is a $m \times m$ Hermitian matrix. Each component is a function of k. (In other words, $\mathcal{H}_k$ is a matrix function of k.)
- The eigenvalues if $\mathcal{H}$ (as a function of k) gives the band structure (the dispersion relation) for each bands $\epsilon_n(k)$ with $n = 1, 2, \ldots m$
- The eigenvectors as a function of k gives the Bloch wave: $u_{n,k}(\alpha)$, where n=1,2,…, m is the band index and α=1,2…m marks different site in a unit cell.
- Using these $u_{n,k}(\alpha)$, we can define the Berry connection, Berry curvature and the Chern number. Just replace the integral in the real space by summing over $\alpha = 1 \ldots m$

Before 80s, people stop at step #2 (computing the eigenvalues), without calculating the eigenvectors. This is because eigenvectors are the wavefunction, which cannot be measured directly in experiments. After the discovery of topological insulators, physicists realized that although the wavefunction can not be measured directly, it contains the topological information, which is a physical observable (e.g. the Hall conductivity).

6.1.7. One band model

In a one-band model, the eigenvector is trivial, which is just identity. Therefore, the Berry connection $\vec{\mathcal{A}}_n = -i \int d^d x \left[u_{n,k}(x)\right]^* \nabla_{\vec{k}} u_{n,k}(x)$ must be trivial $\vec{\mathcal{A}}_n = 0$. This tells us that a one-band model can never show quantum Hall effect. We need at least two bands. In fact, one can prove that the total Chern number summing over all bands must be zero.

6.2. an example of a topologically nontrivial insulator

Let's consider a 2-band model, whose kernel is

$$\mathcal{H}(k) = \begin{pmatrix} -2\,t\cos k_x - 2\,t\cos k_y - \mu & \Delta\left(\sin k_x - i \sin k_y\right) \\ \Delta\left(\sin k_x + i \sin k_y\right) & 2\,t\cos k_x + 2\,t\cos k_y + \mu \end{pmatrix} \tag{6.65}$$

Here, we assume Δ>0 and t>0 and -4 |t|<μ<4 |t|. In real space, this Hamiltonian corresponds to a square lattice and on each site there are two quantum states. Please notice that here I use μ to refer to a control parameter, which is NOT the chemical potential.

6.2.1. Two-band models and Pauli matrices

For any two-band models, $\mathcal{H}$ is a two-by-two Hermitian matrix. For a two-by-two Hermitian matrix, one can always separate it into the identity and Pauli matrices I, and σ_i

$$\mathcal{H} = \mathcal{H}_0(k)\, I \; + \; \mathcal{H}_x(k)\, \sigma_x + \mathcal{H}_y(k)\, \sigma_y + \mathcal{H}_z(k)\, \sigma_z \tag{6.66}$$

where $\mathcal{H}_0(k)$, $\mathcal{H}_x(k)$, $\mathcal{H}_y(k)$ and $\mathcal{H}_z(k)$ are real functions of k. If we define $\vec{\mathcal{H}}(k) = \left(\mathcal{H}_x(k),\ \mathcal{H}_y(k),\ \mathcal{H}_z(k)\right)$, we find that

$$\mathcal{H} = \mathcal{H}_0(k)\, I \; + \; \vec{\mathcal{H}}(k) \cdot \vec{\sigma} \tag{6.67}$$

This Hamiltonian is rather similar to a spin S=1/2 under magnetic fields.

$$H = \text{constant} + \mu\, \vec{B} \cdot \vec{\sigma} \tag{6.68}$$

For the case studied here,

$$\mathcal{H}_0(k) = 0 \tag{6.69}$$

$$\mathcal{H}_x(k) = \Delta \sin(k_x) \tag{6.70}$$

$$\mathcal{H}_y(k) = \Delta \sin(k_y) \tag{6.71}$$

$$\mathcal{H}_z(k) = -2\,t\cos k_x - 2\,t\cos k_y - \mu \tag{6.72}$$

The eigenvalues of $\mathcal{H}$ are

$$E_\pm = \mathcal{H}_0(k) \pm \left|\vec{\mathcal{H}}(k)\right| = \mathcal{H}_0(k) \pm \sqrt{\mathcal{H}_x(k)^2 + \mathcal{H}_y(k)^2 + \mathcal{H}_z(k)^2} \tag{6.73}$$

This is pretty much the same as the spin system, where the eigen-energies of the are

$$E = \text{constant} \pm \mu \left|\vec{B}\right| \tag{6.74}$$

For a two band model, the top band has energy $E_+ \geq \mathcal{H}_0$, while the lower band has $E_- \leq \mathcal{H}_0$. The energy gap between the two bands is:

$$\Delta(k) = E_+(k) - E_-(k) = 2\left|\vec{\mathcal{H}}(k)\right| \tag{6.75}$$

As long as $\vec{\mathcal{H}}(\mathtt{k}) \neq \mathtt{0}$, the two band will not cross with each other. Let's now focus on the lower band (E_-), its eigenvector is

$$u^{(I)}{}_-(k) = \frac{1}{\mathcal{N}^{(I)}}\begin{pmatrix} \mathcal{H}_z(k) - \left|\vec{\mathcal{H}}(k)\right| \\ \mathcal{H}_x(k) + i\,\mathcal{H}_y(k) \end{pmatrix} \tag{6.76}$$

This wavefunction is singular if $\mathcal{H}_x(k) = \mathcal{H}_y(k) = 0$ and $\mathcal{H}_z > 0$,

$$u^{(I)}{}_-(k) = \frac{1}{\mathcal{N}^{(I)}}\begin{pmatrix} \mathcal{H}_z - \left|\vec{\mathcal{H}}(k)\right| \\ \mathcal{H}_x(k) + i\,\mathcal{H}_y(k) \end{pmatrix} = \frac{1}{\mathcal{N}^{(I)}}\begin{pmatrix} \mathcal{H}_z - \sqrt{\mathcal{H}_z(k)^2} \\ 0 \end{pmatrix} = \frac{1}{\mathcal{N}^{(I)}}\begin{pmatrix} \mathcal{H}_z - |\,\mathcal{H}_z(k)\,| \\ 0 \end{pmatrix} = \frac{1}{\mathcal{N}^{(I)}}\begin{pmatrix} 0 \\ 0 \end{pmatrix} \tag{6.77}$$

This is indeed a problem for the Hamiltonian shown above. At $k_x = k_y = \pi$.

$$\mathcal{H}_0(k) = 0 \tag{6.78}$$

$$\mathcal{H}_x(k) = \Delta \sin(k_x) = 0 \tag{6.79}$$

$$\mathcal{H}_y(k) = \Delta \sin(k_y) = 0 \tag{6.80}$$

$$\mathcal{H}_z(k) = -2\,t\cos k_x - 2\,t\cos k_y - \mu = 4\,t - \mu > 0 \tag{6.81}$$

Remember that we assumed $|\,\mu\,| < 4\,t$. In fact, there is another way to write down the same eigenvector (i.e. a phase shift)

$$u^{(\mathrm{II})}{}_-(k) = \frac{1}{\mathcal{N}^{(I)}}\begin{pmatrix} \mathcal{H}_z - \left|\vec{\mathcal{H}}(k)\right| \\ \mathcal{H}_x(k) + i\,\mathcal{H}_y(k) \end{pmatrix} \times \frac{\frac{\mathcal{H}_z + \left|\vec{\mathcal{H}}(k)\right|}{\mathcal{H}_x(k) + i\,\mathcal{H}_y(k)}}{\left|\frac{\mathcal{H}_z + \left|\vec{\mathcal{H}}(k)\right|}{\mathcal{H}_x(k) + i\,\mathcal{H}_y(k)}\right|} = \frac{1}{\mathcal{N}^{(\mathrm{II})}}\begin{pmatrix} \frac{-\mathcal{H}_x(k)^2 - \mathcal{H}_y(k)^2}{\mathcal{H}_x(k) + i\,\mathcal{H}_y(k)} \\ \mathcal{H}_z + \left|\vec{\mathcal{H}}(k)\right| \end{pmatrix} = \frac{1}{\mathcal{N}^{(\mathrm{II})}}\begin{pmatrix} -\mathcal{H}_x(k) + i\,\mathcal{H}_y(k) \\ \mathcal{H}_z + \left|\vec{\mathcal{H}}(k)\right| \end{pmatrix} \tag{6.82}$$

The two wavefunctions $u^{(I)}{}_-(k)$ and $u^{(\mathrm{II})}{}_-(k)$ differ by a phase

$$u^{(\mathrm{II})}{}_-(k) = u^{(I)}{}_-(k)\, e^{i\,\phi(k)} \tag{6.83}$$

where

$$e^{i\,\phi(k)} = \frac{\frac{\mathcal{H}_z + \left|\vec{\mathcal{H}}(k)\right|}{\mathcal{H}_x(k) + i\,\mathcal{H}_y(k)}}{\left|\frac{\mathcal{H}_z + \left|\vec{\mathcal{H}}(k)\right|}{\mathcal{H}_x(k) + i\,\mathcal{H}_y(k)}\right|} \tag{6.84}$$

This new wavefunction is well defined at $\mathcal{H}_x(k) = \mathcal{H}_y(k) = 0$ and $\mathcal{H}_z(k) > 0$. However, it is NOT well defined at $\mathcal{H}_x(k) = \mathcal{H}_y(k) = 0$ and $\mathcal{H}_z(k) < 0$. There, we find

$$u^{(\mathrm{II})}{}_-(k) = \frac{1}{\mathcal{N}^{(\mathrm{II})}}\begin{pmatrix} -\mathcal{H}_x(k) + i\,\mathcal{H}_y(k) \\ \mathcal{H}_z(k) + \left|\vec{\mathcal{H}}(k)\right| \end{pmatrix} = \frac{1}{\mathcal{N}^{(\mathrm{II})}}\begin{pmatrix} 0 \\ \mathcal{H}_z(k) + \sqrt{\mathcal{H}_z(k)^2} \end{pmatrix} = \frac{1}{\mathcal{N}^{(\mathrm{II})}}\begin{pmatrix} 0 \\ \mathcal{H}_z(k) + |\,\mathcal{H}_z(k)\,| \end{pmatrix} = \frac{1}{\mathcal{N}^{(\mathrm{II})}}\begin{pmatrix} 0 \\ \mathcal{H}_z(k) - \mathcal{H}_z(k) \end{pmatrix} = 0 \tag{6.85}$$

For the Hamiltonian shown above, the origin is such a point. $k_x = k_y = 0$

$$\mathcal{H}_0(k) = 0 \tag{6.86}$$

$$\mathcal{H}_x(k) = \Delta \sin(k_x) = 0 \tag{6.87}$$

$$\mathcal{H}_y(k) = \Delta \sin\left(k_y\right) = 0 \tag{6.88}$$

$$\mathcal{H}_z(k) = -2\,t\cos k_x - 2\,t\cos k_y - \mu = -4\,t - \mu < 0 \tag{6.89}$$

Therefore, we need to cut the BZ into two areas and use two different wave functions to describe the Bloch waves. They are connected by a gauge transformation

$$u^{(\mathrm{II})}{}_-(k) = u^{(I)}{}_-(k)\,e^{i\,\phi(k)} \tag{6.90}$$

$$\mathcal{A}_-{}^{(\mathrm{II})}(k) = \mathcal{A}_-{}^{(I)}(k) + \nabla_k\,\phi(k) \tag{6.91}$$

This is in strong analogy to the magnetic monopole case, where we need to different gauge fields.

Some comments:

- These singularities are NOT physical. If one measures any physical observables, there is no singularity anywhere in the momentum space. However. for the wavefunction and the Berry connection, which are not physical observables, there is always some singularity for this Hamiltonian.
- The location of these singularity points depends on the gauge (phase) choice. In other words, the location of the singularities has no physical meaning either.
- Only one thing about these singularities is physical, its existence. There must be some singularity points. This statement is independent of gauge choice and it tells us that the topological index is nonzero.

Let's compute the topological index for this model. For these model, there are four special points which satisfy $\mathcal{H}_x = \mathcal{H}_y = 0$. They are $\vec{k} = (0,\,0)$, $\vec{k} = (\pi,\,\pi)$, $\vec{k} = (0,\,\pi)$ and $\vec{k} = (\pi,\,0)$. At these four points, the values of $\mathcal{H}_z$ are: $\mathcal{H}_z = -4\,t - \mu$, $4\,t - \mu$, $-\mu$ and $-\mu$ respectively.

6.2.2. case I: $\mu < -4\,t$

If $\mu < -4\,t$, all the four special points has $\mathcal{H}_z > 0$. So we can use $u^{(\mathrm{II})}{}_-(k)$ for the whole Brillouin zone, and there is no singularity points.

$$u^{(\mathrm{II})}{}_-(k) = \frac{1}{\mathcal{N}^{(I)}}\begin{pmatrix} \mathcal{H}_z(k) - \left|\vec{\mathcal{H}}(k)\right| \\ \mathcal{H}_x(k) + i\,\mathcal{H}_y(k) \end{pmatrix} \tag{6.92}$$

Then we can get Berry connection

$$\mathcal{A}_-{}^{(\mathrm{II})} = -i\left\langle u^{(\mathrm{II})}{}_-(k)\,\middle|\,\partial_k\,\middle|\,\left(u^{(\mathrm{II})}\right)_-(k)\right\rangle \tag{6.93}$$

The Berry curvature for the lower band is

$$\Omega_- = \nabla \times \mathcal{A}_- \tag{6.94}$$

Thus. the total Berry curvature (2π times the Chern number) is

$$\iint_{\mathrm{BZ}} d\!k\,\Omega_- = \iint_{\mathrm{BZ}} d\!k\,\nabla \times \mathcal{A}_- = \oint_{\partial \mathrm{BZ}} d\!k\,\mathcal{A}_- = 0 \tag{6.95}$$

So, the Chern number is zero

$$C = \frac{1}{2\,\pi}\iint_{\mathrm{BZ}} d\!k\,\Omega_- = 0. \tag{6.96}$$

6.2.3. Marginal case: $\mu = -4\,t$

$\mathcal{H}_x(k) = \mathcal{H}_y(k) = \mathcal{H}_z(k) = 0$ at $k = (\pi,\,\pi)$

Because the gap between the two bands is

$$\Delta(k) = E_+ - E_- = 2\left|\vec{\mathcal{H}}(k)\right| \tag{6.97}$$

$\Delta(k) = 0$ at $k = (\pi, \pi)$.

The gap closes at $k = (\pi, \pi)$, i.e. the two energy bands touch with each other at $k = (\pi, \pi)$. This gives us a Dirac point. Here, the system is not an insulator, but a metal (a semi-metal). So we cannot define a topological index.

Remember that we showed early on, to change the topology of an insulator, one need to close the gap. And here is the gap closing.

6.2.4. case II: $-4\,t < \mu < 0$

For $-4\,t < \mu < 0$, $\mathcal{H}_z < 0$ at $k = (0, 0)$, and $\mathcal{H}_z > 0$ for the other three points $k = (0, \pi)$, $(\pi, 0)$ and (π, π). Therefore, we need two wavefunctions. First, we draw a small circle around the origin. Inside this small circle, which we will call region D_I, we use

$$u^{(I)}{}_-(k) = \frac{1}{\mathcal{N}^{(I)}}\begin{pmatrix} \mathcal{H}_z(k) - \left|\vec{\mathcal{H}}(k)\right| \\ \mathcal{H}_x(k) + i\,\mathcal{H}_y(k) \end{pmatrix} \tag{6.98}$$

Outside the circle, which we will call region D_{II}, we choose

$$u^{(\text{II})}{}_-(k) = \frac{1}{\mathcal{N}^{(I)}}\begin{pmatrix} \mathcal{H}_z(k) - \left|\vec{\mathcal{H}}(k)\right| \\ \mathcal{H}_x(k) + i\,\mathcal{H}_y(k) \end{pmatrix} \tag{6.99}$$

The wavefunctions for these two regions are connected by a gauge transformation

$$u^{(\text{II})}{}_-(k) = u^{(I)}{}_-(k)\, e^{i\,\phi(k)} \tag{6.100}$$

And the Berry connection are related also by the same gauge transformation.

$$\mathcal{A}_-{}^{(\text{II})}(k) = \mathcal{A}_-{}^{(I)}(k) + \nabla_k\,\phi(k) \tag{6.101}$$

The total Berry curvature (2π times the Chern number) is

$$\begin{aligned} \iint_{\text{BZ}} d\,k\,\Omega_- &= \int\int_{D_I} d\,k\,\nabla\times\mathcal{A}_-{}^{(I)} + \int\int_{D_{\text{II}}} d\,k\,\nabla\times\mathcal{A}_-{}^{(\text{II})} = \oint_{\partial D_I} d\,k\,\mathcal{A}_-{}^{(I)} + \oint_{\partial D_{\text{II}}} d\,k\,\mathcal{A}_-{}^{(\text{II})} = \oint_{\partial D_I} d\,k\,\mathcal{A}_-{}^{(I)} - \oint_{\partial D_I} d\,\text{k}\mathcal{A}_-{}^{(\text{II})} = \\ &\oint_{\partial D_I} d\,k\left(\mathcal{A}_-{}^{(I)} - \mathcal{A}_-{}^{(\text{II})}\right) = -\oint_{\partial D_I} d\,k\,\nabla_k\,\phi(k) = -\int_0^{2\,\pi} d\,\theta\,\partial_\theta\,\phi = -[\phi(\theta = 2\,\pi) - \phi(\theta = 0)] = \phi(\theta = 0) - \phi(\theta = 2\,\pi) \end{aligned} \tag{6.102}$$

We know that $\phi\,(\theta = 0) - \phi(\theta = 2\,\pi)$ is quantized

$$\phi(\theta = 0) - \phi(\theta = 2\,\pi) = 2\,\pi\,n \tag{6.103}$$

Therefore, the Chern number is quantized. For the model we considered here:

$$\mathcal{H}_0(k) = 0 \tag{6.104}$$

$$\mathcal{H}_x(k) = \Delta\,\sin\,(k_x) \tag{6.105}$$

$$\mathcal{H}_y(k) = \Delta\,\sin\left(k_y\right) \tag{6.106}$$

$$\mathcal{H}_z(k) = -2\,t\cos k_x - 2\,t\cos k_y - \mu \tag{6.107}$$

If we choose D_I to be a very small circle with radius $k \sim 0$, then around the circle, we can expand everything as a power series of small k.

$$\mathcal{H}_x(k) = \Delta\,\sin\,(k_x) \approx \Delta\,k_x + O\left(k^2\right) \tag{6.108}$$

$$\mathcal{H}_y(k) = \Delta\,\sin\left(k_y\right) \approx \Delta\,k_y + O\left(k^2\right) \tag{6.109}$$

$$\mathcal{H}_z(k) = -2\,t\cos k_x - 2\,t\cos k_y - \mu = -2\,t - \mu + O\left(k^2\right) \tag{6.110}$$

$$e^{i\,\phi(k)} = \frac{\frac{\mathcal{H}_z + \left|\vec{\mathcal{H}}(k)\right|}{\mathcal{H}_x(k) + i\,\mathcal{H}_y(k)}}{\left|\frac{\mathcal{H}_z + \left|\vec{\mathcal{H}}(k)\right|}{\mathcal{H}_x(k) + i\,\mathcal{H}_y(k)}\right|} = \frac{\frac{1}{\mathcal{H}_x(k) + i\,\mathcal{H}_y(k)}}{\left|\frac{1}{\mathcal{H}_x(k) + i\,\mathcal{H}_y(k)}\right|} = \frac{\left|\mathcal{H}_x(k) + i\,\mathcal{H}_y(k)\right|}{\mathcal{H}_x(k) + i\,\mathcal{H}_y(k)} = \frac{\left|\Delta\,k_x + i\,\Delta\,k_y\right|}{\Delta\,k_x + i\,\Delta\,k_y} = \frac{\left|k_x + i\,k_y\right|}{k_x + i\;k_y} = \frac{\left|k\,e^{i\,\theta}\right|}{k\,e^{i\,\theta}} = \frac{1}{e^{i\,\theta}} = e^{-i\,\theta} \tag{6.111}$$

Here, we transfer $k_x + i\, k_y$ into polar coordinate, which is $k\, e^{i\,\theta}$. This result tells us that

$$\phi(k) = -\theta \tag{6.112}$$

Therefore,

$$\oiint_{\mathrm{BZ}} d\!k\, \Omega_- = \phi(\theta = 0) - \phi(\theta = 2\,\pi) = 0 - (-2\,\pi) = 2\,\pi \tag{6.113}$$

So,

$$C = \frac{1}{2\,\pi} \oiint_{\mathrm{BZ}} d\!k\, \Omega_- = 1 \tag{6.114}$$

This is a topological insulator with $C = 1$.

6.2.5. Marginal case: $\mu = 0$

The gap closes at $k = (0,\ \pi)$ and $(\pi,\ 0)$. The system is not an insulator, so we cannot define the Chern number.

6.2.6. case III: $0 < \mu < 4\,t$

For $0 < \mu < 4\,t$, $\mathcal{H}_z > 0$ at k=(π,π), and $\mathcal{H}_z < 0$ for the other three points. Here,we can draw a small circle centered at (π,π) to cut the system into two parts. Inside the circle, near (π,π), we use u^{II} and outside the circle, near (0,0) we use u^{I}. Same as case II, we get a topological insulator with $C = 1$.

6.2.7. Marginal case: $\mu = 4\,t$

The gap closes at $k = (\pi,\ \pi)$. Not an insulator.

6.2.8. case IV: $\mu > 4\,t$

All the four special points has $\mathcal{H}_z < 0$. So we just use u^{I} for the whole BZ. No singularity, and thus $C = 0$.

6.2.9. the top band?

The top band has the opposite Chern number $C_+ = -C_-$. This is because for any tight-binding models, the total Chern number for all the bands is always 0. So here, we have $C_+ + C_- = 0$.

6.2.10. Summary

- This model has four phases. Two topological phases with C=+1 and two trivial insulator phase with C=0.
- A topological phase and a trivial insulator phase are always separated by a phase transition, which is known as a topological phase transition.
  - Across the topological phase transition, the topological index changes its value
  - **Across a topological transition, the insulating gap closes and then reopens (generically true)**
- Gap closing is the necessary condition for a topological transition, but it is not sufficient. One may close and reopen the gap without changing the topological index. e.g. the μ=0 point here.

6.3. The model of Haldane on a honeycomb lattice

6.3.1. the Honeycomb lattice (graphene)

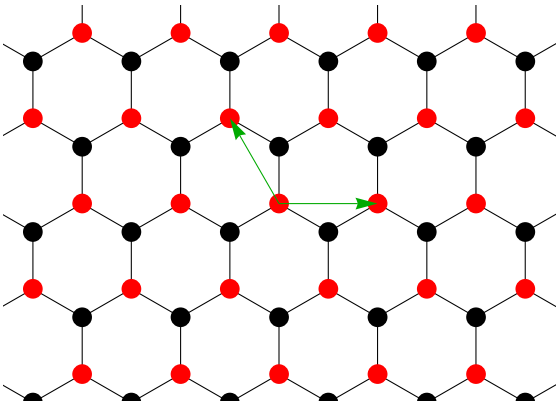

Fig. 1. a honeycomb lattice.

Two sites per unit cell (two sublattices) a and b (red and black respectively). The green arrow indicates the lattice vectors $\vec{v}_1$ and $\vec{v}_2$. If we shift the lattices by $m\ \vec{v}_1 + n\ \vec{v}_2$ with m and n being integers, the lattice is invariant.

$$V\left(\vec{r}\right) = V\left(\vec{r} + m\ \vec{v}_1 + n\ \vec{v}_2\right) \tag{6.115}$$

Define a to the length of the nearest-neighbor (NN) bonds. Then we can show that $\vec{v}_1 = \left(\sqrt{3}\ a,\ 0\right)$ and $\vec{v}_2 = \left(-\sqrt{3}\ \big/ 2\, a,\ \ 3/2\, a\right)$. If we only consider hoppings between the nearest-neighbor (NN) sites, the Hamiltonian is

$$H = -t \sum_{<i,j>} a_i^\dagger\, b_j - t \sum_{<i,j>} b_i^\dagger\, a_j \tag{6.116}$$

where ⟨i,j⟩ means nearest neighbors.

6.3.2. Band structures

There are three type of NN bonds: (1) along the y axis with $\theta=\pi/2$, (2) along $\theta=\pi/2+2\pi/3=7\pi/6$, (3) along $\theta=\pi/2+4\pi/3=11\pi/6$, and we need to write them out separately.

$$H = -t \sum_i a_{\vec{r}_i}^\dagger\, b_{r_i+\vec{e}_1} - t \sum_i a_{\vec{r}_i}^\dagger\, b_{r_i+\vec{e}_2} - t \sum_i a_{\vec{r}_i}^\dagger\, b_{r_i+\vec{e}_3} + h.c. \tag{6.117}$$

with

$$\vec{e}_1 = (0,\ a) \tag{6.118}$$

$$\vec{e}_2 = \left(-\frac{\sqrt{3}}{2}\, a,\ -\frac{a}{2}\right) \tag{6.119}$$

$$\vec{e}_3 = \left(\frac{\sqrt{3}}{2}\, a\,,\ -\frac{a}{2}\right) \tag{6.120}$$

Here the sum $\sum_i$ sums over all unit cells (or say all red sites). If we go to the momentum space, by performing the 2D Fourier transformation,

$$a_k = \frac{1}{\sqrt{N}} \sum_i a_i\, e^{i\,\vec{k}\cdot\vec{r}} \tag{6.121}$$

$$a_i = \frac{1}{\sqrt{N}} \sum_k a_k\, e^{-i\,\vec{k}\cdot\vec{r}} \tag{6.122}$$

$$b_k = \frac{1}{\sqrt{N}} \sum_i a_i\, e^{i\,\vec{k}\cdot\vec{r}} \tag{6.123}$$

$$b_i = \frac{1}{\sqrt{N}} \sum_k a_k\, e^{-i\,\vec{k}\cdot\vec{r}} \tag{6.124}$$

the first term in the Hamiltonian becomes

$$-t\sum_{ii} {a_{r_i}}^{\dagger} b_{r_i+e_1} = -\frac{t}{N}\sum_{ii}\sum_{kk}\sum_{kk'} {a_k}^{\dagger} e^{i\vec{k}\cdot\vec{r_i}} b_{k'} e^{-i\vec{k'}\cdot\left(\vec{r_i}+\vec{e_1}\right)} =$$
$$-t\sum_{kk}\sum_{kk'} {a_k}^{\dagger} b_{k'} e^{-i\vec{k'}\cdot\vec{e_1}} \frac{1}{N}\sum_{ii} e^{i\left(\vec{k}-\vec{k'}\right)\cdot\vec{r_i}} = -t\sum_{kk}\sum_{kk'} {a_k}^{\dagger} b_{k'} e^{-i\vec{k'}\cdot\vec{e_1}} \delta_{k,k'} = -t\sum_{kk} {a_k}^{\dagger} b_k e^{-i\vec{k'}\cdot\vec{e_1}} \tag{6.125}$$

Notice that, as we proved early on, we get a phase factor $\exp\left[i\vec{k}\cdot\left(\vec{r}_a-\vec{r}_b\right)\right]$, which only depends on the separation between a and b sites. If we repeat the same procedure for the other two bonds, we get

$$\begin{aligned} H &= -t\sum_{ii} {a_{r_i}}^{\dagger} b_{r_i+e_1} - t\sum_{ii} {a_{r_i}}^{\dagger} b_{r_i+e_2} - t\sum_{ii} {a_{r_i}}^{\dagger} b_{r_i+e_2} + h.c. \\ &= -t\sum_{kk} {a_k}^{\dagger} b_k\left(e^{-i\vec{k}\cdot\vec{e_1}} + e^{-i\vec{k}\cdot\vec{e_2}} + e^{-i\vec{k}\cdot\vec{e_3}}\right) - t\sum_{kk} {b_k}^{\dagger} a_k\left(e^{i\vec{k}\cdot\vec{e_1}} + e^{i\vec{k}\cdot\vec{e_2}} + e^{i\vec{k}\cdot\vec{e_3}}\right) \\ &= \sum_{kk}\begin{pmatrix} {a_k}^{\dagger} & {b_k}^{\dagger}\end{pmatrix}\begin{pmatrix} 0 & \mathcal{H}_{12}\left(\vec{k}\right) \\ \mathcal{H}_{21}\left(\vec{k}\right) & 0\end{pmatrix}\begin{pmatrix} a_k \\ b_k\end{pmatrix} \end{aligned} \tag{6.126}$$

$$\mathcal{H}_{12}\left(\vec{k}\right) = -t\left[\exp\left(-i\vec{k}\cdot\vec{e_1}\right) + \exp\left(-i\vec{k}\cdot\vec{e_2}\right) + \exp\left(-i\vec{k}\cdot\vec{e_3}\right)\right] \tag{6.127}$$

$$\mathcal{H}_{21}\left(\vec{k}\right) = \mathcal{H}_{12}\left(\vec{k}\right)^* = -t\left[\exp\left(i\vec{k}\cdot\vec{e_1}\right) + \exp\left(i\vec{k}\cdot\vec{e_2}\right) + \exp\left(i\vec{k}\cdot\vec{e_3}\right)\right] \tag{6.128}$$

The kernel of the Hamiltonian:

$$\mathcal{H}\left(\vec{k}\right) = \begin{pmatrix} 0 & \mathcal{H}_{12}\left(\vec{k}\right) \\ \mathcal{H}_{21}\left(\vec{k}\right) & 0\end{pmatrix} \tag{6.129}$$

The eigenvalues of $\mathcal{H}$(k) gives the dispersion relation.

$$\epsilon_{\pm}\left(\vec{k}\right) = \pm\left|\mathcal{H}_{12}\left(\vec{k}\right)\right| = \pm\left|t\right|\sqrt{3 + 2\cos\left(\sqrt{3}\ k_x a\right) + 4\cos\left(\frac{\sqrt{3}}{2}k_x a\right)\mathrm{Cos}\left(\frac{3}{2}k_y a\right)} \tag{6.130}$$

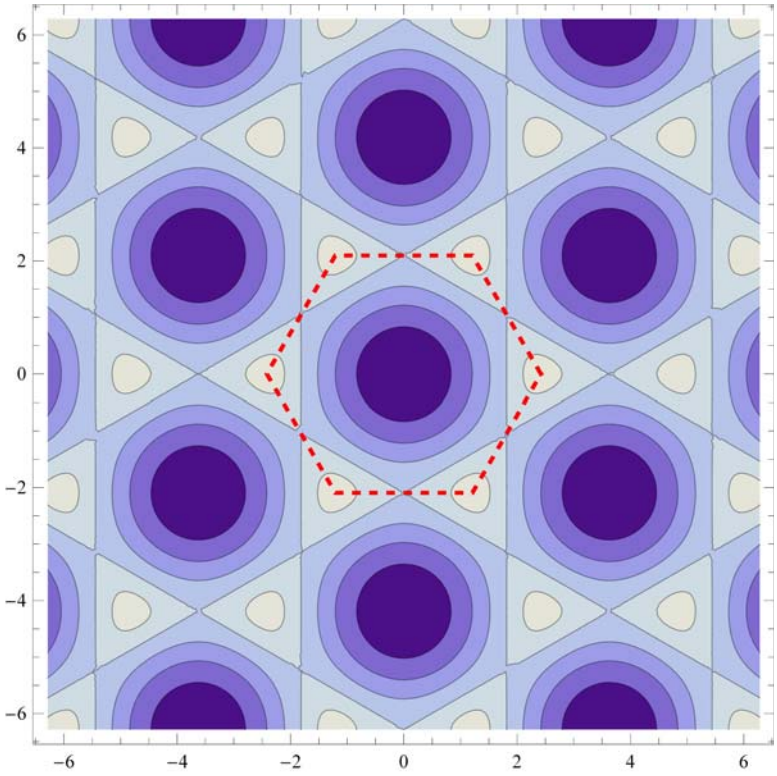

Fig. 2. The contour plot of the energy dispersion for the lower band (ϵ_- as a function of k_x and k_y). The red dash lines mark the first BZ, which is a hexagon.

The dispersion $\epsilon_\pm$ is a periodic function of $\vec{k}$ space (the hexagon repeats itself in the figure shown above). For the lower band, its energy minimum is located at $k = 0$. The maximum of ϵ_- is reached at the corners of the BZ. The first BZ is a hexagon. It has six corners. However, due to the periodic structure in k-space, the three corner points with θ=0, θ=2π/3 and θ=4π/3 are the same point (their momenta differ by a reciprocal vector). Similarly, the three corners with θ=π, θ=π+2π/3 and θ=π+4π/3 are the same point. Therefore, there are only two different corner points and they are known as the K and K' points, where

$$K = \left(\frac{4\pi}{3\sqrt{3}\,a}, 0\right) \tag{6.131}$$

and

$$K' = \left(-\frac{4\pi}{3\sqrt{3}\,a}, 0\right) \tag{6.132}$$

For most of the momentum points, one of the two bands has positive energy $\epsilon_+ > 0$ and the other one has negative energy $\epsilon_- < 0$. However at the corner of the BZ, K and K', the two bands are degenerate $\epsilon_+ = \epsilon_- = 0$.

$$\epsilon_\pm\left(\vec{K}\right) = \pm\left|\mathcal{H}_{12}\left(\vec{K}\right)\right| = \pm|t|\sqrt{3 + 2\cos\left(\sqrt{3}\,k_x a\right) + 4\cos\left(\frac{\sqrt{3}}{2}k_x a\right)\mathrm{Cos}\left(\frac{3}{2}k_y a\right)} =$$

$$\pm|t|\sqrt{\left(3 + 2\cos\left(\sqrt{3}\,\frac{4\pi}{3\sqrt{3}\,a}a\right) + 4\cos\left(\frac{\sqrt{3}}{2}\,\frac{4\pi}{3\sqrt{3}\,a}a\right)\mathrm{Cos}\left(\frac{3}{2}\times 0\right)\right)} = \tag{6.133}$$

$$\pm|t|\sqrt{3 + 2\cos\left(\frac{4\pi}{3}\right) + 4\cos\left(\frac{2\pi}{3}\right)} = \pm|t|\sqrt{3 + 2\times\left(-\frac{1}{2}\right) + 4\times\left(-\frac{1}{2}\right)} = \sqrt{3-3} = 0$$

Two bands have the same energy at K and K', and they are in fact band crossing points.

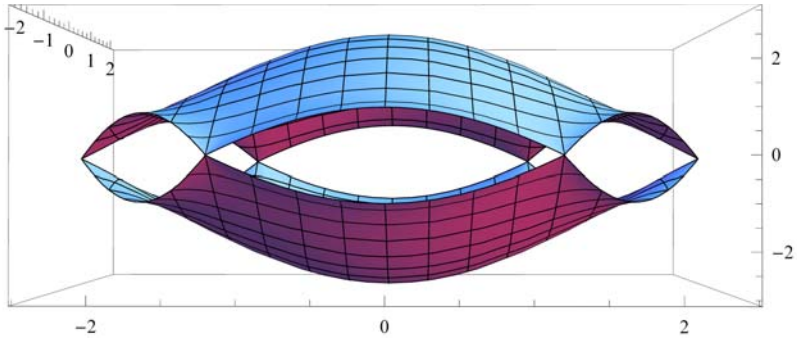

Fig. 3. The energy dispersion for both bands ($\epsilon_\pm$ as a function of k_x and k_y) in the first BZ (within the red dash lines mark in the figure above).

Near K or K' points, the energies of the two bands are linear functions of momentum $\epsilon_\pm \propto \pm(k-K)$. To see this, we expand $\epsilon_\pm\left(\vec{k}\right)$ near $\vec{k}\sim\vec{K}$ and $\vec{k}\sim\vec{K}'$

$$\epsilon_\pm\left(\vec{k}\right)=\epsilon_\pm\left(\vec{K}+\vec{q}\right)=\pm\left|t\right|\sqrt{\left(3+2\cos\left[\sqrt{3}\left(\frac{4\pi}{3\sqrt{3}\,a}+q_x\right)a\right]+4\cos\left[\frac{\sqrt{3}}{2}\left(\frac{4\pi}{3\sqrt{3}\,a}+q_x\right)a\right]\mathrm{Cos}\left(\frac{3}{2}q_y\,a\right)\right)}= \\ \pm\frac{3}{2}\left|t\right|a\sqrt{{q_x}^2+{q_y}^2}+O\left(q^2\right)=\pm\frac{3}{2}\left|t\right|a\,q+O\left(q^2\right) \tag{6.134}$$

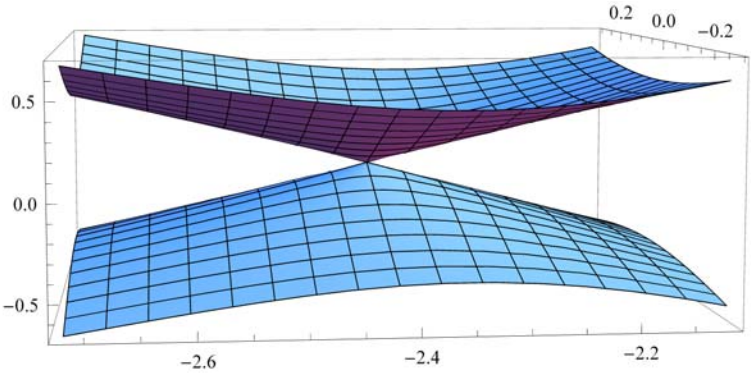

Fig. 4. A zoom in to the Dirac point.

Q: Fermions with linear dispersion $\epsilon \propto k$. What are they?

A: A Massless Dirac fermion.

Near the K point, we have

$$\mathcal{H}(k)=\mathcal{H}(K+q)=\frac{3}{2}t\,a\begin{pmatrix}0 & q_x-i\,q_y\\ q_x+i\,q_y & 0\end{pmatrix}+O\left(q^2\right)\approx\frac{3}{2}t\,a\left(q_x\,\sigma_x+q_y\,\sigma_y\right)=c\,\vec{q}\cdot\vec{\sigma} \tag{6.135}$$

Near the K' point, we have

$$\mathcal{H}(k)=\mathcal{H}(K'+q)=-\frac{3}{2}t\,a\begin{pmatrix}0 & q_x+i\,q_y\\ q_x-i\,q_y & 0\end{pmatrix}+O\left(q^2\right)\approx-\frac{3}{2}t\,a\left(q_x\,\sigma_x-q_y\,\sigma_y\right)=-c\,\vec{q}\cdot\left(\sigma_x\,\vec{\sigma}\,\sigma_x\right) \tag{6.136}$$

Here c is the "speed of light". Each of them is a Weyl fermion. These two Weyl fermions with opposite chirality forms a Dirac fermion.

Dirac fermions we learned in quantum mechanics:

$$i\,\partial_t\psi=\begin{pmatrix}0 & c\,\vec{q}\cdot\vec{\sigma}\\ c\,\vec{q}\cdot\vec{\sigma} & 0\end{pmatrix}\psi \tag{6.137}$$

If one make an unitary transformation (changing the basis, which doesn't change any physics)

$$\psi' = \begin{pmatrix} 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{2}} & 0 & 0 & \frac{1}{\sqrt{2}} \\ 0 & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ -\frac{1}{\sqrt{2}} & 0 & 0 & \frac{1}{\sqrt{2}} \end{pmatrix} \psi$$

The equation turns into

$$i\,\partial_t \psi' = \begin{pmatrix} -c\,\vec{q}\cdot\left(\sigma_x \vec{\sigma}\,\sigma_x\right) & 0 \\ 0 & c\,\vec{q}\cdot\vec{\sigma} \end{pmatrix} \psi' \tag{6.139}$$

The Dirac point (band crossing points) in a honeycomb lattice is stable as long as the space inversion symmetry $r \to -r$ and the time reversal symmetry $t \to -t$ are preserved. We will come back to this point latter. No matter how one perturb the systems (e.g. adding longer-range hoppings), the Dirac point is always there as long as the two symmetries mentioned above are preserved. For graphene, the lower band is filled and the upper band is empty, which is known as "half-filling". The "half" here means that the number of electrons N_e over the number of sites N_s is $N_e/N_S = 1/2$. However, it is worthwhile to notice that the ratio between N_e and the number of unit cells N is 1, because there are two sites in each unit cell. So one of the two bands are totally filled. By gating, one can tune the Fermi energy slightly away from the Dirac point, which is known as a dopped graphene.

6.3.3. Aharonov–Bohm effect and complex hopping

Q: Can hopping strength t be complex?

A: Yes, and the phase can come from the Aharonov–Bohm effect in the presence of a magnetic field or spin-orbital couplings.

The Aharonov–Bohm effect: if one moves a charged particle around a closed contour. The phase difference between the final and initial states is proportional to the magnetic flux enclosed by the contour $\phi = e/\hbar \int\int B\cdot d\!l\, S = e/\hbar \oint A.d\!l\, l$

Now, let's consider a discrete lattice. Consider three sites a, b and c. The hopping strength between these three sites are t_{ab} t_{bc} and t_{ca} respectively. If a particle hops from a to b and then to c, the hopping strength around this loop is:

$$t_{ab}\, t_{bc}\, t_{ca} = \left| t_{ab} \right| e^{i\phi_{ab}} * \left| t_{bc} \right| e^{i\phi_{bc}} * \left| t_{ca} \right| e^{i\phi_{ca}} = \left| t_{ab}\, t_{bc}\, t_{ca} \right| e^{i(\phi_{ab}+\phi_{bc}+\phi_{ca})} \tag{6.140}$$

The phase picked up by the electron is:

$$\phi_{ab} + \phi_{bc} + \phi_{ca} = \frac{e}{\hbar} \int\int B\cdot d\!l\, S \tag{6.141}$$

If B is nonzero inside the triangle formed by these three sites, the phase for these hoppings are nonzero.

Please notice that (1) t_{ab} and t_{ba} has opposite phase, due to the Hermitian condition (One can also prove this using the Fermi's golden rule, which says that the tunneling amplitude from the state |f> to |i> is the complex conjugate of the tuning amplitude from |i> to |f>).

$$t_{ab} = t_{ba}{}^{*} \tag{6.142}$$

(2) The individual phases for t_{ab}, t_{bc} and t_{ca} have no physical meaning and their phases depend on the gauge choice. However, the total phase around a loop is independent of gauge, and in fact it is a physical observable, i.e. the magnetic flux.

Proof:

We know that the phase of the hopping term is

$$\phi_{ab} = e/\hbar \int_a^b A\cdot d\!l\, l \tag{6.143}$$

Underage transformation: $A \to A + \nabla\chi$

$$\phi_{ab} \to \phi_{ab}' = e/\hbar \int_a^b (A + \nabla\chi)\cdot d\!l\, l = e/\hbar \int_a^b A\cdot d\!l\, l + e/\hbar \int_a^b \nabla\chi\cdot d\!l\, l = \phi_{ab} + (\chi_b - \chi_a)\,\frac{e}{\hbar} \tag{6.144}$$

Obviously, ϕ_{ab} is not a physical observables, since it relies on the gauge choice. However, the total phase around a loop is different. It is a loop integral of A, which is gauge independent and the physics meaning of this Integral is the magnetic flux.

$$\phi_{ab} + \phi_{bc} + \phi_{ca} =$$
$$e/\hbar \oint A\cdot d\,l \to \phi_{ab}{}' + \phi_{bc}{}' + \phi_{ca}{}' = e/\hbar \oint (A + \nabla\chi)\cdot d\,l = e/\hbar \oint A\cdot d\,l + e/\hbar \oint \nabla\chi\cdot d\,l = e/\hbar \oint A\cdot d\,l = \phi_{ab} + \phi_{bc} + \phi_{ca} \tag{6.145}$$

The complex hopping strength induced by B fields breaks the time-reversal symmetry because $B \to -B$ under time reversal. (in other words, under time-reversal one needs to flip the sign of all these phases).

6.3.4. Complex next-nearest-neighbor (NNN) hoppings (breaking T-symmetry using *B* fields)

Ref: F. D. M. Haldane, Model for a Quantum Hall Effect without Landau Levels: Condensed-Matter Realization of the "Parity Anomaly", Phys. Rev. Lett. 61, 2015–2018 (1988).

Now let us add some NNN hoppings and assume their hopping amplitudes are complex. For simplicity, we choose the amplitude and the complex phase to be the same for all NNN bonds. If the hopping is along the arrows marked in the figure, the phase of the hopping strength is ϕ. If one hops in the opposite direction, the phase is $-\phi$.

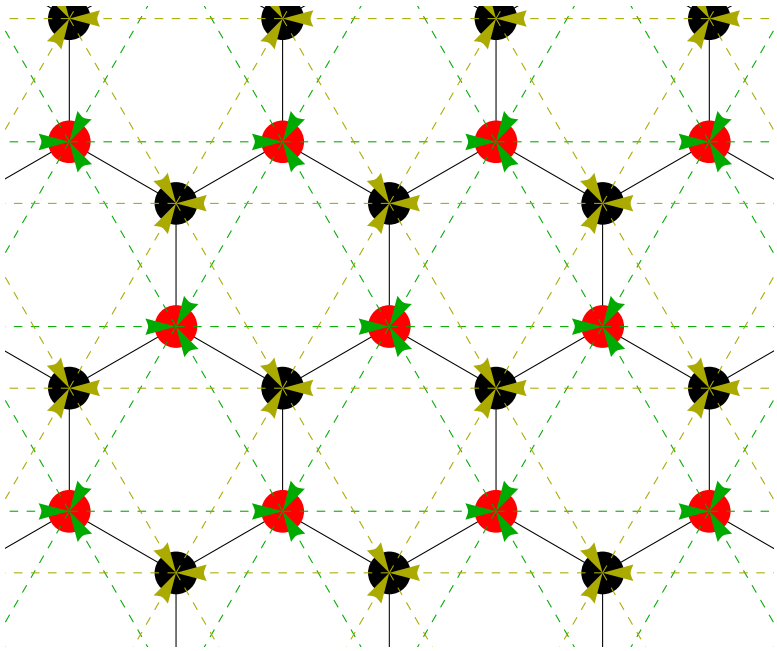

Fig. 5. The model of Haldane.

This complex phases can be realized (in theory) by applying a staggered B field, which is positive near the center of each hexagon and negative near the edges. The NNN hoppings are from an a-site to another a-site (and from a b-site to another b-site). For a-to-a hoppings, there are three different types of NNN bonds, along θ=0, $2\pi/3$ and $4\pi/3$. Same is true for b-to-b hoppings. So the Hamiltonian is:

$$H_{NNN} = -t'\,e^{i\phi}\sum_i a_{r_i}{}^\dagger a_{r_i+v_1} - t'\,e^{i\phi}\sum_i a_{r_i+v_1}{}^\dagger a_{r_i-v_3} - t'\,e^{i\phi}\sum_i a_{r_i-v_3}{}^\dagger a_{r_i} + h.c. + (a \to b \text{ and } \phi \to -\phi) \tag{6.146}$$

Here, v_1 and v_2 are marked on the first figure of this section with $\vec{v}_1 = \left(\sqrt{3}\ a,\ 0\right)$ and $\vec{v}_2 = \left(-\sqrt{3}\ /2\,a,\ \ 3/2\,a\right)$. $\vec{v}_2 = \left(-\sqrt{3}\ /2\,a,\ \ -3/2\,a\right)$

$$\begin{aligned} H_{NNN} &= -t'\,e^{i\phi}\sum_i a_{r_i}{}^\dagger a_{r_i+v_1} - t'\,e^{i\phi}\sum_i a_{r_i+v_1}{}^\dagger a_{r_i-v_3} - t'\,e^{i\phi}\sum_i a_{r_i-v_3}{}^\dagger a_{r_i} + h.c. + (a \to b \text{ and } \phi \to -\phi) \\ &= -t'\,e^{i\phi}\sum_k a_k{}^\dagger a_k \left(e^{-i\,k\cdot v_1} + e^{-i\,k\cdot v_2} + e^{-i\,k\cdot v_3}\right) + h.c. + (a \to b \text{ and } \phi \to -\phi) \\ &= -2\,t'\sum_k a_k{}^\dagger a_k\,[\cos(k\cdot v_1 - \phi) + \cos(k\cdot v_2 - \phi) + \cos(k\cdot v_3 - \phi)]\ - \\ &\quad 2\,t'\sum_k b_k{}^\dagger b_k\,[\cos(k\cdot v_1 + \phi) + \cos(k\cdot v_2 + \phi) + \cos(k\cdot v_3 + \phi)] \end{aligned} \tag{6.147}$$

$$H = H_{NN} + H_{NNN} = \sum_k \begin{pmatrix} a_k{}^\dagger & b_k{}^\dagger \end{pmatrix} \begin{pmatrix} \mathcal{H}_{11}(k) & \mathcal{H}_{12}(k) \\ \mathcal{H}_{21}(k) & \mathcal{H}_{22}(k) \end{pmatrix} \begin{pmatrix} a_k \\ b_k \end{pmatrix} \tag{6.148}$$

$$\mathcal{H}_{12}(k) = -t\left(e^{-i\vec{k}\cdot\vec{e}_1} + e^{-i\vec{k}\cdot\vec{e}_2} + e^{-i\vec{k}\cdot\vec{e}_3}\right) \tag{6.149}$$

$$\mathcal{H}_{21}(k) = \mathcal{H}_{12}(k)^* = -t\left(e^{i\vec{k}\cdot\vec{e}_1} + e^{i\vec{k}\cdot\vec{e}_2} + e^{i\vec{k}\cdot\vec{e}_3}\right) \tag{6.150}$$

$$\mathcal{H}_{11}(k) = -2\,t'\,[\cos(k\cdot v_1 - \phi) + \cos(k\cdot v_2 - \phi) + \cos(k\cdot v_3 - \phi)] \tag{6.151}$$

$$\mathcal{H}_{22}(k) = -2\,t'\,[\cos(k\cdot v_1 + \phi) + \cos(k\cdot v_2 + \phi) + \cos(k\cdot v_3 + \phi)] \tag{6.152}$$

If one computes the eigenvalues of $\mathcal{H}$(k), one find that the two bands never cross with each other for any k (if t' is non-zero and ϕ is NOT an integer multiplied by π).

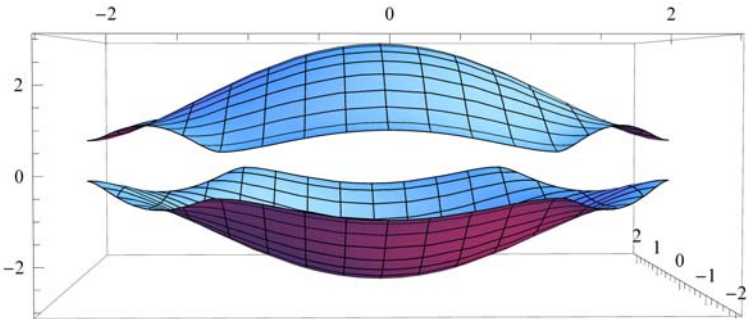

Fig. 6. The energy dispersion for both bands ($\epsilon_\pm$ as a function of k_x and k_y) in the first BZ (within the red dash lines mark in the figure above).

Using Pauli matrices:

$$\mathcal{H} = \mathcal{H}_0(k)\,I + \mathcal{H}_x(k)\,\sigma_x + \mathcal{H}_y(k)\,\sigma_y + \mathcal{H}_z(k)\,\sigma_z \tag{6.153}$$

$$\mathcal{H}_0 = \frac{\mathcal{H}_{11}(k) + \mathcal{H}_{22}(k)}{2} = -2\,t'\cos\phi\,[\cos(k\cdot v_1) + \cos(k\cdot v_2) + \cos(k\cdot v_3)] \tag{6.154}$$

$$\mathcal{H}_z = \frac{\mathcal{H}_{11}(k) - \mathcal{H}_{22}(k)}{2} = -2\,t'\sin\phi\,[\sin(k\cdot v_1) + \sin(k\cdot v_2) + \sin(k\cdot v_3)] \tag{6.155}$$

$$\mathcal{H}_x = \mathrm{Re}[\mathcal{H}_{21}(k)] = -t\left[\cos\left(\vec{k}\cdot\vec{e}_1\right) + \cos\left(\vec{k}\cdot\vec{e}_2\right) + \cos\left(\vec{k}\cdot\vec{e}_3\right)\right] \tag{6.156}$$

$$\mathcal{H}_y = \mathrm{Im}[\mathcal{H}_{21}(k)] = -t\left[\sin\left(\vec{k}\cdot\vec{e}_1\right) + \sin\left(\vec{k}\cdot\vec{e}_2\right) + \sin\left(\vec{k}\cdot\vec{e}_3\right)\right] \tag{6.157}$$

The energy dispersions:

$$\epsilon_\pm\left(\vec{k}\right) = \mathcal{H}_0\left(\vec{k}\right) \pm \sqrt{\mathcal{H}_x\left(\vec{k}\right)^2 + \mathcal{H}_y\left(\vec{k}\right)^2 + \mathcal{H}_z\left(\vec{k}\right)^2} \tag{6.158}$$

The gap between the two bands:

$$\epsilon_+\left(\vec{k}\right) - \epsilon_-\left(\vec{k}\right) = 2\sqrt{\mathcal{H}_x\left(\vec{k}\right)^2 + \mathcal{H}_y\left(\vec{k}\right)^2 + \mathcal{H}_z\left(\vec{k}\right)^2} \tag{6.159}$$

At K or K', $\mathcal{H}_x\left(\vec{k}\right) = \mathcal{H}_y\left(\vec{k}\right) = 0$ and the gap is:

$$\epsilon_+\left(\vec{k}\right) - \epsilon_-\left(\vec{k}\right) = 2\sqrt{\mathcal{H}_x\left(\vec{k}\right)^2 + \mathcal{H}_y\left(\vec{k}\right)^2 + \mathcal{H}_z\left(\vec{k}\right)^2} = 2\left|\mathcal{H}_z\left(\vec{k} = K\right)\right| = 6\sqrt{3}\,t'\sin\phi \tag{6.160}$$

In fact, at the K point,

$$\mathcal{H}_z = -3\sqrt{3}\,t'\sin\phi \tag{6.161}$$

at the K' point

$$\mathcal{H}_z = 3\sqrt{3}\,t'\sin\phi \tag{6.162}$$

They have opposite signs (as long as ϕ is not n π). Based on what we learned last time, this means that one cannot define the wavefunction in

the whole BZ. We need to cut the systems into two regions. The region I contains the K point, the region II contains the K' points. And we need to use different eigenvectors for these two regions. Using the same method we learned in the last lecture, one finds that the Chern number here is ±1.

Here, we can draw a small circle around the K point. Inside the circle, we use wavefunction $u^{(I)}$. Outside it, we use $u^{(II)}$.

$$e^{i\phi(k)} = \frac{\frac{\mathcal{H}_z+|\vec{\mathcal{H}}(k)|}{\mathcal{H}_x(k)+i\mathcal{H}_y(k)}}{\left|\frac{\mathcal{H}_z+|\vec{\mathcal{H}}(k)|}{\mathcal{H}_x(k)+i\mathcal{H}_y(k)}\right|} = \frac{|\mathcal{H}_x(k)+i\mathcal{H}_y(k)|}{\mathcal{H}_x(k)+i\mathcal{H}_y(k)} = \frac{|q_x+iq_y|}{q_x+iq_y} = \frac{1}{e^{i\theta}} = e^{-i\theta} \tag{6.163}$$

6.3.5. Potential energy (breaking inversion symmetry)

Let's keep NNN hoppings to be zero for now and add some potential energy to the Hamiltonian.

$$H_{\text{Potential}} = (V+M)\sum_i a_i^\dagger a_i + (V-M)\sum_i b_i^\dagger b_i = VN + M\sum_i a_i^\dagger a_i - M\sum_i b_i^\dagger b_i \tag{6.164}$$

The V part (average potential between a and b sites) just adds a constant term to the energy, since the total particle number N is conserved. So we can drop the V term and only consider the difference between the potential energies at a and b sites (M).

In k-space

$$H_{\text{Potential}} = M\sum_i a_i^\dagger a_i - M\sum_i b_i^\dagger b_i = M\sum_k a_k^\dagger a_k - M\sum_k b_k^\dagger b_k = \sum_k \begin{pmatrix} a_k^\dagger & b_k^\dagger \end{pmatrix} M\sigma_z \begin{pmatrix} a_k \\ b_k \end{pmatrix} \tag{6.165}$$

It adds a σ_z component to the Hamiltonian. Same as the NNN complex hopping, this term also opens a gap at the Dirac points. At K and K', $\mathcal{H}$ now has the same sign ($\mathcal{H} = M$ for any k). So, we can use one wave-function for the whole BZ and thus the system is a trivial insulator.

Q: What will happen if we have both $H_{\text{Potential}}$ and H_{NNN}?

A: We just need to look at the signs of $\mathcal{H}_z$ at K and K' points. If they have the same sign, we can use one wavefunction to cover the whole BZ, so $C = 0$ (trivial insulator). If they have opposite signs, the system is a topological insulator with C=±1

At the K point,

$$\mathcal{H}_z = M - 3\sqrt{3}\,t'\sin\phi \tag{6.166}$$

at the K' point

$$\mathcal{H}_z = M + 3\sqrt{3}\,t'\sin\phi \tag{6.167}$$

Therefore, as long as $|M| < |3\sqrt{3}\,t'\sin\phi|$, the system is an topological insulator ($\mathcal{H}_z$ flips sign). If $|M| > |3\sqrt{3}\,t'\sin\phi|$, $\mathcal{H}_z$ is always positive (or negative) and thus the system is topologically trivial. The marginal case $|M| = |3\sqrt{3}\,t'\sin\phi|$ is a topological transition, where $\mathcal{H}_z = 0$ at either the K point or the K' point. Because $\mathcal{H}_x = \mathcal{H}_y = 0$ at these two points, the gap must be zero at one of the two points.

Remarks:

- The model of Haldane is the first example of a topological insulator beyond quantum Hall effect.
- It demonstrates that topological insulator is a generic concept, which may appear in any insulating systems (NOT just quantum Hall).
- It also demonstrates that as long as the topological index is nonzero, one will observe all the topological phenomena expected for a quantum Hall state, including the quantized Hall conductivity and the existence of the edge states.
- The key differences between the model of Haldane and the quantum Hall effects are (1) the B field is on average zero in the model of Haldane while the QHE has a uniform B field and (2) there is a very strong lattice background in the model of Haldane while the QHE requires weak lattice potential.
- Systems similar to the Haldane's model are known as topological Chern insulators or Chern insulators (average B is 0 and have a strong lattice potential). But sometimes, Chern insulators are also used to refer to the quantum Hall effect.
- The model of Haldane is also the foundation to explore more complicated and exotic topological states. For example, the time-reversal invariant topological insulators was first proposed using a modified Haldane's models, which we will study latter in the semester.

6.4. Some symmetry properties of the Berry curvature $\mathcal{F}$ and the Chern number.

6.4.1. time-reversal transformation $\mathcal{T}$ and time-reversal symmetry

In this section, we consider systems with time-reversal symmetry and shows that in the presence of the time-reversal symmetry, the Chern number must be zero. Therefore, to get a topological insulator with nonzero C, one must break the time-reversal symmetry.

Dispersion relation

Dispersion relation under time-reversal transformation:

$$\mathcal{T}\,\epsilon_n(k) = \epsilon_n(-k) \tag{6.168}$$

This is because $\mathcal{T}$ turns $\vec{k}$ into $-\vec{k}$. There are two ways to see that $\mathcal{T}$ turns $\vec{k}$ into $-\vec{k}$.

- $\vec{k}$ is the momentum. Under time-reversal, velocity changes sign, so does the momentum.
- $\vec{k} = -i\,\partial_r$. Under $\mathcal{T}$, ∂_r remains invariant. However, because $\mathcal{T}$ is an anti-unitary transformation (which changes a complex number to its complex conjugate), it flips the sign of i. Therefore, $\mathcal{T}\,\vec{k} = \mathcal{T}(-i\,\partial_r) = i\,\partial_r = -\vec{k}$.

If the system is time-reversally invariant, the time-reversal symmetry implies that $\epsilon_n(k)$ is invariant under time-reversal transformation

$$\mathcal{T}\,\epsilon_n(k) = \epsilon_n(k) \tag{6.169}$$

Compare the two equations, we have:

$$\epsilon_n(k) = \epsilon_n(-k) \tag{6.170}$$

Bottom line: for systems with time-reversal symmetry, the dispersion is an even function of the momentum $\vec{k}$.

Bloch wave function $\psi_{n,k}(r) = u_{n,k}(r)\,e^{i\,k\,r}$

Bloch wave under time-reversal transformation:

$$\mathcal{T}\,u_{n,k}(r) = u_{n,-k}(r)^* \tag{6.171}$$

Here, $\mathcal{T}$ changes k to $-k$. In addition, because $u_{n,k}(r)$ is in general a complex function, $\mathcal{T}$ changes the function to its complex conjugate. If the system is time-reversally invariant,

$$\mathcal{T}\,u_{n,k}(r) = e^{i\,\phi(k)}\,u_{n,k}(r) \tag{6.172}$$

By comparing the two equations above, we found that for a system with time-reversal symmetry,

$$u_{n,-k}(r)^* = e^{i\,\phi(k)}\,u_{n,k}(r) \tag{6.173}$$

This arbitrary phase is here because wavefunctions are defined up to an arbitrary phase

Berry connection $\mathcal{A}_n(k) = -i\,\langle u_{n,k}\,|\,\nabla_k\,|\,u_{n,k}\rangle$

The Berry connection is defined as:

$$\mathcal{A}_n(k) = -i\,\langle u_{n,k}\,|\,\nabla_k\,|\,u_{n,k}\rangle = -i\int d\vec{r}\;u_{n,k}\left(\vec{r}\right)^*\,\partial_k\,u_{n,k}\left(\vec{r}\right) \tag{6.174}$$

The under time-reversal, the new $\mathcal{A}_n(k)$ is defined as

$$\begin{aligned}\mathcal{T}\mathcal{A}_n(k) &= -i\,\langle \mathcal{T}\,u_{n,k}\,|\,\nabla_k\,|\,\mathcal{T}\,u_{n,k}\rangle = \\ &-i\int d\vec{r}\;u_{n,-k}\left(\vec{r}\right)\partial_k\,u_{n,-k}\left(\vec{r}\right)^* = i\int d\vec{r}\;\partial_k\,u_{n,-k}\left(\vec{r}\right)u_{n,-k}\left(\vec{r}\right)^* = -i\int d\vec{r}\;u_{n,-k}\left(\vec{r}\right)^*\,\partial_{-k}\,u_{n,-k}\left(\vec{r}\right) = -i\,\langle u_{n,-k}\,|\,\nabla_{-k}\,|\,u_{n,-k}\rangle = \mathcal{A}_n(-k)\end{aligned} \tag{6.175}$$

If the system has time-reversal symmetry

$$\mathcal{T}\mathcal{A}_n(k) = \mathcal{A}_n(k) + \nabla_k\,\chi(k) \tag{6.176}$$

So we have

$$\mathcal{A}_n(-k) = \mathcal{A}_n(k) + \nabla_k\,\chi(k) \tag{6.177}$$

Bottom line: For systems with time-reversal symmetry, $\mathcal{A}_n(-k)$ and $\mathcal{A}_n(k)$ differs by a gauge transformation $\nabla_k\,\chi(k)$.

Berry curvature $\mathcal{F}_n(\boldsymbol{k}) = -i\epsilon_{ij}\left\langle \partial_i u_{n,k} \middle| \partial_j u_{n,k}\right\rangle$

The Berry curvature is defined as:

$$\mathcal{F}_n(k) = -i\epsilon_{\mathrm{ij}}\left\langle \partial_{\mathrm{ki}} u_{n,k} \middle| \partial_{\mathrm{kj}} u_{n,k}\right\rangle = -i\,\epsilon_{\mathrm{ij}}\int d\vec{r}\,\partial_{\mathrm{ki}} u_{n,k}\left(\vec{r}\right)^*\partial_{\mathrm{kj}} u_{n,k}\left(\vec{r}\right) \tag{6.178}$$

The under time-reversal, the new $\mathcal{F}_n(k)$ is defined as

$$\begin{aligned}\mathcal{T}\mathcal{F}_n(k) &= -i\epsilon_{\mathrm{ij}}\left\langle \partial_{\mathrm{ki}}\mathcal{T}\mathrm{u}_{n,k} \middle| \partial_{\mathrm{kj}}\mathcal{T}\mathrm{u}_{n,k}\right\rangle = \\ &-i\,\epsilon_{\mathrm{ij}}\int d\vec{r}\,\partial_{\mathrm{ki}}\mathcal{T}\mathrm{u}_{n,k}\left(\vec{r}\right)^*\partial_{\mathrm{kj}}\mathcal{T}\mathrm{u}_{n,k}\left(\vec{r}\right) = -i\,\epsilon_{\mathrm{ij}}\int d\vec{r}\,\partial_{\mathrm{ki}} u_{n,-k}\left(\vec{r}\right)\partial_{\mathrm{kj}} u_{n,-k}\left(\vec{r}\right)^* = -i\,\epsilon_{\mathrm{ij}}\int d\vec{r}\,\partial_{\mathrm{kj}} u_{n,-k}\left(\vec{r}\right)^*\partial_{\mathrm{ki}} u_{n,-k}\left(\vec{r}\right)\end{aligned} \tag{6.179}$$

If we swap i and j here

$$\mathcal{T}\mathcal{F}_n(k) = -i\,\epsilon_{\mathrm{ji}}\int d\vec{r}\,\partial_{\mathrm{ki}} u_{n,-k}\left(\vec{r}\right)^*\partial_{\mathrm{kj}} u_{n,-k}\left(\vec{r}\right) = i\,\epsilon_{\mathrm{ij}}\int d\vec{r}\,\partial_{\mathrm{ki}} u_{n,-k}\left(\vec{r}\right)^*\partial_{\mathrm{kj}} u_{n,-k}\left(\vec{r}\right) = -\mathcal{F}_n(-k) \tag{6.180}$$

If the system has time-reversal symmetry

$$\mathcal{T}\mathcal{F}_n(k) = \mathcal{F}_n(k) \tag{6.181}$$

So we have

$$\mathcal{F}_n(-k) = -\mathcal{F}_n(k) \tag{6.182}$$

Bottom line: For systems with time-reversal symmetry, $\mathcal{F}_n(k)$ is an odd function of k

Because the integral of an odd function over the whole Brillouin zone must be zero, (the contribution from k and -k cancels each other), the Chern number for a time-reversally invariant system must be C=0.

The Chern number $C = \frac{1}{2\pi}\int\int d\boldsymbol{k}\,\mathcal{F}_n(\boldsymbol{k})$

$$\mathcal{T}\mathrm{C} = \mathcal{T}\,\frac{1}{2\pi}\int\int dk_x\,dk_y\,\mathcal{F}_n(k) = -\frac{1}{2\pi}\int\int dk_x\,dk_y\,\mathcal{F}_n(-k) = -\frac{1}{2\pi}\int\int dk_x\,dk_y\,\mathcal{F}_n(k) = -C \tag{6.183}$$

If the system has time-reversal symmetry,

$$\mathcal{T}\mathrm{C} = C, \tag{6.184}$$

So we have $C = -C$, which means $C = 0$.

Bottom line: to have a nontrivial Chern number, the system must break the time-reversal symmetry (using external B field or some other method).

6.4.2. Space-inversion transformation $\mathcal{I}$ and space-inversion symmetry

Space inversion transformation $\mathcal{I}$ turns a vector $\vec{r}$ into $-\vec{r}$. It also changes $\vec{k}$ into $-\vec{k}$, which is similar to $\mathcal{T}$. However, $\mathcal{I}$ is an unitary transformation, while $\mathcal{T}$ is anti-unitary.

Dispersion relation

Dispersion relation under time-reversal transformation:

$$\mathcal{I}\epsilon_n(k) = \epsilon_n(-k) \tag{6.185}$$

If the system is invariant under space inversion

$$\mathcal{I}\,\epsilon_n(k) = \epsilon_n(k) \tag{6.186}$$

Compare the two equations, we have:

$$\epsilon_n(k) = \epsilon_n(-k) \tag{6.187}$$

Bottom line: For systems invariant under space inversion, the dispersion is an even function of the momentum $\vec{k}$.

If one want make a system where the dispersion is NOT an even function, one need to break both space inversion symmetry and time-reversal symmetry.

Bloch wave function $\psi_{n,k}(r) = u_{n,k}(r)\, e^{i k r}$

$$\mathcal{I}\, u_{n,k}(r) = u_{n,-k}(r) \tag{6.188}$$

If the system is invariant under space inversion

$$\mathcal{I}\, u_{n,k}(r) = e^{i\phi(k)}\, u_{n,k}(r) \tag{6.189}$$

By comparing the two equations above, we found that for a system with time-reversal symmetry,

$$u_{n,-k}(r) = e^{i\phi(k)}\, u_{n,k}(r) \tag{6.190}$$

Berry connection $\mathcal{A}_n(\mathbf{k}) = -i \langle u_{n,k} | \nabla_k | u_{n,k} \rangle$

$$\mathcal{I}\mathcal{A}_n(k) = \mathcal{A}_n(-k) \tag{6.191}$$

If the system is invariant under space inversion

$$\mathcal{I}\mathcal{A}_n(k) = \mathcal{A}_n(k) + \nabla_k \chi(k) \tag{6.192}$$

So we have

$$\mathcal{A}_n(-k) = \mathcal{A}_n(k) + \nabla_k \chi(k) \tag{6.193}$$

Bottom line: For systems invariant under space inversion, $\mathcal{A}_n(-k)$ and $\mathcal{A}_n(k)$ differs by a gauge transformation $\nabla_k \chi(k)$.

Berry curvature $\mathcal{F}_n(\mathbf{k}) = -i\epsilon_{ij} \langle \partial_i u_{n,k} | \partial_j u_{n,k} \rangle$

$$\mathcal{I}\mathcal{F}_n(k) = \mathcal{F}_n(-k) \tag{6.194}$$

If the system is invariant under space inversion

$$\mathcal{I}\mathcal{F}_n(k) = \mathcal{F}_n(k) \tag{6.195}$$

So we have

$$\mathcal{F}_n(-k) = \mathcal{F}_n(k) \tag{6.196}$$

Bottom line: For systems invariant under space inversion, $\mathcal{F}_n(k)$ is an even function of k

Time-reversal symmetry tell us $\mathcal{F}_n(k)$ is an odd function of k. If system have both $\mathcal{T}$ and $\mathcal{I}$ symmetry, $\mathcal{F}$ must be both an even function and an odd function. So $\mathcal{F}$=0 at any k point.

The Chern number $C = \frac{1}{2\pi} \int\int d\mathbf{k} \mathcal{F}_n(\mathbf{k})$

$$\mathcal{I}C = C \tag{6.197}$$

6.4.3. $\mathcal{I}\mathcal{T}$ transformation and $\mathcal{I}\mathcal{T}$ symmetry

Under $\mathcal{I}\mathcal{T}$,

$$\mathcal{I}\mathcal{T}\mathcal{F}_n(k) = \mathcal{I}[-\mathcal{F}_n(-k)] = -\mathcal{F}_n(k) \tag{6.198}$$

If system is invariant under $\mathcal{I}\mathcal{T}$,

$$\mathcal{I}\mathcal{T}\mathcal{F}_n(k) = -\mathcal{F}_n(k) \tag{6.199}$$

So, we have

$$\mathcal{F}_n(k) = -\mathcal{F}_n(k) \tag{6.200}$$

Bottom line: For systems invariant under $\mathcal{I}\mathcal{T}$, $\mathcal{F}_n(k) = 0$ at every k.

For $\mathcal{T}$ symmetry, the integral of $\mathcal{F}$ is zero (C=0), but $\mathcal{F}$ can be nonzero at each k point.

For $\mathcal{I}\mathcal{T}$ symmetry, $\mathcal{F}$ is zero at every single k point. This is a stronger statement!

Please notice that all these conclusions assumes that the system is gapped. For gapless systems with $\mathcal{I}\mathcal{T}$ symmetry, $\mathcal{F}$ can actually be nonzero and quantized, which will be discussed in the next section.

6.5. π flux and Dirac points

6.5.1. magnetic field, Aharonov–Bohm effect and magnetic π- flux

Q: Why B changes sign under $\mathcal{T}$?

A: Many ways to see it. Here, I use the A-B effect.

In the presence of B field, Aharonov and Bohm tells us that if we move a charged-particle around a closed loop, the electron will pick up a phase, and the phase is the total magnetic field enclosed by the loop (times e/$\hbar$)

$$\phi = e/\hbar \int\int B \cdot d\!S = e/\hbar \oint A . d\!l \tag{6.201}$$

Under time-reversal $e^{i\phi}$ turns into $e^{-i\phi}$, because time-reversal is anti-unitary. So the A-B phase changes sign

$$\mathcal{T}\phi = -\phi \tag{6.202}$$

In other words, the integral of B changes sign

$$\mathcal{T} \int\int_D B \cdot d\!S = -\int\int_D B \cdot d\!S \tag{6.203}$$

Because this equation is true for any region D, we have $B = -B$.

If a system has time-reversal symmetry, everything must remain the same before and after we flip the arrow of time, including the A-B phase

$$e^{i\phi} = e^{-i\phi} \tag{6.204}$$

Therefore, we have $\phi = -\phi$. So we have ϕ=0 for any region D, which implies that $B = 0$.

Q: Do we really need to have $B = 0$ to preserve the time-reversal symmetry.

A: NO!

This is because phase is only well-define up to mod 2π. The time-reversal symmetry requires $e^{i\phi} = e^{-i\phi}$, which does NOT implies $\phi = -\phi$. In fact, as long as $\phi = -\phi + 2n\pi$, the A-B phase will be the same before and after the time-reversal transformation. This means that to preserve the time-reversal symmetry we just need $\phi = n\pi$ while n don't need to be 0. So we need to have $\int\int_D B \cdot d\!S = n\pi$ for any D. This means that B can contains some delta functions.

$$B = \frac{\hbar}{e} \sum_i n_i \pi \, \delta(r - r_i) \tag{6.205}$$

For this B field, the integral of B in region D is just

$$\phi = e/\hbar \int\int B \cdot d\!S = \left(\sum_i{}' n_i\right) \pi \tag{6.206}$$

Here the sum $\sum_i{}'$ is over all r_i inside D. A delta function in B is a magnetic flux, which is known as a n×π-flux. These fluxes don't break the time-reversal symmetry. In addition, it worthwhile to point out that for the A-B phase is only well defined up to mod 2π. So all the $2n\pi$ fluxes give the same A-B phase, which is 0. All the $(2n+1)\pi$ fluxes give the same A-B phase, which is π.

6.5.2. Berry flux

We can do the same thing for the Berry curvature $\mathcal{F}$(k). If we goes around a contour in the k-space, the phase change of the Bloch wave function around this contour is

$$\phi = \int\int_D \mathcal{F} \, d\!k^2 \tag{6.207}$$

This is known as the Berry phase.

Under $\mathcal{IT}$, D is invariant so

$$\mathcal{IT}\,\phi = \mathcal{IT} \int\int_D \mathcal{F}(k) \, d\!k^2 = -\int\int_D \mathcal{F}(k) \, d\!k^2 = -\phi \tag{6.208}$$

If the system is $\mathcal{IT}$ invariant, we must have ϕ=-ϕ (but up to mod 2π).

$$\phi = -\phi + 2\,n\,\pi \tag{6.209}$$

so

$$\phi = n\,\pi \tag{6.210}$$

As a result, if we have $\mathcal{IT}$ symmetry, $\int\int_D \mathcal{F}\, d k^2 = n\pi$ for any region D. So $\mathcal{F}$ must be either 0 or some delta functions.

$$\mathcal{F}(k) = \sum_i n_i\,\pi\,\delta(k - k_i) \tag{6.211}$$

These delta functions are known as Berry fluxes. Because the Berry phase is only well defined up to mod 2π, we have in general two types of Berry fluxes: 0 and π. And a π flux is a Dirac point.

6.5.3. Dirac point

Near a Dirac point, the kernel of the Hamiltonian is:

$$\mathcal{H} = v_F \begin{pmatrix} 0 & \mathrm{kx} - i\,\mathrm{ky} \\ \mathrm{kx} + i\,\mathrm{ky} & 0 \end{pmatrix} \tag{6.212}$$

Using polar coordinates

$$\mathcal{H} = v_F\,k \begin{pmatrix} 0 & e^{-i\phi} \\ e^{i\phi} & 0 \end{pmatrix} \tag{6.213}$$

The eigenvalues for this matrix is $\epsilon_\pm = \pm v_F\,k$. The eigenvectors are:

$$u_+ = \frac{1}{\sqrt{2}} \begin{pmatrix} e^{-i\phi} \\ 1 \end{pmatrix} \text{ and } u_- = \frac{1}{\sqrt{2}} \begin{pmatrix} -e^{-i\phi} \\ 1 \end{pmatrix} \tag{6.214}$$

The Berry connection in the polar coordinates has two components, radius and angle.

$$A_{+k} = -i\,\langle u_+ \mid \partial_k \mid u_+ \rangle = 0 \tag{6.215}$$

$$A_{+\phi} = -\frac{i}{k}\left\langle u_+ \middle| \partial_\phi \middle| u_+ \right\rangle = -\frac{i}{k}\,\frac{1}{\sqrt{2}}\,(\,-e^{i\phi}\quad 1\,)\,\partial_\phi\,\frac{1}{\sqrt{2}} \begin{pmatrix} -e^{-i\phi} \\ 1 \end{pmatrix} = -\frac{i}{k}\,\frac{1}{2}\,(\,e^{i\phi}\quad 1\,) \begin{pmatrix} -i\,e^{-i\phi} \\ 1 \end{pmatrix} = -\frac{1}{2\,k} \tag{6.216}$$

Let's choose a small circle around the Dirac point, the flux through this circle is:

$$\chi = \int\int \mathcal{F}\, d k = \oint \mathcal{A}\cdot d k = \int_0^{2\pi} \mathcal{A}_\phi\, k\, d\phi = -\int_0^{2\pi} \frac{1}{2\,k}\, k\, d\phi = -\frac{1}{2}\int_0^{2\pi} d\phi = -\pi \tag{6.217}$$

6.5.4. Why is a Dirac point stable?

This is because the Berry fluxes are very stable. If one has $\mathcal{TI}$ symmetry, the Berry flux is quantized to integer values. Once its value is fixed (say π), no perturbation can change it. The only thing that can change continuously is the location of the delta functions. In other words, without break the $\mathcal{TI}$ symmetry, one can only move the Dirac points around in the k-space, but they cannot just disappear. To get ride of Dirac points, one need to move two Dirac points together and let them annihilate each other (let a π flux meet a -π flux, so that we get zero flux). Another way to get ride of Dirac points is to break either the $\mathcal{T}$ or the $\mathcal{I}$ symmetry.

6.6. Edge states (numerical calculations)

6.6.1. Why energy is a function of k_x and k_y

Consider one electron moving in an infinite 2D lattice (for simplicity, we consider a square lattice here). Because the lattice is invariant under lattice translations ($x \to x + a$ and $y \to y + a$), lattice momentum is a good quantum number (a conserved quantity). Momentum conservation means that

$$[k_x,\, H] = [k_y,\, H] = 0 \tag{6.218}$$

In addition, we know $[k_x,\, k_y] = 0$. With three operators which commute with each other, quantum mechanics tells us that we can find common eigenstates for all these three operators, and we can use these eigenstates as the basis for the Hilbert space. This set of basis looks like:

$$|\psi_{\epsilon_n,k_x,k_y}\rangle \tag{6.219}$$

This state is the eigenstate of H, k_x and k_y with eigenvalues ϵ_n, k_x and k_y respectively. Here, one need a sub-index (n) to distinguish different bands. So we have the band structure:

$$\epsilon_n(k_x,\ k_y) \tag{6.220}$$

6.6.2. What if we have an finite system?

To study edge states, we must have an edge. So we cannot use an infinite 2D plan. Let's consider the simplest case here, a infinitely long stripe. We assumes that the system is infinitely long along x, but it has a finite width along y. Now, the translational symmetry along x is still preserved. But there is no translational symmetry along y (due to the existence of the edge). So k_x is still a good quantum number but k_y is NOT.

$$[k_x,\ H] = 0 \qquad \text{but} \qquad [k_y,\ H] \neq 0 \tag{6.221}$$

So we cannot define common eigenstates for H, k_x and k_y, but we can still find eigenstates for H and k_x, because they commute with each other.

$$|\psi_{\epsilon_m,k_x}\rangle \tag{6.222}$$

Here, again, we need an extra index distinguish different states with the same k_x. Here I emphasize that m is NOT just the band index n. It includes information for both n and k_y. So now, energy is a bunch of function of k_x

$$\epsilon_m(k_x) \tag{6.223}$$

where $m = 1, 2, 3 \ldots N_S$. We have N_S functions of k_x. We can draw these N_S functions on the $k_x - \epsilon$ plane, which gives us N_S curves, which is our 1D dispersion.

6.6.3. Examples

We consider the model of Haldane for a system with infinite size along x, by finite width along y. The Hamiltonian is invariant under translations along the x axis $\left(x \to x + \sqrt{3}\ a\right)$, so the x-component of the momentum is a conserved quantity $[k_x, H] = 0$. So we can find common eigenstates of k_x and H. The system has two edges: one on the top, one at the bottom. We can consider this lattice as a bunch of 1D (horizontal) lines coupled together, so the width are described by the number of chains one have in this infinite stripe (N). For $\epsilon_m(k_x)$, we have m=1,2,…2 N. This factor of 2 comes from the fact that we have two sites in each unit cell for a honeycomb lattice.

First, lets consider the case without NNN hopping (graphene, with two Dirac cones). Here, for the first panel, we show $\epsilon_m(k_x)$ for a system with width N=10 (so there are $N_S = 2\ N = 20$) lines. The second panel is the same calculation for a system with width N=100 ($N_S = 2\ N = 200$). The last panel shows the bulk band structure (side view).

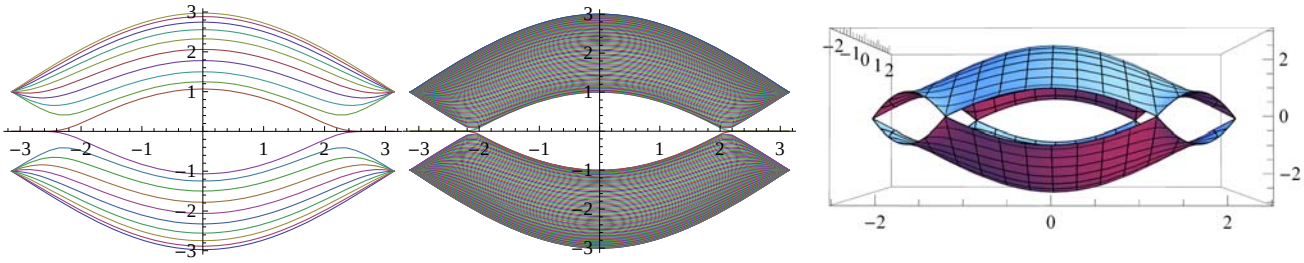

Fig. 7. the energy spectrum for a infinite-long ribbon (the honeycomb lattice)

The 2D figure of $\epsilon_m(k_x)$ is in some sense the 3D band structure $\epsilon_n(k_x,\ k_y)$ projected onto a 2D plane. We can see here an upper band and a lower band and 2 Dirac cones. The system is gapless.

Now, let's consider the topological case, with complex NNN hoppings. The first panel shows the energy dispersion for an infinite stripe and the second one shows the bulk band structure for an infinite 2D system.

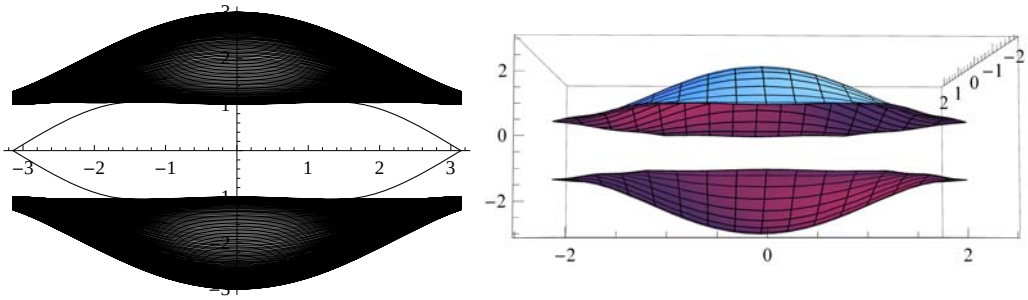

Fig. 8. the energy spectrum for a infinite-long ribbon (the model of Haldane)

Here, I choose the parameter $\phi=\pi/2$, $t_{NN} = 1$ and $t_{NNN} = 0.2$. As one can see, for the infinite stripe, there are two bands, whose energy spectrum coincide with the bulk band structure for infinite systems. But in addition, there are two extra lines shown inside the gap, which are the metallic edge states. By looking at the wavefunctions (see the interactive figures on the course website), one find that one of the in gap state is localized near the top edge, while the other is localized near the bottom edge. If we consider the momentum region $(0,2\pi)$, instead of $(-\pi,\pi)$, we find that one of the two edge states has positive slop and the other one has negative slop (left moving). Because the slop is the velocity of electrons ($v = \partial\epsilon/\partial k$), one of the edge states has positive velocity (right moving) while the other one is right moving (negative velocity).

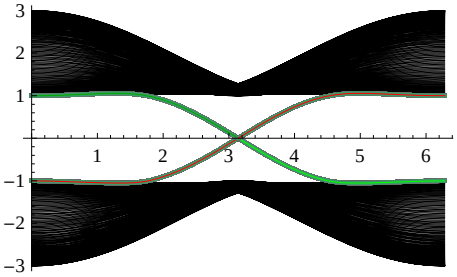

Fig. 9. the velocity of the edge states