Hofstadter physics

Shi Hao

June 20, 2023

1 Hamiltonian

The following note is based on the arXiv paper written by X. Wang & O. Vafek. For a general 2D lattice with unit vector $\boldsymbol{a}_1$ and $\boldsymbol{a}_2$ and reciprocal basis vectors $\boldsymbol{g}_1$ and $\boldsymbol{g}_2$ ($\boldsymbol{a}_i \cdot \boldsymbol{g}_j = 2\pi\delta_{ij}$), the electron Hamiltonian can be written as

$$H(\boldsymbol{p}, \boldsymbol{r}) = T(\boldsymbol{p}) + V(\boldsymbol{r}), \tag{1}$$

where $T(\boldsymbol{p})$ is the kinetic energy which depends purely on $\boldsymbol{p}$, and $V(\boldsymbol{r})$ is the periodic potential energy satisfying $V(\boldsymbol{r}) = V(\boldsymbol{r} + n_1\boldsymbol{a}_1 + n_2\boldsymbol{a}_2)$ for any integers n_1, n_2. The Hamiltonian has the lattice translation symmetry $[H(\boldsymbol{p}, \boldsymbol{r}), T_i] = 0$, where

$$T_1 = e^{-i\boldsymbol{p}\cdot\boldsymbol{a}_1/\hbar}, \quad T_2 = e^{-i\boldsymbol{p}\cdot\boldsymbol{a}_2/\hbar}, \tag{2}$$

so that $T_i f(\boldsymbol{r}) = \langle \boldsymbol{r}|T_i|f\rangle = f(\boldsymbol{r} - \boldsymbol{a}_i)$ for any function $f(\boldsymbol{r})$.

Now under magnetic field $\boldsymbol{B} = B\boldsymbol{e}_z$, the Hamiltonian becomes

$$H(\boldsymbol{p} + e\boldsymbol{A}) = T(\boldsymbol{p} + e\boldsymbol{A}) + V(\boldsymbol{r}) = T(\boldsymbol{\pi}) + V(\boldsymbol{r}), \tag{3}$$

where $\boldsymbol{A}$ is the vector potential, $\boldsymbol{\pi} = \boldsymbol{p} + e\boldsymbol{A}$ is the canonical momentum. In practical calculations we usually choose the Landau gauge, say $\boldsymbol{A} = Bx\boldsymbol{e}_y$, with y-aixs along $\boldsymbol{a}_2$ direction.

2 Landau level basis (no lattice potential)

The absence of the lattice potential $V(\boldsymbol{r})$ reduces the problem to the traditional Landau level problem. In this case we only need to diagonalize $T(\boldsymbol{p} + e\boldsymbol{A})$, and the solution is the Landau level (LL) basis (see my another note called "Landau level formulation")

$$\begin{aligned} \phi_{n,k_y}(\boldsymbol{r}) = \langle \boldsymbol{r}|n, k_y\rangle &= \frac{1}{\sqrt{L_y}} e^{ik_y y}\langle x + k_y l_B^2|n\rangle, \\ \langle x|n\rangle &= \frac{1}{\pi^{1/4}\sqrt{2^n n! l_B}} \exp\left(-\frac{x^2}{2l_B^2}\right) H_n\left(\frac{x}{l_B}\right), \end{aligned} \tag{4}$$

where H_n is the Hermite polynomial, $l_B = \sqrt{\hbar/(eB)}$ is the magnetic length, L_y is the sample length along y direction. LL basis (4) is the common eigenstates of $a^\dagger a$ (with eigenvalue n) and p_y (with eigenvalue $\hbar k_y$), where the ladder operators are defined as ($\boldsymbol{\pi} = \boldsymbol{p} + e\boldsymbol{A}$)

$$a = i\frac{l_B}{\sqrt{2}\hbar}(\pi_x - i\pi_y), \quad a^\dagger = -i\frac{l_B}{\sqrt{2}\hbar}(\pi_x + i\pi_y). \tag{5}$$

The LL degeneracy in the absence of lattice potential is argued as the following. Suppose the sample size is $N_1\boldsymbol{a}_1 \times N_2\boldsymbol{a}_2$, then the quantum number k_y must satisfy $L_x = N_1 a_{1x} = (k_y l_B^2)_{\max} - (k_y l_B^2)_{\min}$, the separation between k_y is $\delta k_y = 2\pi/(N_2 a_2)$, so the LL degeneracy per unit cell is

$$\frac{\mathcal{N}}{N_1 N_2} = \frac{1}{N_1 N_2}\frac{L_x}{l_B^2 \delta k_y} = B\frac{a_{1x}a_2}{h/e} = \frac{\phi}{\phi_0}, \tag{6}$$

where $\phi_0 = h/e$ is the magnetic flux quantum, $\phi = Ba_{1x}a_2$ is the flux through a unit cell.

3 Magnetic translation group

The vector potential will in general break the lattice translation symmetry, but we still can find two magnetic translation operators t_i so that $[H(\boldsymbol{p}+e\boldsymbol{A},\boldsymbol{r}),t_i]=0$. For the present convention, the two operators are

$$t_1 = e^{-i\boldsymbol{q}_\phi\cdot\boldsymbol{r}}T_1, \quad t_2 = T_2, \tag{7}$$

where $\boldsymbol{q}_\phi = \frac{a_{1x}}{l_B^2}\boldsymbol{e}_y = \frac{2\pi}{a_2}\frac{\phi}{\phi_0}\boldsymbol{e}_y$. Although t_i commute with H, they do not commute with each other: $t_2t_1 = \exp(i2\pi\phi/\phi_0)t_1t_2$. However, if $\phi/\phi_0 = p/q$ where p and q are co-prime integers, we have

$$[t_1, t_2^q] = 0. \tag{8}$$

In other words, the eigenstates of $H(\boldsymbol{p}+e\boldsymbol{A},\boldsymbol{r})$ can be chosen as the common eigenstates of t_1 and t_2^q, which is usually called the magnetic translation group (MTG). We can use the LL basis to define a complete and orthonormal set of basis which are eigenstates of t_1 and t_2^q. Suppose the magnetic Brillouin zone (MBZ) is $\boldsymbol{k} = k_1\boldsymbol{g}_1 + k_2\boldsymbol{g}_2$, with $k_1\in[0,1)$, $k_2\in\left[0,\frac{1}{q}\right)$, we define

$$|n,r,k_1,k_2\rangle = \frac{1}{\sqrt{N_1}}\sum_s e^{i2\pi k_1 s}t_1^s\left|n,\frac{2\pi}{a_2}\left(k_2+\frac{r}{q}\right)\right\rangle, \tag{9}$$

for $r=0,1,\ldots,p-1$. We can prove

$$|n,r+p,k_1,k_2\rangle = \exp\left[i2\pi\left(k_1-\left(k_2+\frac{r+p}{q}\right)\frac{a_{1y}}{a_2}\right)\right]|n,r,k_1,k_2\rangle, \tag{10}$$

so $r+p$ and r represent the same state, which is the reason why r should only take p values (rather than q values). Using the identity $t_1^s = e^{i\frac{s(s-1)}{2}\boldsymbol{q}_\phi\cdot\boldsymbol{a}_1}e^{-is\boldsymbol{q}_\phi\cdot\boldsymbol{r}}T_1^s$ (this can be shown iteratively), we have

$$t_1^s|n,k_y\rangle = e^{i\frac{s(s-1)}{2}q_\phi a_{1y}-isk_ya_{1y}}\left|n,k_y-\frac{sa_{1x}}{l_B^2}\right\rangle, \tag{11}$$

$$|n,r,k_1,k_2\rangle = \frac{1}{\sqrt{N_1}}\sum_s e^{i2\pi k_1 s}e^{i\frac{s(s-1)}{2}\frac{2\pi}{a_2}\frac{p}{q}a_{1y}-is\frac{2\pi}{a_2}\left(k_2+\frac{r}{q}\right)a_{1y}}\left|n,\frac{2\pi}{a_2}\left(k_2+\frac{r-sp}{q}\right)\right\rangle. \tag{12}$$

One may understand this basis from a composite tight-binding model view of point: for each k_2, there are p distinguishable atoms indexed by r, and the above basis is just a Fourier transform (Bloch sum) of these basis. We can check ($[p]_1$ represents the fractional part of p)

$$\begin{aligned}
t_1|n,r,k_1,k_2\rangle &= e^{-i2\pi k_1}|n,r,k_1,k_2\rangle,\\
t_2|n,r,k_1,k_2\rangle &= e^{-i2\pi\left(k_2+\frac{r}{q}\right)}\left|n,r,\left[k_1+\frac{p}{q}\right]_1,k_2\right\rangle,\\
t_2^q|n,r,k_1,k_2\rangle &= e^{-i2\pi k_2 q}|n,r,k_1,k_2\rangle.
\end{aligned} \tag{13}$$

The following orthonormality can be proved, which is a special case of the plane wave matrix, see Eq. (17),

$$\langle n',r',k_1',k_2'|n,r,k_1,k_2\rangle = \delta_{n'n}\delta_{r'r}\delta_{k_1'k_1}\delta_{k_2'k_2}. \tag{14}$$

In the absence of lattice potential, the LL degeneracy per unit cell is $\frac{\mathcal{N}}{N_1N_2} = \frac{1}{N_1N_2}\left(p\times N_1\times\frac{N_2}{q}\right) = \frac{\phi}{\phi_0}$, consistent with the previous argument. Here p comes from the degeneracy of index r, N_1 represents the number of k_1, and $\frac{N_2}{q}$ indicate the number of k_2 living in the MBZ.

4 Matrix element

We now write down the matrix element of Hamiltonian under basis (9).

For the kinetic energy $T(\boldsymbol{p}+e\boldsymbol{A})$, the canonical momentum $\boldsymbol{\pi}=\boldsymbol{p}+e\boldsymbol{A}$ is related with the ladder operator a and $a^\dagger$ by

$$\pi_x = i\frac{\hbar}{\sqrt{2}l_B}(a^\dagger - a), \quad \pi_y = \frac{\hbar}{\sqrt{2}l_B}(a + a^\dagger). \tag{15}$$

So we only need to calculate the matrix element of a and $a^\dagger$, which act only on the quantum number n so basis (9) still satisfies standard oscillator algebra. The matrix is diagonal in $\boldsymbol{k}=k_1\boldsymbol{g}_1+k_2\boldsymbol{g}_2$, which under $|n,r,\boldsymbol{k}\rangle = |n,r,k_1,k_2\rangle$ reads

$$\begin{aligned}
\langle n',r',\boldsymbol{k}|a|n,r,\boldsymbol{k}\rangle &= \sqrt{n}\delta_{n'+1,n}\delta_{r'r},\\
\langle n',r',\boldsymbol{k}|a^\dagger|n,r,\boldsymbol{k}\rangle &= \sqrt{n'}\delta_{n',n+1}\delta_{r'r},\\
\langle n',r',\boldsymbol{k}|a^\dagger a|n,r,\boldsymbol{k}\rangle &= n\delta_{n'n}\delta_{r'r}.
\end{aligned} \tag{16}$$

For the potential term $V(\boldsymbol{r})$, which in general can be expanded using plane waves $V(\boldsymbol{r})=\sum_{\boldsymbol{q}}V_{\boldsymbol{q}}e^{i\boldsymbol{q}\cdot\boldsymbol{r}}$, where $\boldsymbol{q}=q_1\boldsymbol{g}_1+q_2\boldsymbol{g}_2$. So we only need to calculate the matrix element of plane waves. Using $t_1^s e^{i\boldsymbol{q}\cdot\boldsymbol{r}} = e^{-is\boldsymbol{q}\cdot\boldsymbol{a}_1}e^{i\boldsymbol{q}\cdot\boldsymbol{r}}t_1^s$ and the matrix element under LL basis $|n,k_y\rangle$ (see another note called "Landau level formulation"), we have

$$\begin{aligned}
&\langle n',r',\boldsymbol{k}'|e^{i\boldsymbol{q}\cdot\boldsymbol{r}}|n,r,\boldsymbol{k}\rangle\\
=&\frac{1}{N_1}\sum_{s's}e^{i2\pi(k_1 s-k_1' s')}e^{is'\boldsymbol{q}\cdot\boldsymbol{a}_1}\left\langle n',\frac{2\pi}{a_2}\left(k_2'+\frac{r'}{q}\right)\middle|e^{i\boldsymbol{q}\cdot\boldsymbol{r}}t_1^{s-s'}\middle|n,\frac{2\pi}{a_2}\left(k_2+\frac{r}{q}\right)\right\rangle\\
=&\frac{1}{N_1}\sum_{s's}e^{is'[2\pi(k_1-k_1')+\boldsymbol{q}\cdot\boldsymbol{a}_1]}e^{i2\pi k_1 s}\left\langle n',\frac{2\pi}{a_2}\left(k_2'+\frac{r'}{q}\right)\middle|e^{i\boldsymbol{q}\cdot\boldsymbol{r}}t_1^{s}\middle|n,\frac{2\pi}{a_2}\left(k_2+\frac{r}{q}\right)\right\rangle\\
=&\delta_{k_1',[k_1+q_1]_1}\sum_s e^{i2\pi k_1 s}\left\langle n',\frac{2\pi}{a_2}\left(k_2'+\frac{r'}{q}\right)\middle|e^{i\boldsymbol{q}\cdot\boldsymbol{r}}t_1^{s}\middle|n,\frac{2\pi}{a_2}\left(k_2+\frac{r}{q}\right)\right\rangle\\
=&\delta_{k_1',[k_1+q_1]_1}\sum_s e^{i2\pi k_1 s}e^{i\frac{s(s-1)}{2}q_\phi a_{1y}-is\frac{2\pi}{a_2}\left(k_2+\frac{r}{q}\right)a_{1y}}\\
&\times\left\langle n',\frac{2\pi}{a_2}\left(k_2'+\frac{r'}{q}\right)\middle|e^{i\boldsymbol{q}\cdot\boldsymbol{r}}\middle|n,\frac{2\pi}{a_2}\left(k_2+\frac{r-sp}{q}\right)\right\rangle\\
=&\delta_{k_1',[k_1+q_1]_1}\sum_s \delta_{k_2'+\frac{r'}{q},k_2+\frac{r-sp}{q}+q_2}e^{i2\pi k_1 s}e^{i\frac{s(s-1)}{2}q_\phi a_{1y}-is\frac{2\pi}{a_2}\left(k_2+\frac{r}{q}\right)a_{1y}}\\
&\times e^{-iq_x\frac{2\pi}{a_2}\left(k_2'+\frac{r'}{q}\right)l_B^2+i\frac{q_x q_y}{2}l_B^2}F_{n'n}\left(\frac{\boldsymbol{q}l_B}{\sqrt{2}}\right),
\end{aligned} \tag{17}$$

where $[k_1+q_1]_1$ is the fractional part of k_1+q_1, i.e., the part in MBZ. The form factor

$$F_{n'n}(\boldsymbol{Q}) = \begin{cases} e^{-\frac{\boldsymbol{Q}^2}{2}}\sqrt{\frac{n'!}{n!}}\left[i(Q_x+iQ_y)\right]^{n-n'}L_{n'}^{n-n'}\left(\boldsymbol{Q}^2\right) & (n\geqslant n')\\ e^{-\frac{\boldsymbol{Q}^2}{2}}\sqrt{\frac{n!}{n'!}}\left[i(Q_x-iQ_y)\right]^{n'-n}L_{n}^{n'-n}\left(\boldsymbol{Q}^2\right) & (n'\geqslant n)\end{cases}, \tag{18}$$

and $L_n^\alpha(x)$ is the associated Laguerre polynomial. Notice that in lattice system, $\boldsymbol{q}=q_1\boldsymbol{g}_1+q_2\boldsymbol{g}_2$ can only take integer values of q_1,q_2, in (17) we then get $\delta_{k_1',[k_1+q_1]_1}=\delta_{k_1',k_1}$, and $\delta_{k_2'+\frac{r'}{q},k_2+\frac{r-sp}{q}+q_2}=\delta_{k_2',k_2}\delta_{r',r-sp+q_2 q}$, i.e., the matrix is diagonal in $\boldsymbol{k}=k_1\boldsymbol{g}_1+k_2\boldsymbol{g}_2$, and

$$\begin{aligned}
\langle n'r'\boldsymbol{k}|e^{i\boldsymbol{q}\cdot\boldsymbol{r}}|nr\boldsymbol{k}\rangle =&F_{n'n}\left(\frac{\boldsymbol{q}l_B}{\sqrt{2}}\right)\sum_s\delta_{r',r-sp+q_2q}e^{i2\pi k_1 s}e^{i\frac{s(s-1)}{2}q_\phi a_{1y}-is\frac{2\pi}{a_2}\left(k_2+\frac{r}{q}\right)a_{1y}}e^{-iq_x\frac{2\pi}{a_2}\left(k_2+\frac{r'}{q}\right)l_B^2+i\frac{q_xq_y}{2}l_B^2}\\
=&F_{n'n}\left(\frac{\boldsymbol{q}l_B}{\sqrt{2}}\right)\sum_s\delta_{r',r-sp+q_2q}e^{i2\pi k_1 s}e^{i\pi\frac{p}{q}\frac{a_{1y}}{a_2}s(s-1)-i2\pi s\frac{a_{1y}}{a_2}(k_2+\frac{r}{q})}e^{i2\pi\frac{q}{p}(\frac{a_{1y}}{a_2}q_2-q_1)(k_2+\frac{r'}{q}-\frac{q_2}{2})},
\end{aligned} \tag{19}$$

while in the summation only at most one s survives. It is clear that Eq. (19) is sparse about momentum index r, but dense about the LL index n.

Finally let's discuss the degeneracy of the Hofstadter band obtained by diagonalizing $H(\boldsymbol{\pi},\boldsymbol{r})$ under basis $|n,r,\boldsymbol{k}\rangle$. At each $\boldsymbol{k}=k_1\boldsymbol{g}_1+k_2\boldsymbol{g}_2\in$ MBZ, where $k_1\in[0,1)$, $k_2\in[0,1/q)$, the eigenstates satisfying

$H|\Phi_{m\boldsymbol{k}}\rangle = \varepsilon_{m\boldsymbol{k}}|\Phi_{m\boldsymbol{k}}\rangle$ is a linear expansion of $|n, r, \boldsymbol{k}\rangle$. Now since $[H, t_2] = 0$, we have $Ht_2|\Phi_{m\boldsymbol{k}}\rangle = t_2H|\Phi_{m\boldsymbol{k}}\rangle = \varepsilon_{m\boldsymbol{k}}t_2|\Phi_{m\boldsymbol{k}}\rangle$, but according to (13) t_2 translates k_1 to another point $[k_1 + p/q]_1$ which is not k_1 since p and q are co-prime (the only exceptional case happens when $p = q = 1$). So $t_2|\Phi_{m\boldsymbol{k}}\rangle$ is another eigenstate with another momentum but with the same energy. More generally, we see all the q states $t_2^s|\Phi_{m\boldsymbol{k}}\rangle, s = 0, 1, \ldots, q-1$ are degenerate. So the spectrum is at least q-fold degenerate and the resulting Hofstadter band is effectively restricted to the reduced domain $k_1, k_2 \in [0, 1/q) \times [0, 1/q)$.

Eqs. (16) and (19) are traditional LL method to diagonalize $H(\boldsymbol{\pi}, \boldsymbol{r})$, which is an exact method whose accuracy only depends on the cutoff of LL basis N_c corresponding to an energy window $\sim N_c\frac{\hbar^2}{ml_B^2}$. In the low-field case, in principle infinite LL cutoff is needed, while the matrix under LL basis is dense. This is the main defect of this method.

5 The Diophantine equation

For the system with flux $\phi/\phi_0 = p/q$ per unit cell, where p and q are co-prime, the energy spectrum will split into a set of Hofstadter bands. The ν-th gap (from low- to high-energy, with a reference gap usually taken as $-\infty$) is characterized by a pair of integers (t_ν, s_ν), which satisfies the Diophantine equation

$$\nu = t_\nu p + s_\nu q. \tag{20}$$

Specifically, t_ν is the Chern number of the ν-th gap, i.e., the Chern number of the ν-th band is $\sigma_\nu = t_\nu - t_{\nu-1}$. In the following we give a proof to the Diophantine equation, which is copied from I. Dana's paper (J. Phys. C: Solid State Phys 18 L679).

5.1 A heuristic proof

We only consider the case the non-degenerate (at each $\boldsymbol{k}$) case. Suppose the ν-th magnetic band with band Chern number σ_ν has the wave function $\psi_{\nu,k_1,k_2}(\boldsymbol{r})$, i.e., $H\psi_{\nu,k_1,k_2} = E_\nu(k_1, k_2)\psi_{\nu,k_1,k_2}$, where $k_1 \in [0, 1)$, $k_2 \in [0, 1/q)$. Then from the translational properties of the LL basis, we have

$$t_1\psi_{\nu,k_1,k_2} = e^{-i2\pi k_1}\psi_{\nu,k_1,k_2}, t_2^q\psi_{\nu,k_1,k_2} = e^{-i2\pi k_2 q}\psi_{\nu,k_1k_2}. \tag{21}$$

In consistency with the band Chern number σ_ν, we may choose the boundary condition on ψ_{ν,k_1,k_2} (one may easily check that the following B.C. could give the correct Chern number by calculating the phase accumulation along the closed path $(k_1, k_2) = (0, 0) \to (1, 0) \to (1, 1/q) \to (0, 1/q) \to (0, 0)$. Other B.C. is also allowable, but I still want to find a more general proof)

$$\psi_{\nu,k_1+1,k_2} = e^{-i2\pi\sigma_\nu k_2 q}\psi_{\nu,k_1,k_2}, \quad \psi_{\nu,k_1,k_2+\frac{1}{q}} = \psi_{\nu,k_1,k_2}. \tag{22}$$

We know that $t_2\psi_{\nu,k_1,k_2} \sim \psi_{\nu,k_1+\frac{p}{q},k_2}$ is still an eigenstate with the same energy, so in general we can assume

$$t_2\psi_{\nu,k_1,k_2} = e^{-i2\pi m_\nu k_2 q}\psi_{\nu,k_1+\frac{p}{q},k_2}, \tag{23}$$

which then gives

$$t_2^q\psi_{\nu,k_1,k_2} = e^{-i2\pi m_\nu k_2 q^2}\psi_{\nu,k_1+p,k_2} = e^{-i2\pi k_2 q(m_\nu q+\sigma_\nu p)}\psi_{\nu,k_1,k_2} = e^{-i2\pi k_2 q}\psi_{\nu,k_1k_2}, \tag{24}$$

meaning $m_\nu q + \sigma_\nu p = 1$ since k_2 can take any numbers. This equation is equivalent to the Diophantine equation above by noticing $\sigma_\nu = t_\nu - t_{\nu-1}$ and $m_\nu = s_\nu - s_{\nu-1}$, once m_ν can be proved to be integers as the following:

$$\begin{aligned} e^{-i2\pi m_\nu k_2 q}\psi_{\nu,k_1+\frac{p}{q},k_2} =& t_2\psi_{\nu,k_1,k_2} = t_2\psi_{\nu,k_1,k_2+\frac{1}{q}} = e^{-i2\pi m_\nu\left(k_2+\frac{1}{q}\right)q}\psi_{\nu,k_1+\frac{p}{q},k_2+\frac{1}{q}} \\ =& e^{-i2\pi m_\nu}e^{-i2\pi m_\nu k_2 q}\psi_{\nu,k_1+\frac{p}{q},k_2}. \end{aligned} \tag{25}$$

The Diophantine equation is a general and simple relation to tell us the Chern number or Hall conductance of the system. It is irrelative about the specific lattice, and is a purely result of the magnetic translation symmetry.

5.2 Relation with the Streda formula

Since in the flux $\phi/\phi_0 = p/q$, in real space the magnetic unit cell has an area q times of the original unit cell (the magnetic translation group elements are t_1 and t_2^q), the electron density, i.e., the number of electrons per original unit cell is $\rho = \nu/q$. From Diophantine equation we know

$$\rho = \frac{\nu}{q} = t_\nu \frac{p}{q} + s_\nu = t_\nu \frac{e}{h} BS + s_\nu. \tag{26}$$

Taking the gradients with respect to the flux, we reach the Streda formula

$$\sigma_{xy}^\nu = \frac{e^2}{h} t_\nu = \frac{e}{S} \frac{\partial \rho}{\partial B}. \tag{27}$$

6 Numerical method to calculate Chern number

Suppose at flux $\phi/\phi_0 = p/q$ the ν-th eigenstate satisfying $H\psi_{\nu\boldsymbol{k}}(\boldsymbol{r}) = E_\nu(\boldsymbol{k})\psi_{\nu\boldsymbol{k}}(\boldsymbol{r})$ is

$$\psi_{\nu\boldsymbol{k}}(\boldsymbol{r}) = \sum_{n,r} \langle \boldsymbol{r}|nr\boldsymbol{k}\rangle P_{nr,\nu}(\boldsymbol{k}). \tag{28}$$

Similar to the Bloch theorem, we can define a specific set of functions $u_{\nu\boldsymbol{k}}(\boldsymbol{r}) = e^{-i\boldsymbol{k}\cdot\boldsymbol{r}}\psi_{\nu\boldsymbol{k}}(\boldsymbol{r})$, which can be proved to satisfy

$$t_1 u_{\nu\boldsymbol{k}}(\boldsymbol{r}) = u_{\nu\boldsymbol{k}}(\boldsymbol{r}), \quad t_2^q u_{\nu\boldsymbol{k}}(\boldsymbol{r}) = u_{\nu\boldsymbol{k}}(\boldsymbol{r}). \tag{29}$$

Correspondingly, the $\boldsymbol{k}$-resolved Hamiltonian reads $H_{\boldsymbol{k}} = e^{-i\boldsymbol{k}\cdot\boldsymbol{r}} H e^{i\boldsymbol{k}\cdot\boldsymbol{r}}$.

The ν-th band Chern number can be numerically calculated as the following. For a series of multiple bands, we can simply replace U matrices below as the determinant in the band index while all other procedures are intact. First we discretize the MBZ $(k_1, k_2) \in [0,1) \otimes [0, 1/q)$ as $\boldsymbol{k}_{ij} = \left(\frac{i}{N_1}, \frac{j}{N_2 q}\right)$, for $i = 0, 1, \ldots, N_1 - 1$, $j = 0, 1, \ldots, N_2 - 1$. Then we can calculate the U and $\mathcal{F}$ matrices as

$$\begin{aligned} U_1^\nu(\boldsymbol{k}_{ij}) &= \langle u_{\nu,\boldsymbol{k}_{ij}} | u_{\nu,\boldsymbol{k}_{i+1,j}}\rangle / |\langle u_{\nu,\boldsymbol{k}_{ij}} | u_{\nu,\boldsymbol{k}_{i+1,j}}\rangle|, \\ U_2^\nu(\boldsymbol{k}_{ij}) &= \langle u_{\nu,\boldsymbol{k}_{ij}} | u_{\nu,\boldsymbol{k}_{i,j+1}}\rangle / |\langle u_{\nu,\boldsymbol{k}_{ij}} | u_{\nu,\boldsymbol{k}_{i,j+1}}\rangle|, \\ \mathcal{F}^\nu(\boldsymbol{k}_{ij}) &= \ln\left[\frac{U_1^\nu(\boldsymbol{k}_{ij}) U_2^\nu(\boldsymbol{k}_{i+1,j})}{U_1^\nu(\boldsymbol{k}_{i,j+1}) U_2^\nu(\boldsymbol{k}_{ij})}\right], \end{aligned} \tag{30}$$

where the phase of $\mathcal{F}$ matrix is confined as $-\pi < \mathcal{F}^\nu(\boldsymbol{k}_{ij})/i < \pi$. Finally the ν-th band Chern number reads

$$t_\nu = \frac{i}{2\pi} \sum_{ij} \mathcal{F}^\nu(\boldsymbol{k}_{ij}). \tag{31}$$

In the above procedure the only different part compared with the normal tight-binding or plane wave basis is the overlap of $u_{\nu\boldsymbol{k}}(\boldsymbol{r})$, which simply again involves the plane wave matrix element

$$\langle u_{\nu',\boldsymbol{k}'} | u_{\nu,\boldsymbol{k}}\rangle = \sum_{n'r'} \sum_{nr} P^*_{n'r',\nu'}(\boldsymbol{k}') \langle n'r'\boldsymbol{k}'| e^{i(\boldsymbol{k}'-\boldsymbol{k})\cdot\boldsymbol{r}} |nr\boldsymbol{k}\rangle P_{nr,\nu}(\boldsymbol{k}) = P^\dagger(\boldsymbol{k}') X(\boldsymbol{k}'-\boldsymbol{k}) P(\boldsymbol{k}), \tag{32}$$

$$\begin{aligned} [X(\boldsymbol{k}'-\boldsymbol{k})]_{n'r',nr} =& \langle n'r'\boldsymbol{k}'| e^{i(\boldsymbol{k}'-\boldsymbol{k})\cdot\boldsymbol{r}} |nr\boldsymbol{k}\rangle \\ =& \delta_{k_1', [k_1 + (k_1' - k_1)]_1} \sum_s \delta_{k_2' + \frac{r'}{q}, k_2 + \frac{r - sp}{q} + k_2' - k_2} e^{i2\pi k_1 s} e^{i\pi s(s-1)\frac{p}{q}\frac{a_{1y}}{a_2} - is2\pi\left(k_2 + \frac{r}{q}\right)\frac{a_{1y}}{a_2}} \\ & \times e^{-i(k_x' - k_x)\frac{2\pi}{a_2}\left(k_2' + \frac{r'}{q}\right) l_B^2 + i\frac{l_B^2}{2}(k_x' - k_x)(k_y' - k_y)} F_{n'n}\left(\frac{(\boldsymbol{k}'-\boldsymbol{k}) l_B}{\sqrt{2}}\right) \\ =& F_{n'n}\left(\frac{(\boldsymbol{k}'-\boldsymbol{k}) l_B}{\sqrt{2}}\right) \delta_{r'r} e^{i2\pi\frac{q}{p}\left[\frac{a_{1y}}{a_2}(k_2' - k_2) - (k_1' - k_1)\right]\left(\frac{r}{q} + \frac{k_2' + k_2}{2}\right)}. \end{aligned} \tag{33}$$

7 Examples

7.1 Square lattice

For a square lattice with $\boldsymbol{a}_1 = (a, 0)$, $\boldsymbol{a}_2 = (0, a)$, the Hamiltonian reads (be careful V_0 should be positive to make the lattice potential a quantum-well-like one near atomic site)

$$H(\boldsymbol{\pi}, \boldsymbol{r}) = \frac{(\boldsymbol{p} + e\boldsymbol{A})^2}{2m} + V(\boldsymbol{r}) = \frac{\hbar^2}{l_B^2 m}\left(a^\dagger a + \frac{1}{2}\right) - V_0(e^{i\boldsymbol{g}_1\cdot\boldsymbol{r}} + e^{i\boldsymbol{g}_2\cdot\boldsymbol{r}} + e^{-i\boldsymbol{g}_1\cdot\boldsymbol{r}} + e^{-i\boldsymbol{g}_2\cdot\boldsymbol{r}}), \tag{34}$$

the matrix elements are

$$\begin{aligned}
\langle n', r', \boldsymbol{k}|\frac{\hbar^2}{l_B^2 m}\left(a^\dagger a + \frac{1}{2}\right)|n, r, \boldsymbol{k}\rangle &= \frac{\hbar^2}{l_B^2 m}\left(n + \frac{1}{2}\right)\delta_{n'n}\delta_{r'r}, \\
\langle n', r', \boldsymbol{k}|e^{i\boldsymbol{g}_1\cdot\boldsymbol{r}}|n, r, \boldsymbol{k}\rangle &= F_{n'n}\left(\frac{\boldsymbol{g}_1 l_B}{\sqrt{2}}\right)\delta_{r'r}e^{-i2\pi\frac{q}{p}\left(k_2+\frac{r}{q}\right)}, \\
\langle n', r', \boldsymbol{k}|e^{i\boldsymbol{g}_2\cdot\boldsymbol{r}}|n, r, \boldsymbol{k}\rangle &= F_{n'n}\left(\frac{\boldsymbol{g}_2 l_B}{\sqrt{2}}\right)\sum_s \delta_{r',r-sp+q}e^{i2\pi k_1 s}.
\end{aligned} \tag{35}$$

7.2 Triangular lattice

The lattice vectors $\boldsymbol{a}_1 = a\left(\frac{\sqrt{3}}{2}, \frac{1}{2}\right), \boldsymbol{a}_2 = a(0, 1)$, $\boldsymbol{g}_1 = \frac{4\pi}{\sqrt{3}a}(1, 0), \boldsymbol{g}_2 = \frac{4\pi}{\sqrt{3}a}\left(-\frac{1}{2}, \frac{\sqrt{3}}{2}\right)$. The kinetic energy is the same as the square one, while the potential term becomes ($\boldsymbol{g}_3 = \boldsymbol{g}_1 + \boldsymbol{g}_2$)

$$V(\boldsymbol{r}) = -V_0[e^{i\boldsymbol{g}_1\cdot\boldsymbol{r}} + e^{i\boldsymbol{g}_2\cdot\boldsymbol{r}} + e^{i\boldsymbol{g}_3\cdot\boldsymbol{r}} + e^{-i\boldsymbol{g}_1\cdot\boldsymbol{r}} + e^{-i\boldsymbol{g}_2\cdot\boldsymbol{r}} + e^{-i\boldsymbol{g}_3\cdot\boldsymbol{r}}]. \tag{36}$$

Its matrix elements are

$$\begin{aligned}
\langle n', r', \boldsymbol{k}|e^{i\boldsymbol{g}_1\cdot\boldsymbol{r}}|n, r, \boldsymbol{k}\rangle &= F_{n'n}\left(\frac{\boldsymbol{g}_1 l_B}{\sqrt{2}}\right)\delta_{r'r}e^{-i2\pi\frac{q}{p}\left(k_2+\frac{r}{q}\right)}, \\
\langle n', r', \boldsymbol{k}|e^{i\boldsymbol{g}_2\cdot\boldsymbol{r}}|n, r, \boldsymbol{k}\rangle &= F_{n'n}\left(\frac{\boldsymbol{g}_2 l_B}{\sqrt{2}}\right)\sum_s \delta_{r',r-sp+q}e^{i2\pi k_1 s}e^{i\frac{\pi}{2}\frac{p}{q}s(s-1)-i\pi s\left(k_2+\frac{r}{q}\right)}e^{i\pi\frac{q}{p}\left(k_2+\frac{r'}{q}-\frac{1}{2}\right)}, \\
\langle n', r', \boldsymbol{k}|e^{i\boldsymbol{g}_3\cdot\boldsymbol{r}}|n, r, \boldsymbol{k}\rangle &= F_{n'n}\left(\frac{\boldsymbol{g}_3 l_B}{\sqrt{2}}\right)\sum_s \delta_{r',r-sp+q}e^{i2\pi k_1 s}e^{i\frac{\pi}{2}\frac{p}{q}s(s-1)-i\pi s\left(k_2+\frac{r}{q}\right)}e^{-i\pi\frac{q}{p}\left(k_2+\frac{r'}{q}-\frac{1}{2}\right)}.
\end{aligned} \tag{37}$$

7.3 Twisted bilayer graphene

The twisted bilayer graphene (TBG) has the BM continuum model, so under magnetic field it can also be diagonalized using the method in this note. However, this time we'd better use the eigenstates of graphene (without the moiré potential, $h_l = v_F(\boldsymbol{\pi} - \hbar\boldsymbol{K}_l)\cdot\boldsymbol{\sigma}$) to form the MTG basis, in order to get a closed basis set for diagonalization. Otherwise, even we have a relatively large cutoff for LLs, the kinetic term itself will scatter the highest LLs into outer space.

Since this system is of great importance, we leave this part in another note called "Note_TBG_Hofstadter".