Band Theory

Energy band theory is based on single-electron approximation. It provides an important framework to understand electronic structures of solids. We shall introduce some of major methods for calculating bands

Nearly free electrons

When the potential U is weak, a reasonable starting point is to expand $\psi_{\vec k}(\vec r)$ in terms of plane waves:

$$\psi_{\vec k}(\vec r) = \sum_{G} \underset{\overset{\uparrow}{\langle \vec k-\vec G | \psi_k \rangle}}{C_{\vec k-\vec G}}\, e^{i(\vec k-\vec G)\cdot \vec r} \quad \text{- - -} ①$$

$\psi_{\vec k-\vec G}$ satisfies

$$\Big(\underbrace{\frac{\hbar^2(\vec k-\vec G)^2}{2m}}_{\varepsilon^0_{k-G}} - \varepsilon\Big) C_{\vec k-\vec G} + \sum_{G'} \underset{\overset{\uparrow}{\langle k-G|U|k-G'\rangle}}{U_{G'-G}}\, C_{k-G'} = 0 \quad L \text{- -} ②$$

Eqs. ① & ② are generally valid not just for weak U.

For weak U, one can expand ψ_k in terms of of U.

Note that in the presence of many atoms in a basis,

$$U(r) = \sum_{R} \sum_{j} \phi_j(r-R-d_j)$$

$$U_G = \frac{1}{v}\int_{cell} d^3\vec{r}\, e^{-i\vec{G}\cdot\vec{r}} \sum_{\vec{R},j} \phi_j(\vec{r}-\vec{R}-\vec{d}_j)$$

↗
volume
of unit
cell

$$= \frac{1}{v}\int_{\text{all space}} d^3\vec{r}\, e^{-i\vec{G}\cdot\vec{r}} \sum_j \phi_j(\vec{r}-\vec{d}_j)$$

$$= \sum_j \frac{1}{v}\int_{\text{all space}} d^3\vec{r}\, e^{-i\vec{G}\cdot\vec{r}}\, \phi_j(\vec{r})\, e^{-i\vec{G}\cdot\vec{d}_j}$$

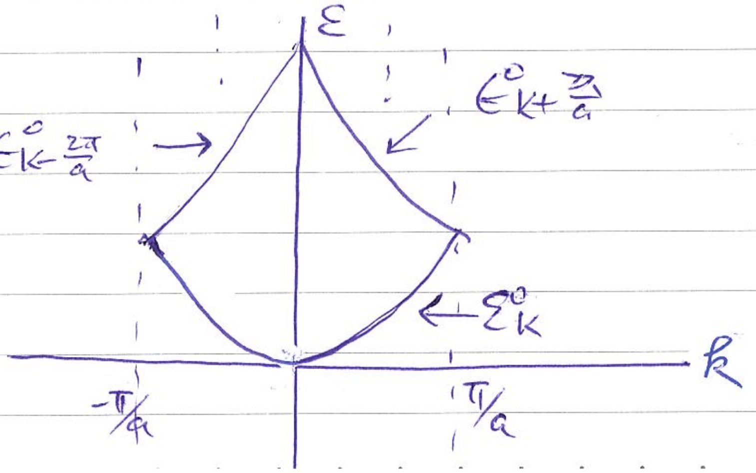
$$= \frac{1}{v}\sum_j \phi_j(G)\, e^{-i\vec{G}\cdot\vec{d}_j}$$

For identical atoms, $U_G = \frac{1}{v}\phi(G) S_G^*$

$S_G = \sum_j e^{i\vec{G}\cdot\vec{d}_j}$ = structure factor.

For $U=0$, eq ② implies $\varepsilon = \varepsilon^0_{k-G}$. This is the case of free electrons.

In the reduce zone scheme (1D as an example)



Therefore, for each k in 1st Brillouin zone,

one can relabel $\varepsilon^0_{k+2\pi/a} \to \varepsilon^0_{n=1}$, $\varepsilon^0_{k+2\times\frac{2\pi}{a}} \to \varepsilon^0_{n=2}, \cdots$

$\varepsilon^0_{k} \to \varepsilon^0_{n=0}$, $\varepsilon^0_{k-\frac{2\pi}{a}} \to \varepsilon^0_{n=-1}, \cdots$

Eq ① $\Rightarrow$ $\overset{|n\rangle}{\overset{\|}{|\psi_k\rangle}} = \sum_n C_n |n^0\rangle \qquad \langle\vec r|n^0\rangle \Leftrightarrow e^{i(\vec k-\vec G)\cdot\vec r}$

Schrödinger eq. $\qquad (\underbrace{\frac{\hat p^2}{2m}}_{H_0} + \hat U)|n\rangle = \varepsilon_n |n\rangle$

For weak $\hat U$, we expand

$$|n\rangle = \underset{\overset{\uparrow}{O(U^0)}}{|n^0\rangle} + \underset{\overset{\uparrow}{O(U^1)}}{|n^1\rangle} + \underset{\overset{\uparrow}{O(U^2)}}{|n^2\rangle} + |n^3\rangle + \cdots$$

$$\boxed{\begin{array}{c}\text{Normalization}\\ \langle n^0|n\rangle = 1\\ \langle n^0|n^k\rangle = 0\end{array}}$$

$$\varepsilon_n = \varepsilon_n^0 + \varepsilon_n^1 + \varepsilon_n^2 + \cdots$$

Formally, one introduces λ, $\hat U \to \lambda \hat U$

and $\qquad |n^k\rangle \to \lambda^k |n^k\rangle$

$\qquad \varepsilon_n^k \to \lambda^k \varepsilon_n^k$

This is the standard perturbation theory in quantum mechancics. One has

$$(H_0 + \lambda U)(|n^0\rangle + \lambda|n^1\rangle + \lambda^2|n^2\rangle + \cdots)$$

$$= (\varepsilon_n^0 + \lambda\varepsilon_n^1 + \lambda^2\varepsilon_n^2 + \cdots)(|n^0\rangle + \lambda|n^1\rangle + \lambda^2|n^2\rangle + \cdots)$$

$O(\lambda^0)$: $\quad (H_0 - \varepsilon_n^0)|n^0\rangle = 0 \quad$ --- ③

$O(\lambda)$: $\quad (H_0 - \varepsilon_n^0)|n^1\rangle + (U - \varepsilon_n^1)|n^0\rangle = 0 \quad$ --- (4)

$O(\lambda^2)$: $(H-\mathcal{E}_n^0)|n^2\rangle + (U-\mathcal{E}_n^1)|n^1\rangle - \mathcal{E}_n^2|n^0\rangle = 0$ -- ⑤

⋮

Using $\langle n^0|n^k\rangle = 0$, we find

$\langle n^0|$④ $\Rightarrow$ $\mathcal{E}_n^1 = \langle n^0|U|n^0\rangle$ -- ⑥

$\langle n^0|$⑤ $\Rightarrow$ $\mathcal{E}_n^2 = \langle n^0|U|n^1\rangle$ --- ⑦

⋮

$\mathcal{E}_K^K = \langle n^0|U|n^K\rangle$

From $(H_0+\hat{U})|n\rangle = \mathcal{E}_n|n\rangle$

$$(\mathcal{E}_n^0-H_0)|n\rangle = (\mathcal{E}_n^0-\mathcal{E}_n)|n\rangle + \hat{U}|n\rangle = (\hat{U}-\underset{\mathcal{E}_n-\mathcal{E}_n^0}{\underset{\uparrow}{\Delta_n}})|n\rangle$$

Projecting $|n^0\rangle$ out, one can inverse $\mathcal{E}_n^0-H_0$,

One gets $|n\rangle = |n^0\rangle + \sum_{m\neq n}\frac{1}{\mathcal{E}_n^0-H_0}|m^0\rangle\langle m^0|\hat{U}-\Delta_n|n\rangle$

$$= |n^0\rangle + \sum_{m\neq n}\frac{1}{\mathcal{E}_n^0-\mathcal{E}_m^0}|m^0\rangle\langle m^0|\underbrace{\lambda\hat{U}-\Delta_n}_{\begin{array}{c}\downarrow\\ \lambda\hat{U}-(\lambda\mathcal{E}_n^1+\lambda^2\mathcal{E}_n^2+\cdots)\end{array}}|n\rangle$$

$$|n^1\rangle = \sum_{m\neq n}\frac{1}{\mathcal{E}_n^0-\mathcal{E}_m^0}|m^0\rangle\langle m^0|U-\mathcal{E}_n^1|n^0\rangle$$

$\langle m^0|n^0\rangle = 0$

$\therefore = \sum_{m\neq n}\frac{1}{\mathcal{E}_n^0-\mathcal{E}_m^0}|m^0\rangle\langle m^0|U|n^0\rangle$ substituting into ⑥

$$\Rightarrow \mathcal{E}_n = \mathcal{E}_n^0 + \langle n^0|U|n^0\rangle + \sum_{m\neq n}\frac{|U_{mn}|^2}{\mathcal{E}_n^0-\mathcal{E}_m^0} + \cdots \quad ⑧$$

$U_{mn} = \langle m^0|U|n^0\rangle$

$$|n\rangle = |n^0\rangle + \sum_{m\neq n} \frac{U_{mn}}{\varepsilon_n^0 - \varepsilon_m^0} |m^0\rangle \quad \text{---} \; ⑨$$

In terms of $\vec{k}$ & G, we find

(shift energy so that $\langle n^0|U|n^0\rangle = 0$)

$$⑧ \Rightarrow \underset{\substack{\uparrow \\ \varepsilon_{k-G}}}{\varepsilon} = \varepsilon^0_{k-G} + \sum_{G'\neq G} \frac{|U_{G-G'}|^2}{\varepsilon^0_{k-G} - \varepsilon^0_{k-G'}} + \dots \qquad ⑩$$

$$\psi_{\vec{k}}(\vec{r}) = e^{i(\vec{k}-\vec{G})\cdot\vec{r}} + \sum_{G'\neq G} \frac{U_{G-G'}}{\varepsilon^0_{k-G} - \varepsilon^0_{k-G'}} e^{i(\vec{k}-\vec{G}')\cdot\vec{r}} + \dots \qquad ⑪$$

Eqs. ⑩ & ⑪ give the energy & wavefunction of energy bands.

<u>Degenerate Perturbation Theory</u> & <u>origin of gap</u>

From eqs ⑩ & ⑪, it is clear that if $\varepsilon^0_{k-G'} = \varepsilon^0_{k-G}$ (degeneracy) for some G', eqs ⑩ & ⑪ break down and are not valid.

In fact, as we have seen, for higher dimensions, degeneracies happen at symmetry points and there may be <u>more than one</u>

In the presence of degeneracy,

$H = H_0 + U$ takes the following form:

$$H = H_0 + U$$

$$= \left(\begin{array}{ccc|cc} \varepsilon_1 & & & & \\ & \varepsilon_2 & & 0 & \\ & & \ddots & & \\ \hline & 0 & & \varepsilon & \\ & & & & \ddots \, \varepsilon \end{array}\right) + \left(\begin{array}{c|c} x & x \\ \hline x & A \end{array}\right)$$

The standard way to deal with degeneracy is to rewrite U as

$$U = \begin{pmatrix} 0 & 0 \\ 0 & A \end{pmatrix} + \underbrace{\begin{pmatrix} x & x \\ x & 0 \end{pmatrix}}_{U'}$$

and combine H_0 with $\begin{pmatrix} 0 & 0 \\ 0 & A \end{pmatrix}$ as a new starting $H_0' = H_0 + \begin{pmatrix} 0 & 0 \\ 0 & A \end{pmatrix}$.

Then the first task is to diagonalized H_0', i.e., $\begin{pmatrix} \varepsilon & & 0 \\ & \ddots & \\ 0 & & \varepsilon \end{pmatrix} + A$. to obtain

$H_0' = \begin{pmatrix} \varepsilon_1 & & & \\ & \ddots & & 0 \\ & & \varepsilon_n & \\ 0 & & & \varepsilon_{n+1}' \end{pmatrix}$ and then treat U' as a perturbation.

The simplest situation is when two free electron levels $\varepsilon^0_{k-G_1}$ & $\varepsilon^0_{k-G_2}$ are degenerate.

In this case, we need to diagonalize the following:

$$(\varepsilon - \varepsilon^0_{K-G_1})\, C_{K-G_1} = U_{G_2-G_1}\, C_{K-G_2}$$

$$(\varepsilon - \varepsilon^0_{K-G_2})\, C_{K-G_2} = U_{G_1-G_2}\, C_{K-G_1}$$

If we set $q = K - G_1$, $G = G_2 - G_1$, we get

$$(\varepsilon - \varepsilon^0_q)\, C_q = U_G\, C_{q-G} \quad ⑫$$

$$(\varepsilon - \varepsilon^0_{q-G})\, C_{q-G} = U_{-G}\, C_q = U^*_G\, C_q \quad ⑬$$

Therefore,
$$\begin{vmatrix} \varepsilon^0_q - \varepsilon & U_G \\ U^*_G & \varepsilon^0_{q-G} - \varepsilon \end{vmatrix} = 0$$

$$(\varepsilon - \varepsilon^0_q)(\varepsilon - \varepsilon^0_{q-G}) = |U_G|^2$$

$$\varepsilon = \frac{1}{2}(\varepsilon^0_q + \varepsilon^0_{q-G}) \pm \left[\left(\frac{\varepsilon^0_q - \varepsilon^0_{q-G}}{2} \right)^2 + |U_G|^2 \right]^{1/2}$$

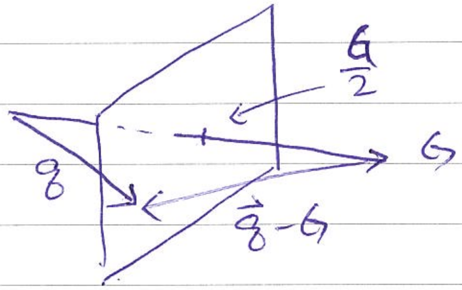
L - ⑭

For $\varepsilon^0_q = \varepsilon^0_{q-G}$ -- $\varepsilon = \varepsilon^0_q \pm |U_G|$

A gap, $\varepsilon_g = 2|U_G|$, arises !

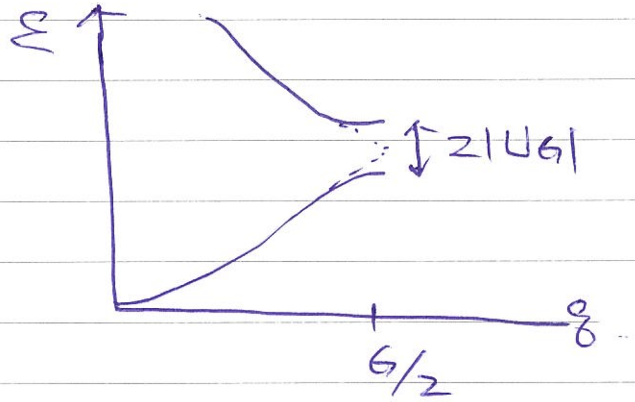
The condition $\varepsilon^0_q = \varepsilon^0_{q-G}$ requires $|\vec{q}| = |\vec{q} - \vec{G}|$.

Therefore, $\vec{q}$ falls on the Bragg plane.



It means that the electron wave undergoes the Bragg reflection near the gap!

For $\underline{\vec{q} \parallel \vec{G}}$, one gets



When $|\vec{q}| = |\vec{q}-\vec{G}|$, $\mathcal{E}_q^0 = \mathcal{E}_{q-G}^0$, $\mathcal{E}-\mathcal{E}_q^0 = \pm|U_G|$,

eqs ⑫ & ⑬ become

$$\pm|U_G|\,C_q = U_G\,C_{q-G}$$
$$\pm|U_G|\,C_{q-G} = U_G^*\,C_q$$

$$\therefore\ C_q = \pm e^{i\phi}\,C_{q-G}.\quad (\text{Here } e^{i\phi}|U_G| = U_G)$$

When the crystal has inversion symmetry,

$^{i\vec{G}\cdot\vec{r}}$

$$U_G^* = \int dr\, U(r)\, e^{-i\vec{G}\cdot\vec{r}}$$

$$= \int d\vec{r}\, U(-\vec{r})\, e^{i\vec{G}\cdot\vec{r}} = \int dr\, U(\vec{r})\, e^{i\vec{G}\cdot\vec{r}} = U_G$$

U_G = real, $\phi = 0$ or π

$C_q = \pm C_{q-G}$, for $q = G/2$

$$\therefore \psi(\vec{r}) = C_{G/2}\, e^{i\frac{\vec{G}}{2}\cdot\vec{r}} + C_{-G/2}\, e^{-i\frac{\vec{G}}{2}\cdot\vec{r}}$$

$$\propto \cos\frac{\vec{G}}{2}\cdot\vec{r}$$

or $\sin\frac{\vec{G}}{2}\cdot\vec{r}$

For $U_G > 0$

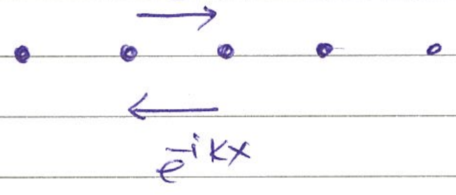
$|\psi(\vec{r})|^2 \propto \cos^2\frac{1}{2}\vec{G}\cdot\vec{r}$ $\quad \varepsilon = \varepsilon_q^0 + |U_G|$

$\propto \sin^2\frac{1}{2}\vec{G}\cdot\vec{r}$ $\quad \varepsilon = \varepsilon_q^0 - |U_G|$

$U_G < 0$ $\quad |\psi(\vec{r})|^2 \propto \sin^2\frac{\vec{G}\cdot\vec{r}}{2}$ $\quad \varepsilon = \varepsilon_q^0 + |U_G|$

$|\psi(\vec{r})|^2 \propto \cos^2\frac{\vec{G}\cdot\vec{r}}{2}$ $\quad \varepsilon = \varepsilon_q^0 - |U_G|$

Physically, consider a 1D lattice, electron's wave e^{ikx} will generally be reflected as e^{-ikx}



When $k = G/2$ $(= \frac{\pi}{a}$ for instance$)$, both

$\pm ikx$ are energetically favored !

$\therefore$ we can form standing waves from $e^{\pm i\frac{\pi}{a}x}$:

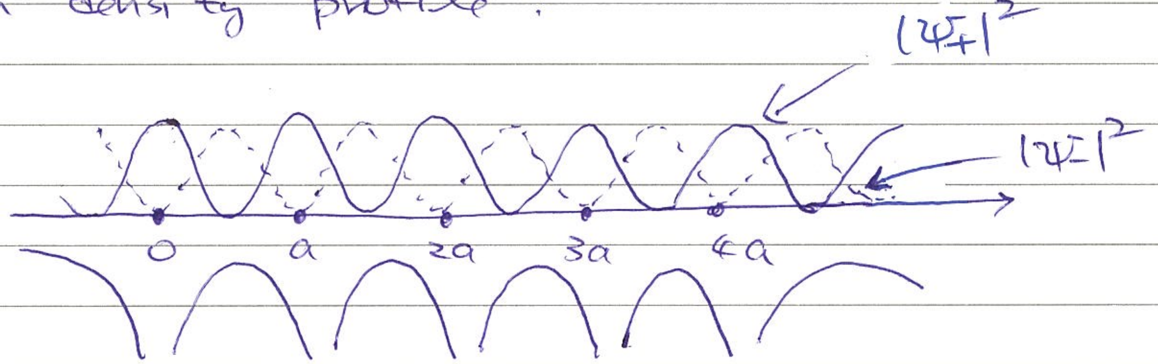
$$\psi_+ = \left(e^{i\frac{\pi}{a}x} + e^{-i\frac{\pi}{a}x}\right) e^{i\frac{E_+}{\hbar}t}$$

$$= 2\cos\frac{\pi x}{a}\, e^{i\frac{E_+}{\hbar}t}$$

$$\psi_- = \left(e^{i\frac{\pi}{a}x} - e^{-i\frac{\pi}{a}x}\right) e^{i\frac{E_-}{\hbar}t}$$

$$= 2i\sin\frac{\pi x}{a}\, e^{i\frac{E_-}{\hbar}t}$$

with density profile:



It's clear that for U<0, $|\psi_+|^2$ is mostly in large negative part of U, $\therefore$ E_+ is smaller!

Their difference in energy

$$E_g \equiv E_+ - E_- \approx \frac{\int_0^a dx\, U(x)\left[|\psi_+|^2 - |\psi_-|^2\right]}{\bar{N}} \leftarrow \text{(normalization)}$$

$$= 4\int_0^a dx\, U(x)\, \frac{1}{\bar{N}}\left[\frac{1+\cos\frac{2\pi}{a}x}{2} - \frac{1-\cos\frac{2\pi}{a}x}{2}\right]$$

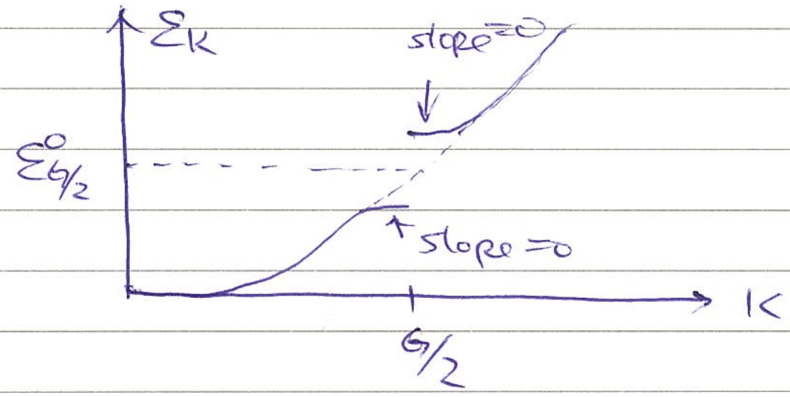
$$= \frac{2}{\bar{N}}\int_0^a dx\, U(x)\left[e^{i\frac{2\pi}{a}x} + e^{-i\frac{2\pi}{a}x}\right]$$

$$= U_G + U_{-G} = 2U_G \qquad G = 2\pi/a$$

$$\bar{N} = \int_0^a dx\, |\psi_+|^2 \quad \int_0^a d\ |\psi_-|^2 = 2a$$

Since, electron's waves are standing waves near the gap, their group velocity $v_g = \frac{\partial \mathcal{E}}{\partial k} = 0$

Generally, one gets



Furthermore, for Coulomb interactions,

$$U(r) \sim \frac{1}{r}, \quad U_G \sim \frac{1}{G^2}.$$

$\mathcal{E}_g$ shrinks as G increases!

Energy bands of common semiconductors

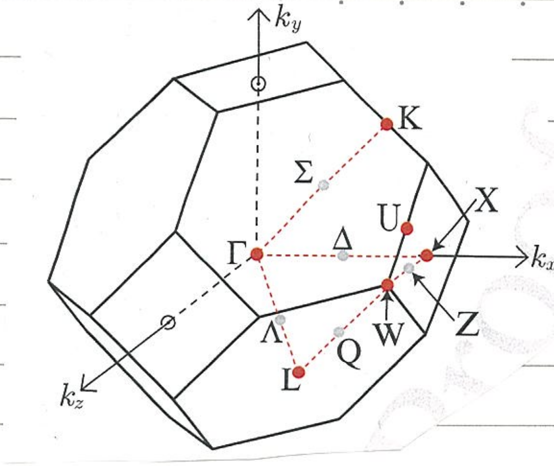
Many common semiconductors are of zincblende structure and electrons are nearly free.

Here degeneracies of free electrons don't just happen at zone boundary and can happen at Γ. That is what actually happens as we have

Zincblende is a FCC structure with primitive translational vectors

$\vec{a}_1 = (\frac{a}{2}, \frac{a}{2}, 0)$, $\vec{a}_2 = (\frac{a}{2}, 0, \frac{a}{2})$

$\vec{a}_3 = (0, \frac{a}{2}, \frac{a}{2})$



1st Brillouin zone of FCC

Its 1st Brillouin zone is shown in the left figure

The primitive vectors of the reciprocal lattice are

$\vec{g}_1 = \frac{2\pi}{a}(-1, -1, 1)$, $\vec{g}_2 = \frac{2\pi}{a}(-1, 1, -1)$, $\vec{g}_3 = \frac{2\pi}{a}(1, -1, -1)$

8 shortest reciprocal lattice vectors : $\frac{2\pi}{a}(\pm1, \pm1, \pm1)$

$|G| = \frac{2\pi}{a}\sqrt{3}$

6 next shortest " " $\frac{2\pi}{a}(\pm2, 0, 0)$ $\frac{2\pi}{a}(0, 0, \pm2)$,

$\frac{2\pi}{a}(0, \pm2, 0)$ $|G| = \frac{4\pi}{a}$

1st Brillouin zone = volume enclosed by planes which are perpendicular bisectors of these 14 reciprocal lattice vectors.

Along (1 00) → Δ along (111) → Λ (line)

along (110) → Σ

$\Gamma = (0,0,0)$ $L = \frac{2\pi}{a}(\frac{1}{2}, \frac{1}{2}, \frac{1}{2})$, $K = \frac{2\pi}{a}(\frac{3}{4}, \frac{3}{4}, 0)$, $W = \frac{2\pi}{a}(\frac{1}{2}, 0, 0)$

At Γ point, $\mathcal{E} = \frac{\hbar^2}{2m}(\vec{k}+\vec{G})^2$

$\vec{G} = \frac{2\pi}{a}(l, m, n)$, $\mathcal{E}_\Gamma = \frac{h^2}{2ma^2}(l^2+m^2+n^2)$ $(\vec{k}=0)$

$\mathcal{E}_\Gamma = 0$ $n=l=m=0 \Rightarrow$ single band

$\mathcal{E}_\Gamma = \frac{h^2}{2ma^2} \times 3$, $(l, m, n) = (\pm 1, \pm 1, \pm 1) \Rightarrow \underbrace{8 \text{ bands}}_{\text{degenerate}}$

Under U, the degeneracy of these 8 bands is lifted.

Most semiconductors will have 8 electrons per unit cell.

For example, $Si = [Ne]\, 3s^2 3p^2$,

$Ge = [Ar]\, 3d^{10} 4s^2 4p^2$

$$\underset{\substack{\uparrow \\ [Ar]3d^{10}4s^24p^1}}{GaAs} \longleftarrow [Ar]3d^{10}4s^24p^3$$ (III-V Compounds)

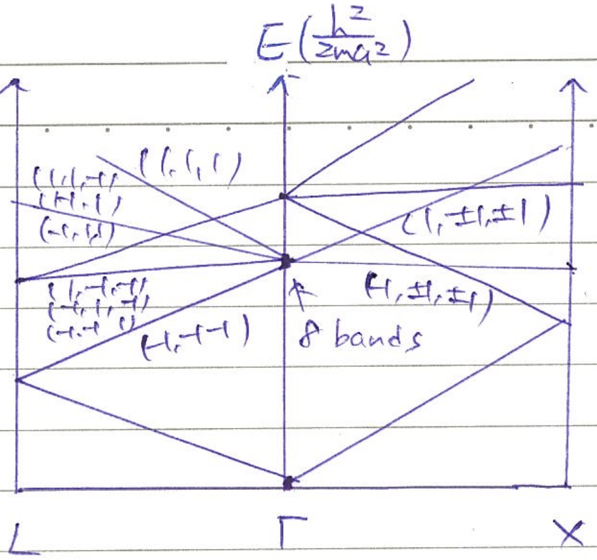
Unit cell: $2Si \Rightarrow 8$ electrons

$2Ge \Rightarrow$ " "

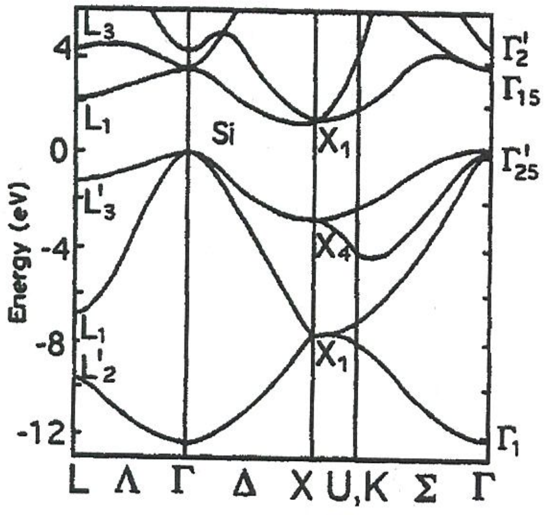
$Ga, As \Rightarrow 8$ "

As one can see from the following figure, the energy bands are similar to free electron bands except for gap opening at Γ.

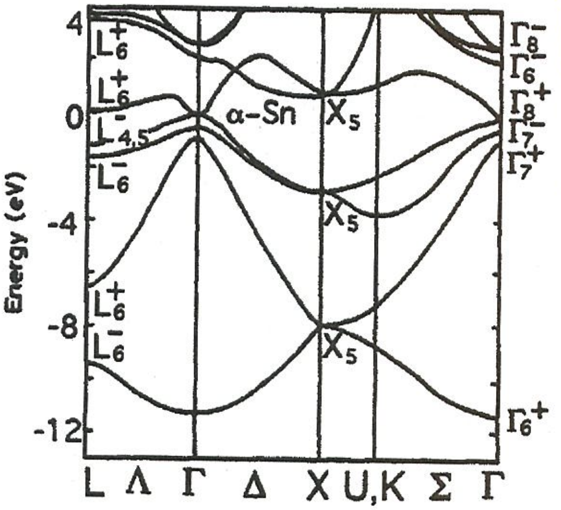
Free electron band



Band structures of typical semiconductors showing similar structures.
(Chelikowsky & Cohen, Phys. Rev. B14, 556, 1976)

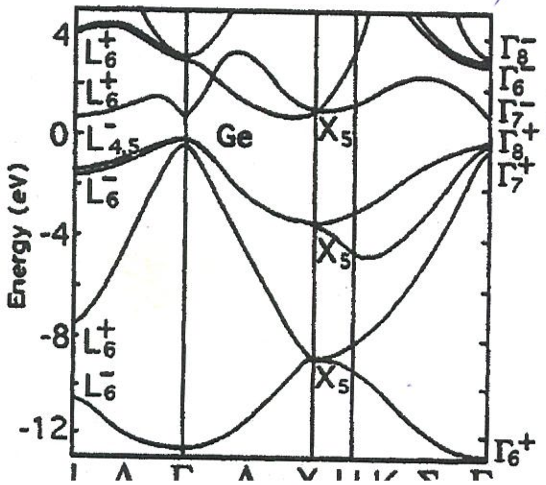


Si (Cubic)

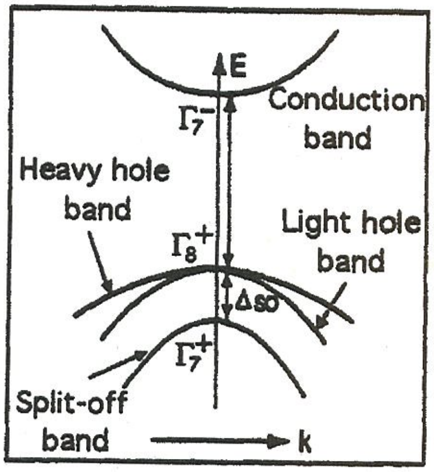


} 8 bands

Ge



} →



No.

Date 10-15.

Brillouin zone and Continuous Fermi surfaces

Just as k space of phonons, we also have many Brillouin zones (B.Z.) for electron.

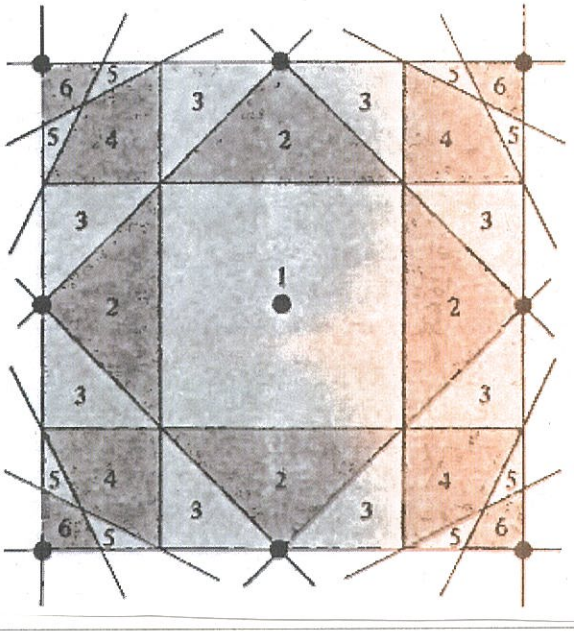
The first B.Z. is $\{\vec{k}\}$ that can be reached from $k=0$ without crossing any Bragg plane.

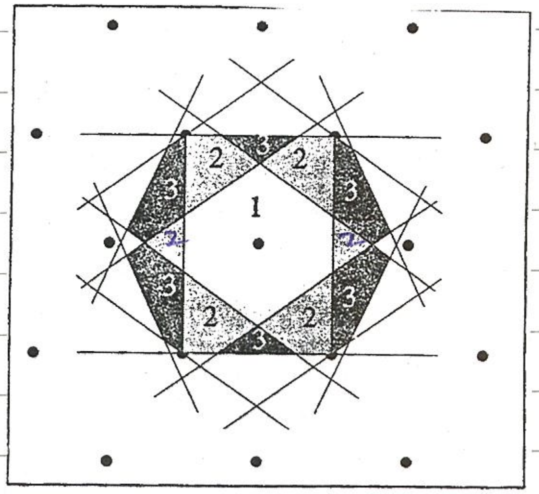
The 2nd B.Z. is $\{\vec{k}\}$ that can be reached from the 1st B.Z. by crossing only one Bragg plane.

⋮

nth B.Z. . . . (n-1) B.Z. . . . one Bragg plane.

Example: Brillouin zones of square lattice & triangle lattice





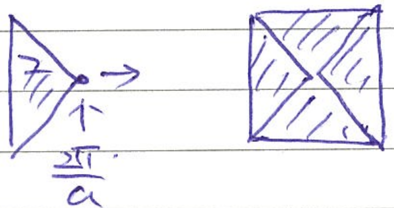
Another way to specify the nth Brillouin zone is : set of points with the origin (K=0) as the nth nearest reciprocal lattice point.

If we translate $\vec{K}=0$ to all possible $\vec{G}$, the nth Brillouin zone will cover all k space:

Given a point $\vec{k}$, there must be only $\vec{G}$ whose relation to $\vec{k}$ is the same as that of $K=0$ to the nth Brillouin zone.

∴ $\vec{k}$ must belong to one of duplicated nth Brillouin zone due to translation.

Example: 2nd Brillouin in square lattice will cover



by translation $\frac{4\pi}{a}\hat{x}$, $\frac{2\pi}{a}\hat{x}+\frac{2\pi}{a}\hat{y}$, $\frac{2\pi}{a}\hat{x}-\frac{2\pi}{a}\hat{y}$

Hence the nth Brillouin zone is a primitive cell with the same volume.

(= N = total # of unit cell in real space)

(# of lattice points)

No.
Date '10-17

Fermi surface

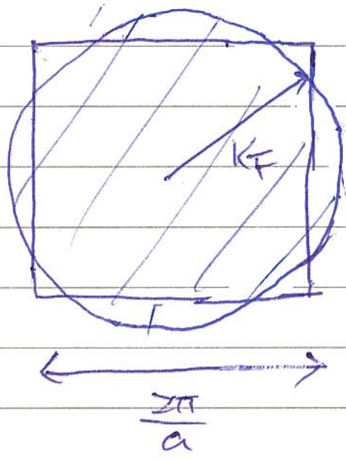
Fermi surface is formed by the k points with energy ε_F.

$2\times$ The total k enclosed by ε_F = total # of electrons

Since energy gaps occur in Bragg planes, it may seem that when Fermi surface crosses the Bragg plane, discontinuity may arise. However, in reality, one can shift sections of Fermi surface appropriately to make the Fermi surface continuous.

As an example, consider the 2D square lattice.

Volume of 1st Brillouin zone $= \left(\frac{2\pi}{a}\right)^2$

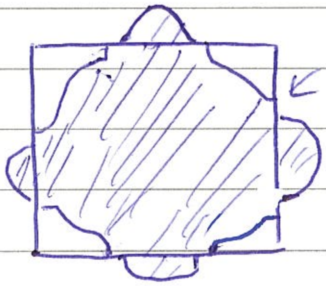


Suppose the lattice has two e's per site.

· then volume enclosed by Fermi energy = volume of 1st Brillouin zone

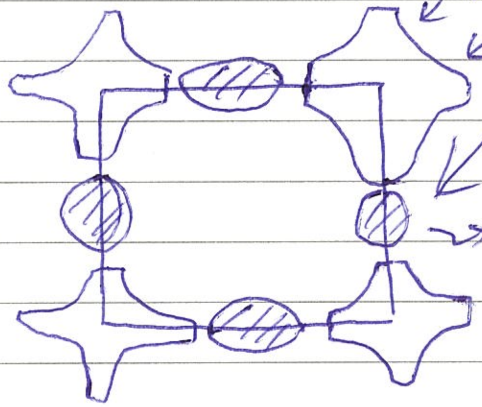
$\pi k_F^2 = 4\pi^2/a^2 \therefore k_F = \frac{2\pi}{\sqrt{\pi}a} = 1.128\,\pi/a$

Under weak U, we have



perpendicular to Bragg plane

It seems to be discontinuous but after appropriate shifting, one get



<u>Tight-binding method</u>

Real crystals are formed by assembling atoms. It is natural to describe bands by starting from atomic wave functions.

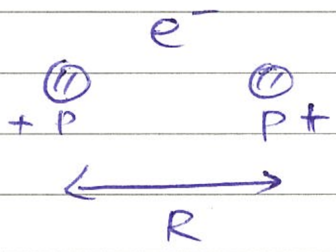
Methods based this view is tight-binding methods.

From this point of view, the energy band

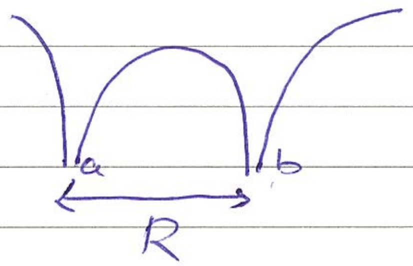
No.
Date 10-19

forms due to linear combinations of atomic orbitals at different sites.

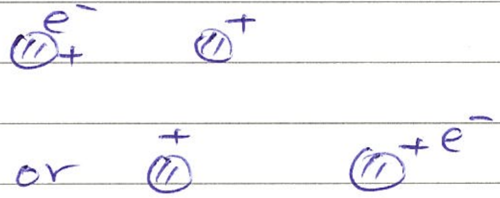
Example : H_2^+



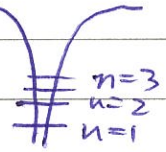
U(r) for e^-



When R = ∞, H_2^+ =



One H-atom



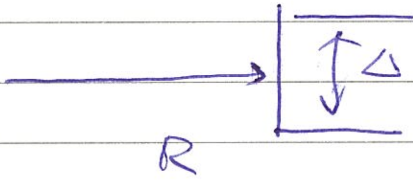
$E_n = -\frac{13.6}{n^2}$ eV

R = finite, $\Psi = a\,\phi_1(r-R_a) + b\,\phi_1(r-R_b)$

minimize $\langle\Psi|H|\Psi\rangle$ w.r.t. a & b

$\because U(-r) = U(r)$ $\therefore$ $\Psi = \phi_1(r-R_a) \pm \phi_1(r-R_b)$

$E_1 = -13.6$ eV
R = ∞

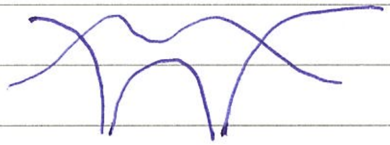


$\Psi \propto \phi_1(r-R_a) - \phi_1(r-R_b)$

$\Psi \propto \phi_1(r-R_a) + \phi_1(r-R_b)$

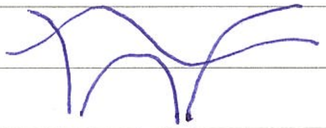
Physically

$a=b$



bonding energy lower

$a=-b$



anti bonding energy higher

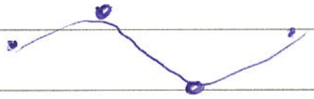
if we represent wavefunctions at each H^+ site by

●, we have

$a=b$

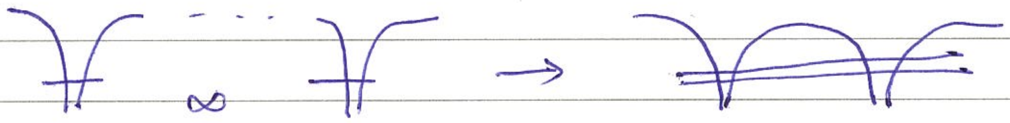
$L=\frac{\lambda}{2}$, large λ, E↓

$a=-b$



$L=\lambda$, small λ, E↑

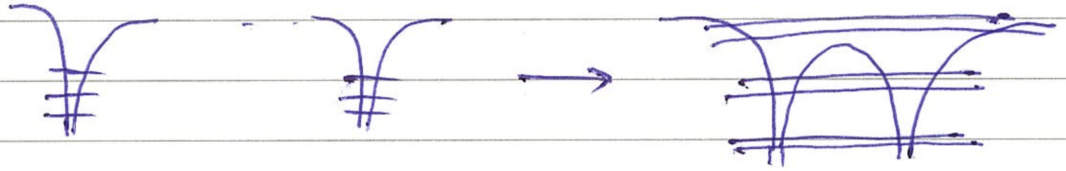
∴



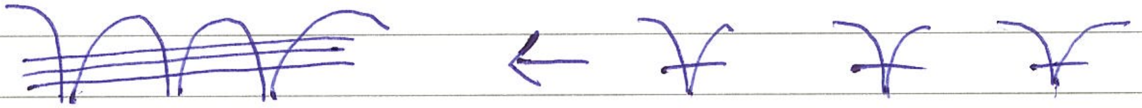
"serial connection" of energy level.

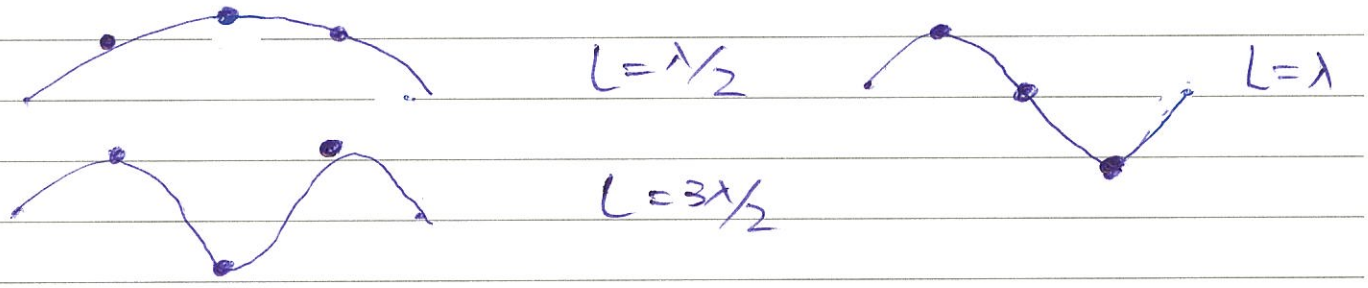
This was also observed in quantum dots.

For other levels, it's similar

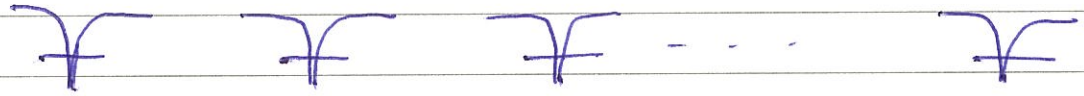


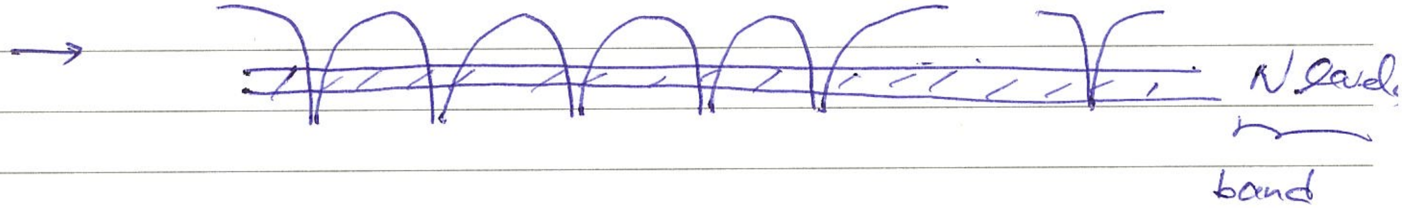
Similarly, for 3 ions, one has



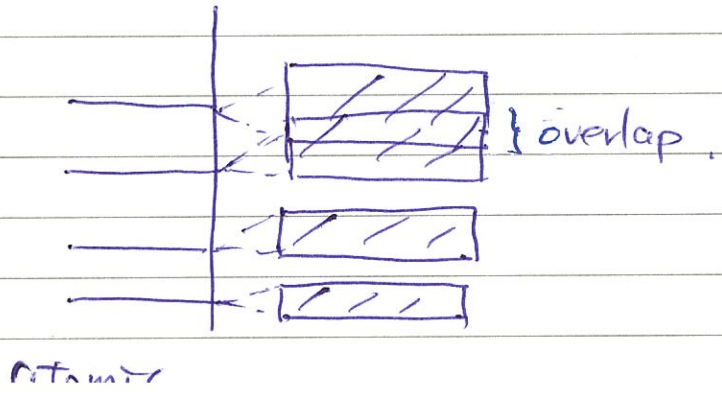


In general, for N atoms, $N >> 1$





Different atomic orbits, they form different bands and may overlap.



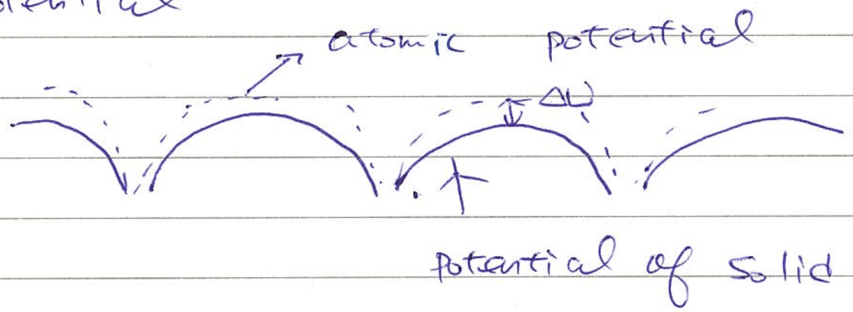
From the above discussions, in general,

we have $$H = H_{atom} + \Delta U(r)$$

and expect

$$\psi_{nk}(\vec{r}) = \sum_{\vec{R}} e^{i\vec{k}\cdot\vec{R}}\,\phi_n(\vec{r}-\vec{R}) \qquad \text{--- ⑮}$$

Here ΔU is the deviation of potential from atomic potential



Form of eq.⑮ satisfies the Bloch theorem

$$\psi_{nk}(\vec{r}+\vec{R}) = \sum_{\vec{R}'} e^{i\vec{k}\cdot\vec{R}'}\,\phi_n(\vec{r}+\vec{R}-\vec{R}')$$

$$= \sum_{\vec{R}''} e^{i\vec{k}\cdot(\vec{R}''+\vec{R})}\,\phi_n(\vec{r}-\vec{R}'') \qquad , \quad \vec{R}''=\vec{R}'-\vec{R}$$

$$= e^{i\vec{k}\cdot\vec{R}}\,\psi_{nk}(\vec{r})$$

Hence it is a valid form of solution to

$$H|\psi\rangle = \varepsilon|\psi\rangle .$$

10-23

Φ_n is an atomic orbital wave function. In general Φ_n's symmetry may not be compatiable with symmetries of crystals. Hence it can be replaced by linear combinations of atomic orbtals. This is then the method of LCAO.

$$\Phi_n \to \sum_n b_n \Phi_n(r-R) \text{ --- } ⑯$$

Wannier function

The wavefunction of an electron in crystals can be generally written in a form similar to eq. ⑮ except that Φ_n is generally not localized to the atom at $\vec{R}$.

For this purpose, one defines

$$W_n(\vec{R},\vec{r}) \equiv \frac{1}{\sqrt{N}} \sum_{\vec{k}} e^{-i\vec{k}\cdot\vec{R}} \psi_{nk}(\vec{r}) \text{ --- } ⑰$$

i.e. W_n is the Fourier transformation on k of ψ_{nk}.

Orthogonality:

$$\int d\vec{r}\, W_n(\vec{R},\vec{r})\, W_n^*(\vec{R}',\vec{r})$$

$$= \frac{1}{N} \sum_{K} \sum_{K'} e^{-i\vec{K}\cdot\vec{R} + i\vec{K}'\cdot\vec{R}'} \int d\vec{r}\, \varphi^*_{nK}(\vec{r})\, \varphi_{mK'}(\vec{r})$$

$$= \frac{1}{N} \sum_{\vec{R},\vec{R}'} e^{-i\vec{K}\cdot\vec{R} + i\vec{K}'\cdot\vec{R}'} \delta_{m,n}\, \delta_{K,K'}$$

$= \delta_{R,R'}\,\delta_{n,m}$ . . . ⑱

Eq. ⑰ can be inverted, one gets

$$\psi_{nk}(\vec r) = \frac{1}{\sqrt N}\sum_{\vec R} e^{i\vec k\cdot\vec R}\, W_n(\vec R,\vec r) \quad \cdots ⑲$$

Furthermore, $\frac{1}{v}\int dk\, e^{-i\vec k\cdot\vec R}$ ⑲ , we get

$$\frac{1}{\sqrt N} W_n(\vec R,\vec r) = \underset{\uparrow}{\frac{1}{v}}\int d\vec k\, e^{-i\vec k\cdot\vec R}\,\psi_{n\vec k}(\vec r) \quad \cdots ⑳$$

Volume of K-space for 1st Brillouin Zone

Hence, if both r & $\vec R$ are shifted by some $\vec R_0$, by using the Bloch theorem, $W_n(\vec R,\vec r)$

Eq⑳ implies

is invariant. $\frac{1}{\sqrt N} W_n(\vec R,\vec r) = \frac{1}{\sqrt N} W_n(\vec r-\vec R)$

$$\psi_{nk}(\vec r) = \frac{1}{\sqrt N}\sum_{\vec R} e^{i\vec k\cdot\vec R}\, W_n(\vec r-\vec R) \quad \cdots ㉑$$

Comparing eq. ㉑ with eq. ⑮, it's clear $\frac{1}{\sqrt N} W_n(\vec r-\vec R)$ plays the role of the atomic orbital $\phi_n(\vec r-\vec R)$!

except that $\int d\vec r\, \phi_n(\vec r-\vec R)\,\phi_n^*(\vec r-\vec R') \neq \delta_{R,R'}$

while $\{W_n\}$ are orthogonal (eq. ⑱)

No.
Date 10-25

In general, $W_n(\vec r-\vec R)$ is not localized as atomic orbital $\phi_n(\vec r-\vec R)$. Its decaying range is generally not easy to obtain. Roughly speak, when energy gap increases, band width decreases, W_n is more localized!

General solutions based on tight-binding method

By using Wannier functions as a basis, we can solve

$$\left(\underbrace{H_{atom}+\Delta U}_{H}\right)\psi(\vec r)=\varepsilon\psi(\vec r) \quad \cdots ㉒$$

This is done by evaluating the matrix element of H, $\langle \vec R'|H|\vec R\rangle \equiv H_{RR'}$

$$\equiv \int d\vec r\, W_n^*(\vec R',\vec r)\,\hat H\, W_n(\vec R,\vec r)$$

Then

$$\hat H=\sum_{R,R'}|\vec R'\rangle\langle\vec R'|H|\vec R\rangle\langle\vec R| \quad \cdots ㉓$$

Here $|\vec R\rangle\equiv W_n(\vec R,\vec r)$

The problem is then to solve the secular eq.

$$\begin{vmatrix} \ddots & \cdots & \cdots \\ H_{R_1R_1}-\varepsilon & H_{R_1R_2} & \cdots \end{vmatrix}=0$$

No.
Date 10-26.

The above is the general scheme by using the Wannier functions. $W_n(\vec{R},\vec{r})$.

If we approximate $W_n(\vec{R},\vec{r})$ by $\phi_n(\vec{r}-\vec{R})$

and consider only one local orbital $\phi(\vec{r}-\vec{R})$,

we have
$$\psi(\vec{r}) = \sum_{\vec{R}} e^{i\vec{k}\cdot\vec{R}}\,\phi(\vec{r}-\vec{R})$$

or simply
$$|\psi\rangle = \sum_{\vec{R}} e^{i\vec{k}\cdot\vec{R}}\,|\phi_{\vec{R}}\rangle$$

To solve
$$(H_{atom} + \Delta U)\,|\psi\rangle = \varepsilon_k\,|\psi\rangle \quad \text{-- ㉖}$$

$\langle\phi_{R=0}|$ ㉖ $\quad \because \quad H_{atom}\,|\phi_{R=0}\rangle = \varepsilon_n\,|\phi_{R=0}\rangle$

$$\Rightarrow \quad (\varepsilon_k - \varepsilon_n)\langle\phi_{R=0}|\psi\rangle = \langle\phi_{R=0}|\Delta U|\psi\rangle$$

$$\text{i.e.}\quad (\varepsilon_k-\varepsilon_n)\int d\vec{r}\,\phi^*(\vec{r})\,\psi(\vec{r}) = \int d\vec{r}\,\phi^*(r)\,\Delta U(r)\,\psi(\vec{r}) \quad \leftarrow \text{-- ㉕}$$

$$\int \phi^*(\vec{r})\,\psi(\vec{r})\,d^3\vec{r} = \int \phi^*(r)\,\phi(\vec{r})\,d^3\vec{r} + \sum_{R\neq 0} e^{i\vec{k}\cdot\vec{R}}\int \phi^*(r)\,\phi(\vec{r}-\vec{R})\,d^3\vec{r}$$

$$\equiv 1 + \sum_{R\neq 0} \alpha(R)\,e^{i\vec{k}\cdot\vec{R}}$$

$$\int \phi^*(\vec{r})\,\Delta U(r)\,\psi(\vec{r}) = \underbrace{\int \phi^*(r)\,\Delta U(\vec{r})\,\phi(\vec{r})}_{\equiv\, -\beta} + \sum_{R\neq 0} e^{i\vec{k}\cdot\vec{R}}\underbrace{\int \phi^*(r)\,\Delta U(r)\,\phi(\vec{r}-\vec{R})}_{\equiv\, -t(R)}$$

No.

Date 10-27

∴ ㉕ ⇒

$$\varepsilon_k = \varepsilon_n - \frac{\beta + \sum_R t(R) e^{i\vec{k}\cdot\vec{R}}}{1 + \sum_{R\neq 0} \alpha(R) e^{i\vec{k}\cdot\vec{R}}}$$

With $\alpha(R) = \int d^3r\, \phi^*(r)\, \phi(r-R)$

$t(R) = -\int d^3r\, \phi^*(r)\, \Delta U(r)\, \phi(r-R)$

$\beta \equiv -\int d^3r\, \phi^*(r)\, \Delta U(r)\, \phi(r)$

$\because \alpha \simeq 0$

$$\therefore \varepsilon_k = \varepsilon_n - \beta - \sum_R t(R)\, e^{i\vec{k}\cdot\vec{R}}$$

If $\underline{\Delta U(-r) = \Delta U(r)}$ (inversion symmetric) & $\underline{\phi = \text{real}}$ (such as s-orbital)

$$t(-R) = -\int d^3r\, \phi^*(r)\, \Delta U(r)\, \phi(r+R)$$

$$= -\int d^3r\, \phi^*(r-R)\, \Delta U(r-R)\, \phi(r)$$

$$= -\int d^3r\, \phi(r)\, \Delta U(r)\, \phi(r-R) = t(R)$$

$$\varepsilon_k = \varepsilon_n - \beta - \sum_R t(R) \cos \vec{k}\cdot\vec{R} \quad \text{---} ㉖$$

Example:

For $\vec{k} \simeq 0$ (Γ point), $\varepsilon_k = \varepsilon_n - \beta - \sum_R t(R)$

(BCC) $\qquad + \frac{1}{2}\left(\sum_R t(R) R_\alpha R_\beta\right) k_\alpha k_\beta$

Since ϕ is localized, one expects

only nearest neighbours contribute.

In this case, $\vec{R} = a(\pm1,0,0),\ a(0,\pm1,0),\ a(0,0,\pm1)$

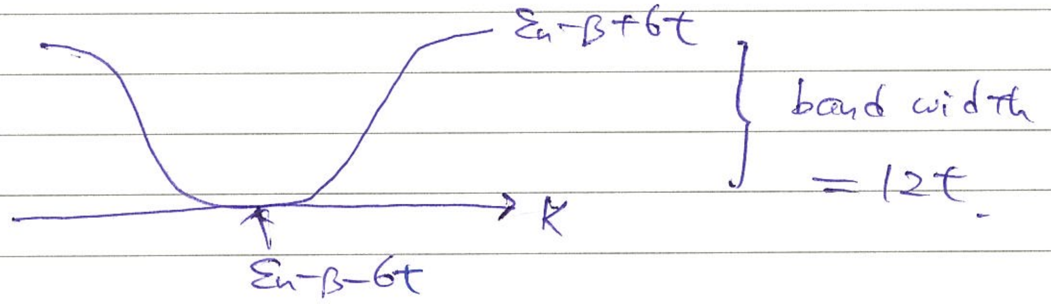
$t(\vec{R}) \equiv t$

$$\varepsilon_k = \varepsilon_n - \beta - 2t(\cos k_x a + \cos k_y a + \cos k_z a)$$

$$= \varepsilon_n - \beta - 6t + ta^2\underbrace{(k_x^2 + k_y^2 + k_z^2)}_{k^2}$$

It's clear that one can identify ta^2k^2 as $\frac{\hbar^2 k^2}{2m_{eff}}$

$$\therefore\ m_{eff} = \frac{\hbar^2}{2ta^2}$$



In general, one has to keep n as what is done in eq. ⑯. This is especially true when there is degeneracy in atomic orbitals. For instance, p-orbitals are degenerate (P_x, P_y, P_z) with ε_p

$$\therefore\ \Phi(\vec{r}) = \sum_m b_m \phi_m(r), \quad m = P_x, P_y, P_z$$

$$|\Psi\rangle = \sum_{m}\sum_{\vec{R}} e^{i\vec{k}\cdot\vec{R}}\, b_m\, |\phi^m_{R}\rangle$$

$$\psi(\vec{r}) = \sum_{m,\vec{R}} e^{i\vec{k}\cdot\vec{R}}\, b_m \phi_m(\vec{r}-\vec{R}) \quad \text{-- ㉗}$$

Eq. ㉕ becomes

$$(\varepsilon_k-\varepsilon_p)\int d^3\vec{r}\, \phi_m^*(r)\,\psi(\vec{r}) = \int d^3\vec{r}\, \phi_m^*(r)\,\Delta U(r)\,\psi(\vec{r})$$

$$\leftarrow \text{-- ㉘}$$

$$\int d^3\vec{r}\, \phi_m^*(r)\,\psi(\vec{r}) = \sum_{m'}\int d^3\vec{r}\, \phi_m^*(r)\phi_{m'}(\vec{r})\cdot b_{m'}$$

$$+\sum_{m'}\left(\sum_{\vec{R}\neq 0} e^{i\vec{k}\cdot\vec{R}}\int d^3\vec{r}\, \phi_m^*(r)\phi_{m'}(\vec{r}-\vec{R})\right)b_{m'}$$

$$= b_m + \sum_{\substack{m' \neq m,\\ \vec{R}\neq 0}} \underbrace{\alpha_{mm'}\, e^{i\vec{k}\cdot\vec{R}}}_{\alpha_{mm'}(k)}\, b_{m'}$$

$$\int \phi_m^*(\vec{r})\,\Delta U(r)\,\psi(\vec{r})\, d^3\vec{r}$$

$$= \sum_{m'}\int \phi_m^*(r)\Delta U(\vec{r})\phi_{m'}(\vec{r})\, b_{m'} + \sum_{\substack{m'\\ \neq m}}\sum_{R\neq 0} e^{i\vec{k}\cdot\vec{R}}\int \phi_m^*(r)\Delta U(r)\phi_{m'}^{(r-R)}$$

$$\times\, b_{m'}$$

$$= -\sum_{m'}\beta_{mm'}\, b_{m'} - \sum_{\substack{R\neq 0\\ m'}} \underbrace{e^{i\vec{k}\cdot\vec{R}}\cdot T^{(R)}_{mm'}}_{T_{mm'}(k)}\, b_{m'}$$

Hence, one gets

$$\left.\begin{aligned}&(\varepsilon_k-\varepsilon_p)\left[b_m + \sum_{m'\neq m}\alpha_{mm'}(k)\, b_{m'}\right]\\ &= -\sum_{m'}\beta_{mm'}b_{m'} - \sum_{m'} t_{mm'}(k)\, b_{m'}\end{aligned}\right\}$$

which becomes a 3x3 matrix

Tight binding in 2nd quantized form

From eq. ㉓, for many electrons, in general one has

$$\hat{H} = \sum_{R,R'} T_{RR'} C^{+}_{R'} C_{R}$$

with $T_{RR'} \equiv \langle R' | H | R \rangle$

$$= \int d^3\vec{r}\; \omega^{*}_{n}(\vec{R}',\vec{r}) \left[-\frac{\hbar^2\nabla^2}{2m} + U(r) \right] \omega_{n}(\vec{R},\vec{r})$$

Note that $T_{R'R} = \langle R | H | R' \rangle = t^{*}_{RR'}$

$$\therefore \hat{H}^{+} = \sum_{RR'} T^{*}_{RR'} C^{+}_{R} C_{R'} = \sum_{RR'} T_{R'R} C^{+}_{R} C_{R'} = \hat{H}$$

In the approximation of $\omega_n(\vec{R},\vec{r})$ by $\phi(\vec{r}-\vec{R})$, (atomic orbital),

$$T_{RR'} = \int d^3\vec{r}\; \phi^{*}(\vec{r}-\vec{R}') \left[H_{atom} + \Delta U \right] \phi(\vec{r}-\vec{R})$$

$$= \int d^3\vec{r}\; \phi^{*}(\vec{r}+\vec{R}-\vec{R}') \left[H_{atom} + \Delta U \right] \phi(\vec{r})$$

$$= \varepsilon_a \int d^3\vec{r}\; \phi^{*}(\vec{r}+\vec{R}-\vec{R}')\, \phi(\vec{r}) + \underbrace{\int d^3\vec{r}\; \phi^{*}(\vec{r}+\vec{R}-\vec{R}') \times \Delta U(\vec{r})\, \phi(\vec{r})}_{\equiv\, -t_{R,R'}}$$

$$\uparrow \quad H_{atom}\, \phi(\vec{r}) = \varepsilon_a\, \phi(\vec{r})$$

$\therefore T_{RR'} = T(R-R')$

$\because \int d^3\vec{r}\; \phi^{*}(\vec{r}+\vec{R})\, \phi(\vec{r}) \cong 0$ for $R \neq 0$.

No.
Date 10-31

$$\therefore\ \hat{H} = \varepsilon_a \sum_{R} c_R^\dagger c_R - \sum_{R\neq R'} \overbrace{t_{RR'}}^{\text{hopping Amplitude}} c_R^\dagger c_{R'}$$

$$= \varepsilon_a \sum_{R} c_R^\dagger c_R - \sum_{\langle R,R'\rangle} t_{R,R'} c_R^\dagger c_{R'} + h.c.$$

Example: 1D tight binding

if only nearest neighbours contribute, one gets

$$\hat{H} = \varepsilon_a \sum_{i,\sigma} c_{i\sigma}^\dagger c_{i\sigma} - \sum_{i\sigma} t\, c_{i\sigma}^\dagger c_{i+1\sigma} + h.c.$$

$(\sigma = \uparrow,\downarrow)$ $(t^* = t = \text{real})$

Energy eigenvalue can be found by Fourier transformation.

$$c_{k\sigma} = \frac{1}{\sqrt{N}} \sum_i c_{i\sigma}\, e^{-ik(ia)}$$

or

$$c_{i\sigma} = \frac{1}{\sqrt{N}} \sum_k c_{k\sigma}\, e^{+ik\cdot(ia)}$$

$$c_{i\sigma}^\dagger = \frac{1}{\sqrt{N}} \sum_k c_{k\sigma}^\dagger\, e^{-ik\cdot(ia)}$$

$$\sum_i c_{i\sigma}^\dagger c_{i+1\sigma} = \sum_{kk'} \sum_i c_{k\sigma}^\dagger c_{k'\sigma} \frac{1}{N} e^{i[-k(ia)+k'(i+1)a]}$$

$$\frac{1}{N} \sum_i e^{i(-k+k')(ia)} = \delta_{k,k'}$$

$$\therefore \sum_i c_{i\sigma}^\dagger c_{i+1\sigma} = \sum_k c_{k\sigma}^\dagger c_{k\sigma}\, e^{+ika}$$

Similarly $\sum_i C_{i+1\sigma}^{+} C_{i\sigma} = \sum_k C_{k\sigma}^{+} C_{k\sigma} e^{-ika}$

$$\sum_i C_{i\sigma}^{+} C_{i\sigma} = \sum_k C_{k\sigma}^{+} C_{k\sigma}$$

$$\therefore \hat{H} = \sum_{k\sigma} \varepsilon_a C_{k\sigma}^{+} C_{k\sigma} - \sum_{k\sigma} (te^{ika} + te^{-ika}) C_{k\sigma}^{+} C_{k\sigma}$$

$$= \sum_{k\sigma} (\varepsilon_k - \varepsilon_a) C_{k\sigma}^{+} C_{k\sigma} \qquad \varepsilon_k = -(te^{ika} + e^{-ika})$$

$$= -2t\cos ka$$

$\therefore$ energy eigenvalues $= -2t\cos ka - \varepsilon_a$.

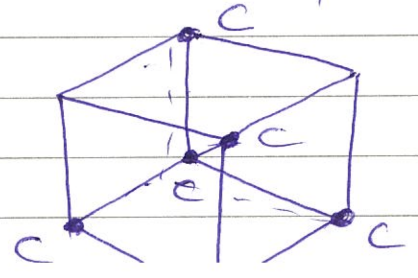
(see over for wave function approach)

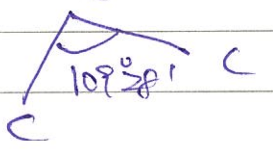
LCAO & graphene/diamond

The linear combination of atomic orbitals is necessary if the symmetry of atomic orbitals is not compatiable with that of crystals!

Important examples include diamond & graphene.

In diamond structure, each carbon is surrounded by 4 carbon atoms as shown in the left figure.

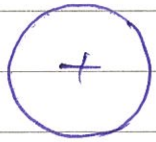




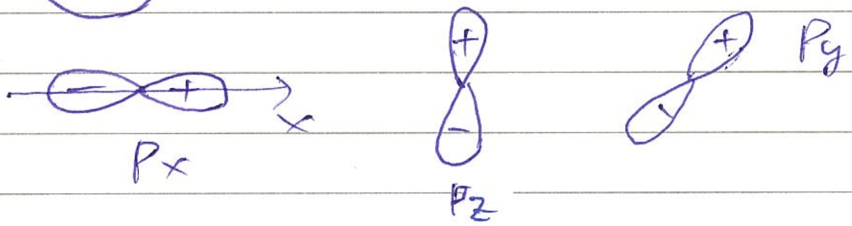
The electronic configuration of a carbon atom is $1s^2 2s^2 2p^2$. To form diamod, each carbon atom contribute 4 electrons.

∴ $\underbrace{2s^2 2p^2}_{\text{electrons in}}$ participate.

s orbital:



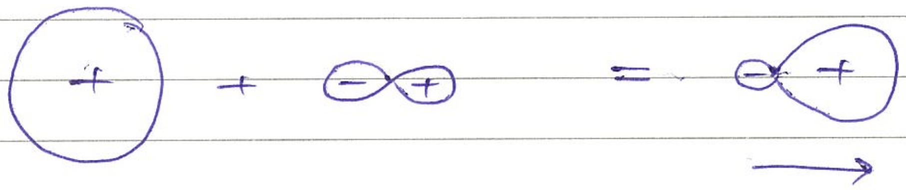
p orbitals



It's clear that pure s or p does not have the symmetry of diamond structure.

For this purpose, one forms linear combination of s, Px, Py & Pz.

It's clear that by combining s & Px, a particular direction can be formed:



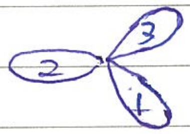
The tetradedron symmetry can be formed by hybridizing s, P_x, P_y & P_z as follows:

$$\begin{aligned} \phi_1 &= \tfrac{1}{2}(S+P_x+P_y+P_z) \\ \phi_2 &= \tfrac{1}{2}(S+P_x-P_y-P_z) \\ \phi_3 &= \tfrac{1}{2}(S-P_x+P_y-P_z) \\ \phi_4 &= \tfrac{1}{2}(S-P_x-P_y+P_z) \end{aligned}$$

these are known as sp3 hybrids

(it's clear, $\langle\phi_i|\phi_j\rangle = \delta_{ij}$)

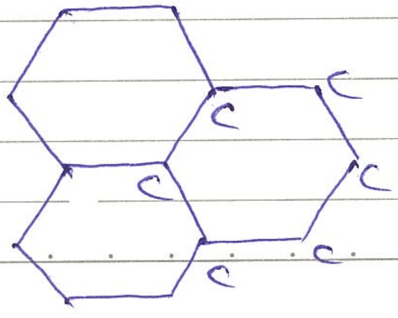
Similarly, for graphene whose lattice is a honeycomb lattice, one forms sp^2 from S, P_x & P_y :

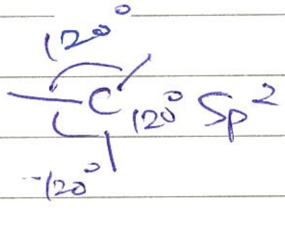


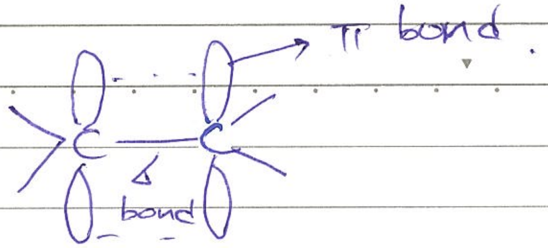
$$\begin{aligned} \phi_1 &= \tfrac{1}{\sqrt{3}} s - \tfrac{1}{\sqrt{6}} P_x - \tfrac{1}{\sqrt{2}} P_y \\ \phi_2 &= \tfrac{1}{\sqrt{3}} s + \tfrac{\sqrt{2}}{\sqrt{3}} P_x \\ \phi_3 &= \tfrac{1}{\sqrt{3}} s - \tfrac{1}{\sqrt{6}} P_x + \tfrac{1}{\sqrt{2}} P_y \end{aligned}$$

these are sp^2 hybrids.

Their connections are know as σ bonds while P_z is left intact!





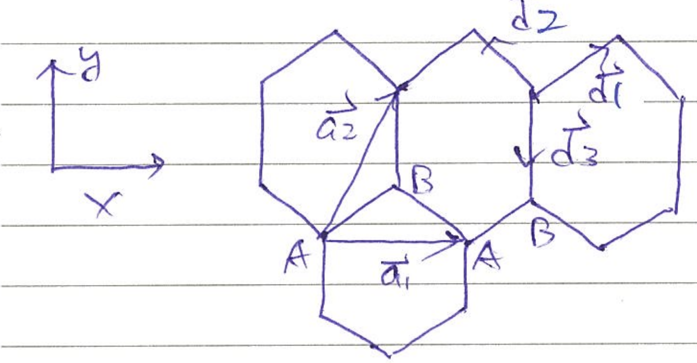


Spectrum of tight-binding model on honeycomb lattice

As an example,

Consider a tight-binding model on honeycomb lattice with hopping between nearest neighbours only. This model is an excellent description of electrons in graphene

$|\vec{a}_1| = |\vec{a}_2| = a$



$$\hat{H} = -t \sum_{\substack{\langle ij \rangle \\ \sigma}} C^{+}_{i\sigma} C_{j\sigma} + h.c.$$

$\langle i,j \rangle$ indicates j is nearest neighbour to i.

$$C^{+}_{i\sigma} = \frac{1}{\sqrt{N}} \sum_{k} C^{+}_{k\sigma} e^{-i\vec{k}\cdot\vec{r}_i}$$

if $i = A$ site, $\quad C_{j\sigma} = \frac{1}{\sqrt{N}} \sum_{k} C^{B}_{k\sigma} e^{+i\vec{k}\cdot\vec{r}_i} e^{i\vec{k}\cdot\vec{d}_m}$

$C_{k\sigma} \to C^{A}_{k\sigma}$

$m = 1, 2, 3$

$$\sum_{\langle ij \rangle} C^{+}_{i\sigma} C_{j\sigma} = \sum_{k\sigma} C^{A+}_{k\sigma} C^{B}_{k\sigma} \left(e^{i\vec{k}\cdot\vec{d}_1} + e^{i\vec{k}\cdot\vec{d}_2} + e^{i\vec{k}\cdot\vec{d}_3} \right)$$

Similarly, if $i=B, \Rightarrow j=A$

$$\sum_{\substack{\langle ij\rangle \\ \sigma}} C_{i\sigma}^{\dagger} C_{j\sigma} = \sum_{k\sigma} C_{k\sigma}^{B\dagger} C_{k\sigma}^{A} \left(e^{-i\vec{k}\cdot\vec{d}_1} + e^{-i\vec{k}\cdot\vec{d}_2} + e^{-i\vec{k}\cdot\vec{d}_3}\right)$$

$$\therefore \hat{H} = \sum_{k\sigma} \Lambda_k C_{k\sigma}^{A\dagger} C_{k\sigma}^{B} + \Lambda_k^{*} C_{k\sigma}^{B\dagger} C_{k\sigma}^{A}$$

$$= \sum_{k\sigma} \underbrace{\left(C_{k\sigma}^{A\dagger}, C_{k}^{B\dagger}\right)}_{\psi_{k\sigma}^{\dagger}} \begin{pmatrix} 0 & \Lambda_k \\ \Lambda_k^{*} & 0 \end{pmatrix} \underbrace{\begin{pmatrix} C_{k\sigma}^{A} \\ C_{k\sigma}^{B} \end{pmatrix}}_{\psi_{k\sigma}}$$

$$\Lambda_k = -t\left(e^{i\vec{k}\cdot\vec{d}_1} + e^{i\vec{k}\cdot\vec{d}_2} + e^{i\vec{k}\cdot\vec{d}_3}\right)$$

$$\vec{d}_1 = \left(\frac{a}{2}, \frac{a}{2\sqrt{3}}\right), \quad \vec{d}_2 = \left(-a/2, \frac{a}{2\sqrt{3}}\right), \quad \vec{d}_3 = \left(0, -\frac{a}{\sqrt{3}}\right)$$

$$\Lambda_k = -t\left(2\cos\frac{ak_x}{2}\, e^{i\frac{ak_y}{2\sqrt{3}}} + e^{-i\frac{ak_y}{\sqrt{3}}}\right)$$

The wave function satisfies

$$\begin{pmatrix} 0 & \Lambda_k \\ \Lambda_k^{*} & 0 \end{pmatrix} \begin{pmatrix} \varphi_k^{A} \\ \varphi_k^{B} \end{pmatrix} = \varepsilon_k \begin{pmatrix} \varphi_k^{A} \\ \varphi_k^{B} \end{pmatrix}$$

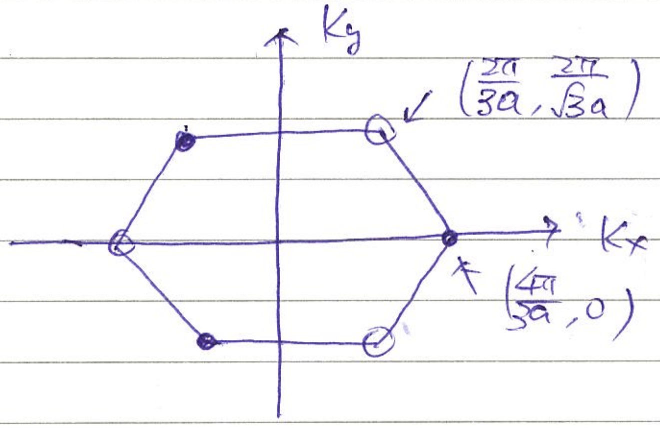
Hence
$$\begin{vmatrix} \varepsilon_k & -\Lambda_k \\ -\Lambda_k^{*} & \varepsilon_k \end{vmatrix} = 0$$

$$\varepsilon_k^2 = |\Lambda_k|^2, \quad \varepsilon_k = \pm \Lambda_k$$

$$= \pm \sqrt{1 + 4\cos^2 k_x a \quad 4\cos \frac{k_x a}{} \quad \sqrt{3}k_y a}$$

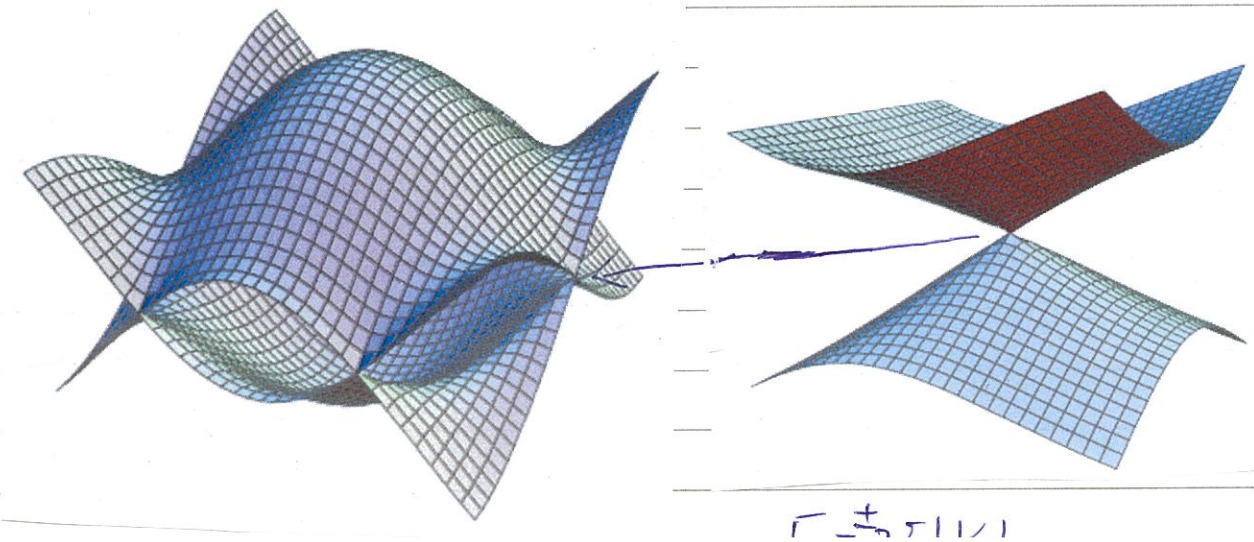
No.
Date 10-31.

E_k vanishes at corners of 1st Brillouin zone.



○ & ● are two inequivalent Dirac points.

Since E_k vanishes linearly at corners, they are known as Dirac Spectrum. with $m=0$. & corner K points are known as Dirac points.
($m=0$, implies mobility $= \infty$, $m\vec{v} = e\vec{E}\tau$,
$\vec{v} = \mu\vec{E}$, $\mu = \frac{e\tau}{m}$, speed is fast.)



$E = \pm \hbar v |k|$

Other methods for calculating bands.

Conceptually, nearly-free and tight-binding electron models capture essential features of energy bands. Quantatively, however, they are often too crude to be useful to be compared with experients (except for particular materals, such as graphene is well described by tight-binding model). There are more reliable methods which we shall briefly introduce them in the followings.

Orthogonalized Plane Waves (OPW)

To find the energy band, one needs a good complete & orthogonal set of basis functions.

The simplest one is the plane waves as we have used in nearly free electro model.

$$\psi_{\vec{k}}(\vec{r}) = \sum_{G} C_{k+G} \frac{e^{i(\vec{k}+\vec{G})\cdot\vec{r}}}{\sqrt{V}}$$

and $\left(\frac{\hbar^2(\vec{k}+\vec{G})^2}{2m} - \varepsilon\right) C_{k+\vec{G}} + \sum_{G'} U_{G'+G}\, C_{k+G'} = 0$

$C_{k+\vec{G}}$ couples to infinite many G', the convergance is often too slow. The reason for this is not a very good description for the region near ions in which the electronic state may become more localized

These tightly bind atomic like states are termed core levels.

Examples:	Core states	Valence states
Na	$1s^2 2s^2 2p^6$	$3s^1$ & higher
Si	$1s^2 2s^2 2p^6$	$3s^2, 3p^2$ & higher
Cu	$1s^2 2s^2 2p^6 3s^2 3p^6$	$3d^{10}, 4s^1$ & higher.

Let us denote core levels by c ($c = 1s, 2s, 2p, \ldots$)

$$\langle \vec{r} | c j \rangle = \phi_c(\vec{r} - \vec{R}_j)$$

$$\left(\frac{-\hbar^2\nabla^2}{2m} + \underset{\substack{\uparrow \\ \text{atomic potential localized at } R_j}}{U_a(\vec{r}-\vec{R}_j)} - E_c\right) \phi_c(\vec{r} - \vec{R}_j) = 0$$

The valence & conduction bands are what we are interested in and must be orthogonal to those core states. This, however, takes a lot of plane waves to get the orthogonality. For this reason, the orthogonalized plane waves (OPW) was introduced by Herring and Hill (PR 57, 1168, 184; 58, 132, 1940).

One defines

$$|\omega_k\rangle = |\vec{k}\rangle - \sum_{cj} \langle cj|k\rangle\, |cj\rangle \quad \text{--- ㉚}$$

$$= \Big(1 - \sum_{cj} |cj\rangle\langle cj|\Big)|k\rangle$$

$$\therefore \langle cj|\omega_k\rangle = 0$$

Here $\langle cj|\vec{k}\rangle = \int d\vec{r}\, \langle cj|\vec{r}\rangle\langle\vec{r}|\vec{k}\rangle$

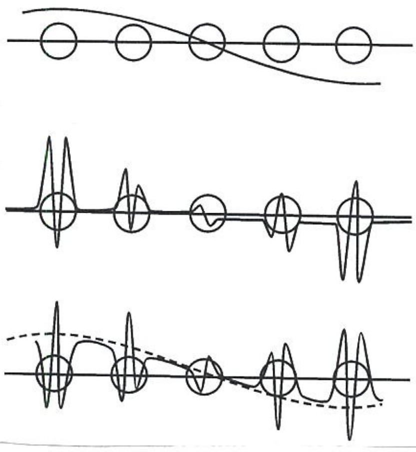
$$= \frac{1}{\sqrt{V}} \int d\vec{r}\, e^{i\vec{k}\cdot\vec{r}}\, \phi_c^*(\vec{r}-\vec{R}_j)$$

$$= \Big[\frac{1}{\sqrt{V}} \int d\vec{r}\,'\, e^{i\vec{k}\cdot\vec{r}\,'}\, \phi_c^*(\vec{r}\,')\Big]\, e^{i\vec{k}\cdot\vec{R}_j}$$

$$\equiv \chi_c^j$$

$$\therefore \omega_k(\vec{r}) = \frac{1}{\sqrt{V}} e^{i\vec{k}\cdot\vec{r}} - \sum_j \chi_c^j e^{i\vec{k}\cdot\vec{R}_j}\, \phi_c(\vec{r}-\vec{R}_j)$$

The picture of OPW is shown in below.



With each ion core, ϕ_c is large and outside the core, ϕ_c is small. Using W_K to expand $\psi_{\vec{K}}$, improves the convergence of calculations. $\left(\psi_k(\vec{r}) = \sum_G a_G W_{k+G}(\vec{r}) \right)$

Pseudopotential method

The OPW method leads in a natural way to the concept of pseudopotential.

Starting from the exact Schrödinger eq.

$$\left[\frac{-\hbar^2}{2m} \nabla^2 + U(r) \right] \psi_k(\vec{r}) = \varepsilon_k \psi_k \, , \quad \cdots ㉛$$

and approximately

$$\left[\frac{-\hbar^2}{2m} \nabla^2 + U(r) \right] \phi_c(\vec{r} - \vec{R}_j) = \varepsilon_c^j \, \phi_c(\vec{r} - \vec{R}_j) \, ,$$

if we construct $\psi_k(\vec{r})$ by

$$|\psi_k\rangle = \sum_G a_G\, |w_{k+G}\rangle$$

$$= \Big(1-\sum_{c_j}|c_j\rangle\langle c_j|\Big)\underbrace{\sum_G a_G\,|k+G\rangle}$$

the problem is to find $|\phi_k\rangle$ $^{|\phi_k\rangle}$.
Similarly to $w_k(\vec{r})$, we have

$$|\psi_k\rangle = |\phi_k\rangle - \sum_{c_j}|c_j\rangle\langle c_j|\phi_k\rangle \quad \text{---} ㉜$$

with $|\phi_k\rangle \simeq$ extended waves ($\simeq |k\rangle$ in $|w_k\rangle$)

Substituting $|\psi_k\rangle$ into eq. ㉛, one gets

$$\Big[\frac{-\hbar^2}{2m}\nabla^2+\hat{U}\Big]|\psi_k\rangle$$

$$=\Big[\frac{-\hbar^2}{2m}\nabla^2+\hat{U}\Big]|\phi_k\rangle - \sum_{c_j}\varepsilon_c^j\,|c_j\rangle\langle c_j|\phi_k\rangle$$

$$=\varepsilon_k|\psi_k\rangle = \varepsilon_k|\phi_k\rangle - \sum_{c_j}\varepsilon_k\,|c_j\rangle\langle c_j|\phi_k\rangle$$

$$\therefore\ \Big[\frac{-\hbar^2}{2m}\nabla^2+\underbrace{\hat{U}+\sum_{c_j}(\varepsilon_k-\varepsilon_c^j)|c_j\rangle\langle c_j|}_{U_{ps}}\Big]|\phi_k\rangle = \varepsilon_k|\phi_k\rangle \quad \text{---} ㉝$$

Therefore, $|\phi_k\rangle$ satisfies the Schrödinger eq. with an effect potential

$$\hat{U}_{ps} = \hat{U} + \sum_{C_j} (\varepsilon_k - \varepsilon_c^j)\,|C_j\rangle\langle C_j| .$$

Therefore, the original problem (eq. ㉛) is replaced by solving eq. ㉝ with different potential which is thus termed pseudopotential.

It's clear that $\hat{U}_{ps}$ is nonlocal:

$$\langle\vec{r}|\hat{U}_{ps}|\vec{r}'\rangle = U\,\delta(r-r') + \underbrace{\sum_{C_j} (\varepsilon_k - \varepsilon_c^j)\,\langle\vec{r}|C_j\rangle\langle C_j|\vec{r}'\rangle}_{v(r,r')}$$

$\neq 0$ if $\vec{r} \neq \vec{r}'$.

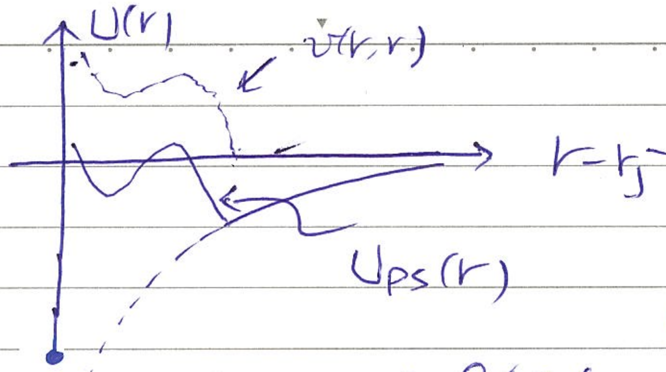
$$\therefore \langle\vec{r}|\hat{U}_{ps}|\phi_k\rangle = U(r)\,\phi_k(\vec{r}) + \int v(r,r')\,\phi_k(\vec{r}')\,d\vec{r}'$$

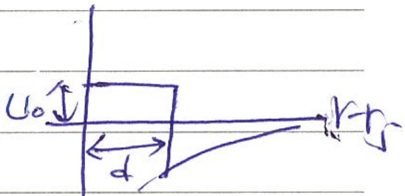
Even though $v(r,r')$ is nonlocal, the local part is still dominating because $\langle\vec{r}|C_j\rangle = \phi_c(\vec{r}-\vec{R}_j)$ is localized in the jth ion.

$$\because v(\vec{r},\vec{r}) = \sum_{C_j} (\varepsilon_k - \varepsilon_c^j)\,|\langle\vec{r}|C_j\rangle|^2 > 0$$

($\varepsilon_k > \varepsilon_c^j$ Valence of conduction e's energy is larger)

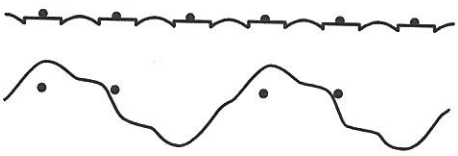
$\therefore$ $v(\vec{r},\vec{r}')$ is repulsive and cancels a portion of attractive periodic potential $U(r)$ near ions as indicated below.





In practice, a 'model potential (left figure) is adopted with appropriately choosen d & U_0' to fit expt. Hence the problem is replaced by a weaker potential $U_{ps}(r)$. as indicated below, the rapid oscillations in the wavefunction have been erased so that it results in a smooth plane-wave-like $\phi_c(\vec{r})$.

Actual potential & corresponding wavefunction



pseudopotential corresponding pseudowavefunction

Now, we can then understand why the valence electron in Na seems to behave as a free particle despite that the crystal potential is very strong at the ionic sites.

If the Pauli exclusion is further included,

the effective potential will become even weaker.
The free-particle behavior thus finds
its explanation in the pseudopotential method.
The method has been used calculate
Be, Na, K, Ge, Si etc with considrable
success.

Unfortunately, it (including OPW) fails
for the transition & rare-earth metals,
where there is no clear dividing line between
tightly-bound core electrons and loosely-bound
valence electrons.

The Muffin-Tin potential & Augmented Plane Waves
(MTO) (APW)

There appear two extreme limits in
the band approach: (i) nearly-free electron
(including OPW & pseudo-potential method)

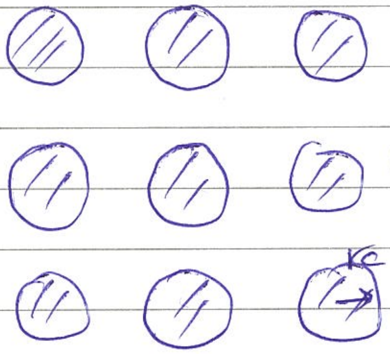
(ii) tight-binding electron (LCAO).

To combine the advantages of each
approach, J. C. Slater (1934) proposed

the concept of a muffin-tin potential to solve this problem:

Periodic potential is divided into two parts: (i) a spherical symmetric atomic potential within ion cores

$$V(r) = V_c(|\vec{r}-\vec{R}_j|) \qquad |\vec{r}-\vec{R}_j| < r_c$$



(ii) a constant potential (usually chosen to be 0)

$$V(r) = 0 \qquad |\vec{r}-\vec{R}_j| > r_c$$

r_c = radius of ion core.

Given the muffin-tin potential, there are many methods to utilize the muffin-tin potential: linear augmented plane waves (LAPW), linear muffin tin orbital (LMTO), & Korringa-Kohn-Rostoker (KKR).

We shall not go into details but just briefly mention the augmented plane wave (APW) method as an example. The wavefunction near a site is taken to be

$$\text{(i)} \quad w_K = \begin{cases} \text{atomic function} & r < r_c \\ \frac{1}{\sqrt{v}} e^{i\vec{K}\cdot\vec{r}} & r \geq r_c \end{cases}$$

(ii) $W_k(r)$ is required to be continuous at $r=r_c$ (by linear combination of atomic-like wavefunction for $r<r_c$)

Clearly, for $r<r_c$, $W_k(r)$ satisfies

$$-\frac{\hbar^2}{2m}\nabla^2 W_k(r) + U_a(r)W_k(r) = \mathcal{E}\, W_k(r)$$

Generally, $W_k(\vec{r}) = \sum_{l=0}^{\infty}\sum_{m=-l}^{l} A_{lm} Y_{lm}(\theta,\phi) R_{l\mathcal{E}}(r)$

doesn't depend on k!

One then form $\psi_k(\vec{r}) = \sum_{G} C_G W_{k+G}(\vec{r})$ as

the trial wavefunction. The wave function $\psi_k(\vec{r})$ appears to combine near-free approach and tight-binding approach.

However, because $W_k(\vec{r})$ usually doesn't have continuous derivatives on the boundary $r=r_c$, instead of working on schrödinger eq directly, one minizes the energy w.r.t C_G via variational principle:

$$E[\psi] = \frac{\int \frac{\hbar^2}{2m}|\nabla\psi|^2 + U(r)|\psi(r)|^2\, d^3\vec{r}}{\int |\psi(r)|^2\, d^3\vec{r}}$$

$\frac{\partial E}{\partial C_G} = 0 \Rightarrow$ homogeneous equations for C_G

coefficients depend on $\mathcal{E}$

$\det(\quad) = 0 \quad \Rightarrow \mathcal{E}(K)$

The non-analytic problem of $W_k(\vec{r})$ can be fixed by introducing the basis function

$$\chi_{\ell m}(\epsilon, r) = \begin{cases} i^\ell Y_{\ell m}(\theta,\phi)\left[R_{\ell\varepsilon}(r) + P_{\ell\varepsilon}\left(\frac{r}{r_c}\right)^\ell\right] & \text{for } r<r_c \\ i^\ell Y_{\ell m}\left(\frac{r_c}{r}\right)^{\ell+1} & \text{for } r>r_c \end{cases}$$

Here $\nabla^2\chi = 0$ for $r > r_c$. $P_{\ell\varepsilon}$ term is introduced to make $\chi_{\ell m}(r)$ & $\frac{d}{dr}\chi_{\ell m}(r)$ continuous at $r = r_c$:

$$R_{\ell\varepsilon}(r_c) + P_{\ell\varepsilon} = 1$$

$$R'_{\ell\varepsilon}(r_c) + \ell\frac{P_{\ell\varepsilon}}{r_c} = -\frac{\ell+1}{r_c}$$

$$\therefore\ P_{\ell\varepsilon} = \frac{-(\ell+1)/r_c - R'_{\ell\varepsilon}(r_c)}{\ell/r_c - R'_{\ell\varepsilon}(r_c)} = \frac{D_\ell(\varepsilon)+\ell+1}{D_\ell(\varepsilon)-\ell}$$

$$D_\ell \equiv \frac{r_c}{R_{\ell\varepsilon}(r_c)} R'_{\ell\varepsilon}(r_c)$$

All information about $\widehat{\ }$ potential is contained in $P_{\ell\varepsilon}$.
$\chi_{\ell m}$ is the Linearized atomic Muffin-Tin Orbitals (LMTO).
By writing $\psi_k(\vec{r}) = \sum_{\substack{\ell m \\ \vec{R}}} A^{\vec{R}}_{\ell m}\, e^{i\vec{k}\cdot\vec{R}}\, \chi_{\ell m}(\varepsilon, \vec{r}-\vec{R})$

and requiring $\psi_k(\vec{r} \cong \vec{R}_0) = 0$, one requires

$P_{\ell\varepsilon}$ term near $\vec{R}_0$ is cancelled by $r>r_c$ terms from other $\vec{R} \neq \vec{R}_0$.

This then becomes a homogeneous equation among $A_{\ell m}^{\vec{K}}$. The vanishing of the determinant then determines $\underline{\varepsilon_K}$!

The APW is a powerful method for calculating band structures, especially for transition metals and rare earths.

$\vec{K}\cdot\vec{p}$ method

For some materials, particular, semiconductors, one is interested in electronic structures close to a conduction band minimum or valence band maximum.

Therefore, instead of working detailed spectrum in the entire Brillouin zone, one focuses on band min or max near energy.

The method $\vec{K}\cdot\vec{p}$ is proven to be very useful in this situation.

From $\psi_K(\vec{r}) = e^{i\vec{K}\cdot\vec{r}}\,u(r)$

& $\left[\frac{\hat{p}^2}{2m} + U(r) - \varepsilon_K\right]\psi_K = 0$, we got

$$\left[\frac{1}{2m}(\hat{p}+\hbar\vec{k})^2 + U(r) - \varepsilon_k\right] u_k(\vec{r}) = 0 \quad \cdots (34)$$

For $k \simeq 0$ (near Γ point), one gets

$$\left(\frac{\hat{p}^2}{2m} + \frac{\hbar}{m}\vec{k}\cdot\vec{p} + \frac{\hbar^2 k^2}{2m} + U(r) - \varepsilon_k\right) u_k(\vec{r}) = 0 \quad \llcorner\cdots (35)$$

It's clear that at Γ, $k=0$, this equation becomes

$$\left(\frac{\hat{p}^2}{2m} + U(r) - \varepsilon_0\right) u_0(\vec{r}) = 0 .$$

There are infinite # of solutions,

$$u_0^{(1)}, u_0^{(2)}, \ldots, u_0^{(n)}, \ldots$$

with $\varepsilon_0^{(1)}, \varepsilon_0^{(2)}, \ldots, \varepsilon_0^{(n)}, \ldots$

They form a complete set and we can use them to expand $u_k(\vec{r})$ at $k \neq 0$.

$$u_k(\vec{r}) = \sum_n a_n(k)\, u_0^{(n)}(r) \quad \cdots (36)$$

Substituting eq. (36) into eq. (35), we got

$$\left(\varepsilon_0^{(n)} + \frac{\hbar^2 k^2}{2m} - \varepsilon_k\right)\sum_n a_n u_0^{(n)} + \frac{\hbar}{m}\vec{k}\cdot\vec{p}\sum_n a_n u_0^{(n)} = 0 \quad \llcorner\cdots (37)$$

$\because \langle u_0^{(m)} | u_0^{(n)} \rangle = \delta_{nm}$

$\therefore$ Eq. (37) implies

$$\left(\varepsilon_0^{(n)} + \frac{\hbar^2 k^2}{2m} - \varepsilon_k\right) a_m + \sum_n \langle u_0^{(m)} | \frac{\hbar}{m}\vec{k}\cdot\vec{p} | u_0^{(n)} \rangle a_n = 0 \quad \llcorner\cdots (38)$$

Therefor, $\hat{H}_1 \equiv \frac{\hbar}{m}\vec{K}\cdot\vec{p}$ becomes a perturbation !

Eq. (38)

$$\Rightarrow \begin{pmatrix} \varepsilon_0^{(1)}+\frac{\hbar^2K^2}{2m}-\varepsilon_K & \langle u_0^{(1)}|H_1|u_0^{(2)}\rangle & \langle u_0^{(1)}|H_1|u_0^{(3)}\rangle & \cdots \\ \langle u_0^{(2)}|H_1|u_0^{(1)}\rangle & \varepsilon_0^{(2)}+\frac{\hbar^2K^2}{2m}-\varepsilon_K & \cdots & \\ \langle u_0^{(3)}|H_1|u_0^{(1)}\rangle & \langle u_0^{(3)}|H_1|u_0^{(2)}\rangle & - & \\ - & \cdot & \cdot & \end{pmatrix}\begin{pmatrix}a_1\\ a_2\\ a_3\\ \vdots\end{pmatrix}=0$$

For K small, one can use perturbation theory.

We get to the 2nd order

$$\varepsilon_K^{(n)\prime} = \varepsilon_0^{(n)} + \frac{\hbar^2K^2}{2m} + \frac{\hbar^2}{m^2}\sum_{n'\neq 0}\frac{|\vec{K}\cdot\langle u_0^{(n)}|\vec{p}|u_0^{(n')}\rangle|^2}{\varepsilon_0^{(n)}-\varepsilon_0^{(n')}}$$

which can be written as

$$\varepsilon_K^{(n)} = \varepsilon_0^{(n)} + \frac{\hbar^2}{2}\vec{K}\cdot(m^*)^{-1}\cdot\vec{K}$$

with the inverse mass tensor (of band n) being given by

$$(m^*)^{-1}_{ij} = m^{-1}\delta_{ij} + \frac{2}{m^2}\sum_{n'\neq n}\frac{\langle u_0^{(n)}|\hat{p}_i|u_0^{(n')}\rangle\langle u_0^{(n')}|\hat{p}_j|u_0^{(n)}\rangle}{\varepsilon_0^{(n)}-\varepsilon_0^{(n')}}$$

L - - (38)

The inverse mass tensor describes how mass of electron is renormalized due to the periodic potential.

Since $\hat{p} \sim \frac{\hbar}{a}$ $a =$ lattice spacing,

$\langle u_0^{(n)} | \hat{p}_i | u_0^{(n')} \rangle \sim \hbar/a$,

one gets $\frac{m}{m^*} \simeq 1 + 2 \frac{\hbar^2/ma^2}{\Delta G}$

$\Delta G \sim$ energy gap.

For $a \sim 3 \times 10^{-8}$ cm, $\frac{\hbar^2}{ma^2} \sim 10$ eV, typical energy gaps at Γ (semiconductors) $\sim 10^{-1}$ eV

$\frac{m}{m^*} \sim 10^{-2}$, i.e. $m^* \sim 10^{-2} m$.

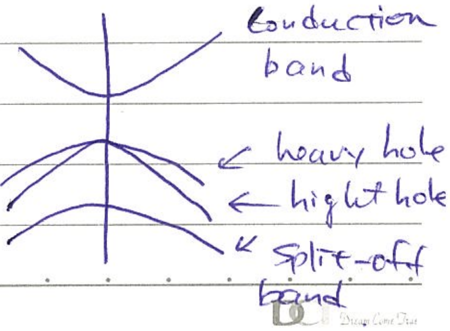
for small-gap semiconductors.

(e.g. Ge, $m^* = 0.55 m$, GaAs, $m^* = 0.067 m$

ZnO, $m^* = 0.19 m$ ZnSe, $m^* = 0.17 m$

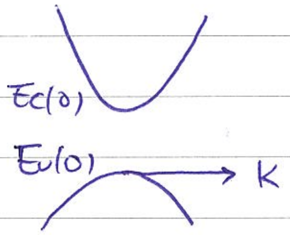
InSb, $m^* = 0.013$ m)

The above perturbation breaks down when there is degeneracy at Γ point. This indeed happens for zincblende semiconductors (such as InAs, InSb, GaAs, ...) as we have seen. In that case, there are 4 bands that need to be included:



This is Known as Kano model.
(J. Phys. Chem. Solids. 1, 82, 1956;
1, 249, 1957)

Example : $\vec{k}\cdot\vec{p}$ for two non-degenerate band



$$\begin{vmatrix} E_c(0)+\frac{\hbar^2k^2}{2m}-\varepsilon_k & \frac{\hbar}{m}\vec{k}\cdot\vec{P}_{cv} \\ \frac{\hbar}{m}\vec{k}\cdot\vec{P}_{cv} & E_v(0)+\frac{\hbar^2k^2}{2m}-\varepsilon_k \end{vmatrix}=0$$

$$\vec{P}_{cv}=\langle c|\vec{p}|v\rangle$$

$$\therefore\ \varepsilon_k=\frac{1}{2}\left[E_c(0)+E_v(0)+\frac{\hbar^2k^2}{m}\right]\pm\frac{1}{2}\left[(E_c(0)-E_v(0))^2+\frac{4\hbar^2}{m}|\vec{k}\cdot\vec{P}_{cv}|^2\right]^{\frac{1}{2}}$$

Set $E_v(0)=0$, $E_v(0)=\varepsilon_g$ (energy gap)

$$\varepsilon_k=\frac{1}{2}\left(\varepsilon_g+\frac{\hbar^2k^2}{m}\right)\pm\frac{1}{2}\left[\varepsilon_g^2+\frac{4\hbar^2}{m}|\vec{k}\cdot\vec{P}_{cv}|^2\right]^{\frac{1}{2}}$$

$$\underset{\text{small } \vec{k}\cdot\vec{P}_{cv}}{\approx}\begin{cases}\varepsilon_g+\frac{\hbar^2k^2}{2m}+\frac{\hbar^2}{\varepsilon_g m^2}|\vec{k}\cdot\vec{P}_{cv}|^2 \rightarrow \text{Conduction band}\\ \frac{\hbar^2k^2}{2m}-\frac{\hbar^2}{\varepsilon_g m^2}|\vec{k}\cdot\vec{P}_{cv}|^2 \rightarrow \text{Valence band}\end{cases}$$

Assum $\vec{P}_{cv}$ isotropic, ∴ $\vec{k}\cdot\vec{P}_{cv}=kP_{cv}$

then conduction band $\frac{\hbar^2k^2}{2m_c^*}=\frac{\hbar^2k^2}{2m}+\frac{\hbar^2k^2}{\varepsilon_g m^2/P_{cv}^2}$

$$\therefore\ \frac{m}{m_c^*}=1+\frac{2P_{cv}^2}{m\varepsilon_g}$$

Valence band $\frac{\hbar^2k^2}{2m_v^*}=\frac{\hbar^2k^2}{2m}-\frac{\hbar^2k^2P_{cv}^2}{\varepsilon_g m^2}$

$$\frac{m}{m_v^*}=1-\frac{2P_{cv}^2}{m\varepsilon_g}$$

$\varepsilon_g<\frac{2P_{cv}^2}{m}$, valence band → ⌒

No.
Date.

For small bandgap

$$E_g \ll \frac{2P_{cv}^2}{m}, \quad |m_v^*|, m_c^* \propto E_g$$

It is expected

$$\frac{m_c^*(GaAs)}{m_c^*(Ge)} = \frac{E_g(GaAs)}{E_g(Ge)}$$

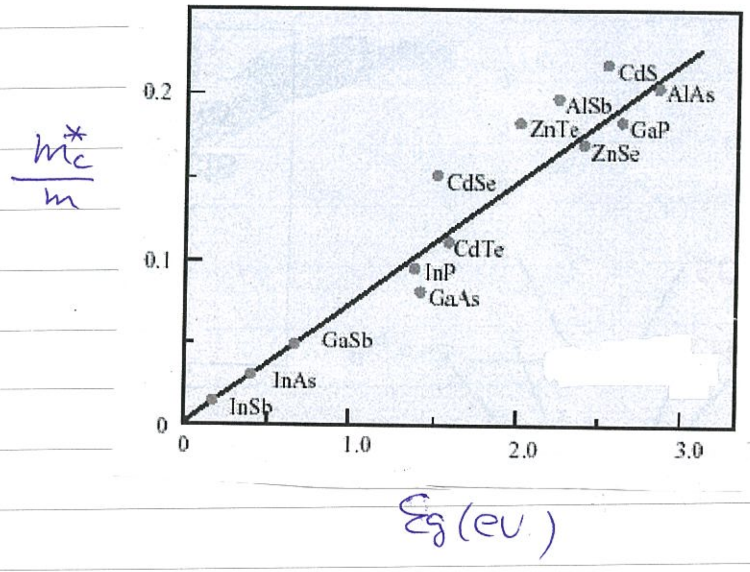
expt. $E_g(GaAs) = 1.52$ eV

$E_g(Ge) = 0.881$ eV

$m_c^*(Ge) = 0.041$

$\Rightarrow m_c^*(GaAs) = 0.070$, expt. $m_c^*(GaAs) = 0.067$

Comparison of m_c^*/m with E_g :



Relativistic effects

Typical energy scales between the lowest and highest filled conduction bands are of the order of a few eV.

$\because mc^2 \sim 0.5\,\mathrm{MeV}$

Therefore, relativistic effect $\sim 10^{-5}$.

It might seem that relativistic effects are not important.

However, the potentials well near the ion of heavy atoms are very deep. If $\Delta E \sim 10\,\mathrm{eV}$, $10^{-4}\Delta E \sim 1\,\mathrm{meV}$, $10^{-5}\Delta E \sim 0.1\,\mathrm{meV}$ which can be now measured and make qualitative change at points such as Γ point.

This is what happens at Γ point of semiconductors of zincblende structure.

Lowest order relativistic corrections are

$$H_{rel} = \underbrace{-\frac{p^4}{8m^3c^2}}_{\text{mass-velocity term}} + \underbrace{\frac{1}{2m^2c^2}\,\vec{s}\cdot\left[(\nabla U)\times\vec{p}\,\right]}_{\text{spin-orbital interaction}} + \underbrace{\frac{\hbar^2}{8m^2c^2}\nabla^2 U}_{\text{Darwin term}}$$

They give the same order of energy shifts and must be included in materials such as PbTe.

In the case of zincblende semiconductors, one needs to include spin. The $\vec{k}\cdot\vec{p}$ term is then replaced by $\vec{\pi}\cdot\vec{p}$ with

$$\vec{\pi} = \hbar\vec{k} + \frac{\hbar}{4mc^2}\vec{\sigma}\times\nabla U \, . \qquad \left(\vec{s}=\frac{\hbar}{2}\vec{\sigma}\right)$$

The degenerate perturbation now requires to diagonalize a 8×8 matrix, to obtain a good description of bands near Γ point.

Specifically, eq. ㉟ becomes

$$\left[\underbrace{\frac{\hat{p}^2}{2m} + U(r)}_{H_0} + \frac{\hbar}{m}\vec{k}\cdot\vec{p} + \frac{\hbar}{4m^2c^2}\vec{\sigma}\times\nabla U\cdot\vec{p}\right]u_k(\vec{r})$$

$$= \varepsilon_k' u_k(\vec{r}) \, . \qquad \varepsilon_k' = \varepsilon_k - \hbar^2k^2/2m \quad \text{---} ㊴$$

$\vec{k}$ = crystal momentum, $\vec{p}$ = actual momentum

For IV semiconductors

C $1s^2\,\boxed{2s^22p^2}$

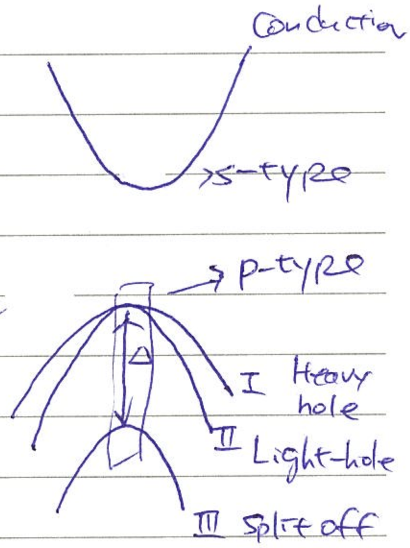
Si $1s^2\,2s^2\,2p^6\,\boxed{3s^23p^2}$

Ge $1s^22s^22p^63s^23p^6\,3d^{10}\boxed{4s^24p^2}$

III − V

Ga $1s^22s^22p^63s^23p^63d^{10}\,\boxed{4s^24p^1}$

As $\quad$, , $\quad \boxed{4s^23p3}$



Δ = split-off energy

As we indicated before, at Γ point, bands can be classified according to irreducible representations of point groups. For semiconductors, it turns out $u(k=0,\vec{r})$ possesses the symmetry as $|s\rangle$ state (invariant under all point group operations) for conduction band, while for valence band, $u(k=0,\vec{r})$ behaves as $|X\rangle$, $|Y\rangle$, $|Z\rangle$.

To include effects of spins, one forms the following basis:

$$|u_1\rangle = |is\downarrow\rangle$$

$$|u_2\rangle = |\tfrac{X-iY}{\sqrt{2}}\uparrow\rangle = |Y_{1,-1}\uparrow\rangle$$

$$|u_3\rangle = |Z\downarrow\rangle = |Y_{10}\downarrow\rangle$$

$$|u_4\rangle = |-\tfrac{X+iY}{\sqrt{2}}\uparrow\rangle = |Y_{11}\uparrow\rangle$$

$$|u_5\rangle = |is\uparrow\rangle$$

$$|u_6\rangle = |-\tfrac{X+iY}{\sqrt{2}}\downarrow\rangle = |Y_{11}\downarrow\rangle$$

$$|u_7\rangle = |Z\uparrow\rangle = |Y_{10}\uparrow\rangle$$

$$|u_8\rangle = |\tfrac{X-iY}{\sqrt{2}}\downarrow\rangle = |Y_{1,-1}\downarrow\rangle$$

By evaluating the matrix element (using symmetry of u_n), one can write down 8×8 H. For instance,

$$H_{11} = \langle iS\downarrow | H_0 + \frac{\hbar}{m}\vec{k}\cdot\vec{p} + \frac{\hbar}{4m^2c^2}\sigma\cdot\nabla U\times p | iS\rangle$$

$$= \underbrace{\langle S|H_0|S\rangle}_{E_s}\underbrace{\langle\downarrow|\downarrow\rangle}_{1} + \frac{\hbar}{m}\vec{k}\cdot\underbrace{\langle S|\vec{p}|S\rangle}_{\int d^3\vec{r}\, S(\vec{r})\frac{\hbar}{i}\nabla S(\vec{r}) = 0}$$

$$+\frac{\hbar}{4m^2c^2}\langle S\downarrow|\frac{\hbar}{4m^2c^2}\underbrace{\sigma_z\left(\frac{\partial U}{\partial x}p_y - \frac{\partial U}{\partial y}p_x\right)}_{\text{Changes sign when } x\leftrightarrow y}|S\downarrow\rangle$$

$$= 0$$

(σ_x, σ_y will flip ↓ to ↑ = 0).

$$H = \begin{pmatrix} h & 0 \\ 0 & h \end{pmatrix}$$

with
$$h = \begin{pmatrix} E_s & 0 & kp & 0 \\ 0 & E_p - \Delta/3 & \sqrt{2}\Delta/3 & 0 \\ kp & \sqrt{2}\Delta/3 & E_p & 0 \\ 0 & 0 & 0 & E_p + \Delta/3 \end{pmatrix}$$

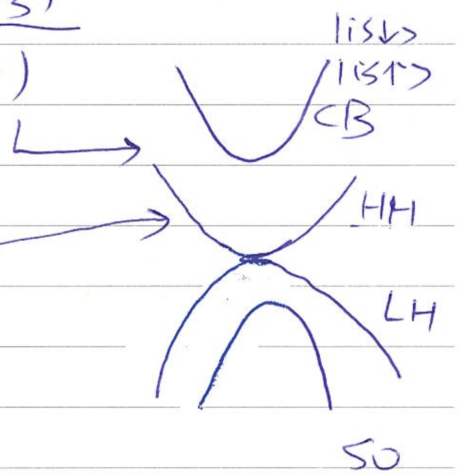
$p = -i\frac{\hbar}{m}\langle S|p_z|Z\rangle$ (Kane's parameter)

$\Delta = \frac{3\hbar i}{4m^2c^2}\langle X|\frac{\partial U}{\partial x}p_y - \frac{\partial U}{\partial y}p_x|Y\rangle$ (split off energy)

This is the Kane's model.

For small K, the Kane's model predicts

$$\mathcal{E}_c(K) = \mathcal{E}_g + \frac{\hbar^2 K^2}{2m} + \frac{K^2 P^2 (\mathcal{E}_g + 2\Delta/3)}{\mathcal{E}_g(\mathcal{E}_g+\Delta)}$$



$$\mathcal{E}_{HH}(K) = \frac{\hbar^2 K^2}{2m}$$

$$\mathcal{E}_{LH}(K) = \frac{\hbar^2 K^2}{2m} - \frac{2K^2P^2}{3\mathcal{E}_g}$$

$$\mathcal{E}_{SO}(K) = -\Delta + \frac{\hbar^2 K^2}{2m} - \frac{K^2P^2}{3(\mathcal{E}_g+\Delta)}$$

eigenfunctions : CB $|iS\uparrow\rangle, |iS\downarrow\rangle$

HH $\frac{-1}{\sqrt{2}}(|X+iY\rangle|\uparrow\rangle = |\underset{j}{3/2}, \underset{m_j}{3/2}\rangle$

$\frac{1}{\sqrt{2}}|(X-iY\rangle|\downarrow\rangle = |3/2, -3/2\rangle$

LH $\frac{1}{\sqrt{6}}|X-iY\rangle|\uparrow\rangle + \sqrt{\frac{2}{3}}|Z\downarrow\rangle = |\frac{3}{2}, \frac{-1}{2}$

$-\frac{1}{\sqrt{6}}|X+iY\rangle|\downarrow\rangle + \sqrt{\frac{2}{3}}|Z\uparrow\rangle = |3/2, \frac{1}{2}$

SO $\frac{1}{\sqrt{3}}|X-iY\rangle\overset{|\uparrow\rangle}{\,} - \frac{1}{\sqrt{3}}|Z\downarrow\rangle = |1/2, -\frac{1}{2}\rangle$

$\frac{1}{\sqrt{3}}|X+iY\rangle|\downarrow\rangle + \frac{1}{\sqrt{3}}|Z\rangle|\uparrow\rangle = |\frac{1}{2}, \frac{1}{2}\rangle$

Note that notation $|j, m\rangle$ simply means the state's symmetry is the same as $|j\,m\rangle$, it doesn't mean the state is a real eigenstate to the angular momentum!

The Kane's model has wrong prediction for the heavy hole band. This is fixed by the Luttinger-Kohn model in which 6 Valance Bands are considered with corrections from remote bands.

The effective Luttinger model for $\overset{\text{HH \& LH}}{}$ holes are given by

$$H = \frac{\hbar^2}{2m}\left[\left(\gamma_1 + \tfrac{5}{2}\gamma_2\right)k^2 - 2\gamma_2\left(\vec{k}\cdot\vec{S}\right)^2\right]$$

$\vec{S}$ = spin 3/2 matrices. Here γ_1, γ_2 are Luttinger parameters satisfying

$$\frac{-\hbar^2}{2m}\gamma_1 = \tfrac{1}{3}\left(A_0 + 2B_0\right), \quad \frac{-\hbar^2}{2m}\gamma_2 = \frac{1}{6}\left(A_0 - B_0\right)$$

$$A_0 = \frac{\hbar^2}{2m} + \frac{\hbar^2}{m}\underset{\underset{\text{remote bands}}{\uparrow}}{\sum_n}\frac{P^x_{\underline{X}n}P^x_{n\underline{X}}}{\varepsilon_0 - \varepsilon_n}$$

$\varepsilon_0 = 0$ for LH, HH bands

$$B_0 = \quad '' \qquad '' \sum_n \frac{P^x_{\underline{X}n}P^y_{n\underline{X}}}{\varepsilon_0 - \varepsilon_n}$$

$\varepsilon_0 = -\Delta$ for SO band

$$S_x = \begin{pmatrix} 0 & \sqrt{3}i/2 & 0 & 0 \\ -\sqrt{3}i/2 & 0 & i & 0 \\ 0 & -i & 0 & \sqrt{3}i/2 \\ 0 & 0 & -\sqrt{3}i/2 & 0 \end{pmatrix} \qquad S_y = \begin{pmatrix} 0 & \sqrt{3}/2 & 0 & 0 \\ \sqrt{3}/2 & 0 & 1 & 0 \\ 0 & 1 & 0 & \sqrt{3}/2 \\ 0 & 0 & \sqrt{3}/2 & 0 \end{pmatrix}$$

$$S_z = \begin{pmatrix} 3/2 & 0 & 0 & 0 \\ 0 & \frac{1}{2} & 0 & 0 \\ 0 & 0 & -1/2 & 0 \\ 0 & 0 & 0 & -3/2 \end{pmatrix}$$

As a reference, in the following, we also outline the Luttinger-Kohn model.

As mentioned, to treat the degenerate unperturbed states, one first diagonalize

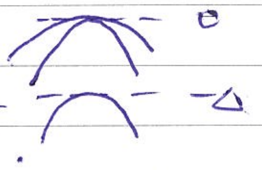
$$\underbrace{H_0 + \begin{pmatrix} 0 & 0 \\ 0 & A \end{pmatrix}}_{H_0'} . \quad \text{where } V = \begin{pmatrix} 0 & 0 \\ 0 & A \end{pmatrix} + \begin{pmatrix} \times & \times \\ \times & 0 \end{pmatrix}$$

is the perturbation.

This is what is done in the Kane model. and we have seen that curvature of the HH band is wrong. We need to include $\begin{pmatrix} \times & \times \\ \times & 0 \end{pmatrix}$.

In the Luttinger-Kohn model, one starts from the 6 valence bands and include the second order perturbation of $\begin{pmatrix} \times & \times \\ \times & 0 \end{pmatrix}$ from other bands.

$$\left. \begin{aligned} H_0' &= \left(0 + \frac{\hbar^2 k^2}{2m}\right) \ (4 \text{ states}) \\ &= \left(-\Delta + \frac{\hbar^2 k^2}{2m}\right) \ (2 \text{ states}) \end{aligned} \right\} = \left(E_i + \frac{\hbar^2 k^2}{2m}\right)\delta_{ij}$$



$V =$ perturbation $= \frac{\hbar}{m} \vec{k} \cdot \vec{p}$. (not $\vec{k} \cdot \vec{\pi}$.)

$$(H_I')_{ij} = \frac{\hbar^2}{m^2}\sum_{n\neq i,j} \frac{(\vec{k}\cdot\vec{P}_{in})(\vec{k}\cdot\vec{P}_{nj})}{\mathcal{E}_0 - E_n^0} \qquad n = \text{other bands}$$

where $i, j = 1, 2, 3, 4, 5, 6$ with

$$u_{10} = \frac{-1}{\sqrt{2}}|x+iy\rangle|\uparrow\rangle = |\tfrac{3}{2},\tfrac{3}{2}\rangle$$

$$u_{20} = \frac{-1}{\sqrt{6}}|x+iy\rangle|\downarrow\rangle + \sqrt{\tfrac{2}{3}}|z\uparrow\rangle = |\tfrac{3}{2},\tfrac{1}{2}\rangle$$

$$u_{30} = \frac{1}{\sqrt{6}}|x-iy\rangle|\uparrow\rangle + \sqrt{\tfrac{2}{3}}|z\downarrow\rangle = |3/2, -\tfrac{1}{2}\rangle$$

$$u_{40} = \frac{1}{\sqrt{2}}|x-iy\rangle|\downarrow\rangle = |3/2, -3/2\rangle$$

$$u_{50} = \frac{1}{\sqrt{3}}|x+iy\rangle|\downarrow\rangle + \frac{1}{\sqrt{3}}|z\uparrow\rangle = |\tfrac{1}{2},\tfrac{1}{2}\rangle$$

$$u_{60} = \frac{1}{\sqrt{3}}|x-iy\rangle|\uparrow\rangle - \frac{1}{\sqrt{3}}|z\downarrow\rangle = |\tfrac{1}{2},-\tfrac{1}{2}\rangle$$

Except γ_1, γ_2, one has $\frac{-\hbar^2}{2m}\gamma_3 = \frac{\hbar^2}{6m^2}\sum_n \frac{P^x_{Xn}P^y_{nY} + P^y_{Xn}P^x_{nY}}{\mathcal{E}_0 - E_n}$

The result Hamiltonian, is

$$H_{LK} = -\begin{array}{l} \left[\begin{array}{cccccc} P+\Theta & -S & R & 0 & -S/\sqrt{2} & \sqrt{2}R \\ -S^+ & P-\Theta & 0 & R & -\sqrt{2}\Theta & \sqrt{3/2}\,S \\ R^+ & 0 & P-\Theta & S & \sqrt{3/2}\,S^+ & \sqrt{2}\Theta \\ 0 & R^+ & S^+ & P+\Theta & \sqrt{2}R^+ & -S^+/\sqrt{2} \\ -S^+/\sqrt{2} & -\sqrt{2}\Theta^+ & \sqrt{3/2}\,S & -\sqrt{2}R & P+\Delta & 0 \\ \sqrt{2}R^+ & \sqrt{3/2}\,S^+ & \sqrt{2}\Theta^+ & -S/\sqrt{2} & 0 & P+\Delta \end{array}\right] \end{array}\begin{array}{l} HH \\ LH \\ LH \\ HH \\ SO \\ SO \end{array}$$

where $P = \frac{\hbar^2\gamma_1}{2m}(k_x^2 + k_y^2 + k_z^2)$ $\qquad S = \frac{\hbar^2\gamma_3\sqrt{2}}{m}(k_x - ik_y)k_z$

$\Theta = \frac{\hbar^2\gamma_2}{2m}(k_x^2 + k_y^2 - k_z^2)$ $\qquad$ (Luttinger & Kohn, PR 97. 869, 1955)

$R = \frac{\hbar^2}{2m}[-\sqrt{3}\,\gamma_2(k_x^2 - k_y^2) + i2\sqrt{3}\,\gamma_3 k_x k_y]$