Physics 555 Fall 2003 – Note # 5
Band Theory for the Graphite π-System
(it also applies to MgB$_2$!)
(Dated: 24 October 2003)

The lattice structure is shown in Fig. 1. This is a 2-d network made from real 3-d carbon atoms. Graphite is a very stable material (highest melting temperature known, more stable than diamond.) Obviously it peels easily in layers, which is why it is used as a lubricant and for pencils. A single free-standing layer would be hard to peel off, but if it could be done, no doubt it would be quite stable except at the edges. In fact, carbon nanotubes are just this – layers of graphite which solve the edge problem by curling into closed cylinders. They come in "single-walled" and "multiwalled" forms, with quantized circumferences of many sizes, and with quantized helical pitch of many types. Their electrical properties are fascinating and can be understood in much the same way that a layer of graphite is understood.

As explained in class, the atomic carbon s functions as well as the atomic carbon p_x and p_y functions (abbreviated from now on as x and y) form strong bonding orbitals which are doubly occupied and lie well below the Fermi energy. They also form strongly antibonding orbitals which are high up and empty. This leaves "space" on the energy axis near the Fermi level for the z orbitals (technically called π orbitals because they point perpendicular to the direction of the bond between them.) We want to study how these z orbitals form two bands, one bonding band lower in energy which is doubly occupied, and one antibonding band higher in energy which is unoccupied. And we want to understand why these levels are not separated by a gap, but have a tendency to overlap by a small amount leading to a "semimetal."

The arguments above show why we expect to be able to truncate the basis to a single z function per atom (as long as we only care about the π bands near the Fermi level.) Since there are two atoms in the unit cell (let's call them A and B), we can form two candidate Bloch orbitals. The term "Bloch" indicates that they are eigenstates of translation; they are basis functions from which we can construct approximate Bloch eigenstates of the Hamiltonian. The Bloch orbitals are

$$|\vec{k}A> = \sqrt{\frac{1}{N}} \sum_{\vec{\ell}} e^{i\vec{k}\cdot\vec{\ell}} \phi_A(\vec{r} - \vec{\ell})$$

$$|\vec{k}B> = \sqrt{\frac{1}{N}} \sum_{\vec{\ell}} e^{i\vec{k}\cdot\vec{\ell}} \phi_B(\vec{r} - \vec{\ell})$$

where N is the number of unit cells in the sample (half the number of atoms) and $\phi_A(\vec{r})$ is a z function centered on atom A in the central ($\vec{\ell} = 0$) unit cell. We can take the position of this atom to be the origin. Then $\phi_B(\vec{r})$ is just $\phi_A(\vec{r} - \vec{\tau})$ where $\vec{\tau}$ is the vector which separates the two atoms in the central unit cell.

Now we pretend that there are no other orbitals which are likely to be mixed into the energy eigenstates near the Fermi energy – and crudely this is true. Furthermore, even though an error is necessarily being introduced at this point, it is still exactly the right way to get insight into the nature of the true orbitals, their symmetry, etc. Then the Bloch eigenstates must be

$$|\vec{k}n> = \alpha|\vec{k}A> + \beta|\vec{k}B>$$

where the expansion coefficients (normalized by $|\alpha|^2 + |\beta|^2 = 1$) are eigenvectors of the two by two Hamiltonian matrix

$$\mathcal{H}(\vec{k}) = \begin{pmatrix} <\vec{k}A|\mathcal{H}|\vec{k}A> & <\vec{k}A|\mathcal{H}|\vec{k}B> \\ <\vec{k}B|\mathcal{H}|\vec{k}A> & <\vec{k}B|\mathcal{H}|\vec{k}B> \end{pmatrix}$$

Now we must evaluate these matrix elements. First we have

$$<\vec{k}A|\mathcal{H}|\vec{k}A> = \frac{1}{N} \sum_{\vec{\ell}\vec{\ell}'} e^{i\vec{k}\cdot(\vec{\ell}-\vec{\ell}')} \int d\vec{r}\, \psi_A(\vec{r} - \vec{\ell}')\mathcal{H}\psi_A(\vec{r} - \vec{\ell}).$$

Now two different carbon atoms which are both of type A are necessarily fairly far apart – the closest three neighbors are always on the opposite sublattice. So we will make the nearest neighbor approximation which says that integrals are zero unless the functions are on the same atom or on nearest neighbor atoms. Then the integral becomes

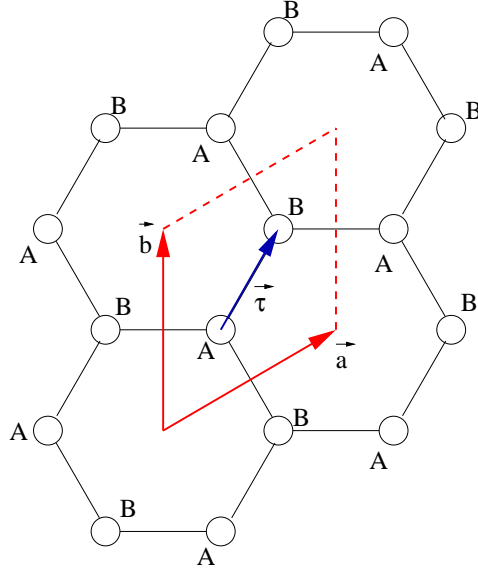$$<\vec{k}A|\mathcal{H}|\vec{k}A> = \int d\vec{r}\, \psi_A(\vec{r})\mathcal{H}\psi_A(\vec{r}).$$

FIG. 1: Lattice structure of a graphite layer (or the boron layer of MgB_2). There are two carbon atoms per cell, designated as the A and B sublattices. The vector $\vec{\tau}$ connects the two sublattices, and is not a translation vector. Primitive translation vectors $\vec{a}$ and $\vec{b}$ are shown.
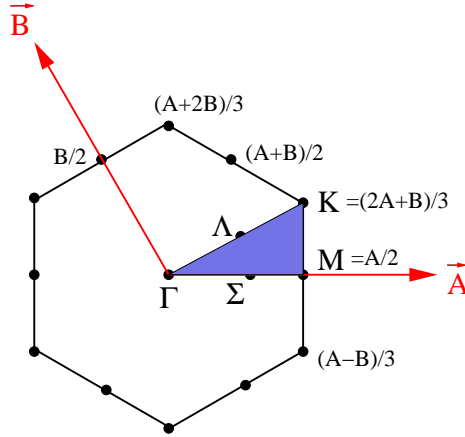

FIG. 2: The Brillouin zone, drawn by the symmetric Wigner-Seitz construction, starting from the primitive reciprocal lattice vectors $\vec{A}$ and $\vec{B}$. The "irreducible wedge" is shown shaded. Points of high symmetry are labeled.

The other diagonal element of the Hamiltonian matrix has the same form,

$$< \vec{k}B|\mathcal{H}|\vec{k}B > = \int d\vec{r}\psi_B(\vec{r})\mathcal{H}\psi_B(\vec{r})$$

and is numerically identical. Therefore these two matrix elements can be chosen to define the zero of energy. They both equal the same constant which is chosen to be zero.

Now we need to do the off-diagonal elements, which are complex conjugates of each other. They will involve integrals over functions ϕ_A and ϕ_B on different sublattices. Each A atom has three B first neighbors and vice versa. All farther neighbors are assumed to contribute nothing. Here is the answer:

$$< \vec{k}A|\mathcal{H}|\vec{k}B > = \int d\vec{r}\phi_A(\vec{r})\mathcal{H}\phi_B(\vec{r}) + e^{-i\vec{k}\cdot\vec{a}} \int d\vec{r}\phi_A(\vec{r})\mathcal{H}\phi_B(\vec{r}+\vec{a}) + e^{-i\vec{k}\cdot\vec{b}} \int d\vec{r}\phi_A(\vec{r})\mathcal{H}\phi_B(\vec{r}+\vec{b})$$

where $\vec{a}$ and $\vec{b}$ are the primitive translations shown in Fig. 1. Now each orbital ϕ_A or ϕ_B is cylindrically symmetric around the local $\hat{z}$ axis, which means that the three integrals in the previous formula all have the same numerical value, traditionally written as $-t$. The minus sign makes the value of t positive. That is because $\mathcal{H}$ is dominantly
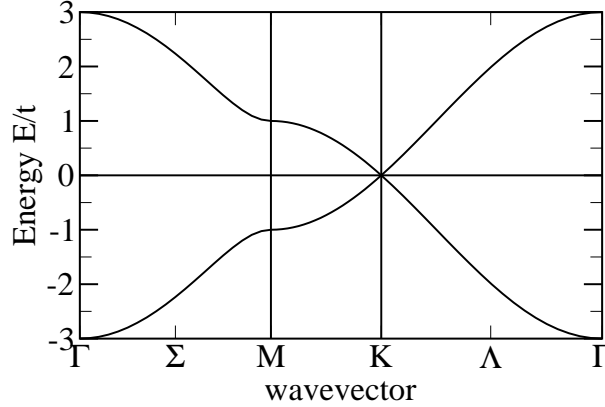
FIG. 3: Energy eigenvalues of the 2×2 π subspace tight-binding Hamiltonian for a graphite layer.

negative between two near atoms. So the off-diagonal matrix element is

$$< \vec{k}A|\mathcal{H}|\vec{k}B >= -t[1 + e^{-i\vec{k}\cdot\vec{a}} + e^{-i\vec{k}\cdot\vec{b}}].$$

The energy eigenvalues are then just plus or minus the absolute value of the off-diagonal element,

$$\epsilon(\vec{k}\pm) = \pm t|1 + e^{-i\vec{k}\cdot\vec{a}} + e^{-i\vec{k}\cdot\vec{b}}|$$

which can be written as

$$\epsilon(\vec{k}\pm) = \pm t[3 + 2\cos(\vec{k}\cdot\vec{a}) + 2\cos(\vec{k}\cdot\vec{b}) + 2\cos(\vec{k}\cdot(\vec{a}-\vec{b}))]^{1/2}$$

Now that we have the formal answer, what can we do? The eigenvalue is defined over the points $\vec{k}$ which lie in the two dimensional Brillouin zone, the unit cell defined by primitive translation vectors $\vec{A}$ and $\vec{B}$ which obey $\vec{A}\cdot\vec{a} = 2\pi$, $\vec{B}\cdot\vec{b} = 2\pi$, and $\vec{A}\cdot\vec{b} = 0$, $\vec{B}\cdot\vec{a} = 0$. The Brillouin zone is shown in Fig. 2. One convenient representation would be to draw contours of constant energy. Another is to plot the eigenvalues versus $\vec{k}$ for some special directions in $\vec{k}$-space. There are three directions of special symmetry which outline the "irreducible wedge" of the Brillouin zone. Any other point $\vec{k}$ of the zone which is not in this wedge can be rotated into a $\vec{k}$-vector inside the wedge by a symmetry operation that leaves the crystal invariant. One of these special directions is called the Σ direction, and goes from Γ (or $\vec{k} = 0$) to M (or $\vec{k} = \vec{A}/2$). Another (whose name is too obscure to be remembered) goes from M to K (or $\vec{k} = 2\vec{A}/3 + \vec{B}/3$). The third, called the Λ direction, goes from K back to Γ (or vice-versa, it is up to you how you want to plot it.) The eigenvalues are plotted along these directions in Fig. 3. It is not hard to verify that formulas along these directions are:

$$\text{along } \Sigma: \quad \epsilon(\vec{k}\pm) = \pm t[5 + 4\cos(2\pi\zeta)]^{1/2} \quad (0 < \zeta < 1)$$

$$\text{middle section}: \quad \epsilon(\vec{k}\pm) = \pm t[3 - 4\cos(2\pi\zeta/3) + 2\cos(4\pi\zeta/3)]^{1/2} \quad (0 < \zeta < 1)$$

$$\text{along } \Lambda: \quad \epsilon(\vec{k}\pm) = \pm t[3 + 2\cos(4\pi\zeta/3) + 4\cos(2\pi\zeta/3)]^{1/2} \quad (1 > \zeta > 0).$$

Note that at the K point ($\zeta = 1$ of the middle section or of the Λ line) the eigenvalues are zero because the off-diagonal element vanishes by destructive interference of the three terms. This is a robust effect that holds also in higher order versions of the theory. The gap is truly zero for a single plane, and weak three-dimensional effects turn graphite into a semimetal.