Lecture 12
Band Gap

Today:

1. Series solutions to the cosine potential Hamiltonian.

2. Derivation of the band-diagrams – the graphical representation of single electron solutions

3. Analytical expression for bandgaps

Questions you should be able to address after today's lecture:

1. What is the resulting eigenvalues problem and how are its solutions related to the eigenfunction of the periodic Hamiltonian?

2. How to derive an approximate analytical expression for bandgaps?

3. How is the electronic distribution related to the energy of a solution?

4. How does the bandgap scale with the potential parameters?

5. Bandgap trends in the periodic table.

6. How are the band gaps measured?

References:

1. "Introduction to Solid State Physics", Kittel
2. "Solid State Physics", Ashcroft and Mermin

Last time we have derived the Central Equation:

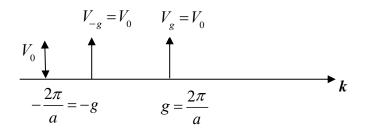$$\frac{\hbar^2}{2m}k^2 C_k + \sum_G V_G C_{k-G} = E C_k$$

This equation relates all the coefficients C_k that are separated by an inverse lattice vector to each other forming a set of coupled algebraic equations (which in essence is a recursion formula). Let's consider a simple example to clarify the importance of central equation:

Example: Cosine Potential

For simplicity lets consider a periodic potential, which is a simple cosine:

$$V(x) = 2V_0 \cos\left(\frac{2\pi}{a}x\right) = 2V_0 \cos gx = V_0\left(e^{igx} + e^{-igx}\right) = V_0 e^{igx} + V_0 e^{-igx}$$

We can place the two Fourier coefficients of the potential on the same k axis we used previously as both have inverse length units:

We have used the definition of the inverse lattice vector G and defined a "primitive" inverse lattice vector g: $G = n\frac{2\pi}{a} = ng$. Then system of the equations above can be written as:

$$\frac{\hbar^2}{2m}k^2 C_k + VC_{k-g} + VC_{k+g} = E C_k$$

$$\frac{\hbar^2}{2m}(k-g)^2 C_{k-g} + VC_{k-2g} + VC_k = E C_{k-g}$$

$$\frac{\hbar^2}{2m}(k+g)^2 C_{k+g} + VC_k + VC_{k+2g} = E C_{k+g}$$

....

Let's rearrange the equations above and add a couple more just for you to see the pattern:

...

$$V_0 C_{k-3g} + \frac{\hbar^2}{2m}(k-2g)^2 C_{k-2g} + V_0 C_{k-g} = E C_{k-2g}$$
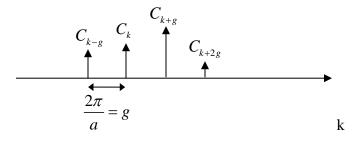
$$V_0 C_{k-2g} + \frac{\hbar^2}{2m}(k-g)^2 C_{k-g} + V_0 C_{k} = E C_{k-g}$$

$$V_0 C_{k-g} + \frac{\hbar^2}{2m}k^2 C_{k} + V_0 C_{k+g} = E C_{k}$$

$$V_0 C_{k} + \frac{\hbar^2}{2m}(k+g)^2 C_{k+g} + V_0 C_{k+2g} = E C_{k+g}$$

$$V_0 C_{k+g} + \frac{\hbar^2}{2m}(k+2g)^2 C_{k+2g} + V_0 C_{k+3g} = E C_{k+2g}$$

....

$$\frac{2\pi}{a} = g$$

k

Since **g** and **V** are known, once **k** is specified then a set (infinite) of coupled algebraic equations is completely determined, thus the problem has been transformed from a differential operator to a matrix operator eigenvalues problem.

For a given **k**, the eigenvectors define the energy eigenfunctions, while the eigenvalues are the corresponding energies.

$$
\begin{bmatrix}
\frac{\hbar^2}{2m}(k-2g)^2 & V & 0 & 0 & 0 \\
V & \frac{\hbar^2}{2m}(k-g)^2 & V & 0 & 0 \\
0 & V & \frac{\hbar^2}{2m}k^2 & V & 0 \\
0 & 0 & V & \frac{\hbar^2}{2m}(k+g)^2 & V \\
0 & 0 & 0 & V & \frac{\hbar^2}{2m}(k+2g)^2
\end{bmatrix}
\begin{pmatrix}
C_{k-2g} \\
C_{k-g} \\
C_{k} \\
C_{k+g} \\
C_{k+2g}
\end{pmatrix}
= E_k
\begin{pmatrix}
C_{k-2g} \\
C_{k-g} \\
C_{k} \\
C_{k+g} \\
C_{k+2g}
\end{pmatrix}
$$

The matrix above was truncated after 5 terms (the effects of this truncation will be explored shortly) keep in mind that the full matrix is infinite as all the coefficients separated by an inverse lattice vector are related.

Discussion:

1. The eigenfunctions have the following form:

$$\varphi_{k,E_{k,n}}(x)=...C_{k-2g}e^{i(k-2g)x}+C_{k-g}e^{i(k-g)x}+C_ke^{ikx}+C_{k+g}e^{i(k+g)x}+C_{k+2g}e^{i(k+2g)x}+...$$

$$=\sum_G C_{k-G}e^{i(k-G)x}=e^{ikx}\underbrace{\sum_G C_{k-G}e^{-iGx}}_{f_k(x)}$$

Associated with a corresponding eigenvalue: $E_{k,n}$

Which is just a restatement of Bloch's Theorem, where *f(x)* is a periodic function with the periodicity of the lattice.

This reflects the fact that once a *specific k* is defined all the coefficients corresponding to that k can be generated thus defining a solution labeled by *k*.
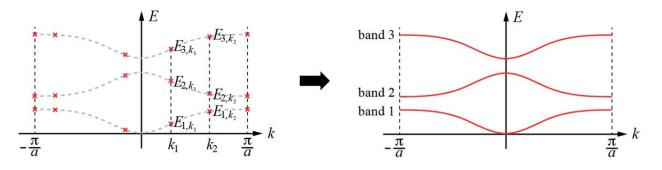
2. By considering the full matrix form you can convince yourself that choosing two specific *k*'s that are separated by any G would generate the exact same matrix yielding identical solutions: $\varphi_{k+G}(x)=\varphi_k(x)$

This means that the eigenvectors and energy eigenvalues are periodic in *G*: $E_{k+G,n}=E_{k,n}$

One can therefore choose *k*'s that are restricted to the first BZ: $-\dfrac{\pi}{a}\le k\le\dfrac{\pi}{a}$ and completely capture all the solutions!

3. A hierarchy of solutions is generated at each *k* point. These solutions are called *bands* and labeled by an index *n*=1, 2, 3... thus one needs to specify both *k* and the band index *n* to completely specify the solution.

4. The number of nonzero matrix elements in the equations is determined by the number of Fourier components that are needed to express the potential function.

Band gaps: what happens near the edge of the BZ?

The edge of the BZ is defined by: $k = \pm \dfrac{\pi}{a} = \pm \dfrac{g}{2}$

Review the master equation for $V(x) = 2V_0 \cos gx$:

$$\left(\frac{\hbar^2}{2m} k^2 - E \right) C_k + VC_{k-g} + VC_{k+g} = 0$$

And consider only the first terms in the sum and focus on C_k and C_{k+g}:

$$
\left.
\begin{aligned}
VC_{k-g} + \left(\frac{\hbar^2}{2m} k^2 - E \right) C_k + VC_{k+g} &= 0 \\
VC_k + \left(\frac{\hbar^2}{2m} (k+g)^2 - E \right) C_{k+g} + VC_{k+2g} &= 0
\end{aligned}
\right\}
\Rightarrow
\begin{pmatrix}
\left(\dfrac{\hbar^2}{2m} k^2 - E \right) & V \\
V & \left(\dfrac{\hbar^2}{2m} (k+g)^2 - E \right)
\end{pmatrix}
\begin{pmatrix} C_k \\ C_{k+g} \end{pmatrix}
=
\begin{pmatrix} 0 \\ 0 \end{pmatrix}
$$

For $k = -\dfrac{g}{2}$ we can write:
$$
\begin{pmatrix}
\left(\dfrac{\hbar^2}{2m} \left(-\dfrac{g}{2} \right)^2 - E \right) & V \\
V & \left(\dfrac{\hbar^2}{2m} \left(\dfrac{g}{2} \right)^2 - E \right)
\end{pmatrix}
\begin{pmatrix} C_{-g/2} \\ C_{g/2} \end{pmatrix}
=
\begin{pmatrix} 0 \\ 0 \end{pmatrix}
$$

Analogously for $k = \dfrac{g}{2}$:
$$
\begin{pmatrix}
\left(\dfrac{\hbar^2}{2m} \left(\dfrac{g}{2} \right)^2 - E \right) & V \\
V & \left(\dfrac{\hbar^2}{2m} \left(-\dfrac{g}{2} \right)^2 - E \right)
\end{pmatrix}
\begin{pmatrix} C_{g/2} \\ C_{-g/2} \end{pmatrix}
=
\begin{pmatrix} 0 \\ 0 \end{pmatrix}
$$

The only way this homogeneous system of equation can have a non-zero solution; the determinant of the matrix has to be equal to zero.

$$
\begin{vmatrix}
\dfrac{\hbar^2}{2m} \left(\dfrac{g}{2} \right)^2 - E & V \\
V & \dfrac{\hbar^2}{2m} \left(-\dfrac{g}{2} \right)^2 - E
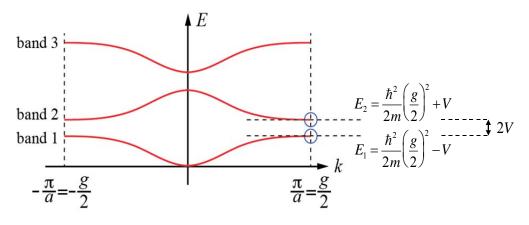\end{vmatrix}
= 0
$$

Then the energy eigenvalues are: $E_\pm = \dfrac{\hbar^2}{2m} \left(\dfrac{g}{2} \right)^2 \pm V$

Thus the energy has two roots, which are separated by $\Delta E = E_+ - E_- = 2V$.

Keep in mind that this result is valid for the particular form of the potential that we used:

$$V(x) = 2V \cos gx = V\left(e^{igx} + e^{-igx}\right)$$

So we are able to predict the existence of a band gap, which is approximately proportional to the magnitude of the periodic potential.

The form of the eigenfunctions near the band edge is: $u_{\frac{g}{2}}(x) = e^{\frac{igx}{2}} \mp e^{-\frac{igx}{2}}$

One solution gives the eigenfunctions near the bottom of the gap the other gives it near the top of the gap. The differences in energies can be explained by looking at the charge distribution relative to the "ion" positions.

$$u_-(x) = e^{i\frac{\pi x}{a}} + e^{i\frac{\pi x}{a}} = 2\cos\frac{\pi x}{a}$$

$$u_+(x) = e^{i\frac{\pi x}{a}} - e^{i\frac{\pi x}{a}} = 2i\sin\frac{\pi x}{a}$$

The two standing waves concentrate their electrons in different regions – the probability density for a particle is: $u_-^*(x)u_-(x) = \cos^2\frac{\pi x}{a}$

This function "piles" electrons (with negative charge) on the positive ions centered at $x=0$, a, $2a$… where the potential energy is lowest.

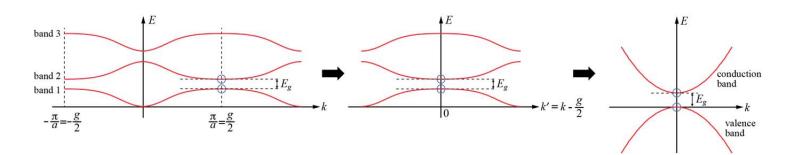For the other standing wave: $u_+^*(x)u_+(x) = \sin^2\frac{\pi x}{a}$

This function concentrates the electrons precisely in the middle between the ions, i.e. on $x=a/2$, $3a/2$…
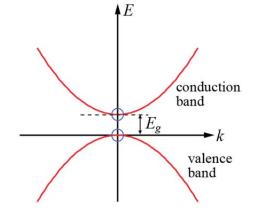
Figure removed due to copyright restrictions. Fig. 3: Kittel, Charles. *Introduction to Solid State Physics*. 8th ed. Wiley, 2004, p. 178.

What information can we get from the band diagrams?

Let's focus on the band shape near the band edges. If we shift the **k**-axis by **g/2** so that the band edge is at **k'=0** (see the picture bellow), then we can approximate the first two bands with parabolas near the band edge:

Then the simplest form of the band diagram has the following shape:

1. Allowed and forbidden bands. Identification of the gap energy.

2. Slope of the bands – group velocity.

3. Curvature of bands – effective mass.

1. <u>Magnitude of the band gaps</u>

Crystal	Gap	E_g, eV		Crystal	Gap	E_g, eV	
		0 K	300 K			0 K	300 K
Diamond	i	5.4		HgTe[a]	d	-0.30	
Si	i	1.17	1.11	PbS	d	0.286	0.34-0.37
Ge	i	0.744	0.66	PbSe	i	0.165	0.27
αSn	d	0.00	0.00	PbTe	i	0.190	0.29
InSb	d	0.23	0.17	CdS	d	2.582	2.42
InAs	d	0.43	0.36	CdSe	d	1.840	1.74
InP	d	1.42	1.27	CdTe	d	1.607	1.44
GaP	i	2.32	2.25	ZnO		3.436	3.2
GaAs	d	1.52	1.43	ZnS		3.91	3.6
GaSb	d	0.81	0.68	SnTe	d	0.3	0.18
AlSb	i	1.65	1.6	AgCl		-	3.2
SiC(hex)	i	3.0	-	AgI		-	2.8
Te	d	0.33	-	Cu_2O	d	2.172	-
ZnSb		0.56	0.56	TiO_2		3.03	-

Observations:

1. **Larger atoms $\Rightarrow$ smaller bandgaps.** Why?

   Larger atoms have more electron shells and the ionic charge is screened from the free flying electron. Consequently electrons will be experiencing lower effective potential in lattices of large atoms.

   From our simple model we have seen that in the simplest case of cosine potential the bandgap is proportional to the amplitude of the potential. Consequently: $V\downarrow \Rightarrow E_g = 2V\downarrow$.

   While in real crystals the relationship between the potential and the bandgap is more complex, the same logic applies:
   larger atoms $\Rightarrow$ smaller potential $\Rightarrow$ smaller bandgap.

2. **Energy gaps span 0.08-5.4 eV**.

3. **The wider the gap the heavier the electron.**

Optical measurement of the bandgap.

Bandgaps are related to the optical transparency and electrical conductivity of the material. Electronic bandgaps are generally measured by optical absorption spectroscopy:
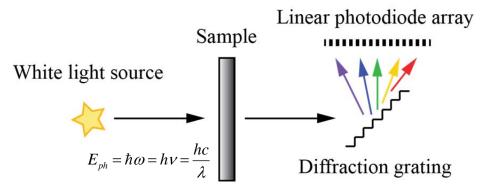
Figure removed due to copyright restrictions. Fig. 4-5: Kittel, Charles. *Introduction to Solid State Physics*. 8th ed. Wiley, 2004, p. 202.

(from "Introduction to Solid State Physics" by Kittel)

While in class we have discussed a very simple cosine potential, in real materials potentials look significantly more complicated leading to more complicated forms of band diagram. For some materials it turns out that the valence band maximum does not coincide with the conduction band minimum in k. This means that in order for an optical transition to happen (i.e. for electron to jump from valence to conduction band with a help of a photon), the electron needs to also obtain extra crystal momentum.
Where would the electron get extra crystal momentum? – Lattice vibrations, phonons have very low energy comparing to photons but large crystal momentum. However this means that for absorption to happen the photon, the electron and the phonon need to all meet at the same place and time, which is a low probability occurrence and hence the absorption is poor for the indirect bandgap materials between E_g and E_{vert} (veritical gap – the energy difference between the conduction and valence bands at $k = 0$).

The experimental setup:

Linear photodiode array

Sample

White light source

$E_{ph} = \hbar\omega = h\nu = \dfrac{hc}{\lambda}$

Diffraction grating

3.024 Electronic, Optical and Magnetic Properties of Materials
Spring 2013