PX446: Condensed Matter Physics II

Chung Xu

Last edited: April 21, 2025

Contents

0.1 Introduction, Credits and Notation

0.1.1 Credits

A big thank you to

- Chung.

0.1.2 How to use the guide

Anything in a white box with a blue title frame like

Title here

Content placed in these contains information or principles that **must be remembered for the exams**. This applies information placed in the same colour box, but without a title!

Any equation which contains a regular, black box like $\boxed{this}$ is important information but you shouldn't need to memorise it.

This guide is very detailed, almost like its own set of lecture notes. It aims to answer as many questions as possible regarding both the maths and the physics in this module. Any parts non-examinable will be explicitly marked non-examinable. Beware that this can change over the years if this guide isn't updated and therefore check with the lecturer.

0.1.3 Tips

- Strong chance (with the removal of magnetism from this year's course) that the majority of the module will be examined.
- Ensure you are familiar with PX3A3 Electrodynamics (gauge transform, Maxwell's equations) and thermodynamical quantities from PX265 Thermal Physics II or PX284 SMETO.
- For the optical properties, Fox's book *Optical Properties of Solids* has quite good practice questions
- For superconductivity, Blundell's, *A Short Introduction to Superconductivity*, also has good practice problems
- As always: **past papers go hard**!
- There are a lot of figures in these notes, I recommend being able to interpret and understand them, maybe even recall the figures relating to experimental setups.

Chapter 1

Superconductivity

Some notation for this chapter (please note, some changes later in this chapter to keep consistent with the lecture course, **watch out**):

- R, resistance; ρ, resistivity; $\sigma = 1/\rho$: conductivity.
- $\mathbf{E}$: electric field; $\mathbf{B}$, magnetic flux density; $\mathbf{H}$, magnetic field; $\mathbf{J}$, electric current density.

1.1 Basic Properties of Superconductors

There are multiple characteristics that describe whether a material is in a superconducting state (click the items to visit their respective subsections):

Properties of Superconductors

- **Zero** electrical resistance below a **critical temperature** T_c. hence, $R, \rho = 0$. By Ohm's law, $\mathbf{E} = \rho\mathbf{J} \implies \mathbf{E} = 0$ in the **bulk of the superconductor**.
- Observes the **Meissner effect**.
- Has surface currents.
- Has a critical field.
- Has a critical current.
- May have persistent currents.
- Superconductors have poor thermal conductivity, since the charge carriers do not carry heat or entropy.

1.1.1 Meissner effect

If a material in the superconducting state is subjected to a magnetic field (say, using an electromagnet to raise the external field from 0 T to a field $B_{\text{ext}} < B_c$, a critical field), a supercurrent is induced at the very surface of the superconductor which exclude and expel B_{ext}, such that

Meissner effect

$\mathbf{B} = 0$ in the bulk of a superconductor

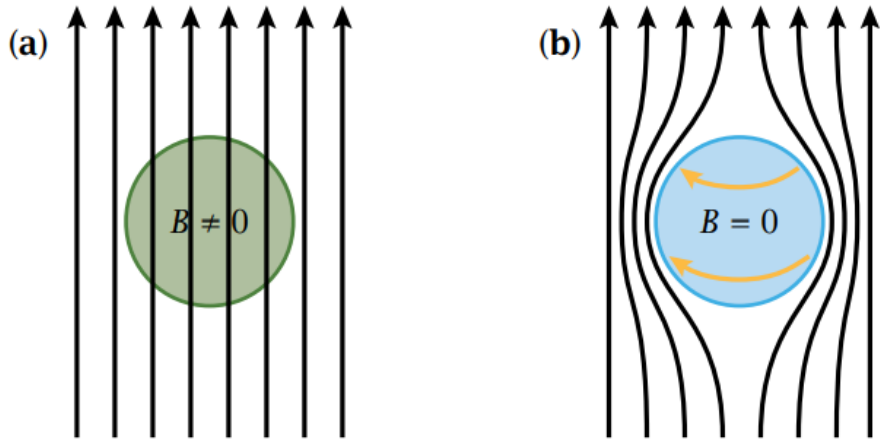


Figure 1.1: (a) Above T_c, external fields are able to penetrate the material. (b) shows the same material now in the superconducting state (below T_c). Non-examinable: Supercurrents flow horizontally around the material (shown in yellow) and cause the external fields to be repulsed from the surface.

Inside the superconductor

$$\mathbf{B} = \mu_0(1 + \chi_{SC})\mathbf{H} = 0 \tag{1.1}$$

$$\implies \chi_{SC} = -1 \tag{1.2}$$

Superconductors are perfect diamagnets.

It is important to note that the diamagnetism is due to **surface currents** and not atomic orbits.

However, the expelled flux density close to the surface of the superconductor **can depend on sample shape**. For example, take a sphere and place it in a uniform field. Then the magnetic field lines are more dense closer to the equator than they are at the poles.

This can lead to complicated intermediate states, where part of the sample remains superconducting, but other parts remain normal (i.e., regions of no superconductivity). For the sphere example, this results in planes of non-superconducting regions parallel to the field lines.

Moreover, whether you cool a conductor before or after does not matter if the material is a superconductor:

1.1.2 Critical Fields

There are two types of superconductors, each with different phase transitions.

Type I Superconductors

For a Type I, the critical magnetic field at temperature T is

Critical magnetic field

$$H_c(T) = H_c(0)\left[1 - \left(\frac{T}{T_c}\right)^2\right] \tag{1.3}$$

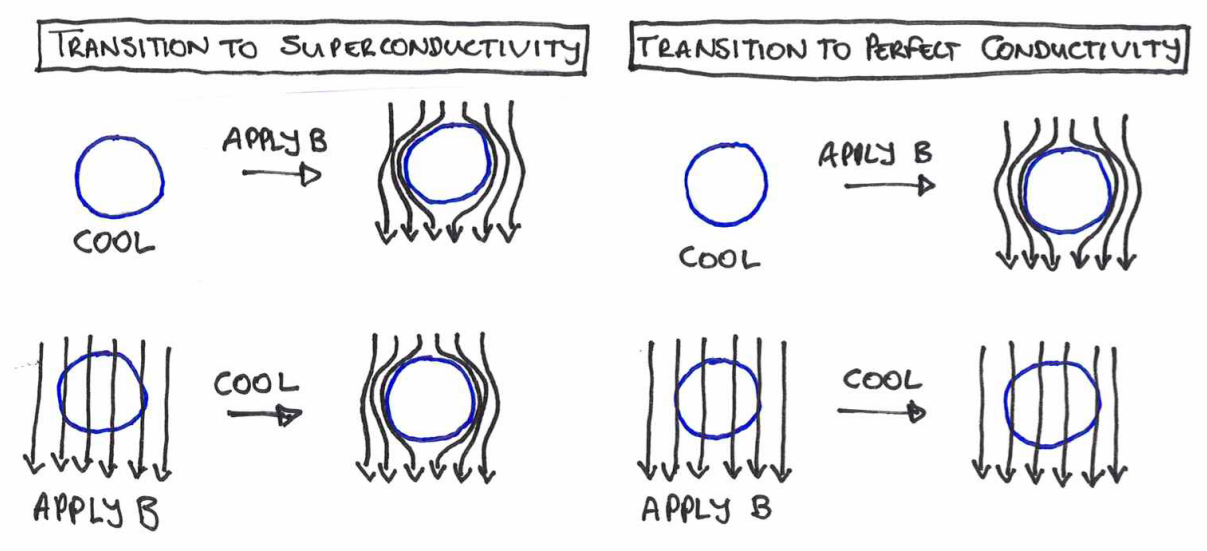


Figure 1.2: Cartoon showing the difference between a superconductor and perfect conductor.

(a) Phase diagram for Type I superconductors. The boundary is a parabola.

(b) Magnetisation of Type I superconductor against applied field. The gradient is 1.

Figure 1.3: Graphs displaying the important properties of Type 1 superconductors.

Type II Superconductors

Type II Superconductors

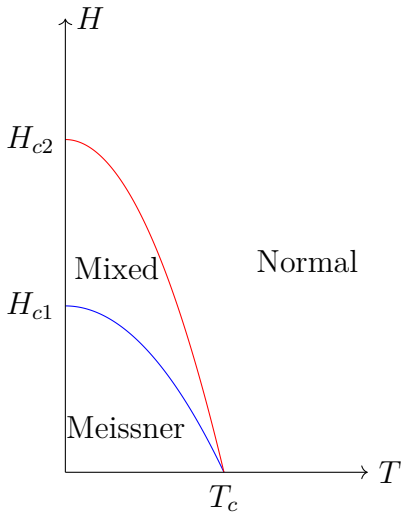
Type II Superconductors have 2 critical phases.

- At $H < H_{c1}$, we have the **Meissner state** which is the same as in Type 1: $\rho = 0, \chi = -1$.
- At $H_{c1} < H < H_{c2}$, we have the **Mixed state** where $\rho = 0, \chi \neq -1$.
- At $H > H_{c2}$, superconductivity is destroyed.

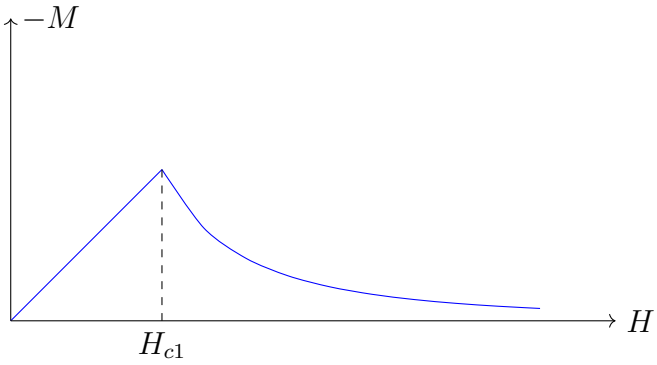
The phase diagram is shown in Fig. 1.4(a) and the magnetisation in Fig. 1.4(b).

The most important things to note are

- In a mixed state, fine filaments (regions) of the material become normal. This means that the external magnetic flux can penetrate the material.
- The magnetisation graph in Fig. 1.4(b) is linear from the origin to H_{c1} with gradient 1. After that, it tapers down towards 0 as $H \to H_{c2}$ in a hyperbolic fashion - i.e. draw a reciprocal graph like $1/x$.
- $\mu_0 H_{c1} \sim 10^{-2}$T but $\mu_0 H_{c2}$ can be large.



(a) Phase diagram for Type II superconductors. The boundaries are parabolas.



(b) Magnetisation of Type II superconductor against applied field.

Figure 1.4: Graphs displaying the important properties of Type 2 superconductors.

1.1.3 Surface Currents

In Section 1.1.1, we briefly mentioned supercurrents - but why do they arise and why only at the surface? This is explained by Ampere's Law. Recall:

$$\oint_C \mathbf{B} \cdot d\mathbf{l} = \mu_0 I. \tag{1.4}$$

where C is a closed loop and I is the current **enclosed by that loop**. We can then try evaluating along two loops: a loop which is entirely in the bulk, and another loop which is "half-in, half-out" of the superconductor.

Indeed, by the Meissner effect, there is zero $\mathbf{B}$ inside the superconductor, so Ampere's law evaluates to 0. Hence there is no current inside the bulk at all.

However with a non-zero $\mathbf{B}$ outside the material, the loop that is half-i the superconductor receives a non-zero contribution to Ampere's law, so there is indeed a surface supercurrent.

1.1.4 Critical Currents

Silsbee's Rule

The superconducting state is destroyed above when the current exceeds a critical current I_c.[1].

For a wire of radius a, using Ampere's law, the current at the surface is

$$B(a) = \frac{\mu_0 I}{2\pi a}, \tag{1.5}$$

This can be arranged to get the critical current, which occurs at the critical field

$$I_c = \frac{2\pi a B_c}{\mu_0} \tag{1.6}$$

[1]Silsbee, F. B. (1918). "Note on electrical conduction in metals at low temperatures". Bulletin of the Bureau of Standards. 14 (2): 301. doi:10.6028/bulletin.335. ISSN 0096-8579.

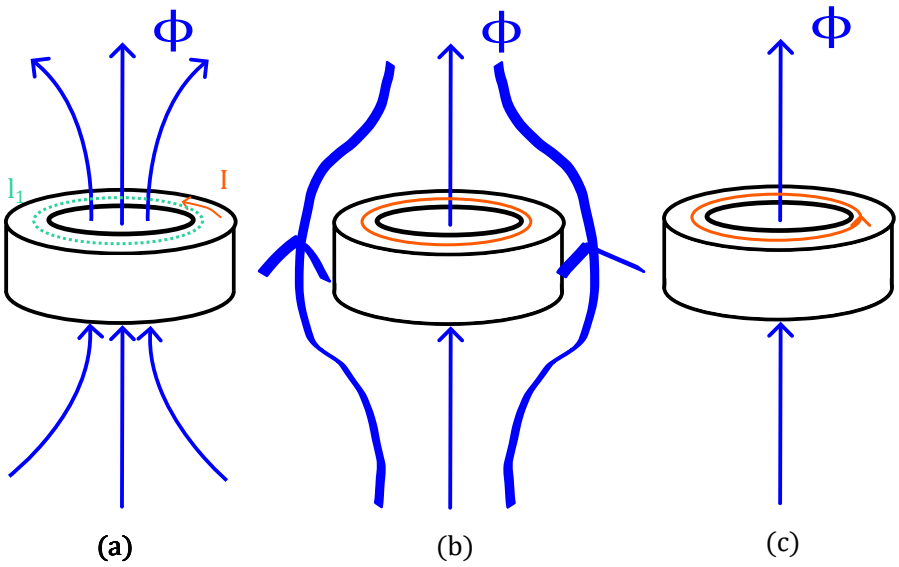


Figure 1.5: (a) Cylindrical conductor with applied magnetic field **B**, causing a current loop I. Then dotted loop l_1 is on the surface of the conductor. (b) Same system now cooled to be a superconductor. (c) Superconductor when the field **B** is removed.

1.1.5 Persistent Currents

Suppose we have the following setup in Fig. 1.5(a). We know the magnetic flux is given by

$$\Phi = \int \mathbf{B} \cdot d\mathbf{S} \tag{1.7}$$

But we also know Faraday's Law. Thus, we differentiate both sides with respect to time:

$$\frac{\partial \Phi}{\partial t} = -\int \frac{\partial \mathbf{B}}{\partial t} \cdot d\mathbf{S} = -\int (\nabla \times \mathbf{E}) \cdot d\mathbf{S} = -\oint \mathbf{E} \cdot d\mathbf{l}, \tag{1.8}$$

where in the last equality, we applied **Green's theorem**. We know that for a conductor, $\mathbf{E} = 0$ in any loop so the above equation tells us that there is **flux conservation**. Now, let's cool the conductor to $T < T_c$. Since a superconductor is still a conductor, flux conservation still applies. This causes some flux to be shifted out of the concentric hole as in Fig. 1.5(b). Keeping the temperature fixed and removing **B** gives us Fig. 1.5(c). Since flux conservation holds, the flux through the concentric hole must be fixed. The external fields disappear since they are not trapped in the confines of the superconductor.

1.2 Thermodynamics of Superconductors

At constant pressure $\mathrm{d}G = -S\ \mathrm{d}T - m\ \mathrm{d}B$, where G is the Gibbs free energy and m is the magnetic moment. Therefore, when T is constant and less than the superconducting critical temperature T_C

$$G_\mathrm{s}(B) - G_\mathrm{s}(0) = -\int_0^B m\ \mathrm{d}B$$

Here subscript s (or n) represents the superconducting (or normal) state, $m = MV$ and $M = -I = -B/\mu_0$ for a superconductor. This implies that

$$G_\mathrm{s}(B) - G_\mathrm{s}(0) = -\int_0^B m\ \mathrm{d}B = \frac{VB^2}{2\mu_0}$$

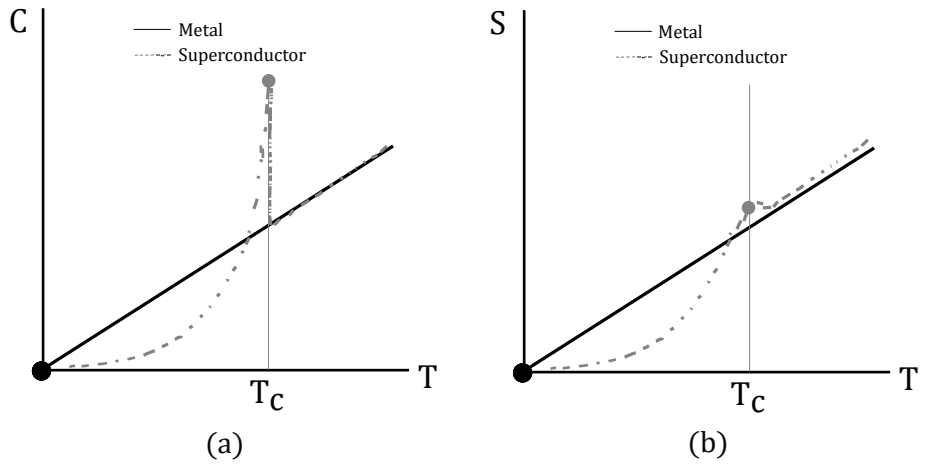


Figure 1.6: (a) Graph showing heat capacity C against temperature T (b) Graph showing entropy S against temperature T. All temperatures are in Kelvin.

At $B = B_{\mathrm{c}}$, we have that

$$G_{\mathrm{s}}\left(B_{\mathrm{c}}\right) = G_{\mathrm{n}}\left(B_{\mathrm{c}}\right) = G_{\mathrm{n}}(0)$$

because - the superconducting and normal states are in equilibrium, - we assume no field dependence in G_{n}.

Hence

$$G_{\mathrm{n}}(0) - G_{\mathrm{s}}(0) = \frac{VB^2}{2\mu_0}$$

and therefore

$$S_{\mathrm{n}} - S_{\mathrm{s}} = -\frac{V}{\mu_0} B_{\mathrm{c}} \frac{\mathrm{d}B_{\mathrm{c}}}{\mathrm{d}T} > 0$$

because

$$\frac{\mathrm{d}B_{\mathrm{c}}}{\mathrm{d}T} < 0$$

Differentiation yields

$$C_{\mathrm{n}} - C_{\mathrm{s}} = -\frac{TV}{\mu_0}\left[B_{\mathrm{c}} \frac{\mathrm{d}^2 B_{\mathrm{c}}}{\mathrm{d}T^2} + \left(\frac{\mathrm{d}B_{\mathrm{c}}}{\mathrm{d}T}\right)^2\right]$$

Remark.

$$\frac{B_{\mathrm{c}}^2}{2\mu_0}, \tag{1.9}$$

is the **condensation energy** which is the energy needed to get the superconducting state

We see that in Fig. 1.6(a), the superconducting metal has a non-linear increase in C, but at $T > T_c$, it behaves like a regular metal. The straight line contribution for the metal is the **Debeye contribution** and shows a proportional relationship to T. The entropy S can also be plotted and is shown in Fig. 1.6(b). Moreover, we se at T_c that is it **continuous**, meaning we have a **second-order phase transition**.

1.3 London Equations

The London brothers wanted to effectively have an 'Ohm's Law' for superconductors relating the vector potential and current density. They realised that there was some long-range order phenomenon regarding the momentum vector. The rigidity of the superconducting

wavefunction ψ was also responsible for perfect diamagnetism. Indeed, they made two important assumptions

Assumptions of the London Equation

1. Superconducting charge carriers are **NOT** normal state carriers and only exist below T_c.
2. The superconducting charge carriers all occupy the same energy state (bosonic). This implies 0 canonical momentum.

1.3.1 Experimental justification

- No hysteresis on the magnetisation M against H curve compared to normal state carriers. This means there is no 'lag' in the system's response.
- No dissipation of supercurrent

No hysteresis already told the London brothers that these charge carriers were not just regular electrons. Importantly, there was no continuum of states (unlike electrons) but a single $\mathbf{p} = 0$ state which the second assumption states.

The superconducting current density is given by

$$\mathbf{J} = n_s q \mathbf{v}, \tag{1.10}$$

where

- n_s is the superconducting charge density
- q is the electric charge of each charge carrier
- $\mathbf{v}$ is the carrier velocity

The canonical momentum is defined as the sum of the kinetic and electromagnetic momentum, so

$$\mathbf{p} = m\mathbf{v} + q\mathbf{A}, \tag{1.11}$$

where $\mathbf{A}$ is the magnetic vector potential satisfying $\mathbf{B} = \nabla \times \mathbf{A}$ as you know from electromagetism. Setting it to be zero, we then get $\mathbf{v} = -q\mathbf{A}$ and thus

London Equation

$$\mathbf{J} = -\frac{n_s q^2}{m}\mathbf{A} \tag{1.12}$$

1.3.2 Consequences

- The superconducting currents can be entirely caused by a magnetic field $\mathbf{B}$.
- London Penetration Depth λ:
- Macroscopic quantum state

Theorem 1.3.1. *The **London Penetration Depth** λ characterizes the distance to which a magnetic field penetrates into a superconductor and becomes equal to e^{-1} times that of the magnetic field at the surface of the superconductor. It satisfies the equation*

$$\nabla^2 \mathbf{B} = \frac{\mathbf{B}}{\lambda^2} \tag{1.13}$$

Proof. Take the curl of Eq, (1.12):

$$\nabla \times \mathbf{J} = -\frac{n_s q^2}{m}(\nabla \times \mathbf{A}) \tag{1.14}$$

By Maxwell equations,

$$\nabla \times \mathbf{B} = \mu_0 \mathbf{J} + \mu_0 \epsilon_0 \frac{\partial \mathbf{E}}{\partial t} = \mu_0 \mathbf{J} \tag{1.15}$$

since $\mathbf{E} = 0$. Then rearranging for $\mathbf{J}$ and substituting into Eq. (1.14):

$$\nabla \times (\nabla \times \mathbf{B}) = -\frac{n_s q^2 \mu_0}{m}\mathbf{B} = \nabla(\nabla \cdot \mathbf{B}) - \nabla^2 \mathbf{B}, \tag{1.16}$$

$$\implies \nabla^2 \mathbf{B} = \frac{n_s q^2 \mu_0}{m}\mathbf{B} \text{ as } \nabla \cdot \mathbf{B} = 0, \tag{1.17}$$

$$= \frac{\mathbf{B}}{\lambda^2}, \tag{1.18}$$

where

London Penetration Depth Formula

$$\lambda = \sqrt{\frac{m}{\mu_0 n_s q^2}} \tag{1.19}$$

□

Notice then we have derived a differential equation in Theorem 1.3.1. Writing it in component form:

$$\frac{\partial^2 B_i}{\partial x_i^2} = \frac{B_i}{\lambda^2}, \tag{1.20}$$

where $i = x, y, z$. In order to create a well-posed PDE, we require boundary conditions. First, we let the z coordinate run parallel to the superconductor surface and x perpendicular but into the superconductor. A general solution to the PDE is

$$B_z = B_1 e^{-x/\lambda} + B_2 e^{x/\lambda} \tag{1.21}$$

with boundary conditions $B_z(x = 0) = B_0$ and $B_z \to 0$ as $x \to \infty$.[2] The second boundary condition eliminates the $e^{x/\lambda}$ term since it diverges as x increases. Therefore the equation for the magnetic field is

$$B_z = B_0 e^{-x/\lambda} \tag{1.22}$$

Indeed we see that if we set $x = \lambda$, the magnetic field is reduced by a factor of e in strength.

Experiments in the variation of λ look like Fig. 1.8. Specifically, there is always a non-zero penetration depth at 0 K. This depends on the material, its geometry and so on.

In Section 1.2, we noted the existence of a phase transition. In PX3A7 Statistical Physics, this suggests we can describe the system by some order parameter. For superconductors, this is

[2] Obviously a real superconductor doesn't have infinite thickness, but I simply mean into the bulk.

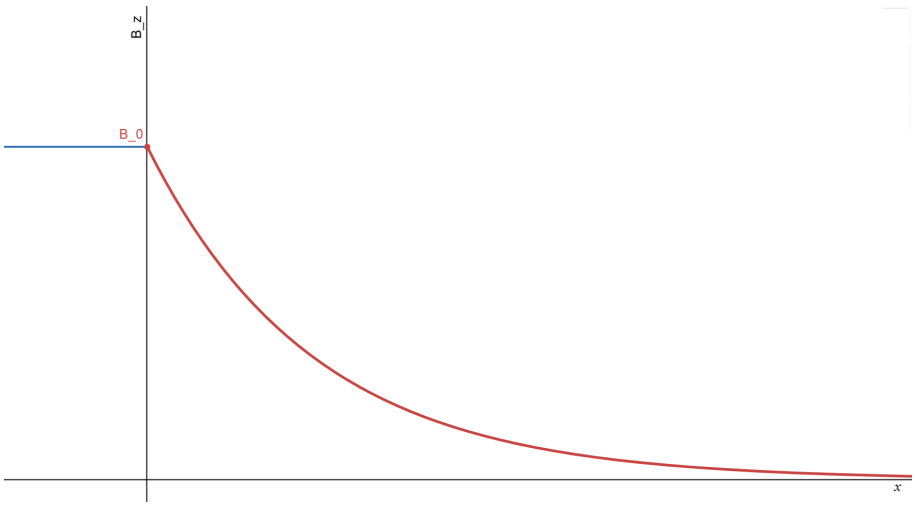


Figure 1.7: Graph of B_z against x from the surface of a superconductor.

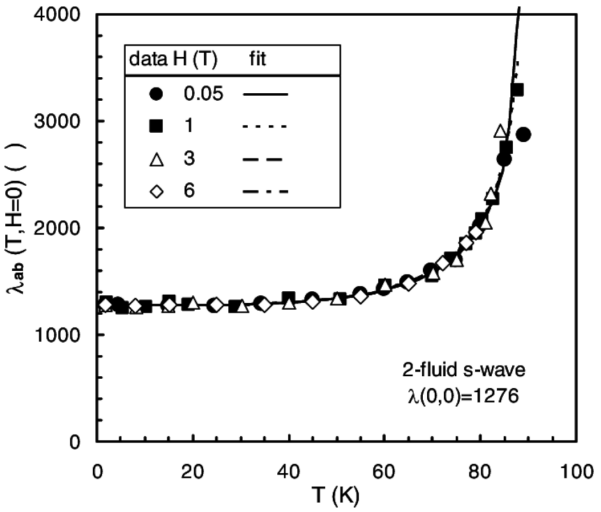


Figure 1.8: Penetration depth variation with temperature for $YBa_2Cu_3O_7$ [3].

Superconducting Wavefunction (order parameter)

$$\psi = \psi_0 e^{i\theta}, \tag{1.23}$$

where ψ is the wavefunction of the superconducting charge carriers.

1.3.3 Lack of gauge invariance of the London Equation

The London equation is unfortunately not gauge invariant (meaning we can't transform the physical system mathematically and still have it represent the same system). The London equation is valid in the **London gauge**, $\nabla \cdot \mathbf{A} = 0$.

The canonical momentum is also not gauge invariant. Simply shift it and $\mathbf{A}$ by some scalar derivative $\nabla\chi$:

$$\mathbf{A} \to \mathbf{A} + \nabla\chi \qquad\qquad \mathbf{p} \to \mathbf{p} + q\nabla\chi \tag{1.24}$$

Assuming a spatial-dependence on the superconducting wavefunction $\psi(r) = \psi_0 e^{i\theta(\mathbf{r})}$, we

can apply the momentum operator to ψ:

$$\mathbf{p}\psi = -i\hbar\nabla\psi = \hbar(\nabla\theta)\psi,$$

thus the gauge transform is $\hbar\nabla\theta \to \hbar\nabla\theta + q\nabla\chi$ which means $\theta \to \theta + q\chi/\hbar$ so the phase is also not gauge invariant (sadge).

Is this an issue though? Luckily it isn't entirely an issue. The kinetic energy operator of the Hamiltonian $\mathbf{p}\cdot\mathbf{p}/2m$ *is* gauge-invariant. Considering only the numerator, we have $m\mathbf{v} = p - q\mathbf{A} = \hbar\nabla\theta - q\mathbf{A}$ all squared. But from the gauge transforms, the $q\nabla\chi$ terms cancel and we are fine.

1.4 Flux quantisation

The existence of an order parameter leads to a remarkable observation in superconductors, that the flux has distinct levels. We assume a spatially-dependent $\psi(\mathbf{r}) = \psi_0 e^{i\theta(\mathbf{r})}$. Then we associate $|\psi|^2 = n_\mathrm{s}$, the charge carrier density. This motivation is from Ginzburg and Landau, and we will see a more complete Ginzburg-Landau theory later.

Using this wavefunction, $\mathbf{J}$ can be rewritten as

$$\mathbf{J} = \Re(\psi^* q\mathbf{v}\psi). \tag{1.25}$$

where $\Re$ denotes the real part and $*$ denotes complex conjugation. Now, unlike in the derivation of the London equation, we will not assume $\mathbf{p} = 0$ to consider excited states and so $q\mathbf{v} = (q/m)(\mathbf{p} - q\mathbf{A})$. Substituting this back into the expression for $\mathbf{J}$, we have two parts to deal with

$$\mathbf{J} \propto \psi^*\mathbf{p}\psi + \psi^* q\mathbf{A}\psi, \tag{1.26}$$

where we have temporarily dropped the q/m prefactor and will add it back on at the end.

$$\psi^*\mathbf{p}\psi = \psi e^{-i\theta}(-i\hbar)\left(i\nabla\theta\psi_0 e^{i\theta}\right) = n_s\hbar\nabla\theta \tag{1.27}$$

$$\psi^*\mathbf{A}\psi = n_\mathrm{s}\mathbf{A} \implies \mathbf{J} = \propto n_s(\hbar\nabla\theta - q\mathbf{A}) \tag{1.28}$$

Now remember we are in a superconductor and so we take an integral of a closed loop. We re-add the factor q/m in front of everything and

$$\oint \mathbf{J}\cdot d\mathbf{l} = \frac{n_s q}{M}\left(\hbar\oint\nabla\theta\cdot d\mathbf{l} - q\oint\mathbf{A}\cdot d\mathbf{l}\right) \tag{1.29}$$

But we know the definition of magnetic flux, it's just the total flux density enclosed by a surface. However we also know that $\mathbf{B} = \nabla\times\mathbf{A}$, so we can apply Stokes' theorem:

$$\Phi = \int_S \mathbf{B}\cdot d\mathbf{S} = \oint_{\partial S}\mathbf{A}\cdot d\mathbf{l} \tag{1.30}$$

We can therefore substitute for Φ directly and get

$$\oint \mathbf{J}\cdot d\mathbf{l} = \frac{n_s q}{M}\left(\hbar\oint\nabla\theta\cdot d\mathbf{l} - q\Phi\right). \tag{1.31}$$

Now take any generic superconducting material with flux going through it (e.g. a hole like in Fig. 1.5) and consider two loops l_1 and l_2 both in the bulk, but the latter loop

surrounds some Φ whereas the former doesn't. Obviously, for l_2,
$mathbf{J} = 0$ but $\Phi \neq 0$ and we can rearrange the above equation for Φ:

$$\Phi = \frac{\hbar}{q} \oint \nabla\theta \cdot d\mathbf{l} \tag{1.32}$$

For ψ to be single-valued then the integral term must be $2\pi N$ with $N \in \mathbb{N}$. Substituting this into Φ gets

$$\Phi = \frac{Nh}{q}. \tag{1.33}$$

Namely, **flux is quantised**.

> Experiments showed that $q = 2e$ and $\Phi = Nh/2e$, i.e. superconducting charge carriers are electron pairs.

1.5 Bose-Einstein Condensates

The Maxwell-Boltzmann distribution can be used to describe identical,, distinguishable particles. In contrast, the Fermi-Dirac (-) and Bose-Einstein (+) distributions can be used to describe identical, indistinguishable particles

$$f(\epsilon) = \frac{1}{e^{\frac{\epsilon-\mu}{k_B T}} \mp 1} \tag{1.34}$$

where μ is the **chemical potential**, which can be interpreted as the energy absorbed or released to change the particle number by 1, assuming entropy and volume are fixed:

$$\mu = \left.\frac{\partial U}{\partial N}\right|_{S,V} \tag{1.35}$$

where the system has internal energy U.
Definition 1.5.1. Let E, T represent energy and temperature (in Kelvin) respectively. Then E_F, T_F are the **Fermi energy** and **Fermi temperature** respectively with $E_F = k_B T_F$. As you know, the Fermi energy is the energy of the highest occupied state and is well-defined at $T = 0$.

1.5.1 Ideal Fermi gas

A graph of chemical potential against temperature (rescaled by E_F and T_F) is shown in Fig. 1.9. Now, fermions obey the Pauli Exclusion Principle (PEP). At $T = 0$, all states are filled accordingly: In particular, there are different scenarios based on the temperature

$0 < T < T_F$**:**

- Energy is now added to the system.
- Some of the initial fermions close to E_F in Fig. 1.10 are excited (they go to higher rungs of the energy ladder), leaving vacant states for particles to fill.
- Excited states mean more available ways to *arrange* the particles
- To keep entropy S fixed, internal energy U must decrease $\implies$ μ **gets smaller**.

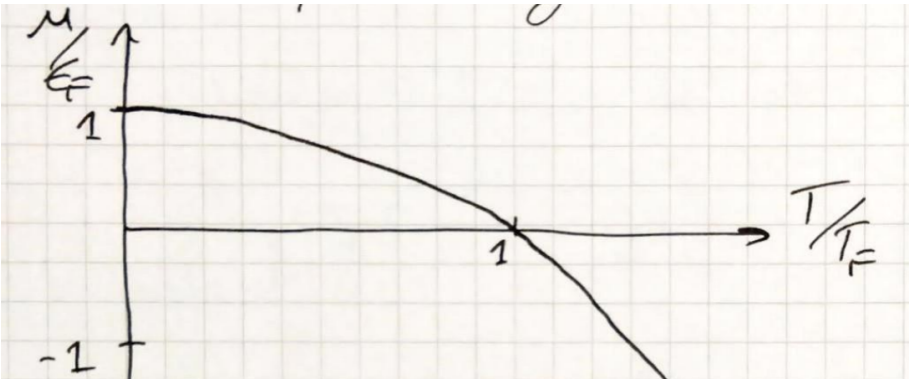


Figure 1.9: Graph of chemical potential against temperature for an ideal Fermi gas.

$T = T_F$**:** Remember the Fermi energy is the energy of the highest occupied ground state $T = 0$ which occurs at $T = T_F$. Thus, at this temperature, all initial ground states are **unoccupied**. Therefore, particles can be added to the ground state.

The number of ways to organise these particles offsets the energy, which requires $\mu < 0$! This is because, if we add a particle to the system, the available configuration space increases which increases S. But to make sense of μ, we require S to be fixed, so the additional particle can be thought of as adding a quantum of negative energy into the system, thus $\Delta U < 0$ whilst $\Delta N > 0$. This is the **classical limit** and explains why in Fig. 1.9, above T_F, it goes negative. Hence, the PEP dominates at low T in a Fermi gas.

1.5.2 Ideal Bose gas

Now, the occupancy of bosons $f_{\text{BE}} > 0$ since numerous bosons can occupy the same quantum state, and the ground state therefore will always contain at least one boson.

High T : At these temperatures, $\mu < 0$ but it is also the classical limit as for the fermionic case.

$T = 0$: Bosons do not obey the PEP. Therefore, there is only 1 ground state which contains all bosons. As a result, there is **no cost** to adding a particle, hence $\mu = 0$ as in Fig. 1.11 at $T = 0$. This is the **Bose-Einstein Condensate**

$0 < T < T_c$: Here, $T_c$is the phase transition temperature. Suppose a particle is added to the ground state. This is an excited state, and there are more possible arrangements. However, most of the bosons are in the ground state, so μ is **slightly negative**. A mathematical argument follows from requiring $f_{\text{BE}} > 0$, then $E - \mu > 0$ for every energy value. Hence, if $E = 0$, we require $\mu < 0$.

$T = T_{\mathbf{c}}$: the ground state is now unoccupied, so there is a high number of available states and configurations. By same justification for the fermionic case, $\mu < 0$.

Not all particles are i the $E = 0$ ground state when $0 < T < T_{\text{c}}$. This is captured in the particle density function:

$$\underbrace{n}_{\text{total particle density}} = \underbrace{n_0}_{\text{ground state density}} \underbrace{+ 2.61 \left(\frac{mk_BT}{2\pi\hbar^2}\right)^{3/2}}_{\text{normal state density}} \tag{1.36}$$

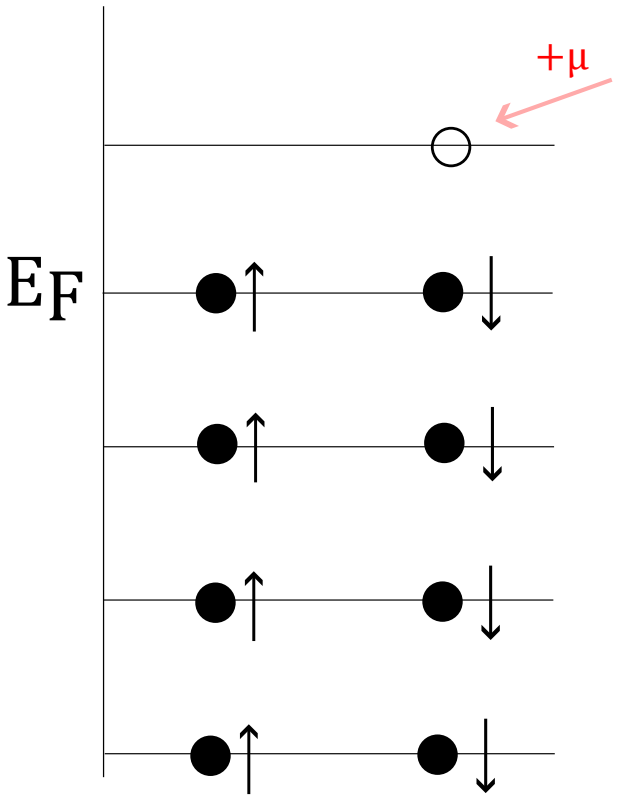


Figure 1.10: Schematic showing fermion arrangement (black-filled circles) in the ground state at $T = 0$. Potential space for new fermion (white circle) requires added $\mu = E_F$ into the system.

1.6 Ginzburg-Landau Theory

Ginzburg-Landau theories (GLTs) attempt to describe phase transitions in terms of an order parameter and phenonomelogical quantities/functions, as well as temperature and pressure etc. This module (PX446) is not a course in that - we will take GLTs as valid to study superconductivity (in fact, superconductivity is a special case in GLTs). I recommend looking at PX3A7 Statistical Physics and this pdf to refresh your memory.

Anyways, we start by defining a complex order parameter similar to before, that is spatially-dependent

$$\Psi = \begin{cases} 0 & T > T_c \\ \psi_0(\mathbf{r})e^{i\theta(\mathbf{r})} & T < T_c \end{cases}, \tag{1.37}$$

where $\psi_0(\mathbf{r}) : \mathbb{R}^3 \to \mathbb{R}$ is the amplitude such that $|\psi_0|^2 = n_s$. We will consider the free energy density near T_c and then minimise the energy w.r.t Ψ.

1.6.1 Bulk phase transition

Assume no defects and away from the surface. Then $\Psi \neq \Psi(\mathbf{r})$, i.e. it is not spatially dependent as the material looks identical everywhere in the bulk. Now, from GLTs, we need to construct the free energy density by writing generic terms for all possible scalar invariants. For example, for a magnetisation vector $\mathbf{m}$, the first order scalar invariant is $\mathbf{m} \cdot \mathbf{m}$, i.e. the inner product. We take the same idea ("guess") and write the free energy

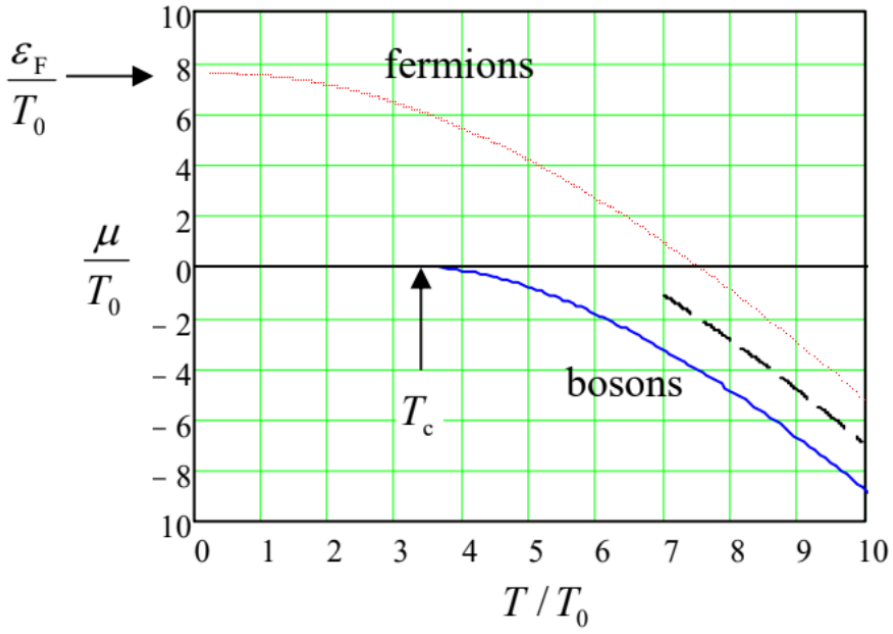


Figure 1.11: Chemical potential of fermions and bosons. The dashed line is the chemical potential computed from the Maxwell Boltzmann distribution for classical particles. Reproduced from Ref. [4].

density for the superconducting bulk to be

$$f_s(T) = \underbrace{f_n(T)}_{\text{free energy density in normal stae}} + a(T)|\Psi|^2 + \frac{1}{2}b(T)|\Psi|^4 \tag{1.38}$$

where we have cut off the series after the fourth order (higher order terms can be included but don't make a consequential difference) and the functions $a(T) : \mathbb{R}^{\geq 0} \to \mathbb{R}$ and $b(T) : \mathbb{R}^{\geq 0} \to \mathbb{R}^{>0}$.

Remark. There is no $|\Psi|^3$ term. This is because the inner product of a complex function space is defined as $\langle\Psi|\Psi\rangle = \Psi^*\Psi = |\Psi|^2$. We cannot produce $|\Psi|^3$ from a linear combination or product of scalar invariants (inner product). This would be different if the physics could be represented in a real scalar order parameter (such as the isotropic-nematic phase transition).

Plotting $f_s - f_n$ we have in Fig. 1.12. In particular, $a(T) = a_0(T - T_c)$ where $a_0 > 0$. Remember we are expanding around T_c. We also have $b(T) = b > 0$. We can now differentiate f w.r.t $|\Psi|$ to find the stationary points:

$$\frac{d(f_s - f_n)}{d|\Psi|} = 2a_0(T - T_c)|\Psi| + 2b|\Psi|^3 \tag{1.39}$$

$$|\Psi| = \begin{cases} 0 & T > T_c \\ \left(\frac{a_0}{b}\right)^{1/2}(T_c - T)^{1/2} & T < T_c \end{cases} \tag{1.40}$$

So it is a decreasing square root, and as a function of T, it is zero at $T = T_c$ and so looks like an order parameter. At the beginning of this section, we had $|\Psi| = \sqrt{n_s}$ and $\lambda = 1/\sqrt{n_s}$ so $\lambda \propto (T_c - T)^{-1/2}$. Substituting our $|\Psi|$ back in gives

$$f_s - f_n = -\frac{a_0^2}{2b}(T - T_c)^2 < 0 \tag{1.41}$$

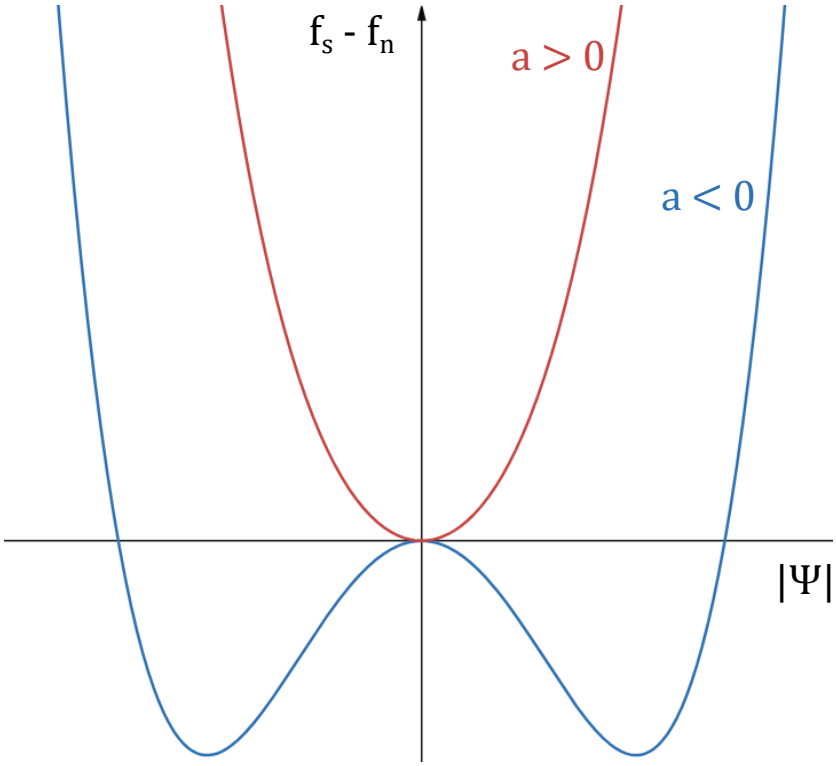


Figure 1.12: Plot of $f_s - f_n$ for different values of $|\Psi|$.

This is the **condensation energy**, the energy saved upon becoming a superconductor.

Now, as you may remember from statistical mechanics, once we have the free energy, we can find other thermodynamical quantities. Recall:

$$S = -\left(\frac{\partial f}{\partial T}\right)\Big|_{V,N} \qquad C = T\left(\frac{\partial S}{\partial T}\right)\Big|_{V,N} \tag{1.42}$$

which are the entropy and specific heat capacity respectively. Since derivatives are linear, we can also just directly differentiate the free energy difference $f_\mathrm{s} - f_\mathrm{n}$ directly to get the entropy difference and heat capacity difference respectively. This is a simple differentiation exercise and we get

$$S_\mathrm{s} - S_\mathrm{n} = \begin{cases} 0 & T > T_\mathrm{c} \\ -\frac{a_0^2}{b}(T_\mathrm{c} - T) & T < T_\mathrm{c} \end{cases} \tag{1.43}$$

$$C_\mathrm{s} - C_\mathrm{n} = \begin{cases} 0 & T > T_\mathrm{c} \\ \frac{a_0 T}{b} & T < T_\mathrm{c} \end{cases} \tag{1.44}$$

- Entropy **decreases** when entering the superconducting phase, but there is *no discontinuous jump.*
- C discontinuously jumps from 0 to a non-zero value at T_c. This matches with experiments!

1.6.2 Inhomogeneous + Charged case

We now consider doing the exact same thing with the complete Ginzburg-Landau expression for superconductivity, where now $\Psi = \Psi(\mathbf{r})$, there is the inclusion of a charged term and a term for the energy stored in an applied field B_0:

$$F_\mathrm{s} = F_\mathrm{n} + \int \mathrm{d}^3 r \left[a(T)|\Psi(\mathbf{r})|^2 + \frac{b}{2}|\Psi(\mathbf{r})|^4 + \frac{1}{2m}| - i\hbar\nabla\Psi(\mathbf{r}) + 2e\mathbf{A}\Psi(\mathbf{r})|^2 + \frac{(B(\mathbf{r}) - B_0)^2}{2\mu_0} \right] \tag{1.45}$$

(Don't memorise this). The cases where $\mathbf{A}$ are zero and non zero are both examinable.

$\underline{\mathbf{A} = 0}$

Setting $\mathbf{A} = 0$

$$F_\mathrm{s} = F_\mathrm{n} + \int f \mathrm{d}^3 r$$

where

$$f = a(T)|\Psi|^2 + \frac{b}{2}|\Psi|^4 + \frac{\hbar^2}{2m}|\nabla\Psi|^2$$

This equation is entirely real and so here we only consider variations in the amplitude of the order parameter. If Ψ is allowed to vary $\Psi \to \Psi + \mathrm{d}\Psi$ and then

$$\mathrm{d}f = 2a\Psi\ \mathrm{d}\Psi + 2b\Psi^3\ \mathrm{d}\Psi + \frac{\hbar^2}{2m}\ \mathrm{d}|\nabla\Psi|^2$$

Now, to first order

$$\mathrm{d}|\nabla\Psi|^2 = |\nabla(\Psi + \mathrm{d}\Psi)|^2 - |\nabla\Psi|^2 = 2\nabla\Psi \cdot \nabla(\ \mathrm{d}\Psi)$$

and

$$\nabla\Psi \cdot \mathrm{d}\nabla\Psi = \nabla \cdot [\nabla\Psi \mathrm{d}\Psi] - \left(\nabla^2\Psi\right) \mathrm{d}\Psi$$

The first term on the right-hand side of this last equation would represent a surface contribution in our superconductor that we can take to be zero. Substituting this back into the equation for $\mathrm{d}f$ above we have,

$$\mathrm{d}f = 2\ \mathrm{d}\Psi \left[\left(a + b\Psi^2\right)\Psi - \frac{\hbar^2}{2m}\nabla^2\Psi\right] = 0$$

for any Ψ. Minimising this implies $\mathrm{d}f = 0$ for any $\mathrm{d}\Psi$, and leaves us with,

$$-\frac{\hbar^2}{2m}\nabla^2\Psi + \left(a + b\Psi^2\right)\Psi = 0$$

which looks like a non-linear Schrödinger equation. Near T_c we can neglect the $b\Psi^2$ term because $\Psi \to 0$ and then the equation takes the form

$$\nabla^2\psi = \frac{\psi}{\xi^2}$$

where $\xi = \sqrt{\frac{\hbar^2}{2m|a(T)|}}$.

$\underline{\mathbf{A} \neq 0}$

Now we turn on the magnetic field. Just like the lecturer, I cannot be asked to type out about 2 pages of maths even though I *could*. You must minimise w.r.t. variations in both Ψ and $\mathbf{A}$. This (from what I can tell) has not appeared in exams. We can therefore skip to the main result

Ginzburg-Landau Equations

$$\begin{aligned} \frac{\hbar^2}{2m}\left(-i\nabla + \frac{2e}{\hbar}\mathbf{A}\right)^2\Psi + a(T)\Psi + b\Psi^3 = 0, \\ \frac{\nabla \times \mathbf{B}}{\mu_0} = \mathbf{J} = -i\frac{e\hbar}{m}\left[\Psi^*\nabla\Psi - \Psi\nabla\Psi^*\right] - \frac{4e^2}{m}|\Psi|^2\mathbf{A} \end{aligned} \tag{1.46}$$

Consider $\nabla\psi = \nabla\psi^* = 0$, then the second GL equation $\implies$ $\mathbf{J} = \frac{-(2e)^2}{m}|\psi|^2\mathbf{A}$ (when assuming $\mathbf{p} = 0$) This is the London equation with $q = 2e$ and $|\Psi|^2 = n_\mathrm{s}$.

1.6.3 Consequences

The GL-equations predict Type-I and Type-II superconductors.
Definition 1.6.1. The **Ginzburg-Landau parameter** is defined as

$$\kappa = \frac{\lambda(T)}{\xi(T)} \tag{1.47}$$

where $\lambda(T)$ is the penetration depth of a magnetic field into a superconductor, and $\xi(T)$ is the recovery distance of the order parameter Ψ from the superconducting edge to the bulk ψ_0 (GL coherence length).

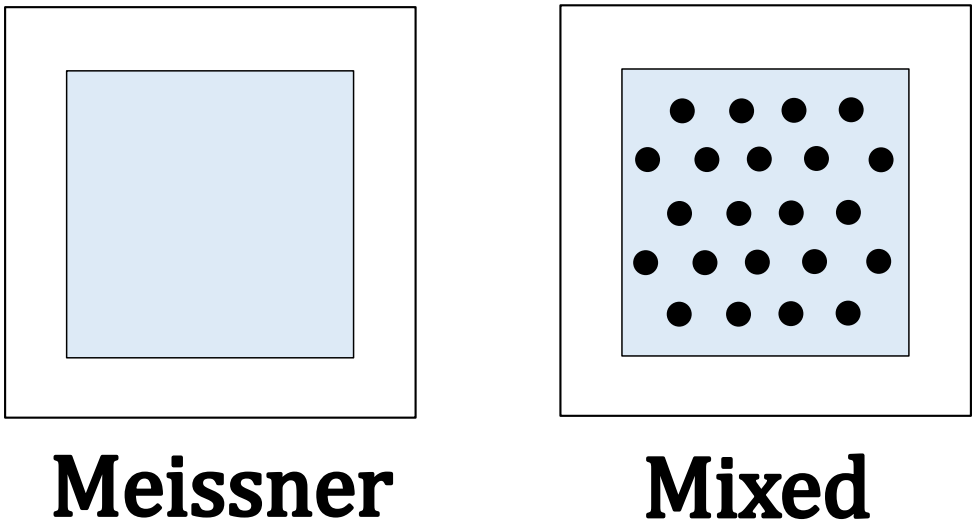


Figure 1.13: Diagram showing a Meissner state (left) and a mixed superconducting state (right). The black holes represent **normal cores/vortices**.

The expressions for these are predicted by the GL equations:

$$\lambda = \sqrt{\frac{mb}{4\mu_0 e^2 |a(T)|}} \qquad \xi = \sqrt{\frac{\hbar^2}{2m|a(T)|}}. \tag{1.48}$$

The GL equations also predict the Meissner effect, flux trapping and flux quantisation as we saw earlier.
Remark. The Pippard coherence length is the maximal length from a magnetic perturbation which is experienced by charge carriers and is denoted ξ_0. At very low $T, \xi \to \xi_0$ and ξ_0 was predicted in BCS theory.

In κ, there are competing energy scales:

- The condensation energy saved on becoming a superconductor: $|a^2/2b|$ which is related to ξ
- The diamagnetic energy cost to expel the magnetic field: $B^2/2\mu_0$ related to λ

- For a type-I superconductor, $\kappa < 1/\sqrt{2}$
- For a type-II superconductor, $\kappa > 1/\sqrt{2}$

Experiments can visualise what a type I and II look like.

Fig. 1.13 shows that the mixed superconducting state is a mix of a pure Meissner state (shaded) and vortices. Each vortex contains **one flux quantum** and each has a radius $\sim \xi$.

Fig. 1.14(left) above shows the behaviour of magnetic field (B) and order parameter Ψ close to the surface of a type I superconductor that exists for $x > 0$. Here the surface costs energy because there is an extended region for which field is being expelled, but where the order parameter has not reached its bulk value. The middle figure is the same plot for a type II superconductor showing that the surface saves energy because there are regions that have the full bulk value of the order parameter, but are not completely expelling the magnetic field. This qualitatively accounts for the formation of the mixed or vortex phase.

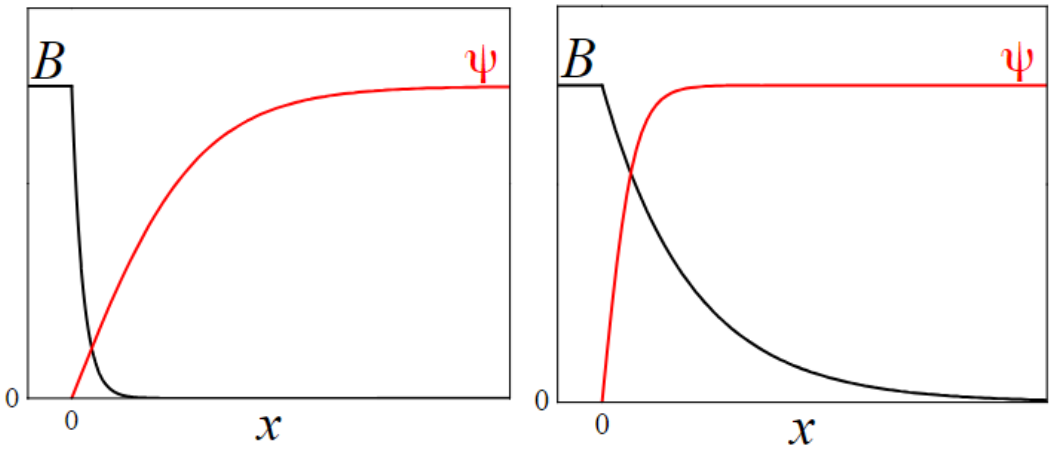


Figure 1.14: .

All the way back in Fig. 1.3a (for Type 1) and Fig. 1.4a, the quantities H_c, H_{c1} and H_{c2} are mentioned but never defined. Indeed, the GL equations predict them, though we will use B instead of H:

- B_c is the field associated with the condensation energy.
- $B_{c1} = \mu_0 H_{c1} \approx \frac{\Phi_0}{4\pi\lambda^2} \ln(\kappa)$
- $B_{c2} = \mu_0 H_{c2} \approx \frac{\Phi_0}{2\pi\xi^2}$

For a square vortex of spacing d, $\Phi_0 = Bd^2 \implies d = (\Phi_0/B)^{1/2}$. For the triangular vortex, $d = \sqrt{2\Phi_0/\sqrt{3}B}$

1.7 Phase Coherence

The GL order parameter is identical to the macroscopic superconducting wavefunction. In the bulk, we had $\psi_0 = \sqrt{n_s}$, a constant. From the GL free energy Eq. (1.45),

$$F_s = F_s^0 + \frac{\hbar^2 n_s}{2m} \int \left(\nabla\theta + \frac{2e}{\hbar}\mathbf{A} \right) d^3r, \tag{1.49}$$

where F_s^0 is the ground state free energy and $\nabla\theta$ represents an **energy cost for variations in phase**. In the ground state all superconducting charge carriers have the same phase. Hence we can do a bit of spontaneous symmetry breaking

$$\text{Broken Symmetry in superconducting state} \rightarrow \begin{matrix} \text{Macroscopic phase} \\ \text{coherence} \end{matrix}$$
$$\text{Rigidity} \rightarrow \text{energy cost of } \nabla\theta$$

Using Eq. 1.46, in particular the second one, we can use $\Psi = \sqrt{n_s(\mathbf{r})} e^{i\theta(\mathbf{r})}$ and evaluate the terms

$$\begin{aligned} \Psi^* \nabla\Psi &= (n_s^{1/2} e^{-i\theta}) \left(e^{i\theta} \nabla (n_s)^{1/2} + n_s^{1/2} \nabla (e^{i\theta}) \right) \\ &= \frac{1}{2} \nabla n_s + n_s \nabla\theta \\ \Psi \nabla\Psi^* &= \frac{1}{2} \nabla n_s - n_s \nabla\theta \end{aligned}$$

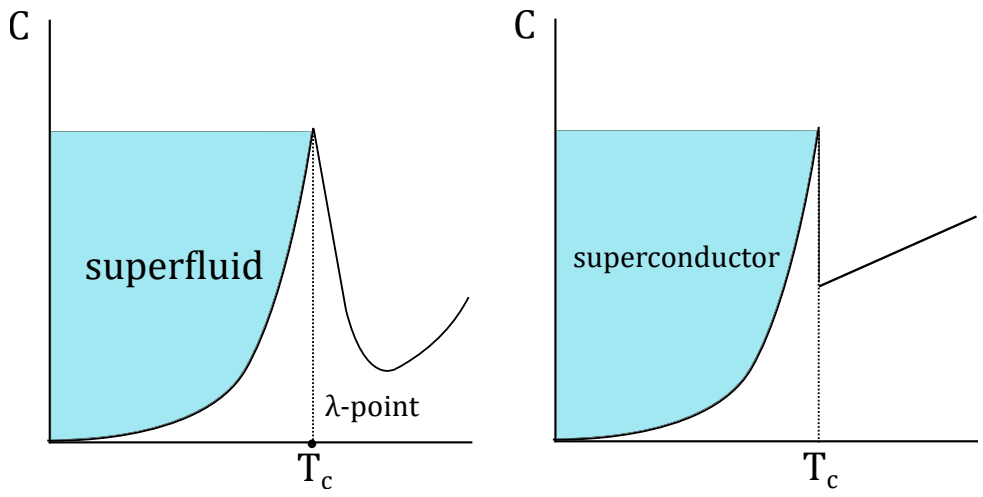


Figure 1.15: Plots of heat capacity against temperature.

Thus, the current density is

$$\mathbf{J} = \frac{2e\hbar}{m} n_s \nabla\theta - \frac{(2e)^2}{m} n_s \mathbf{A}$$

You always expect a magnetic field to induce a current, so $\mathbf{A}$ appearing is fine, but now what we learn is **phase gradients can cause currents**. To calculate the flux, we just integrate both sides around a closed loop

$$\oint \mathbf{A} \cdot d\mathbf{l} = \Phi = \frac{\hbar}{2e} \oint \nabla\theta \cdot d\mathbf{l} \tag{1.50}$$

$$\implies \boxed{\Phi = \frac{Nh}{2e}} \tag{1.51}$$

We get **flux quantisation**, and $N \in \mathbb{Z}^+$ is called the **winding number** and θ changes by 2π around a flux quantum. Note in this formula, it is the regular Planck's constant h NOT the reduced constant $\hbar$.

1.8 Superfluidity

Superfluids are cool. We focus on Helium-4 (^{4}He). This enters the superfluid state at 2.17 K. Superfluid flow, *flow* quantisation and a macroscopic quantum state are all observed. The specific heat capacity of a superconductor and superfluid are similar and are shown in Fig. 1.15. This phase transition is evidence for a macroscopic wave function and order parameter as before

$$\psi_0(\mathbf{r}) = \sqrt{n_0} e^{i\theta(\mathbf{r})}, \tag{1.52}$$

where $n_0 = |\psi_0(\mathbf{r})|^2$ is the density of particles in the condensed superfluid state. This is similar to the superconducting version. This order parameter undergoes **macroscopic phase coherence** as it becomes a superfluid. Just like with superconductors, we go through their properties.

Properties of superfluids

Click any links in this box to get taken to the relevant subsection.

- Superflow: flow of particles with no dissipation $\implies$ no viscosity.
- Two-fluid model: viscous drag found in ^{4}He, decreasing with temperature, suggesting a normal and superfluid coexistence.
- Leads to Two-fluid hydrodynamics to analyse this.
- Flow quantisation
- Has a critical velocity.

1.8.1 Superflow

This is zero viscosity flow. If the phase is constant everywhere, then only the density of superflow condensate particles is spatially-dependent, which isn't particularly interesting because that is expected to happen. So what if θ is spatially-dependent? We can perform similar analysis to superconductors, but instead of the electric current density, we must use the **quantum-mechanical probability current** for free spin-0, neutral particles

$$\mathbf{J}_0 = \frac{\hbar}{2mi}\left(\psi_0^*\nabla\psi_0 - \psi_0\nabla\psi_0^*\right) \tag{1.53}$$

Remember that $n_0 = n_0(\mathbf{r})$. This is a standard differentiation exercise, and we get

$$\mathbf{J}_0 = n_0 \underbrace{\frac{\hbar}{m}\nabla\theta}_{\text{velocity}} \tag{1.54}$$

Since $\mathbf{J} = q\mathbf{v}$ (particle density times velocity), we get the **superfluid velocity $\mathbf{v_s}$** on the RHS. Now, the consequences of zero viscosity are cool, but can it be observed? If you had access to your own supercooled ^{4}He, you too can

1. Submerge a stack of rotating disks in superfluid
2. Measure the drag force
3. Using standard fluid dynamics, can get viscosity
4. For ^{4}He, they found non-zero viscosity, meaning not all the helium-4 was a superfluid. This suggests there are 2 phases coexisting

Link back to properties of superfluids.

1.8.2 Two-fluid model

Viscosity measurements *decreased* with temperature in helium-4. Since it was found some parts of it were superfluid and others normal, we can consider to model the system with both phases together:

$$n = n_\mathrm{s} + n_\mathrm{n}, \tag{1.55}$$

where

- n is the total particle density,
- n_s is the superfluid particle density, and
- n_n is the normal particle density

The variation of these densities is shown in Fig. 1.16 and the original paper by C. J. Gorter is here. Link back to properties of superfluids.

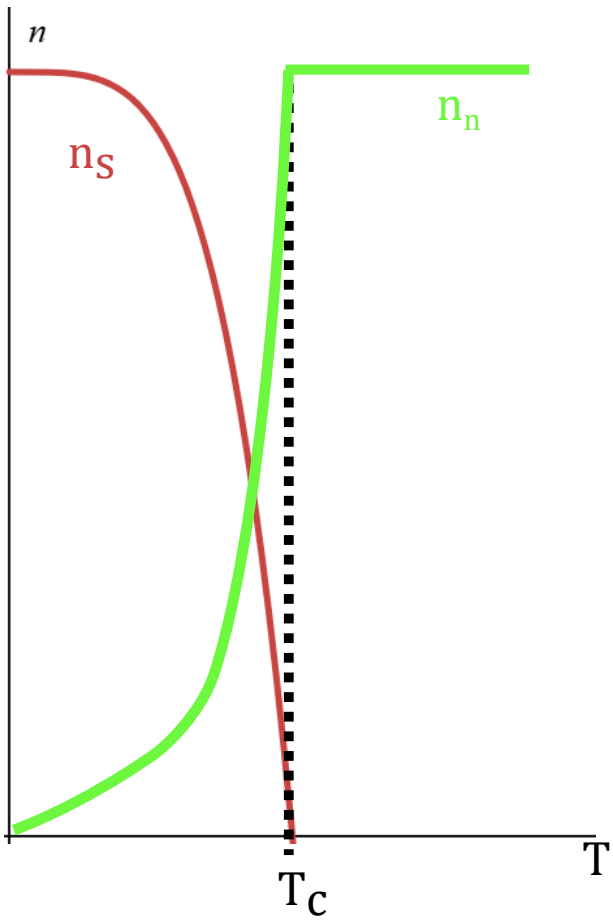


Figure 1.16: Variation of the normal (n_{n}) and superfluid (n_{s}) densities as a function of temperature.

1.8.3 Two-fluid hydrodynamics

Now we want to study how these 2 fluids interact. Suppose we have a capillary (thin tube). The superfluid will flow through it without friction, whilst the normal part experiences friction. In the stacked disks experiment, the normal phase experiences drag whilst the superfluid stays at rest, so we have **two types of current flow**.

$$\mathbf{J} = \mathbf{J}_\mathrm{s} + \mathbf{J}_\mathrm{n} \tag{1.56}$$

The superfluid carries **no entropy** (it is a condensate, same thing for superconducting charge carriers), so the flow of heat is entirely due to the normal (non-superfluid) particles. Hence we can get a **fountain effect** by manipulating the flow:

1. Insert a capillary (dense material) into liquid ^{4}He below T_c with the top half of the capillary exposed outside the fluid.
2. Heat up the capillary tube $\implies$ temperature difference ΔT
3. ^{4}He a condensate, must have same chemical potential throughout it.
4. Temperature difference leads to pressure difference, since change in chemical potential must be 0, so $\Delta\mu = 0 = \Delta P - (S/V)\Delta T$.
5. Superfluid can flow, but cannot equalise the temperature difference because it cannot carry heat.

Link back to properties of superfluids.

1.8.4 Flow quantisation

We saw earlier that $\mathbf{v}_\mathrm{s} = \frac{\hbar}{m}\nabla\theta$. In particular,

Proposition 1.8.1. *The* ***flow circulation*** $\kappa = \frac{\hbar}{m}2\pi N$ ***where*** $N \in \mathbb{Z}$ ***around a closed path***

Proof. The flow circulation is defined as

$$\kappa = \oint \mathbf{v}_s \cdot d\mathbf{s} = \frac{\hbar}{m}\oint \nabla\theta \cdot d\mathbf{r} = \frac{\hbar}{m}2\pi N \tag{1.57}$$

□

Link back to properties of superfluids.

1.8.5 Critical velocity

Definition 1.8.1. The **critical velocity** is the velocity at which superfluidity is **lost**.

Consider large mass M of fluid flowing down a pipe with velocity $\mathbf{v} = \mathbf{p}/m$ where $\mathbf{p}$ is the fluid momentum. Suppose there is some roughness on the pipe. This causes particles to scatter, forming a quasi-particle, and the fluid *recoils* by momentum $\mathbf{q}$.

By energy conservation, the energy of the fluid after cannot exceed energy before (before quasiparticle formation) so

$$\frac{p^2}{2M} \geq \frac{(p-q)^2}{2M} + Aq$$

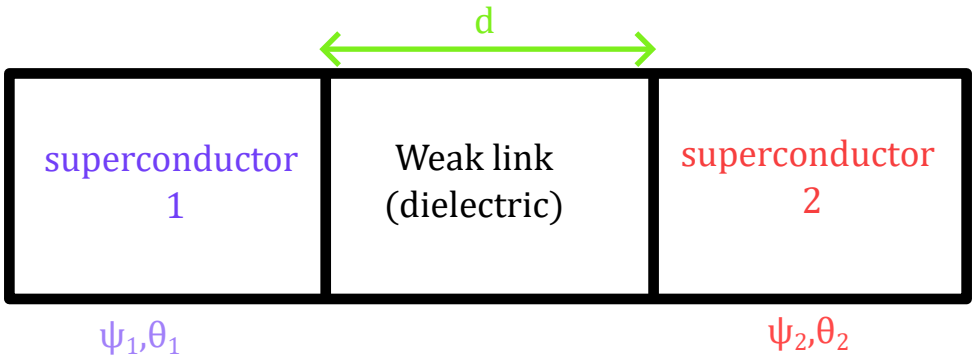


Figure 1.17: Schematic of Josephson junction.

where $A \in \mathbb{R}$ is a factor from a dispersion relationship and Aq is the energy of the quasiparticle. Rearranging the inequality we get

$$0 \geq \frac{q^2}{2m} - \frac{pq}{M} + Aq \tag{1.58}$$

$$\implies Aq \leq vq - \frac{q^2}{2M} \tag{1.59}$$

$$\overset{M \text{ large}}{\implies} \boxed{v \geq A} \text{ to form quasiparticle} \tag{1.60}$$

We set $A = v_{\text{crit}}$, the **critical velocity**. If the velocity is less than this, the quasiparticle cannot form, hence there is no friction. This quasiparticle is called a **roton** and can be uniquely identified from its dispersion relationship, where it initially displays a linear dispersion - these are phonons, before attaining a maximum, then a minimum (the roton). Figures 8 and 9 of Yarnell *et al.* show experimental validation of rotons (same paper displayed in lectures). Link back to properties of superfluids.

1.9 Josephson Effect

A typical Josephson junction is shown in Fig. 1.17, which is a dielectric surrounded by 2 superconductors, 1 and 2. Each has their own macroscopic wave function and phase.
Definition 1.9.1. There are two types of Josephson junctions depending on the weak link.

- If the weak link is a **dielectric**, it is called SIS junction.
- If the weak link is a **normal metal**, it is called SNS junction. The weak link is usually very narrow

Remark. **Proximity effect**. If you don't have a weak link, and instead place two superconductors nearby, the wavefunction is smeared and weird things happen.

So, superconducting carriers will **tunnel** through the weak link because they are sufficiently narrow , and because there is a phase difference, needed by GL equations $\phi \sim \theta_1 - \theta_2$.

Josephson Equations

Theorem 1.9.1.

$$I = I_J \sin(\phi) \qquad V = \frac{\hbar}{2e}\frac{\partial \phi}{\partial t}, \tag{1.61}$$

where I, V are the total current and voltage respectively flowing across the junction. The equations are called the first and second Josephson equations respectively, or you may sometimes see them referred to as the ***Josephson current-phase*** *and* ***Josephson voltage-phase*** *relationships.*

The module presents the handwavy argument, and so will be presented here since it is examinable. The derivation (from quantum mechanics) is shown in Appendix TODO.

Proof. The current flowing left to right $I_{\mathrm{L}\to\mathrm{R}} \sim e^{i\phi}$. Similarly, right to left is $I_{\mathrm{R}\to\mathrm{L}} \sim e^{-i\phi}$. Thus the total current is the sum, and using complex trig identities we get

$$I = I_{\mathrm{J}} \sin\phi. \tag{1.62}$$

I_{J} is an empirically-determined prefactor. It depends on the junction, with factors such as: materials (of both superconductors and the weak link), the temperature, pressure, defects etc.

Now, let's attach a voltmeter in parallel, connected to the 2 superconductors. We want to find the voltage. The wavefunction we propose to be that of electrons in a stationary state:

$$\Psi(\mathbf{r}, t) = \psi(\mathbf{r}) e^{-i\epsilon t/\hbar} \tag{1.63}$$

where ϵ is the energy associated with an electron pair. Then if we say $\psi_1 = |\psi_1| e^{i\theta(t)}$ then

$$-\hbar\partial_t\theta_1 = \epsilon_1 \qquad -\hbar\partial_t\theta_2 = \epsilon_2 \tag{1.64}$$

Subtracting the right equation from the left:

$$\Delta\epsilon = \epsilon_1 - \epsilon_2 = \hbar\partial_t\phi = qV \tag{1.65}$$

with $q = 2e$ and V the applied voltage. Rearranging gets us the second Josephson equation. □

We can now look at different scenarios of the Josephson junction

1.9.1 DC Josephson effect

We suppose there is **no applied voltage**.

- Spontaneous supercurrent flows
- Decreasing of I_{J} as field strength increases, see Fig. 1.18.

1.9.2 AC Josephson effect

Apply a DC voltage across the junction. We can directly integrate the Josephson voltage relationship:

$$\phi = \frac{2eV}{\hbar} t + \phi_0. \tag{1.66}$$

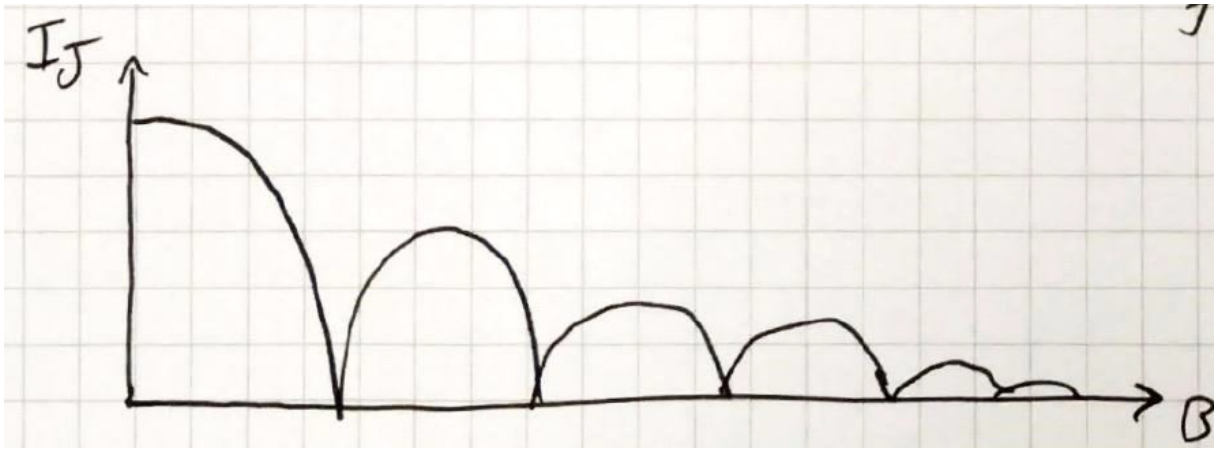


Figure 1.18: Field dependence of I_J. Reproduced from Prof. Alex Robertson's lectures.

Substituting this into the Josephson current equation, we get

$$I = I_J \sin(\omega_J t + \phi_0) \qquad \omega_J = \frac{2eV}{\hbar} \tag{1.67}$$

We get an **AC supercurrent from DC voltage**! ω_J the **Josephson frequency**. For a rough scale, if $V \sim 100\mu V$ then $\omega_J \sim 50$ GHz.

1.9.3 Inverse AC Josephson effect

We now apply a **DC** voltage V_0 AND a **radio-frequency AC voltage** $V_{rf}\cos(\omega t)$, so $V = V_0 + V_{rf}\cos(\omega t)$. Doing the same steps as for the AC effect:

$$\phi = \left(\frac{2eV_0}{\hbar}\right) t + \frac{2eV_{rf}}{\hbar\omega}\sin(\omega t) + \phi_0 \tag{1.68}$$

$$I = I_J \sin\left[\omega_J t + \frac{2eV_{rf}}{\hbar\omega}\sin(\omega t) + \phi_0\right] \tag{1.69}$$

This is the **inverse AC effect**, where the phase of the AC supercurrent **oscillates**. If you are bothered, you can use Fourier series and Bessel functions and rewrite the expression to remove the nested sin functions, allowing an easier fit in experiments:

$$I = I_J \sum_{n=-\infty}^{\infty} (-1)^n \left[J_n\left(\frac{2eV_{rf}}{\hbar\omega}\right)\right] \sin\left[(\omega_J - n\omega)\, t + \phi(0)\right], \tag{1.70}$$

where J_n is a Bessel function of the first kind. Whenever $\omega_J = n\omega$ or $2eV_{rf} = \hbar\omega$, we get DC voltage again.

1.9.4 DC vs Inverse AC

We can compare the 2 effects using an IV characteristic. It is hard to measure the AC supercurrents for both the AC and Inverse AC effect, they are often in the high GHz range. It is possible, but measurement of the AC current is no longer 'loss-less'. Despite this, we can still measure the DC current I_{DC} for both the DC and inverse AC effects, which are shown in Fig. 1.19. The amplitude of the Shapiro spikes in Fig. 1.19 is given by the Bessel functions, and we see that the amplitude decreases generally, but still oscillates.

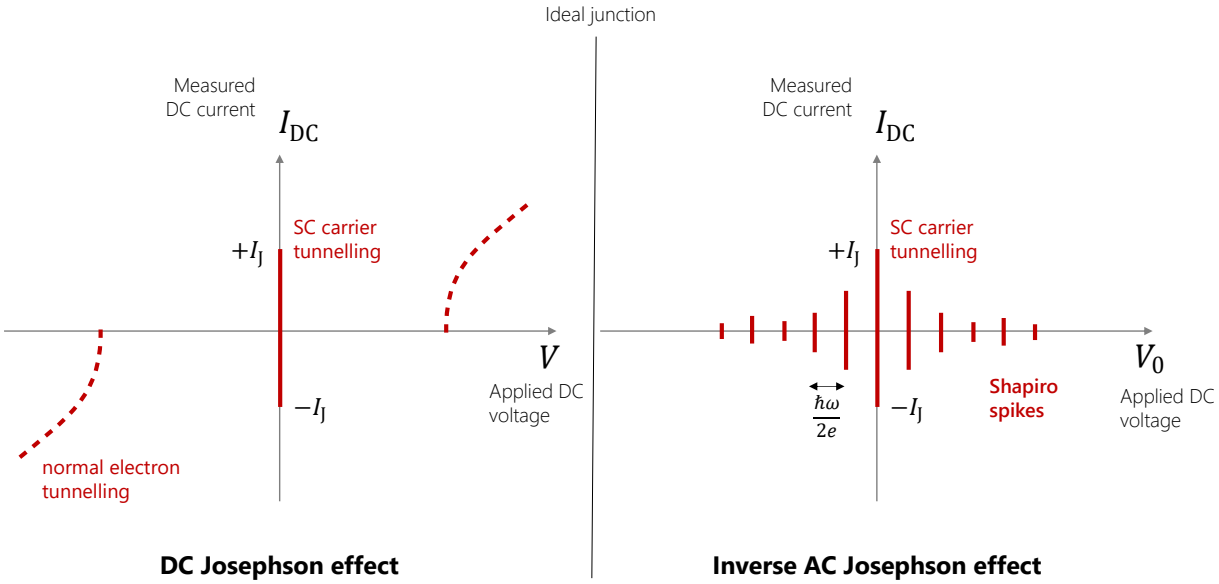


Figure 1.19: $I-V$ characteristic for ideal Josephson junctions. Modified from Prof. Alex Robertson's lectures.

1.10 Real Josephson Junctions

This is sometimes called the current-source model. A real circuit has resistive, capacitive and inductive effects, and potentially many more. Figure 1.20 shows the circuit we will be considering in this section Now, the total current is given by the sum from each component, that is if you remember your circuit theory

$$I_0 = I_J \sin(\phi) + \frac{V}{R} + C\frac{dV}{dt} \tag{1.71}$$

Combined with the second Josephson equation, Eq. (1.61), we can use it to form a second-order ODE in terms of ϕ

$$\frac{\hbar C}{2e}\frac{d^2\phi}{dt^2} + \frac{\hbar}{2eR}\frac{d\phi}{dt} = I_0 - I_J \sin(\phi). \tag{1.72}$$

Remark. If $V = 0$ is fixed, we just get the DC Josephson effect.

$\underline{V \neq 0 \textbf{ but } C = 0}$ We have a first-order autonomous ODE in ϕ

$$\frac{d\phi}{dt} = \frac{2eR}{\hbar} I_0 - \frac{2eRI_J}{\hbar}\sin\phi \tag{1.73}$$

Plotting this against ϕ for different values of I_0 is shown in Fig. 1.21.

(i) If $I_0 \leq I_J$, we get a steady-state solution (derivative oscillates between positive and negative).

(ii) If $I_0 > I_J$: we get a dynamic case, as the derivative is oscillating but always positive, so there is never a return to steady state.

Thus, we get a **step** at $V_0 = 0$, not a Shapiro spike. If we then apply a radio-frequency current on top of the DC, we get **Shapiro steps** (NOT peaks) whenever $2eV_0 = n\hbar\omega$, see Fig. 1.22.

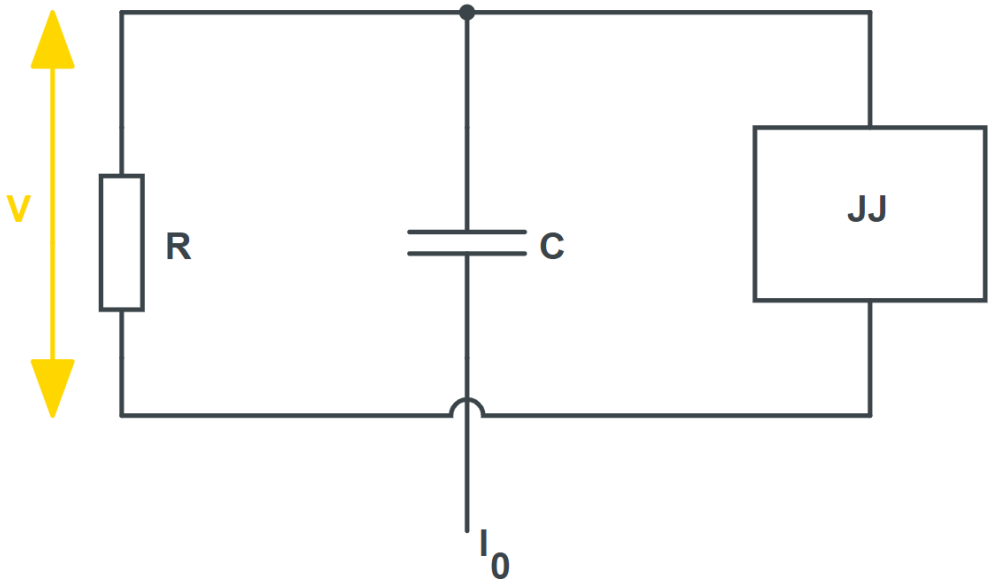


Figure 1.20: Example circuit with an ideal Josephson junction (JJ), a capacitor and resistor.

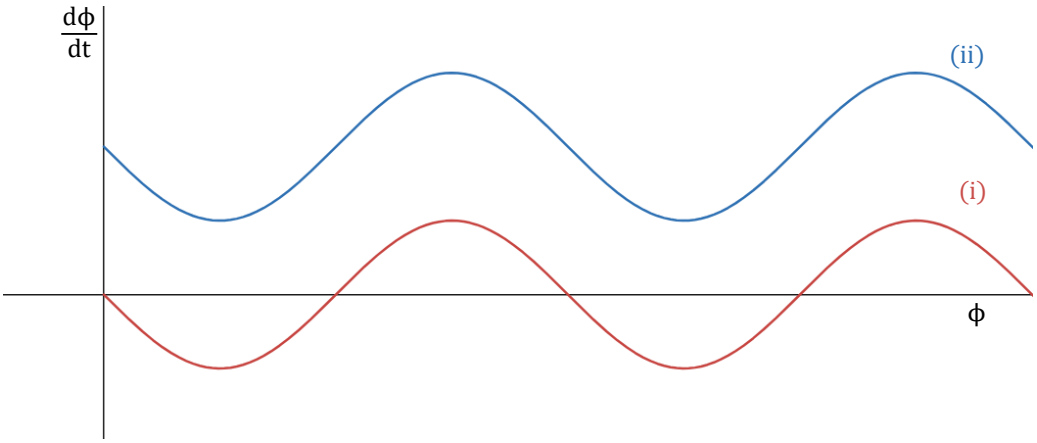


Figure 1.21: Plot of the first derivative of ϕ against ϕ for the case when $V \neq 0, C = 0$. There are 2 sub-cases (i) and (ii) to consider.

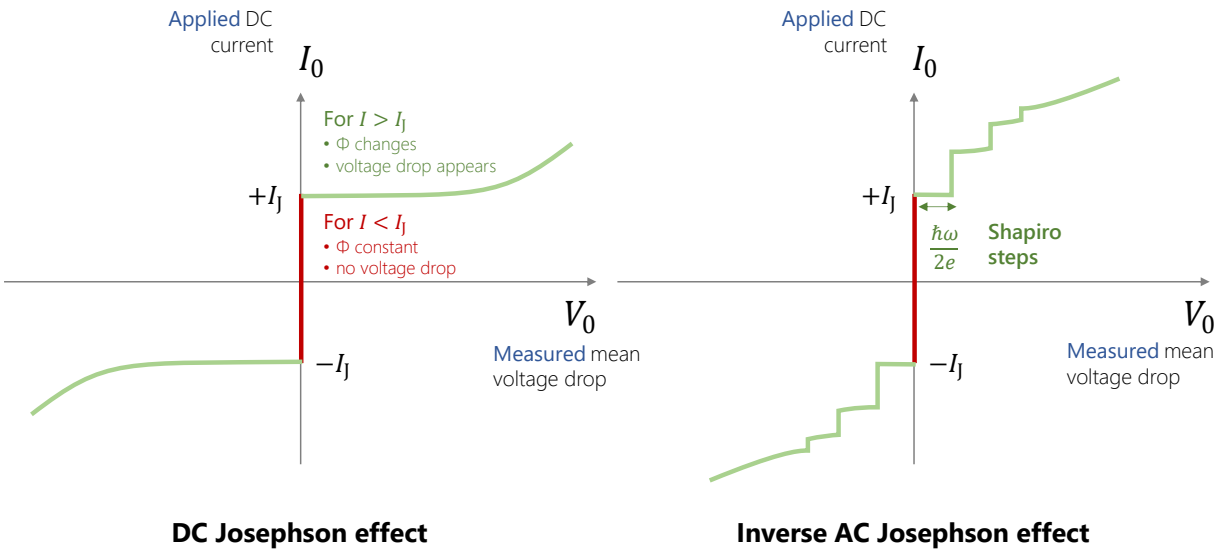


Figure 1.22: DC $I-V$ characteristic for an ideal Josephson junction placed in a resistive-capacitative circuit, for the DC and Inverse AC effect. Modified from Prof. Alex Robertson's lectures.

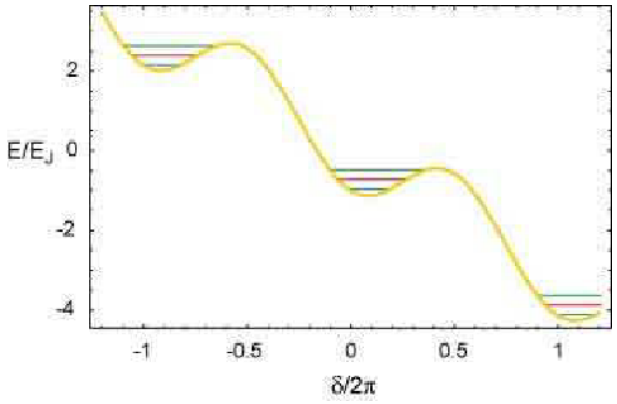


Figure 1.23: Tilted-washboard potential of a current-biased Josephson junction. Reproduced from Ref. [2].

$\underline{V, C \neq 0}$ We will rewrite the ODE in a more suggestive form

$$m\frac{d^2\phi}{dt^2} = -\frac{\partial U}{\partial \phi} - \gamma\frac{d\phi}{dt}, \tag{1.74}$$

where $m = \hbar C/2e, \gamma = \hbar/2eR$ and

$$U = \underbrace{-I_0\phi}_{\text{tilt}} - \underbrace{I_J\cos(\phi)}_{\text{bumps}} \tag{1.75}$$

This potential is called a **tilted-washboard potential**, an example is shown in Fig. 1.23.

We again get two situations:

(i) $I_0 < I_J$: steady state solution is *possible* depending on size of m. But they always occur when $\frac{\partial U}{\partial \phi} = 0 \implies I_0 = I_J \sin\phi$.

(ii) No steady state solution.

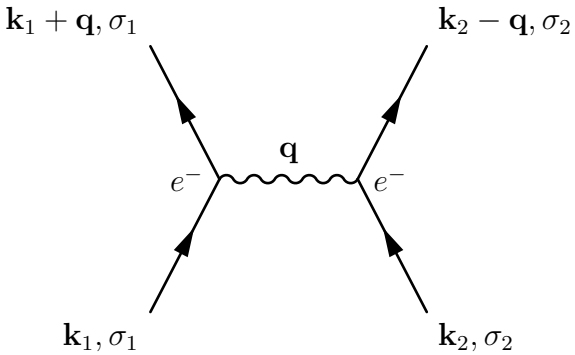


Figure 1.24: Electron-phonon coupling Feynman diagram. $\mathbf{k}_i$ are the wavevectors of the electrons, $\mathbf{q}$ the phonon momentum and σ_i the spin of the electrons.

1.11 BCS Theory

This section is an overview of BCS theory. For a more in-depth (mathematical) look, please see PX453 Advanced Quantum Theory. There are 2 main ideas in BCS theory:

The Isotope Effect: the critical temperature $T_c \propto m^{-\alpha}$, where m is the atomic mass and $\alpha \sim 1/2$. This implies **phonons** are involved in superconductivity.

Cooper pairs:

- Bare electrons repel
- In a metal, the Fermi sea screens quasiparticles, reducing the repulsion.
- Quasiparticles interact via the lattice

This interaction is

- Attractive for quasiparticles within $\pm\hbar\omega_D$ of the Fermi energy E_F, where ω_D is the Debye frequency.
- Leads to a bound state - **Cooper pairs** are bound states of electron pairs as shown in Fig. 1.24.

The ground state wavefunction of a Cooper pair

$$\psi_{\mathrm{CP}}\left(r_1, \sigma_1, r_2, \sigma_2\right) = e^{i\mathbf{k}_{\mathrm{TOT}}\cdot\mathbf{R}}\psi(\mathbf{r}_1 - \mathbf{r}_2)\phi^{\mathrm{spin}}_{\sigma_1\sigma_2}, \tag{1.76}$$

where

- $\mathbf{k}_{\mathrm{TOT}}$ is the crystal momentum (total wave)
- $\mathbf{R}$ is the centre-of-mass of the electrons
- The exponential part describes kinetic motion of the crystal
- $\psi(\mathbf{r}_1 - \mathbf{r}_2)$ is the orbital part
- $\phi^{\mathrm{spin}}_{\sigma_1\sigma_2}$ is the spin wavefunction

In the ground state ψ_{CP} links quasiparticles of momentum $\pm\mathbf{k}$ and so their sum is $\mathbf{k}_{\mathrm{TOT}} = 0$ if it is a symmetric situation (quasiparticles are bosons). However, the quasiparticles are fermions, so the wavefunction is antisymmetric:

$$\psi_{\mathrm{CP}}\left(r_1, \sigma_1, r_2, \sigma_2\right) = -\psi_{\mathrm{CP}}\left(r_2, \sigma_2, r_1, \sigma_1\right) \tag{1.77}$$

so one of the kinetic part or the orbital part must be antisymmetric. Cooper proposed for the $L = 0, s$ state, the orbital part is symmetric (not entirely true...) so the $\phi^{\text{spin}}_{\sigma_1\sigma_2}$ must be antisymmetric. Therefore,

$$\phi^{\text{spin}}_{\sigma_1\sigma_2} = \frac{1}{\sqrt{2}}\left(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle\right) \tag{1.78}$$

We can solve the time-independent Schrödinger equation for this state. The derivation is non-examinable, but the key steps in the derivation are

1. Cast the problem in terms of annihilation and creation operators (for electrons) and the phonon. See handout or PX453.
2. Get the BCS gap equation at $T = 0$

$$\Delta = |g_{\text{eff}}|^2 \sum_{\mathbf{k}} \frac{\Delta}{2E_{\mathbf{k}}} \tag{1.79}$$

3. Invoke electron-phonon coupling parameter $\lambda = |g_{\text{eff}}|^2 g(E_F) \ll 1$ where $g(E_F)$ is density of states at the Fermi level
4. $\Delta \ll \hbar\omega_D$ and the **BCS approximation (gap not a function of k)** so Taylor expand the result of the integral to first order

From this, we get the BCS gap/Cooper pair binding energy as

$$E_{\text{CP}} = -2\hbar\omega_D e^{-1/\lambda} \tag{1.80}$$

Remark. The $e^{-1/\lambda}$ explains why good metals are not superconductors, because they have a very small λ,and small λ makes E_{CP} very small, i.e. electrons never really interact with the lattice.

Additionally, $\omega_D \propto m^{-1/2}$ which is necessary for the isotope effect.
Definition 1.11.1. The **BCS energy gap** is simply $\Delta = -E_{\text{CP}}$

Now, if $T \neq 0$, the integral solved in the above list is different and more complicated.

$$1 = \lambda \int_0^{\hbar\omega_D} \frac{1}{E(\epsilon)} \tanh \frac{E(\epsilon)}{2k_B T} d\epsilon, \tag{1.81}$$

where $E(\epsilon) = \sqrt{\Delta^2 + \epsilon^2}$ is measured **relative to** μ. You can try to do this by hand (not examinable), but the key point is the energy gap, whilst independent of $\mathbf{k}$, is dependent on temperature:

- It roughly looks like the $n_{\text{s}}(T)$ curve in Fig. 1.16.
- The size of Δ depends on the number of superconducting charge carriers.
- So superconductivity is a **cooperative effect**

Evaluating the new integral as $\Delta \to 0$ and Taylor expanding around T_{c}

$$k_B T_{\text{c}} = 1.13\hbar\omega_D e^{-1/\lambda} \tag{1.82}$$

- Can enhance $T_{\text{c}}$by increasing λ or ω_{D}
- Since λ is usually small ($\lambda \ll 1$ usually), BCS theory predicts upper limit on $T_{\text{c}}$to be 30 - 40 Kelvin.

Combining Δ at 0 K, we get

$$2\Delta(T = 0) = 3.52 k_B T_{\text{c}}. \tag{1.83}$$

1.11.1 Evidence for the gap

- Heat capacity C
  - Cooper pairs don't contribute
  - Caused entirely by normal carriers excited above Δ
  - Process is more or less random/stochastic, so $C \propto e^{-\Delta/k_B T}$

  Experiments and theory both look like Fig. 1.15 (right). Further verification comes from plotting $\ln C$ against $1/T$ - we get a linear plot with negative gradient, and you find Δ from that
- Optical absorption
  - Only photons with energy $hf > \Delta$ can be absorbed by the superconductor
  - For pure elements, typical $2\Delta \sim 2$ meV, so $f \sim 10$ cm^{-1}, which is far-infrared.

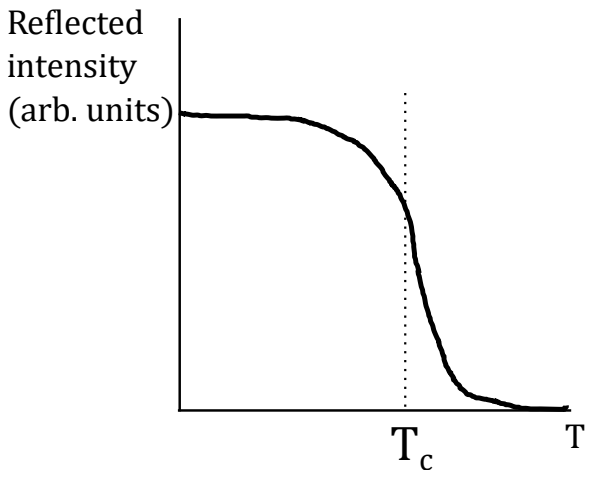


Figure 1.25: Relative intensity plot for a typical pure element in superconducting state.

- Electron tunnelling: Suppose we have a normal metal placed next to a superconductor with a surface barrier between them. Observations show
  - quasiparticles, *not pairs*, tunnel into and out of the superconductor
  - There are no quasiparticle states in the superconductor within Δ of E_F

  Experimentally, a plot of the conductance dI/dV against applied bias voltage V is done, see Fig. 1.26.
- Photoemission spectroscopy - similar plots to absorption
- Andreev reflections - suppose you place a metal and superconductor directly next to each other. You can get reflections where an incident electron is at the interface, but what is reflected back into the metal is a **hole** (opposite spin, same momentum) whilst the superconductor gains a charge of $2e$. This only occurs at energies less than Δ.

1.12 Unconventional Superconductors and Beyond

Unconventional superconductors are those whose behaviours cannot be predicted or explained by BCS theory. For example, high-temperature superconductors cannot be ex-

plained by BCS theory, since it predicts a limit of around 40 K. Cuprates for example, have a max $T_c \sim$ 150 K.

That being said, unconventional superconductors still exhibit

- Electron pairing
- zero resistance
- Meissner effect
- Enter mixed state with vortices for higher B fields than conventional superconductors.

1.12.1 Why are they unconventional?

There are possible explanations (which haven't theoretically gone far) with some observations - maybe one of you guys can figure it out?

- New pairing mechanism: in BCS theory we only assumed $L = 0$ state for the orbital wavefunction $\psi(\mathbf{r}_1 - \mathbf{r}_2)$, but we could have higher angular momentum states: Observations back this, with a generalised phase diagram of superconductivity to involve antiferromagnetism (AFM), such as in Fig. 1.28.[3]

1.12.2 Helium-3

Helium is unique because it is unable to solidify at $T = 0$ at standard pressure. There are two reasons for this

- Low atomic mass[4] and high zero-point motion

  The zero-point energy $E_0 = 1.5\hbar\omega_0$, where ω_0 is the vibrational frequency of an atom displaced from equilibrium, and $\omega_0 \propto m^{-1/2}$

- Weak attractive interaction due to high symmetry, so there are very weak van-der-Waals interactions.

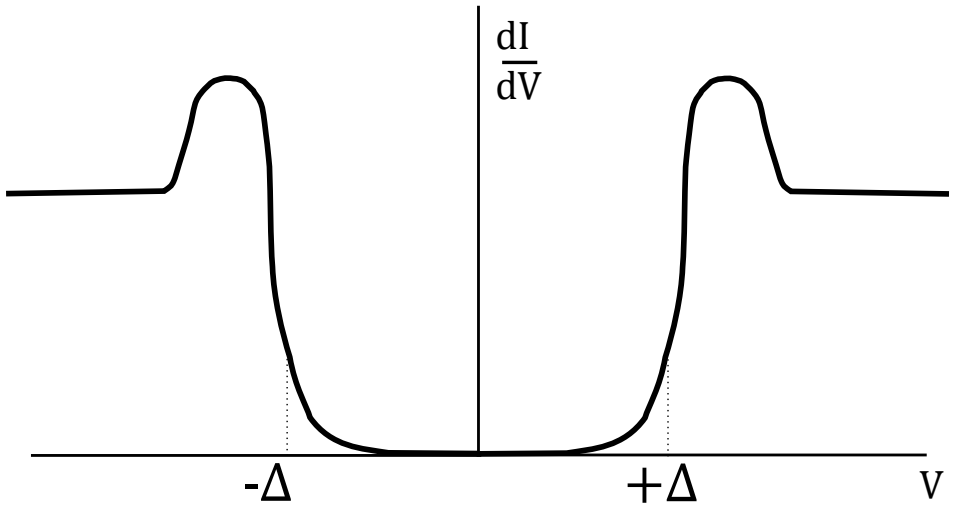


Figure 1.26: Graph of conductance against applied bias voltage V.

[3] Some examples of antiferromagnetism in superconductors in Bazarnik *et. al.* and Buzdin *et. al.*
[4] Of course, hydrogen with a lower atomic mass does solidify at 14 K and standard pressure.

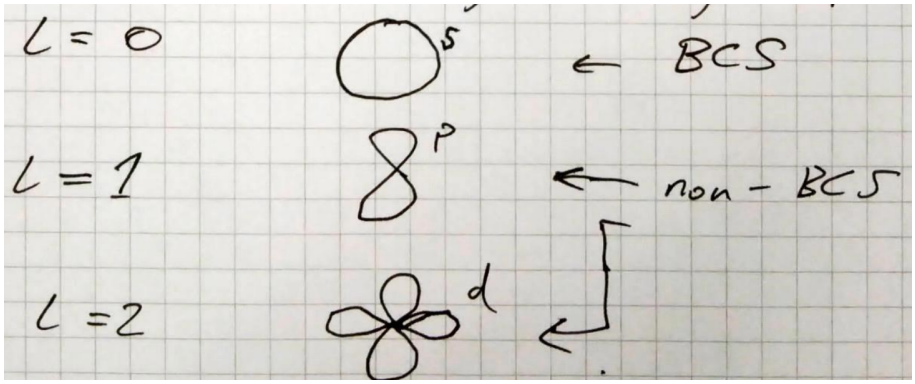


Figure 1.27: Depiction of different angular momentum states. Screenshot from Prof. Alex Robertson's lectures.

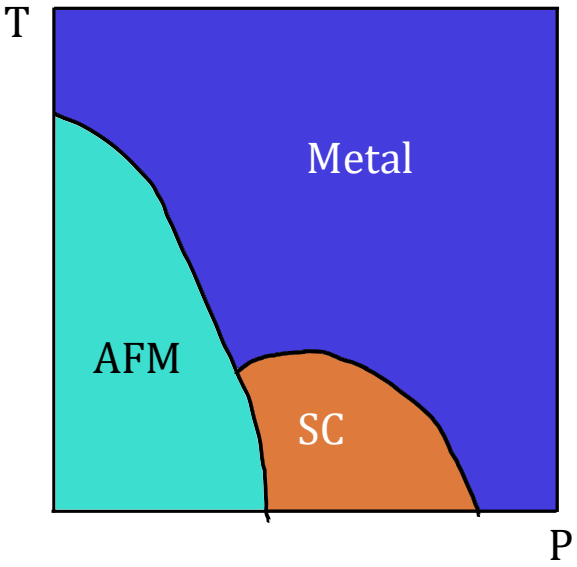


Figure 1.28: Typical phase diagram of unconventional superconductor.

The Lennard-Jones potential

$$V(r) = \epsilon \left(\frac{d^1 2}{r^1 2} - 2\frac{d^6}{r^6} \right) \tag{1.84}$$

where ϵ is the depth of the potential and d the position of the potential, then He and Ne are similar but Ne is more attractive. What do the Cooper pairs do? In ^{3}He, they are thought to form their own composite bosons, but if this is the case, the crystal lattice and thus phonons do not mediate the interaction. To allow vdW, we require $L > 0$, thus definitely making ^{3}He unconventional.

The Phase Diagram

There are three superfluid states in helium-3, with an intricate phase diagram as in Fig. 1.29. If there is **no** B field, then there are 2 phases

- A phase: described by ABM (Anderson-Brinkman-Morel) theory and high pressures
- B phase: Balian-Werthamer state (BW) theory.

If we turn on a magnetic field, the A phase splits:

- A_1 phase: only with B-field

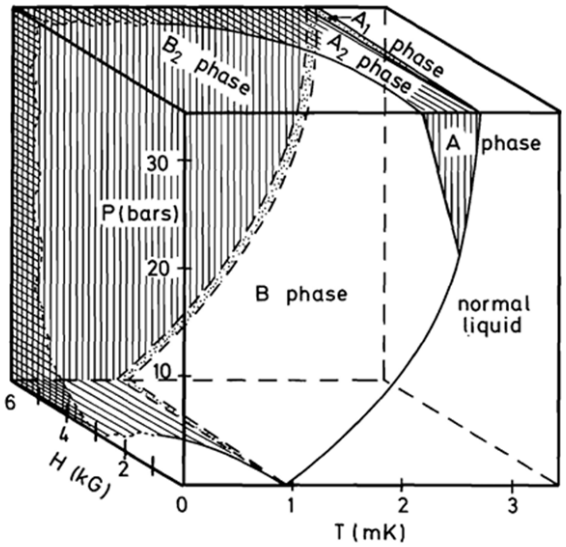


Figure 4.2 The P–T–H phase diagram of ^{3}He, showing the normal Fermi liquid as well as the superfluid A, A_1 and B phases. The superfluid A_2 and B_2 phases correspond to the A and B phases in the presence of a magnetic field. The values of the transition temperature shown in this figure are a composite of the results of several groups and should only be taken as semiquantitative (see the discussion in Section 4.1). A complete list of T_c values obtained by Greywall (1986) is shown in Table 4.1.

Figure 1.29: Helium-3 phase diagram, reproduced from *The Superfluid Phases of Helium 3* Dieter Vollhardt and Peter Wölfle.

- A_1 phase: same to ABW just with the field.
- B_2 phase: same to B phase but the field is turned on

All of these phases consist of Cooper pairs of ^{3}He quasiparticles with pairing wavefunction satisfying $S = 1, L = 1$ - this is a **spin-triplet** p-wave pairing . Contrast this to BCS superconductors, which are spin-singlet, s-wave pairing.

For $S = 1$, there are 3 allowed states $m_S = -1, +1, 0$, which correspond to the 3 states:

$$|\downarrow\downarrow\rangle \qquad |\uparrow\uparrow\rangle \qquad \frac{1}{\sqrt{2}}\left(|\downarrow\uparrow\rangle + |\uparrow\downarrow\rangle\right) \tag{1.85}$$

Therefore, the pair wavefunction is a linear superposition of all 3:

$$\begin{aligned}\Psi = \psi_{1,+}(\mathbf{k})\,|\uparrow\uparrow\rangle + \psi_{1,0}\,(\mathbf{k})\,(|\downarrow\uparrow\rangle + |\uparrow\downarrow\rangle) \\ + \psi_{1,-}\,|\downarrow\downarrow\rangle\end{aligned} \tag{1.86}$$

where $\psi_{1,+}, \psi_{1,0}, \psi_{1,-} : \mathbb{R}^3 \to \mathbb{C}$ are complex-valued amplitudes. However, the pairing wavefunction should also account for $L = 1 \implies m_L = 0 \pm 1 \implies 9$ substitutes. This makes the wavefunction **highly anisotropic**, leading to a complex phase diagram, a different beast to BCS entirely.

Chapter 2

Optical Properties

This chapter of the module studies light-matter interaction. Studying this has led to numerous advancements

- Widgets (illumination, lasers, sensors, internet etc.)
- Probing condensed matter: x-ray diffraction, dielectrics, bandstructures (ARPES), ultrafast dynamics
- Fundamental physics: QM, spontaneous symmetry breaking, renormalisation, topological variants, emergent behaviour

All of these are **testable** (in the experimental sense, though some of these will be examinable!).

2.1 Optical Processes

A rundown of different optical processes is shown in Fig. 2.1

Definition 2.1.1. The **refractive index** n is a ratio of the speed of light in vacuum, c, to the speed of light in a material v, such that $n = c/v$. In particular, $n \geq 1$.

Definition 2.1.2. Scattering is the process of light changing direction. **Elastic scattering** is when $\lambda_{\text{in}} = \lambda_{\text{out}}$, i.e. wavelength unchanged. **Inelastic scattering** is when they are not equal.

Definition 2.1.3. Luminescence is the emission of light by excited states of atoms or defects in matter. Excited states may be produced by incident light, called **photoluminescence** (PL).

Definition 2.1.4. Fluorescence is a longer-lived excited state, and so there is a delay between Δt between excitation and light emission. Fluorescence often emits a lower wavelength of light but not always.

Definition 2.1.5. The **absorption coefficient** α describes how far into a material light of a particular wavelength can travel before being absorbed.

Lemma 2.1.1. ***Beer's law*** *describes optical absorption using α. Let $I(z)$ be the* ***optical power per unit area****. Then*

$$I = I_0 e^{-\alpha z}, \tag{2.1}$$

assuming the material is aligned along z. Also, $I \propto E^2$ where $E = |\mathbf{E}|$ is the magnitude of the electric field of light, so

$$E = E_0 e^{-\alpha z/2} \tag{2.2}$$

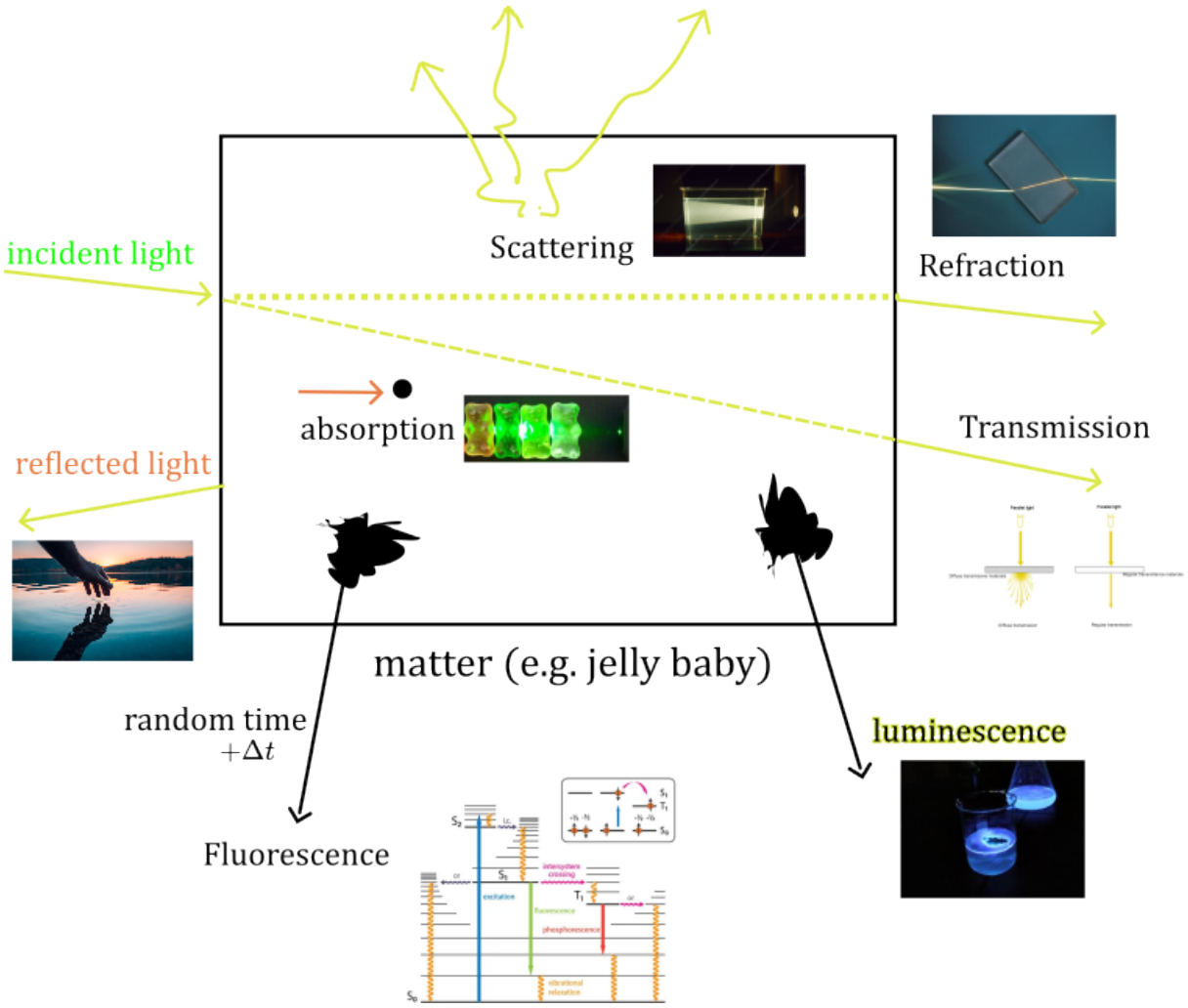


Figure 2.1: A schematic of optical processes.

2.1.1 Classical vs Quantum

In classical physics, transmission, reflection, absorption and refraction were all describable using a **complex refractive index**

$$\tilde{n} = n + i\kappa \tag{2.3}$$

where n is the regular refractive index as in Definition 2.1.1 and κ is the **extinction coefficient**, which is related to the attenuation of light through an optical medium, and therefore is related to α. For the rest of this chapter, the real part of $\tilde{n}$ is denoted as n or $\Re\tilde{n}$. The imaginary component is denoted just κ or $\Im\tilde{n}$

The **Stokes shift** is a change of λ of scattered light in (photo-)luminescence and **requires QM to explain it**. A sudden change in wavelength $\Delta\lambda$ implies some quantum of energy ΔE_{photon} to satisfy conservation of energy. But a quantum of energy could be many things: energy levels in atoms, molecules, defects; or energetic quasiparticles like phonons and plasmons.

2.1.2 Complex refractive index

Electromagnetic radiation is characterised by Maxwell's equations. In classical physics, light is thought of as a wave, and this leads to the wave equation

$$\frac{\partial^2 E}{\partial z^2} = \frac{1}{c^2}\frac{\partial^2 E}{\partial t^2} \tag{2.4}$$

where $E = E(z,t) = E_0 e^{i(kz-\omega t)}$ is the electric field propagating along z. In a non-vacuum medium, both the permittivity ϵ_r and permeability μ_r scale by relative factors:

$$\epsilon_0 \to \epsilon_r = \epsilon_0 \tilde{\epsilon}_r \qquad\qquad \mu_0 \to \mu_r = \mu_0 \tilde{\mu}_r \tag{2.5}$$

where $\tilde{\epsilon}_r$ and $\tilde{\mu}_r$ are the **relative permittivity and permeability** respectively. For this section, we will assume $\tilde{\mu}_r = 1$, that the material is non-magnetic at optical frequencies. The speed of light in the material then changes

$$v = \frac{1}{\sqrt{\epsilon_0 \mu_0}} \to \frac{1}{\sqrt{\epsilon_r \mu_0}} \tag{2.6}$$

Remark. Of course, if $\tilde{\mu}_r \neq 1$, then $\mu_0 \to \mu_r$.

Now, $\Re\tilde{n} = n = c/v = \sqrt{\tilde{\epsilon}_r} = \sqrt{\tilde{\epsilon}_{\mathrm{r}}(\omega)}$. Now, what actually is $\tilde{\epsilon}_{\mathrm{r}}$? It describes how much the electric field is weakened compared to the strength of true vacuum. It is therefore a **material property** and has to be experimentally determined for each material. $\tilde{\epsilon(\omega)}_{\mathrm{r}}$ depends on how the electric charges in the material react to the oscillating electric field of light, and there are multiple responses

- For a **bound charge**, we get **polarisation**. You can see this in a dipole
- For a **free charge**, you can either get **oscillations** (plasmons) or **absorption**. The latter has both a classical and QM treatment.

We know that Eq. (2.3) will lead to a dampening of $E(z)$, so

$$E = E_0 e^{i(kz-\omega t)},$$

where $\omega = vk \iff k = \frac{\omega}{c}(n + i\kappa) = \sqrt{\tilde{\epsilon}_{\mathrm{r}}}$. Thus our expression for $E(z)$ generalises to

$$E(z) = E_0 \exp\left[-\frac{\kappa\omega z}{c}\right] \exp\left[i\omega\left(\frac{nz}{c} - t\right)\right] \tag{2.7}$$

- The first exponential $\exp[-\kappa\omega z/c]$ is an **absorption term**. Referring back to Eq. (2.2), we see that
$$\alpha = \frac{4\pi\kappa}{\lambda} \tag{2.8}$$
where λ is the wavelength of light in the medium.
- The second exponential is a phase shift.

Therefore, $\tilde{\epsilon}_{\mathrm{r}}$ must be complex and we write it as $\tilde{\epsilon}_{\mathrm{r}} = \epsilon_1 + i\epsilon_2$. Expanding both sides:

$$\tilde{\epsilon}_{\mathrm{r}} = (n + i\kappa)^2 = n^2 + \kappa^2 - 2in\kappa = \epsilon_1 + i\epsilon_2. \tag{2.9}$$

Equating coefficients,

$$\epsilon_1 = n^2 - \kappa^2 \qquad\qquad \epsilon_2 = 2n\kappa \tag{2.10}$$

Since n, κ can be rewritten in terms of ω or λ, we can completely determine ϵ_1 and ϵ_2 as functions of these quantities too. We sometimes call $\tilde{\epsilon}_{\mathrm{r}}$ the **dielectric function** in this context.

Definition 2.1.6. The **weakly absorbing limit** is defined as a medium where $n \gg \kappa$. Here, $n \sim \sqrt{\tilde{\epsilon}_{\mathrm{r}}}$ and $\kappa = \epsilon_2/2n$, then $\alpha = 2\pi\epsilon_2/n\lambda$. Namely, $\Re\tilde{\epsilon}_{\mathrm{r}}$ determines refractive index and $\Im\tilde{\epsilon}_{\mathrm{r}}$ determines the absorption coefficient.

2.2 Calculating the dielectric function

There are multiple ways to find $\tilde{\epsilon}_\mathrm{r} = \epsilon_1 + i\epsilon_2$

1. Pure classical model: EM theory plus dipole oscillator model
2. Semiclassical model: use QM to define energy levels, excitations, quasiparticles and model light classically.
3. Sidestep $\tilde{\epsilon}_\mathrm{r}$ by treating light and matter with QM.

In this section, we will focus on method 1.

Assumptions:

- Light is an oscillating electric field (ignore magnetic component)
- Charges in condensed matter respond by oscillating

The most general method is everyone's favourite: the classical dipole oscillator shown in Fig. 2.2. The simplified atom model has a resonant frequency ω_0

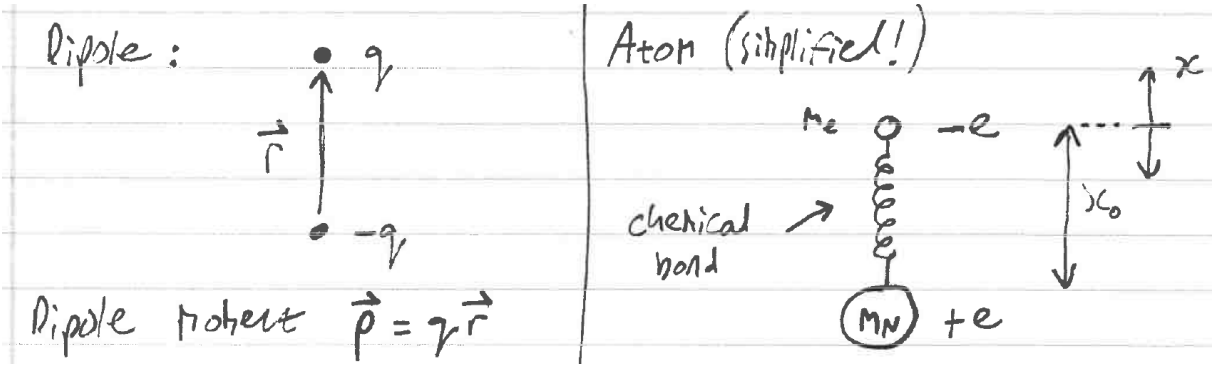


Figure 2.2: Classical dipole oscillator model. Screenshot from Prof. Gavin Bell's lectures.

$$\omega_0 = \sqrt{\frac{k'}{\mu}} \tag{2.11}$$

where k' is the **spring constant** (bond stiffness) and μ is the **reduced mass**, defined as **Definition 2.2.1.** The **reduced mass** is the effective mass that a multi-body system has as if it was behaving as one mass. For two-bodies (nucleus and electron), it is defined as

$$\mu = \frac{1}{m_\mathrm{n}^{-1} + m_\mathbf{e}^{-1}} \tag{2.12}$$

Since $m_\mathrm{n} \gg m_\mathrm{e}, \mu \sim m_\mathrm{e}$.

Now, if we *apply* an electric field $E(t) = E_0 e^{-i\omega t}$, we then solve the inhomogeneous second-order ODE for the classical damped oscillator:

$$m_e \frac{d^2x}{dt^2} + m_e\gamma\frac{dx}{dt} + m_e\omega_0^2 x = -eE \tag{2.13}$$

- γ is the **damping factor** and the entire term $m_\mathrm{e}\gamma dx/dt$ is the **drag force**.
- $m_\mathrm{e}\omega_0^2 x$ is the **spring force**.

- $m_e d^2x/dt^2$ is the resultant force.
- The entire ODE is a generalised $f = ma$ situation.

We try the solution $x(t) = \tilde{x}_0 e^{-i\omega t}$[1]

$$-m_e\omega^2\tilde{x}_0 - im_e\gamma\omega\tilde{x}_0 + m_e\omega_0^2\tilde{x}_0 = -eE_0. \tag{2.14}$$

Rearranging for $\tilde{x}_0$ gives

$$\tilde{x}_0 = \frac{-eE_0}{m_e\left(\omega_0^2 - \omega^2\right) - i\gamma\omega m_e}. \tag{2.15}$$

Definition 2.2.2. The **dipole moment** $p(t) = qr(t)$ where $r(t)$ is the position. Note this is generally a vector quantity.

Indeed, our $p(t) = -e\tilde{x}_0 e^{-i\omega t}$. Now, a real material is usually not made up of a single oscillator, but many - let's suppose there are N oscillators per unit volume. Then the **bulk polarisation** $P(t)$ is

$$P(t) = p(t)N = \frac{Ne^2}{m_e}\frac{1}{\omega_0^2 - \omega^2 - i\gamma\omega}E_0 e^{-i\omega t} \tag{2.16}$$

From electromagnetism, we know

Lemma 2.2.1. *The* ***displacement field*** $\mathbf{D} = \epsilon_0\mathbf{E} + \mathbf{P} = \epsilon_0\epsilon_r\mathbf{E}$. $\mathbf{P}$ *here is the total polarisation.*

Indeed, we will break down the total polarisation into two parts: *static* and *dynamic*:

$$\mathbf{P} = \underbrace{\epsilon_0\chi\mathbf{E}}_{\text{static}} + \underbrace{\mathbf{P}(t)}_{\text{dynamic}}, \tag{2.17}$$

and χ is the electric susceptibility. We substitute Eq. (2.17) into the expression inside Lemma 2.2.1, and rearrange for $P(t)$

$$\begin{aligned} \epsilon_0\mathbf{E}(1+\chi) + \mathbf{P}(t) &= \epsilon_0\tilde{\epsilon}_\mathrm{r}\mathbf{E} \\ P(t) &= \tilde{\epsilon}_\mathrm{r}\epsilon_0 E - \epsilon_0 E(1+\chi) \end{aligned} \tag{2.18}$$

Now, we equate this to Eq. (2.16). Notice in Eq. (2.18), we can factor out $\epsilon_0 E$. But then this factor cancels with the same factor in Eq. (2.16). Hence, rearranging for $\tilde{\epsilon}_\mathrm{r}$ gives

$$\tilde{\epsilon}_\mathrm{r} = 1 + \chi + \frac{Ne^2}{\epsilon_0 m_e}\frac{1}{\omega_0^2 - \omega^2 - i\gamma\omega} \tag{2.19}$$

It is now the job to find ϵ_1 and ϵ_2. This can be done by multiplying the fraction term by

$$\frac{\omega_0^2 - \omega^2 + i\gamma\omega}{\omega_0^2 - \omega^2 + i\gamma\omega}$$

to cancel out imaginary components in the denominator, expanding and rearranging. If you do that, you will get

$$\begin{aligned} \varepsilon_1 &= 1 + \chi + \frac{Ne^2}{\varepsilon_0 m_e}\frac{{\omega_0}^2 - \omega^2}{\left({\omega_0}^2 - \omega^2\right)^2 + (\gamma\omega)^2} \\ \varepsilon_2 &= \frac{Ne^2}{\varepsilon_0 m_e}\frac{\gamma\omega}{\left({\omega_0}^2 - \omega^2\right)^2 + (\gamma\omega)^2} \end{aligned} \tag{2.20}$$

[1] You may recall that this ODE has 2 solutions with 2 different frequencies. We brush this under the rug for the time being.

Figure 2.3: Example plot of n and α against ω from the expressions derived above. Screenshot from Prof. Gavin Bell's lectures because I couldn't get a nice enough looking one in Desmos or Tikz with good parameters.

Behaviour of ϵ_1, ϵ_2

- ϵ_2 peaks at $\omega = \omega_0$ with FWHM of γ - **maximum absorption**.
- ϵ_1 steadily increases when approaching ω_0, drops sharply and then climbs again. As $\omega \to 0, \epsilon_1 \to 1 + \chi + \Delta$ but as $\omega \to \infty, \epsilon_1 \to 1 + \chi$. Here, $\Delta \in \mathbb{R}$.
- This type of model is sometimes called a Lorentzian harmonic oscillator.

Now, real materials are not just made up of many copies of a single oscillator, but they are made up of multiple *types* of oscillators, corresponding to different mechanisms, each with some 'oscillator strength' f_j. Since this is just a prefactor, the expressions for ϵ_1, ϵ_2 immediately generalise:

$$\begin{aligned} \epsilon_1 &= 1 + \sum_j \frac{Ne^2}{\varepsilon_0 m_e} \frac{f_j(\omega_j^2 - \omega^2)}{\left(\omega_j^2 - \omega^2\right)^2 + (\gamma_j \omega)^2} \\ \epsilon_2 &= \sum_j \frac{Ne^2}{\varepsilon_0 m_e} \frac{f_j \gamma_j \omega}{\left(\omega_j^2 - \omega^2\right)^2 + (\gamma_j \omega)^2} \end{aligned} \tag{2.21}$$

The χ is constructed from all the **low-frequency** limits of the oscillator tails, that is all of the Δ_j. Using Eq. (2.10), we can plot the behaviour of n and α and use the weakly absorbing limit, Definition 2.1.6. Example graphs look like in Fig. **??**. Material examples: The first example that could be described by this model is **uric acid**, shown in Fig. 2.4

- Uric acid has 3 clear peaks due to electronic transitions in and around the heterocyclic ring
- Strong chance to be modelled successfully by oscillators
- The FWHM is about 25 nm, the question is which γ_j does that correspond to?

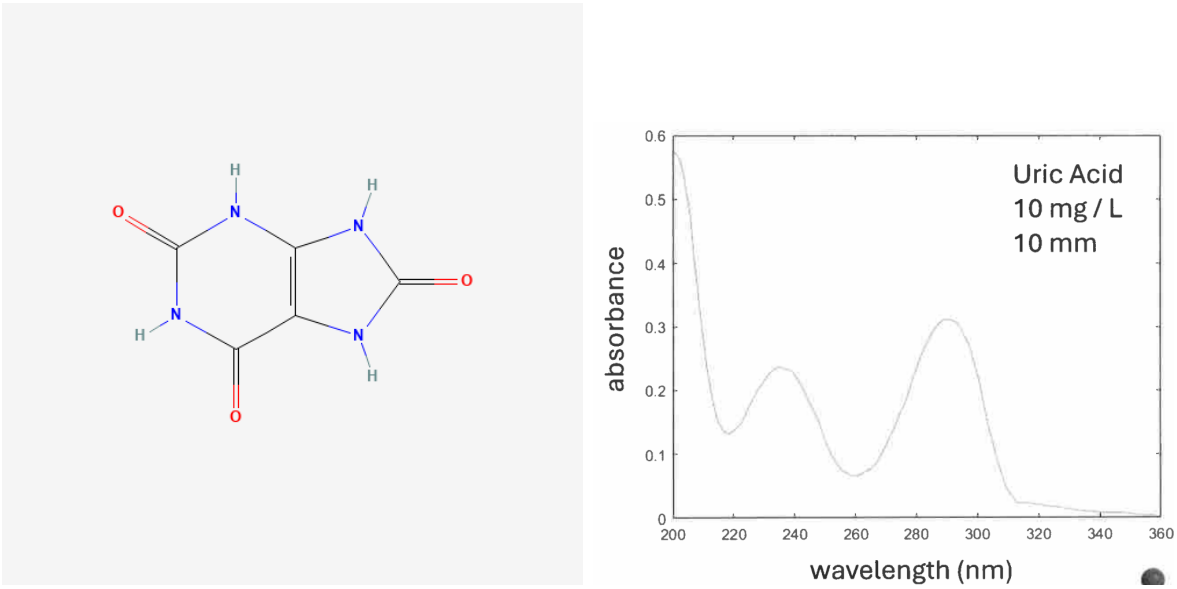


Figure 2.4: Uric acid molecule (left) and its absorption spectrum.

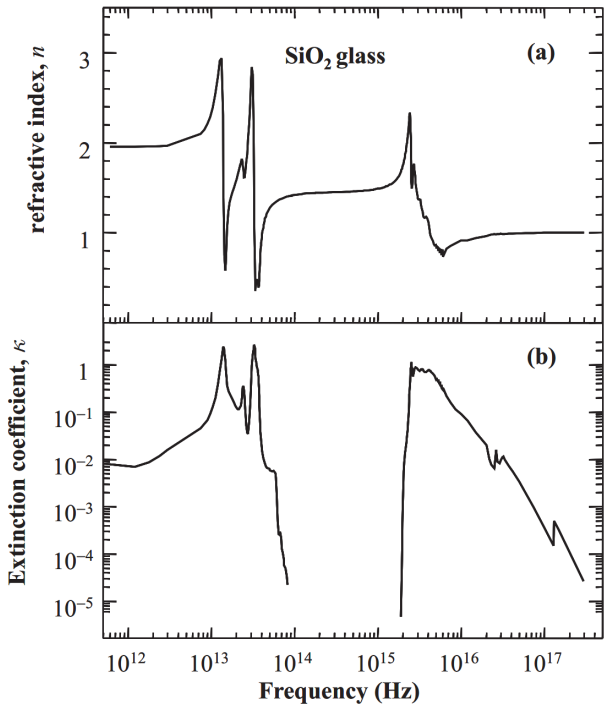


Figure 2.5: Refractive index (top) and extinction coefficient (bottom) plot against frequency for silica glass, SiO_2.

Another example is **silica glass**, whose n and κ is shown in Fig. 2.5.

- $n \to 1$ at higher frequencies, this is good for our model
- Before 10^14 Hz, there are 2 strong 'wiggles' in n which have peaks in κ, corresponding to **vibrational modes**
- There is an additional peak at 2×10^15 Hz but this is an *absorption edge* rather than a peak - our model doesn't have absorption edges and this peak occurs due to interband electronic absorption
- Minor peaks in κ around $10^16, 10^17$ Hz but no corresponding peak in n - **core electron transitions**. Our model doesn't have this feature either
- $n \gg \kappa$ except near absorption peaks, good for the weakly absorbing limit we used
- Transparency across the visible range, where $\kappa = 0$

2.2.1 Refractive index

Suppose you have 2 optical materials at an interface with refractive index n_1, n_2 and light is incident from medium 1. In general, there will be **both** transmission and reflection. The proportion of light that goes into each mode is given by the **transmission** and **reflection coefficients** T, R respectively. By conservation of energy

$$R + T = 1 \tag{2.22}$$

and in particular,

$$R = \left| \frac{n_1 - n_2}{n_1 + n_2} \right|^2 \tag{2.23}$$

Now, $n = c/v$ but $n < 1$ is allowed, implying $v > c$. This seems like a dilemma, but we need to be careful about which velocity we are considering. Light travels in **wave packets** so the velocity of information transfer is the **group velocity**

$$v_g = \frac{\partial \omega}{\partial k} \tag{2.24}$$

and *not* the phase velocity ω/k. We now aim to find an expression for v_g in terms of the refractive index:

$$\begin{aligned}
&v = \frac{\omega}{k} = \frac{c}{n}; \quad d\omega = \frac{\partial \omega}{dk} dk + \frac{\partial \omega}{\partial n} dn; \quad \omega = \frac{kc}{n} \\
&\frac{\partial \omega}{\partial k} = \frac{c}{n}; \quad \frac{\partial \omega}{\partial n} = -\frac{kc}{n^2}; \quad d\omega = \frac{c}{n} dk - \frac{kc}{n^2} dn \\
&v_g = \frac{d\omega}{dk} = \frac{c}{n} - \frac{kc}{n^2} \frac{dn}{dk} = v \left(1 - \frac{k}{n} \frac{dn}{dk} \right) \\
&v_g = v \left(1 - \frac{\omega n}{cn} \frac{dn}{d\omega} \frac{c}{n} \right) = v \left(1 - \frac{\omega}{n} \frac{dn}{d\omega} \right)
\end{aligned} \tag{2.25}$$

Both $n, dn/d\omega$ usually positive in the visible part of the spectrum, so $v_g < v$.

Another property of n is that it changes across the visible spectrum and is affected by the tails of the phonon and electron absorption features. We can see this through **Snell's law**

$$\frac{\sin \theta_1}{\sin \theta_2} = \frac{n_2}{n_1} = \frac{v_1}{v_2} \tag{2.26}$$

with $n_2 > n_1, v_2 < v_1$.

Since $n = n(\lambda)$ (because $v = v(\lambda)$), different wavelengths are refracted by different amounts. In particular, **shorter-wavelength light is refracted more strongly than longer-wavelengths**. This is called **dispersion**.

There is also the notion of a **time domain** with n. A pulse of light containing multiple frequencies will spread out in time, i.e. diverge. This is an important phenomenon to take into account for **fibre-optic cables** and picosecond pulses for data storage.

A pulse is a wave packet of light, so it is characterised by a group velocity v_g. A fractional change $1/v_g$ with wavelength λ disperses travel times τ. We would like the 'spread time' $\Delta\tau$ as a function of the wavelength spread $\Delta\lambda$

$$\frac{1}{v_y} = \frac{dk}{d\omega} = \frac{d}{d\omega}\left[\omega\frac{n(\omega)}{c}\right] = \frac{n(\omega)}{c} + \frac{u}{c}\frac{dn(\omega)}{d\omega} \tag{2.27}$$

$$\tau = \frac{L}{v_g} = \frac{L}{c}\left[n(\omega) + \omega\frac{dn}{d\omega}\right] \tag{2.28}$$

$\omega = 2\pi c/\lambda$:

$$\omega\frac{dn(\omega)}{d\omega} = \frac{2\pi c}{d\lambda}\frac{dn}{d\lambda}\left(\frac{-2\pi c}{\omega^2}\right) = \frac{-\lambda^2}{\lambda}\frac{dn}{d\lambda} = -\frac{\lambda dn(\lambda)}{d\lambda} \tag{2.29}$$

$$\tau = \frac{L}{c}\left[n(\lambda) - \lambda\frac{dn(\lambda)}{d\lambda}\right], \tag{2.30}$$

The square brackets term can be thought of as a *group refractive index*. Now,

$$\Delta\tau = \frac{d\tau}{d\lambda}\Delta\lambda \tag{2.31}$$

$$\frac{d\tau}{d\lambda} = \frac{L}{c}\left[\frac{dn(\lambda)}{d\lambda} - \lambda\frac{d^2n(\lambda)}{d\lambda^2} - \frac{dn(\lambda)}{d\lambda}\right] \tag{2.32}$$

giving

$$\Delta\tau = L|D|\Delta\lambda \qquad\qquad |D| = \frac{\lambda}{c}\frac{d^2n}{d\lambda^2}, \tag{2.33}$$

where D is the **material dispersion parameter** having units of (ps nm^{-1} km^{-1}). This can be interpreted as the picosecond-time dispersion for every nanometre-change in wavelength over a km-distance.

Since fibre-optic cables are many kilometres long, it is an important aspect to consider, with sub-picosecond data pulses. An example graph of $dn/d\lambda$ is shown in

2.3 Band structure

Atoms have discrete electronic states (1s, 2p etc.), each with a well-defined E_B, called the **binding energy**, the energy needed to remove an electron from that energy level.

Electronic states with multiple angular momentum values like 3p, also have a **spin-orbit splitting** energy Δ. For p-states, the total angular momentum $j = 1/2, 3/2$; for d-states it is $j = 3/2, 5/2$ etc. The pattern of E_B's across elements inspired the development of QM and observations led to things such as Moseley's law,

$$\sqrt{f} \propto Z \tag{2.34}$$

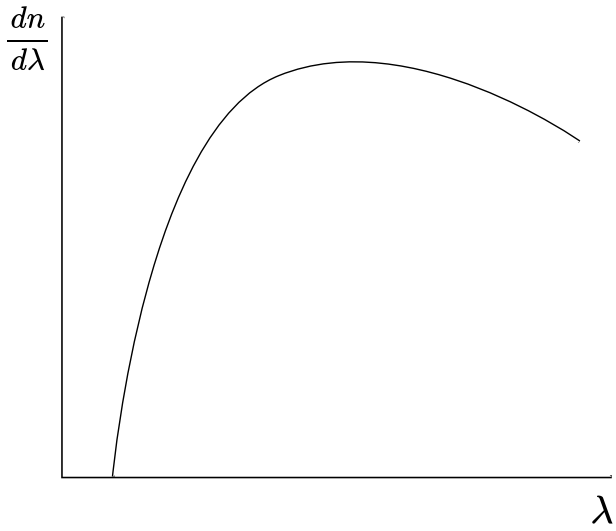


Figure 2.6: Graph of $dn/d\lambda$ against λ for silica.

where f is the frequency of emitted x-ray photon and Z the atomic number.

In a general condensed matter system, the *outer* (valence) electrons interact with each other which forms new quantum states. In this section, we will think about **bulk, crystalline materials**. We are thus excluding

- glasses (they are amorphous)
- materials with long-range order
- molecules or nanostructures

In a bulk, crystalline material, the **core levels** are **not** involved in bonding and thus do not interact to form new states. The valence levels *do*, and now our job is to figure out *how* they mix.

A crystal is nothing more than a **periodic arrangement of atoms**. In 3D space, each atom's position can be described by a **lattice vector R**

$$\vec{R} = n_1\vec{a} + n_2\vec{b} + n_3\vec{c}, \tag{2.35}$$

where $n_1, n_2, n_3 \in \mathbb{Z}$ and $\vec{a}, \vec{b}, \vec{c} \in \mathbb{R}^{1,3}$ defines the **unit cell** of the crystal. Please see PX385 if you need a recap of unit cells, but it is not necessary right now.

Because crystals are periodic, translating by **R** will get you to the same place relative to the entire universe - in particular, you will get identical local environment.

$$V(\vec{r} + \vec{R}) = V(\vec{r}) \Rightarrow |\psi(\vec{r} + \vec{n})|^2 = |\psi(\vec{r})|^2 \tag{2.36}$$

$$\psi(\vec{r} + \vec{R}) = \psi(\vec{r})e^{i\vec{k}\cdot\vec{R}} \tag{2.37}$$

This is nothing more than **Bloch's theorem**, which states that potential of a periodic system is equal to the product of a periodic function $\psi(\vec{r})$ times a plane wave.

So to actually work out the quantum states? Using **perturbation theory** (every undergraduate's favourite) we can get approximate solutions.

- Nearly-free electron
- Tight binding

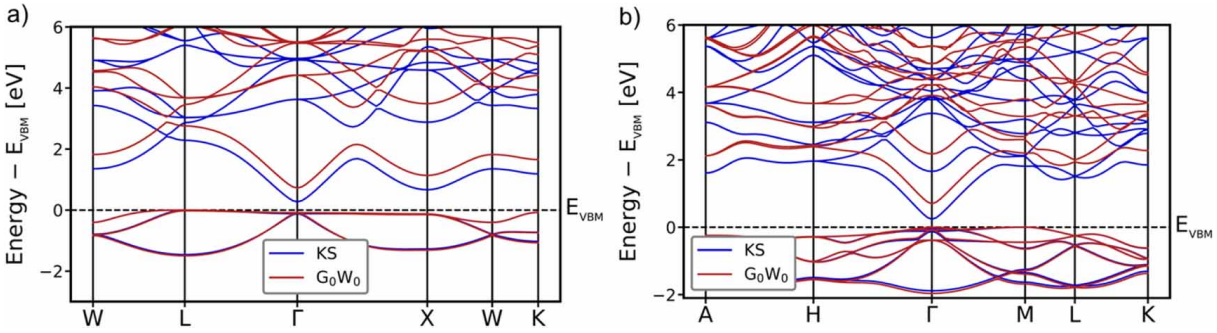


Figure 2.7: Computationally predicted bandstructure of a single unit cell of two phases of Na-K-Sb, alsigned at their respective valence band maxima. Reproduced from C. Xu *et al* [7].

- $k \cdot p$ model
- density-functional theory (great for ground state solutions, also my research!)
- many-body perturbation theory (excited states analysis)

2.3.1 Recap: interpreting band structures

Brief recap of bandstructures. These are plots of energy along **paths through k-space**. These paths are denoted by various letters like in Fig. 2.7. These letters are called **high-symmetry points** and correspond to special points in the Brillouin Zone.

In crystal theory, there is an underlying group theory of operations - rotations, reflections and translations, which you can do to certain crystal structures that leave particular points invariant - the classification of these crystals are called **space groups**, which you do not need to know about for the exam.

The invariant points are given special letters. Typically:

- $\Gamma - (0, 0, 0)$
- $X - (1/2, 0, 1/2)$
- $L - (1/2, 1/2, 1/2)$

Note: these letters (other than Γ) may be slightly different depending on the order of the reciprocal lattice vectors, the crystal structure and literature.

If a path through the BZ has the potential for occupation (or has occupied states), a line will be plotted. At these lines, $g(E)$, the density of states, is definitely not 0. A point in the bandstructure tells you where in the BZ you are and how much energy you are expected to have.

it is important to note you *cannot* directly extract which electrons in which orbitals correspond to each location in the BZ just by looking - you need to calculate the partial density of states to do that.

2.3.2 Tight-binding model

To make the maths easier, we go back to 1D and assume only nearest-neighbour (NN) interactions. We assume the atoms are spaced by a distance a:

$$\begin{aligned}\psi(k+x) &= \psi(x)e^{ikx}; \quad x = \pm na \\ &= \psi(x)e^{\pm ikna}\end{aligned} \tag{2.38}$$

The total wavefunction is a linear superposition of all wavefunctions:

$$\Phi(x) \propto \sum_M \psi(k+mka)e^{imka} \tag{2.39}$$

where m is every *allowed* n, i..e the distance of the interaction. Since we are assuming NN interactions, $m = \pm 1$ (the neighbour to the left and right). A typical Φ is then

$$\Phi(x) = e^{-ika}\phi(x-a) + e^{ika}\phi(x+a) + \phi(x) \tag{2.40}$$

and this state satisfies the time-independent Schrödinger equation

$$\hat{H}\phi = E\phi \tag{2.41}$$

To make writing the following maths easier, we use Dirac notation, so $\phi(x) \to |x\rangle$. Then

$$\hat{H}\,|x\rangle = E_0\,|x\rangle \implies \langle x|H|x\rangle = E_0 \tag{2.42}$$

Rewriting Eq. (2.40) gives

$$e^{-ika}\,\langle x|\hat{H}|x-a\rangle + e^{ika}\,\langle x|\hat{H}|x+a\rangle = E_k \tag{2.43}$$

$$E_k = E_0 + e^{-ika}(-t) + e^{ika}(t) \tag{2.44}$$

$$t = \langle x|\hat{H}|x+a\rangle\,, \tag{2.45}$$

where t is the **overlap integral**. We can combine the exponential terms into a trigonometric term, leading us to

$$E_k = E_0 - 2t\cos(ka) \tag{2.46}$$

Definition 2.3.1. The **band width** is the difference between the peaks and the troughs.

In the tight-binding case, the amplitudes are at $\pm 2t$, so the band width is $4t$. Additionally,

Definition 2.3.2. The **Brillouin zone** (BZ) is a unit cell in reciprocal (wavevector) space

For the tight-binding model in 1D, it is the interval $[-\pi/a, \pi/a]$.

Definition 2.3.3. The energy function $E = E(\vec{k})$ is the **band structure**. It tells you where states are distributed in the BZ.

Since t is the overlap, stringer NN interactions lead to a larger band width.

Near $k = 0$, we can perform a Taylor expansion

$$\begin{aligned}E(E) &= E_0 - 2t\left[1 - \frac{(ka)^2}{2} + \cdots\right] \\ E(k) &\simeq E_0 - 2t + ta^2k^2\end{aligned} \tag{2.47}$$

and we see we get **parabolic dispersion** around $k = 0$.

2.3.3 Tight-binding vs. free electron

In the free-electron model, the energy function is

$$E = \frac{\hbar^2 k^2}{2m_e} \qquad \frac{\partial^2 E}{\partial k^2} = \frac{\hbar^2}{m_e} \tag{2.48}$$

the electron mass is then

$$m_e = \hbar^2 = \left(\frac{\partial^2 E}{\partial k^2}\right)^{-1} \tag{2.49}$$

However, if we repeat the same calculation for the tight-binding model above:

$$E(L) = t_0 - 2E + ta^2K^2 \quad \frac{\partial E}{\partial K} = 2ta^2K \quad \frac{\partial^2 E}{\partial K^2} = 2ta^2 \tag{2.50}$$

Then our **effective mass** is

$$m^* = \frac{\hbar^2}{2ta^2} \tag{2.51}$$

In particular, the 'mass' of the particle in a 1D crystal depends strongly on the separation between the atoms, and behaves like a free electron except with a mass m^*. This is the general rule in a 1D, 2D or 3D crystal and are **quasiparticles**.

The overlap integral t encodes lots of hidden information

- If we have two s-orbitals nearby, the electrons are distributed roughly equally around the centre, meaning each atom has a cloud of electrons, which thus repel each other and $t < 0$
- If we have two p-orbitals interacting, well the electrons can move about between the lobes, thus at any moment there can be a net dipole moment across an atom. This means the atom has a slight locally positive and negative charge, which influences nearby atoms. Thus there will be minor attraction and $t > 0$.

The quasiparticle electron has a momentum

$$p \propto \frac{\partial E}{\partial k} \tag{2.52}$$

For our bandstructure, we have $p = 0$ at $k = 0, \pm\pi/a$ - the BZ boundary. This produces standing-wave-like solutions, so our NNTB model is very simple.

In reality, real bandstructure are complicated because there are many interactions. In Fig. 2.7,

2.3.4 Nearly-free electron model

This time, we *start* with with delocalised free electrons and *apply* a perturbation arising from the crystal potential $V(\vec{r}) = V(\vec{r} + \vec{R})$. To again make the maths easier, we look at it in 1D (you only need to replace everything with vector operations later)

$$\hat{H} = \hat{H}_0 + V(x) \tag{2.53}$$

$$\hat{H}_0 = \frac{\hat{p}^2}{2m} = \frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} \tag{2.54}$$

$$V(x) = \sum_j V_j e^{iq_j x} \tag{2.55}$$

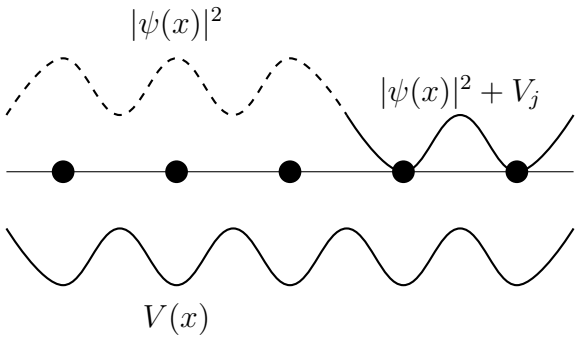


Figure 2.8: Example of a perturbed electron wave in a lattice.

We will skip over all the long perturbation theory calculation. It tells us that the final energy is simply the ground state energy plus or minus the perturbation energy, that is if $E(k) = \hbar^2 k^2/2m$ is the free electron energy, then $E(k)$ is perturbed by $\pm V_j$ when $k = q_j$.

When $k = q_j$, the electron wave has the same periodicity as the periodicity at one point in the lattice. This means that $|\psi|^2$ can be **in-phase** or **out-of-phase** with $V_j(x)$ as shown in Fig. 2.8. Whether the wave becomes in-phase of out-of-phase, that is $|\psi|^2$ is $\pm V_j$ in energy depends on whether V_j is repulsive or attractive.

When $k \neq q_j$, there is no extended phase relation to $V(x)$ so the electron does not 'see' the potential on average, leading to a free-electron-like dispersion.

Returning back to 3D, the free electron dispersion (Fermi surface) has spherical symmetry about $\vec{k} = 0$. At this constant energy isosurface, the energy is E_0 and

$$E_0 = \frac{\hbar^2 |\vec{k}|^2}{2m_e} \iff |\vec{k}| = \frac{1}{\hbar}\sqrt{2m_e E_0} \tag{2.56}$$

in all directions. However, as $\vec{k} \to \vec{q}_j$, the sphere becomes **distorted**.

2.3.5 Density Functional Theory (DFT)

DFT is based on the assumption that any property of a many electron system can be obtained as a functional of its ground state charge density $\rho = \rho(\vec{r})$. The existence of such a potential and its accuracy is derived by the Hohenburg-Kohn (HK) theorems; DFT now is performed by the Kohn-Sham scheme where a many-body system is broken down into a system of many time-independent, non-interacting Kohn-Sham orbitals ϕ_i, such that

$$\rho(\vec{r}) = \sum_{i=\text{occupied}} e|\phi_i(\vec{r})|^2 \qquad\qquad E_{\text{TOT}} = E[\rho(\vec{r})] \tag{2.57}$$

The total energy is a **unique** functional, of ρ (HK theorems). However, once you obtain this, you will be able to calculate any ground state propertyn you wish.

But, there is a big problem: you need an approximate way to treat electron-electron interactions, and these approximations can give big errors, especially in semiconductor bandgaps.

Additionally, DFT is very computationally expensive compared to tight-binding or nearly-free electron models. It is difficult to scale from a few atoms (like, say 10) to a nanostructures or bulk crystals with hundreds or even more.

- Indeed, this procedure can be reduced down to a lot of linear algebra.
- However, the matrices and vectors are often ridiculously large
- You also need to perform diagonalisation routines, matrix inversion etc., which even on very optimised implementations, still scaled poorly with the number of atoms
- Ignoring linear-scaling DFT, which is very nifty, most optimised DFT implementations have a time-complexity scaling as the number of atoms, *cubed*.

So for optical materials we prefer TB-type approaches.[2]

Another popular technique is $\vec{k} \cdot \vec{p}$. If we know $E(\vec{k}_0)$ at some point $\vec{k}_0$, we can calculate $E(\vec{k}_0 + \vec{k})$. This might seem like it wouldn't work, but the underlying mechanism is again perturbation theory, where by you treat the $+\vec{k}$ term as coming from some perturbing Hamiltonian $H_{\vec{k}}$.

- This can be quite accurate using semi-empirical parameters
- Simplest model is the '2-band' model featuring 1 CB and 1 VB
- More complex models use more bands
- Works for nanostructures

2.3.6 Group III-IV semiconductors

These are compounds formed from a combination of

- Group 13 elements (metals): Al, Ga, In
- Group 15 elements (non-metals and semi-metals): N, P, As, Sb

and make up a lot of optoelectronic (light-interacting electronics) devices. Examples are GaN, GaAs, InSb, InP, $Al_{0.7}Ga_{0.3}As$ etc. We can manipulate the materials and their *doping* to engineer bandgaps E_g of different sizes, for different uses:

- GaN: $E_g = 3.4$ eV - **wide-gap**; blue LEDS and lighting
- GaAs: $E_g = 1.4$ eV - **medium gap**; red or near infrared; solar panels
- InSb: $E_g = 0.17$ eV - **narrow gap**; mid-infrared; electron source

Now, consider the electronic structure of Ga and As:

$$\text{Ga}: \ [\text{Ar}]3d^{10}4s^2 4p^1 \qquad\qquad \text{As}: \ [\text{Ar}]3d^{10}4s^2 4p^3 \tag{2.58}$$

Now if you combine them together, all the valence electrons are in the $n = 4$ energy level, that is the all the $4s, 4p$ electrons, giving 8 valence electrons - the orbitals combine to effectively give 4 identical hybrid orbitals. This process is $\mathbf{sp^3}$ hybridisation. GaAs has a cubic unit cell, but the atoms within are **tetrahedrally-bonded** as in Fig. 2.9.

These are 4 covalent bonds with a minor amount of ionic character. Next, is how the electrons organise themselves. Since the orbitals merge, this means the electrons must rearrange into a new electronic configuration. Unfortunately, bonding is not perfect and 2 types of orbitals form from hybridisation

[2] As in Ref. [7], you may apply excited state calculations on top of DFT ground states, like GW, to correct the bandgaps. This is still an approximation, but it gets us a lot of the way there,

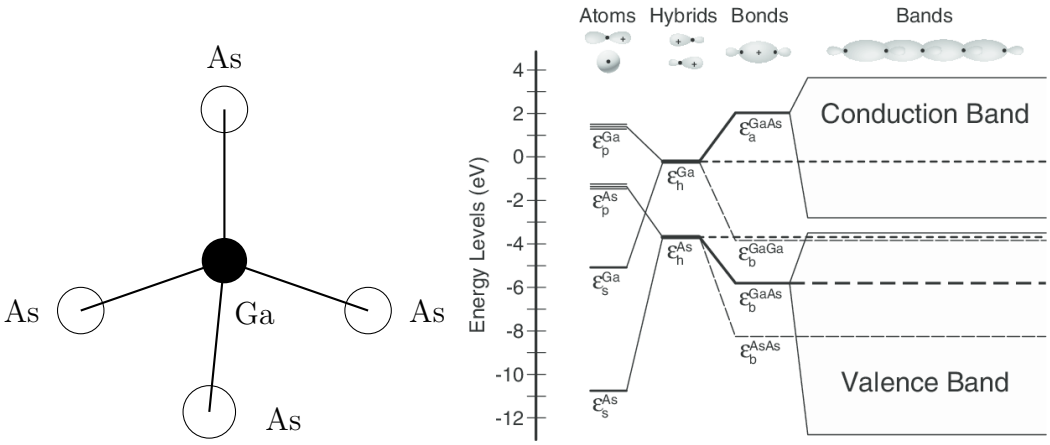


Figure 2.9: (left) Bonding for GaAs. (right) Energy-level diagram for sp^3-hybridised GaAs reproduced from Ref. [5].

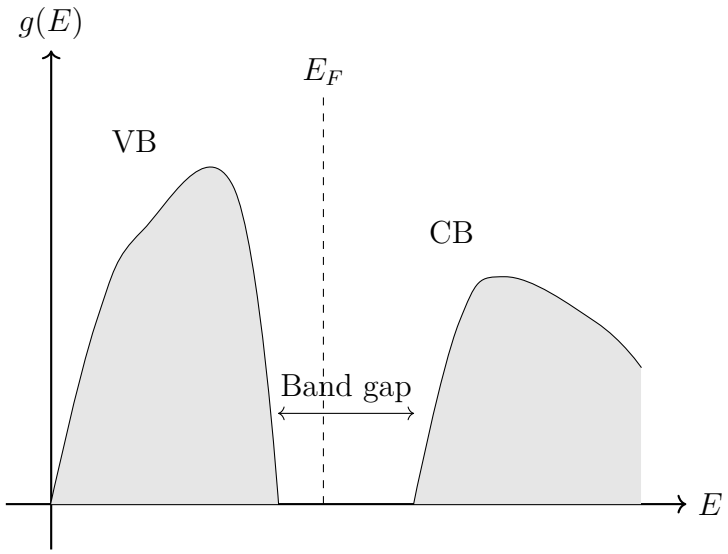


Figure 2.10: Density of states for GaAs.

Definition 2.3.4. An **antibonding** orbital is a molecular orbital which *weakens* the molecular bond and increases the energy (decreasing stability) of the molecule.

As you can imagine then, the polar opposite is
Definition 2.3.5. A **bonding** orbital is a molecular orbital which *increases the strength* the molecular bond.

Now, you cannot have identical quantum states by Pauli exclusion, we cannot have 2 identical $4s^2$ orbitals, so the $4s^2$ orbital actually *splits* into a lower component that remains in the VB and an upper component in the CB in the sp^3-hybridised molecule. Hence, the other $4s^2$ from the other atom are *promoted* into p-orbitals since there are empty states there. Again though, the same logic applies. The $4p$ state again splits such that the now 6 electrons are lower in energy (but still above the new $4s^2$) and an additional new unoccupied band in the CB. This is captured in Fig. 2.9(right). Hence, the antibonding orbitals lie in the CB and bonding ones in the VB.

Experimentally, this can be seen in the **density of states** (DOS) which will show a gap or zero between the CB and VB as in Fig. 2.10.

The band structure near Γ is shown in Fig. 2.11 for GaAs and InSb.

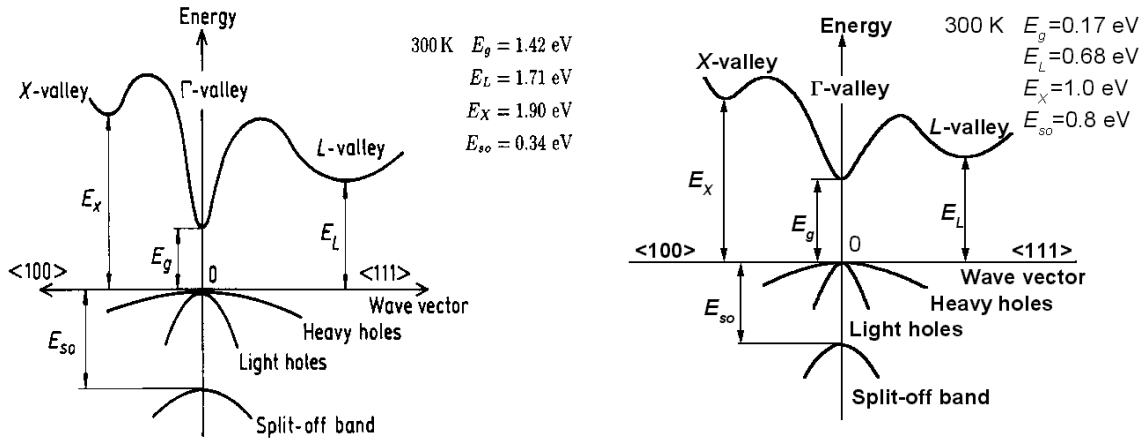


Figure 2.11: Bandstructure of (left) GaAs and (right) InSb, around the high-symmetry point Γ. Images reproduced without modification from this page for GaAs and this page for InSb.

- Light holes have low effective masses. If you imagine them as a lightweight ball, when you throw them, their trajectory is sharper. In this case, the light hole is a lightweight ball through the BZ.
- Similar for heavy holes - they have heavy effective mass so traverse the BZ in a broader fashion
- the split-off band is due to spin-orbit coupling
- The CB in GaAs is very parabolic but in InSb, it is not very parabolic.

For InSb, the VB is well described by a 2-band $\vec{k}\cdot\vec{p}$ 'Kane' model. We will compare the validity of a parabolic isotropic band vs Kane model for describing the dispersion around Γ.

We expand $E(\vec{k})$ in a power series

$$E(k) = a_1k^2 + a_2k^4 + a_3k^6 + \dots \tag{2.59}$$

Remark. We assume even powers only because around Γ, the dispersion looks like an even function.

For a parabolic isotropic band, $a_1\hbar^2/2m^*$ and $a_2 = a_3 = \dots = 0$. This is expected from the name alone. In the Kane model,

$$\frac{\hbar^2k^2}{2m^*} = E(1+\alpha E), \tag{2.60}$$

where $\alpha = E_g^{-1}$ for a 2-band $\vec{k}\cdot\vec{p}$ model. Then the dispersion is

$$E(k) = \frac{-E_g}{2} + \left[\left(\frac{E_g}{2}\right)^2 + E_g\frac{\hbar^2k^2}{2m_e^*}\right]^{1/2} \tag{2.61}$$

The Kane model is valid when the energy is less than about $4E_g$. Without derivation, we state the result for the density of states $g(E)$

$$g(E) = \frac{2}{(2\pi)^D}\int_S \frac{ds}{|\nabla E(\vec{k})|} \tag{2.62}$$

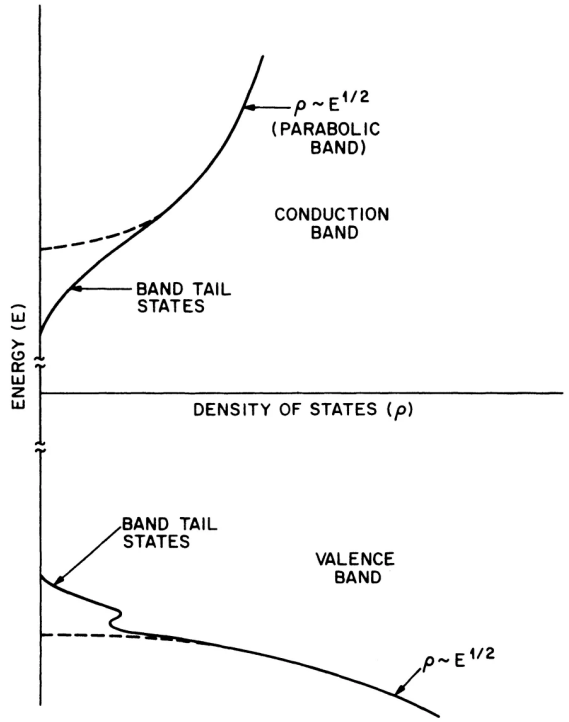


Figure 2.12: DOS for the parabolic (dashed) and Kane models (solid). The solid lines will continue down to zero. Reproduced from this site without modification.

where D is the number of spatial dimensions and S is a constant energy surface in $\vec{k}$ space. This expression is evaluated for the isotropic (Eq. (2.63)), isotropic parabolic (Eq. (2.64)) and Kane models (Eq. (2.65)):

$$g_{\text{iso}}(k) = \frac{k^2}{\pi^2}\left(\frac{\partial E}{\partial K}\right)^{-1} \tag{2.63}$$

$$g_{\text{para}}(E) = \frac{1}{2\pi^2}\left(\frac{2m_e^*}{\hbar^2}\right)^{3/2} E^{1/2} \tag{2.64}$$

$$g_{\text{kane}}(E) = \left(\frac{2E+E_g}{E_g}\right) g_{\text{para}}\left[\left(\frac{E+E_g}{E_g}\right)^{1/2}\right] \tag{2.65}$$

A plot of the parabolic and Kane models is shown in Fig. 2.12

2.3.7 Interband absorption

In this section, we will study the mechanism by which electrons in the VB get excited into the CB. We suppose the electron starts in an initial state $|i\rangle$, absorbs an incident photon of energy $\hbar\omega$, ending up at a final state $|f\rangle$.

To study the transition rate, we use **Fermi's Golden Rule**

$$W_{i\to f} = \frac{2\pi}{\hbar}\left|M_{if}\right|^2 \delta\left(E_f - E_i - \hbar\omega\right), \tag{2.66}$$

where

- $W_{i\to f}$ is the **transition rate**
- $M_{if} = \langle f|\hat{t}|i\rangle$ is the transition matrix element and $\hat{t}$ is the **electric dipole transition operator**, which you don't need to know in detail.

- $\delta\left(E_f - E_i - \hbar\omega\right)$ enforces the correct selection rule
- If you really wanted to, you could add the Heaviside function to explicitly enforce energy conservation but this is not necessary

In this section, we ignore interference effects, namely $\lambda_{\text{light}} \gg$ unit cell dimensions. Then the electromagnetic vector potential can be approximated to first order (linear):

$$\vec{A} = A_0 \sum_{j=0} (\vec{q}\cdot\vec{r})^j \to \vec{A} \sim A_0\left(1 + \vec{q}\cdot\vec{r}\right)$$

where $\vec{q}$ is the photon wavevector. Since the photon also carries an electric field $\vec{E}$, then the electric dipole transition operator $\hat{t}$ is

$$\hat{t} = \vec{p}\cdot\vec{E}, \tag{2.67}$$

where $\vec{p}$ is the **electric dipole moment** defined as $\vec{p} = -e\vec{r}$. Generally, $\vec{E} = E_0\vec{e}f(\vec{q}\cdot\vec{r})$ (some sort of wave), so $\hat{t} \propto E_0\vec{r}$. We substitute this into M_{if}

$$M_{it} \propto \int_{\text{medium}} \psi_f^*(\vec{r})\vec{r}\psi_i(\vec{r})d^3\vec{r} \tag{2.68}$$

We are integrating over the entire medium, so the region of integration is symmetric. Now, $\vec{r}$ is an odd function, so we require that ψ_f and ψ_i must have **opposite parity** for non-zero M_{if}.

For example, in GaAs, VB states are derived from p-states whereas in the CB, it is from s-states. These states have opposite parity, so electronic transitions between these states are *allowed*. there are further transition rules. Here, let l denote the orbital angular momentum of the electron, and L, S, J the total orbital angular momentum, the total spin and total angular momentum respectively. Then in addition to requiring opposite parity:

- $\Delta l = \pm 1$
- $\Delta L = 0, \pm 1$ but $L = 0 \to 0$ is forbidden
- $\Delta J = 0, \pm 1$ but $J = 0 \to 0$ is forbidden
- $\Delta S = 0$

When electric-dipole transitions are forbidden, other types of processes may be possible. For example, magnetic-dipole and electric-quadrupole transitions are possible between states of the same parity. This comes from including higher-order terms in our expansion for $\vec{A}$.

Using Bloch's theorem, in a crystalline solid

$$\psi_i \propto u_i(\vec{r})e^{i\vec{k}_i\cdot\vec{r}} \qquad\qquad \psi_f \propto u_f(\vec{r})e^{i\vec{k}_f\cdot\vec{r}} \tag{2.69}$$

We substitute this into Eq. (2.68) and get

$$M_{if} \propto \int_{\text{medium}} e^{-i\vec{k}_f\cdot\vec{r}}u_r^*\vec{r}e^{i\vec{k}_i\cdot\vec{r}}u_i e^{i\vec{q}\cdot\vec{r}}d^3\vec{r} \tag{2.70}$$

$$M_{if} \propto \int_{\text{rens}} \left(u_f{}^* u_i\right)\vec{r}\exp\left(-i\left[\vec{k}_f - \left(\vec{k}_i + \vec{q}\right)\right]\cdot\vec{r}\right)d^3\vec{r} \tag{2.71}$$

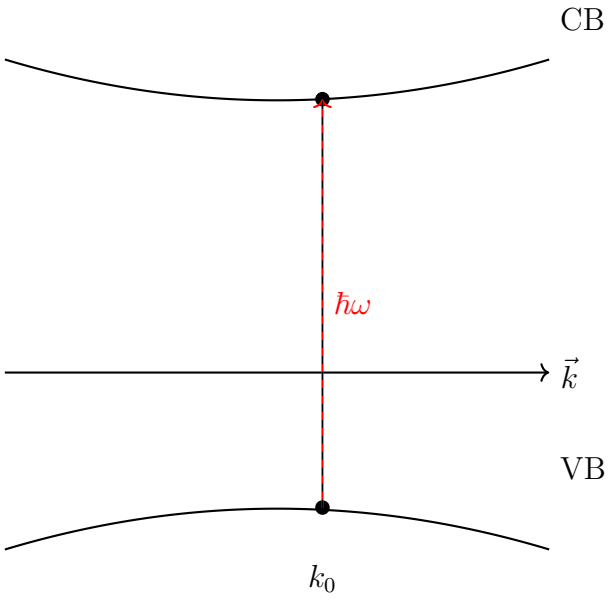


Figure 2.13: Zoomed in bandstructure showing a vertical dipole transition (red).

We have an exponential factor that is causing a phase shift. Since the complex exponential is just some trig functions, in order to avoid $M_{if} = 0$, we want $\vec{k}_f = \vec{k}_i + \vec{q}$, so that $\hbar\vec{k}_f = \hbar\vec{k}_i + \hbar\vec{q}$ - this is nothing more than **conservation of momentum** (or energy, if you prefer).

For dipole transitions where $\lambda \gg a$, then $q \ll q_j$ (photon wavevector is smaller than the wavevector of the modes). The energy is smaller too, and so dipole transitions are **purely vertical in the BZ**. An example is shown in Fig. 2.13.

To get the *total* absorption α, we need to sum over all possible absorbing transitions given a photon frequency of $\hbar\omega$. Naturally then, we need a function that can produce the states available at a particular energy and accounts for the number of available states to excite into, and the number of valence electrons - this is the **joint density of states** (JDOS) $g(E)$

$$W_{\text{TOT}} = \frac{2\pi}{\hbar}|M|^2 g(\hbar\omega) \tag{2.72}$$

such that $g(E)dE$ is the number of available states in the interval $[E, E + dE]$.

Consider a single VB to CB vertical transition for an isotropic, parabolic band. This is typically the optical excitation corresponding to from the VBM (valence band maximum) to the conduction band minimum (CBm) at Γ. Then the electron in the VB has to reach different energies:

- Must have enough energy to reach the edge of the valence band (binding energy) based on where it is in the BZ. Implies we need effective electron mass m_e^*.
- Go across the bandgap E_g
- electron leaves behind a bound state (negative energy) which is an effective hole
- Therefore $\hbar\omega = E_e - E_h$

Being more explicit now

$$E_e = E_g + \frac{\hbar^2 k^2}{2m_e^*} \qquad\qquad E_h = -\frac{\hbar^2 k^2}{2m_h^*} \tag{2.73}$$

Then the photon energy satisfies

$$\hbar\omega = E_e - E_h = E_g + \frac{\hbar^2 k^2}{2m_e^*} + \frac{\hbar^2 k^2}{2m_h^*} = E_g + \frac{\hbar^2 k^2}{2\mu}, \tag{2.74}$$

where $\mu^{-1} = m_e^{*-1} + m_h^{*-1}$ is the **reduced effective mass**.

For an isotropic, 3D material, the density of states $g(k)$ satisfies

$$g(E) = \frac{2g(k)}{(\partial E/\partial k)} \quad g(k)dk = \frac{1}{(2\pi)^3} 4\pi k^2 dk \tag{2.75}$$

The derivative $\partial E/\partial k$ is obtained from Eq. (2.74):

$$\frac{\partial E}{\partial k} = \frac{\hbar^2 k}{\mu}$$

Substitute this back into the expression for $g(E)$, ensuring you then substitute Eq. (2.74) and you get Then we have 2 cases for $g(E)$

$$g(E) = \begin{cases} \frac{1}{2\pi^2}\left(\frac{2\mu}{\hbar^2}\right)^{3/2}\sqrt{E - E_g} & \hbar\omega > E_g \\ 0 & \hbar\omega < E_g \end{cases} \tag{2.76}$$

If you plot a graph of α^2 against $\hbar\omega$, you will get a straight line graph which is then 0 for $\hbar\omega < E_g$.

2.3.8 Indirect gap semiconductors

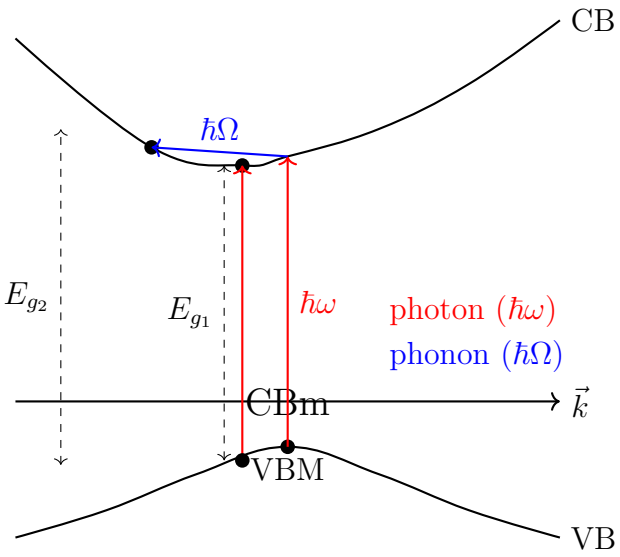


Figure 2.14: Depiction of typical indirect gap semiconductor.

Many Group 3-5 semiconductors are direct gap - meaning the global CBm lies vertically above the global VBM. There exist **indirect-bandgap** semiconductors where the VBM and CBm are not aligned, as in Fig.2.14. There are 2 bandgaps $E_{g_1} < E_{g_2}$, where E_{g_1} is from the true CBm to somewhere on the VBM, whereas E_{g_2} is from, the VBM to a local CBm.

This means in general, to go from VBM to CBm, we need a change of $\vec{k}$, and the photon's momentum $\vec{q}$ is too small. Thus,

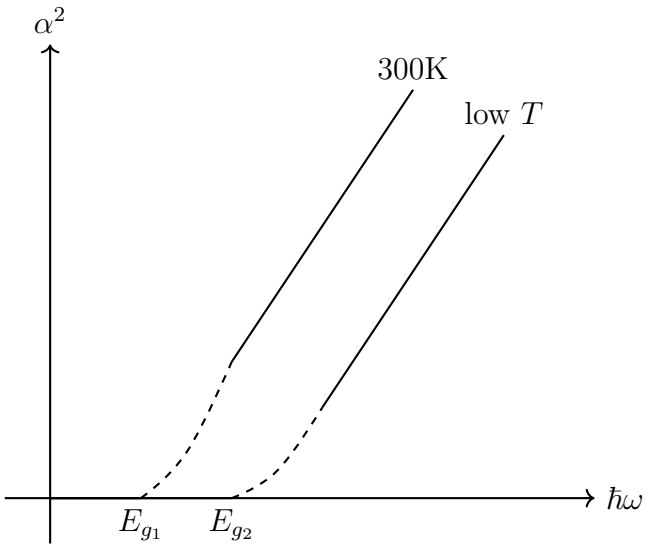


Figure 2.15: Plot of α^2 against incident photon energy $\hbar\omega$ for an indirect gap semiconductor at 2 temperature regimes.

Phonons are involved in transitions across the BZ to conserve momentum.

Hence $E_{g_1} = \hbar\omega \pm \hbar\Omega$ where the $+/-$ corresponds to absorption or emission of a phonon respectively. Since $\hbar\Omega \sim 10$ meV, so they are usually **thermally generated**. As such, at lower temperatures, they are a smaller factor in total absorption, as quantified in Fig. 2.15.

2.3.9 Absorption over indirect gaps

In order to find the transition rate for a "phonon + photon" interaction, you need to sell your soul to mathematics. Roughly speaking, the absorption coefficient

$$\alpha_{\text{ind}}(\hbar\omega) \propto (\hbar\omega - E_g \pm \hbar\Omega)^2 \tag{2.77}$$

$$\alpha_{\text{dir}}(\hbar\omega) \propto (\hbar\omega - E_g)^{1/2} \tag{2.78}$$

i.e., it is a completely different power law dependence between indirect and direct gap semiconductor, and this is useful to distinguish between them.

- Indirect transitions tend to give *much* weaker absorption than direct gap. Germanium is an example of an indirect gap semiconductor.
- Indirect transitions have a strong temperature dependence. This is because phonons exist by thermal activation. A typical $\hbar\Omega \sim 5 - 50$ meV, but $k_B T \sim 50$ meV at room temperature, so they are always present. However, at low T, phonons are frozen out (atoms vibrate less) so really only get phonons during the transition.

2.3.10 Other direct transitions

Other points in the BZ may have available direct transitions, not just the direct gap at the VBM or whatever. For example, in Fig. 2.7, I could have transitions between the red lines at L, or along the $\Gamma - X$ path on the left figure. These transitions would correspond to extra peaks in $\alpha(\hbar\omega)$. These peaks will be emphasised by bands which are locally parallel to each other, over some range of $\vec{k}$.

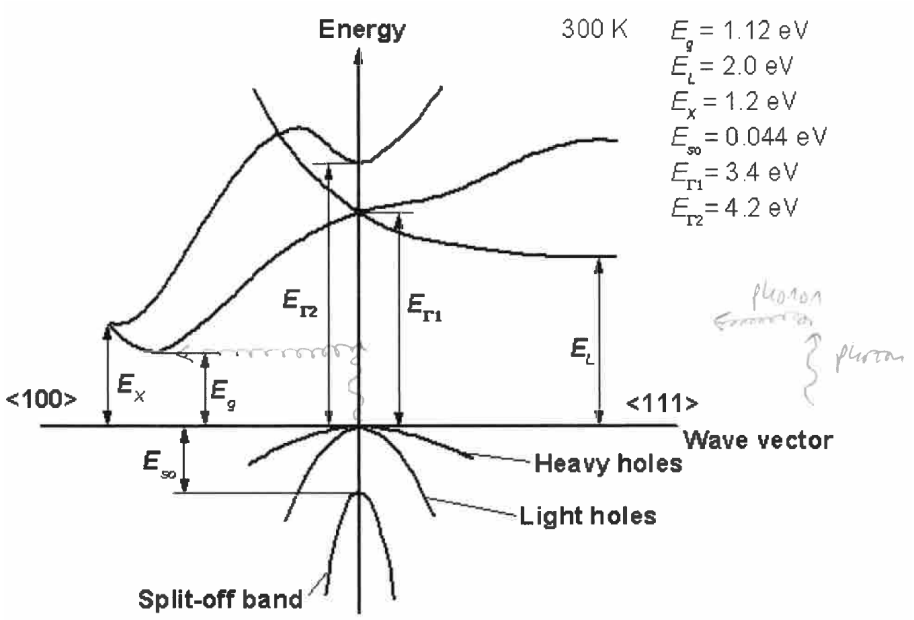


Figure 2.16: Bandstructure of silicon, which is an indirect gap semiconductor, showing an example of phonon scattering (pencil lines). Screenshot taken from Prof. Gavin Bell's lectures, and original diagram from this page.

Flat bands also enhance the DOS, leading to increased probability of absorption by Fermi's Golden Rule. Since there is a low curvature, they have a high effective mass (so E_e is potentially smaller) and still come with a high $g(E)$, leading to a **sharper** peak.

2.4 Excitons

When a transition occurs, the electron leaves behind a **hole**. Here, there is a chance to form an **exciton** and there are 2 types.

In very flat bands, the electron or hole can get 'stuck', i.e. they don't move around the BZ as much - these are high m^* particles. In contrast, high curvature bands have a low m^* and are very mobile. In both cases, there is the chance for the electron and hole to form a **bound state** called an **exciton**.

- A **bound/Frenkel exciton** has a binding energy usually around 100 meV and forms a radius of about the lattice parameter a.
- A **free/Wannier-Mott exciton** has a binding energy around 10 meV and its radius $R \gg a$.

In general, for an exciton to be stable, we need the **exciton binding energy** $\boxed{E_x > k_B T}$. Free excitons are typically seen at low $T < 77$ K.
Remark. You can still see bound excitons at low temperatures, see for example, this paper by Wei *et al*. Additionally, if you do any optical simulations at 0 K, you can still see excitons with a $E_x \sim 70 - 80$ meV, see Ref. [7].

Now, an exciton is a single negative and positive charged particle orbiting around each other - this is like the hydrogen atom, so we can model the exciton as such:

$$E = -\frac{R_x}{n^2}, \tag{2.79}$$

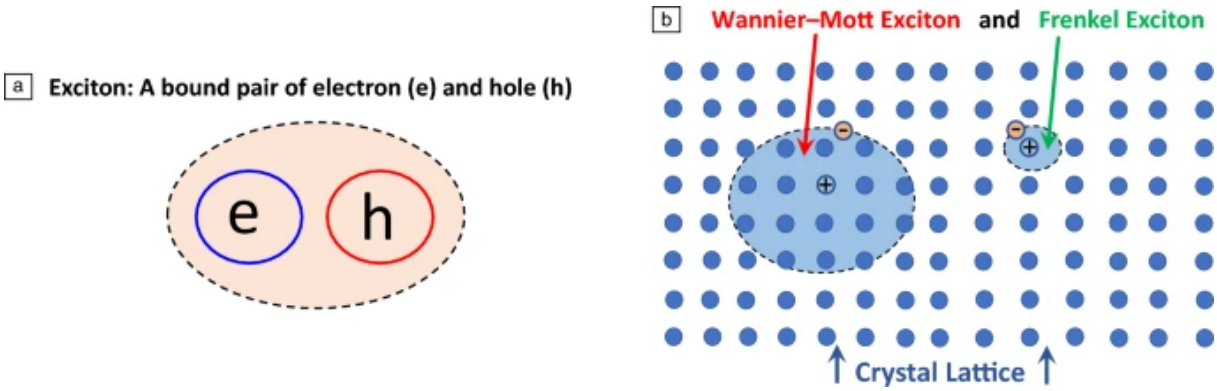


Figure 2.17: (a) A schematic of an exciton as a quasiparticle of negatively charged electron (e) and positively charged hole (h) bound together through their mutual Coulomb attraction. (b) Schematic of the spatial extents of WannierMott excitons and Frenkel (bound) excitons within a crystal lattice. Reproduced from Ref. [1] without modification.

where R_x is the effective Rydberg energy. Much like how $R_H = -13.6$ eV sets the scale for the highest energy photon that can be emitted, R_x does the same but for a exciton. In this model, we have to take account two more ideas

- e^-, h^+ are oppositely charged and so form an internal electric field that will anti-align with an applied external one $\rightarrow \epsilon(0)$, non-vacuum dielectric constant
- Assume e^-, h^+ have their normal effective masses m_e^*, m_h^*

Hence, $R_H \rightarrow R_x$ by setting $m \rightarrow \mu, \epsilon_0 \rightarrow \epsilon(0)$:

$$R_H = \frac{me^4}{8h^2\varepsilon_0^2} \rightarrow R_x = \frac{\mu e^4}{8h^2\varepsilon(0)^2} \tag{2.80}$$

Typically, $\epsilon(0)$ is around 10 or more, and $m_e > \mu \sim 0.1m_e$ giving us $R_x \sim 10$ meV as anticipated. We also scale the Bohr radius to the **exciton radius** R

$$R = \frac{a_B\epsilon(0)}{\mu} \tag{2.81}$$

where a_B is a constant around 0.05 nm, giving $R \sim 5$ nm. Plotting α against $\hbar\omega$ for GaAs is shown in Fig. 2.18. We see that there is an initial exciton peak at low temperatures that gets sharper at lower temperatures. It is hard to see, but after the peak, the square root behaviour follows from Eq. (2.78). At high temperature, we see that there is basically no peak at all. This is characteristic of a **bound exciton** causing enhanced absorption at the band edge.

2.4.1 Excitons in external fields

If an applied electric field E_{ext} is stronger than the exciton's internal electric field E_{int}, the electron splits off from the hole.

The internal electric field is roughlty $E_{\text{int}} \sim 2R_x/eR$, which is *smaller* than the typical fields in most devices such as p-n junctions.

Inside a p-n junction, $E_{\text{ext}} = \Delta V/L > E_{\text{int}}$, so this easily breaks up the electron-hole pair and in most devices, you can treat the electron and hole separately. An example of this is solar cells.

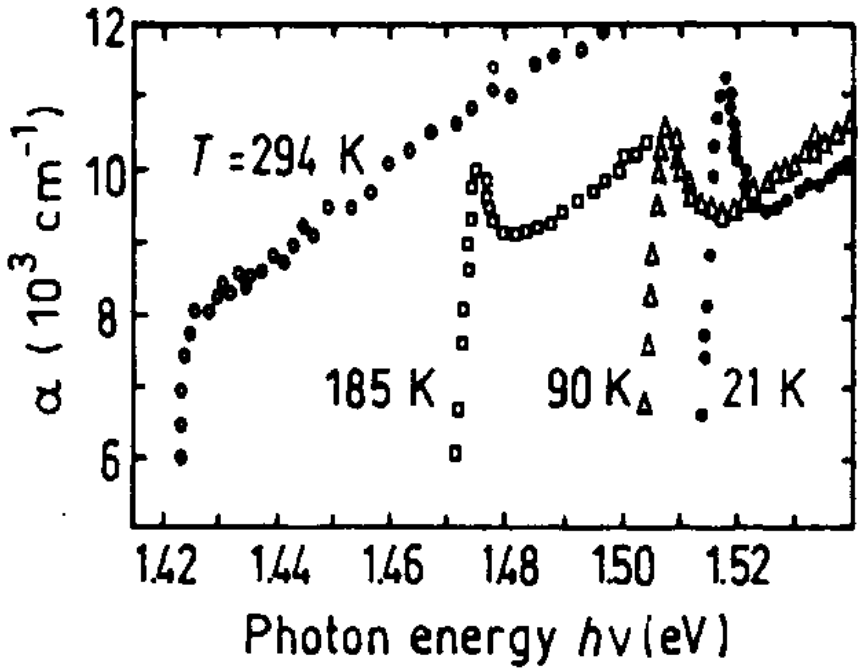


Figure 2.18: Optical absorption of GaAs at different temperatures. Reproduced from Ref. [6].

2.5 Luminescence and LEDs

We have been studying the case where photons are absorbed. Let's study the opposite, where they are released. In this case, an excited state electron drops down to the ground state and emits a photon. This is **luminescence**.

There are 3 methods to produce excited states (and thus 3 ways to do luminescence). These are

- Photoluminescence (PL) - emission after photon absorption
- Cathodoluminescence (CL) - emission after electron collisions
- Electroluminescence (EL) - emission due to a current or electric field. This is what is used in LEDs and laser diodes.

We first compare direct gap and indirect gap materials for use in luminescence.

Direct gap	**Indirect gap**
Efficient absorption for $\hbar\omega > E_g$	Phonon momentum $\hbar\Delta k$ needed for radiative electron-hole recombination - **inefficient**
Efficient emission across band extrema	τ_{P} long, luminescent efficiency small so strong combination with non-radiative recombination τ_{NR}
Short radiative lifetime $\tau_{\mathrm{R}} \sim 10^{-8}, 10^{-9}$ seconds	$$\text{Efficiency} = \frac{1}{1+\frac{\tau_{\mathrm{R}}}{\tau_{\mathrm{NR}}}} \quad (2.82)$$ We do NOT use indirect gap semiconductors for luminescence.

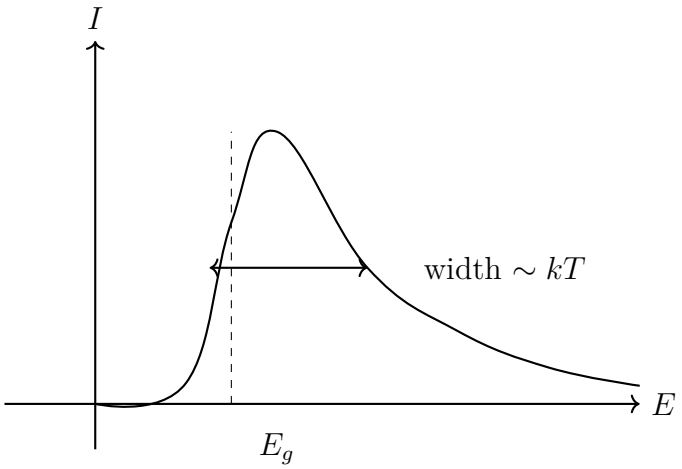


Figure 2.19: The luminescent intensity for a non-degenerate direct gap semiconductor with bandgap E_g.

Definition 2.5.1. The **luminescence intensity** $I(\hbar\omega) = W_{i\to f} f_c f_v$, where

- f_c is the probability of electron occupying the CB
- f_v is the probability of hole occupying the VB
- Both f_c, f_v are Fermi-Dirac (occupation) functions
- $W_{i\to f} = |M_{if}|^2 g(\hbar\omega)$ is the transition rate, and is the same for absorption in that it is $\propto \sqrt{\hbar\omega - E_g}$ for direct gap materials.

The Fermi-Dirac function is

$$f(E) = \frac{1}{e^{(E-\mu)/k_B T} + 1} \tag{2.83}$$

where μ here is the chemical potential and we will be ignoring this for the rest of the section. We can use a Taylor series expansion to first order (i.e, the Boltzmann factors) so that

$$f_c = e^{-E_e/k_B T} \qquad\qquad f_c = e^{-E_h/k_B T}$$

Then the intensity is

$$I(\hbar\omega) \propto \sqrt{\hbar\omega = E_g}\exp\left[-\left(E_e + E_h\right)/k_B T\right] \tag{2.84}$$

But by Eq. (2.74), the term $E_e + E_h = \hbar\omega - E_g$ so substituting into Eq. (2.84) we get

$$I\left(\hbar\omega\right) \propto \left(\hbar\omega - E_g\right)^{1/2} e^{-(\hbar\omega - E_g)/k_B T} \tag{2.85}$$

Figure 2.19 shows a plot of this. We see that E_g is just a bit to the left of the emission peak. You can find the location of this peak by differentiation. Indeed, we get

$$\begin{aligned}
\frac{dI}{d(\hbar\omega)} &= \frac{1}{2}\left(\hbar\omega - E_g\right)^{-1/2} e^{-(\hbar\omega - E_g)/k_B T} + \frac{-1}{k_B T}\left(\hbar\omega - E_g\right)^{1/2} e^{-(\hbar\omega - E_g)/k_B T} = 0 \\
&\implies \frac{1}{2} - \frac{\hbar\omega - E_g}{k_B T} = 0 \\
&\implies \hbar\omega = E_g + \frac{k_B T}{2}
\end{aligned}$$

as the location of the intensity maximum.

In degenerate direct-gap semiconductors, there are two situations

- For n-type, many electrons fill states in CB
- For p-type, many holes fill states in VB

For example, for n-type GaAs, set $E_F = 0$. This may push part of the conduction band into the negatives. Then the energy emitted $\Delta E = E_F - E_{\text{CBm}}$ where E_{CBm} is the energy of the conduction band minimum.

2.5.1 CL-imaging and spectroscopy

See Fig. 2.20(left).

- Generate electron-hole pairs with nm-sized beams in scanning electron microscopy.
- Beam reaches several keV
- Useful for looking at (semiconductor) devices

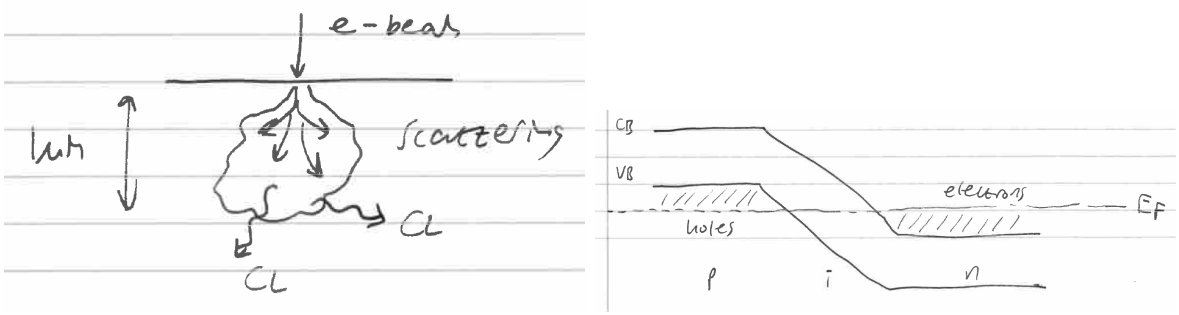


Figure 2.20: (left) Diagram of cathodoluminiscence. (right) p-n junction for EL. The shaded region denotes filled states in the VB and CB. Both diagrams are screenshots of Prof. Gavin Bell's lectures.

2.5.2 EL-devices

We take a good-ol p-n junction and apply a bias voltage V. This injects electrons (and by extension, holes) by a bias voltage, which is how an LED is made. The setup is specifically p-i-n (positive-interface-negative) and is shown in Fig. 2.20(right).

- Apply V: the electrons and holes move into the depletion region 'i' which narrows.
- If the material of 'i' is a direct gap material $\rightarrow$ strong emission and narrow range of photon energies around E_g emitted.
- If 'i' has many defects, τ_{NR} is short and luminescence is weak (using table from before).
- Therefore, materials in an LED must have **similar lattice constants**, otherwise you get dislocations (bonds become unstable) and LED doesn't work.
- Canonical example: GaAs - (Al,Ga)As-AlAs. All 3 materials have very close lattice constants (by coincidence really). AlAs has a higher E_g. Embedding GaAs between p-(Al,Ga)As and n-(Al,Ga)As gives an efficient red LED. You would imagine then, using GaN with wide bandgaps gives blue LEDs.

If we are careful, we can engineer bandgaps by adjusting the doping of Al. The compound is $\text{Al}_x\text{Ga}_{1-x}\text{As}$ over 1.4 - 1.9 eV for $x < 0.4$. GaAs is a near-infrared emitter, and if $x = 0.4$, it is a red LED.

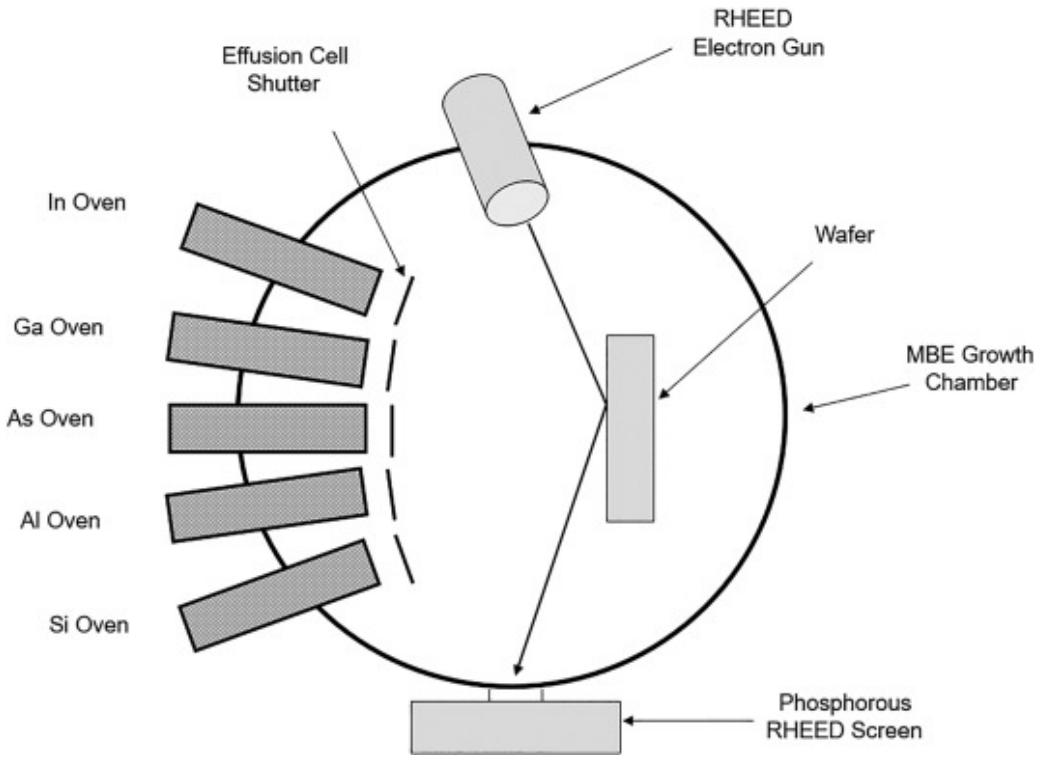


Figure 2.21: Schematic setup of MBE.

2.6 Molecular Beam Epitaxy (MBE)

We turn our attention to the growth of nanomaterials. Molecular Beam Epitaxy (MBE) is a modern, high-quality technique to grow **low defect density, high purity** material layers, and allows for **doping control** and **variable layer thicknesses** from the micron to single-atomic scale. MBE is extensively used for Group 3-5 semiconductors, though other methods are available for other compounds.
Definition 2.6.1. Heteroepitaxy is the growth of different materials with the same crystal structure.

For example, GaAs and AlAs can be layered together. The interface they share (the As) is the **heterojunction**. Both GaAs has a zincblende structure and AlAs has an FCC structure, more specifically they both are in the F-43m space group.

In order for MBE to work, we need the lattice constants of the 2 materials to be as close to each other as possible. Even a difference of 7% like InAs/GaAs is enough to cause **crystallographic defects**.

These defects as previously mentioned, leads to strong non-radiative recombination, so $\tau_{\mathrm{NR}} \ll \tau_{\mathrm{R}}$, which is optically very bad.

2.6.1 Setup of MBE

For the exam, you will need to know how MBE works and why each component is necessary. A schematic is shown in Fig. 2.21.

- **effusion cells**: elements are heated up into a vapour. The shutters are open to let out materials, and allows thickness adjustments. They are all aimed at the wafer (sample plate).

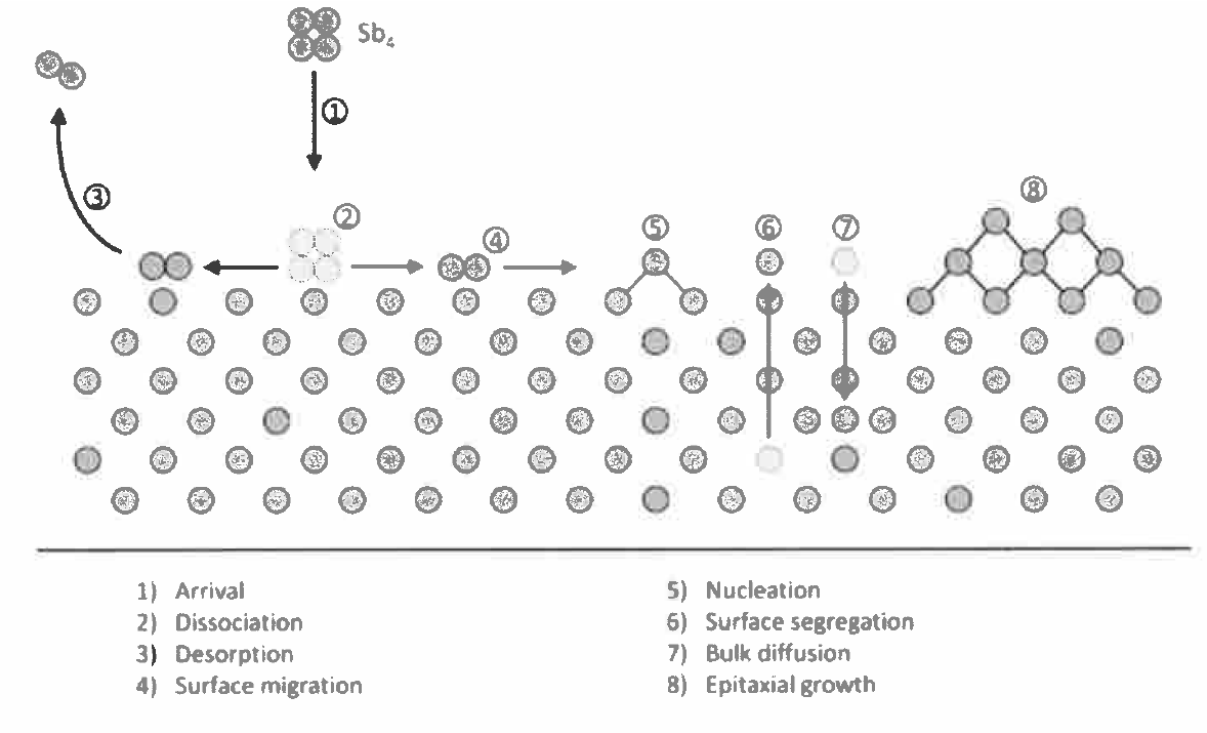


Figure 2.22: Close-up of what can happen at the surface of the sample during MBE.

- MBE requires **ultra-high vacuum** and **sterile environments** to avoid contamination
- The wafer has a **heater filament** to keep the sample at the correct temperature to allow for deposition
- There is also a **manipulator** (not labelled) on the wafer to rotate and move the sample around
- A RHEED (Reflection high-energy electron diffraction) allows you to do electron diffraction of the sample surface to see what the sample looks like.

At the material level, multiple things can happen, which is shown in Fig. 2.22. Note that naturally you will not have isolated atoms even in vapour form - *molecules* or small groups of atoms will generally make their way to the sample, and under the right conditions, the thing you want will happen (hopefully).

1. Arrival - self-explanatory, the structure arrives near the surface
2. Dissociation - atoms hit the surface and may *dissociate*. in Fig. 2.22, Sb_4 is incident and dissociates into 2 pairs of Sb.
3. Desorption - excess material may leave the surface and be collected in waste.
4. Surface migration - due to interacting edge potentials and momentum conservation, incident atoms may end up moving along the sample surface
5. Nucleation: bonds formed with the surface
6. Surface segregation: material breaks bonds - this could be because of defects or lattice mismatches, and what you have is just an atom sitting there
7. Bulk diffusion - incident atoms may move *beyond the surface* into the sample.

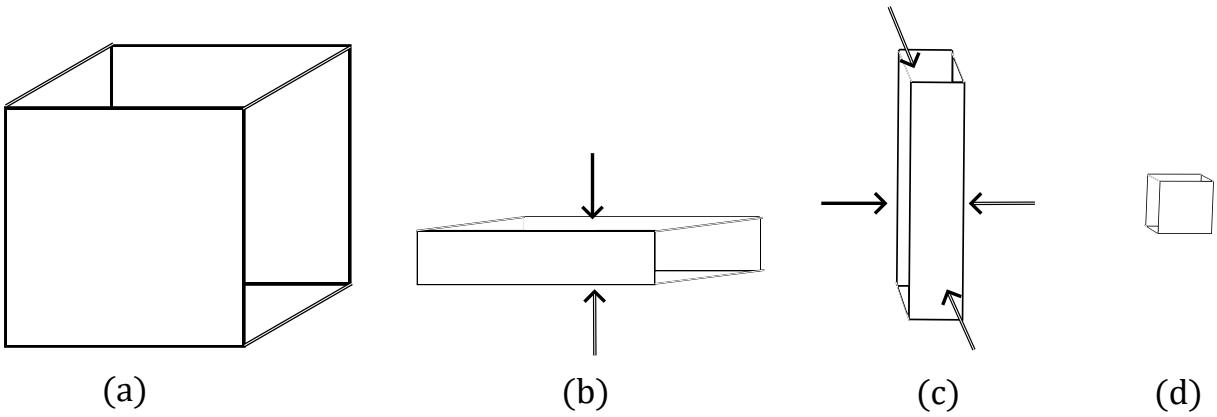


Figure 2.23: Schematic of typical nanomaterials in different dimensions. (a) 3D bulk crystal, (b) 2D quantum well (confinement in 1 axis), (c) 1D quantum wire (confinement in 2 axes), and (d) quantum dot (QD), confinement in all axes. Each pair of black arrows that are opposite each other represent confinement in that direction.

8. Epitaxial growth - the end goal - atoms form bonds and stably sit on top of layers on the sample surface.

2.6.2 Nanomaterials

Quantum confinement affects the wavefunction and energy eigenvalues of electrons and holes in semiconductors. This has lead to physicists across the world abusing this to make cool, small quantum devices spanning different spatial dimensions. The typical categories are shown in Fig.2.23. 2D and 1D materials are real. Graphene is a 2D quantum well (it has carbon chains confined to a plane), and carbon nanotubes are wires (the chain axis is along one direction). It is important to note that the dimension here only refers to the unconstrained dimensions - not their actual dimension in 3D space as we see it.

2.6.3 Particle in a box

In any confinement direction, the wavefunction behaves like a particle in a box with discrete eigenenergies E_n and stationary wave-like solutions, so that

$$\psi_n \propto \cos(k_n z) = \cos \frac{n\pi z}{l} \tag{2.86}$$

assuming confinement in the z direction (so the material is in the $x-y$ plane). The energies are then

$$E_n = \frac{\hbar k_n^2}{2m^*} = \frac{n^2\pi^2 k_n^2}{2m^* L^2} \tag{2.87}$$

where L is the width of the quantum well in that axis.

Now, we can do the same for the free (unconfined) axes by separation of variables. In this case, the energy in the unconfined axes is $\hbar^2 k_i^2/2m^*$ where i is the unconfined directions, whilst it is of the form Eq. (2.87) for the confined direction(s). Defining the constant $\hbar^2/2m^* := K$

$$E(k_x, k_y, k_z) = \begin{cases} K(k_x^2 + k_y^2 + k_z^2) & \text{3D} \\ K(k_x^2 + k_y^2) + E_{z,n} & \text{2D} \\ K(k_x^2) + E_{y,m} + E_{z,n} & \text{1D} \\ E_{z,l} + E_{y,m} + E_{z,n} & \text{QD} \end{cases} \tag{2.88}$$

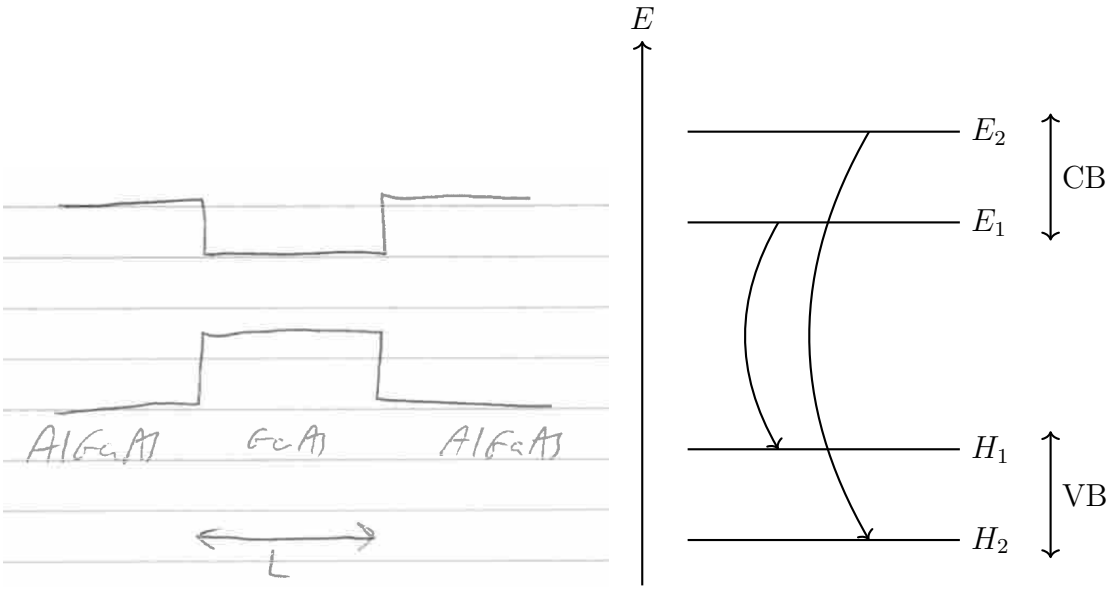


Figure 2.24: Left: GaAs/AlGaAs quantum well of width L. The hills of GaAs are caused by band offsets. Right: energy level diagram for the CB and VB, relative to the band edges.

Let's look at this in action with a GaAs/AlGaAs quantum well. It is confined in one axis as shown in Fig. 2.24.

Suppose the energy levels are arranged as in Fig. 2.24(right). Then there will be strong emission or absorption when

$$\hbar\omega_i = E_g + E_i + H_i; \quad i = 1, 2, \ldots \tag{2.89}$$

Here, we have drawn **finitely-many states**, which is not an infinite square well. We can solve the finite square well instead. You should refresh your memory of how to do so for the exam:

1. Split the finite well into 2 (or 3) regimes. The case when $V = 0$ inside the well, and the cases $V = V_0$ outside the well.
2. Always recall that $E = \hbar^2 k^2 / 2m^*$
3. Rearrange the Schroödinger equation into something you can solve, it should be second-order and easy to solve
4. Inside the well, the general solution will be $\psi = A\sin(kx) + B\cos(kx)$
5. Outside the well, the left and right side will have the form $\psi = Xe^{-\alpha x} + Ye^{\alpha x}$. where X, Y will be different constants for the left and right sides
6. Find all constants from boundary conditions and requiring differentiability at the places where the wavefunctions meet.

Anyhow, we can move on to find the transition rate $W_{i\to f}$. Recall Fermi's Golden Rule

$$W_{i\to f} = \frac{2\pi}{\hbar}\left|\langle f|\hat{H}|i\rangle\right|^2 g(\hbar\omega) \tag{2.90}$$

Due to confinement, the JDOS and matrix element will change. Here, $\hat{H} = -e\vec{r}\cdot\vec{E}$ the Hamiltonian of the system. Assume the system is confined along z. The general electron

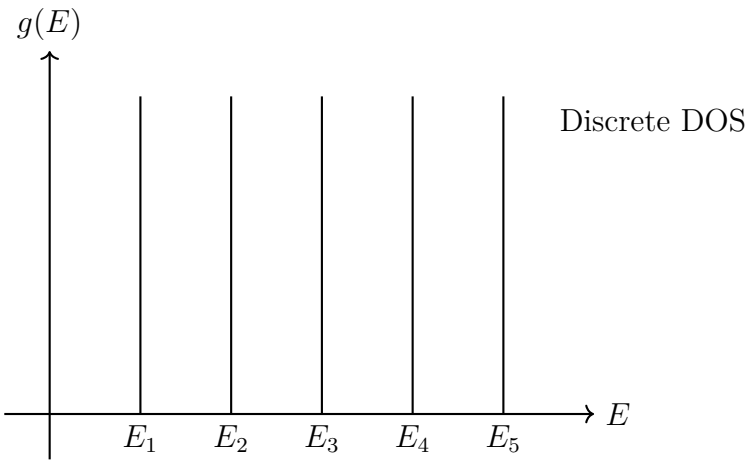


Figure 2.25: Example of a discrete density of states (DOS) for a quantum dot.

wavefunction by Bloch's theorem is

$$\Psi(x,y,z) = u(x,y)\phi(z)$$

where the functions are different by separation of variables. Indeed, this can be rewritten as

$$\Psi(\vec{k},n) = u_{\vec{k}(\vec{k}_\parallel)\phi_n}$$

- $\vec{k}_\parallel$ is the 2D wavevector in the x,y plane.
- n is the quantum number describing confined states along z
- $g(\hbar\omega)$ is as in Eq. (2.76) for a direct gap semiconductor.

Absorption of a photon takes us to the final state $\Psi \to |f\rangle$. This excites an electron into the CB, so since a hole will be left behind in the VB, the initial state $|i\rangle$ is a **hole**. The electron and hole wavefunctions are

$$\begin{aligned} |i\rangle &\propto u_h(\vec{r})e^{i\vec{k}\cdot\vec{r}}\phi_n^{(h)}(z) \\ |f\rangle &\propto u_e(\vec{r})e^{i\vec{k}'\vec{r}}\phi_n^{(e)}(z) \end{aligned} \tag{2.91}$$

where of course $u_e,, u_h$ are electron and hole Bloch functions.

For a direct-gap semiconductor, transitions are vertical in the BZ, so the wavevectors before and after are the same,

$$\boxed{\vec{k}' = \vec{k}}$$

The matrix element in Fermi's Golden Rule is

$$M \propto \langle f|z|i\rangle = \int u_e^* u_n z \phi_n^{(e)} \phi_n^{(h)} e^{i\vec{k}\cdot\vec{r}} d^3\vec{r} \tag{2.92}$$

2.6.4 DOS for confined states

Definition 2.6.2. A **quantum dot** is a material which has been confined in all three directions. Sometimes these are termed **artificial atoms**.

Because of this, all wavefunction states in each direction are discretized, leading to a discrete DOS rather than one with some continuity: You can have QDs with more complex

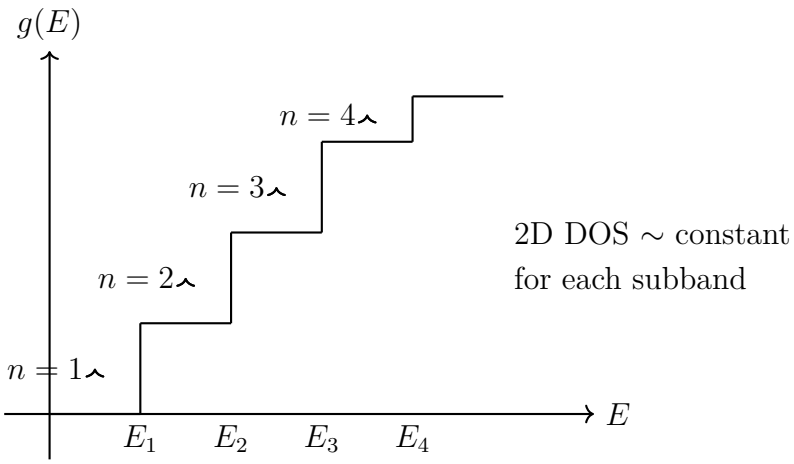


Figure 2.26: Density of states for a generic quantum well. It exhibits a step like pattern, and starts off at zero.

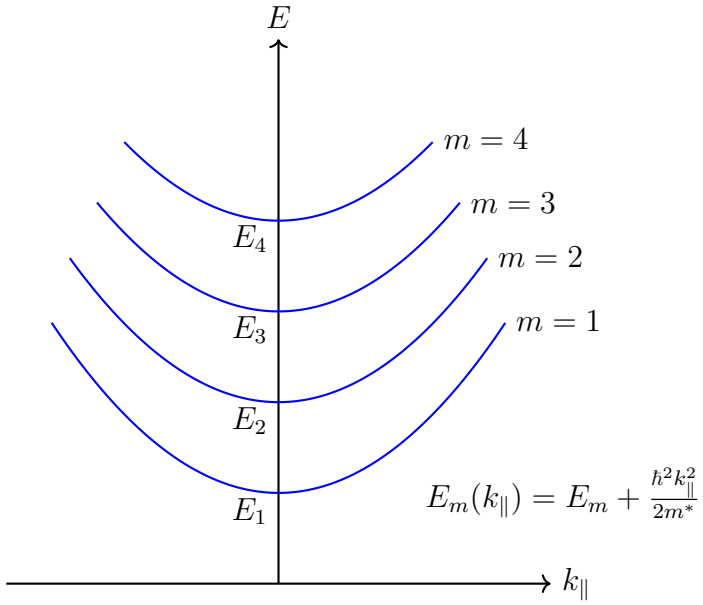


Figure 2.27: Parabolic dispersion along $x - y$ (unconfined) region.

potentials, thus more complex wavefunction indexing and energy levels, but they are **always discrete**.

In a quantum well,you have regular dispersion in 2D and confinement in 1D

$$\begin{aligned} g(E) &= \frac{2g(k)}{(\partial E/\partial k)}; \quad g(k)dk_\parallel = \frac{1}{\pi} k dk_\parallel \text{ in 2D} \\ E &= \frac{\hbar^2 k_\parallel^2}{2m^*}; \quad \frac{\partial E}{\partial k_\parallel} = \frac{\hbar^2 k_\parallel}{m^*}; \quad g(E) = \frac{2m^*}{\hbar^2 \pi} = \text{const} \end{aligned} \tag{2.93}$$

So what can we get from these equations? We see that $g(E)$ is a constant, but the derivative is positive, and depends on $k_\parallel$, therefore, we have **steps** for every n as shown in Fig. 2.26. For each discrete state n in the confined direction, each level m in the dispersion directions have branches. These are **subbands**. Plotting the energy against $k_\parallel$, we get parabolic dispersion

The absorption spectrum is also interesting, comparing it to the spectra of 2D and 3D

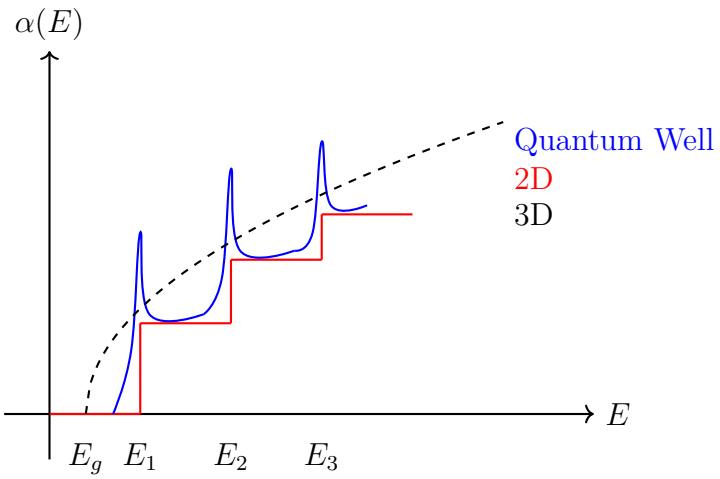


Figure 2.28: Plot of the absorption coefficient against incident photon energy for quantum well against different dispersions.

dispersion, it is a smoothened out 2D dispersion with peaks at every jump. The magnetic component of angular momentum $M \neq 0$ for $n = m$ so the allowed transitions are from hole 1 to electron 1, hole 2 to electron 2 etc. This is notated as $1 \rightarrow 1$ etc. This is shown in Fig. 2.28. For GaAs (and most group 3-5 semiconductors), there are **pairs of transitions** for light and heavy holes. Spikes at the onset of every $n \rightarrow n$ transition are due to excitons, which enhance α.

2.6.5 QDs by MBE

MBE is very good at 2D structures - is it good for QDs?

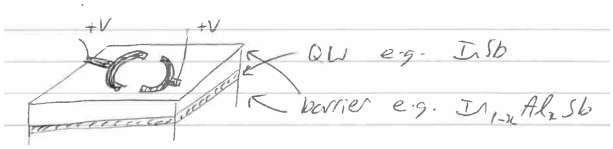


Figure 2.29: Schematic of electrically-defined QD.

Electrically-defined QDs

1. Create a quantum well with a **high mobility, narrow gap semiconductor** surrounded by a wide-gap material as in Fig. 2.29.
2. Fabricate nano-scale electrical contacts on the QW of a horseshoe shape
3. Apply a voltage V, which generates a nanoscale potential well electrically, leading to confinement in x, y - see Fig
4. z-confinement by the QW
5. Disadvantages: slow, expensive, difficult, low density of QDs
6. Applications: qubits, spin-charge devices

Strain-driven, self-assembled QDs Classic example: InAs on GaAs(001). There is a high strain due to a 6.7% lattice mismatch, but instead of forming dislocations, by pure chance, the following forms instead

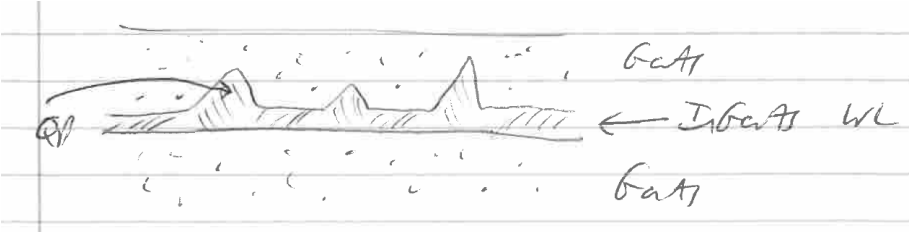


Figure 2.30: Strain-driven quantum dot islands.

- alloyed wetting layer (WL)
- After, forms 3D nanoscale islands, as seen in Fig. 2.30.
- Disadvantages: need to control strain - too much strain leads to dislocations, so τ_{NR} shortened and emission killed.
- Applications: QD lasers with low temperature sensitivity.

This process is **thermodynamically-driven**. There is lots of flexibility over the size, composition and spacing BUT since the islands are self-assembled, the positions and sizes are random. This means we have a size distribution, and so the energy levels from dot to do will be different - this is **inhomogeneous broadening**.

Droplet epitaxy for low-strain QDs

- Deposit Ga onto GaAlAs - forms nano-scale droplets
- Open As shutter, forms GaAs 3D islands
- Cap with AlGaAs, forms QDs
- Advantages: these are low strain and more homogeneous in shape and size
- Fine structure in optical emission suppressed - great source of entangled photons
- Applications: quantum key distribution

2.6.6 Ideal rectilinear QD

Suppose we have a QD with dimensions $L_x \times L_y \times L_z$. Since there is confinement in all 3 directions, the energy levels in each direction are discrete and index by n_x, n_y, n_z respectively. The energy of each state is then

$$E = \frac{t_1^2}{2n^*}\left[\left(\frac{n_x\pi}{L_x}\right)^2 + \left(\frac{n_y\pi}{L_y}\right)^2 + \left(\frac{n_z\pi}{L_z}\right)^2\right] \tag{2.94}$$

Unfortunately, real QDs are much more complicated and needs the tight-binding/ $\vec{k}\cdot\vec{p}$/perturbation theory to model them.

2.6.7 Colloidal QD

Spherical QDs have atomic-like eigenfunctions. Recall from the solution of the hydrogen atom, you had the radial component $R_{n,l}(\vec{r})$ and the spherical harmonics $Y_{l,m}(\theta,\phi)$ where l, m are the orbital and magnetic quantum numbers respectively. effectively for spherical QDs, we take the hydrogen solution and scale it by $\epsilon(0)$ and m^*.

Figure 2.31: Inverted lens QD. Screenshot from Prof. Gavin Bell's lectures.

For CdSe (a group 2-6 material), the energy

$$E = E_g + E(r, \theta, \phi) \tag{2.95}$$

where E_g is the bulk bandgap, around 1.7 eV which is red light. $E(r, \theta, \phi)$ is the **confinement energy**, derived from the ydrogen-like solution, which can go up to around 1 eV, taking us to an energy around 2.8 eV, which is blue light emission. The varying QD size, means varying confinement energies, so the entire visible spectrum is covered.

2.6.8 Inverted lens QD

Most QDs are **wider than they are tall** - e.g. InAs/GaAs. This necessitates a better model of them, called the **inverted lens QD** shown in Fig. 2.31. The lateral $(x - y)$ potential is $V(r) \propto r^2$ for small $r^2 = x^2 + y^2$. This approximation is sufficient because carriers can spill onto the InGaAs WL. Using separation of variables and cylindrical polar coordinate ϕ we have

$$\Psi(x, y, z) = \Psi(r, z) = \psi_{\|}(r, \phi)\psi_z(z) \tag{2.96}$$

By separation of variables, each term satisfies its own Schrödinger equation

$$\left[\frac{-\hbar^2}{2m^*}\nabla_{\|}^2 + \frac{1}{2}m^*\omega_0 r^2\right]\psi_{\|}(r, \phi) = E^{\|}\psi\,(r, \phi) \tag{2.97}$$

$$\left[\frac{-\hbar^2}{2m^*}\frac{d^2}{dz^2} + V(z)\right]\psi_z(z) = E^z\psi_z\,(z) \tag{2.98}$$

The total energy $E = E^{\|} + E^z$. As you know from quantum mechanics, the harmonic oscillator in 2D has 2 degree of freedoms n_x, n_y so

$$E^{\|} = (n_x + n_y + 1)\,\hbar\omega_0 \tag{2.99}$$

where $n_x, n_y \in \mathbb{N}$ including 0. Considering only the *ground state energy of* $z, E^{z,1}$ (flat QD approximation, because QDs generally broader than taller), the total energy is

$$E = E = (n_x + n_y + 1)\,\hbar\omega_0 + E^{z,1} \tag{2.100}$$

- If QDs **symmetrical** in x, y (e.g. droplet epitaxy) then E is **degenerate** in n_x, n_y: $E_{01} = E_{10}$
- If QDs **elliptical** in x, y ()e.g. strain driven InAs/GaAs) then no degeneracy - leads to **fine structure**, which is bad for high-purity single photon emission
- Possible to isolate single QDs, even if grown by self-assembly or droplet epitaxy - leads to **sharp emission lines**

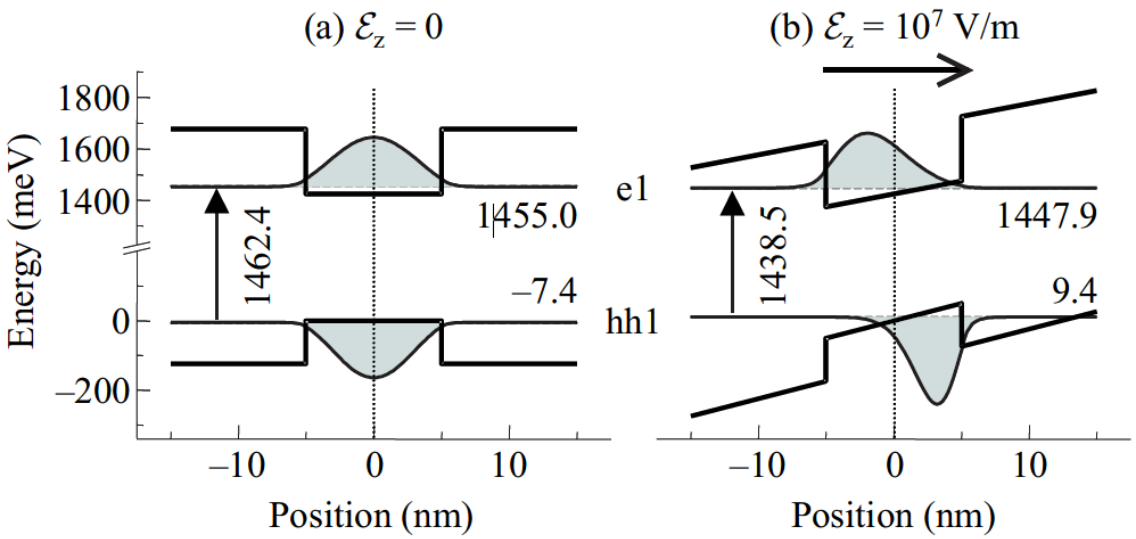


Figure 2.32: Shift in wavefunction after applying an electric field - these are calculated from approximations but you see the trend happening.

2.6.9 Quantum-confined Stark effect

Definition 2.6.3. The atomic **Stark effect** is the splitting of spectral lines after applying an external electric field $\vec{E}$

Definition 2.6.4. The **quantum-confined Stark effect** is the shifting of quantum-confined electron and hole levels after applying an external electric field $\vec{E}$ in QW/QDs

Suppose the electric field is along z, so $\vec{E} \to E_z$. The electron energy change $\Delta E_e = eE_z z$. Then:

1. Electron wavefunction shifts to the one side - e.g. left, see Fig. 2.32
2. Hole wavefunction shifts to the opposite side - e.g. right, see Fig. 2.32
3. Total overlap is **reduced** since the matrix element

$$M \propto \int_0^L \phi_h^* \phi_e dz$$

4. Luminescence and absorption **reduced**

Confinement of excitons in QWs means that electric fields **exceeding** the exciton ionisation field (*in bulk material*) can be applied.

Now, since an electric field changes energies, it must change transition energies too

$$\Delta(\hbar\omega) = \Delta E_e + \Delta E_h = -E_z\left[-e\Delta z_e + e\Delta z_h\right] \tag{2.101}$$

Since electrons and holes are oppositely-charged, they move in opposite directions under an applied field.m and $\delta z_h > 0$ and $\delta z_e < 0$

$$\Delta(\hbar\omega) = -E_z e\left(\Delta z_h - \Delta z_e\right) < 0 \tag{2.102}$$

QC Stark effect leads to **red-shift of emission/absorption**

From perturbation theory, we find $\Delta E \propto E_z^2 m^* L^4$ for small (weak) fields. For stronger fields, ΔE **saturates** because ψ_h, ψ_e 'pile-up' at opposite sides of the well. This leads to applications like tunable photodetectors.

2.6.10 Intersubband detectors

We assumed so far that the Fermi level is in the midgap where there are no available states. However, if we dope materials, we can shift the Fermi level to a different energy, say into the CB above energy E_1, so $n = 1$ is filled, but $n \geq 2$ is empty. The question is how do we calculate the transition rate from $n = 1 \to 2$

Since the binding energies in a QW are several tens of meV, the transition energy is the same and this produces an infrared spectrum.

The selection rule polarisation in-plane means $M = 0$. Intersubband detectors need a component of light's electromagnetic field in the z direction

$$M = \langle 1|z|2\rangle = \int_0^L \phi_1^* z \phi_2 dz \tag{2.103}$$

ϕ_1 and ϕ_2 must have opposite parity for $M \neq 0$, so the transition *is* allowed. This allows us to push the spectrum into far infrared or even THz frequencies!

2.7 Angle-Resolved Photo-emission Spectroscopy (ARPES)

So we have talked about the band structure and how we can use it to understand the optical response of materials. However, we want to be able to **measure** the dispersion $E(\vec{k})$ - this is the role of Angle-Resolved PhotoEmission Spectroscopy (ARPES).

2.7.1 Requirements

Not all techniques are without a catch. You will need to know the numbered bullet points and the details highlighted in **bold** may be keywords to use in a written question for the exam.

1. A good photon source
    - Photon energy $\hbar\nu > \varphi$ the **workfunction** of the material.
    - Can use **second harmonic generation/frequency doubling** to get the right frequencies. This involves using a non-linear optical process (i.e. use second order susceptibility tensor and vector potentials). For more extreme frequencies, can use **high-harmonic generation** (HHG) for target frequencies around 6-12 eV.
    - An ultraviolet (UV) gas discharge lamp is usually used, with photon frequencies $\hbar\nu = 21.2$ or 40.8 eV
    - For large-scale, **high-intensity** experiments, can use **synchrotrons** (e.g. Diamond Light Source) to produce energies in the range $30 - 1000$ eV
    - Synchrotron also allows for **polarisation control**.
2. Measurement of electron energy and momenta
    - ARPES setups have a **hemispherical analyser** with a 2D detector
    - A **time-of-flight** analyser, usually a pulsed source, to allow for accurate electron trajectory analysis
    - Movable point detector

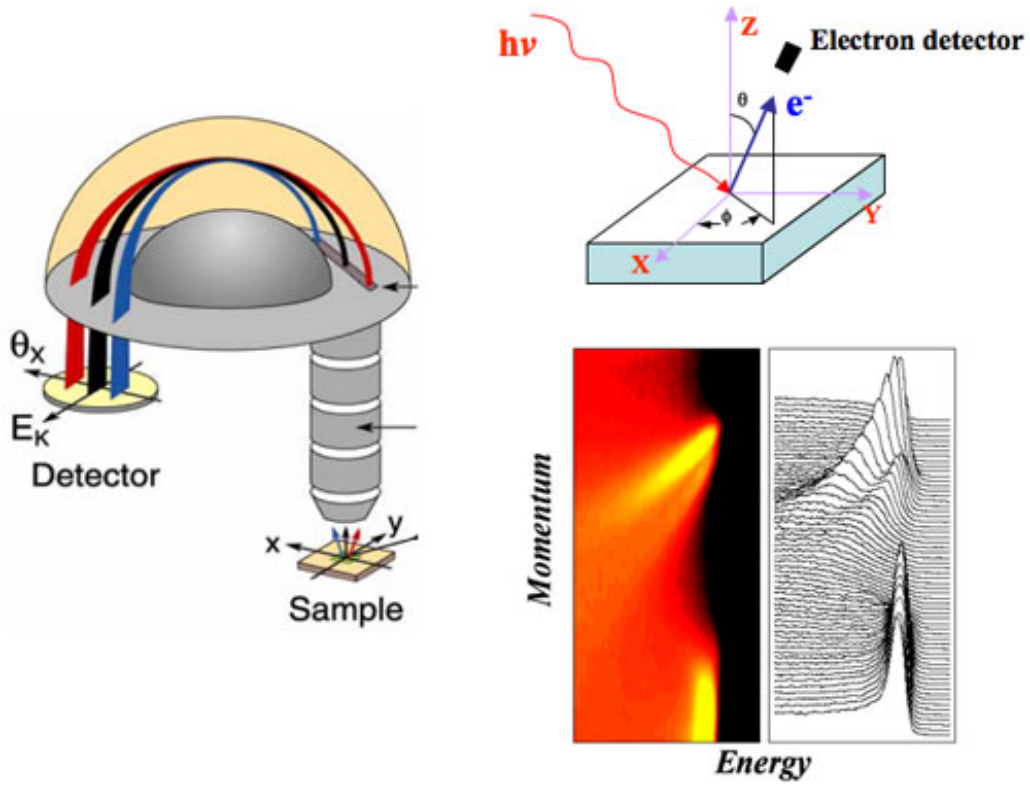


Figure 2.33: Left: typical ARPES schematic with a hemispherical analyser. Top right: coordinate system at the sample surface. Bottom right: Sample ARPES spectra in energy-momentum space. Bright areas indicate energy and momentum values where high photoelectron flux is observed and correspond to the bandstructure of the material from Stanford's Shen Laboratory.

3. Ultra-high vacuum
    - **Allows electron transport** (so they don't collide with, e.g. air molecules)
    - **Clean sample surface**
    - A short inelastic mean free path (short IMFP) of low energy electrons can lead to an **oxide layer/adsorbates** on the surface, killing ARPES measurements - means the sample has to be re-prepared or a new one used entirely (can be expensive and time-consuming)
4. **Low** $T < 77$ K usually. Allows for optimal energy resolution from 20 meV down to a few meVs.

2.7.2 Setup of ARPES

Here, $\hbar\nu = \hbar k_{\text{photon}}$ in Fig. 2.33 is small compared to the (parallel) energy $k_{\|}$ of the electrons. Additionally, the electrons are focused by an **iinput lens** into the hermispherical analyser. The orientation of the sample maps θ, ϕ to the BZ. By moving the sample angles relative to the analyser, you can map out the **intensity** $I(E_{\text{KE}}, \theta, \phi)$ where E_{KE} is the electron kinetic energy.

At the surface, we must conserve **energy** and **parallel momentum**

$$\begin{aligned} E_{\text{KE}} &= \hbar\omega - \varphi - E_B \\ p_{\|} &= \hbar\vec{k}_{\|} = \sqrt{2m_e E_{\text{KE}}}\sin\theta \end{aligned} \tag{2.104}$$

Important: $rk_\perp$ is **not conserved**. At the surface, the potential is non-periodic (unlike in the bulk). At high θ, the electron is emitted close to parallel with the surface. This means $\sin\theta \sim 1$. In this case, the parallel momentum is very large, and the electron scattering here will be due to **phonons**, and the electrons are pushed 'outside of the first BZ'. This means you can get parallel momenta

$$k_= k_\parallel + g_\parallel \tag{2.105}$$

where $g_\parallel$ is the reciprocal lattice vector. Since BZs are also periodic, you can always fold back to the first BZ. This large scattering is called **Umklapp scattering**.

As well as large momentum measurements, we also have limitations on small momentum and energy measurements:

- **Energy resolution** - this will be due to a calibration and possibly combination of the photon source, the analyser and possibly temperature
- **Momentum resolution** - if you consider a change in Eq. (2.104), you get

$$\Delta k_\parallel = \sqrt{2m_e E_{\mathrm{KE}}}/\hbar \cos(\theta)\Delta\theta \tag{2.106}$$

Now, the change in angle $\Delta\theta$ is fixed by both the input lens and analyser, but we want to improve $\Delta k_\parallel$. Well Eq. (2.106) tells us we should **lower the electron kinetic energy, so we lower the photon energy**. This means we collect more **grazing exits**. This isn't good for states at Γ but is good for other states in the BZ. The typical method for grazing-ARPES is **laser ARPES** where we can just focus a low-energy photon beam.

For 1D and 2D materials, we only care about the parallel components of energy and momentum because there is no perpendicular dispersion. For example: graphene, hexagonal boron nitride, $MoSe_2$ and surface accumulation layers in InAs.

However, for 3D materials, there *is* a perpendicular component and we would like to measure the full $E(\vec{k})$. To accomplish this, we need a technique that can satisfy the following three steps:

1. Knocks electrons out of an occupied VB state to an unoccupied state
2. Electron travels to the surface in this excited state
3. Electron overcomes potential barrier and travels to analyser as a plane wave with momentum components $(k_\perp, k_\parallel)$

Normally, the unoccupied state is modelled as a **free electron** with a minimal energy eV_0 below the vacuum level. This isn't a complex model, but it's quite accurate. V_0 is the **inner potential** - it is a parameter (found by fitting $\hbar\nu$-dependent ARPES data) and its not a physical quantity (in the sense that it isn't attached to the setup, material directly). Note that $eV_0 \neq \varphi$, the workfunction.

$$\frac{\hbar^2 k^2}{2m_e} = E_{\mathrm{KE}} = \frac{\hbar^2}{2m_e}\left(k_\parallel^2 + k_u^2\right), \tag{2.107}$$

where k_u is unconserved momentum.

However, we know

$$\frac{\hbar^2 k_\parallel^2}{2m_e} = E_{\mathrm{KE}} \sin^2\theta$$

$$\implies E_{\mathrm{KE}} \cos^2\theta = \frac{\hbar^2 k_u^2}{2m_e} = \frac{\hbar^2 k_\perp^2}{2h} - eV_0$$

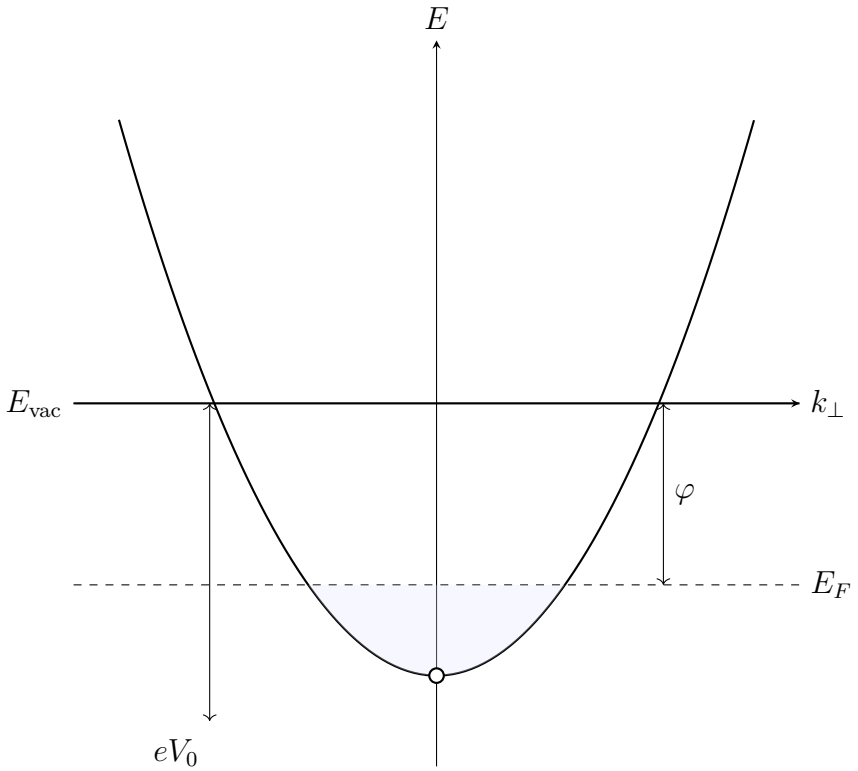


Figure 2.34: Parabolic dispersion for $k_\perp$ labelling the Fermi energy E_F, inner potential V_0. The shaded blue area represents potential occupied states

where the first term is the free-electron energy. The perpendicular component of momentum is then

$$\sqrt{(E_{\text{KE}}\cos^2\theta + eV_0)\,2m_e}\frac{1}{\hbar} = k_\perp \tag{2.108}$$

and $\varphi = E_{\text{vac}} - E_F$.

2.7.3 ARPEs Fermi-Surface Mapping

A metal is a solid with a Fermi surface. Hence, $g(E_F) \neq 0$ and the Fermi surface is just the band dispersion $E(\vec{k})$ at $E = E_F$.

For example, we can look at ARPES measurements for Cu in Fig. 2.35. What we see in Fig. 2.35(right) is only a 2D slice of the 3D Fermi surface.

In contrast, **semiconductors do not have a Fermi surface** because E_F usually lies in the **bandgap** between the VB and CB, hence $g(E_F) = 0$.

However, as mentioned previously, we can **dope** semiconductors to move E_F into the CB (p-type) or VB (n-type). AN example is heavily n-doped InSb or InAs. In this case, $E_F > E_g$ leading to **degenerate states**. Near the CBm, Cb is mostly isotropic ($E = E(|\vec{k}|)$) leading to a 3D Fermi surface and any 2D slice is circular.

However, we can get non–circular Fermi surface slices. This would mean an **anisotropic band structure**.

The prototypical example is Si, silicon, shown in Fig. 2.36.

An indirect bandgap semiconductor, it has Fermi surface **pockets not centred at** Γ for a degenerate n-type. If you look at $X \to \Gamma$ path, its curvature is *different* from $X \to K$.

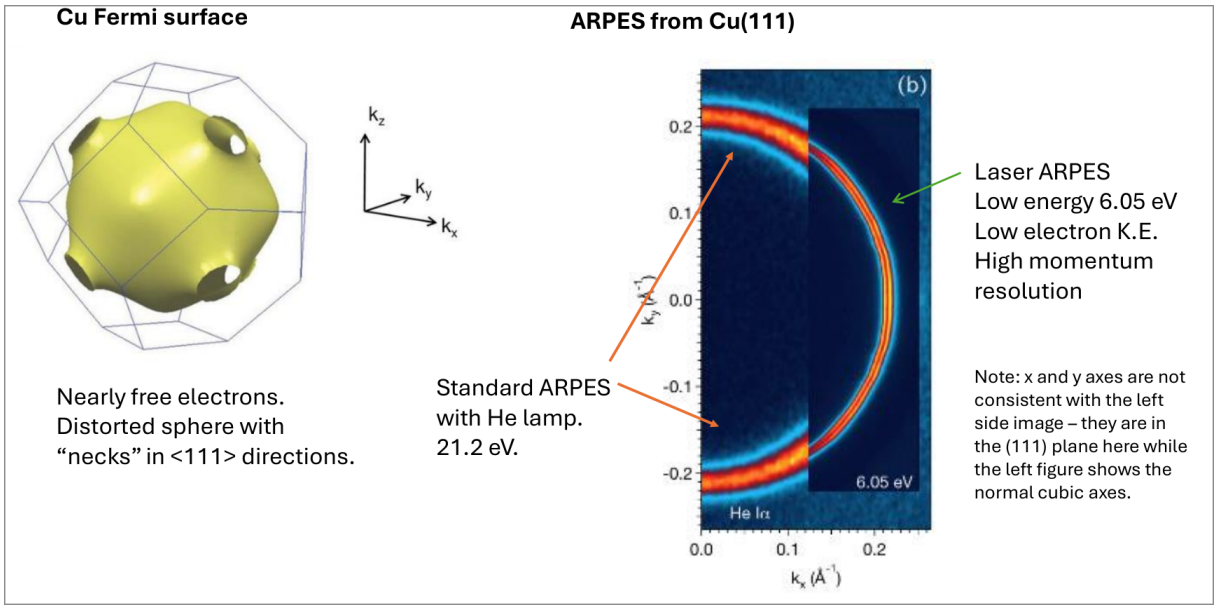


Figure 2.35: Left: Fermi surface for Cu in Cartesian reciprocal space. Right: Laser ARPES measurements in $\langle 111 \rangle$ plane. Screenshot from Prof. Gavin Bell's slides.

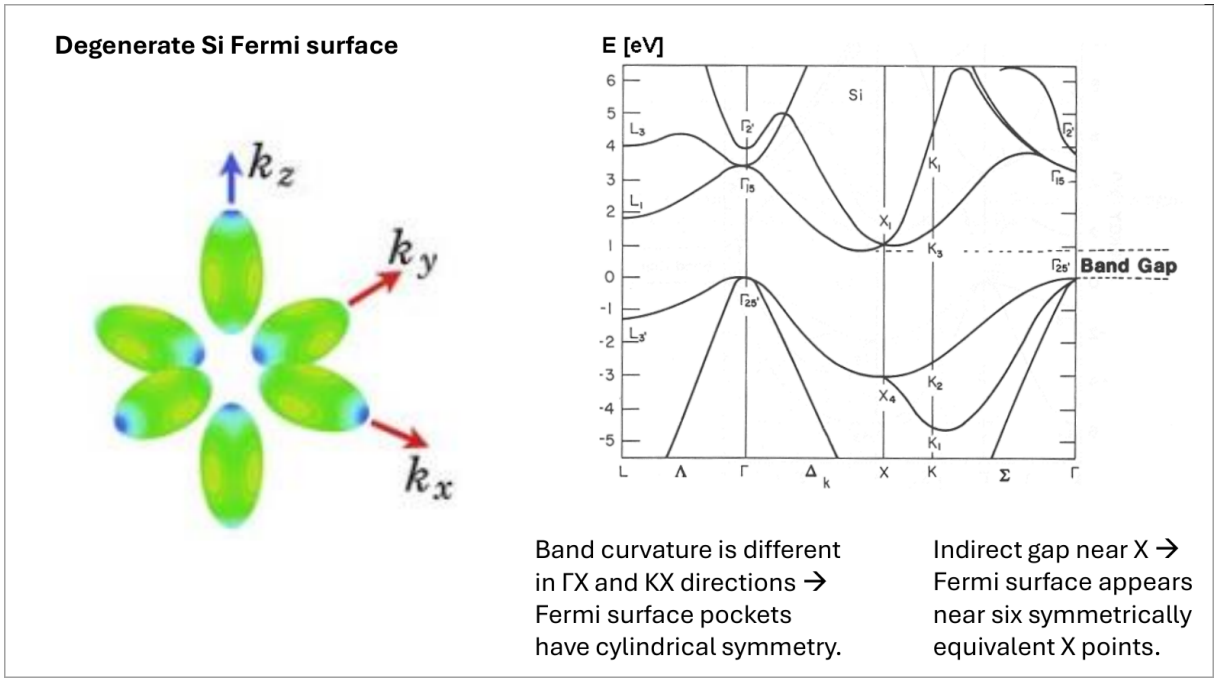


Figure 2.36: Silicon degenerate fermi surface (left) and band structure (right).

This also means the effective mass m^* is anisotropic (varies in the BZ).

Bibliography

[1] Shouvik Datta and Xavier Marie. "Excitons and excitonic materials". In: *MRS Bulletin* 49.9 (Sept. 2024), pp. 852–861. DOI: 10.1557/s43577-024-00766-x. URL: https://doi.org/10.1557/s43577-024-00766-x.

[2] M. Devoret, Andreas Wallraff, and J.M. Martinis. "Superconducting Qubits: A Short Review". In: (Dec. 2004).

[3] Dale Harshman et al. "Verification of Nodeless Superconducting Pairing in Single-Crystal YBa2Cu3O7". In: *International Journal of Modern Physics B* 17 (Aug. 2003), pp. 3582–3593. DOI: 10.1142/S0217979203021447.

[4] Konstantin K. Likharev. acessed: 2025-03-28. URL: https://phys.libretexts.org/Bookshelves/Thermodynamics_and_Statistical_Mechanics/Essential_Graduate_Physics_-_Statistical_Mechanics_(Likharev)/03%3A_Ideal_and_Not-So-Ideal_Gases/3.02%3A_Calculating_mu#mjx-eqn-46.

[5] Dewey Murdick et al. "Predicting surface free energies with interatomic potentials and electron counting". In: *J. Phys.: Condens. Matter* 17 (Oct. 2005), pp. 6123–6137. DOI: 10.1088/0953-8984/17/39/002.

[6] M. D. Sturge. "Optical Absorption of Gallium Arsenide between 0.6 and 2.75 eV". In: *Phys. Rev.* 127 (3 Aug. 1962), pp. 768–773. DOI: 10.1103/PhysRev.127.768. URL: https://link.aps.org/doi/10.1103/PhysRev.127.768.

[7] Chung Xu, Richard Schier, and Caterina Cocchi. "Electronic and optical properties of computationally predicted NaKSb crystals". In: *Electronic Structure* 7.1 (Jan. 2025), p. 015001. DOI: 10.1088/2516-1075/adaab8. URL: https://dx.doi.org/10.1088/2516-1075/adaab8.