Tight-binding Theory

1. Localized and Bloch basis sets

We begin our discussion with the concept of primitive cell. As indicated by its name, primitive cell is the primary building block of macroscopic materials, and cannot be split into more fundamental parts. Any primitive cell should have 3 lattice vectors, namely $\boldsymbol{a}_1$, $\boldsymbol{a}_2$ and $\boldsymbol{a}_3$, whose Cartesian coordinates can be arranged into a 3×3 matrix as

$$A=\begin{pmatrix} a_{1x} & a_{1y} & a_{1z} \\ a_{2x} & a_{2y} & a_{2z} \\ a_{3x} & a_{3y} & a_{3z} \end{pmatrix} \tag{1}$$

A primitive cell should also contain atoms located at different positions, each carrying at least one orbital. We label the atomic orbitals as $\phi_i=\phi_i(\boldsymbol{r}-\tau_i)$, with i being the orbital index and τ_i being the position of corresponding atom. Take bulk silicon for example. The primitive cell of silicon has two Si atoms, each carrying one 3s and three 3p orbitals. We may label them as $\phi_1-\phi_8$, with $\phi_1-\phi_4$ and $\phi_5-\phi_8$ sharing the same atomic positions, respectively.

Bulk materials are formed by replicating the primitive cell in 3 directions. We denote the number of repeated primitive cells along different directions as N_1-N_3. The total number of primitive cells is thus $N=N_1\times N_2\times N_3$. We label each cell in the bulk with a tuple of integers (r_1,r_2,r_3), and define the position vector $\boldsymbol{R}$ as

$$\boldsymbol{R}\equiv r_1\boldsymbol{a}_1+r_2\boldsymbol{a}_2+r_3\boldsymbol{a}_3,\ r_i\in[0,N_i-1] \tag{2}$$

Now the atomic orbitals have an additional label $\boldsymbol{R}$ and their spatial distribution becomes

$$\phi_{\boldsymbol{R}i}(\boldsymbol{r})\equiv\phi_i(\boldsymbol{r}-\boldsymbol{R}-\tau_i) \tag{3}$$

We define the atomic orbitals in eqn. (3) as *localized basis*. The size of localized basis equals to the number of orbitals per cell multiplied by the number of primitive cells. Since there are so-many cells in macroscopic materials, e.g., a silicon single crystal with volume of 1 mm^3 has approximately 2.5×10^{19} primitive cells, the actual size of localized basis is extremely large.

From the localized basis, we can construct the *Bloch basis* via unitary transforms, following different conventions:

$$\chi_{\boldsymbol{k}i}=\frac{1}{\sqrt{N}}\sum_{\boldsymbol{R}}\mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot(\boldsymbol{R}+\tau_i)}\phi_{\boldsymbol{R}i} \tag{4}$$

$$\chi_{\boldsymbol{k}i}=\frac{1}{\sqrt{N}}\sum_{\boldsymbol{R}}\mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{R}}\phi_{\boldsymbol{R}i} \tag{5}$$

with

$$\boldsymbol{k}\equiv\frac{k_1}{N_1}\boldsymbol{b}_1+\frac{k_2}{N_2}\boldsymbol{b}_2+\frac{k_3}{N_3}\boldsymbol{b}_3,\ k\in[0,N_i-1] \tag{6}$$

being the wave vector and $\boldsymbol{b}_i$ being reciprocal lattice vectors defined as

$$\boldsymbol{a}_i\cdot\boldsymbol{b}_j=2\pi\delta_{ij} \tag{7}$$

We label eqn. (4) as convention I and eqn. (5) as convention II, respectively. For the difference of these conventions, see the note of PythTB. TBPlaS uses convention I most of the time, and switches to convention II when needed. It is straightforward to show that the number of k-points equals to the number of primitive cells. Since there are enormous primitive cells within macroscopic materials, the k-points are so dense in the Brillouin zone that they become quasi-continual.

The orthogonality between Bloch functions in convention II follows

$$\begin{aligned}\langle\chi_{\boldsymbol{k}i}|\chi_{\boldsymbol{k}'j}\rangle&=\frac{1}{N}\left(\sum_{\boldsymbol{R}}\mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{R}}\langle\phi_{\boldsymbol{R}i}|\right)\left(\sum_{\boldsymbol{R}'}\mathrm{e}^{\mathrm{i}\boldsymbol{k}'\cdot\boldsymbol{R}'}|\phi_{\boldsymbol{R}'j}\rangle\right)\\&=\frac{1}{N}\sum_{\boldsymbol{R}}\sum_{\boldsymbol{R}'}\mathrm{e}^{\mathrm{i}(\boldsymbol{k}'\cdot\boldsymbol{R}'-\boldsymbol{k}\cdot\boldsymbol{R})}\langle\phi_{\boldsymbol{R}i}|\phi_{\boldsymbol{R}'j}\rangle\end{aligned} \tag{8}$$

Since the localized orbitals are localized enough, and they are often chosen to be the eigenstates of atomic Hamiltonian of some sort, we can safely make the assumption that $\langle\phi_{\boldsymbol{R}i}|\phi_{\boldsymbol{R}'j}\rangle=\delta_{\boldsymbol{R}\boldsymbol{R}'}\delta_{ij}$. Then eqn. (8) becomes

$$\begin{aligned}\langle\chi_{\boldsymbol{k}i}|\chi_{\boldsymbol{k}'j}\rangle&=\frac{1}{N}\sum_{\boldsymbol{R}}\sum_{\boldsymbol{R}'}\mathrm{e}^{\mathrm{i}(\boldsymbol{k}'\cdot\boldsymbol{R}'-\boldsymbol{k}\cdot\boldsymbol{R})}\delta_{\boldsymbol{R}\boldsymbol{R}'}\delta_{ij}\\&=\frac{1}{N}\sum_{\boldsymbol{R}}\mathrm{e}^{\mathrm{i}(\boldsymbol{k}'-\boldsymbol{k})\cdot\boldsymbol{R}}\delta_{ij}\\&=\delta_{\boldsymbol{k}\boldsymbol{k}'}\delta_{ij}\end{aligned} \tag{9}$$

Similar conclusions can be drawn for Bloch basis constructed following convention I.

2. Hamiltonian in localized and Bloch basis sets

We begin with some properties and concepts of the matrix elements of Hamiltonian in localized basis. The most important property is translational symmetry

$$\langle\phi_{\boldsymbol{R}i}|\hat{H}|\phi_{\boldsymbol{R}'j}\rangle=\langle\phi_{\boldsymbol{0}i}|\hat{H}|\phi_{\boldsymbol{R}'-\boldsymbol{R}j}\rangle\equiv\langle\phi_{\boldsymbol{0}i}|\hat{H}|\phi_{\Delta\boldsymbol{R}j}\rangle \tag{10}$$

which means we only need take the matrix elements between $(0,0,0)$ and other cells into consideration. When eqn. (10) is used in tandem with the conjugate relation, it yields

$$\langle\phi_{\boldsymbol{0}i}|\hat{H}|\phi_{\Delta\boldsymbol{R}j}\rangle=\langle\phi_{\Delta\boldsymbol{R}j}|\hat{H}|\phi_{\boldsymbol{0}i}\rangle^*=\langle\phi_{\boldsymbol{0}j}|\hat{H}|\phi_{-\Delta\boldsymbol{R}i}\rangle^* \tag{11}$$

indicating that we can reduce the number of matrix elements between $(0,0,0)$ and other cells by a half. We further define the matrix elements satisfying $\Delta\boldsymbol{R}=\boldsymbol{0}$ and $i=j$ as *on-site energies*, and other elements as *hopping terms*.

The matrix elements of Hamiltonian in Bloch basis of convention II follow

$$\begin{aligned}\langle\chi_{\boldsymbol{k}i}|\hat{H}|\chi_{\boldsymbol{k}'j}\rangle&=\frac{1}{N}\left(\sum_{\boldsymbol{R}}\mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{R}}\langle\phi_{\boldsymbol{R}i}|\right)\hat{H}\left(\sum_{\boldsymbol{R}'}\mathrm{e}^{\mathrm{i}\boldsymbol{k}'\cdot\boldsymbol{R}'}|\phi_{\boldsymbol{R}'j}\rangle\right)\\&=\frac{1}{N}\sum_{\boldsymbol{R}}\sum_{\boldsymbol{R}'}\mathrm{e}^{\mathrm{i}(\boldsymbol{k}'\cdot\boldsymbol{R}'-\boldsymbol{k}\cdot\boldsymbol{R})}\langle\phi_{\boldsymbol{R}i}|\hat{H}|\phi_{\boldsymbol{R}'j}\rangle\\&=\frac{1}{N}\sum_{\boldsymbol{R}}\sum_{\Delta\boldsymbol{R}}\mathrm{e}^{\mathrm{i}(\boldsymbol{k}'\cdot\boldsymbol{R}+\boldsymbol{k}'\cdot\Delta\boldsymbol{R}-\boldsymbol{k}\cdot\boldsymbol{R})}\langle\phi_{\boldsymbol{R}i}|\hat{H}|\phi_{\boldsymbol{R}+\Delta\boldsymbol{R}j}\rangle\\&=\frac{1}{N}\sum_{\boldsymbol{R}}\mathrm{e}^{\mathrm{i}(\boldsymbol{k}'-\boldsymbol{k})\cdot\boldsymbol{R}}\sum_{\Delta\boldsymbol{R}}\mathrm{e}^{\mathrm{i}\boldsymbol{k}'\cdot\Delta\boldsymbol{R}}\langle\phi_{\boldsymbol{0}i}|\hat{H}|\phi_{\Delta\boldsymbol{R}j}\rangle\\&=\delta_{\boldsymbol{k}\boldsymbol{k}'}\sum_{\Delta\boldsymbol{R}}\mathrm{e}^{\mathrm{i}\boldsymbol{k}'\cdot\Delta\boldsymbol{R}}\langle\phi_{\boldsymbol{0}i}|\hat{H}|\phi_{\Delta\boldsymbol{R}j}\rangle\end{aligned} \tag{12}$$

And similarly for convention I

$$\langle\chi_{\boldsymbol{k}i}|\hat{H}|\chi_{\boldsymbol{k}'j}\rangle=\delta_{\boldsymbol{k}\boldsymbol{k}'}\sum_{\Delta\boldsymbol{R}}\mathrm{e}^{\mathrm{i}\boldsymbol{k}'\cdot(\Delta\boldsymbol{R}+\Delta\tau)}\langle\phi_{\boldsymbol{0}i}|\hat{H}|\phi_{\Delta\boldsymbol{R}j}\rangle,\ \Delta\tau=\tau_j-\tau_i \tag{13}$$

The meaning of eqn. (12) and (13) are clear: the Hamiltonian is block-diagonal in Bloch basis. This allows us to solve the eigenvalue problem for each k-point individually, with the Hamiltonian defined as

$$H_{ij}(\boldsymbol{k})=\sum_{\Delta\boldsymbol{R}}\mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot(\Delta\boldsymbol{R}+\Delta\tau)}\langle\phi_{\boldsymbol{0}i}|\hat{H}|\phi_{\Delta\boldsymbol{R}j}\rangle \tag{14}$$

for convention I and

$$H_{ij}(\boldsymbol{k})=\sum_{\Delta\boldsymbol{R}}\mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\Delta\boldsymbol{R}}\langle\phi_{\boldsymbol{0}i}|\hat{H}|\phi_{\Delta\boldsymbol{R}j}\rangle \tag{15}$$

in convention II. Both Hamiltonian yield the same eigenvalues, while the eigenvectors satisfy

$$c_{n\boldsymbol{k}}^{\mathrm{II}}(j)=\mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\tau_j}c_{n\boldsymbol{k}}^{\mathrm{I}}(j) \tag{16}$$

where n and j are indices of the eigenvectors and coefficients, respectively.