Series - Electronic and optical semiconductor devices

Series 3 - Brillain zone, band structure (introduction),

Exercise 1: Introduction	lo Brillouin Zones
--------------------------	--------------------

Number of 1st nearest neighbors: 6

2. 1 2 Namber of 2nd nearest neighbors: 6

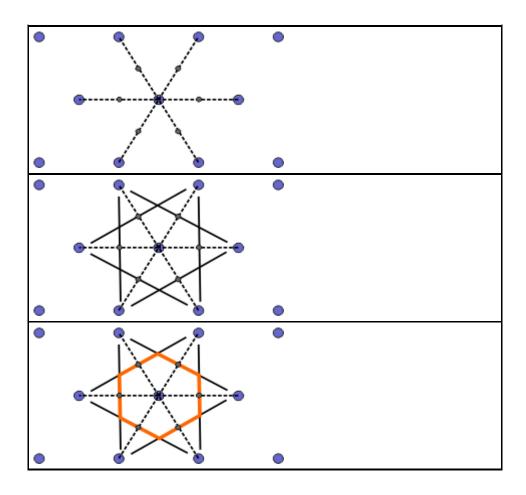
Note that the surface of Brillouin zones (BZ) in reciprocal space is identical: Space = Signare = Signare (He same remark applied to the hexagonal lettice). This is also trae in the 3D case (BZ = VIZ = VIZ;)

Some more details on BZ: Let G be a vector of the reciprocal lattice such that eight is a periodic function of the crystal lattice. Then, the

1st BZ is defined as the primitive Wigner- Seitz cell of the reciprocal lattice, i.e. the ensemble of points located closer to G'= o' than to any other point of the reciprocal latice. Since Braggplanes are the bisecting planes of the lines connecting the origin to the points of the reciprocal lattice, the 1st BZ can also be defined as the ensemble of points that can be reached from the origin without crossing any Bragg plane. More generally, the nth BZ can be defined as the ensemble of points that can be reached from the origin by crossing (n-1) Bragg planes (but not less).

Series 3, exercise 1 : Supplementary information

 1^{st} and $2^{\text{nd}}\,$ Brillouin zones in the case of a two-dimensional hexagonal lattice (case of graphene)



PHYS-433 – Semiconductor physics and light-matter interaction

Series 3 Ex. 2 – SOLUTIONS

This question is quite challenging, and may seem laborious; however, by the end you will have by hand plotted a band structure for graphene which matches very well to reality. As such, this question is designed to clear up any confusion you may still experience upon viewing 2-D or 3-D band structure diagrams. This is important, since these diagrams describe the electronic and optical properties of nearly all materials.

1. The vectors in real space $\vec{a_i}$ are related to the reciprocal vectors $\vec{b_j}$ through the relation $\vec{a_i} \cdot \vec{b_j} = 2\pi \delta_{ij}$. This allows us to determine the $\vec{b_j}$ from the set of equations

$$\vec{a_1} \cdot \vec{b_1} = 2\pi, \quad \vec{a_2} \cdot \vec{b_1} = 0, \quad \vec{a_1} \cdot \vec{b_2} = 0, \quad \vec{a_2} \cdot \vec{b_2} = 2\pi.$$

Expressing $\vec{b_1}$ and $\vec{b_2}$ in a convenient form, and using the information in the question

$$\vec{b_1} = \begin{pmatrix} b_{1x} \\ b_{1y} \end{pmatrix}, \quad \vec{b_2} = \begin{pmatrix} b_{2x} \\ b_{2y} \end{pmatrix}, \quad \vec{a_1} = a_0 \sqrt{3} \begin{pmatrix} 1/2 \\ \sqrt{3}/2 \end{pmatrix}, \quad \vec{a_2} = a_0 \sqrt{3} \begin{pmatrix} -1/2 \\ \sqrt{3}/2 \end{pmatrix},$$

we have

$$\vec{a_1} \cdot \vec{b_1} = \frac{\sqrt{3}}{2} a_0 (b_{1x} + \sqrt{3}b_{1y}) = 2\pi$$

$$\vec{a_2} \cdot \vec{b_1} = \frac{\sqrt{3}}{2} a_0 (-b_{1x} + \sqrt{3}b_{1y}) = 0$$

$$\vec{a_1} \cdot \vec{b_2} = \frac{\sqrt{3}}{2} a_0 (b_{2x} + \sqrt{3}b_{2y}) = 0$$

$$\vec{a_2} \cdot \vec{b_2} = \frac{\sqrt{3}}{2} a_0 (-b_{2x} + \sqrt{3}b_{2y}) = 2\pi.$$

Solving, we obtain

$$\vec{b_1} = \frac{2\pi}{\sqrt{3}a_0} \begin{pmatrix} 1\\ 1/\sqrt{3} \end{pmatrix}, \quad \vec{b_2} = \frac{2\pi}{\sqrt{3}a_0} \begin{pmatrix} -1\\ 1/\sqrt{3} \end{pmatrix}.$$

Using these as basis vectors, we can then map out the reciprocal lattice knowing that any sum of these vectors leads to a lattice point, resulting in a hexagonal plane. Constructing the smallest Wigner-Seitz cell of this lattice gives the 1st Brillouin zone (see Fig. 1(a)), which you should recognise from your solution to Ex. 1 part 3. This figure has the high-symmetry points Γ , M, and K indicated, and labels the closest reciprocal lattice points as A, B, ... G.

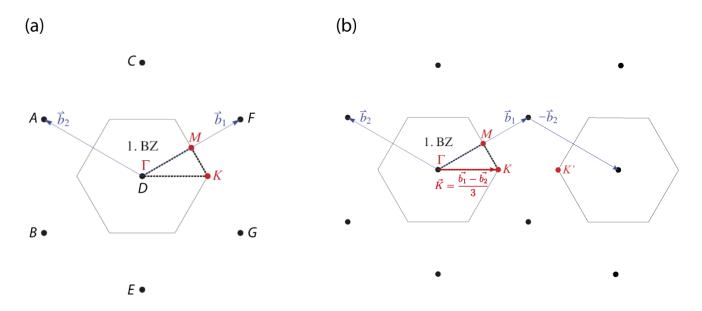


Figure 1: (a) Reciprocal lattice of graphene and first Brillouin zone. (b) Diagram showing more of the reciprocal lattice, allowing for clear visualisation of the \vec{K} expression.

2. If your spatial reasoning is good, for this part of the question it is possible to sketch the free electron band structure (also called the "empty lattice" band structure) purely by carefully considering the parabolas originating from reciprocal lattice points D, F, and G as depicted in Fig. 2 of the exercise series. However, going through the calculations is of course more rigorous. The coordinates of the high-symmetry point K are given by $\vec{K} = (\vec{b_1} - \vec{b_2})/3$ (see Fig. 1(b)). As such, we can express the path $\Gamma \to K$ as the vector $\vec{k_K}$

$$\vec{k_K} = k_K \frac{2\pi}{\sqrt{3}a_0} \begin{pmatrix} 1\\0 \end{pmatrix},$$

where $k_K = [0, \frac{2}{3}]$ is a convenient scalar to use as our plotting variable. Now, using equation (2) from the exercise series considering points D, F, and G (i.e., with $\vec{G} = 0$, $\vec{b_1}$, and $-\vec{b_2}$), we obtain

$$D: E_D(k_K) = \frac{\hbar^2}{2m_0} \left[\vec{k_K} - 0 \right]^2 = \frac{\hbar^2}{2m_0} \frac{4\pi^2}{3a_0^2} k_K^2$$

$$F: E_F(k_K) = \frac{\hbar^2}{2m_0} \left[\vec{k_K} - \frac{2\pi}{\sqrt{3}a_0} \begin{pmatrix} 1\\ 1/\sqrt{3} \end{pmatrix} \right]^2 = \frac{\hbar^2}{2m_0} \frac{4\pi^2}{3a_0^2} \left[(k_K - 1)^2 + \frac{1}{3} \right]$$

$$G: E_G(k_K) = \frac{\hbar^2}{2m_0} \left[\vec{k_K} - \frac{2\pi}{\sqrt{3}a_0} \begin{pmatrix} 1\\ -1\sqrt{3} \end{pmatrix} \right]^2 = \frac{\hbar^2}{2m_0} \frac{4\pi^2}{3a_0^2} \left[(k_K - 1)^2 + \frac{1}{3} \right]$$

here we can already note that the branches arising from F and G are degenerate. Meanwhile, the coordinates of the high-symmetry point M are given by $\vec{M} = \vec{b_1}/2$, so the path $\Gamma \to M$ is expressed as $\vec{k_M}$

$$\vec{k_M} = k_M \frac{2\pi}{\sqrt{3}a_0} \begin{pmatrix} \sqrt{3}/2\\ 1/2 \end{pmatrix},$$

with $k_M = [0, \frac{1}{\sqrt{3}}]$. Applying the same treatment, this leads to

$$\begin{split} D:E_D(k_M) &= \frac{\hbar^2}{2m_0} \left[\vec{k_M} - 0 \right]^2 \\ &= \frac{\hbar^2}{2m_0} \frac{4\pi^2}{3a_0^2} k_M^2 \\ F:E_F(k_M) &= \frac{\hbar^2}{2m_0} \left[\vec{k_M} - \frac{2\pi}{\sqrt{3}a_0} \left(\frac{1}{1/\sqrt{3}} \right) \right]^2 \\ &= \frac{\hbar^2}{2m_0} \frac{4\pi^2}{3a_0^2} \left[\left(\frac{k_M\sqrt{3}}{2} - 1 \right)^2 + \left(\frac{k_M}{2} - \frac{1}{\sqrt{3}} \right)^2 \right] \\ G:E_G(k_M) &= \frac{\hbar^2}{2m_0} \left[\vec{k_M} - \frac{2\pi}{\sqrt{3}a_0} \left(\frac{1}{-1\sqrt{3}} \right) \right]^2 \\ &= \frac{\hbar^2}{2m_0} \frac{4\pi^2}{3a_0^2} \left[\frac{3k_M^2}{4} + \left(\frac{k_M}{2} - \frac{2}{\sqrt{3}} \right)^2 \right] \end{split}$$

Note that while it is possible to choose any plotting scalar, is is easiest to do the calculations with this choice of k_K and k_M . We can now plot these expressions to obtain most of the band structure asked for in the question (see Fig. 2(a)). Here we can already see that there exists a three-fold degeneracy in the lowest energy at K, and a two-fold degeneracy in the lowest energy at M.

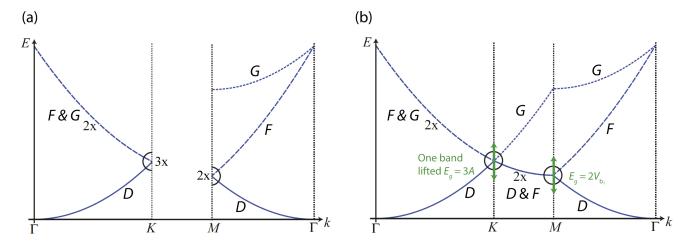


Figure 2: (a) Free electron dispersion for the graphene lattice, calculated from the equations above. Degeneracies are indicated, along with which reciprocal lattice point (\vec{G}) each branch arises from. The lowest energy states at K and M are circled. (b) Full free electron dispersion, including along $K \to M$ by deduction. Lowest energy states at K and M are circled, and green arrows indicate the magnitude of the band gaps that will open here on including the effect of a small crystal potential.

Finally, we need to consider the path $K \to M$. Of course we could carry out the same calculations as above, but in reality this would be a painstaking route to achieve something that is quite obvious to sketch. We simply have to note that the branches need to be joined across this path, since there can't be any discontinuities. To get the exact forms correct, we realise that $K \to M$ is part of a perpendicular bisector of D and F, so there must be a turning point

at the midpoint M for these branches. For G, we can realise that $K \to M$ is just a continuation of the parabola, so there will be no turning point. This gives the final free electron dispersion along $\Gamma \to K \to M \to \Gamma$ as shown in Fig. 2(b).

3. We have the secular equation:

$$\left(\frac{\hbar^2(\vec{k}-\vec{G})^2}{2m_0} - E_G(\vec{k})\right)C_{\vec{k}-\vec{G}} + \sum_{\vec{G}'} V_{\vec{G}'-\vec{G}}C_{\vec{k}-\vec{G}'} = 0.$$

Since we are interested in the new energies at the M point, we use the value of \vec{k} here: $\vec{k} = \frac{\vec{b_1}}{2}$. We consider the values of \vec{G} which lead to the degeneracy at the M point, i.e., $\vec{G} = \vec{0}$ and $\vec{G} = \vec{b_1}$. The question then arises: which \vec{G}' values should we consider? The secular equation implies \vec{G}' can be any reciprocal lattice vector, and so we should have an infinite sum. However, to first order we only need to consider the reciprocal lattice vectors which lead to the degeneracy, i.e., the same values we are using for \vec{G} : $\vec{G}' = \vec{0}$ and $\vec{G}' = \vec{b_1}$. This is because terms in the sum over \vec{G}' which are associated with free electron energies at M far away from the degenerate energy we are considering will only have contribution $O(V^2)$, while those associated with free electron energies close to the degeneracy will have contribution O(V) (see Ashcroft & Mermin, beginning of Section 9). Since we are solving for a weak potential we can neglect the $O(V^2)$ terms. Using our values of \vec{G} and \vec{G}' we obtain (first equation with $\vec{G} = \vec{0}$, second with $\vec{G} = \vec{b_1}$)

$$\left(\frac{\hbar^2 |\vec{b_1}|^2}{8m_0} - E\left(\frac{\vec{b_1}}{2}\right)\right) C_{\frac{1}{2}\vec{b_1}} + V_0 C_{\frac{1}{2}\vec{b_1}} + V_{\vec{b_1}} C_{-\frac{1}{2}\vec{b_1}} = 0$$

$$\left(\frac{\hbar^2 |\vec{b_1}|^2}{8m_0} - E\left(\frac{\vec{b_1}}{2}\right)\right) C_{-\frac{1}{2}\vec{b_1}} + V_{-\vec{b_1}} C_{\frac{1}{2}\vec{b_1}} + V_0 C_{-\frac{1}{2}\vec{b_1}} = 0$$

As stated in the exercise, $V_{\vec{G}} = V_{-\vec{G}}$ (due to the potential of graphene being real-valued and exhibiting inversion-symmetry). Meanwhile $V_{\vec{0}}$ can be chosen to be 0, since the first Fourier coefficient is simply the spatial average of the potential, and we are at liberty to change the potential energy by an additive constant. Using this, and rearranging, we obtain

$$\left(\frac{\hbar^2 |\vec{b_1}|^2}{8m_0}\right) C_{\frac{1}{2}\vec{b}_1} + V_{\vec{b}_1} C_{-\frac{1}{2}\vec{b}_1} = C_{\frac{1}{2}\vec{b}_1} E\left(\frac{\vec{b}_1}{2}\right)$$

$$\left(\frac{\hbar^2 |\vec{b_1}|^2}{8m_0}\right) C_{-\frac{1}{2}\vec{b_1}} + V_{\vec{b_1}} C_{\frac{1}{2}\vec{b_1}} = C_{-\frac{1}{2}\vec{b_1}} E\left(\frac{\vec{b_1}}{2}\right)$$

These two equations can be expressed simultaneously in matrix form

$$\begin{pmatrix} \hbar^2 |\vec{b_1}|^2 / 8m_0 & V_{\vec{b}_1} \\ V_{\vec{b}_1} & \hbar^2 |\vec{b_1}|^2 / 8m_0 \end{pmatrix} \begin{pmatrix} C_{\frac{1}{2}\vec{b}_1} \\ C_{-\frac{1}{2}\vec{b}_1} \end{pmatrix} = E \begin{pmatrix} \vec{b_1} \\ 2 \end{pmatrix} \begin{pmatrix} C_{\frac{1}{2}\vec{b}_1} \\ C_{-\frac{1}{2}\vec{b}_1} \end{pmatrix}$$

which is clearly recognisable as an eigenvalue equation, and solving the characteristic equation of the left matrix we see the off-diagonal elements lift the degeneracy leading to a bandgap at M of $2|V_{\vec{b}_i}|$ (see Fig. 2(b)).

$$E_M^{\pm} = \frac{\hbar^2 |\vec{b_1}|^2}{8m_0} \pm |V_{\vec{b_1}}|$$

4. The initial part of this trickier question follows the exact same method as part 3: we apply the secular equation with the value of $\vec{k} = (\vec{b_1} - \vec{b_2})/3$ and the reciprocal lattice vectors which lead to the degeneracy $\vec{G} = 0$, $\vec{b_1}$, and $-\vec{b_2}$, with the same values of \vec{G}' . This leads to

$$\begin{pmatrix} \hbar^2 |\vec{K}|^2 / 2m_0 & V_{\vec{b}_1} & V_{\vec{b}_2} \\ V_{\vec{b}_1} & \hbar^2 |\vec{K}|^2 / 2m_0 & V_{\vec{b}_1 + \vec{b}_2} \\ V_{\vec{b}_2} & V_{\vec{b}_1 + \vec{b}_2} & \hbar^2 |\vec{K}|^2 / 2m_0 \end{pmatrix} \begin{pmatrix} C_{\vec{K}} \\ C_{\vec{K} - \vec{b}_1} \\ C_{\vec{K} + \vec{b}_2} \end{pmatrix} = E(\vec{K}) \begin{pmatrix} C_{\vec{K}} \\ C_{\vec{K} - \vec{b}_1} \\ C_{\vec{K} + \vec{b}_2} \end{pmatrix}$$

after using our usual rules $(V_{\vec{G}} = V_{-\vec{G}}, V_{\vec{0}} = 0)$ and realising that $|\vec{K}|^2 = |\vec{K} - \vec{b}_1|^2 = |\vec{K} + \vec{b}_2|^2$ is exactly the degeneracy we showed in part 2. We could apply the characteristic equation to the left matrix at this stage, but doing so would be uninformative when we have a lot of terms with no clear relation $(V_{\vec{b}_1}, V_{\vec{b}_2}, V_{\vec{b}_1 + \vec{b}_2})$. As such, to simplify the matrix, let us inspect these Fourier coefficients more carefully

$$V_{\vec{G}} = \frac{1}{\Omega} \int_{PC} V(\vec{r}) e^{-i\vec{G}\cdot\vec{r}} d^2r,$$

where Ω is the area of the real-space primitive unit cell (PC) and \vec{r} is a real-space vector. $V(\vec{r}) = V_c(\vec{r} - \vec{r_A}) + V_c(\vec{r} - \vec{r_B})$ where V_c refers to the potential of a single carbon atom, and $\vec{r_A}$ and $\vec{r_B}$ refer to the vectors leading to the atoms labelled A and B in Fig. 1(a) of the exercise series (hence using the origin in this figure $\vec{r_A} = (0, a_0)$ and $\vec{r_B} = \vec{0}$). Using this expression for $V(\vec{r})$

$$V_{\vec{G}} = \frac{1}{\Omega} \int_{PC} V_c(\vec{r} - \vec{r_A}) e^{-i\vec{G} \cdot \vec{r}} d^2r + \frac{1}{\Omega} \int_{PC} V_c(\vec{r} - \vec{r_B}) e^{-i\vec{G} \cdot \vec{r}} d^2r.$$

Now that the integrals are separate, we can redefine $\vec{r} = \vec{r'} + \vec{r_A}$ in the first integral, and $\vec{r} = \vec{r'} + \vec{r_B}$ in the second (since any translation of the unit cell will not affect the integral values due to the periodicity)

$$V_{\vec{G}} = e^{-i\vec{G}\cdot\vec{r_A}} \frac{1}{\Omega} \int_{PC} V_c(\vec{r'}) e^{-i\vec{G}\cdot\vec{r'}} d^2r' + e^{-i\vec{G}\cdot\vec{r_B}} \frac{1}{\Omega} \int_{PC} V_c(\vec{r'}) e^{-i\vec{G}\cdot\vec{r'}} d^2r'$$

$$V_{\vec{G}} = A(e^{-i\vec{G}\cdot\vec{r_A}} + e^{-i\vec{G}\cdot\vec{r_B}}),$$

where A is equal to the value of the integral term. For the $V_{\vec{G}}$ we are considering A is a constant as long as the carbon potential V_c is circularly symmetric (a reasonable assumption) since $|\vec{b_1}| = |\vec{b_2}| = |\vec{b_1} + \vec{b_2}|$. If we change our origin to the midpoint between atoms A and B, we can set $\vec{r_A} = -\vec{r_B} = \vec{d}/2 = (0, a_0/2)^1$

$$V_{\vec{G}} = A(e^{-i\vec{G}\cdot\vec{d}/2} + e^{i\vec{G}\cdot\vec{d}/2})$$

$$V_{\vec{G}} = 2A\cos\left(\frac{\vec{G}\cdot\vec{d}}{2}\right).$$

Since A is independent of \vec{G} , we can now compute the relations between all the unknown Fourier coefficients

$$V_{\vec{b}_1} = 2A\cos\left(\frac{1}{2}\frac{2\pi}{\sqrt{3}a_0}\begin{pmatrix}1\\1/\sqrt{3}\end{pmatrix}\cdot\begin{pmatrix}0\\a_0\end{pmatrix}\right) = 2A\cos\left(\frac{\pi}{3}\right) = A$$

$$V_{\vec{b}_2} = 2A\cos\left(\frac{1}{2}\frac{2\pi}{\sqrt{3}a_0}\begin{pmatrix}-1\\1/\sqrt{3}\end{pmatrix}\cdot\begin{pmatrix}0\\a_0\end{pmatrix}\right) = 2A\cos\left(\frac{\pi}{3}\right) = A$$

$$V_{\vec{b}_1+\vec{b}_2} = 2A\cos\left(\frac{1}{2}\frac{2\pi}{\sqrt{3}a_0}\begin{pmatrix}0\\2/\sqrt{3}\end{pmatrix}\cdot\begin{pmatrix}0\\a_0\end{pmatrix}\right) = 2A\cos\left(\frac{2\pi}{3}\right) = -A$$

which simplifies the original eigenvalue equation to

$$\begin{pmatrix} \hbar^{2} |\vec{K}|^{2} / 2m_{0} & A & A \\ A & \hbar^{2} |\vec{K}|^{2} / 2m_{0} & -A \\ A & -A & \hbar^{2} |\vec{K}|^{2} / 2m_{0} \end{pmatrix} \begin{pmatrix} C_{\vec{K}} \\ C_{\vec{K} - \vec{b}_{1}} \\ C_{\vec{K} + \vec{b}_{2}} \end{pmatrix} = E(\vec{K}) \begin{pmatrix} C_{\vec{K}} \\ C_{\vec{K} - \vec{b}_{1}} \\ C_{\vec{K} + \vec{b}_{2}} \end{pmatrix}$$

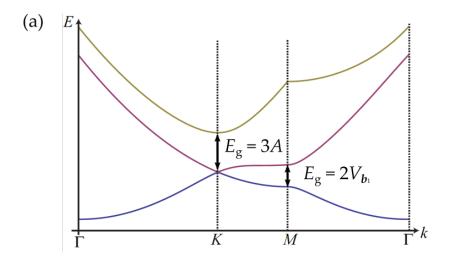
and solving (which you may wish to do in Mathematica or another program to save you time), we obtain

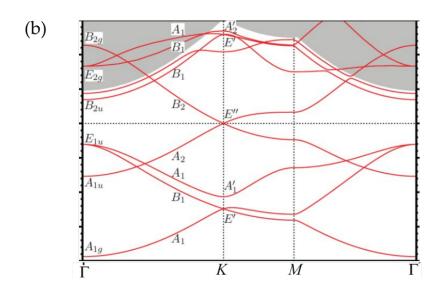
$$E_{\vec{K}}^{D,F} = \frac{\hbar^2 |\vec{K}|^2}{2m_0} - 2A, \quad E_{\vec{K}}^G = \frac{\hbar^2 |\vec{K}|^2}{2m_0} + A$$

where we can see the particular symmetry of graphene (which we accounted for with our expressions of the different $V_{\vec{G}}$) has kept a two-fold degeneracy at the lower energy, and created a bandgap of 3A to the next energy. Putting this all together, we can finally plot the band structure of graphene with a weak crystal potential in Fig. 3(a).

Extra information: Fig. 3(b) shows the full band-structure calculated for graphene using density functional theory. You should be able to spot the bands you have calculated by-hand in this much more sophisticated diagram! In these band structures, each separate band is fully occupied by two electrons in the unit cell (see Ashcroft & Mermin)². Since graphene has two carbon atoms per unit cell, it has a total of 12 electrons per unit cell. The 1s electrons can be neglected since they are so tightly bound (the 1s bands would be far below the lowest band in Fig. 3(b)), leaving us with 8 electrons. These fill the lowest four bands, placing the Fermi energy at the horizontal dashed line in Fig. 3(b)—directly at the degeneracies, known as Dirac points. This is emphasised in Fig. 3(c), which expands our condensed 2-D band structure along the path $\Gamma \to K \to M \to \Gamma$ to full 3-D for the bands below and above the Fermi energy (the valence band and conduction band), showing the Dirac cones clearly.

These Dirac cones are the source of the unusual electron transport properties of graphene. For instance, our usual expression for the effective mass of carriers in semiconductors is not valid for graphene, since we have a non-parabolic E(k) dispersion around the Fermi energy.





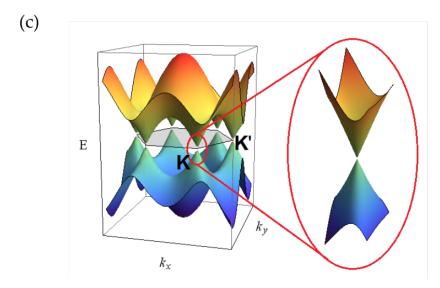


Figure 3: (a) Nearly-free electron dispersion for the graphene lattice. The lowest energy state at K remains degenerate and the dispersion around this point is linear (Dirac cone), whereas at M the degeneracy is fully lifted. (b) Full graphene band structure as calculated by density functional theory. (c) Full 3-D structure of the highest occupied band (valence band) and lowest unoccupied band (conduction band), highlighting the presence of the Dirac cones at the K and K' points.

Instead the carriers are treated as massless fermions, which leads to various quantum Hall effects, and ultra-high carrier mobility.

Overall, this question sheet should have made the origins of band structures clear. Every band structure has its origins in that same method: (i) the reciprocal lattice is calculated from the real-space lattice of the crystal, (ii) the free electron band structure is found by repeating a parabolic dispersion at every reciprocal lattice point, and (iii) applying the crystal potential leads to degeneracies in the free electron band structure being lifted. The degree to which bands are shifted from their free electron values is related to the strength of the potential, while whether a degeneracy is lifted or not will depend on the symmetry of the crystal. The main complexity of visualising band structures is due to the number of dimensions required to plot them. For 3-D materials, the band structure is four dimensional (k_x, k_y, k_z, E) which means they can only ever be partially visualised by plotting E(k) along particular high-symmetry paths, as we have done in this question (or as constant-energy surfaces in 3-D). While this is a bit unsatisfying, you should be more comfortable with this concept after completing this exercise.

¹ For those familiar with crystallography, calculating $V_{\vec{G}}$ is identical to calculating the 2-D structure factor F_{hk} :

$$F_{hk} = \sum_{j=1}^{N} f_j e^{[-2\pi i(hx_j + ky_j)]}$$

where the sum is over all atoms in the unit cell, x_j, y_j are the positional coordinates of the j-th atom, and f_j is the atomic form factor of the j-th atom (our A). The coordinate system used is the natural one for a crystal: real-space vector $\vec{r} = [x_j y_j] = x_j \vec{a_1} + y_j \vec{a_2}$ and reciprocal vector $\vec{G} = (hk) = h\vec{b_1} + k\vec{b_2}$ (i.e., Miller indices).

² This fact is why a material with an odd number of electrons per unit cell is **always** a metal, never a semiconductor or an insulator. You should be able to justify this to yourself. Extra question: is a material with an even number of electrons per unit cell always a semiconductor/insulator?