

PHYS 502 (Condensed Matter I) Notes

Rio Weil

This document was typeset on December 5, 2022

Introduction:

This is a set of lecture notes taken from UBC's PHYS 502 (Graduate Condensed Matter I) course, taught by Dr. Marcel Franz. The course covers second quantization notation, electrons in solids, boson systems, electrons in periodic potentials, semiclassical theory of metallic conduction, electron-phonon interactions, and elements of superconductivity. The course textbook is "Solid State Physics" by Ashcroft and Mermin. If any errors are found in the notes, feel free to email me at ryoheiweil@phas.ubc.ca.

Contents

1 Solids as Interacting Quantum Many-Body Systems	5
1.1 A Condensed Matter Theory of Everything	5
1.2 The Born-Oppenheimer Approximation	6
2 Second Quantization	8
2.1 Motivation	8
2.2 The Central Idea	8
2.3 The Boson Case	8
2.4 Many-Body Hilbert Space, Creation/Annihilation Operators	9
2.5 Second Quantization Result	9
2.6 The Fermion Case	10
2.7 Field Operators	11
3 Degenerate Electron Gas	13
3.1 Introducing the Degenerate Electron Gas	13
3.2 Simplifying the background terms	14
3.3 Second Quantization of the Electron Term	15
3.4 Rescaling the Hamiltonian	16
3.5 Perturbation Theory (High Density)	17
3.5.1 Zeroth Order	17
3.5.2 First Order	18
3.5.3 Combining Results	20
3.6 The Variational Viewpoint	20
3.7 Perturbation Theory (Low Density)	21
5 The Hartree-Fock Approximation	22
5.1 Motivation and Main Idea	22
5.2 Heuristic Approach	22
5.3 Applying Hartree-Fock to Coulomb	23
5.4 Hartree-Fock as a Variational Bound	25

6	Screening	27
6.1	General Definitions and Setup of Problem	27
6.2	Thomas-Fermi Theory	27
6.3	Implications of Thomas-Fermi Theory	28
6.4	Lindhard Theory	29
6.5	Limits of the Lindhard Results	30
7	Bosons, Bose-Einstein Condensation, and Helium-4	32
7.1	Housekeeping	32
7.2	Boson types and statistics	32
7.3	Deriving the Bose-Einstein Distribution	32
7.4	Bose-Einstein Condensation	33
7.5	Bogoliubov Theory of Helium-4	34
7.6	Bogoliubov Transformations and Quasiparticle Spectrum	35
7.7	Quasiparticle Spectrum - Special Cases	36
7.7.1	Non-Interacting	36
7.7.2	Contact Repulsion	37
7.7.3	Typical Helium-4 Interaction	37
7.8	Landau Argument for Superfluidity	37
8	Lattice Vibrations & Phonons	39
8.1	Phonons in One Dimension	39
8.2	Diagonalizing the Potential	40
8.3	Translation-Invariant Systems	41
8.4	Example 1 - Monoatomic chain	42
8.5	Example 2 - Diatomic Chain	43
9	Phonons in Three Dimensions	45
9.1	Review - The Real and Reciprocal 3D Lattice	45
9.2	Writing down the 3D Hamiltonian	45
9.3	Translation Invariant Solution	46
9.4	Debye Model, Specific Heat of Phonons	47
9.5	Debye Frequency, Momentum, Temperature	48
9.6	Back to Heat Capacity	49
9.6.1	Low T behavior	49
9.6.2	High T behavior	50
9.7	Einstein Model	50
9.8	Anharmonic Effects and Phonon Interactions	51
9.9	Thermal Conductivity	53
10	Magnons	55
10.1	The Heisenberg Model	55
10.2	Magnon Variables	55
10.3	Ferromagnetic Case	56
10.4	An Example: Cubic Lattice in 3D	57
10.5	Magnetization Reversal	58
11	Electrons in a Periodic Potential: Band Theory of Solids	60
11.1	Review: Bloch's Theorem	60
11.2	Weak Periodic Potential	60
11.2.1	Zeroth-order perturbation theory	61
11.2.2	First-order perturbation theory	61

11.2.3 Second-order perturbation theory	62
11.3 Degenerate Perturbation Theory	62
12 Electrons in a Periodic Potential - Tight Binding Approximation	63
12.1 Review and Examples of Band Structure	63
12.2 Tight-Binding Approximation	63
12.3 Example - Single non-degenerate orbital in 2D	66
12.3.1 Effective Mass	67
12.3.2 Band Velocity	68
12.4 Example - Lattice with a Basis (Dimerized chain)	68
13 Density Functional Theory	71
13.1 Motivation	71
13.2 The Hohenberg-Kohn Theorem	71
13.3 Proof of the Hohenberg-Kohn Theorem	71
13.3.1 Part 1	71
13.3.2 Part 2	72
13.4 Variational Principle	72
13.5 The Kohn-Sham Formulation	73
13.5.1 DFT for the Reference System	73
13.5.2 Remarks	74
13.6 Choices for the Exchange-Correlation Potential	74
14 Semiclassical Theory of Conduction I	75
14.1 Semiclassical QM in Crystals	75
14.2 Limits of Validity	75
14.3 Filled Bands are Inert	75
14.4 Motion in a Uniform Electric Field	76
14.5 The Hole Picture	77
14.6 Motion in Uniform Magnetic Field	77
15 Semiclassical Theory of Conduction II	78
15.1 Non-equilibrium Distribution Function	78
16 Semiclassical Theory of Conduction III	80
16.1 Solving for Non-Equilibrium Distribution	80
16.2 DC electrical conductivity	80
16.3 Thermal Conductivity	81
17 Electron-Phonon Interactions	84
17.1 Deriving the Electron-Phonon Interaction Hamiltonian	84
17.2 Kohn Anomaly and Peierls Instability	85
17.2.1 3 dimensions - Kohn Anomaly	86
17.2.2 1 dimension - Peierls Instability	87
17.3 Polyacetylene - A Classic Peierls Example	88
17.4 Polarons and Mass Enhancement	90
18 Phonon-Mediated Attraction Between Electrons	93
18.1 Canonical Transformations	93
18.2 Applying Canonical Transformations to the Frolich Hamiltonian	94
18.3 A physical picture for attractive electrons	95

19 Superconductivity	97
19.1 Phenomenology and Historical Background	97
19.2 Cooper Instability	97
20 The BCS Ground State	100
20.1 Ground State Wavefunction	100
20.2 Calculation of u_k, v_k	100
20.3 BCS Condensation Energy	104

1 Solids as Interacting Quantum Many-Body Systems

In this class we will largely discuss the theory of solids. A solid is anything that is rigid, but specifically we will discuss solids as ions arranged in a regular lattice plus electrons.

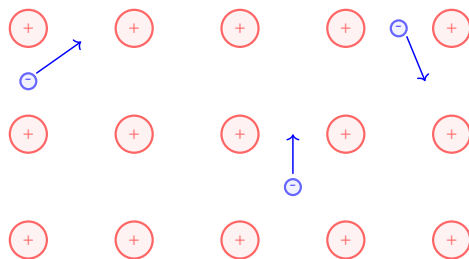


Figure 1.1: A cartoon visualization of a solid, here a square regular lattice with free electrons.

1.1 A Condensed Matter Theory of Everything

Consider the Hamiltonian:

$$H = \sum_i \frac{p_i^2}{2M} + \frac{(Ze)^2}{2} \sum_{i,i'} \frac{1}{|\mathbf{R}_i - \mathbf{R}_{i'}|} + \sum_j \frac{p_j^2}{2m} + \frac{e^2}{2} \sum_{j,j'} \frac{1}{|\mathbf{r}_j - \mathbf{r}_{j'}|} - Ze^2 \sum_{i,j} \frac{1}{|\mathbf{R}_i - \mathbf{r}_j|. \quad (1.1)$$

First term is ion KE, second term is ion-ion Coulomb interaction, third term is electron KE, fourth term is electron-electron Coulomb interaction, fifth term is ion-electron Coulomb interaction. This is in principle the theory of everything, which encompasses all that there is need to know in a solid. Note that spin is missing here; we should add two copies of everything (spin up, spin down) and relativistic effects (spin orbit coupling) but for most solids these are relatively small corrections. However, there is a large problem; this is a largely intractable problem. The main problem is that N (the number of electrons in a given solid) is extremely large; $N \sim 10^{23}$. Let's consider some cases of N .

- $N = 1$ is the hydrogen atom; this has been solved by Schrodinger (and in undergraduate QM) exactly.
- $N = 2$ is the Helium atom; already there exists no exact solution. But there are approximate methods that work well (e.g. variational principle for finding the ground state energy)
- $N = 1 - 100$ is the whole of chemistry; there are more sophisticated approximation techniques here.
- $N \sim 10^{23}$ is the theory of solids.

The key issue of the problem is the size of the corresponding Hilbert space is *enormous*. It's even hard to estimate how large, as position and momentum are continuous. But just to illustrate the size of \mathcal{H} for $N = 10^{23}$, let's consider a simpler setting where we only consider spin and ignore all of the motional degrees of freedom. For spin, there are two states; \uparrow and \downarrow per electron. So the total number of basis states is $2^N = 2^{10^{23}} \approx 10^{10^{23}/3}$. There is no computer possible that can store this much information! In fact as an amusing comparison, there are only 3.8×10^{50} atoms on Earth, 1.2×10^{57} atoms on the sun, and 1.3×10^{79} atoms in the visible universe; our brute force method is destined to fail. Our conclusion is that drastic approximations are required in order to make progress in any valid description of solids. And note that they may be drastic, but these approximations turn out to be quite good; there is some simplicity that emerges from what seems to be a hopelessly large and complex Hilbert space. We can achieve a very good understanding of many things; e.g. the physics necessary to construct the device on which this document was written.

1.2 The Born-Oppenheimer Approximation

The Born-Oppenheimer, or adiabatic approximation was originally developed as an approximation method to describe complex molecules; however it applies to our current discussion of solids. It is based on the observation that $M \gg m$ (where M is the ion mass and m the electron mass), namely $\frac{m}{M} \sim 10^{-3} - 10^{-5}$. We imagine that in a complicated system of electrons and ions we have equipartition of energy¹; because the energy scales of electrons and ions are comparable, the electrons will be moving much faster. Therefore it is possible to decouple the problem of electrons and phonons, by solving the electron motion on a static background of ions.

One can deduce that $v_{ion} \sim \left(\frac{m}{M}\right)^{3/4} v_F \sim 10^{-2} - 10^{-3} v_F$. Also, $v_F \sim 3 \times 10^6 \text{m/s} \sim 10^{-2} c$ so the physics we consider is non-relativistic (and we can add corrections to the order of 1%). There are various supposedly intuitive arguments for why we have a power of 3/4 on $\frac{m}{M}$, but most are not at all obvious or really reasonable; we will derive it after going further into our discussion of solids.

We explore the consequences of $v_{ion} \ll v_F$ for solutions of the Schrodinger equation:

$$H\psi(\mathbf{r}, \mathbf{R}) = E\psi(\mathbf{r}, \mathbf{R}) \quad (1.2)$$

where $\mathbf{r} = \{\mathbf{r}_j\}_j$ and $\mathbf{R} = \{\mathbf{R}_i\}_i$. We make the ansatz:

$$\psi(\mathbf{r}, \mathbf{R}) = \sum_n \phi_n(\mathbf{R}) \psi_{e,n}(\mathbf{r}, \mathbf{R}) \quad (1.3)$$

where $\psi_{e,n}$ are solutions to the *electron* problem at fixed ion positions. In other words:

$$\boxed{(T_e + V_{ee} + V_{ei})\psi_{e,n}(\mathbf{r}, \mathbf{R}) = E_{e,n}(\mathbf{R})\psi_{e,n}(\mathbf{r}, \mathbf{R})} \quad (1.4)$$

This in itself is an intractable problem, but it will be useful for our analysis to assume a solution of this form. Let us substitute our ansatz into the SE. We then obtain:

$$(T_i + T_e + V_{ii} + V_{ee} + V_{ei})\psi = E\psi.$$

We can rewrite this as:

$$(T_i + V_{ii})\psi + \sum_n \phi_n(T_e + V_{ee} + V_{ei})\psi_{e,n} = E\psi$$

But the term in brackets of the sum is the electronic part, so:

$$(T_i + V_{ii})\psi + \sum_n E_{e,n}(\mathbf{R})\phi_n\psi_{e,n}(\mathbf{r}, \mathbf{R}) = E\psi \quad (1.5)$$

We can not multiply by $\psi_{e,m}^*(\mathbf{r}, \mathbf{R})$ and integrate over \mathbf{r} . We then have many simplifications that arise from orthonormality (namely in the second term and the RHS). But the first term on the RHS is nontrivial as T_i contains $\nabla_{\mathbf{R}}$. In any case, we are left with:

$$\sum_n \int d\mathbf{r} \psi_{e,m}^*(\mathbf{r}, \mathbf{R}) T_i \phi_n(\mathbf{R}) \psi_{e,n}(\mathbf{r}, \mathbf{R}) + (V_{ii} + E_{e,m}(\mathbf{R}) - E)\phi_m(\mathbf{R}) = 0. \quad (1.6)$$

where we have used the orthonormality of $\psi_{e,n}$ to collapse most of the terms. Let us now analyze the troublesome term. We rewrite this as $\sum_i \langle em | \frac{p_i^2}{2M} \phi_n(\mathbf{R}) | en \rangle$. P_i^2 is a second derivative, so we end up getting

¹Equipartition is a result from classical physics, but it applies suprisingly well.

three terms; one term where both derivatives act on ϕ_n , a term where one acts on ϕ_n and the other on $|en\rangle$, and the last where both act on $|en\rangle$. Explicitly, we can write it as:

$$\sum_i \langle em | \frac{p_i^2}{2M} \phi_n(\mathbf{R}) | en \rangle = -\frac{\hbar^2}{2M} \sum_i \int d\mathbf{r} \psi_{e,m}^*(\mathbf{r}, \mathbf{R}) \left[(\nabla_{R_i}^2 \phi_n(\mathbf{R})) + 2(\nabla_{R_i} \phi_n(\mathbf{R})) \nabla_{R_i} + \phi_n(\mathbf{R}) \nabla_{R_i}^2 \right] \psi_{e,n}(\mathbf{r}, \mathbf{R}). \quad (1.7)$$

The first term can be evaluated (as before) using orthonormality. The other two are not as convenient, but in the B-O approximation we may neglect the other two terms (and we will discuss shortly why this is a good idea).

We obtain the following equation:

$$\boxed{[T_i + V_{ii} + E_{e,n}(\mathbf{R})] \phi_n(\mathbf{R}) = E_n \phi_n(\mathbf{R})} \quad (1.8)$$

Note we can solve Eq. (1.4) assuming the ions are static/in fixed positions. From there we obtain $E_{e,n}(\mathbf{R})$ which allows us to solve Eq. (1.8) (which is known as a phonon equation), which allows us to obtain E_n and $\phi_n(\mathbf{R})$, which gives us the solution of the whole problem. We have decoupled one very complex problem into two connected but separately solvable equations. $E_{e,n}(\mathbf{R})$ is called the effective ionic potential; without it a crystal would blow apart (via repulsive interaction), but it holds things together.

As a last step, we must still demonstrate that the two neglected terms in (1.7) make negligible contributions. One can show that:

1. The first term is order $\left(\frac{m}{M}\right)^{1/2} \epsilon_F$.
2. The second term is order $\left(\frac{m}{M}\right)^{3/4} \epsilon_F$.
3. The third term is order $\left(\frac{m}{M}\right) \epsilon_F$.

However these estimates are for now opaque; we will confirm them later by further analysis. For now though, we recall that $\frac{m}{M} \sim 10^{-3} - 10^{-5}$ so the second/third terms tend to be at least an order of magnitude smaller than the first (and can be neglected to first order). However, some important properties of crystalline solids are actually derived from these terms. For example if we analyze how electron motion couples to lattice vibration, then we have to start to worry about them. Additionally, the terms also contribute to resistivity of metals (phonon-electron scattering, especially important at high temperatures; why resistivity drops at lower temps)

So to start, we will study the electron and lattice degrees of freedom separately, but as we go further into our study we will have to revisit these coupling terms. Next week,

2 Second Quantization

2.1 Motivation

The goal is to re-state the familiar Schrodinger equation:

$$i\hbar \frac{\partial}{\partial t} \psi(\mathbf{x}_1, \dots, \mathbf{x}_N, t) = H\psi(\mathbf{x}_1, \dots, \mathbf{x}_N, t). \quad (2.1)$$

in a more convenient format for $N \sim 10^{23}$. Second quantization is a bit of a misnomer; we will not quantize any further, but we will just recast the SE into a more convenient basis. Here we will give a summary of the derivation, and the gory mathematical details left to self-study; refer to the Chapter 1 handout of Fetter and Walecka.

We will consider the following Hamiltonian as an example:

$$H = \sum_{k=1}^N T(\mathbf{x}_k) + \frac{1}{2} \sum_{k < l}^N V(\mathbf{x}_k, \mathbf{x}_l) \quad (2.2)$$

where we have the single-particle operator T (kinetic energy) and the two-particle operator V (interaction; e.g. Coulomb).

2.2 The Central Idea

The problem is that the number of variables that this wavefunction depends on is absolutely astronomical. The key will be that any two electrons are *fundamentally indistinguishable*; instead of keeping track of $N \sim 10^{23}$ positions, it is sufficient to specify how many particles occupy a given single-particle state. To this end we choose a basis of single particle states $\psi_{E_k}(\mathbf{x}_k)$ where E_k represents a complete set of single-particle quantum numbers² (e.g. momentum \mathbf{p} for spinless bosons in a 3d box, or n, l, m, s_z for an electron in a hydrogen atom). We then write the many-body wavefunction in this basis as:

$$\psi(\mathbf{x}_1, \dots, \mathbf{x}_N, t) = \sum_{E_1, \dots, E_N} C(E_1, \dots, E_N, t) \psi_{E_1}(\mathbf{x}_1) \dots \psi_{E_N}(\mathbf{x}_N) \quad (2.3)$$

We must distinguish two possible cases for these particles; namely they can either be bosons or fermions i.e. take care of the “exchange statistics”. This is encoded in the many body wavefunction as a property of how the wavefunction behaves under exchange of any two particles:

$$\psi(\dots, \mathbf{x}_i, \dots, \mathbf{x}_j, \dots, t) = \pm \psi(\dots, \mathbf{x}_j, \dots, \mathbf{x}_i, \dots, t) \quad (2.4)$$

with $+$ corresponding to bosons and $-$ corresponding to fermions. This has far-reaching consequences for the nature of many-body states. If this wavefunction rule is obeyed, the coefficients must obey the same rule:

$$C(\dots, E_i, \dots, E_j, \dots, t) = \pm C(\dots, E_j, \dots, E_i, \dots, t). \quad (2.5)$$

Bosons are a bit easier, so we discuss them first.

2.3 The Boson Case

For the sake of simplicity, we will imagine that the E_j s are represented by integers, namely $E_j \in \mathbb{N}$. Suppose we have coefficient $C(12134115\dots, t)$. Since we are free to exchange any of the integers as we like, we may arrange it as:

$$C(12134115\dots, t) = C(1111\dots 2222\dots 333\dots, t)$$

² E_k does *not* represent energy

where we have n_1 1s, n_2 2s, n_3 3s and so on. It should be immediately clear that it is not necessary to keep track of all 10^{23} numbers, but just the number of particles in each state (each number). We then define:

$$C(1111 \dots 2222 \dots 333 \dots, t) \equiv \bar{C}(n_1, n_2, \dots, n_\infty, t). \quad (2.6)$$

In analogy, when we think about our bank account, we do not care about the individual dollars or what they look like; we only care about the total number of dollars in each of our accounts. We can then write the wavefunction in terms of \bar{C} , and then massage the resulting expressions to obtain convenient equations (as is done in the text).

2.4 Many-Body Hilbert Space, Creation/Annihilation Operators

We introduce a many-body Hilbert space and creation/annihilation operators that act on states in the space. States in the space look like. These states are orthonormal and complete:

$$\begin{aligned} \langle n'_1 n'_2 \dots n'_\infty | n_1 n_2 \dots n_\infty \rangle &= \prod_{i=1}^{\infty} \delta_{n_i, n'_i} \\ \sum_{n_1, n_2, \dots, n_\infty} |n_1 n_2 \dots n_\infty\rangle \langle n_1 n_2 \dots n_\infty| &= 1. \end{aligned} \quad (2.7)$$

We then define the creation/annihilation operators by defining their commutation relations. For bosons, we have:

$$\begin{aligned} [b_k, b_{k'}^\dagger] &= \delta_{kk'} \\ [b_k, b_{k'}] &= 0 \\ [b_k^\dagger, b_{k'}^\dagger] &= 0 \end{aligned} \quad (2.8)$$

where b_k^\dagger is said to create a boson in state $\psi_{E_k}(\mathbf{x})$. We record the notation:

$$|n_1 n_2 \dots n_\infty\rangle = |n_1\rangle \otimes |n_2\rangle \otimes \dots \otimes |n_\infty\rangle. \quad (2.9)$$

We can now use the commutation relations to count the number of particles, as well as create and annihilate them:

$$\begin{aligned} b_k^\dagger b_k |n_k\rangle &= n_k |n_k\rangle \\ b_k |n_k\rangle &= \sqrt{n_k} |n_k - 1\rangle \\ b_k^\dagger |n_k\rangle &= \sqrt{n_k + 1} |n_k + 1\rangle. \end{aligned} \quad (2.10)$$

and if there is no boson to destroy (i.e. we have the vacuum state $|0\rangle$), we have the special case of:

$$b_k |0\rangle = 0.$$

Most of the states of interest in CM physics is low-temperature states where there are limited number of states with large occupancies. E.g. Bose-Einstein condensation, where all particles go into the single-particle ground state (n_1 is huge, n_i for $i > 1$ are zero). When you heat up this condensate a little, n_1 will still be large, and the excited states will start to be occupied.

2.5 Second Quantization Result

With these definitions, one can show (see F&W) that Eq. (2.1) becomes:

$$\begin{aligned} i\hbar \frac{\partial}{\partial t} |\psi(t)\rangle &= H |\psi(t)\rangle \\ H &= \sum_{i,k} \langle i|T|j\rangle b_i^\dagger b_j + \frac{1}{2} \sum_{ijkl} \langle ij|V|kl\rangle b_i^\dagger b_j^\dagger b_l b_k. \end{aligned} \quad (2.11)$$

Note the order of $b_i b_k$ above. This does not matter for bosons (as the two are seen to commute via the commutation relations), but it will matter for fermions, as we will soon see. As a reminder, $|\psi(t)\rangle$ lives in the many-body Hilbert space:

$$|\psi(t)\rangle = \sum_{n_1 n_2 \dots n_\infty} f(n_1, n_2, \dots, n_\infty, t) |n_1 n_2 \dots n_\infty\rangle.$$

In second quantization, the important quantities of interest to calculate will be the matrix elements of T and V with respect to the chosen multi-particle basis.

2.6 The Fermion Case

For fermions, the anti-symmetry under exchange implies the Pauli exclusion principle; that is, at most one fermion can occupy a given state. To see this in terms of the coefficients C , we have the relation:

$$C(11\dots) = -C(11\dots) \quad (2.12)$$

where we have interchanged the 1s. The only way this can be satisfied is if $C(11\dots) = 0$. So for the coefficient to be nonzero, all of the fermions must be in different states. In second quantization, this is implemented by the anti-commutation relations of creation and destruction operators:

$$\begin{aligned} \{c_s, c_{s'}^\dagger\} &= \delta_{kk'} \\ \{c_s, c_{s'}\} &= 0 \\ \{c_s^\dagger, c_{s'}^\dagger\} &= 0 \end{aligned} \quad (2.13)$$

Where $\{A, B\} = AB + BA$ is the anticommutator. We can derive the following properties:

1. $\{c_s^\dagger, c_s^\dagger\} = 2c_s^\dagger c_s^\dagger = 0$, so:

$$c_s^{\dagger 2} = 0 \implies c_s^{\dagger 2} |0\rangle = 0 \quad (2.14)$$

this is a restatement of the Pauli principle. We cannot create two fermions in the same state. Analogously, $c_s^2 = 0$.

2. We have the number operator (as in the boson case) of $\hat{n} = c^\dagger c$. We then have that:

$$(\hat{n})^2 = (c^\dagger c)^2 = c^\dagger c c^\dagger c = c^\dagger (1 - c^\dagger c) c = c^\dagger c = \hat{n} \quad (2.15)$$

where in the second-to-last equality we use the anticommutation relation and for the last equality we use that $c^{\dagger 2} = 0$. So, the number operator has the property of idempotency. From this we can conclude that \hat{n} has eigenvalues of 0 and 1 (as these are the only values that square to 1). This again is consistent with the Pauli exclusion principle; either we have zero or one fermions in a given quantum state.

3. It is easy to deduce:

$$\begin{aligned} c^\dagger |0\rangle &= |1\rangle & c |1\rangle &= |0\rangle \\ c^\dagger |1\rangle &= c^\dagger c^\dagger |0\rangle = 0 \\ c |1\rangle &= c c |0\rangle = 0 \end{aligned} \quad (2.16)$$

A note with bookkeeping; because of the anti-commutation rules, it becomes necessary to track signs in many-body states. We have the following many-particle state which we apply c_s to:

$$\begin{aligned} |n_1 n_2 \dots n_\infty\rangle &= (c_1^\dagger)^{n_1} (c_2^\dagger)^{n_2} \dots (c_\infty^\dagger)^{n_\infty} |0\rangle. \\ c_s |n_1 n_2 \dots n_\infty\rangle &= (-1)^{s_s} (c_1^\dagger)^{n_1} (c_2^\dagger)^{n_2} \dots (c_s c_s^\dagger) \dots (c_\infty^\dagger)^{n_\infty} |0\rangle \end{aligned} \quad (2.17)$$

where $s_s = \sum_{i=1}^{s-1} n_i$; the sign has been accumulated by pushing the c_s through. This implies the following rules for many-body fermion states:

$$\begin{aligned} c_s |\dots n_s \dots\rangle &= \begin{cases} (-1)^{s_s} \sqrt{n_s} |\dots n_s - 1 \dots\rangle & \text{if } n_s = 1 \\ 0 & \text{if } n_s = 0 \end{cases} \\ c_s^\dagger |\dots n_s \dots\rangle &= \begin{cases} (-1)^{s_s} \sqrt{n_s + 1} |\dots n_s + 1 \dots\rangle & \text{if } n_s = 0 \\ 0 & \text{if } n_s = 1 \end{cases} \\ c_s^\dagger c_s |\dots n_s \dots\rangle &= n_s |\dots n_s \dots\rangle. \end{aligned} \quad (2.18)$$

Note however the quantities in square roots are always one, so we can just forget about them.

Similar to bosons, we can rewrite Eq. (2.1) as:

$$\begin{aligned} i\hbar \frac{\partial}{\partial t} |\psi(t)\rangle &= H |\psi(t)\rangle \\ H &= \sum_{rs} \langle r|T|s\rangle c_r^\dagger c_s + \frac{1}{2} \sum_{rstn} \langle rs|V|tn\rangle c_r^\dagger c_s^\dagger c_u c_t. \end{aligned} \quad (2.19)$$

where we again note the order of the annihilation operators.

2.7 Field Operators

It is often convenient to create a particle at a point \mathbf{x} . To this end, we define field operators:

$$\hat{\psi}(\mathbf{x}) = \sum_k \psi_k(\mathbf{x}) c_k, \quad \hat{\psi}^\dagger(\mathbf{x}) = \sum_k \psi_k^\dagger(\mathbf{x}) c_k^\dagger \quad (2.20)$$

These can be viewed as a kind of Fourier transform, or more generally as a change of basis. As an example, consider spin-1/2 fermions. We can label them by the momentum \mathbf{k} and the spin s_z . The index k can be thought as $k = (\mathbf{k}, s_z)$. Then, $\psi_k(\mathbf{x})$ can be thought of as two-component spinors, where:

$$\psi_k(\mathbf{x}) = \begin{pmatrix} \psi_{\mathbf{k}}(\mathbf{x})_1 \\ \psi_{\mathbf{k}}(\mathbf{x})_2 \end{pmatrix} \quad (2.21)$$

which can be thought of the wavefunctions for the spin up and down projections. We can easily deduce commutation (and anti-commutation) relations for the field operators:

$$\begin{aligned} [\hat{\psi}_\alpha(\mathbf{x}), \hat{\psi}_\beta^\dagger(\mathbf{x}')]_\pm &= \sum_k \psi_k(\mathbf{x})_\alpha \psi_k^\dagger(\mathbf{x}')_\beta^* = \delta_{\alpha\beta} \delta(\mathbf{x} - \mathbf{x}'). \\ [\hat{\psi}_\alpha(\mathbf{x}), \hat{\psi}_\beta(\mathbf{x}')]_\pm &= [\hat{\psi}_\alpha^\dagger(\mathbf{x}), \hat{\psi}_\beta^\dagger(\mathbf{x}')]_\pm = 0 \end{aligned} \quad (2.22)$$

Similarly, the Hamiltonian can be written as:

$$H = \int d^3x \hat{\psi}^\dagger(\mathbf{x}) T(\mathbf{x}) \hat{\psi}(\mathbf{x}) + \frac{1}{2} \int d^3x d^3y \hat{\psi}^\dagger(\mathbf{x}) \hat{\psi}^\dagger(\mathbf{y}) V(\mathbf{x}, \mathbf{y}) \hat{\psi}(\mathbf{y}) \hat{\psi}(\mathbf{x}). \quad (2.23)$$

and we invite the reader to check that this is indeed the case. Note that in general there should always be the same number of creation and annihilation operators; except towards the end of the course when we will look at superconducting systems, where there will be more creation than annihilation (and we will think about how to understand this). There are other operators that can be discussed in the same way.

1. The current operator. The form of the operator in first and second quantization is given below:

$$\begin{aligned}
 J(\mathbf{x}) &= \sum_{i=1}^N J(\mathbf{x}_i) \\
 \hat{J} &= \sum_{rs} \langle r | J | s \rangle c_r^\dagger c_s = \int d^3x \sum_{rs} \psi_r^\dagger(\mathbf{x}) J(\mathbf{x}) \psi_s(\mathbf{x}) c_r^\dagger c_s = \int d^3x \hat{\psi}^\dagger(\mathbf{x}) J(\mathbf{x}) \hat{\psi}(\mathbf{x})
 \end{aligned} \tag{2.24}$$

We see why the second quantization notation is so useful/economical; we simply take the first-quantized operator, sandwich it between field operators, and integrate.

2. The number operator. This follows much of the same logic as the current operator above. The first expression gives the number density, and the second the total particle number (which is obtained by integrating the number density over all space).

$$\begin{aligned}
 \hat{n}(\mathbf{x}) &= \sum_{r,s} \psi_r^\dagger(\mathbf{x}) \psi_s(\mathbf{x}) c_r^\dagger c_s = \hat{\psi}^\dagger(\mathbf{x}) \hat{\psi}(\mathbf{x}) \\
 \hat{N} &= \int d^3x \hat{n}(\mathbf{x}) = \int d^3x \hat{\psi}^\dagger(\mathbf{x}) \hat{\psi}(\mathbf{x}).
 \end{aligned} \tag{2.25}$$

3 Degenerate Electron Gas

The office hours for this course will be Monday 4-5pm with Marcel, and 4-5pm on Tuesday with Oguzhan.

3.1 Introducing the Degenerate Electron Gas

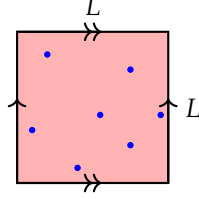


Figure 3.1: A cartoon depiction of the degenerate electron gas model. We consider a fixed, finite number of electrons N in a three-dimensional box with side length L and periodic boundary conditions. The electrons feel a uniformly distributed background of positive charge.

Also known as the “Jellium Model”, we consider a gas of electrons moving in a uniformly distributed background of positive charge. We begin with a 3d box of size L , and then take the thermodynamic limit of $L \rightarrow \infty$.

We will use periodic boundary conditions, so if an electron leaves the box from one side, it comes in from the other. This is convenient as then the model admits plane waves solutions. We could use hard boundaries, and the description should agree in the $L \rightarrow \infty$ limit, but this set of boundary conditions is harder to work with; we would have cosine and sines instead of plane waves.

The plane-wave basis will be a natural choice given the periodic BCs. Explicitly, we can write this as:

$$\psi_{k,\lambda}(\mathbf{x}) = \frac{1}{\sqrt{V}} e^{i\mathbf{k} \cdot \mathbf{x}} \eta_{\lambda} \quad (3.1)$$

with $\lambda = (\uparrow, \downarrow)$ is the spin index. We have the spinors:

$$\eta_{\uparrow} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \eta_{\downarrow} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad (3.2)$$

The momentum is given by:

$$\mathbf{k} = (k_x, k_y, k_z), k_i = \frac{2\pi}{L} n_i \quad (3.3)$$

where $n_i \in \mathbb{Z}$. The Hamiltonian is given by:

$$\begin{aligned} H &= H_{el} + H_b + H_{el-b} \\ H_{el} &= \sum_{i=1}^N \frac{p_i^2}{2m} + \frac{1}{2} e^2 \sum_{i \neq j} \frac{e^{-\mu|\mathbf{r}_i - \mathbf{r}_j|}}{|\mathbf{r}_i - \mathbf{r}_j|} \\ H_b &= \frac{1}{2} e^2 \int d^3x d^3x' \frac{n(\mathbf{x}) n(\mathbf{x}') e^{-\mu|\mathbf{x} - \mathbf{x}'|}}{|\mathbf{x} - \mathbf{x}'|} \\ H_{el-b} &= -e \sum_{i=1}^N \int d^3x \frac{n(\mathbf{x}) e^{-\mu|\mathbf{x} - \mathbf{r}_i|}}{|\mathbf{x} - \mathbf{r}_i|}. \end{aligned} \quad (3.4)$$

H_{el} is just the kinetic energy of the electrons and the point-charge on point-charge interactions. H_b is the electrostatic interaction of the background field with itself. And H_{el-b} is the electrostatic interactions of

the electrons with the background field. N is the number of electrons, $V = L^3$ is the volume, $n = N/V$ is the electron density, and μ is a convergence factor which we send $\mu \rightarrow 0$ ³.

3.2 Simplifying the background terms

We want to rewrite this in second quantization notation. Only H_e will have nontrivial structure in the second quantization notation, but nevertheless the other terms are necessary for the stability of the system.

Let us start with the second term, which is the simplest. We deal with a uniform background density, namely $n(\mathbf{x}) = n = N/V$. H_b then becomes a simple integral:

$$\begin{aligned} H_b &= \frac{1}{2}e^2 \left(\frac{N}{V}\right)^2 \int d^3x \int d^3x' \frac{e^{-\mu|\mathbf{x}-\mathbf{x}'|}}{|\mathbf{x}-\mathbf{x}'|} \\ &= \frac{1}{2}e^2 \left(\frac{N}{V}\right)^2 \int d^3x \int d^3z \frac{e^{-\mu z}}{z} \\ &= \frac{1}{2}e^2 \left(\frac{N^2}{V}\right) \frac{4\pi}{\mu^2} \end{aligned} \tag{3.5}$$

Where in the second equality we use $z = \mathbf{x}' - \mathbf{x}$ in order to make the integrals independent, and evaluate the integrals in the third equality (the inner integral evaluating to $\frac{4\pi}{\mu^2}$, the outer to V). We can see why it was useful to introduce the $e^{-\mu}$; the integral would have diverged otherwise due to the long-range nature of the Coloumb interaction. Note also that we performed the integral assuming $\mu^{-1} \ll L$. We can similarly calculate H_{el-b} :

$$\begin{aligned} H_{el-b} &= -e \frac{N}{V} \sum_{i=1}^N \int d^3x \frac{e^{-\mu|\mathbf{x}-\mathbf{r}_i|}}{|\mathbf{x}-\mathbf{r}_i|} \\ &= -e \frac{N}{V} \sum_{i=1}^N \int d^3z \frac{e^{-\mu z}}{z} \\ &= -e^2 \frac{N}{V} N \int d^3z \frac{e^{-\mu z}}{z} \\ &= -e^2 \frac{N^2}{V} \frac{4\pi}{\mu^2} \end{aligned} \tag{3.6}$$

We note the partial cancellation of the H_b and the H_{el-b} terms of the Hamiltonian; we will see another cancellation later.

A reasonable question is why the $\mu^{-1} \ll L$ assumption is necessary. It boils down to the the fact that we have periodic boundary conditions, and we do *not* want the electric field of a given electron to interact with itself (or at least, not in a way that is exponentially insignificant and hence ignorable). See Fig. 3.2 below for a visual demonstration of the importance of this assumption.

With these simplifications,

$$H = -\frac{1}{2}e^2 \frac{N^2}{V} \frac{4\pi}{\mu^2} + H_{el} \tag{3.7}$$

Note that all of the interesting physics is contained in H_{el} , but we need $H_b + H_{el}$ to get a finite theory; we can see that if we took $\mu \rightarrow 0$ now, the energy would diverge; presumably there will be some sort of cancellation that occurs with H_{el} that will regularize the theory.

³In nuclear physics this has significance as the *Yukawa potential*; here we just use it as a convenient trick to remove some diverging integrals

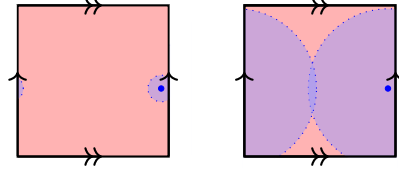


Figure 3.2: Comparisons of “spheres of interaction” of electrons when $\mu^{-1} \ll L$ (left) and $\mu^{-1} \sim L$ (right). We can see that in the former case, the electron does not interact with itself through the periodic boundary condition (beyond a negligible exponentially small contribution), and so the system is physically sound. In the latter case, the electron does have a nontrivial interaction with its own electric field, as seen through the overlap of the sphere of interaction; this is not physical. Hence in our analysis of the Jellium model, we make the assumption that $\mu^{-1} \ll L$.

3.3 Second Quantization of the Electron Term

Let us now transform the H_{el} term into second quantized notation.

(i) We start with the kinetic energy term:

$$\begin{aligned} \langle \mathbf{k}_1 \lambda_1 | T | \mathbf{k}_2 \lambda_2 \rangle &= \frac{1}{2mV} \int d^3x e^{-i\mathbf{k} \cdot \mathbf{x}} \eta_{\lambda_1}^\dagger (-\hbar^2 \nabla^2) e^{i\mathbf{k}_2 \cdot \mathbf{x}} \eta_{\lambda_2} \\ &= \frac{\hbar^2 k_2^2}{2mV} \delta_{\lambda_1 \lambda_2} \int d^3x e^{-i\mathbf{x} \cdot (\mathbf{k}_1 - \mathbf{k}_2)} \\ &= \frac{\hbar^2 k_2^2}{2m} \delta_{\lambda_1 \lambda_2} \delta_{\mathbf{k}_1 \mathbf{k}_2} \end{aligned} \quad (3.8)$$

where we use that $\int d^3x e^{-i\mathbf{x} \cdot (\mathbf{k}_1 - \mathbf{k}_2)} = V \delta_{\mathbf{k}_1 \mathbf{k}_2}$. We therefore obtain:

$$\hat{T} = \sum_{\mathbf{k}, \lambda} \frac{\hbar^2 k^2}{2m} c_{\mathbf{k}\lambda}^\dagger c_{\mathbf{k}\lambda}. \quad (3.9)$$

(ii) We now look at the potential term:

$$\langle k_1 \lambda_1 k_2 \lambda_2 | V | k_3 \lambda_3 k_4 \lambda_4 \rangle = \frac{e^2}{V} \delta_{\lambda_1 \lambda_3} \delta_{\lambda_2 \lambda_4} \delta_{\mathbf{k}_1 + \mathbf{k}_2, \mathbf{k}_3 + \mathbf{k}_4} \frac{4\pi}{(\mathbf{k}_1 - \mathbf{k}_3)^2 + \mu}. \quad (3.10)$$

See F&W for details; there is nothing conceptually new in the calculation above, it is only slightly more annoying at there are four plane wave terms.

We therefore obtain:

$$\hat{H} = \hat{T} - \frac{1}{2} \frac{e^2 N^2}{V} \frac{4\pi}{\mu} + \frac{e^2}{2V} \sum_{\mathbf{k}, \lambda} \delta_{\lambda_1 \lambda_3} \delta_{\lambda_2 \lambda_4} \delta_{\mathbf{k}_1 + \mathbf{k}_2, \mathbf{k}_3 + \mathbf{k}_4} \frac{4\pi}{(\mathbf{k}_1 - \mathbf{k}_3)^2 + \mu} c_{\mathbf{k}_1 \lambda_1}^\dagger c_{\mathbf{k}_2 \lambda_2}^\dagger c_{\mathbf{k}_4 \lambda_4} c_{\mathbf{k}_3 \lambda_3}. \quad (3.11)$$

Now we have to think a little bit; we have three delta functions, two for spin, one for momenta. We can explicitly two summations over λ and one over momentum. Instead of doing so blindly, we will find it useful to make the following change of variables:

$$\begin{pmatrix} \mathbf{k}_1 = \mathbf{k} + \mathbf{q} & \mathbf{k}_3 = \mathbf{k} \\ \mathbf{k}_2 = \mathbf{p} - \mathbf{q} & \mathbf{k}_4 = \mathbf{p} \end{pmatrix} \quad \begin{pmatrix} \lambda_1 = \alpha \\ \lambda_2 = \beta \end{pmatrix} \quad (3.12)$$

We may notice that we express the four momenta in terms of three, but this is ok; we have the extra constraint on the momentum already, and it was designed to satisfy this constraint. Substituting, the potential term becomes:

$$\frac{e^2}{2V} \sum_{\mathbf{k}\mathbf{p}\mathbf{q}} \sum_{\alpha\beta} \frac{4\pi}{\mathbf{q}^2 + \mu^2} c_{\mathbf{k}+\mathbf{q}\alpha}^\dagger c_{\mathbf{p}-\mathbf{q}\beta}^\dagger c_{\mathbf{p}\beta} c_{\mathbf{k}\alpha}. \quad (3.13)$$

We now want to send $\mu \rightarrow 0$. Note that we can do this for any term in the sum for which $\mathbf{q} \neq \mathbf{0}$. The only singular term is $\mathbf{q} = \mathbf{0}$, so let us study that term:

$$\begin{aligned} \frac{e^2}{2V} \sum_{\mathbf{k}\mathbf{p}} \sum_{\alpha\beta} \frac{4\pi}{\mu^2} c_{\mathbf{k}\alpha}^\dagger c_{\mathbf{p}\beta}^\dagger c_{\mathbf{p}\beta} c_{\mathbf{k}\alpha} &= \frac{e^2}{2V} \sum_{\mathbf{k}\mathbf{p}} \sum_{\alpha\beta} \frac{4\pi}{\mu^2} c_{\mathbf{k}\alpha}^\dagger (c_{\mathbf{k}\alpha} c_{\mathbf{p}\beta}^\dagger - \delta_{\mathbf{k}\mathbf{p}} \delta_{\alpha\beta}) c_{\mathbf{p}\beta} \\ &= \frac{e^2}{2V} \frac{4\pi}{\mu^2} (\hat{N}^2 - \hat{N}) \\ &= \frac{e^2}{2} \frac{N^2}{V} \frac{4\pi}{\mu^2} - \frac{e^2}{2} \frac{N}{V} \frac{4\pi}{\mu^2} \end{aligned} \quad (3.14)$$

where in the first equality we have commuted the $c_{\mathbf{k}\alpha}$ between the two c^\dagger s (being careful to respect the commutation relations), in the second equality we have used the definition of the number operator, and in the third equality we used the fact that we work with a system with a finite, fixed number of electrons and hence we can replace the operators \hat{N} with the number of particles N .

We see that the first term in Eq. (3.14) and the second term in Eq. (3.11) cancel. We argue that the second term in Eq. (3.14) is vanishingly small in the thermodynamic limit. We argue this as follows; since N/V (the number density) is constant with the system size, the term is constant with the system size; however $\langle H \rangle$ is extensive, and scales with the system size ($\langle H \rangle \sim V \sim N$). Hence we may choose to ignore it. So let us conclude by stating our final Hamiltonian:

$$\hat{H} = \sum_{\mathbf{k}\alpha} \frac{\hbar^2 \mathbf{k}^2}{2m} c_{\mathbf{k}\alpha}^\dagger c_{\mathbf{k}\alpha} + \frac{e^2}{2V} \sum'_{\mathbf{k}\mathbf{p}\mathbf{q}} \sum_{\alpha\beta} \frac{4\pi}{\mathbf{q}^2} c_{\mathbf{k}+\mathbf{q}\alpha}^\dagger c_{\mathbf{p}-\mathbf{q}\beta}^\dagger c_{\mathbf{p}\beta} c_{\mathbf{k}\alpha} \quad (3.15)$$

where the prime on the summation denotes that we do not include the $\mathbf{q} = \mathbf{0}$ term.

3.4 Rescaling the Hamiltonian

It is possible to gain important insights by introducing “natural” dimensionless variables. We define the inter-electron spacing r_0 as the radius of the sphere corresponding to the volume per electron. We then define the dimensionless quantity r_s as the ratio of r_0 with the Bohr radius a_0 :

$$\begin{aligned} \frac{V}{N} &= \frac{4}{3} \pi r_0^3 \\ a_0 &= \frac{\hbar^2}{me^2} \\ \boxed{r_s = \frac{r_0}{a_0} \approx 2 - 6 \text{ for metals}} \end{aligned} \quad (3.16)$$

The above are very good things to remember; there will be questions about them on the midterm and final! Based on these, let us define:

$$\bar{V} = V/r_0^3, \bar{\mathbf{k}} = \mathbf{k}r_0 \quad (3.17)$$

So our rescaled Hamiltonian can be written as:

$$\hat{H} = \frac{e^2}{a_0 r_s^2} \left(\sum_{\bar{\mathbf{k}}\alpha} \frac{\bar{\mathbf{k}}^2}{2} c_{\bar{\mathbf{k}}\alpha}^\dagger c_{\bar{\mathbf{k}}\alpha} + \frac{e^2}{2\bar{V}} \sum'_{\bar{\mathbf{k}}\bar{\mathbf{p}}\bar{\mathbf{q}}} \sum_{\alpha\beta} \frac{4\pi}{\bar{\mathbf{q}}^2} c_{\bar{\mathbf{k}}+\bar{\mathbf{q}}\alpha}^\dagger c_{\bar{\mathbf{p}}-\bar{\mathbf{q}}\beta}^\dagger c_{\bar{\mathbf{p}}\beta} c_{\bar{\mathbf{k}}\alpha} \right) \quad (3.18)$$

where $\frac{e^2}{a_0} \approx 13.6\text{eV}$ is the Rydberg constant/hydrogen binding energy. This result shows that in the $r_s \rightarrow 0$ (high density) limit, the electron-electron interaction becomes weak. This is very counterintuitive; in classical physics, the electron-electron interaction would dominate! Another thing to note; starting from the high-density limit, we can solve an easy problem (one that just contains the kinetic energy of the electrons) and then treat the electron-electron interactions as a small perturbation (can be reasonably treated in perturbation theory, expanding in powers of r_s). The actual series for the ground-state energy reads:

$$E_G = \frac{Ne^2}{a_0 r_s^2} \left(a + br_s + cr_s^2 \log(r_s) + dr_s^2 + \dots \right) \quad (3.19)$$

the $\log(r_s)$ term is perhaps a bit peculiar, but indeed if we do the perturbation expansion diligently we can confirm that it shows up. In the following, we will find a and b . We also remark that c may be similarly obtained, but d and higher powers require more advanced techniques, namely Green's function techniques⁴. We now proceed with the perturbation theory.

3.5 Perturbation Theory (High Density)

We return to our non-rescaled Hamiltonian so as to avoid having to write bars all the time. We split the Hamiltonian into two parts (the kinetic energy term and the perturbing electron-electron term):

$$\begin{aligned} \hat{H}_0 &= \sum_{\mathbf{k}\alpha} \frac{\hbar^2 \mathbf{k}^2}{2m} c_{\mathbf{k}\alpha}^\dagger c_{\mathbf{k}\alpha} \\ \hat{H}_1 &= \frac{e^2}{2V} \sum'_{\mathbf{k}\mathbf{p}\mathbf{q}} \sum_{\alpha\beta} \frac{4\pi}{q^2} c_{\mathbf{k}+\mathbf{q}\alpha}^\dagger c_{\mathbf{p}-\mathbf{q}\beta}^\dagger c_{\mathbf{p}\beta} c_{\mathbf{k}\alpha} \end{aligned} \quad (3.20)$$

3.5.1 Zeroth Order

The ground state of \hat{H}_0 can be written as:

$$|F\rangle = \prod_{|\mathbf{k}| < k_F} c_{\mathbf{k}\uparrow}^\dagger c_{\mathbf{k}\downarrow}^\dagger |0\rangle. \quad (3.21)$$

For N electrons, the Fermi momentum is determined by:

$$N = \langle F | \hat{N} | F \rangle = \sum_{\mathbf{k}\lambda} \langle F | n_{\mathbf{k}\lambda} | F \rangle = \sum_{\mathbf{k}\lambda} \theta(k_F - |\mathbf{k}|) = 2V \int \frac{d^3k}{(2\pi)^3} \theta(k_F - |\mathbf{k}|) = \frac{2V}{(2\pi)^3} \left(\frac{4}{3} \pi k_F^3 \right) = \frac{V}{3\pi^2} k_F^3 \quad (3.22)$$

We have therefore obtained k_F defined by electron density. In the above we have used the standard prescription:

$$\frac{1}{V} \sum_{\mathbf{k}} \rightarrow \int \frac{d^3k}{(2\pi)^3} \quad (3.23)$$

and the step function:

$$\theta(x) = \begin{cases} 1 & x > 0 \\ 0 & x < 0. \end{cases}$$

We can now solve for k_F :

$$k_F = \left(\frac{3\pi^2 N}{V} \right)^{1/3} = \left(\frac{9\pi}{4} \right)^{1/3} r_0^{-1} \approx 1.92 r_0^{-1} \quad (3.24)$$

⁴Green's functions give us a way to do a more formalized version of perturbation theory. With them, the expansion has been computed to seventh order; but the computation becomes much more difficult as we add higher order terms.

which tells us that k_F is (up to a factor of order unity) equal to the inverse of r_0 . One can use the above to relate conduction bands and k_F to the lattice spacing, among other useful applications. So now doing the ground state energy calculation, we have:

$$\begin{aligned}
E^{(0)} = \langle F | \hat{H}_0 | F \rangle &= \frac{\hbar^2}{2m} \sum_{\mathbf{k}, \alpha} \langle F | n_{\mathbf{k}\alpha} | F \rangle \mathbf{k}^2 = \frac{\hbar^2}{2m} \sum_{\mathbf{k}, \alpha} \mathbf{k}^2 \theta(k_F - |\mathbf{k}|) = \frac{\hbar^2}{2m} 2 \frac{V}{(2\pi)^3} \int d^3k \mathbf{k}^2 \theta(k_F - \mathbf{k}) \\
&= \frac{3}{5} \frac{\hbar^2 k_F^2}{2m} N \\
&= \frac{3}{5} \epsilon_F N \\
&= \left(\frac{e^2}{2a_0} \right) N \frac{2.21}{r_s^2}
\end{aligned} \tag{3.25}$$

where in the fourth equality the factor of 2 comes from the summation over spin, and the integral is performed by going into spherical coordinates. Again $\frac{e^2}{2a_0} = 13.6\text{eV}$ is the Rydberg constant. This tells us that for metals (i.e. $r_s \approx 2 - 6$), the energy of electrons is on the order of eV.

3.5.2 First Order

To first order, we simply calculate the expectation value:

$$E^{(1)} = \langle F | \hat{H}_1 | F \rangle = \frac{e^2}{2V} \sum_{\mathbf{k}\mathbf{p}\mathbf{q}} \sum_{\alpha\beta} \langle F | c_{\mathbf{k}+\mathbf{q}\alpha}^\dagger c_{\mathbf{p}-\mathbf{q}\beta}^\dagger c_{\mathbf{p}\beta} c_{\mathbf{k}\alpha} | F \rangle. \tag{3.26}$$

The above is more complicated; we must deal with the four-fermion operator acting on the ground state. We can do this the hard way; we can substitute in the ground state and use commutation relations amongst the operators. Or, we can give a slick argument that does the same job but with less writing; let's do it this way. We start with a Fermi sphere filled with electrons. The annihilation operators can only remove two electrons from inside the spheres, and the creation operators can either create electrons where they are removed, or create them elsewhere. If they are created elsewhere, the state we end up with will be orthogonal to the ground state $|F\rangle$. So, the only nonvanishing contributions will come from the terms where the annihilation operators remove electrons from inside the sphere, and the creation operators fill back in the holes (giving us back $|F\rangle$, up to some prefactor).

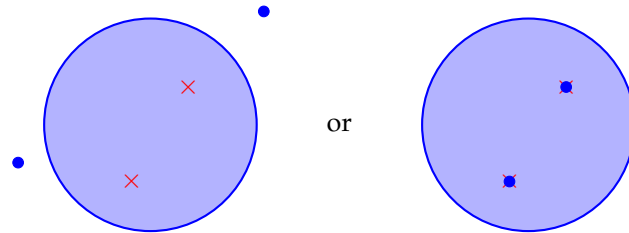


Figure 3.3: Action of the 4 fermion operators on the ground state $|F\rangle$. The two annihilation operators remove some pair of fermions within the Fermi sphere. The creation operators can then create fermions elsewhere (left), in which case the final state is orthogonal to $|F\rangle$ and does not contribute to the expectation value. Alternatively, the creation operators can fill the holes created by the annihilation operators (right), in which case the final state is proportional to $|F\rangle$ and hence does contribute to the expectation value.

There are two pairings of the creation/annihilation operators for which the above can occur. The first

possibility is:

$$\begin{pmatrix} \mathbf{k} + \mathbf{q}, \alpha = \mathbf{k}, \alpha \\ \mathbf{p} - \mathbf{q}, \beta = \mathbf{p}, \beta \end{pmatrix} \quad \text{or} \quad \begin{pmatrix} \mathbf{k} + \mathbf{q}, \alpha = \mathbf{p}, \beta \\ \mathbf{p} - \mathbf{q}, \beta = \mathbf{k}, \alpha \end{pmatrix} \quad (3.27)$$

If we look at the first possibility, we immediately obtain the constraint that $\mathbf{q} = 0$. However, we have already removed all such terms in our sum (note the prime); so the only the second possibility contributes. We call this the “exchange term”, because the spins are exchanged. The conclusion of this argument is that the terms of the sum are only nonzero when the exchange conditions of $\mathbf{k} + \mathbf{q} = \mathbf{p}, \alpha = \beta$ are satisfied. We can now use this to compute the first order correction to the GS energy:

$$\begin{aligned} E^{(1)} &= \frac{e^2}{2V} \sum'_{\mathbf{k}\mathbf{p}\mathbf{q}} \sum_{\alpha\beta} \delta_{\mathbf{k}+\mathbf{q},\mathbf{p}} \delta_{\alpha\beta} \langle F | c_{\mathbf{k}+\mathbf{q}\alpha}^\dagger c_{\mathbf{p}-\mathbf{q}\beta}^\dagger c_{\mathbf{p}\beta} c_{\mathbf{k}\alpha} | F \rangle \frac{4\pi}{\mathbf{q}^2} \\ &= \frac{e^2}{2V} \sum'_{\mathbf{k}\mathbf{p}\mathbf{q}} \sum_{\alpha\beta} \delta_{\mathbf{k}+\mathbf{q},\mathbf{p}} \delta_{\alpha\beta} \langle F | c_{\mathbf{k}+\mathbf{q}\alpha}^\dagger c_{\mathbf{k}\alpha}^\dagger c_{\mathbf{k}+\mathbf{q}\alpha} c_{\mathbf{k}\alpha} | F \rangle \frac{4\pi}{\mathbf{q}^2} \\ &= -\frac{e^2}{2V} \sum'_{\mathbf{k}\mathbf{q}} \sum_{\alpha} \langle F | \hat{n}_{\mathbf{k}+\mathbf{q}\alpha} \hat{n}_{\mathbf{k}\alpha} | F \rangle \frac{4\pi}{\mathbf{q}^2} \\ &= -\frac{e^2}{2V} \sum'_{\mathbf{k}\mathbf{q}} \sum_{\alpha} \frac{4\pi}{\mathbf{q}^2} \theta(k_F - |\mathbf{k} + \mathbf{q}|) \theta(k_F - k) \\ &= -\frac{e^2}{2} \frac{4\pi V}{(2\pi)^6} \int d^3k \int d^3q \frac{1}{\mathbf{q}^2} \theta(k_F - |\mathbf{k} + \mathbf{q}|) \theta(k_F - k) \end{aligned} \quad (3.28)$$

We make a change of variables $\mathbf{k} \rightarrow \mathbf{p} = \mathbf{k} + \frac{1}{2}\mathbf{q}$. This yields:

$$E^{(1)} = -\frac{4\pi e^2 V}{(2\pi)^6} \int d^3q \frac{1}{\mathbf{q}^2} \int d^3p \theta(k_F - |\mathbf{p} + \frac{1}{2}\mathbf{q}|) \theta(k_F - |\mathbf{p} - \frac{1}{2}\mathbf{q}|) \quad (3.29)$$

This integral is much more symmetric; if we think about this geometrically, we are looking at the overlap of spheres when we calculate the inner integral over q :

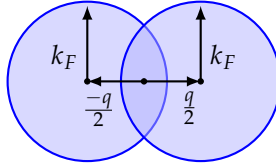


Figure 3.4: Visualization of the integral $I(q) = \int d^3p \theta(k_F - |\mathbf{p} + \frac{1}{2}\mathbf{q}|) \theta(k_F - |\mathbf{p} - \frac{1}{2}\mathbf{q}|)$. The two step functions correspond to the two spheres shown above, and the overall integral calculates their overlap in volume.

The integral is a standard exercise in multivariable calculus. One finds:

$$\int d^3p \theta(k_F - |\mathbf{p} + \frac{1}{2}\mathbf{q}|) \theta(k_F - |\mathbf{p} - \frac{1}{2}\mathbf{q}|) = \frac{4\pi}{3} k_F^3 \left(1 - \frac{3}{2}x + \frac{1}{2}x^2 \right) \theta(1-x), \quad x = \frac{q}{2k_F}. \quad (3.30)$$

where the θ function is there to account for the fact that when q is sufficiently large the spheres do not touch. We now calculate the outer integral. We make the observation that in spherical coordinates, $d^3q \rightarrow 4\pi q^2 dq$ and so the q^2 in the denominator cancels. All we have to do is just an integral of a polynomial; child’s play. We are left with:

$$E^{(1)} = -\frac{e^2}{2a_0} \frac{N}{r_s} \left(\frac{9\pi}{4} \right)^{1/2} \frac{3}{2\pi} = -\frac{e^2}{2a_0} N \frac{0.916}{r_s}. \quad (3.31)$$

3.5.3 Combining Results

In summary, the total ground state energy (to first order) is:

$$\frac{E}{N} = \frac{e^2}{2a_0} \frac{1}{r_s^2} (2.21 - 0.916r_s + \dots) \quad (3.32)$$

We identify the first term as the free Fermi gas energy, and the second term as the exchange energy.

A comment: We have done perturbation theory in r_s which we have treated as “small”, but really r_s is 2 – 6 for metals so it is not really small. It surprisingly works fairly well, anyway.

3.6 The Variational Viewpoint

We now switch perspectives a little bit, and find that we can learn something more interesting about the calculation we just did. We viewed it as a perturbation theory expansion in r_s . But we can view it instead as a variational calculation.

Recall the variational principle:

$$E_{GS} \leq \langle \psi | H | \psi \rangle$$

where H is a Hamiltonian, $|\psi\rangle$ is any state, and E_{GS} is the ground state energy of H . In our case, we have calculated $\langle F | \hat{H}_0 + \hat{H}_1 | F \rangle$, which can also be viewed as a variational energy parametrized by electron density r_s . Taking our result from Eq. (3.32), we can plot an energy landscape as a function of r_s :

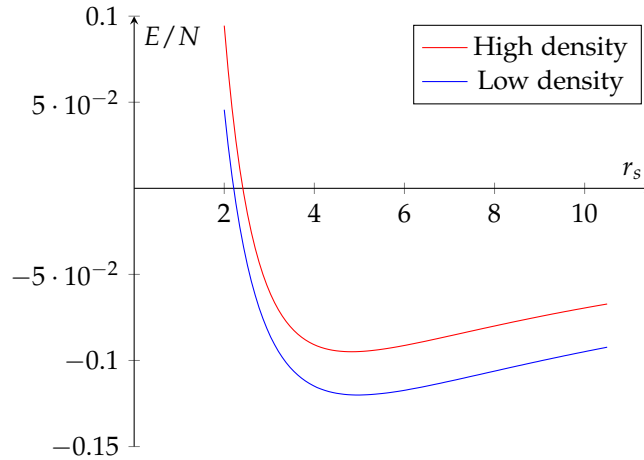


Figure 3.5: Plot of the variational energy landscape for. E/N is in units of $e^2/2a_0$, and we plot in a range of typical r_s for metals. In red we plot the first-order perturbation expansion for E/N in the high-density limit (Eq. (3.32)). In blue we plot the first-order perturbation expansion for E/N in the low-density limit (Eq. (3.33)). We can find the r_s that minimizes E/N to approximate the true ground state energy (i.e. the binding energy per electron in metals).

We find that the r_s that minimizes this energy is $(r_s)_{\min} = 4.83$ and $E_{\min}/N = -0.095 \frac{e^2}{2a_0} \approx -1.29\text{eV}$. For comparison, the binding energy per electron in sodium (found experimentally) is $r_s = 3.86$ and $E/N = -1.13\text{eV}$. Even this very simple calculation gets us the correct order of magnitude.

3.7 Perturbation Theory (Low Density)

It turns out that we can do perturbation theory also in the large r_s /low-density limit, where we take:

$$\hat{H}_0 = \frac{e^2}{2V} \sum_{\mathbf{k}\mathbf{p}\mathbf{q}}' \sum_{\alpha\beta} \frac{4\pi}{\mathbf{q}^2} c_{\mathbf{k}+\mathbf{q}\alpha}^\dagger c_{\mathbf{p}-\mathbf{q}\beta}^\dagger c_{\mathbf{p}\beta} c_{\mathbf{k}\alpha}$$

$$\hat{H}_1 = \sum_{\mathbf{k}\alpha} \frac{\hbar^2 \mathbf{k}^2}{2m} c_{\mathbf{k}\alpha}^\dagger c_{\mathbf{k}\alpha}$$

i.e. we exchange which is the dominant and which is the perturbing Hamiltonian. In this limit, we find (though the calculation is more difficult):

$$\frac{E}{N} = \frac{e^2}{2a_0} \frac{1}{r_s} \left(-1.79 + \frac{2.66}{\sqrt{r_s}} + \dots \right) \quad (3.33)$$

This is plotted as the dashed line in the figure above. This is the controversial “Wigner crystal”; it is not known if this actually applied, as the crystallization of electrons has never been observed in three dimensions.

Next class, we will look at the Hartree-Fock approximation.

5 The Hartree-Fock Approximation

Note: Section 4 is not missing, but rather Degenerate Electron Gases spanned two lectures.

5.1 Motivation and Main Idea

The Hartree-Fock, or mean-field approximation is useful when we have an intractable 4-fermion term in the Hamiltonian, for example the Coulomb interaction term.

The idea is quite simple. We approximate the full Hamiltonian $H = H_0 + H_1$ by the “best possible” approximation of the form:

$$H_{HF} = \sum_{\mathbf{k}, \alpha} \epsilon_{\alpha}^{MF}(\mathbf{k}) c_{\mathbf{k}\alpha}^{\dagger} c_{\mathbf{k}\alpha}. \quad (5.1)$$

This looks like H_0 from before, but the modified dispersion relation allows for a tractable calculation.

5.2 Heuristic Approach

We first go for an intuitive approach, and then proceed with an approach that casts the problem in a variational setting (placing it on better footing). We consider:

$$\begin{aligned} H_0 &= \sum_{\mathbf{k}\alpha} \epsilon_0(\mathbf{k}) c_{\mathbf{k}\alpha}^{\dagger} c_{\mathbf{k}\alpha} \\ H_1 &= \frac{1}{2} \sum_{\mathbf{k}\mathbf{p}\mathbf{q}\alpha\beta} V(\mathbf{q}) c_{\mathbf{k}+\mathbf{q}\alpha}^{\dagger} c_{\mathbf{p}-\mathbf{q}\beta}^{\dagger} c_{\mathbf{p}\beta} c_{\mathbf{k}\alpha} \end{aligned} \quad (5.2)$$

Next, we “decouple” H_1 using the operator identity:

$$AB = A \langle B \rangle + \langle A \rangle B - \langle A \rangle \langle B \rangle + (A - \langle A \rangle)(B - \langle B \rangle) \quad (5.3)$$

This is useful as if we just focus on the last term, this is just the fluctuation of operator A around its mean field value time the fluctuation of B around its mean field value. The idea of mean-field theory is to say that the product of (small) fluctuations is small, and we therefore can neglect it. The central idea of the mean-field approach is to replace such products with the first three terms in the above operator identity.

Remark: in some sense, we assume that the fluctuations are small before we know what the solutions are, so we do not necessarily a priori know this will be the case. However, in practice this approach is greatly successful for the majority of problems in CM physics (the remaining 1% of problems for which it fails tend to be interesting research problems).

Looking at our interaction term H_1 , we have two ways to pair up our operators ($A = 13$ and $B = 24$ or $A = 14$ and $B = 23$). Let’s go ahead with trying this simplification for our Hamiltonian.

$$\begin{aligned} H_1^{MF} &= \frac{1}{2} \sum_{\mathbf{k}\mathbf{p}\mathbf{q}\alpha\beta} V(\mathbf{q}) \left[c_{\mathbf{k}+\mathbf{q}\alpha}^{\dagger} c_{\mathbf{k}\alpha} \langle c_{\mathbf{p}-\mathbf{q}\beta}^{\dagger} c_{\mathbf{p}\beta} \rangle + \langle c_{\mathbf{k}+\mathbf{q}\alpha}^{\dagger} c_{\mathbf{k}\alpha} \rangle c_{\mathbf{p}-\mathbf{q}\beta}^{\dagger} c_{\mathbf{p}\beta} - \langle c_{\mathbf{k}+\mathbf{q}\alpha}^{\dagger} c_{\mathbf{k}\alpha} \rangle \langle c_{\mathbf{p}-\mathbf{q}\beta}^{\dagger} c_{\mathbf{p}\beta} \rangle \right. \\ &\quad \left. - c_{\mathbf{k}+\mathbf{q}\alpha}^{\dagger} c_{\mathbf{p}\beta} \langle c_{\mathbf{p}-\mathbf{q}\beta}^{\dagger} c_{\mathbf{k}\alpha} \rangle - + \langle c_{\mathbf{k}+\mathbf{q}\alpha}^{\dagger} c_{\mathbf{p}\beta} \rangle c_{\mathbf{p}-\mathbf{q}\beta}^{\dagger} c_{\mathbf{k}\alpha} + \langle c_{\mathbf{k}+\mathbf{q}\alpha}^{\dagger} c_{\mathbf{p}\beta} \rangle \langle c_{\mathbf{p}-\mathbf{q}\beta}^{\dagger} c_{\mathbf{k}\alpha} \rangle \right] \end{aligned} \quad (5.4)$$

Why do we include both pairings (the other pairings are dropped, but are zero in the Fermi sea)? Aren’t we overcounting? In fact we are not (and careful thought shows this is not the case); this will become clearer when we do this variationally.

This looks intimidating, but there are some nice simplifications; we have two terms that come out: for example, (s):

$$\begin{aligned} \langle c_{\mathbf{p}-\mathbf{q}\beta}^{\dagger} c_{\mathbf{p}\beta} \rangle &= \delta_{\mathbf{q}=0} \langle c_{\mathbf{p}\beta}^{\dagger} c_{\mathbf{p}\beta} \rangle \\ \langle c_{\mathbf{p}-\mathbf{q}\beta}^{\dagger} c_{\mathbf{k}\alpha} \rangle &= \delta_{\alpha\beta} \delta_{\mathbf{p}-\mathbf{q}=\mathbf{k}} \langle c_{\mathbf{p}\beta}^{\dagger} c_{\mathbf{p}\beta} \rangle \end{aligned} \quad (5.5)$$

where the delta shows up as the product vanishes unless we are looking at the same spin with the same momentum. The first line is the direct, or *Hartree* term. The second line is the exchange, or *Fock* term. Taking into account these simplifications, the mean field for the Hartree-Fock Hamiltonian takes the form:

$$H^{MF} = \sum_{\mathbf{k}\alpha} \left[\epsilon_0(\mathbf{k}) + V(0) \sum_{\mathbf{p}\beta} \langle c_{\mathbf{p}\beta}^\dagger c_{\mathbf{p}\beta} \rangle - \sum_{\mathbf{p}} V(\mathbf{p} - \mathbf{k}) \langle c_{\mathbf{p}\alpha}^\dagger c_{\mathbf{p}\alpha} \rangle \right] c_{\mathbf{k}\alpha}^\dagger c_{\mathbf{k}\alpha}. \quad (5.6)$$

So we have that:

$$\epsilon_\alpha^{MF}(\mathbf{k}) = \epsilon_0(\mathbf{k}) + V(0) \sum_{\mathbf{p}\beta} \langle c_{\mathbf{p}\beta}^\dagger c_{\mathbf{p}\beta} \rangle - \sum_{\mathbf{p}} V(\mathbf{p} - \mathbf{k}) \langle c_{\mathbf{p}\alpha}^\dagger c_{\mathbf{p}\alpha} \rangle = \epsilon_0(\mathbf{k}) + \epsilon_{dir} + \epsilon_{ex}(\mathbf{k})$$

where the second term is the Hartree term and the third term is the Fock term.

5.3 Applying Hartree-Fock to Coulomb

Let's work this out for the Coulomb interaction:

$$V(\mathbf{q}) = \frac{e^2}{V} \frac{4\pi}{q^2}$$

Note that there is no Hartree term as $q \neq 0$. We work out the exchange term:

$$\begin{aligned} -\epsilon_{ex}(\mathbf{k}) &= \frac{4\pi e^2}{V} \sum_{\mathbf{p}} \frac{\theta(k_F - \mathbf{p})}{(\mathbf{p} - \mathbf{k})^2} = 4\pi e^2 \int \frac{d^3 p}{(2\pi)^3} \frac{\theta(k_F - \mathbf{p})}{(\mathbf{p} - \mathbf{k})^2} \\ &= \frac{e^2}{2\pi^2} 2\pi \int_0^{k_F} \int_0^\pi d\theta \sin\theta \frac{1}{p^2 + k^2 - 2pk \cos\theta} \\ &= \frac{e^2}{2\pi} \int_0^{k_F} p^2 dp \int_{-1}^1 \frac{dz}{p^2 + k^2 - 2pkz} \\ &= \frac{e^2}{k\pi} \int_0^{k_F} dp p [\ln|\mathbf{k} - \mathbf{p}| - \ln|\mathbf{k} + \mathbf{p}|] \\ &= \frac{e^2 k_F}{\pi} \left[1 + \frac{k_F^2 - k^2}{2k_F k} \ln \left| \frac{k + k_F}{k - k_F} \right| \right] \end{aligned} \quad (5.7)$$

where we use the cosine rule in the second line, and make the substitution $\cos\theta = z$ in the third. So, we can write the final result as:

$$\begin{aligned} H^{MF} &= \sum_{\mathbf{k}\alpha} \epsilon(\mathbf{k}) c_{\mathbf{k}\alpha}^\dagger c_{\mathbf{k}\alpha} \\ \epsilon(\mathbf{k}) &= \epsilon_0(\mathbf{k}) - \frac{2e^2}{\pi} F\left(\frac{k}{k_F}\right) \\ F(x) &= \frac{1}{2} + \frac{1-x^2}{4x} \ln \left| \frac{1+x}{1-x} \right| \end{aligned} \quad (5.8)$$

F is plotted in Fig. 5.1.

We also recall the dispersion relation for free electrons:

$$\epsilon_0(\mathbf{k}) = \frac{\hbar^2 \mathbf{k}^2}{2m}.$$

We can plot the two dispersion relations side-by-side to compare them; this is done in Fig. 5.2.

Note that at $k = k_F$, there is an unphysical divergence in $\nabla_{\mathbf{k}} \epsilon(\mathbf{k})$. See HW1Q3 for more discussion of this problem, and how it can be remedied; namely, the source of the issue is the long-range Coulomb interaction. If we take into account screening, the problem goes away.

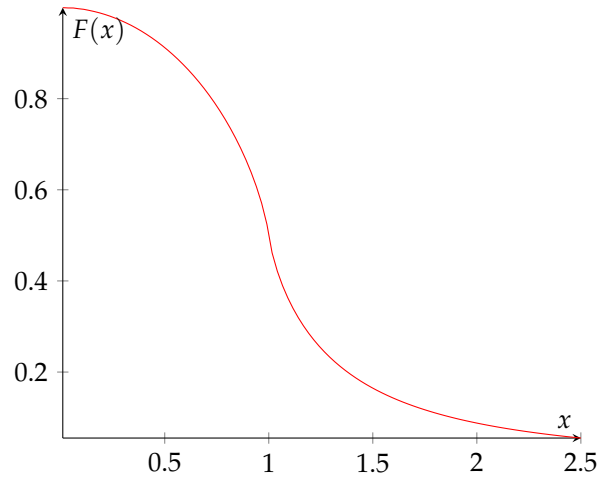


Figure 5.1: Plot of the function $F(x) = \frac{1}{2} + \frac{1-x^2}{4x} \ln \left| \frac{1+x}{1-x} \right|$ which appears in the mean-field Hamiltonian dispersion relation. Note the divergence of F' at $x = 1$ ($k = k_F$).

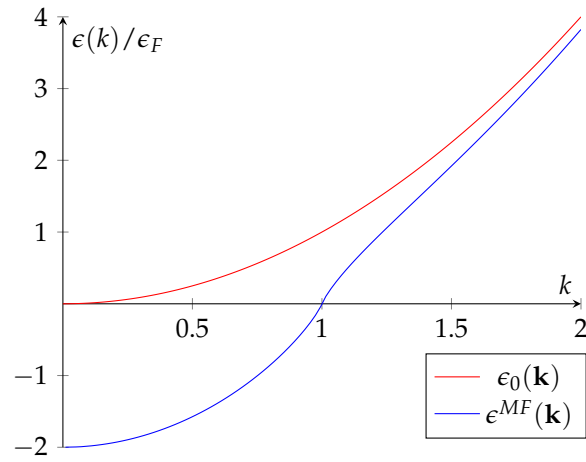


Figure 5.2: Plot of the free electron dispersion relation and the mean-field Hamiltonian dispersion relation. For the sake of plotting, we set $k_F = 1$.

5.4 Hartree-Fock as a Variational Bound

Let us try to find the best possible $\epsilon(\mathbf{k})$ through the variational principle. We will look for $\epsilon_0(\mathbf{k}) + \eta_{\alpha\mathbf{k}}$ with $\eta_{\alpha\mathbf{k}}$ the variational parameter such that $\langle H \rangle_{MF}$ field is minimized. This expectation value is with respect to the ground state of the mean-field Hamiltonian.

This seems strange; we normally fix a Hamiltonian and then pick states over which we conduct an energy minimization. Here, the parameters appear in the mean-field Hamiltonian, but then the MF hamiltonian is sufficiently simple such that we know what the ground state is. So, the variational parameters directly determine the ground state of H_{MF} which is the state we look at the expectation values for. Explicitly:

$$\langle H_0 \rangle_{MF} = \sum_{\mathbf{k}\alpha} \epsilon_0(\mathbf{k}) \langle c_{\mathbf{k}\alpha}^\dagger c_{\mathbf{k}\alpha} \rangle_{MF} = \sum_{\mathbf{k}\alpha} \epsilon_0(\mathbf{k}) n_{\mathbf{k}\alpha} \quad (5.9)$$

$$\begin{aligned} \langle H_1 \rangle_{MF} &= \frac{1}{2} \sum_{\mathbf{k}\mathbf{p}\mathbf{q}\alpha\beta} V(\mathbf{q}) \left[\langle c_{\mathbf{k}+\mathbf{q}\alpha}^\dagger c_{\mathbf{k}\alpha} \rangle_{MF} \langle c_{\mathbf{p}-\mathbf{q}\beta}^\dagger c_{\mathbf{p}\beta} \rangle_{MF} - \langle c_{\mathbf{k}+\mathbf{q}\alpha}^\dagger c_{\mathbf{p}\beta} \rangle_{MF} \langle c_{\mathbf{p}-\mathbf{q}\beta}^\dagger c_{\mathbf{k}\alpha} \rangle_{MF} \right] \\ &= \frac{1}{2} \sum_{\mathbf{k}\mathbf{p}\mathbf{q}\alpha\beta} V(\mathbf{q}) \left[\delta_{\mathbf{q}=0} n_{\mathbf{k}\alpha} n_{\mathbf{p}\alpha} - \delta_{\mathbf{k}+\mathbf{q}=\mathbf{p}} \delta_{\mathbf{p}-\mathbf{q}=\mathbf{k}} \delta_{\alpha\beta} n_{\mathbf{k}\alpha} n_{\mathbf{p}\beta} \right] \\ &= \frac{1}{2} \left[V(0) \left(\sum_{\mathbf{k}\alpha} n_{\mathbf{k}\alpha} \right)^2 - \sum_{\mathbf{k}\mathbf{p}\alpha} V(\mathbf{k}-\mathbf{p}) n_{\mathbf{k}\alpha} n_{\mathbf{p}\alpha} \right] \end{aligned} \quad (5.10)$$

We get these pairings via the same “destroy an electron in the Fermi sphere and then recover it” argument we covered when discussing the second quantized form of the Jellium model. We now wish to minimize these expectation values w.r.t. our variational parameter η . One must remember that η is implicitly hidden inside of the n s. We minimize via a chain rule:

$$0 = \frac{\partial \langle H \rangle_{MF}}{\partial \eta_{\mathbf{q}\lambda}} = \sum_{\mathbf{q}'\lambda'} \frac{\partial \langle H \rangle_{MF}}{\partial n_{\mathbf{q}'\lambda'}} \frac{\partial n_{\mathbf{q}'\lambda'}}{\partial \eta_{\mathbf{q}\lambda}} \quad (5.11)$$

and from this we obtain:

$$\sum_{\mathbf{k}\alpha} \left[\epsilon_0(\mathbf{k}) + V(0) \sum_{\mathbf{p}\beta} n_{\mathbf{p}\beta} - \sum_{\mathbf{p}} V(\mathbf{k}-\mathbf{p}) n_{\mathbf{p}\alpha} \right] \frac{\partial n_{\mathbf{q}'\lambda'}}{\partial \eta_{\mathbf{q}\lambda}} = 0 \quad (5.12)$$

To solve this, consider:

$$\frac{\partial \langle H_{MF} \rangle_{MF}}{\partial \eta_{\mathbf{q}\lambda}} = n_{\mathbf{q}\lambda} + \sum_{\mathbf{k}\alpha} [\epsilon_0(\mathbf{k}) + \eta_{\mathbf{k}\alpha}] \frac{\partial n_{\mathbf{q}'\lambda'}}{\partial \eta_{\mathbf{q}\lambda}} \quad (5.13)$$

So subtracting (5.13) from (5.12), we obtain:

$$\sum_{\mathbf{k}\alpha} \left[-\eta_{\mathbf{k}\alpha} + V(0) \sum_{\mathbf{p}\beta} n_{\mathbf{p}\beta} - \sum_{\mathbf{p}} V(\mathbf{k}-\mathbf{p}) n_{\mathbf{p}\alpha} \right] \frac{\partial n_{\mathbf{k}\alpha}}{\partial \eta_{\mathbf{q}\lambda}} = n_{\mathbf{q}\lambda} - \frac{\partial \langle H_{MF} \rangle_{MF}}{\partial \eta_{\mathbf{q}\lambda}} \quad (5.14)$$

The solution is:

$$\begin{aligned} \eta_{\mathbf{k}\alpha} &= V(0) \sum_{\mathbf{p}\beta} n_{\mathbf{p}\beta} - \sum_{\mathbf{p}} V(\mathbf{k}-\mathbf{p}) n_{\mathbf{p}\alpha} \\ n_{\mathbf{q}\lambda} &= \frac{\partial \langle H_{MF} \rangle_{MF}}{\partial \eta_{\mathbf{q}\lambda}} = \langle c_{\mathbf{q}\lambda}^\dagger c_{\mathbf{q}\lambda} \rangle_{MF} \end{aligned} \quad (5.15)$$

It is easy to lose sight of what is being done through the many lines of writing, but one notices with the result that we have obtained the heuristic result using a very different method!

One way of justifying why the two approaches coincide; if we assume that the Hamiltonian takes the simple variational form, then the product of fluctuations that we have neglected in the heuristic approach exactly vanish; the operator approximation becomes exact.

We will not use this result immediately, but when we discuss tight-binding models and superconductivity, these approaches will be exceedingly useful.

Next time, we will discuss screening; when a test charge is put into an electron gas, there will be no long range interactions as the Coulomb force is screened by the cloud of electrons. Please by Wednesday read p.337-339 of A&M; this is the "Screening (General)" section.

6 Screening

6.1 General Definitions and Setup of Problem

For definitions and general discussion, see Ashcroft and Mermin p.337-339.

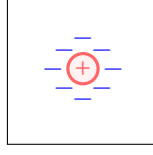


Figure 6.1: Cartoon of electron screening; when a positive test charge is placed in a cloud of negative charge, the negative charges are attracted to the positive charge. This has the net effect of screening out the positive charge, as it gets neutralized to some capacity.

If we place a (positive) test charge in an cloud of negative charge, then it attracts negative charges around it, which results in a net screening of the positive charge.

We have the potentials ϕ^{ext} , ϕ and the charge densities ρ^{ext} , ρ , $\rho^{ind} = \rho - \rho^{ext}$.

$$\begin{aligned}\phi(\mathbf{q}) &= \frac{1}{\epsilon(\mathbf{q})} \phi^{ext}(\mathbf{q}) \\ \rho^{ind}(\mathbf{q}) &= \chi(\mathbf{q}) \phi(\mathbf{q}) \\ \epsilon(\mathbf{q}) &= 1 - \frac{4\pi}{q^2} \chi(\mathbf{q})\end{aligned}\tag{6.1}$$

where $\epsilon(\mathbf{q})$ is the dielectric function and $\chi(\mathbf{q})$ is the dielectric susceptibility.

Today, we will go through two different calculations of the susceptibility.

6.2 Thomas-Fermi Theory

We have the second-quantized Hamiltonian:

$$\begin{aligned}\hat{H} &= \int d^3x \hat{\psi}^\dagger(\mathbf{x}) T(\mathbf{x}) \hat{\psi}(\mathbf{x}) \\ T(\mathbf{x}) &= \frac{\hbar^2 \nabla^2}{2m} - e\phi(\mathbf{x})\end{aligned}\tag{6.2}$$

we neglect $e - e$ interactions (including them makes this much more difficult). Of course electrons will see the total screened potential rather than the bare test charge potential, which this formula accounts for (ϕ vs. ϕ^{ext}).

For $\phi(\mathbf{x}) = \phi_0$ a constant, we can solve the problem exactly. The Hamiltonian becomes:

$$H = \sum_{\mathbf{k}} \left(\frac{\hbar^2 \mathbf{k}^2}{2m} - e\phi_0 \right) c_{\mathbf{k}}^\dagger c_{\mathbf{k}} = \sum_{\mathbf{k}} \epsilon_{\mathbf{k}} c_{\mathbf{k}}^\dagger c_{\mathbf{k}}.\tag{6.3}$$

Thomas-Fermi theory assumes that $\phi(\mathbf{x})$ varies slowly on the scale k_F^{-1} ; therefore in a semiclassical approximation, it is appropriate to write:

$$\epsilon_{\mathbf{k}} \rightarrow \frac{\hbar^2 \mathbf{k}^2}{2m} - e\phi_0\tag{6.4}$$

We will calculate the local electron density at \mathbf{r} . It will be given by:

$$\begin{aligned}
\rho(\mathbf{r}) &= -e \langle \hat{\psi}^\dagger(\mathbf{r}) \hat{\psi}(\mathbf{r}) \rangle \\
&= -2e \sum_{\mathbf{k}\mathbf{k}'} \langle \hat{\psi}_{\mathbf{k}}^\dagger(\mathbf{r}) \hat{\psi}_{\mathbf{k}'}(\mathbf{r}) c_{\mathbf{k}}^\dagger c_{\mathbf{k}'} \rangle \\
&= -2e \sum_{\mathbf{k}\mathbf{k}'} \hat{\psi}_{\mathbf{k}}^\dagger(\mathbf{r}) \hat{\psi}_{\mathbf{k}'}(\mathbf{r}) \langle c_{\mathbf{k}}^\dagger c_{\mathbf{k}'} \rangle \\
&= -2e \sum_{\mathbf{k}\mathbf{k}'} \hat{\psi}_{\mathbf{k}}^\dagger(\mathbf{r}) \hat{\psi}_{\mathbf{k}'}(\mathbf{r}) \delta_{\mathbf{k}\mathbf{k}'} f(\epsilon_{\mathbf{k}}(\mathbf{r})) \\
&= -2e \sum_{\mathbf{k}} |\psi_{\mathbf{k}}(\mathbf{r})|^2 f(\epsilon_{\mathbf{k}}(\mathbf{r}))
\end{aligned} \tag{6.5}$$

where the factor of 2 comes from counting the two possible spin states, $f(\epsilon_{\mathbf{k}}(\mathbf{r}))$ is the Fermi-Dirac distribution (which comes from the fact that at any temperature, the expectation value of the number of fermions of any state at a given energy is given from stat mech to be the FD distribution; previously we denoted this as $n_{\mathbf{k}}$. For bosons next week it will be the BE-distribution), and in the last equality we use the Kronecker delta to perform the last of the summations. For the plane wave basis, we have $\psi_{\mathbf{k}}(\mathbf{r}) = \frac{1}{\sqrt{V}} e^{-i\mathbf{k}\cdot\mathbf{r}}$, so:

$$\begin{aligned}
\rho(\mathbf{r}) &= -\frac{2e}{V} \sum_{\mathbf{k}} \frac{1}{e^{\beta(\epsilon_{\mathbf{k}}(\mathbf{r})-\mu)+1} + 1} \\
&= -en_0(\mu + e\phi(\mathbf{r}))
\end{aligned} \tag{6.6}$$

where $\beta = \frac{1}{k_B T}$ as usual, and we define:

$$n_0(\mu) = \int \frac{d^3k}{4\pi^3} \frac{1}{e^{\beta(\frac{\hbar^2 \mathbf{k}^2}{2m} - \mu) + 1} + 1}. \tag{6.7}$$

We can now write the induced charge density as:

$$\rho^{ind}(\mathbf{r}) = -e [n_0(\mu + e\phi(\mathbf{r})) - n_0(\mu)] \approx -e^2 \phi(\mathbf{r}) \frac{\partial n_0}{\partial \mu}. \tag{6.8}$$

which is just the difference between the total charge density and the average charge density in the sample (with no test charge), and we make the approximation that not only is ϕ slowly varying, but in itself small, allowing us to take the first-order Taylor approximation to ρ^{ind} and neglect the higher-order terms.

Now, comparing Eqs. (6.1) and (6.8), we can read off the dielectric susceptibility (and the function):

$$\boxed{
\begin{aligned}
\chi(\mathbf{q}) &= -e^2 \frac{\partial n_0}{\partial \mu} \\
\epsilon(\mathbf{q}) &= 1 + \frac{4\pi e^2}{q^2} \frac{\partial n_0}{\partial \mu}
\end{aligned}
} \tag{6.9}$$

this is the main result of Thomas-Fermi theory of electron screening.

6.3 Implications of Thomas-Fermi Theory

Some housekeeping; Eq. (6.9) is often stated in the form:

$$\epsilon(\mathbf{q}) = 1 + \frac{k_{TF}^2}{q^2}, \quad k_{TF} = 4\pi e^2 \frac{\partial n_0}{\partial \mu}. \tag{6.10}$$

k_{TF} is called the Thomas-Fermi momentum, and gives us a length scale for which the test charge is screened out. We illustrate the significance of k_{TF} by considering the screened potential of a point charge:

$$\phi^{ext}(\mathbf{r}) = \frac{Q}{r} \rightarrow \phi^{ext}(\mathbf{q}) = \frac{4\pi Q}{q^2} \quad (6.11)$$

$$\phi(\mathbf{r}) = \frac{1}{\epsilon(\mathbf{q})} \phi^{ext}(\mathbf{q}) = \frac{4\pi Q}{q^2 + k_{TF}^2} \quad (6.12)$$

In real space, this becomes:

$$\phi(\mathbf{r}) = \int \frac{d^3}{(2\pi)^3} e^{i\mathbf{q}\cdot\mathbf{r}} \frac{4\pi Q}{q^2 + k_{TF}^2} = \frac{Q}{r} e^{-k_{TF}r} \quad (6.13)$$

where the integral can be done by going into spherical coordinates. This is a very important result which tells us how screening acts in our simple Thomas-Fermi theory (though it is true more generally). This is known as the screened Coulomb, or Yukawa potential (often used in nuclear physics, though its origin is not electrostatic but nuclear). The presence of electrons exponentially suppresses the test charge with distance, in particular the potential rapidly vanishes at distances $r > k_{TF}^{-1} = \lambda_{TF}$ which we call the screening length.

We may recall that we found a singularity in the derive in the effective dispersion when we covered Hartree-Fock theory. This was due to the long range interactions; but if one accounts for the screening/exponential suppression, then this singularity goes away.

As a last step, let us estimate the size of λ_{TF} in a metal. We have that:

$$\frac{\partial n_0}{\partial \mu} = - \int \frac{d^3 k}{4\pi^3} \frac{\partial}{\partial \mu} \left[\frac{1}{e^{\beta(\epsilon_{\mathbf{k}} - \mu)} + 1} \right] \xrightarrow{T \rightarrow 0} \int \frac{d^3 k}{4\pi^3} \delta(\epsilon_{\mathbf{k}} - \mu) = g(\mu) = \frac{mk_F}{\hbar^2 \pi^2} \quad (6.14)$$

where the last quantity is just the density of states at the Fermi level in 3-D. Sticking this into the result for k_{TF} , we get:

$$\frac{k_{TF}^2}{k_F^2} = \frac{4}{\pi} \frac{me^2}{\hbar^2 k_F} = \frac{4}{\pi} \frac{1}{a_0 k_F} = \left(\frac{16}{3\pi^2} \right)^{3/2} r_s \implies k_{TF} = 0.815 k_F \sqrt{r_s} \implies \lambda_{TF} \approx k_F^{-1} = \frac{\sqrt{r_s}}{2.95} A^\circ \quad (6.15)$$

This tells us that in a metal (where $r_s = 2 - 6$), the Coulomb interaction is screened at very short distances, comparable to the ionic spacings.

A comment: We have made the semiclassical assumption that ϕ varies slowly; but it does not vary slowly on the given length scale. It's not valid in real materials, but it is still a useful result.

6.4 Lindhard Theory

In the Lindhard theory, we assume that $\phi(\mathbf{r})$ is small and can be treated as a perturbation:

$$H = \frac{\hbar^2 \nabla^2}{2m} - e\phi(\mathbf{r}) \quad (6.16)$$

where the first term is H_0 and the second is H_1 . This approximation is almost always valid. We calculate the charge density using Eq. (6.5). We will need to find the correction to $\psi_{\mathbf{k}}(\mathbf{r})$ due to perturbation. This correction looks like:

$$\psi_{\mathbf{k}}(\mathbf{r}) = \psi_{\mathbf{k}}^0(\mathbf{r}) + \sum_{\mathbf{k}'} \frac{\langle \psi_{\mathbf{k}'}^0 | -e\phi(\mathbf{r}) | \psi_{\mathbf{k}}^0 \rangle}{\epsilon_{\mathbf{k}} - \epsilon_{\mathbf{k}'}} \psi_{\mathbf{k}'}^0(\mathbf{r}) + \dots \quad (6.17)$$

Now, since:

$$\psi_{\mathbf{k}}^0(\mathbf{r}) = \frac{1}{\sqrt{V}} e^{i\mathbf{k}\cdot\mathbf{r}}, \quad \epsilon_{\mathbf{k}}^0 = \frac{\hbar^2 \mathbf{k}^2}{2m}$$

we have:

$$\langle \psi_{\mathbf{k}}'^0 | -e\phi(\mathbf{r}) | \psi_{\mathbf{k}}^0 \rangle = -\frac{e}{V} \int d^3r e^{-i\mathbf{k}' \cdot \mathbf{r}} \phi(\mathbf{r}) e^{i\mathbf{k} \cdot \mathbf{r}} = -\frac{e}{V} \phi(\mathbf{k} - \mathbf{k}') \quad (6.18)$$

where in the last equality we recognize the operation being carried out as a Fourier transform. We now have:

$$\psi_{\mathbf{k}}(\mathbf{r}) = \psi_{\mathbf{k}}^0(\mathbf{r}) - \frac{e}{V} \sum_{\mathbf{k}'} \frac{\phi(\mathbf{k} - \mathbf{k}')}{\epsilon_{\mathbf{k}}^0 - \epsilon_{\mathbf{k}'}^0} \psi_{\mathbf{k}'}^0(\mathbf{r}) \quad (6.19)$$

We then substitute $\psi_{\mathbf{k}}(\mathbf{r})$ into Eq. (6.5) and retain terms up to first order in $\phi(\mathbf{k} - \mathbf{k}')$. This yields:

$$\rho(\mathbf{r}) = -e \left[\sum_{\mathbf{k}} f_{\mathbf{k}} |\psi_{\mathbf{k}}^0|^2 - \frac{e}{V} \sum_{\mathbf{k}} \left(f_{\mathbf{k}} \psi_{\mathbf{k}}^{0*} \sum_{\mathbf{k}'} \frac{\psi_{\mathbf{k}'}^0}{\epsilon_{\mathbf{k}}^0 - \epsilon_{\mathbf{k}'}^0} \phi(\mathbf{k} - \mathbf{k}') + c.c. \right) \right] \quad (6.20)$$

Where the first term is $\rho_0(\mathbf{r})$ and the second term is $\rho^{ind}(\mathbf{r})$. $f_{\mathbf{k}}$ is the Fermi-dirac distribution for energy $\epsilon_{\mathbf{k}}$. We can evaluate this integral using the substitution:

$$(\mathbf{k}, \mathbf{k}') \rightarrow (\mathbf{k} + \frac{1}{2}\mathbf{q}, \mathbf{k} - \frac{1}{2}\mathbf{q})$$

We then obtain:

$$\rho^{ind}(\mathbf{r}) = -\frac{e^2}{V} \sum_{\mathbf{q}} e^{i\mathbf{q} \cdot \mathbf{r}} \left(\sum_{\mathbf{k}} \frac{f_{\mathbf{k}+\frac{1}{2}\mathbf{q}} - f_{\mathbf{k}-\frac{1}{2}\mathbf{q}}}{\epsilon_{\mathbf{k}+\frac{1}{2}\mathbf{q}}^0 - \epsilon_{\mathbf{k}-\frac{1}{2}\mathbf{q}}^0} \right) \phi(\mathbf{q}) \quad (6.21)$$

we can therefore conclude that $\rho^{ind}(\mathbf{q}) = \chi(\mathbf{q})\phi(\mathbf{q})$, where:

$$\chi(\mathbf{q}) = -\frac{e^2}{V} \sum_{\mathbf{k}} \frac{f_{\mathbf{k}+\frac{1}{2}\mathbf{q}} - f_{\mathbf{k}-\frac{1}{2}\mathbf{q}}}{\epsilon_{\mathbf{k}+\frac{1}{2}\mathbf{q}}^0 - \epsilon_{\mathbf{k}-\frac{1}{2}\mathbf{q}}^0} \quad (6.22)$$

This is our final result, the Lindhard susceptibility function. We note that $\chi(\mathbf{q})$ looks different from the Thomas-Fermi case, and there is an explicit dependence of $\chi(\mathbf{q})$ on \mathbf{q} .

6.5 Limits of the Lindhard Results

We look at two limits. At low T and small q , the numerator is small unless $|\mathbf{k}| \approx k_F$ (we subtract two near-step functions in the numerator, and the difference is only nonzero for $|\mathbf{k}| \approx k_F$). Therefore we can expand:

$$f_{\mathbf{k} \pm \frac{1}{2}\mathbf{q}} \approx f_{\mathbf{k}} \pm \frac{\hbar^2 \mathbf{k} \cdot \mathbf{q}}{2m} \frac{\partial f_{\mathbf{k}}}{\partial \mu} + O(q^2) \quad (6.23)$$

so we get:

$$\chi(\mathbf{q}) \approx -\frac{e^2}{V} \sum_{\mathbf{k}} \frac{\partial f_{\mathbf{k}}}{\partial \mu} \quad (6.24)$$

we recover the TF result! This makes sense, as the small q limit is the large-wavelength limit, i.e. ϕ varies slowly.

At $T = 0$, the \mathbf{k} integral can be evaluated exactly:

$$\chi(\mathbf{q}) = -e^2 \left(\frac{mk_F}{\pi^2 \hbar^2} \right) \left[\frac{1}{2} + \frac{1-x^2}{4x^2} \ln \left| \frac{1+x}{1-x} \right| \right] \quad (6.25)$$

where $x = q/2k_F$. This is the same function that appeared in our discussion of Hartree-Fock, and it is non-analytic at $x = 1$. Because the dielectric function $\epsilon(\mathbf{q}) = 1 - \frac{4\pi}{q^2}\chi(\mathbf{q})$ is non-analytic at $q \rightarrow 2k_F$, it is possible to show that the screened potential of a point charge contains a term:

$$\phi(\mathbf{r}) \approx \frac{1}{r^3} \cos(2k_F r) \quad (6.26)$$

for $r \gg k_F^{-1}$. This is known as the “Friedel”, or “RKKY” oscillation (so to recap, we have an exponentially decaying term, and we also have a power-law decaying oscillatory term). It is observable experimentally, using STM.

Why is this? Electrons come in to screen the test charge, but “overscreen”, creating a region of high negative charge density. So then the electrons go away from this region to compensate, and this leads to this oscillatory effect.

Next class, we study bosonic excitations in solids.

7 Bosons, Bose-Einstein Condensation, and Helium-4

7.1 Housekeeping

Reading assignment: Fermi liquid theory in Ashcroft and Mermin p345-357. We don't have quite the mathematical tools to explain it, but the excerpt above gives a short qualitative introduction.

Midterm and Final: MT on First week of November. Both exams have the same structure; short in class portion (10-15 minutes), questions concerning things we are expected to know (e.g. $r_s = 2 - 6$ in metals, or temperature resistivity in Fermi liquid theory is T^2 etc.). Then there is a takehome portion with homework-level questions.

7.2 Boson types and statistics

But now we switch gears from fermions to bosons. There are two types of bosonic particles we consider in our discussion:

- (a) "Real" (number-conserving): ^4He , Rb, Na, K, ...
- (b) "Emergent": Phonons (quanta of lattice vibrations), Magnons, ...

The main difference is that for the first class, the number of particles is conserved; the number of Helium atoms in a closed container is fixed. On the other hand, phonons can be created and destroyed out of thin air, so to speak. So there is no conservation law for them.

An important distinction between bosons and fermions is the type of statistics they obey. Mathematically, this is based on commutation relations; in terms of second quantization, bosons obey the commutation relations:

$$\begin{aligned} [a_{\mathbf{k}}^\dagger, a_{\mathbf{k}'}] &= \delta_{\mathbf{k}\mathbf{k}'} \\ [a_{\mathbf{k}}^\dagger, a_{\mathbf{k}'}^\dagger] &= [a_{\mathbf{k}}, a_{\mathbf{k}'}] = 0. \end{aligned} \quad (7.1)$$

Recall the Pauli exclusion principle for fermions (which could be derived from the anticommutation relations); no such principle exists for bosons, of which any number are free to occupy the same quantum state. These commutation relations give rise to the BE distribution function:

$$\bar{n}_{\mathbf{k}} = \frac{1}{e^{\beta(\epsilon_{\mathbf{k}} - \mu)} - 1}$$

where compared to the FD distribution, the sign in the denominator has flipped.

7.3 Deriving the Bose-Einstein Distribution

Let's review how this distribution function comes about. Consider a non-interacting system:

$$H = \sum_{\mathbf{k}} (\epsilon_{\mathbf{k}} - \mu) a_{\mathbf{k}}^\dagger a_{\mathbf{k}} \quad (7.2)$$

The thermal average of any operator \hat{O} is given by:

$$\begin{aligned} \langle \hat{O} \rangle_\beta &= \frac{1}{Z} \sum_j \langle j | \hat{O} | j \rangle e^{-\beta(E_j - \mu N)} \\ Z &= \sum_j e^{-\beta(E_j - \mu N)}. \end{aligned} \quad (7.3)$$

where the matrix elements of \hat{O} are weighted by the boltzmann factor, of which the partition function Z is the sum of: Note we use the grand canonic ensemble so we have the μN . We then have:

$$\begin{aligned}
\langle \hat{n}_{\mathbf{k}} \rangle &= \frac{1}{Z} \sum_j \langle j | a_{\mathbf{k}}^\dagger a_{\mathbf{k}} e^{-\beta(\hat{H} - \mu \hat{N})} | j \rangle \\
&= \frac{1}{Z} \text{Tr}(a_{\mathbf{k}}^\dagger a_{\mathbf{k}} e^{-\beta(\hat{H} - \mu \hat{N})}) \\
&= \frac{1}{Z} \text{Tr}(a_{\mathbf{k}} e^{-\beta(\hat{H} - \mu \hat{N})} a_{\mathbf{k}}^\dagger) \\
&= \frac{1}{Z} \text{Tr}(a_{\mathbf{k}} a_{\mathbf{k}}^\dagger e^{-\beta(\hat{H} - \mu \hat{N})} e^{-\beta(\epsilon_{\mathbf{k}} - \mu)}) \\
&= \frac{1}{Z} \text{Tr}((1 + a_{\mathbf{k}}^\dagger a_{\mathbf{k}}) e^{-\beta(\hat{H} - \mu \hat{N})} e^{-\beta(\epsilon_{\mathbf{k}} - \mu)}) \\
&= (1 + \langle \hat{n}_{\mathbf{k}} \rangle) e^{-\beta(\epsilon_{\mathbf{k}} - \mu)}
\end{aligned} \tag{7.4}$$

where in the first line, we take advantage of the fact that when \hat{H} acts on its eigenstate, it gives back the eigenvalue. So we can pull the constant into the matrix element and convert the energy into the Hamiltonian. The second line we cast this expression as a trace. In the third line we use the cyclicity of the trace. In the fourth line we skip some math (but it can be done in detail) but this can be viewed as creating a single boson with energy $\epsilon_{\mathbf{k}}$. In the fifth line we use the bosonic commutation relations. In the sixth line we evaluate the expression. We can then solve for $\langle \hat{n}_{\mathbf{k}} \rangle$ to obtain:

$$\langle \hat{n}_{\mathbf{k}} \rangle = \frac{1}{e^{\beta(\epsilon_{\mathbf{k}} - \mu)} - 1}. \tag{7.5}$$

Note that a similar derivation can be done for Fermions to get the Fermi-Dirac distribution.

7.4 Bose-Einstein Condensation

BE Condensation occurs in real (or number conserving) bosons, most famously Helium-4 at low temperature. The easiest way to see this occurs is to consider the total number of bosons N :

$$N = \sum_{\mathbf{k}} \bar{n}_{\mathbf{k}} = \sum_{\mathbf{k}} \frac{1}{e^{\beta(\epsilon_{\mathbf{k}} - \mu)} - 1}, \quad \epsilon_{\mathbf{k}} = \frac{\hbar^2 \mathbf{k}^2}{2m} \tag{7.6}$$

Note that for real bosons $\mu \leq 0$; otherwise we would have $\bar{n}_{\mathbf{k}} < 0$ for some \mathbf{k} which is forbidden. This implies that:

$$e^{\beta(\epsilon_{\mathbf{k}} - \mu)} \geq e^{\beta \epsilon_{\mathbf{k}}}$$

For this reason, we can bound the total number of bosons from above:

$$\begin{aligned}
N &\leq \sum_{\mathbf{k}} \frac{1}{e^{\beta \epsilon_{\mathbf{k}}} - 1} \\
&= N_0 + \frac{\Omega}{(2\pi)^3} \int_0^\infty dk \frac{4\pi k^2}{e^{\beta \hbar^2 k^2 / 2m} - 1}
\end{aligned} \tag{7.7}$$

where we have separated the sum into two terms; the $\mathbf{k} = 0$ term (which is problematic as it formally diverges; it only comes about as we have discarded the chemical potential) and the rest of the sum rewritten as an integral, which we call $N'(T)$. Ω here is the volume. We leave the N_0 term for now and evaluate the $N'(T)$ by substitution. We let $x = \beta \frac{\hbar^2 k^2}{2m}$ and $dx = \beta \frac{\hbar^2}{m} k dk$ so:

$$N'(T) = \frac{\Omega}{(2\pi)^3} 4\pi \sqrt{2} \left(\frac{m}{\beta \hbar^2} \right)^{3/2} \int_0^\infty \frac{\sqrt{x} dx}{e^x - 1} = C \Omega T^{3/2} \tag{7.8}$$

The actual integral on the right is a finite constant (formally it can be evaluated by considering the Riemann zeta function), but we are really only interested in the temperature dependence, so we've lumped things into a constant C . It's also important that $N'(T)$ is extensive/grows proportionally to the system volume. Now we look at how BE condensation comes about from this. We rewrite the inequality as:

$$N \leq N_0 + N'(T)$$

We plot $N'(T)$ as a function of T . N is fixed. There is some T_c for which $N'(T)$ intersects N ; above T_c , we have $N_0 = 0$, and below T_c we have $N_0 > 0$. We have an extensive number of bosons in the $\mathbf{k} = 0$ state. This is what is known as BE condensation. In particular as we take $T \rightarrow 0$ all of the bosons occupy the ground state. This is not surprising; in the view of QM, if we try to minimize the energy of a set of bosons, we can just cram them all into the ground state (there is nothing preventing us from doing this; no Pauli exclusion)! However in terms of classical physics this phenomena was unusual.

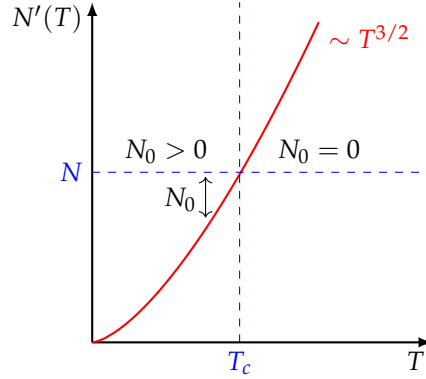


Figure 7.1: Plot of the upper bound of excited state bosons $N'(T)$ as a function of temperature T . Above some T_c , we have zero bosons in the ground state and so $N_0 = 0$. Below some T_c , we have that $N_0 > 0$ and so an extensive number of bosons occupy the $\mathbf{k} = 0$ ground state; we therefore have Bose-Einstein condensation.

7.5 Bogoliubov Theory of Helium-4

This is a classic theory; 1946 (but still valid to this day)! Helium-4 was interesting from the early days of physics at it had interesting superfluid properties. Bogoliubov started off this explanation, and Landau later would give an argument for why Helium-4 is superfluidic (which we cover next lecture).

We consider weakly interacting (spinless) bosons described by the Hamiltonian:

$$H = \sum_{\mathbf{k}} \epsilon_{\mathbf{k}} a_{\mathbf{k}}^\dagger a_{\mathbf{k}} + \frac{1}{2} \sum_{\mathbf{k}, \mathbf{p}, \mathbf{q}} V_{\mathbf{q}} a_{\mathbf{k}-\mathbf{q}}^\dagger a_{\mathbf{p}+\mathbf{q}}^\dagger a_{\mathbf{p}} a_{\mathbf{k}} \quad (7.9)$$

where $V_{\mathbf{q}}$ is a Fourier transform of a short-range interatomic potential; it is a short range interaction that dies off quickly outside of the Helium atom (on the order of an Angstrom). As usual, $\epsilon_{\mathbf{k}} = \frac{\hbar^2 \mathbf{k}^2}{2m}$. We assume (following Bogoliubov) that $V_{\mathbf{q}}$ is weak and $T \ll T_c$. Then we expect the ground state to be close to a perfect BEC, where:

$$|\Phi_0^N\rangle = (a_0^\dagger)^N |0\rangle \quad (7.10)$$

Now we perform the following approximation:

$$\begin{aligned} a_0^\dagger a_0 &\rightarrow \langle a_0^\dagger a_0 \rangle = N_0 \approx N \\ a_0^\dagger a_0^\dagger &\rightarrow N_0 \end{aligned} \quad (7.11)$$

The first line: whenever we see the number operator, we take it to be its average value, which is large/close to the total number. The second line is less obvious, but consider $a_0^\dagger a_0^\dagger |\Phi_0^N\rangle = \sqrt{(N_0+1)(N_0+2)} |\Phi_0^{N+2}\rangle$ and we can then take $\sqrt{(N_0+1)(N_0+2)} \approx N_0$ for N_0 large. The assumption being made is that the Hamiltonian is always acting on a state close to the perfect BEC ground state, so we can approximate these operators as we have above. In the interaction term, we split all sums as:

$$\sum_{\mathbf{k}} = \sum_{\mathbf{k}=0} + \sum'_{\mathbf{k}}$$

and retain only terms containing at least only one power of N_0 . This is a bit of a mess (as we have 8 sums to work with...); we will not go through it explicitly, but we justify this approximation by saying that since N_0 is large, all terms without powers of N_0 are relatively small and hence can be neglected. The result is as follows:

$$H \approx \sum_{\mathbf{k}} \epsilon_{\mathbf{k}} a_{\mathbf{k}}^\dagger a_{\mathbf{k}} + \frac{1}{2} N_0^2 V_0 + N_0 V_0 \sum'_{\mathbf{k}} a_{\mathbf{k}}^\dagger a_{\mathbf{k}} + N_0 \sum'_{\mathbf{q}} V_{\mathbf{q}} a_{\mathbf{q}}^\dagger a_{\mathbf{q}} + \frac{1}{2} N_0 \sum'_{\mathbf{q}} V_{\mathbf{q}} (a_{\mathbf{q}} a_{-\mathbf{q}} + a_{\mathbf{q}}^\dagger a_{-\mathbf{q}}^\dagger) \quad (7.12)$$

The first (kinetic energy) term remains unchanged. The second term comes from $\mathbf{k} = v\mathbf{q} = \mathbf{p} = \mathbf{0}$, The third term comes from $\mathbf{p} = \mathbf{q} = \mathbf{0}$ or $\mathbf{k} = \mathbf{q} = \mathbf{0}$ and so on. To simplify, we define:

$$\begin{aligned} \eta_{\mathbf{k}} &= N_0 V_{\mathbf{k}} \\ \hbar\Omega_{\mathbf{k}} &= \epsilon_{\mathbf{k}} + \eta_{\mathbf{k}}. \end{aligned} \quad (7.13)$$

Notice also that $N_0 + \sum'_{\mathbf{k}} a_{\mathbf{k}}^\dagger a_{\mathbf{k}} \approx N$, where $N_0, N \gg N' = \sum'_{\mathbf{k}} a_{\mathbf{k}}^\dagger a_{\mathbf{k}}$. With this, let us combine some terms:

$$\frac{1}{2} N^2 V_0 = \frac{1}{2} V_0 \left[N_0^2 + 2N_0 \sum'_{\mathbf{k}} a_{\mathbf{k}}^\dagger a_{\mathbf{k}} + \dots \right] \quad (7.14)$$

Hence we can write the entire Hamiltonian as:

$$H \approx \frac{1}{2} N^2 V_0 + \sum_{\mathbf{k}} \left[\hbar\Omega_{\mathbf{k}} a_{\mathbf{k}}^\dagger a_{\mathbf{k}} + \frac{1}{2} \eta_{\mathbf{k}} (a_{\mathbf{k}} a_{-\mathbf{k}} + a_{\mathbf{k}}^\dagger a_{-\mathbf{k}}^\dagger) \right] \quad (7.15)$$

We draw our attention to the last term(s); these are “anomalous terms”, which *do not* conserve particle number. This is a consequence of the Bogoliubov approximation. Physically, $a_{\mathbf{k}} a_{-\mathbf{k}}$ represent bosons $(\mathbf{k}, -\mathbf{k})$ “disappearing” into the condensate. The number of bosons in the condensate is so large that you do not have to keep track of the bosons in the condensate itself; we only need to keep track of the other particles as they disappear and appear out of it.

7.6 Bogoliubov Transformations and Quasiparticle Spectrum

Bogoliubov also developed a theory of how to treat such Hamiltonians. They can be diagonalized by means of Bogoliubov transformations:

$$\begin{aligned} a_{\mathbf{k}} &= \mu_{\mathbf{k}} \alpha_{\mathbf{k}} + \nu_{\mathbf{k}} \alpha_{-\mathbf{k}}^\dagger & \alpha_{\mathbf{k}} &= \mu_{\mathbf{k}} a_{\mathbf{k}} - \nu_{\mathbf{k}} a_{-\mathbf{k}}^\dagger \\ a_{\mathbf{k}}^\dagger &= \mu_{\mathbf{k}} \alpha_{\mathbf{k}}^\dagger + \nu_{\mathbf{k}} \alpha_{-\mathbf{k}} & \alpha_{\mathbf{k}}^\dagger &= \mu_{\mathbf{k}} a_{\mathbf{k}}^\dagger - \nu_{\mathbf{k}} a_{-\mathbf{k}} \end{aligned} \quad (7.16)$$

where we have the forwards transformations on the left and the inverse transformation on the right. Here, $\alpha_{\mathbf{k}}$ are our new bosonic “quasiparticle” operators and $(\mu_{\mathbf{k}}, \nu_{\mathbf{k}})$ are real coefficients. We want this to be a canonical transformation, so we want these new α operators to satisfy the same commutation relations as the a s. This is important as it will place constraints on the $\mu_{\mathbf{k}}$ and $\nu_{\mathbf{k}}$. Let's see what happens:

$$\begin{aligned} [\alpha_{\mathbf{k}}, \alpha_{\mathbf{k}'}^\dagger] &= [\mu_{\mathbf{k}} a_{\mathbf{k}} - \nu_{\mathbf{k}} a_{-\mathbf{k}}^\dagger, \mu_{\mathbf{k}'} \alpha_{\mathbf{k}'}^\dagger - \nu_{\mathbf{k}'} a_{-\mathbf{k}'}^\dagger] \\ &= \mu_{\mathbf{k}} \mu_{\mathbf{k}'} [a_{\mathbf{k}}, a_{\mathbf{k}'}^\dagger] + \nu_{\mathbf{k}} \nu_{\mathbf{k}'} [a_{-\mathbf{k}}^\dagger, a_{-\mathbf{k}'}^\dagger] \\ &= (\mu_{\mathbf{k}}^2 - \nu_{\mathbf{k}}^2) \delta_{\mathbf{k}\mathbf{k}'} \end{aligned} \quad (7.17)$$

where we have used the known commutation relations for the a s. So we obtain the constraint:

$$[\alpha_{\mathbf{k}}, \alpha_{\mathbf{k}'}^\dagger] = \mu_{\mathbf{k}}^2 - \nu_{\mathbf{k}}^2 = 1. \quad (7.18)$$

A comment: we want to find $(\mu_{\mathbf{k}}, \nu_{\mathbf{k}})$ such that the resulting H is diagonal, that is:

$$H = \sum_{\mathbf{k}} \hbar \omega_{\mathbf{k}} \alpha_{\mathbf{k}}^\dagger \alpha_{\mathbf{k}} + E_0. \quad (7.19)$$

So to this end we compute:

$$\begin{aligned} \alpha_{\mathbf{k}}^\dagger \alpha_{\mathbf{k}} &= (\mu_{\mathbf{k}} a_{\mathbf{k}}^\dagger - \nu_{\mathbf{k}} a_{-\mathbf{k}})(\mu_{\mathbf{k}} a_{\mathbf{k}} - \nu_{\mathbf{k}} a_{-\mathbf{k}}^\dagger) \\ &= \mu_{\mathbf{k}}^2 a_{\mathbf{k}}^\dagger a_{\mathbf{k}} + \nu_{\mathbf{k}}^2 a_{-\mathbf{k}} a_{-\mathbf{k}}^\dagger - \mu_{\mathbf{k}} \nu_{\mathbf{k}} (a_{\mathbf{k}} a_{-\mathbf{k}} + a_{\mathbf{k}}^\dagger a_{-\mathbf{k}}^\dagger) \end{aligned} \quad (7.20)$$

Comparing this with the form of the Hamiltonian, we can see we are on the right track. The only terms to seem to not appear there are the $a_{-\mathbf{k}} a_{-\mathbf{k}}^\dagger$ s but we recognize the sum to run over all \mathbf{k} and so we can make a substitution there. We make an assumption that $\omega_{\mathbf{k}} = \omega_{-\mathbf{k}}$ and $\nu_{\mathbf{k}}^2 = \nu_{-\mathbf{k}}^2$ (we can check later that this is consistent) then:

$$\sum_{\mathbf{k}} \hbar \omega_{\mathbf{k}} \alpha_{\mathbf{k}}^\dagger \alpha_{\mathbf{k}} = \sum_{\mathbf{k}} \hbar \omega_{\mathbf{k}} (\mu_{\mathbf{k}}^2 + \nu_{\mathbf{k}}^2) a_{\mathbf{k}}^\dagger a_{\mathbf{k}} + \sum_{\mathbf{k}} \hbar \omega_{\mathbf{k}} \nu_{\mathbf{k}}^2 - \sum_{\mathbf{k}} \hbar \omega_{\mathbf{k}} \mu_{\mathbf{k}} \nu_{\mathbf{k}} (a_{\mathbf{k}} a_{-\mathbf{k}} + a_{\mathbf{k}}^\dagger a_{-\mathbf{k}}^\dagger) \quad (7.21)$$

Comparison with the original Hamiltonian (Eq. (7.15)) implies:

$$\begin{aligned} \hbar \omega_{\mathbf{k}} (\mu_{\mathbf{k}}^2 + \nu_{\mathbf{k}}^2) &= \hbar \Omega_{\mathbf{k}} \\ -2\hbar \omega_{\mathbf{k}} \mu_{\mathbf{k}} \nu_{\mathbf{k}} &= \eta_{\mathbf{k}} \end{aligned} \quad (7.22)$$

To solve for $\omega_{\mathbf{k}}$ we square both equations and subtract:

$$(\hbar \omega_{\mathbf{k}})^2 [(\mu_{\mathbf{k}}^2 + \nu_{\mathbf{k}}^2)^2 - 4\mu_{\mathbf{k}}^2 \nu_{\mathbf{k}}^2] = (\hbar \Omega_{\mathbf{k}})^2 - \eta_{\mathbf{k}}^2$$

but on the LHS We have $(\mu_{\mathbf{k}}^2 - \nu_{\mathbf{k}}^2)^2 = 1$ and so the dependence on μ/ν drops out and we have:

$$\hbar \omega_{\mathbf{k}} = \sqrt{\hbar^2 \Omega_{\mathbf{k}}^2 - \eta_{\mathbf{k}}^2} = \sqrt{\epsilon_{\mathbf{k}} (\eta_{\mathbf{k}} + 2N V_{\mathbf{k}})} \quad (7.23)$$

which is one of the main results of Bogoliubov theory, the “spectrum of quasiparticle excitations”. This spectrum will have interesting implications; among other things, it will be the basis for superfluidity in liquid He-4 according to the Landau argument. Let us discuss some special cases of this spectrum.

7.7 Quasiparticle Spectrum - Special Cases

7.7.1 Non-Interacting

If the bosons are noninteracting, then $V_{\mathbf{k}} = 0$ and so:

$$\hbar \omega_{\mathbf{k}} = \sqrt{\epsilon_{\mathbf{k}}^2} = \frac{\hbar^2 \mathbf{k}^2}{2m}$$

so we indeed reproduce our prior result for free bosons.

7.7.2 Contact Repulsion

We take $V(\mathbf{r}) = U\delta(\mathbf{r})$ (i.e. the bosons repel when they are on top of each other). Since the FT of a dirac delta is just a constant, we have:

$$V_{\mathbf{k}} = \frac{U}{\Omega}$$

for all \mathbf{k} . We then have:

$$\hbar\omega_{\mathbf{k}} = \sqrt{\epsilon_{\mathbf{k}}(\epsilon_{\mathbf{k}} + 2(NU/\Omega))} = \begin{cases} \sqrt{\epsilon_{\mathbf{k}}E_0} \sim |\mathbf{k}|, & \epsilon_{\mathbf{k}} \ll E_0 \text{ (long wavelength = sound-like dispersion)} \\ |\epsilon_{\mathbf{k}}| \sim \mathbf{k}^2, & \epsilon_{\mathbf{k}} \gg E_0 \text{ (short wavelength = particle-like dispersion)} \end{cases} \quad (7.24)$$

where we define $2(NU/\Omega) = E_0$. So the weak interactions of Helium modify the dispersion from particle like to sound-like at long wavelengths with linear dispersion relation.

7.7.3 Typical Helium-4 Interaction

The Fermi velocity for a typical interaction in Helium-4 scales with the wavevector as:

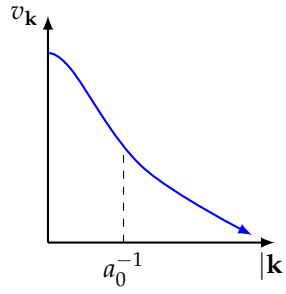


Figure 7.2: $v_{\mathbf{k}}$ vs. $|\mathbf{k}|$ for a typical Helium-4 interaction.

From this we obtain a dispersion relation that looks as follows:

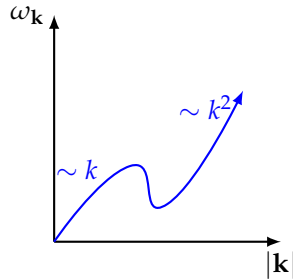


Figure 7.3: Dispersion relation for a typical Helium-4 interaction. The dispersion scales as k for small $|\mathbf{k}|$ and as k^2 for large $|\mathbf{k}|$. The minima present in between the two regimes is known as the roton minimum.

This concludes the Bogoliubov theory of Helium-4 in a nutshell. In the next assignment we will explicitly calculate the coefficients in the Bogoliubov transformation and other fun exercises.

7.8 Landau Argument for Superfluidity

To conclude, we cover Landau's argument for superfluidity for Liquid Helium 4. At the time, it was a known experimental fact that liquid Helium-4 would flow in a tube without friction. Experiments

were also done where one would start spinning a bucket of Helium-4 at temperature above T_c , then the friction would cause it to rotate; then if one cooled the bucket down, the Helium would continue to rotate indefinitely.

Landau considered the following scenario. He considered a vessel of liquid Helium-4 and an object of large mass $M \gg m$ moving with velocity \mathbf{v} . The question Landau asked was can the object relax its energy and momentum by emitting quasiparticle excitations?

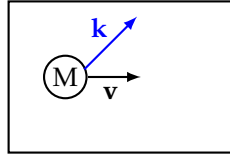


Figure 7.4: Setting of the Landau argument for the superfluidity criterion. A heavy object with mass M travels through a vessel of liquid Helium-4 with velocity \mathbf{v} ; can it relax its energy/momentum by emitting a quasiparticle of wavevector \mathbf{k} ?

We will show that it is impossible to conserve both energy and momentum and for the particle to relax so long as the object does not exceed a critical velocity v_c . Energy conservation and momentum conservation tell us that:

$$\begin{aligned}\frac{1}{2}M\mathbf{v}^2 &= \frac{1}{2}M\mathbf{v}'^2 + \hbar\omega_{\mathbf{k}} \\ M\mathbf{v} &= M\mathbf{v}' + \hbar\mathbf{k}\end{aligned}\tag{7.25}$$

These are the two equations we must satisfy if this process is allowed. Squaring the second equation, we have:

$$M^2\mathbf{v}'^2 = M^2\mathbf{v}^2 + \hbar^2\mathbf{k}^2 - 2M\hbar\mathbf{k} \cdot \mathbf{v}$$

We can divide out by $2M$ and combine the two conservation equations to find:

$$\hbar\mathbf{v} \cdot \mathbf{k} = \frac{\hbar^2\mathbf{k}^2}{2M} + \hbar\omega_{\mathbf{k}}\tag{7.26}$$

We consider $M \gg m$, and so we can neglect the $\frac{\hbar^2\mathbf{k}^2}{2m}$ term. Further we can let $\mathbf{v} \cdot \mathbf{k} = |\mathbf{v}||\mathbf{k}| \cos \theta$ (where θ is the angle between the velocity and the wavevector of the emitted quasiparticle) and so:

$$\hbar|\mathbf{v}||\mathbf{k}| \cos \theta = \hbar\omega_{\mathbf{k}}\tag{7.27}$$

For low enough velocities, the LHS will always be less than the RHS.

The graphical solution shows that slow moving objects cannot dissipate energy and momentum in liquid Helium-4. We therefore have dissipationless motion, or superflow.

Note that one can solve for a critical velocity:

$$v_c = \min_{\mathbf{k}} \left(\frac{\omega_{\mathbf{k}}}{|\mathbf{k}|} \right) \approx 1 \text{cm s}^{-1} \text{ in Helium-4}\tag{7.28}$$

Note that the actual critical velocity is 10 times slower, due to neglected effects such as the creation of vortex rings etc. but nevertheless Landau's argument is an argument that provides some nice explanatory power!

Question: Why does argument not predict superfluid flow in general life situations? This is left as an exercise for the reader...

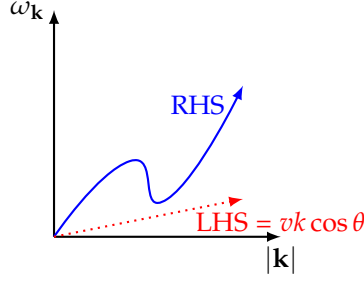


Figure 7.5: Since $\cos \theta \leq 1$, for all θ the LHS ($\hbar v k \cos \theta$) of Eq. (7.27) will be less than the RHS ($\hbar \omega_k$).

8 Lattice Vibrations & Phonons

8.1 Phonons in One Dimension

We consider a simple monoatomic chain: We label the atoms by index l and the displacements from

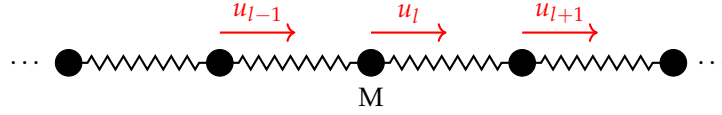


Figure 8.1: A monoatomic chain. The atoms are of mass M and are labelled with index l , with u_l denoting the l th-atoms' displacement from equilibrium.

equilibrium for the l th atom by u_l . We assume we have a crystal lattice with individual atoms oscillating around equilibrium positions. The general Hamiltonian we can write for this system is:

$$\begin{aligned}
 H &= \sum_l \frac{p_l^2}{2M} + V(u_1, u_2, \dots) \\
 p_l &= -i\hbar \frac{\partial}{\partial u_l} \\
 [u_l, p_{l'}] &= i\hbar \delta_{ll'}
 \end{aligned} \tag{8.1}$$

Where we take the momentum to be canonically conjugate to the displacement and hence it satisfies the canonical commutation relation. We assume that the potential has a minimum for $u_l = 0$ for all l . We expand V in a Taylor series about these minima:

$$V(u_1, u_2, \dots) = V(0, 0, \dots) + \sum_l u_l \left[\frac{\partial V}{\partial u_l} \right]_{u_1=u_2=\dots=0} + \frac{1}{2!} \sum_{l,l'} u_l u_{l'} \left[\frac{\partial^2 V}{\partial \mu_l \partial \mu_{l'}} \right]_{u_1=u_2=\dots=0} + \frac{1}{3!} \dots \tag{8.2}$$

The first term can be eliminated by suitable choice of energy zero (a constant in the Hamiltonian changes none of the physics). The second term vanishes by virtue of $u_1 = u_2 = \dots = 0$ being a minimum of V . We are left with just the second-order term, and so:

$$H \approx \sum_l \frac{p_l^2}{2M} + \frac{1}{2} \sum_{l,l'} u_l V_{ll'} u_{l'} \tag{8.3}$$

where $V_{ll'} = \left[\frac{\partial^2 V}{\partial \mu_l \partial \mu_{l'}} \right]_{u_1=u_2=\dots=0}$ is the dynamical matrix. We neglect all higher order terms in Eq. (8.2) which amounts to a harmonic approximation.

8.2 Diagonalizing the Potential

We next diagonalize the V -term. To this end, we define a vector of displacements:

$$Y = \begin{pmatrix} \mu_1 \\ \mu_2 \\ \vdots \end{pmatrix}$$

and then we have:

$$V = \begin{pmatrix} V_{11} & V_{12} & \dots \\ V_{21} & \ddots & \\ \vdots & & \end{pmatrix}$$

We now write that the V -term is:

$$Y^T V Y = Y^T U^\dagger U V U^\dagger U Y = (Y^T U^\dagger)(U V U^\dagger)(U Y) = \tilde{Y}^\dagger \tilde{V} \tilde{Y} = \sum_q \tilde{Y}_q^\dagger \tilde{V}_{qq} \tilde{Y}_q \quad (8.4)$$

where U is the unitary matrix that diagonalizes V . This is of course nothing more than a change of basis. Since it is diagonal in the new basis, we can index it with a new index q .

It is easy to see that this transformation leaves the kinetic energy unchanged (i.e. it remains diagonal!). To this end we consider:

$$p_l = -i\hbar \frac{\partial}{\partial \mu_l} = -i\hbar \sum_q \frac{\partial \tilde{\mu}_q}{\partial \mu_l} \frac{\partial}{\partial \tilde{\mu}_q} = -i\hbar \sum_q U_{lq} \frac{\partial}{\partial \tilde{\mu}_q} = \sum_q U_{lq} (-i\hbar \frac{\partial}{\partial \tilde{\mu}_q}) \quad (8.5)$$

or in other words: $\tilde{P} = U^{-1}P$. Therefore:

$$\sum_l p_l^2 = P^T P = (U\tilde{P})^\dagger (U\tilde{P}) = \tilde{P}^\dagger U^\dagger U \tilde{P} = \tilde{P}^\dagger \mathbb{I} \tilde{P} = \tilde{P}^\dagger \tilde{P} = \sum_q \tilde{p}_q^\dagger \tilde{p}_q \quad (8.6)$$

So the kinetic energy remains diagonal! An important remark; in our canonical representation, P is a Hermitian operator. But the transformed P may no longer have this property, so we keep the \tilde{p}_q^\dagger in the last expression (instead of assuming it will be equal to \tilde{p}_q). In the new coordinates, the Hamiltonian reads:

$$H = \sum_q \left(\frac{1}{2M} p_q^\dagger p_q + \frac{1}{2} M \omega_q \mu_q^\dagger \mu_q \right) \quad (8.7)$$

where we define $\omega_q^2 = V_q / M$ (and dropped the \sim s so we don't have to keep writing it; the new coordinates are distinguished via their index). So our H is diagonal! Let us also note that the unitary transformation preserves the commutation rules (Check!), so:

$$[\mu_q, p_{q'}] = i\hbar \delta_{qq'}. \quad (8.8)$$

Eq. (8.7) is seen to describe a collection of decoupled harmonic oscillators; we can solve this by the usual raising and lowering operator method. For each mode we define:

$$\begin{aligned} a_q &= \frac{1}{\sqrt{2M\hbar\omega_q}} (M\omega_q \mu_q + i p_q^\dagger) \\ a_q^\dagger &= \frac{1}{\sqrt{2M\hbar\omega_q}} (M\omega_q \mu_q^\dagger - i p_q) \end{aligned} \quad (8.9)$$

these satisfy the usual algebra of raising and lowering operators:

$$[a_q, a_q^\dagger] = \delta_{qq'}.$$

In this representation the Hamiltonian acquires a simple form:

$$H = \sum_q \hbar \omega_q \left(a_q^\dagger a_q + \frac{1}{2} \right) \quad (8.10)$$

Often the goal is to know what the ω_q is; this will give us the spectrum of excitations and how they depend on q . We will comment on this shortly when we do an example. In principle this is not difficult as $\omega_q = V_q/M$, and we can get V_q from transforming the dynamical matrix.

8.3 Translation-Invariant Systems

This entire discussion so far was quite general; however since most solid-state systems are translation invariant, we can now consider such systems to analyze. In particular, this has the effect that the dynamical matrix depends only on differences $l - l'$:

$$V_{ll'} = V_{l-l'} \quad (8.11)$$

This is useful as we can guess immediately what the transformation will be; a Fourier transform!

$$U_{ql} = \frac{1}{\sqrt{N}} e^{iql} \quad (8.12)$$

therefore:

$$\tilde{V}_{qq'} = [U^\dagger V U]_{qq'} = \frac{1}{N} \sum_{ll'} e^{-iq'l} V_{l-l'} e^{iq'l'} = \left(\frac{1}{N} \sum_{l'} e^{-il'(q-q')} \right) \left(\sum_l e^{-iq'l} V_l \right) = \delta_{qq'} V_q \quad (8.13)$$

where in the third equality we make the substitution $l \rightarrow l + l'$. In the fourth equality the δ comes from basic Fourier analysis (one can imagine the oscillations exactly cancelling if $q \neq q'$, and exactly adding up if $q = q'$).

We will now see what are the consequences we can draw from this. One thing to specify before we go on: let us use periodic boundary conditions, so that the last atom in the chain of atoms is connected back to the first chain. Mathematically, this is expressed as:

$$\mu_{l+Na} = \mu_l$$

where in the above l represents a distance rather than an index. Note that this implies $e^{iq'l} = e^{iq(l+Na)}$. This further implies a restriction on the form of the q s, namely:

$$1 = e^{iqNa} \implies q = \frac{2\pi n}{Na}, \quad n \in \mathbb{Z}. \quad (8.14)$$

From this we have:

$$\mu_q = \frac{1}{\sqrt{N}} \sum_l e^{-iq'l} \mu_l, \quad p_q = \frac{1}{\sqrt{N}} \sum_l e^{iq'l} p_l \quad (8.15)$$

It follows that:

$$\mu_{q+G} = \mu_q, \quad p_{q+G} = p_q, \quad \text{with } G = \frac{2\pi}{a}. \quad (8.16)$$

Therefore; momentum is only defined in the "1st Brillouin zone" with $q \in (-\frac{\pi}{a}, \frac{\pi}{a})$ and there are only N distinct values,

$$q = \frac{2\pi n}{Na}, \quad n = -\frac{N}{2} + 1, \dots, \frac{N}{2}. \quad (8.17)$$

Much like the case with electrons in a periodic potential, for phonons everything happens in the first Brillouin zone.

One more comment before proceeding with an example. It also holds that:

$$\mu_q^\dagger = \mu_{-q}, \quad p_q^\dagger = p_{-q}$$

further accenting that the transformed μ/p s are not Hermitian, unlike their canonical counterparts? Taking into account $\omega_{-q} = \omega_q$, Eqs. (8.9) can be inverted as:

$$\begin{aligned} \mu_q &= \sqrt{\frac{\hbar}{2M\omega_q}}(a_{-q}^\dagger + a_q) \\ p_q &= i\sqrt{\frac{M\hbar\omega_q}{2}}(a_{-q}^\dagger - a_q) \end{aligned} \quad (8.18)$$

This is not profound (it is just an inversion) but will be quite useful when we proceed to calculate expectation values of operators.

8.4 Example 1 - Monoatomic chain

We consider again a monoatomic chain as in Fig. 8.1; a system of masses M connected by springs of spring constant K with periodic boundary conditions. The potential energy (and from it the dynamical matrix) is given by:

$$\begin{aligned} V &= \sum_l \frac{1}{2} K (\mu_l - \mu_{l+1})^2 = \sum_l \frac{1}{2} K (2u_l^2 - 2u_l u_{l+1}) \\ \Rightarrow V_{ll'} &= \begin{cases} 2K & \text{if } l = l' \\ -K & \text{if } l = l' \pm 1 \\ 0 & \text{otherwise} \end{cases} \end{aligned} \quad (8.19)$$

note there is no factor of 2 in front of the $-K$ because there are two entries in the matrix; one above/below the diagonal. The fourier transform gives:

$$V_q = 2K - K(e^{iqa} + e^{-iqa}) = 2K(1 - \cos(qa)) = 4K \sin^2 \frac{qa}{2} \quad (8.20)$$

from this we can read off the spectrum of normal modes:

$$\omega_q = \sqrt{\frac{V_q}{M}} = 2\sqrt{\frac{K}{M}} \left| \sin \frac{qa}{2} \right| \approx \sqrt{\frac{K}{M}} |qa| \text{ for } |qa| \ll 1 \quad (8.21)$$

so the dispersion is linear (for small qa), signifying a sound-like or acoustic mode. This should be true, because sound propagates through solids! Let us now sketch the spectrum, as a function of q ; known as the “phonon dispersion” of the “phonon energy spectrum”. We often focus on a phonon of a particular wavenumber q and specific energy; and this is indeed something that is experimentally measurable! (E.g. by neutron scattering).

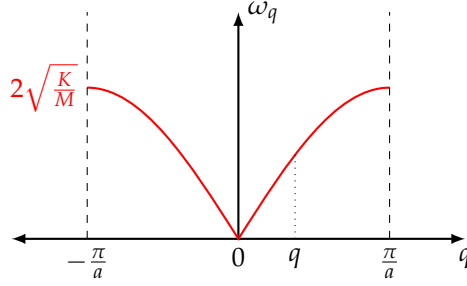


Figure 8.2: Dispersion relation for the monoatomic chain.

8.5 Example 2 - Diatomic Chain

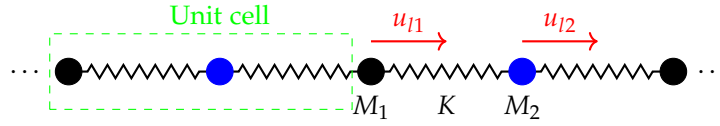


Figure 8.3: A diatomic chain. We have alternating atoms of mass M_1/M_2 joined by springs of spring constant K . We index the displacement of the M_1 atoms with u_{l_1} and the M_2 atoms as u_{l_2} . A unit cell contains one atom of each type and two of the springs.

We now consider a diatomic chain. We now have two different types of atoms in our chain; mass M_1 and mass M_2 . The springs between them are still of spring constant K . The unit cell contains one atom of each type and two of the springs. We denote the displacement of the M_1 atoms as u_{l_1} and the M_2 atoms as u_{l_2} .

Similar analysis as before gives the momentum space dynamical matrix (Check!):

$$V_q = \begin{pmatrix} 2K & -K(1 - e^{iqa}) \\ -K(1 + e^{-iqa}) & 2K \end{pmatrix} \quad (8.22)$$

The Hamiltonian that follows from this is:

$$H = \sum_q \left(\frac{p_{q1}^\dagger p_{q1}}{2M_1} + \frac{p_{q2}^\dagger p_{q2}}{2M_2} \right) + \frac{1}{2} \sum_q \begin{pmatrix} \mu_{q1}^\dagger & \mu_{q2}^\dagger \end{pmatrix} V_q \begin{pmatrix} \mu_{q1} \\ \mu_{q2} \end{pmatrix} \quad (8.23)$$

this is typical for a more complicated chain, where N different atoms would give an $N \times N$ V_q matrix in momentum space. Another complication is that $M_1 \neq M_2$, and so it is not immediate to read off ω_q ; we need to get rid of unequal masses in the kinetic energy term. This is done through the rescaling of momenta:

$$p_{q1} \rightarrow p_{q1} \left(\frac{M_1}{M_2} \right)^{1/4} \quad p_{q2} \rightarrow p_{q2} \left(\frac{M_2}{M_1} \right)^{1/4} \quad (8.24)$$

This implies a corresponding scaling of the displacements:

$$\mu_{q1} \rightarrow \mu_{q1} \left(\frac{M_2}{M_1} \right)^{1/4} \quad \mu_{q2} \rightarrow \mu_{q2} \left(\frac{M_1}{M_2} \right)^{1/4} \quad (8.25)$$

so the Hamiltonian becomes:

$$H = \sum_q \frac{p_{1q}^\dagger p_{1q} + p_{2q}^\dagger p_{2q}}{2\sqrt{M_1 M_2}} + \frac{1}{2} \sum_q \begin{pmatrix} \mu_{q1}^\dagger & \mu_{q2}^\dagger \end{pmatrix} \begin{pmatrix} 2K\sqrt{\frac{M_1}{M_2}} & -K(1 + e^{iqa}) \\ -K(1 + e^{-iqa}) & 2K\sqrt{\frac{M_2}{M_1}} \end{pmatrix} \begin{pmatrix} \mu_{q1} \\ \mu_{q2} \end{pmatrix} \quad (8.26)$$

the normal modes are given by eigenvalues of \tilde{V}_q (the matrix in the potential energy term above):

$$0 = \det(\tilde{V}_q - \omega \sqrt{M_1 M_2}) = M_1 M_2 \omega^4 - 2K(M_1 + M_2)\omega^2 + 2K^2(1 - \cos qa) \quad (8.27)$$

This is just a quadratic equation in ω^2 , so:

$$\omega_{12}^2 = \frac{2K}{M_1 M_2} \left((M_1 + M_2) \pm \sqrt{(M_1 + M_2)^2 - 2M_1 M_2(1 - \cos qa)} \right) \quad (8.28)$$

If we sketch the dispersion relation, we have:

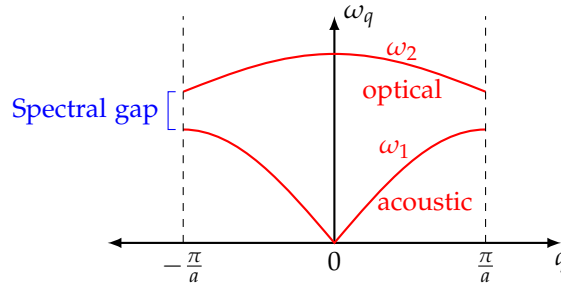


Figure 8.4: Dispersion relations for diatomic chain. Note the presence of two different branches/phonon modes, and the spectral gap between them.

The acoustic branch is linear and the basis of sound waves inside of the material. The optical branch is excited by shining light on the material. If we shine waves of an intermediate energy in the spectral gap (in between the two branches), none will propagate!

Next class, we will discuss phonons further; in particular how they behave in 3D, and thermodynamic properties of vibrating crystals.

9 Phonons in Three Dimensions

9.1 Review - The Real and Reciprocal 3D Lattice

We consider atoms located at positions \mathbf{R} of a Bravais lattice in 3D space. Such a Bravais lattice can be written via:

$$\mathbf{R} = n_1 \mathbf{a}_1 + n_2 \mathbf{a}_2 + n_3 \mathbf{a}_3 \quad (9.1)$$

where $n_i \in \mathbb{Z}$ and \mathbf{a}_i are primitive vectors.

The corresponding reciprocal lattice vectors \mathbf{G} are defined by $e^{i\mathbf{R} \cdot \mathbf{G}} = 1$, which implies:

$$\mathbf{G} = m_1 \mathbf{b}_1 + m_2 \mathbf{b}_2 + m_3 \mathbf{b}_3. \quad (9.2)$$

Where $m_j \in \mathbb{Z}$ and \mathbf{b}_j are primitive vectors of the reciprocal lattice, satisfying the usual relation:

$$\mathbf{a}_i \cdot \mathbf{b}_j = 2\pi \delta_{ij}. \quad (9.3)$$

We will also need a relation:

$$\sum_{\mathbf{R}} e^{i\mathbf{R} \cdot \mathbf{q}} = N \sum_{\mathbf{G}} \delta_{\mathbf{q}, \mathbf{G}} = N \Delta(\mathbf{q}) \quad (9.4)$$

9.2 Writing down the 3D Hamiltonian

The Hamiltonian for the 3d lattice with the harmonic approximation can be written as:

$$H = \sum_{\mathbf{R}, i} \frac{(p_{\mathbf{R}}^i)^2}{2M} + \frac{1}{2} \sum_{\mathbf{R}, \mathbf{R}'} \mu_{\mathbf{R}}^i V_{\mathbf{R}\mathbf{R}'}^{ij} \mu_{\mathbf{R}'}^j \quad (9.5)$$

This is the expected generalization from 1D, where we see that the atoms can move and vibrate in three dimensions; $i, j \in \{x, y, z\}$. $V_{\mathbf{R}\mathbf{R}'}^{ij}$ (as before) is the dynamical matrix $V_{\mathbf{R}\mathbf{R}'}^{ij} = \left. \frac{\partial^2 V}{\partial \mu_{\mathbf{R}}^i \partial \mu_{\mathbf{R}'}^j} \right|_{\mu=0}$. We define the displacement/momenta at each lattice site as:

$$\begin{aligned} \mathbf{Y}_{\mathbf{R}} &= \begin{pmatrix} \mu_{\mathbf{R}}^x \\ \mu_{\mathbf{R}}^y \\ \mu_{\mathbf{R}}^z \end{pmatrix} \\ \mathbf{P}_{\mathbf{R}} &= \begin{pmatrix} p_{\mathbf{R}}^x \\ p_{\mathbf{R}}^y \\ p_{\mathbf{R}}^z \end{pmatrix} \end{aligned} \quad (9.6)$$

And therefore the Hamiltonian becomes:

$$H = \frac{1}{2M} \sum_{\mathbf{R}} \mathbf{P}_{\mathbf{R}}^T \mathbf{P}_{\mathbf{R}} + \frac{1}{2} \sum_{\mathbf{R}, \mathbf{R}'} \mathbf{Y}_{\mathbf{R}}^T \begin{pmatrix} V_{\mathbf{R}\mathbf{R}'}^{xx} & V_{\mathbf{R}\mathbf{R}'}^{xy} & V_{\mathbf{R}\mathbf{R}'}^{xz} \\ V_{\mathbf{R}\mathbf{R}'}^{yx} & V_{\mathbf{R}\mathbf{R}'}^{yy} & V_{\mathbf{R}\mathbf{R}'}^{yz} \\ V_{\mathbf{R}\mathbf{R}'}^{zx} & V_{\mathbf{R}\mathbf{R}'}^{zy} & V_{\mathbf{R}\mathbf{R}'}^{zz} \end{pmatrix} \mathbf{Y}_{\mathbf{R}} \quad (9.7)$$

note that due to the definition of the dynamical matrix elements, the matrix appearing in the above expression is a real symmetric matrix.

9.3 Translation Invariant Solution

With the assumption of translation invariance, we obtain the condition $V_{\mathbf{R}\mathbf{R}'}^{ij} = V_{\mathbf{R}-\mathbf{R}'}^{ij}$. As we did for the 1D case, we will use a fourier transform:

$$\mathbf{Y}_{\mathbf{q}} = \frac{1}{\sqrt{N}} \sum_{\mathbf{R}} e^{-i\mathbf{q}\cdot\mathbf{R}} \mathbf{Y}_{\mathbf{R}}, \quad \mathbf{P}_{\mathbf{q}} = \frac{1}{\sqrt{N}} \sum_{\mathbf{R}} e^{i\mathbf{q}\cdot\mathbf{R}} \mathbf{P}_{\mathbf{R}} \quad (9.8)$$

As in 1D, we also have the periodicity:

$$\mathbf{Y}_{\mathbf{q}+\mathbf{G}} = \mathbf{Y}_{\mathbf{q}}, \quad \mathbf{P}_{\mathbf{q}+\mathbf{G}} = \mathbf{P}_{\mathbf{q}} \quad (9.9)$$

This defines the 3D Brillouin zone \mathcal{B} as the fundamental domain of \mathbf{q} . The Inverse FT gives:

$$\mathbf{Y}_{\mathbf{R}} = \frac{1}{\sqrt{N}} \sum_{\mathbf{q} \in \mathcal{B}} e^{i\mathbf{q}\cdot\mathbf{R}} \mathbf{Y}_{\mathbf{q}}, \quad \mathbf{P}_{\mathbf{R}} = \frac{1}{\sqrt{N}} \sum_{\mathbf{q} \in \mathcal{B}} e^{-i\mathbf{q}\cdot\mathbf{R}} \mathbf{P}_{\mathbf{q}} \quad (9.10)$$

The Hamiltonian reduces to:

$$H = \sum_{\mathbf{q} \in \mathcal{B}} \left(\frac{1}{2M} \mathbf{P}_{\mathbf{q}}^\dagger \mathbf{P}_{\mathbf{q}} + \frac{1}{2} \mathbf{Y}_{\mathbf{q}}^\dagger \begin{pmatrix} V_{\mathbf{q}}^{xx} & V_{\mathbf{q}}^{xy} & V_{\mathbf{q}}^{xz} \\ V_{\mathbf{q}}^{yx} & V_{\mathbf{q}}^{yy} & V_{\mathbf{q}}^{yz} \\ V_{\mathbf{q}}^{zx} & V_{\mathbf{q}}^{zy} & V_{\mathbf{q}}^{zz} \end{pmatrix} \mathbf{Y}_{\mathbf{q}} \right) \quad (9.11)$$

where the $3N \times 3N$ matrix in the original basis has now reduced to a 3×3 matrix. The matrix elements in this basis are:

$$V_{\mathbf{q}}^{ij} = \sum_{\mathbf{R}'} e^{i\mathbf{q}\cdot(\mathbf{R}-\mathbf{R}')} V_{\mathbf{R}\mathbf{R}'}^{ij}$$

To complete this solution, we have to diagonalize the 3×3 dynamical matrix. It is a Hermitian matrix and thus has three orthogonal vectors $\mathbf{s}_1, \mathbf{s}_2, \mathbf{s}_3$ belonging to eigenvalues $\mathbf{v}_{\mathbf{q}}^1, \mathbf{v}_{\mathbf{q}}^2, \mathbf{v}_{\mathbf{q}}^3$. In this basis defined by these three orthogonal vectors, we have:

$$H = \sum_{\mathbf{q}, \mu=1,2,3} \left(\frac{p_{\mathbf{q}}^{\mu\dagger} p_{\mathbf{q}}^{\mu}}{2M} + \frac{1}{2} V_{\mathbf{q}}^{\mu} \mu_{\mathbf{q}}^{\mu\dagger} \mu_{\mathbf{q}}^{\mu} \right) \quad (9.12)$$

where $\mu_{\mathbf{q}}^{\mu} = \boldsymbol{\mu}_{\mathbf{q}} \cdot \mathbf{s}_{\mu}$ and $p_{\mathbf{q}}^{\mu} = \mathbf{p}_{\mathbf{q}} \cdot \mathbf{s}_{\mu}$.

The 3 directions described by $\mathbf{s}_{\mu}(\mathbf{q})$ define phonon *polarization*. In addition, for \mathbf{q} along a high-symmetry axis, we typically have $\mathbf{s} \parallel \mathbf{q}$ which is a "longitudinal phonon" and $\mathbf{s} \perp \mathbf{q}$ which is two "transverse phonons".

Why a high symmetry axis? If we take an arbitrary \mathbf{q} pointing in some random direction in reciprocal space, in general none of the \mathbf{s} will be parallel to \mathbf{q} or orthogonal to it. But when the phonon propagates along a high symmetry axis, we do typically see this separation.

Phonon frequencies are given by:

$$\omega_{\mathbf{q}\mu} = \sqrt{\frac{V_{\mathbf{q}}^{\mu}}{M}} \quad (9.13)$$

and the corresponding raising/lowering operators are:

$$a_{\mathbf{q}\mu} = \frac{1}{\sqrt{2M\hbar\omega_{\mathbf{q}\mu}}} (M\omega_{\mathbf{q}\mu} \mu_{\mathbf{q}}^{\mu} + ip_{\mathbf{q}}^{\mu\dagger}), \quad a_{\mathbf{q}\mu}^{\dagger} = \dots \quad (9.14)$$

which leads to:

$$H = \sum_{\mathbf{q}, \mu} \hbar\omega_{\mathbf{q}\mu} (a_{\mathbf{q}\mu}^{\dagger} a_{\mathbf{q}\mu} + \frac{1}{2}) \quad (9.15)$$

One more comment before moving onto the next topic; this was all for a simple Bravais lattice. This is the simplest possible type of crystal structure where you have one type of atom periodically repeating in space. Most solids are not quite so simple and have a larger unit cell, such as Sodium Chloride. Things would work out exactly the same in this case, there would just be an additional index labelling the position of the atom inside the unit cell (of which there are n_b). The 3×3 matrix would become a larger $3n_b \times 3n_b$ matrix with $3n_b$ eigenvalues/eigenvectors, which describe the internal degrees of freedom in the unit cell. Previously, we had 3 polarizations for the one atom type. Now we have 3 polarizations per atom type, so there is a more complex structure with more modes (e.g. different atoms can vibrate in different directions). Of these, 3 will be acoustic modes (frequency goes to zero as $\mathbf{q} \rightarrow \mathbf{0}/\lambda \rightarrow \infty$), and $3(n_b - 1)$ will be optical modes.

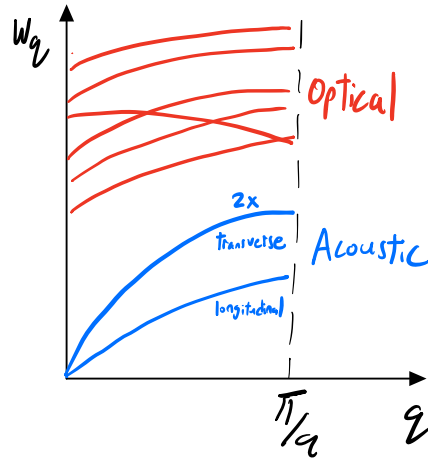


Figure 9.1: Plot of the phonon dispersion curves for a complex 3d solid. We have 3 acoustic modes (1 longitudinal mode and 2 degenerate transverse modes) and $3(n_d - 1)$ optical modes.

9.4 Debye Model, Specific Heat of Phonons

This was one of the main puzzles in solid-state physics before the advent of QM! Let's figure out how to calculate this. The internal energy in lattice vibrations is given by:

$$U = \langle H \rangle_\beta = \sum_{\mathbf{q}\mu} \hbar \omega_{\mathbf{q}\mu} \left(\langle a_{\mathbf{q}\mu}^\dagger a_{\mathbf{q}\mu} \rangle + \frac{1}{2} \right) \quad (9.16)$$

Where $\langle a_{\mathbf{q}\mu}^\dagger a_{\mathbf{q}\mu} \rangle = \bar{n}_{\mathbf{q}\mu} = \frac{1}{e^{\beta \hbar \omega_{\mathbf{q}\mu}} - 1}$ is the bose-einstein occupation factor. It will actually be easier to go straight to the heat capacity:

$$C_V = \frac{dU}{dT} = \frac{1}{k_B T^2} \sum_{\mathbf{q},\mu} \frac{(\hbar \omega_{\mathbf{q}\mu})^2 e^{\beta \hbar \omega_{\mathbf{q}\mu}}}{(e^{\beta \hbar \omega_{\mathbf{q}\mu}} - 1)^2} \quad (9.17)$$

to evaluate the sum we would need to know the form of $\omega_{\mathbf{q}\mu}$, and even then there usually does not exist a closed-form solution for the summation.

To evaluate Eq. (9.17), it is useful to define the phonon density of states:

$$D(\omega) = \sum_{\mathbf{q},\mu} \delta(\omega - \omega_{\mathbf{q}\mu}) \quad (9.18)$$

This implies:

$$C_V = \frac{1}{k_B T^2} \int_0^\infty d\omega D(\omega) \frac{(\hbar\omega)^2 e^{\beta\hbar\omega}}{(e^{\beta\hbar\omega} - 1)^2} \quad (9.19)$$

The Debye model assumes $\omega_{\mathbf{q}\mu} = c_\mu |\mathbf{q}|$, approximating the acoustic modes as straight lines. It should be accurate at low T when only low frequency acoustic phonons are thermally excited. For simplicity, we further assume equal velocities $c_\mu = c$ for $\mu = 1, 2, 3$ (but this is not essential, and the calculation can still be done). We then get:

$$\begin{aligned} D(\omega) &= \frac{3V}{(2\pi)^3} \int d^3q \delta(\omega - cq) \\ &= \frac{3V}{(2\pi)^3} 4\pi \int q^2 dq \delta(\omega - cq) \\ &= \frac{3V}{3\pi^2} \frac{\omega^2}{c^3}. \end{aligned} \quad (9.20)$$

where the 3 comes from the 3 directions summed over and we evaluate the integral by going into spherical coordinates.

9.5 Debye Frequency, Momentum, Temperature

There is one more constraint for us to accomodate. The density of states has the property that if we integrate over it, we should get the total number of modes in the entire system (which in this case should be $3N$; N particles moving in three dimensions). Therefore, the above expression cannot go on forever; and this is evident from the acoustic spectra where we see there are no states above some energy. So, we introduce a Debye frequency ω_D :

$$D(\omega) = \begin{cases} \frac{3V}{2\pi^2} \frac{\omega^2}{c^3} & \Omega < \omega_D \\ 0 & \omega > \omega_D \end{cases} \quad (9.21)$$

where ω_D is determined by:

$$\int_0^{\omega_D} D(\omega) d\omega = 3N \implies \omega_D^3 = 4\pi^2 c^3 \frac{N}{V} \quad (9.22)$$

,

This notion of a Debye frequency turns out to be very useful. It also defines Debye momentum:

$$k_D = \frac{\omega_D}{c} = \left(6\pi^2 \frac{N}{V} \right)^{1/3} \sim \frac{4}{a} \quad (9.23)$$

where a is the lattice spacing. It also defines the Debye temperature:

$$\Theta_D = \frac{\hbar\omega_D}{k_B} \sim 74\text{K} - 1800\text{K} \quad (9.24)$$

where 74K corresponds to Pr and 1800K corresponds to diamond. On average it is on the order of a few hundred Kelvin. These are all characteristic scales of phonons. Of all of them, Debye temperature tends to be the most useful as it is in the most tangible units. It distinguishes a low and high temperature regime for phonon excitations (i.e. the frequency for which below it the acoustic approximation is reasonable)

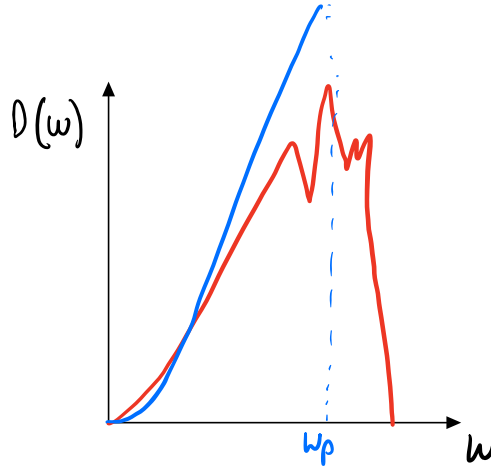


Figure 9.2: Plot of a realistic/experimentally measured phonon density of states (red) and our approximate calculated value (blue). We cut off the ω^2 prediction for the density of the states at the Debye frequency ω_D such that the area under the curve (i.e. the total number of modes) is $3N$ (and the same for the two plots). It may seem like we neglect a lot of structure in making our approximation, but since we are interested in the heat capacity C_V , this complex structure generally gets smeared out regardless; hence our approximation can give reasonable predictions for heat capacity.

9.6 Back to Heat Capacity

We have:

$$\begin{aligned} C_V &= \frac{3V\hbar^2}{2\pi^2 c^2 k_B T^2} \int_0^{\omega_D} d\omega \frac{\omega^4 e^{\beta\hbar\omega}}{(e^{\beta\hbar\omega} - 1)^2} \\ &= g N k_B \left(\frac{T}{\theta_D} \right)^3 \int_0^{\theta_D/T} dx \frac{x^4 e^x}{(e^x - 1)^2} \end{aligned} \quad (9.25)$$

where we have made the substitution $x = \beta\hbar\omega$. We call the integral $f(\theta_D/T) = f(x_D)$ the Debye function.

We now analyze some consequences. In principle we can solve the integral numerically, but we can also start by studying the integral analytically in two limits.

9.6.1 Low T behavior

In this limit we have $T \ll \theta_D$ and hence $x_D = \frac{\theta_D}{T} \gg 1$. We can therefore write:

$$f(x_D) = \int_0^\infty \frac{x^4 e^x}{(e^x - 1)^2} - \int_{x_D}^\infty \frac{x^4 e^x}{(e^x - 1)^2} \approx \frac{4\pi^4}{15} - \int_{x_D}^\infty x^4 e^{-x} \approx \frac{4\pi^4}{15}$$

where we have evaluated the first term analytically (calculus exercise) and the second integral we have that $x \gg 1$ over the range of integration and we may therefore approximate $\frac{x^4 e^x}{(e^x - 1)^2} \approx x^4 e^{-x}$. Since this is an exponentially small term, we can drop it. The specific heat then becomes:

$$C_V \approx \frac{12\pi^4}{5} N k_B \left(\frac{T}{\theta_D} \right)^3 \quad (9.26)$$

so for $T \ll \theta_D$ we find $C_V \sim T^3$. This is obeyed in many solids.

9.6.2 High T behavior

In this limit, we have $T \gg \theta_D$ and so $x_D \ll 1$. We therefore may expand e^x to leading order in the integrand. This yields:

$$f(x_D) \approx \int_0^{x_D} dx \frac{x^4}{(1+x-1)^2} = \int_0^{x_D} dx x^2 = \frac{1}{3} x_D^3$$

where in the numerator we have approximated $e^x \sim 1$ and in the denominator we have approximated $e^x \sim 1 + x$.

So plugging this back into our specific heat expression:

$$C_V \approx 3Nk_B \quad (9.27)$$

so for $T \gg \theta_D$ we find that C_V is independent of temperature. This is the Dulong-Petit law; this is the heat capacity of a harmonic crystal in classical theory (i.e. no quantum effects). If we recall the equipartition theorem, we have N atoms oscillating in 3 directions, so 3 vibrational degrees of freedom and 3 kinetic degrees of freedom and so $6N$ degrees of freedom in total. Equipartition associates $\frac{1}{2}k_B T$ per energy per degree of freedom and so we have $U = 6N \cdot \frac{1}{2}k_B T = 3Nk_B T$ of energy and hence $C_V = 3Nk_B$ specific heat. This of course in some sense makes sense as at high temperatures we expect quantum effects to be washed out and the system behaves basically classically (though an interesting point to note here - the quantum effect of the suppression of specific heat persists up to room temperature, as θ_D is on the order of hundreds of Kelvin; we do not have to cool our system to very low temperature for the quantum effects to become significant). This was known before QM was invented, but disagreed with experiment, where the specific heat went to zero at zero temperature (though it described the high- T behavior well). With quantum mechanics, we are able to obtain a prediction that described experimental results very well.

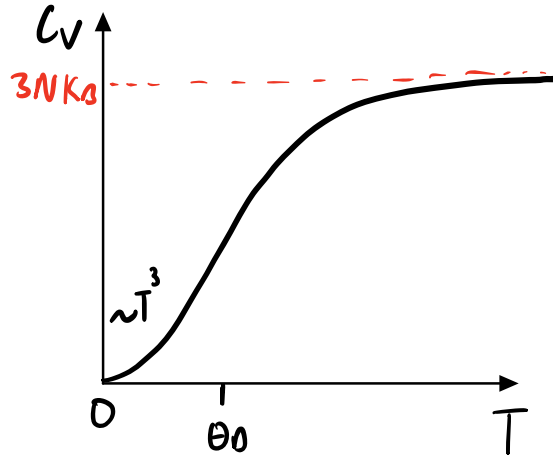


Figure 9.3: Specific heat for phonons as a function of temperature T . At low T we see $C_V \sim T^3$ behavior. At high T we see that C_V tends to a constant of $3Nk_B$.

9.7 Einstein Model

Is a model suitable for the study of optical phonons. We replace the dispersion relation for an optical phonon with an average ω_0 .

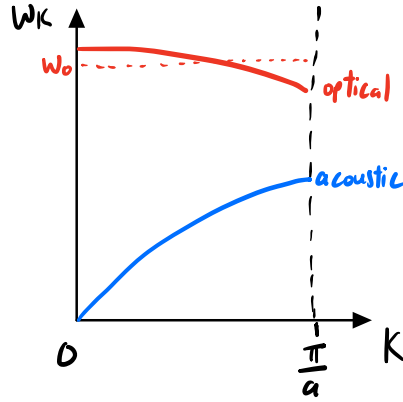


Figure 9.4: Dispersion relation for acoustic and optical modes of phonons, and Einstein's approximation for replacing the optical mode dispersion relation with a straight line at the average ω_0 .

We consider an approximate phonon density of states:

$$D(\omega) = N\delta(\omega - \omega_0)$$

and from this we can calculate contribute to the heat capacity:

$$C_V = \frac{1}{k_B T^2} \int_0^\infty d\omega \frac{(\hbar\omega)^2 e^{\beta\hbar\omega}}{(e^{\beta\hbar\omega} - 1)^2} D(\omega) = \frac{N}{k_B T^2} \frac{(\hbar\omega_0)^2 e^{\beta\hbar\omega_0}}{(e^{\beta\hbar\omega_0} - 1)^2} \quad (9.28)$$

At low T , we have $\beta\hbar\omega_0 \gg 1$ and so:

$$C_V \approx Nk_B \left(\frac{\hbar\omega_0}{k_B T} \right)^2 e^{-\frac{\hbar\omega_0}{k_B T}} \quad (9.29)$$

so we have “exponentially activated behavior” here.

Interestingly, if we compare the two plots for the specific heat, they agree pretty well. Some differences at the very low T limits, however. For the low T einstein model at very low temperatures we do not even have one quantum of energy to excite, and so we have $C_V = 0$. Conversely, for the acoustic modes we are able to excite things at even very low temperatures (so long as the temperature is not zero).

9.8 Anharmonic Effects and Phonon Interactions

We take a step back and regard a solid (metal) as a collection of electrons (fermions) and phonons (bosons).

At low T , both the electrons and phonons contribute to thermodynamic (and other) properties, e.g. specific heat:

$$C_V^{ph} \approx 6T^3 \text{ (Debye)}$$

$$C_V^{el} \approx aT \text{ (Sommerfield)}$$

both are crucially quantum mechanical; if we forget quantum mechanics then these become temperature dependent constants. The total specific heat is the sum of the two:

$$C_V = C_V^{el} + C_V^{ph} \approx aT + bT^3$$

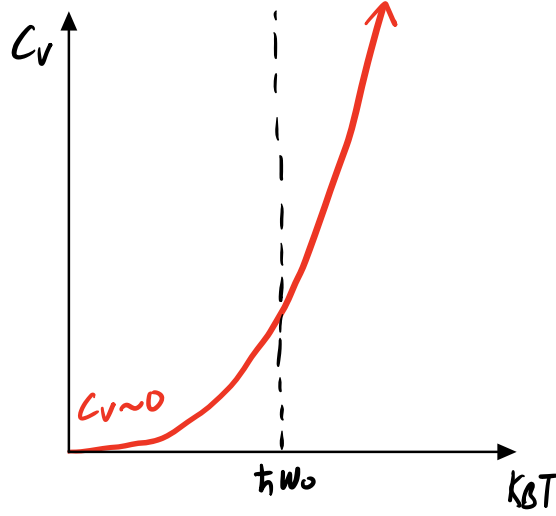


Figure 9.5: Plot of the low- T specific heat as calculated by the Einstein model. We see the exponentially activated behavior, where past $\hbar\omega_0$ the specific heat sharply spikes.

or:

$$\frac{C_V}{T} \approx a + bT^2 \quad (9.30)$$

and many (most) metals behave in this way. This is a huge success of quantum theory and modern solid-state physics.

However, there are some significant failures of this model:

1. Thermal expansion (Within the harmonic expansion, you can prove this does not occur)
2. Thermal conductivity (Phonons to not carry heat in this model)
3. Sound attenuation (Sound attenuates forever)

but whenever our theories fail, we can look back to what approximations we have made, and see what things to improve. These three failures can be understood by incorporating anharmonic effects. In the harmonic approximation, we threw away all terms past the second derivative, but we could consider successive terms of order:

$$H_3 = \frac{1}{3!} \sum_{RR'R''} \mu_R^i \mu_{R'}^j \mu_{R''}^k V_{RR'R''}^{ijk} \quad (9.31)$$

where:

$$V_{RR'R''}^{ijk} = \left. \frac{\partial^3 V}{\partial \mu_R^i \partial \mu_{R'}^j \partial \mu_{R''}^k} \right|_{\mu=0} \quad (9.32)$$

Note that the inclusion of these third-order terms is sufficient to describe thermal expansion and sound attenuation. To obtain thermal conductivity from the model, one needs to go to fourth order.

In translation-invariant systems, this can be expressed as:

$$H_3 = \frac{1}{6\sqrt{N}} \left(\frac{\hbar}{2M} \right)^{3/2} \sum_{qq'q''ijk\mu\nu\lambda} \frac{S_\mu^i S_\nu^j S_\lambda^k}{(\omega_{q\mu} \omega_{q'\nu} \omega_{q''\lambda})^{1/2}} V_{qq'q''}^{ijk} \Delta(\mathbf{q} + \mathbf{q}' + \mathbf{q}'') (a_{-q\mu}^\dagger + a_{q\mu}) (a_{-q'\nu}^\dagger + a_{q'\nu}) (a_{-q''\lambda}^\dagger + a_{q''\lambda}) \quad (9.33)$$

This looks absolutely terrible, but people have come up with clever ways to analyze such Hamiltonians, namely through the graphical means of Feynman diagrams. Each of the product of three terms contains an annihilation and creation operator. If we were to multiply them out, we would get six terms, which could be represented as follows:

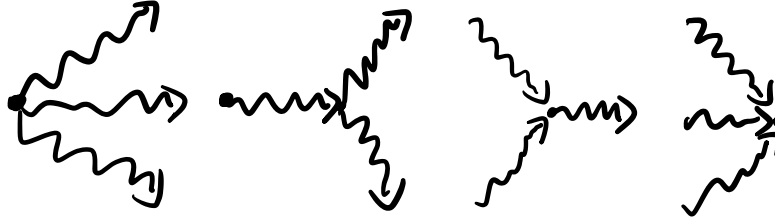


Figure 9.6: The four Feynman diagrams for three phonon interactions that come up in the third order term. In each diagram time runs from left to right. The leftmost diagram represents the creation of three phonons. The second diagram represents the destruction of one phonon and creation of two phonons. The third diagram represents the destruction of two phonons and creation of one phonon. Finally, the fourth diagram represents the destruction of three phonons.

we can then assign a mathematical quantity to each of these diagrams, and thus evaluate each of their contributions in a clever way.

Without writing down the Hamiltonian, we can also look at the diagrams that arise in the fourth-order Hamiltonian H_4 . We would have a product of four terms, which creates a sum with four phonons. Again, each of these diagrams could be assigned a mathematical quantity to evaluate their contribution.



Figure 9.7: Feynman diagrams for four phonon interactions that come up in the fourth order term. In each diagram time runs from left to right. The leftmost diagram represents the most dominantly contributing interaction, which is the scattering of two phonons. There are other interactions that also contribute, for example the destruction of one phonon and creation of three phonons as depicted in the right diagram.

9.9 Thermal Conductivity

In Drude theory, electrons contribute:

$$\kappa_{el} = \frac{1}{3} C_V^e v_F l \quad (9.34)$$

with C_V^{el} the electronic specific heat, v_F the Fermi velocity, and l is the mean-free path - this is where many different details and factors come in (and generically very hard to calculate); this is where anharmonic effects enter. An electron that travelled forever would contribute an infinite conductivity - this does not happen so there must be features in our metal that slow it down, e.g. impurities. For phonons, we have (analogously)

$$\kappa_{ph} = \frac{1}{3} C_V^{ph} c l' \quad (9.35)$$

and again l' the mean free path is where anharmonic effects enter (e.g. collisions of phonons with electrons).

10 Magnons

Magnons occur as excitations in spin Hamiltonians describing magnetic systems.

10.1 The Heisenberg Model

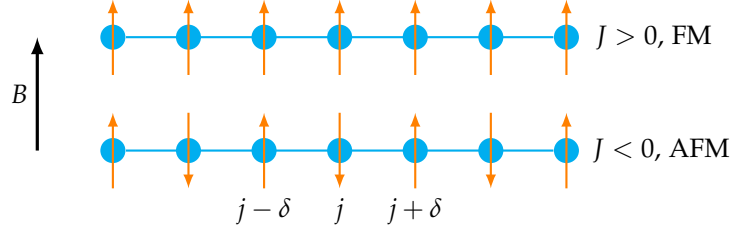


Figure 10.1: Cartoon depiction of the 1-D Heisenberg chain. For $J > 0$ we have a ferromagnet with aligned spins. For $J < 0$ we have an antiferromagnet with alternating spins. Nearest neighbour spins are coupled, and there is an external magnetic field that affects each spin.

Here, we will discuss the simplest such Hamiltonian describing a magnetic system, the Heisenberg model. It is written as:

$$H = -J \sum_{j,\delta} \mathbf{S}_j \cdot \mathbf{S}_{j+\delta} - 2B \sum_j S_j^z \quad (10.1)$$

Where the first term represents the coupling between spins, parameterized by a coupling constant J (where $J > 0$ gives rise to a ferromagnet and $J < 0$ to an antiferromagnet), and the second term represents the contribution from an external magnetic field with strength B . Note that this Hamiltonian is completely invariant under the choice of z -axis; more formally, it is invariant under $SU(2)$ rotations.

Note that $\mathbf{S}_j = (S_j^x, S_j^y, S_j^z)$ is a spin operator at site j that obeys:

$$[S_i^\alpha, S_j^\beta] = i\delta_{ij}\epsilon^{\alpha\beta\lambda}S_j^\lambda \quad (10.2)$$

The second property is that $\mathbf{S}_j \cdot \mathbf{S}_j$ has eigenvalues $s(s+1)$ where $s = \frac{1}{2}, 1, \frac{3}{2}, \dots$. Note that despite its simple form, the Heisenberg Hamiltonian is generally not solvable; it is only analytically solvable (in a difficult method; Bethe ansatz) in 1D and treatable via approximate methods in higher dimensions. In our class, we will solve the 1D chain approximately

10.2 Magnon Variables

We will proceed with the approximate solve by transforming into “magnon variables” using the Holstein-Primakoff transformation. Recall the raising/lowering operators:

$$S_i^\pm = S_i^x \pm iS_i^y$$

In Holstein-Primakoff, these operators are expressed using bosonic operators:

$$\begin{aligned} S_j^+ &= \sqrt{2S}(1 - n_j/2s)^{1/2}a_j \\ S_j^- &= \sqrt{2S}a_j^\dagger(1 - n_j/2s)^{1/2} \end{aligned} \quad (10.3)$$

with $[a_j, a_k^\dagger] = \delta_{jk}$ and $n_j = a_j^\dagger a_j$ the number operator. This seems like a strange, nonlinear transformation, but in fact is necessary to satisfy the $SU(2)$ algebra. Raising/lowering operators allow for raising/lowering

the states to infinity, but spins have a finite number of states. For certain eigenvalues, $n_j = 2s$ and so we get zero, i.e. we cannot raise the spin state beyond its maximal value. We have S^+ , S^- , and to finish the set we have:

$$S_z = s(s-1) - (S^x)^2 - (S^y)^2 = s(s-1) - \frac{1}{2}(S^+S^- + S^-S^+) = S - a_j^\dagger a_j \quad (10.4)$$

Note that the transformations Eqs. (10.3), (10.4) is *exact* in that it satisfies the original spin algebra⁵ Eq. (10.2). However, it is not very useful as written due to the $\sqrt{\cdot}$ in (10.3). However, it is useful when we study excited states close to the ordered ground state of H , where the spins are completely saturated. In this limit, we assume that only a small number of spins deviate from their perfect arrangement, that is:

$$s - \langle S_j^z \rangle = \langle a_j^\dagger a_j \rangle \quad (10.5)$$

is small, and so:

$$\frac{\langle n_j \rangle}{s} \ll 1. \quad (10.6)$$

We can therefore expand the square roots in Eq. (10.3):

$$\begin{aligned} S^+ &= \sqrt{2S} \left[a_j - \frac{n_j}{4s} + \dots \right] \\ S^- &= \sqrt{2S} \left[a_j^\dagger - a_j^\dagger \frac{n_j}{4s} + \dots \right] \end{aligned} \quad (10.7)$$

and retain only the leading term in this expansion, and write:

$$\mathbf{S}_i \mathbf{S}_j = \frac{1}{2}(S_i^+ S_j^- + S_i^- S_j^+) + S_i^z S_j^z \approx s(a_i^\dagger a_j + a_i a_j^\dagger) + s - n_i - n_j \quad (10.8)$$

from here, the ferromagnetic and anti-ferromagnetic lines diverge, so let's analyze the two cases separately.

10.3 Ferromagnetic Case

Here we have $J > 0$, and we write the Hamiltonian as:

$$H = -JS \sum_{\langle i,j \rangle} (a_i^\dagger a_j + a_i a_j^\dagger + s - n_i - n_j) - 2V \sum_j (s - n_j) \quad (10.9)$$

where $\langle i, j \rangle$ means that i, j are nearest neighbour sites. We can solve this by Fourier transforming:

$$\begin{aligned} b_k &= \frac{1}{\sqrt{N}} \sum_j e^{i\mathbf{k} \cdot \mathbf{r}_j} a_j \\ b_k^\dagger &= \frac{1}{\sqrt{N}} \sum_j e^{-i\mathbf{k} \cdot \mathbf{r}_j} a_j^\dagger \end{aligned} \quad (10.10)$$

The momentum-space H assumes the form:

$$H = -JNzs^2 - 2BNs + \mathcal{H}_0 + \mathcal{H}_1 \quad (10.11)$$

where z denotes the coordination number (the number of nearest neighbours) and:

$$\mathcal{H}_0 = \sum_k (2Jzs(1 - \gamma_k) + 2B) b_k^\dagger b_k \quad (10.12)$$

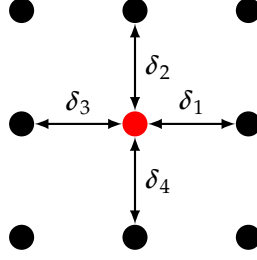


Figure 10.2: The four δ vectors for the 2D square lattice. Each atom has neighbours at $\delta = \pm\hat{x}$, $\delta = \pm\hat{y}$.

with $\gamma_k = \frac{1}{z} \sum_{\delta} e^{i\mathbf{k} \cdot \delta}$. If we had a two-dimensional square lattice, then the δ vectors would be as follows:

Note that this form of \mathcal{H}_0 is valid for lattices with an inversion center which implies $\gamma_{\mathbf{k}} = \gamma_{-\mathbf{k}}$. Without this condition the form of \mathcal{H}_0 would be more complicated. One way to think about this is a lattice has an inversion center if for every lattice site, there is both a δ and $-\delta$ nearest neighbour. This would not be the case (e.g.) for a honeycomb lattice.

\mathcal{H}_1 contains higher order terms in b_k, b_k^\dagger and represents magnon interactions.

The situation is analogous to phonons; we made a harmonic approximation, which gave us a nice quadratic Hamiltonian. The higher order terms represented interactions of the phonons. The mechanics is different here but the idea is the same; we have a “nice” quadratic Hamiltonian \mathcal{H}_0 and then the higher order terms in \mathcal{H}_1 representing magnon interactions. Similar to phonons where the harmonic approx was sufficient to describe lattice vibrations but not expansions, we will find that thermodynamics is described well by the lower-order expansion, but (e.g.) thermal conduction will require the higher order terms to analyze.

We write things suggestively as:

$$\begin{aligned} \mathcal{H}_0 &= \sum_k \omega_k n_k \\ \omega_k &= 2Js(1 - \gamma_k) + 2B \end{aligned} \quad (10.13)$$

where ω_k is the magnon spectrum. This tells us about the low energy excitations.

10.4 An Example: Cubic Lattice in 3D

In this example, we have:

$$\delta = \pm a\hat{x}, \pm a\hat{y}, \pm a\hat{z}. \quad (10.14)$$

So then:

$$z(1 - \gamma_{\mathbf{k}}) = 6 - \sum_{\delta} e^{i\mathbf{k} \cdot \delta} = 2(3 - \cos(ak_x) - \cos(ak_y) - \cos(ak_z))$$

We are interested in the low-temperature/energy and hence long wavelength excitations, so $ka \ll 1$ and so we can expand $\cos(k_i a) \approx 1 - \frac{1}{2}k_i^2 a^2$. We then find:

$$\omega_k \approx 2B + 2Js(ka)^2 \quad (10.15)$$

note that the same result holds for FCC and BCC lattices. At zero magnetic field, the magnons exhibit particle like spectra:

$$\omega_k = \frac{k^2}{2m^*}, \quad m^* = \frac{1}{4Jsa} \quad (10.16)$$

⁵“It is a bitch to work out.” - Marcel Franz 2022

for conventional ferromagnets, it is found that $m^* \approx 10m_e$.

To find the magnon heat capacity, we take $\omega_k = Dk^2$ with $D = 2sJa^2$ and calculate the internal energy $U(T)$:

$$\begin{aligned}
 U(T) &= \sum_{\mathbf{k}} \omega_{\mathbf{k}} \langle n_{\mathbf{k}} \rangle_T = \sum_{\mathbf{k}} \frac{\omega_{\mathbf{k}}}{e^{\beta\omega_{\mathbf{k}}} - 1} \\
 &= \frac{1}{(2\pi)^3} \int_{|\mathbf{k}| < k_{max}} d^3k \frac{Dk^2}{e^{\beta Dk^2} - 1} \\
 &= \frac{(k_B T)^{5/2}}{4\pi^2 D^{3/2}} \int_0^{x_m} dx \frac{x^{3/2}}{e^x - 1} \quad x = \beta Dk^2, x_m = D\beta k_{max}^2 \\
 &= \frac{0.45}{\pi^2} \frac{(k_B T)^{5/2}}{D^{3/2}}
 \end{aligned} \tag{10.17}$$

The integral is dimensionless, so the temperature dependence is entirely in the prefactor. In the last line we have taken the $x_m \rightarrow \infty$ to carry out the integral, an approximation that is valid at low temperatures. From this we easily obtain the heat capacity:

$$C_V = \frac{dU}{dT} = 0.113 k_B (k_B T / D)^{3/2} \tag{10.18}$$

we have another different temperature dependence of the heat capacity! For electrons, $C_V \sim T$, for phonons, $C_V \sim T^3$, and for magnons, we have $C_V \sim T^{3/2}$. This suggests that if we have an insulating (i.e. no electronic contribution) antiferromagnet, to observe C_V we include a phonon contribution $\sim bT^3$ which means that:

$$C_V^{tot} = cT^{3/2} + bT^3$$

we can then plot $C_V^{tot} / T^{3/2}$ vs. $T^{3/2}$, which gives us a straight line with intercept c as:

$$\frac{C_V^{tot}}{T^{3/2}} = c + bT^{3/2}.$$

Reading Assignment: Antiferromagnetic magnons - see pages 58-62 in the handout (Kittel's *Quantum Theory of Solids*).

10.5 Magnetization Reversal

We expect the total magnetization:

$$M_s = 2\mu_0 \sum_j \langle S_j^z \rangle$$

to decrease as T is raised and more magnons are thermally excited. We can calculate this by going into magnon variables:

$$M_S(T) = 2\mu_0 \left(NS - \sum_{\mathbf{k}} \langle b_{\mathbf{k}}^\dagger b_{\mathbf{k}} \rangle \right) \tag{10.19}$$

We are interested in $\Delta M(T) = M_S(0) - M_S(T)$:

$$\Delta M(T) = 2\mu_0 \sum_{\mathbf{k}} \langle \hat{n}_{\mathbf{k}} \rangle = \frac{2\mu_0 V}{(2\pi)^d} \int_0^{k_{max}} d^d k \frac{1}{e^{\beta Dk^2} - 1} \tag{10.20}$$

Let's calculate the integral. Note that k_{max} is the momentum cutoff analogous to the Debye momentum for phonons. The integrand only depends on the magnitude of \mathbf{k} , so we can carry out the angular integrals to give us the surface S_d of the d -dimensional unit sphere (e.g. $S_3 = 4\pi, S_2 = 2\pi$):

$$\Delta M(T) = S_d V \int_0^{k_{max}} dk k^{d-1} \frac{1}{e^{\beta D k^2} - 1}$$

substituting $x = \beta D k^2$, $dx = 2\beta D k dk$ we find:

$$\begin{aligned} \Delta M(T) &= S_D V \int_0^{x_m} \frac{dx}{\beta D} \left(\frac{x}{\beta D} \right)^{\frac{d-2}{2}} \frac{1}{e^x - 1} \\ &= S_D V \left(\frac{k_B T}{D} \right)^{d/2} \int_0^{x_m} \frac{x^{(d-2)/2}}{e^x - 1} dx \end{aligned} \quad (10.21)$$

The behaviour near $x = 0$ is interesting. Because $e^x \sim 1 + x$ for $x \ll 1$, the integral diverges at the lower bound when $d \leq 2$. This implies that thermal fluctuations tend to destabilize ferromagnetic order in low dimensions (that is to say, $d \leq 2$). This of course is an approximate treatment, but you can analyze this much more carefully and this is indeed a correct result. Strictly speaking, in 2D ferromagnetic order only exists at zero temperature, and as soon as you raise the temperature above zero, you get long wavelength fluctuations and the average magnetization is immediately zero. The same is true of 1 dimension. In three dimensions this is not a problem, as the integral is not convergent. At $T = 0$ we have a perfectly ordered ferromagnetic, and as we increase the temperature the magnetization decreases, as $\Delta M(T) \sim T^{3/2}$. This is experimentally observed.

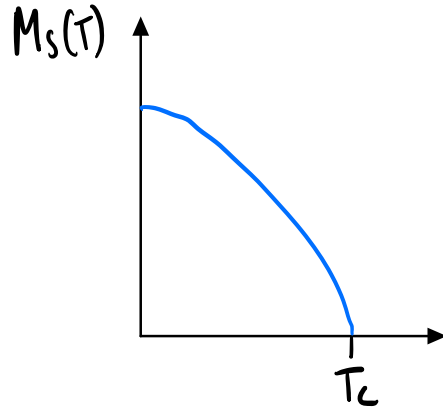


Figure 10.3: Plot of the total magnetization $M_S(T)$ as a function of temperature T . The deviation from total magnetization goes as $T^{3/2}$, and the total magnetization goes to zero at the critical (Curie) temperature T_c .

A concrete example that you may try at home: Fridge magnets lose their magnetic properties when heated up. Interestingly, when you cool them back down, you find that the magnetization is still zero. This is because the rise of domains within the magnet when cooled past the critical temperature, which makes the net average magnetization zero. It can however be re-magnetized in the presence of a second magnet.

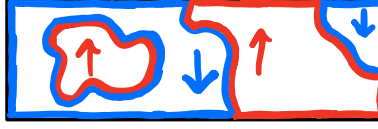


Figure 10.4: Cartoon sketch of the magnetic domains that appear in a bar magnet when cooling it back down from the critical temperature. As a result, the total magnetization is zero.

11 Electrons in a Periodic Potential: Band Theory of Solids

11.1 Review: Bloch's Theorem

See A&M Ch. 8 for proof(s) and extended discussion.

Theorem (Bloch). Consider the eigenstates ψ of the one-electron Hamiltonian:

$$H = \frac{\hbar^2 \nabla^2}{2m} + U(\mathbf{r})$$

where $U(\mathbf{r} + \mathbf{R}) = U(\mathbf{r})$ for all \mathbf{R} in the Bravais lattice. These eigenstates can be chosen to have the form:

$$\psi_{nk}(\mathbf{r}) = e^{i\mathbf{k} \cdot \mathbf{r}} \mu_{nk}(\mathbf{r}) \quad (11.1)$$

where:

$$\mu_{nk}(\mathbf{r} + \mathbf{R}) = \mu_{nk}(\mathbf{r}) \quad (11.2)$$

The interesting part of this statement is that even though the Hamiltonian has a full symmetry under translation by $+\mathbf{R}$. The eigenstates do not; there is a part that possesses the symmetry $\mu_{nk}(\mathbf{r})$ and a part that does not, $e^{i\mathbf{k} \cdot \mathbf{r}}$ which is translation invariant. This should not be totally unexpected; for example the QHO is symmetric under inversion, but the wavefunctions do not have all of this symmetry (n even is even, n odd is odd).

A remark: Note that Eqs. (11.1) and (11.2) imply that:

$$\psi_{nk}(\mathbf{r} + \mathbf{R}) = e^{i\mathbf{k} \cdot \mathbf{R}} \psi_{nk}(\mathbf{r}). \quad (11.3)$$

In Eq. (11.1), \mathbf{k} is known as a crystal momentum and n is the band index. In free space, complete translation invariance implies the conservation of momentum. In a lattice, we have translation invariance w.r.t the Bravais lattice \mathbf{R} , only, which implies that momentum is not conserved, but the crystal momentum is conserved.

11.2 Weak Periodic Potential

This is also known as “nearly free” electrons in a periodic lattice. We consider the same Hamiltonian $H = \frac{\hbar^2 \nabla^2}{2m} + U(\mathbf{r})$ where $U(\mathbf{r})$ is “weak”, sufficiently weak enough to be treated in perturbation theory. We call $H_0 = \frac{\hbar^2 \nabla^2}{2m}$ and $H' = U(\mathbf{r})$.

First, we transform H into second-quantized notation:

$$\begin{aligned} H_0 &= \sum_{\mathbf{k}, \sigma} \epsilon_{\mathbf{k}} c_{\mathbf{k}\sigma}^\dagger c_{\mathbf{k}\sigma}, \quad \epsilon_{\mathbf{k}} = \frac{\hbar^2 \mathbf{k}^2}{2m} \\ H' &= \sum_{\mathbf{k}, \mathbf{k}', \sigma, \sigma'} \langle \mathbf{k}\sigma | U | \mathbf{k}'\sigma' \rangle c_{\mathbf{k}\sigma}^\dagger c_{\mathbf{k}'\sigma'} \end{aligned} \quad (11.4)$$

where we note that as usual we work in the plane wave basis $\psi_{\mathbf{k}} = \frac{1}{\sqrt{V}}e^{i\mathbf{k}\cdot\mathbf{r}}$, and so:

$$\sum_{\mathbf{k}\mathbf{k}'\sigma\sigma'} \langle \mathbf{k}\sigma | U | \mathbf{k}'\sigma' \rangle = \frac{1}{V} \delta_{\sigma\sigma'} \int d^3r e^{-i\mathbf{r}\cdot(\mathbf{k}-\mathbf{k}')} U(\mathbf{r}) \quad (11.5)$$

Note that due to its periodic property, U can be written as:

$$U(\mathbf{r}) = \sum_{\mathbf{G}} e^{i\mathbf{r}\cdot\mathbf{G}} U_{\mathbf{G}} \quad (11.6)$$

where \mathbf{G} are reciprocal lattice vectors satisfying $e^{i\mathbf{G}\cdot\mathbf{R}} = 1$ for all \mathbf{R} in the Bravais lattice (check)! $U_{\mathbf{G}}$ is the fourier transform of our potential, evaluated at \mathbf{G} .

We further assume that $U_{\mathbf{G}=0} = 0$, which only redefines the overall energy zero. We can therefore write:

$$\langle \mathbf{k}\sigma | U | \mathbf{k}'\sigma' \rangle = \frac{\delta_{\sigma\sigma'}}{V} \int d^3r \sum_{\mathbf{G}} U_{\mathbf{G}} e^{-i\mathbf{r}\cdot(\mathbf{k}-\mathbf{k}'-\mathbf{G})} = \delta_{\sigma\sigma'} \sum_{\mathbf{G}} U_{\mathbf{G}} \delta_{\mathbf{k}-\mathbf{k}',\mathbf{G}} \quad (11.7)$$

when the dust settles, we find:

$$H' = \sum_{\mathbf{k}\mathbf{G}\sigma} U_{\mathbf{G}} c_{\mathbf{k}+\mathbf{G}\sigma}^\dagger c_{\mathbf{k}\sigma} \quad (11.8)$$

which is a very suggestive rewrite; for each term corresponds to the destruction of an electron with wavevector \mathbf{k} and the creation of an electron with wavevector $\mathbf{k} + \mathbf{G}$. This also explains the conservation of crystal momentum; the electron undergoes scattering processes which changes the momentum from \mathbf{k} to $\mathbf{k} + \mathbf{G}$ (so momentum is not conserved) but the crystal momentum (i.e. momentum up to reciprocal lattice vectors) is. This also connects back to the Bruillion zone; the wavevector \mathbf{k} appearing here can always be confined to the first Bruillion zone.

In the following, we suppress the spin index, and focus on 1D systems for simplicity. So, we have:

$$H_0 = \sum_k \epsilon_k c_k^\dagger c_k, \quad H' = \sum_{k,G} U_G c_{k+G}^\dagger c_k$$

we ask how is an electron in an eigenstates $|k\rangle = c_k^\dagger |0\rangle$ of H_0 perturbed by H' .

11.2.1 Zeroth-order perturbation theory

To start, in zeroth order perturbation theory, we have:

$$E_k^{(0)} = \epsilon_k \quad (11.9)$$

Of course, nothing exciting here...

11.2.2 First-order perturbation theory

We have:

$$E_k^{(1)} = \langle k | H' | k \rangle = \langle k | \sum_{q,G} U_G c_{q+G}^\dagger c_q | k \rangle = \langle k | \sum_G U_G c_{k+G}^\dagger | 0 \rangle = \sum_G U_G \langle k | k+G \rangle = U_0 = 0. \quad (11.10)$$

so the first order contribution is zero.

11.2.3 Second-order perturbation theory

We have:

$$E_k^{(2)} = \sum_{k' \neq k} \frac{|\langle k|H'|k' \rangle|^2}{\epsilon_{k'} - \epsilon_k} \quad (11.11)$$

So we calculate:

$$\langle k|H'|k' \rangle = \sum_{qG} U_G \langle k|c_{q+G}^\dagger c_q|k' \rangle = \sum_G U_G \langle k|c_{k'+G}^\dagger|0 \rangle = \sum_G U_G \langle k|k' + G \rangle = \sum_G U_G \delta_{k,k'+G} = \quad (11.12)$$

Therefore:

$$E_k^{(2)} = \sum_{k' \neq k, G, G'} \frac{U_G U_{G'}^* \delta_{k,k'+G} \delta_{k,k'+G'}}{\epsilon_{k'} - \epsilon_k} = \sum_{G \neq 0, G'} \frac{U_G U_{G'}^* \delta_{GG'}}{\epsilon_{k+G} - \epsilon_k} = \sum_{G \neq 0} \frac{|U_G|^2}{\epsilon_{k+G} - \epsilon_k} \quad (11.13)$$

For small $|U_G|$, this correction is small, EXCEPT when $|\epsilon_{k-G} - \epsilon_k|$ is also small. This occurs at specific values of $\mathbf{k} = \frac{1}{2}\mathbf{G}$ (valid also in 3D). However this is not the final answer; in reality there will not be an infinite correction. What's the catch? When we do perturbation theory as we have, we assume that the energy levels are non-degenerate; but here there is a degeneracy. Hence, at and near $\mathbf{k} = \frac{1}{2}\mathbf{G}$, we must apply degenerate perturbation theory because $\epsilon_{k-G} = \epsilon_k$ and this could be viewed as a degeneracy.

11.3 Degenerate Perturbation Theory

We define two near degenerate states:

$$|1\rangle = |k\rangle, \quad |2\rangle = |k - G_1\rangle \quad (11.14)$$

and construct the Hamiltonian matrix in this basis:

$$H = \begin{pmatrix} \langle 1|H|1\rangle & \langle 1|H|2\rangle \\ \langle 2|H|1\rangle & \langle 2|H|2\rangle \end{pmatrix} = \begin{pmatrix} \epsilon_k & U_G \\ U_G^* & \epsilon_{k-G} \end{pmatrix} \quad (11.15)$$

The perturbed energies are given by eigenvalues:

$$\det \begin{pmatrix} \epsilon_k - E & U_G \\ U_G^* & \epsilon_{k-G} - E \end{pmatrix} = 0 \implies (\epsilon_k - E)(\epsilon_{k-G} - E) - |U_G|^2 = 0 \quad (11.16)$$

so then:

$$E_k = \frac{1}{2}(\epsilon_k + \epsilon_{k-G}) \pm \sqrt{\frac{1}{4}(\epsilon_k - \epsilon_{k-G})^2 + |U_G|^2} \quad (11.17)$$

Note for $k = \frac{1}{2}G$, i.e. the degeneracy point, this implies:

$$E_k = \epsilon_k \pm |U_G|. \quad (11.18)$$

So we have an energy gap! This quantifies the distinction of metals and insulators, using quantum mechanics; we will go further into this discussion next class.

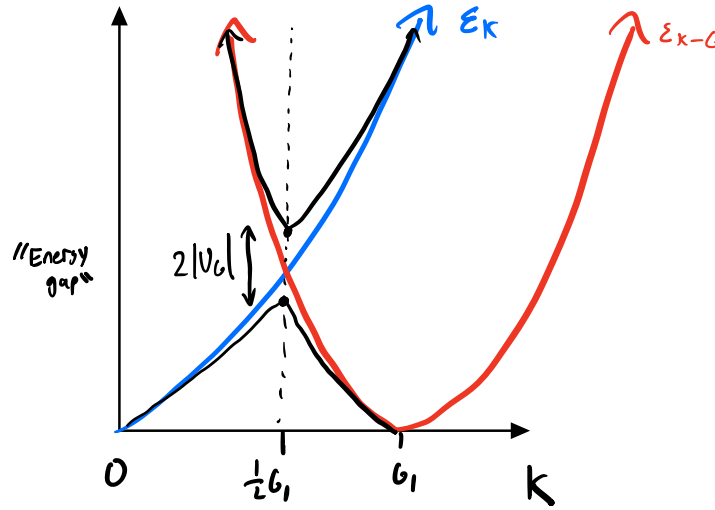


Figure 11.1: Plot of the energy ϵ_k and ϵ_{k-G} as a function of k . At $k = \frac{1}{2}G$ the two energy functions coincide. At and near this point, non-degenerate perturbation theory breaks down. A more careful treatment using degenerate perturbation theory shows that there are two energy levels, separated by a gap $2|U_G|$.

12 Electrons in a Periodic Potential - Tight Binding Approximation

12.1 Review and Examples of Band Structure

We recall that a weak perturbation has no effect on the electron spectrum except at the degeneracy points. At these points, the perturbation results in a gap at these degeneracy points of magnitude $2|U_G|$.

We did this for one specific degeneracy point (Bragg plane), but there will be many such points. The free electron parabola gets broken up into many such disjoint pieces. We can picture this in the extended zone scheme (pictures $-\infty < k < \infty$), or the reduced zone scheme (pictures the first Brillouin zone with many levels), or the repeated zone scheme (copying over the structure in the first Brillouin zone multiple times throughout space) as pictured below.

In three dimensions, the analysis proceeds in the same way (but there are more levels at the same k). If the Fermi energy lies inside of the band, we have a metal (e.g. aluminum), and if we have a Fermi energy inside of the gap, we have an insulator (e.g. silicon).

An important 2D example is graphene; we consider carbon atoms in a honeycomb lattice, which can be viewed as a triangular Bravais lattice with a two point basis. It has a hexagonal Brillouin zone, and the band structure looks fairly free electron. Near the K point (corner of the Brillouin zone), there is a protection (from time reversal and inversion symmetry) and there is the realization of Dirac dispersion near the K points; here we see behavior like massless particles, with zero curvature. Electrons near these K points are in some sense massless. This is a kind of band structure that we could *not* get from the weak periodic potential approximation, as the periodic potential is not weak. It also motivates our next topic, which is the tight-binding approximation (which is successful for (e.g.) graphene).

12.2 Tight-Binding Approximation

We consider again a periodic potential, this time the case where we have a periodic Coulomb potential. If we focus on the highly excited wavefunctions at each site, we have an exponentially decaying wavefunction centered at each well. If only the exponentially small tails of the wavefunctions overlap, then we can think

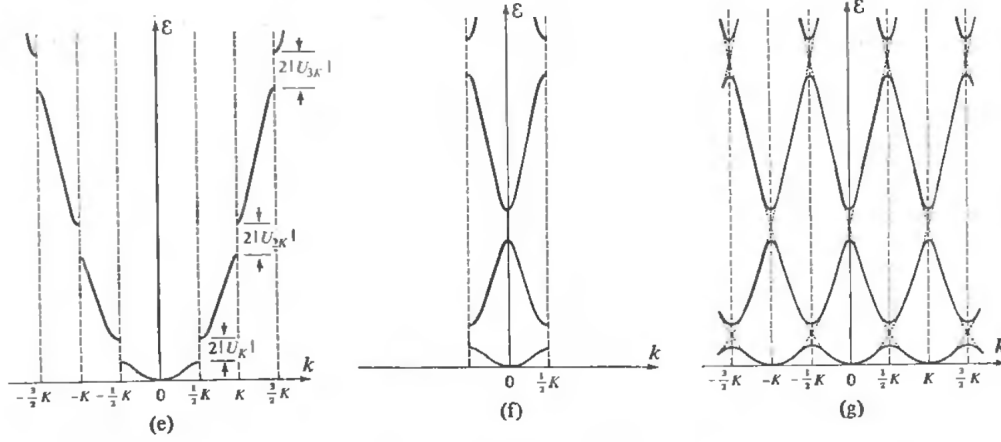


Figure 12.1: Sketch of the extended (left), reduced (center), and repeated (right) zone schemes for describing the electron dispersion under a weak periodic perturbation. Figures reproduced without permission from Ashcroft and Mermin.

of the system as electrons sitting in their orbitals, and being subject to a small perturbation that allows them to tunnel from one well to another. In some sense it is the complete opposite of the weak potential approximation. There we started with free electrons and treated the potential as a perturbation. Here, we assume electrons are tightly bound in atomic orbitals; tunnelling (a.k.a. “hopping”) occurs due to small wavefunction overlap (the perturbation).

We have the Hamiltonian:

$$H = -\frac{\hbar^2 \nabla^2}{2m} + \sum_i U_{at}(\mathbf{r} - \mathbf{R}_i) \quad (12.1)$$

We work in the basis of atomic wavefunctions $\phi_n(\mathbf{r})$, which satisfy:

$$H_{at}\phi_n = E_n\phi_n, \quad H_{at} = -\frac{\hbar^2 \nabla^2}{2m} + U_{at}(\mathbf{r}) \quad (12.2)$$

we assume that the ϕ_n s are known. This is our starting point. We write the full Hamiltonian in second quantization notation using the ϕ_n basis. We introduce operators $c_n^\dagger(\mathbf{R}_i)$ which creates an electron in state $\phi_n(\mathbf{r} - \mathbf{R}_i)$ represented as $|n, \mathbf{R}_i\rangle$. Then, the most general tight binding Hamiltonian can be written as:

$$\hat{H} = \sum_{m,n} \sum_{i,j} \langle n, \mathbf{R}_i | H | m, \mathbf{R}_j \rangle c_n^\dagger(\mathbf{R}_i) c_m(\mathbf{R}_j) \quad (12.3)$$

$c_n^\dagger(\mathbf{R}_i) c_m(\mathbf{R}_j)$ precisely destroys one electron at energy E_m at position \mathbf{R}_j and creates one at energy E_n at position \mathbf{R}_i . The matrix elements $\langle n, \mathbf{R}_i | H | m, \mathbf{R}_j \rangle = t_{ij}^{mn}$ give the amplitudes for this process to occur.

This general form is generally not super useful unless more about the matrix elements are known. Fortunately, because the atomic wavefunctions $\phi_n(\mathbf{r})$ decay as $\sim e^{-r/a_0}$, we expect the overlap between the wavefunctions is only appreciable between a few nearest neighbours, and the rest are negligible. So, the matrix elements t_{ij}^{mn} are negligible when $|\mathbf{R}_i - \mathbf{R}_j| \gg a_0$. In practice, it is usually sufficient to include first (potentially second or third) neighbour hopping.

Tight binding models can be solved for any order of these hoppings, but low order ones have very simple solutions. Note that we will not go into the band structure calculations of the t_{ij}^{mn} matrix elements here; but there are many textbooks and packages written specifically about this topic. There are well-established techniques, but they require detailed knowledge of atomic orbitals. This is often difficult to do by hand; however, they can often be constrained by symmetries.

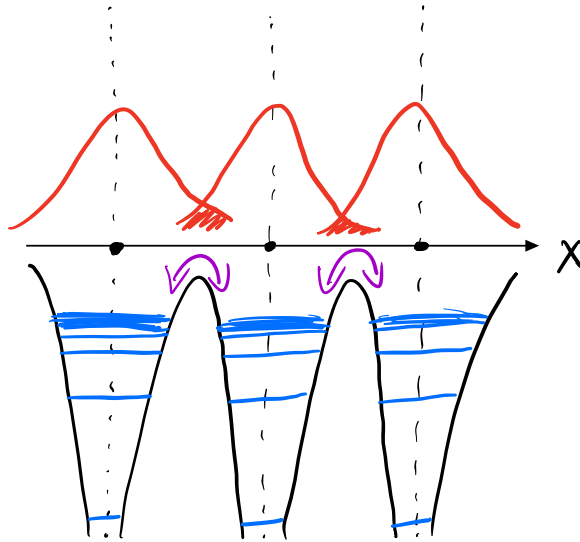


Figure 12.2: Cartoon sketch of the tight-binding approximation. We have per-site wavefunctions (arising from a Coloumb potential, or some other strong per-site potential) that are centered at each site of the lattice that “strongly bind” the electrons. There is overlap between the per-site wavefunctions that leads to hopping in between the sites.

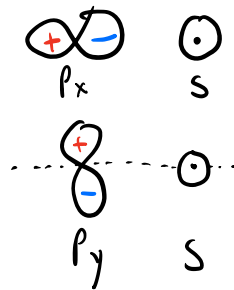


Figure 12.3: The overlap matrix elements t_{ij}^{mn} can often be constrained by symmetries. When looking at the overlap of p_x and s , symmetry constraints give us no information. However, when looking at the overlap of p_y and s , the overlap integral is odd under the interchange $y \leftrightarrow -y$ and so we conclude that the matrix element vanishes.

12.3 Example - Single non-degenerate orbital in 2D

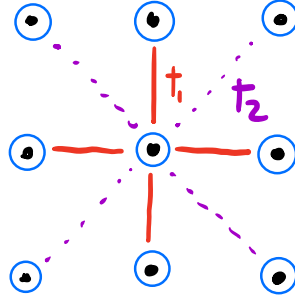


Figure 12.4: Sketch of a 2D square lattice with a single non-degenerate orbital at each site. The interactions to consider will be the nearest neighbour interactions (t_1) and the next-nearest interactions (t_2). The next-nearest neighbour interactions and further can be neglected.

In this scenario our tight-binding Hamiltonian looks like:

$$\hat{H} = \sum_i E_i c^\dagger(\mathbf{R}_i) c(\mathbf{R}_i) - \sum_{i \neq j} t_{ij} c^\dagger(\mathbf{R}_i) c(\mathbf{R}_j) \quad (12.4)$$

here, E_i is the “on-site” energy. We will take $E_i = \epsilon_0$ for all i in our monoatomic lattice. for t_{ij} we consider only first and second neighbours as indicated in the figure:

$$t_{ij} = \begin{cases} t_1 & \text{first neighbour} \\ t_2 & \text{second neighbour} \\ 0 & \text{otherwise} \end{cases}$$

With this we can rewrite:

$$\hat{H} = \epsilon_0 \sum_i c^\dagger(\mathbf{R}_i) c(\mathbf{R}_i) - \sum_{i, \delta} t_\delta c^\dagger(\mathbf{R}_i + \delta) c(\mathbf{R}_i) \quad (12.5)$$

this is often a useful rewrite as in many cases the hopping terms do not depend on both lattice positions i, j but only the difference between them δ . δ here are vectors that point from a site to its first and second neighbour. We will see that a Fourier transform immediately solves this type of simple problem:

$$\begin{aligned} c(\mathbf{R}_i) &= \frac{1}{\sqrt{N}} \sum_{\mathbf{k}} e^{i\mathbf{k} \cdot \mathbf{R}_i} c_{\mathbf{k}} \\ c^\dagger(\mathbf{R}_i) &= \frac{1}{\sqrt{N}} \sum_{\mathbf{k}} e^{-i\mathbf{k} \cdot \mathbf{R}_i} c_{\mathbf{k}}^\dagger \end{aligned} \quad (12.6)$$

our Fourier transformed Hamiltonian takes the form:

$$\begin{aligned} \hat{H} &= \epsilon_0 \sum_{\mathbf{k}, \mathbf{k}'} c_{\mathbf{k}}^\dagger c_{\mathbf{k}'} \frac{1}{N} \sum_{\mathbf{R}_i} e^{-i\mathbf{R}_i \cdot (\mathbf{k} - \mathbf{k}')} - \sum_{\mathbf{k}, \mathbf{k}'} c_{\mathbf{k}}^\dagger c_{\mathbf{k}'} \frac{1}{N} \sum_{i, \delta} t_\delta e^{-i\mathbf{R}_i \cdot (\mathbf{k} - \mathbf{k}') - i\delta \cdot \mathbf{k}} \\ &= \epsilon_0 \sum_{\mathbf{k}} c_{\mathbf{k}}^\dagger c_{\mathbf{k}} - \sum_{\mathbf{k}} c_{\mathbf{k}}^\dagger c_{\mathbf{k}} \left(\sum_{\delta} t_\delta e^{-i\delta \cdot \mathbf{k}} \right) \end{aligned} \quad (12.7)$$

and so:

$$\hat{H} = \sum_{\mathbf{k}} \epsilon(\mathbf{k}) c_{\mathbf{k}}^{\dagger} c_{\mathbf{k}}, \quad \epsilon(\mathbf{k}) = \epsilon_0 - \sum_{\delta} t_{\delta} e^{-i\mathbf{k} \cdot \delta} \quad (12.8)$$

Actually the solution we write above works for any tight binding model. We can evaluate the energies for our case specifically using that $\delta_1 = a(\pm\hat{x}, \pm\hat{y})$ and $\delta_2 = a(\hat{x} \pm \hat{y}, -\hat{x} \pm \hat{y})$. This yields:

$$\epsilon(\mathbf{k}) = \epsilon_0 - 2t_1(\cos(ak_x) + \cos(ak_y)) - 2t_2(\cos(a(k_x + k_y)) + \cos(a(k_x - k_y))) \quad (12.9)$$

For simplicity, let us illustrate the case with $t_2 = 0$. We sketch the lines of constant energy for the dispersion:

$$\epsilon(\mathbf{k}) = \epsilon_0 - 2t_1(\cos(ak_x) + \cos(ak_y))$$

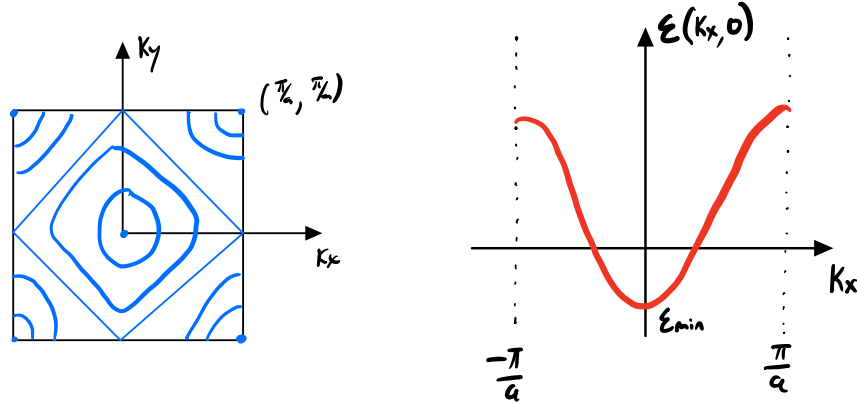


Figure 12.5: Lines of constant energy for the 2D tight-binding square lattice with one non-degenerate orbital (left) and plot of energy vs. k_x for fixed $k_y = 0$ (right).

We find that $\epsilon_{\min} = \epsilon(\mathbf{k} = 0) = \epsilon_0 - 4t$ and $\epsilon_{\max} = \epsilon(\mathbf{k} = (\frac{\pi}{a}, \frac{\pi}{a})) = \epsilon_0 + 4t$. The bandwidth is $W = \epsilon_{\max} - \epsilon_{\min} = 8t$.

A standard undergraduate question: if we have a 2D square lattice and each atom contributes one electron, do we have a metal or an insulator? To answer this, we recall that the number of k states in the band is equal to the number of unit cells in the crystal.

On the face of it, it seems like all of the N states would be filled with electrons so we would have an insulator. But taking into account spin, there are $2N$ states, so we only fill half of the states, corresponding to all of the states inside of the diamond. This meets the definition of a metal!

12.3.1 Effective Mass

The effective mass for these electrons can be defined near either the minima or the maxima of the dispersion by expanding it to second order in momenta and study the long wavelength behavior.

(i) Near $\mathbf{k} = 0$, we can expand:

$$\epsilon(\mathbf{k}) = \epsilon_0 - 2t_1(1 - \frac{(ak_x)^2}{2} + \dots + 1 - \frac{(ak_y)^2}{2} + \dots) = (\epsilon_0 - 4t_1) + t_1 a^2 k^2 + \dots \quad (12.10)$$

In the literature, often $t_1 a^2 k^2$ is written as $\frac{\hbar^2 k^2}{2m^*}$ with $m^* = \frac{\hbar^2}{2t_1 a^2}$ the effective mass. In principle this could be larger or smaller than the real electron mass, and in the literature when things such as

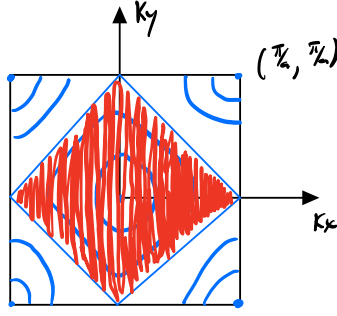


Figure 12.6: If each atom contributes one electron, the band is half filled, corresponding to all of the states inside of the diamond (shaded in red). Thus the fermi energy falls within the band and the material will be a metal.

“heavy electrons” are discussed, these are compounds where the band mass of the electrons is heavy. In conclusion, near $\mathbf{k} = 0$ (the Γ point) the tight-binding electron behaves as a particle of mass m^* .

(ii) Near $\mathbf{k} = \left(\frac{\pi}{a}, \frac{\pi}{a}\right) \equiv \mathbf{K}$, we expand:

$$\epsilon(\mathbf{k} + \mathbf{K}) = (e_0 + 4t_1) - t_1 a^2 k^2 + \dots \quad (12.11)$$

where we now have the effective mass $m^* = -\frac{\hbar^2}{2t_1 a^2}$. The effective mass is negative! The carriers are holes.

12.3.2 Band Velocity

The group velocity is given by:

$$\mathbf{v}_{\mathbf{k}} = \frac{1}{\hbar} \frac{\partial \epsilon(\mathbf{k})}{\partial \mathbf{k}} = \frac{2t_1 a}{\hbar} (\sin(ak_x), \sin(ak_y)) \quad (12.12)$$

Near $\mathbf{k} = 0$, we have $\mathbf{v}_{\mathbf{k}} \approx \frac{2t_1 a^2}{\hbar} \mathbf{k}$. This is as expected for a free particle of mass m^* . Near $k_x = \frac{\pi}{a}$, we have $\mathbf{v}_{\mathbf{k}} = -\frac{2t_1 a^2}{\hbar} \delta k_x$. This is very odd. The velocity is opposite to the change in momentum. This is however consistent with the effective mass being negative.

One may ask what makes electrons behave in this strange way. The intuition of free space is violated here as the electron is not moving in free space (it is in the tight-binding model)! In the long wavelength limit the free space analogy may hold, but in other parts of the Brillouin zone this is simply not the case.

Another point - this shows why it is important to distinguish between momentum and crystal momentum. Here we have crystal momentum, and we can see it really does not behave as we conventionally understand momentum.

12.4 Example - Lattice with a Basis (Dimerized chain)

Our unit cell now contains two atoms, and is of length $2a$. Every atom in the chain can be labelled by the position of the unit cell and with $1/2$. We therefore write the Hamiltonian in second quantization notation as:

$$H = \epsilon_0 \sum_i \left[c_1^\dagger(R_i) c_1(R_i) + c_2^\dagger(R_i) c_2(R_i) \right] + t \sum_i \left[c_1^\dagger(R_i) c_2(R_i + a) + h.c. \right] + t' \sum_i \left[c_1^\dagger(R_i) c_2(R_i - a) + h.c. \right] \quad (12.13)$$

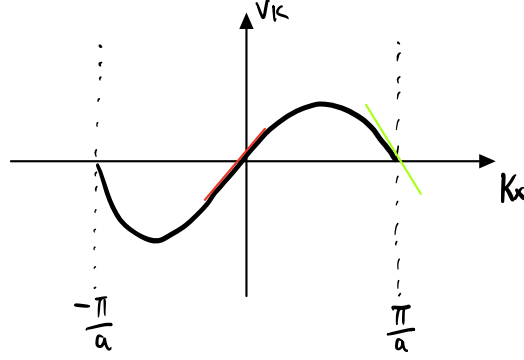


Figure 12.7: Group velocity of electrons for the tight-binding square lattice, as a function of k_x . Near $\mathbf{k} = 0$ the electrons behave as free particles with positive effective mass m^* , and the group velocity is aligned with the change in momentum (red) as would be expected for a free particle. Near $k_x = \frac{\pi}{a}$, the electrons have negative effective mass, and the group velocity is opposite to the change in momentum (green).

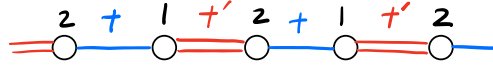


Figure 12.8: Cartoon of dimerized tight-binding chain, with alternating bond strengths t/t' .

where $c_1^\dagger c_2$ describes a hop from site 2 to site 1 and h.c. the opposite. We solve this via a Fourier transform:

$$c_\alpha(r) = \frac{1}{\sqrt{N}} \sum e^{ikr} c_{\alpha k}$$

$\alpha = 1, 2$:

$$H = \epsilon_0 \sum_{\alpha, k} c_{k, \alpha}^\dagger c_{k, \alpha} + t \sum_k \left(c_{1k}^\dagger c_{2k} e^{ika} + h.c. \right) + t' \sum_k \left(c_{1k}^\dagger c_{2k} e^{-ika} + h.c. \right) \quad (12.14)$$

Two solve, define a two-component spinor:

$$\chi_k = \begin{pmatrix} c_{1k} \\ c_{2k} \end{pmatrix}, \quad \chi_k^\dagger = \begin{pmatrix} c_{1k}^\dagger & c_{2k}^\dagger \end{pmatrix}$$

and express the Hamiltonian (setting $a = 1$):

$$H = \sum_k \chi_k^\dagger h_k \chi_k, \quad h_k = \begin{pmatrix} \epsilon_0 & te^{ik} + t'e^{-ik} \\ te^{-ik} + t'e^{ik} & \epsilon_0 \end{pmatrix} \quad (12.15)$$

We diagonalize this matrix, setting $\det(j_k - \epsilon \mathbb{I}) = 0$. We get a secular equation:

$$(\epsilon_0 - \epsilon)^2 - (te^{ik} + t'e^{-ik})(te^{-ik} + t'e^{ik}) = 0$$

When we solve the quadratic equation for ϵ , we find:

$$\epsilon = \epsilon_0 \pm \sqrt{t^2 + t'^2 + 2tt' \cos(2k)} \quad (12.16)$$

so the band structure will be composed of two bands.

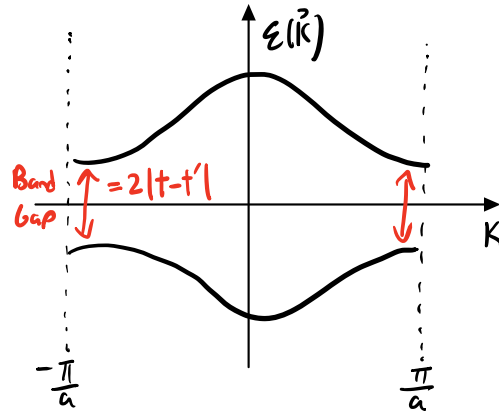


Figure 12.9: Energy as a function of k for the dimerized tight-binding chain. The band structure is composed of two bands.

Now, suppose that each atom in this dimerized chain donates exactly one electron. Do we have a metal or insulator? Each unit cell has two atoms. So, if each atom contributes one electron, the bottom band is completely filled (with N electrons) and the top band is completely unfilled. So we have an insulator.

Another question - what happens if the atoms we make them now identical? The gap closes - but this is something for the reader to think about. This ends the discussion of the tight binding model, and we do other things starting monday.

13 Density Functional Theory

13.1 Motivation

A brief recap - at the beginning of the term we discussed the Jellium model, where we replaced the ionic lattice with a uniform positive background, and we then calculated the energy of electrons amongst other properties. We then began to discuss electrons moving in a periodic potential of the lattice, which lead to band structure theory - but in this approximation we neglected interactions between the electrons. In a real solid that we have both a periodic lattice and electronic interactions. Trying to combine the two leads to an intractable problem, as as soon as we abandon the independent electron approximation, suddenly the wavefunction becomes a function of 10^{23} variables, and there is no good method to get this wavefunction in a system that does not have full translational invariance.

DFT is then the only practical way to perform calculations for electrons structure in real solids. It was developed bhy Walter Kohn in the 60s. We briefly survey the topic in a single class today, but there exist entire books written on the topic and it is a technique well-used in research today.

13.2 The Hohenberg-Kohn Theorem

QFT is grounded in the Hohenberg-Kohn theorem, which applies to the family of electron Hamiltonians of the form:

$$\hat{H} = \sum_{\mathbf{k},\sigma} \frac{\hbar \mathbf{k}^2}{2m} c_{\mathbf{k}\sigma}^\dagger c_{\mathbf{k}\sigma} + \sum_{\mathbf{q},\mathbf{q},\sigma} U(\mathbf{q}) c_{\mathbf{k}+\mathbf{q},\sigma}^\dagger c_{\mathbf{k}\sigma} + \frac{1}{2} \sum_{\mathbf{k}\mathbf{p}\mathbf{q}\sigma\sigma'} V_{\mathbf{q}} c_{\mathbf{k}-\mathbf{q},\sigma}^\dagger c_{\mathbf{p}+\mathbf{q},\sigma'}^\dagger c_{\mathbf{p},\sigma'} c_{\mathbf{k},\sigma} \quad (13.1)$$

where the first term is the electron kinetic energy which we call \hat{T} , the second term is the potential term \hat{U} , and the third term is the Coulomb interactions \hat{V} .

The Hohenberg-Kohn theorem says that the expectation value of any operator \hat{O} is a unique functional of the ground-state electron density $n_0(\mathbf{r})$. We don't need the full many-body wavefunction of the system; we only need the ground state electron density. Quite incredible! From this theorem unfolds the apparatus of DFT. the simplification is that the density is a function only of three spatial variables $\mathbf{r} = (r_x, r_y, r_z)$. We only need to couch calcaultions in terms of 3 variables, not 10^{23} !

Comment: There do exist extensions of the theorem to thermal ensembles at finite temperature, though the HK theorem in the above form only addresses the ground state.

A few other remarks. The \hat{T}, \hat{V} pieces of Eq. (13.1) are universal across solids, and \hat{U} is the only part that differs between solids. In principle, given \hat{U} we can calcalte the electron density $n_0(\mathbf{r})$.

The HK theorem states that the mapping:

$$U(\mathbf{r}) \leftrightarrow n_0(\mathbf{r}) \quad (13.2)$$

is actually reversible. Because from $U(\mathbf{r})$ we can in principle obtain the full many-body wavefunction $\psi(\mathbf{r}_1, \dots, \mathbf{r}_n)$ (and thus the expectation value $\langle \hat{O} \rangle$) this implies that $\langle \hat{O} \rangle$ is uniquely determined by $n_0(\mathbf{r})$. Also note that everything we will say is true of non-degenerate ground states (degeneracy can add complications to these uniqueness statements), though we can extend the results to the case with degeneracy.

13.3 Proof of the Hohenberg-Kohn Theorem

13.3.1 Part 1

We first show that two potentials $U(\mathbf{r})$ and $U'(\mathbf{r})$ that differ by more than a trivial constant, necessarily lead to different ground states $\psi_0 \neq \psi'_0$. We have:

$$\begin{aligned} (\hat{T} + \hat{V} + \hat{U})\psi_0 &= \epsilon_0 \psi_0 \\ (\hat{T} + \hat{V} + \hat{U}')\psi'_0 &= \epsilon'_0 \psi'_0 \end{aligned} \quad (13.3)$$

We proceed by proof by contradiction. Assume $\psi_0 = \psi'_0$, then subtracting the second equation from the first:

$$(\hat{U} - \hat{U}')\psi_- = (\epsilon_0 - \epsilon'_0)\psi_0 \quad (13.4)$$

So \hat{U}, \hat{U}' only differ by a trivial constant, contradicting our assertion that they differ by more than a trivial constant. \square

13.3.2 Part 2

It is also clear that two different densities $n_0(\mathbf{r}) \neq n'_0(\mathbf{r})$ require different potentials - this follows from the uniqueness of solutions to the Schrodinger equation (in the non-degenerate case).

It therefore follows that the ground state wavefunction is uniquely specified by the electron density. One can also show that $\psi_0 \neq \psi'_0$ implies $n_0(\mathbf{r}) \neq n'_0(\mathbf{r})$ which completes the proof. \square

13.4 Variational Principle

There is an important variational principle associated with the HK theorem. HK applies for expectation values of all operators, so in particular let us pick the Hamiltonian (i.e. the ground state energy):

$$\epsilon[n] = \langle \psi_0[n] | \hat{T} + \hat{V} + \hat{U} | \psi_0[n] \rangle \quad (13.5)$$

where $\epsilon[n]$ is the ground state energy functional, such that $\epsilon_0 = \epsilon[n_0]$ is the true ground state energy. As a refresher, a functional is a mapping from the space of functions to the space of numbers. For example the definite integral over $[0, 1]$ is a functional. We distinguish them by angular brackets, and their arguments are functions.

It then holds that:

$$\epsilon_0 < \epsilon[n] \quad (13.6)$$

for any $n(\mathbf{r}) \neq n_0(\mathbf{r})$. This is nothing more than a restatement of the familiar variational principle of $E_0 \leq \langle \psi | H | \psi \rangle$.

This variational principle gives the practical machinery for doing DFT calculations; the ground state energy and $n_0(\mathbf{r})$ can be found by minimizing this functional. $\epsilon[n]$ is usually written as:

$$\epsilon[n] = F_{HK}[n] + \int d^3r n(\mathbf{r})U(\mathbf{r}) \quad (13.7)$$

where $F_{HK}[n] = \langle \psi_0[n] | \hat{T} + \hat{V} | \psi_0[n] \rangle$, which importantly is the same for all systems on the account of universality! F_{HK} has to be determined only once! Of course anything that is this useful is not so easily obtained, and various DFT approaches differ in what they use for the F_{HK} functional. Let us explain the second term - why can it be written as the simple integral? First, we write:

$$n(\mathbf{r}) = \int \{d^3r\} \psi^*(\mathbf{r}_1, \dots, \mathbf{r}_N) \sum_{i=1}^N \delta(\mathbf{r} - \mathbf{r}_i) \psi(\mathbf{r}_1 \dots \mathbf{r}_N)$$

Where $\{d^3r\} = d^3r_1 \dots d^3r_N$. Then:

$$\langle \psi_0 | U | \psi_0 \rangle = \int \{d^3r\} \psi^*(\mathbf{r}_1, \dots, \mathbf{r}_N) \sum_{i=1}^N U(\mathbf{r}_i) \psi(\mathbf{r}_1, \dots, \mathbf{r}_N) = \int \{d^3r\} d^3r_p \psi^*(\dots) \sum_i \delta(\mathbf{r}_p - \mathbf{r}_i) U(\mathbf{r}_p) \psi(\dots) \quad (13.8)$$

and separating this into two integrals, and the integral over all wavefunction coordinates gives exactly $n(\mathbf{r})$, giving precisely $\langle \psi_0 | U | \psi_0 \rangle = \int d^3r n(\mathbf{r})U(\mathbf{r})$.

13.5 The Kohn-Sham Formulation

This formulation allows for practical applications of the HK theorem. The idea - one imagines that there is a non-interacting reference system $\hat{H}_S = \hat{T} + \hat{U}_S$ whose potential \hat{U}_S is chosen such that the ground state density is precisely the same as the actual interacting system $\hat{H} = \hat{T} + \hat{V} + \hat{U}$ we are interested in. Because computations of ψ_0 and n_0 are “easy” in a non-interacting system \hat{H}_S (as the independent electron approximation can be used), we can then practically implement this whole scheme.

13.5.1 DFT for the Reference System

Similarly for the actual system, we have a ground state energy function:

$$\epsilon_S[n] = T_S[n] + \int d^3r U_S(\mathbf{r})n(\mathbf{r}) \quad (13.9)$$

where here $n_s(\mathbf{r})$ can be easily obtained from the wavefunctions:

$$n_s(\mathbf{r}) = \sum_{i=1}^N |\phi_i(\mathbf{r})|^2 \quad (13.10)$$

and we solve for the ϕ_i s by solving the Schrodinger equation:

$$\left[-\frac{\hbar^2 \nabla^2}{2m} + U_S(\mathbf{r}) \right] \phi_i(\mathbf{r}) = \epsilon_i \phi_i(\mathbf{r}) \quad (13.11)$$

The hard part is we do not know what U_S is - we need to find its form. We write down again the ground state energy functional for the full system, but in a peculiar way:

$$\begin{aligned} \epsilon[n] &= T[n] + V[n] + \int d^3r n(\mathbf{r})U(\mathbf{r}) \\ &= T_S[n] + \left[T[n] - T_S[n] - \frac{e^2}{2} \int d^3r d^3r' \frac{n(\mathbf{r})n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} \right] + \frac{e}{2} \int d^3r d^3r' \frac{n(\mathbf{r})n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} + \int d^3r U(\mathbf{r})n(\mathbf{r}) \end{aligned} \quad (13.12)$$

Where the term in brackets is the exchange-correlation functional:

$$\epsilon_{xc}[n] = F_{HK}[n] - \frac{e^2}{2} \int d^3r d^3r' \frac{n(\mathbf{r})n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} - T_S[n] \quad (13.13)$$

We minimize $\epsilon[n]$ by variational calculus:

$$0 = \frac{\delta \epsilon[n]}{\delta n(\mathbf{r})} = \frac{\delta T_S[n]}{\delta n(\mathbf{r})} + e^2 \int d^3r' \frac{n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} + U(\mathbf{r}) + v_{xc}[n(\mathbf{r})] \quad (13.14)$$

where $v_{xc}[n(\mathbf{r})] = \frac{\delta \epsilon_{xc}[n]}{\delta n(\mathbf{r})}$ is the exchange-correlation potential. We minimize the reference system:

$$0 = \frac{\delta \epsilon_n[n]}{\delta n(\mathbf{r})} = \frac{\delta T_S}{\delta n(\mathbf{r})} + U_S(\mathbf{r}) \quad (13.15)$$

and then subtracting from the minimization of the full system in Eq. (13.14):

$$U_s(\mathbf{r}) = U(\mathbf{r}) + e^2 \int d^3r' \frac{n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} + v_{xc}(\mathbf{r}) \quad (13.16)$$

This is the important result! In practical calculations we can therefore implement Kohn-Sham by following the three-step program:

1. Choose an initial trial density $n(\mathbf{r})$. Substitute in Eq. (13.16) and calculate $U_S(\mathbf{r})$.
2. Solve the SE in Eq. (13.11) for $\phi_i(\mathbf{r})$ and find the new density $n(\mathbf{r})$ from Eq. (13.10)
3. Iterate until $n(\mathbf{r})$ stops changing.

The enormous advantage of this procedure is that at no point in this calculation are we ever forced to calculate the full many-body wavefunction. Of course, it comes at a cost - we have hidden all of the interaction physics inside of the exchange-correlation potential v_{xc} . Despite this - DFT remains as one of the most successful techniques in CM physics today, and is the basis for much of our understanding of solids.

13.5.2 Remarks

- (i) The above procedure is implemented in various software packages (Wien 2K, Quantum Espresso...)
- (ii) These differ mainly in the form of v_{xc} .
- (iii) Packages provide “band structures”, i.e. energy eigenvalues $\epsilon_i(\mathbf{k})$ of Eq. (13.11). If we construct the full many-body wavefunctions from the obtained observable expectation values, we often obtain very good approximations to the true systems - and systems for which DFT fail are unusual/interesting.
- (iv) There are many useful generalizations of DFT, such as TD-DFT which adds time-dependence, CDFT (current DFT) which includes an external magnetic field, and EDFT (ensemble DFT) which deals with degeneracies.

13.6 Choices for the Exchange-Correlation Potential

Many sophisticated approximations for $\epsilon_{xc}[n]$ have been developed and implemented over the years.

The simplest such is the local density approximation, which assumes that ϵ_{xc} only depends on the local density:

$$\epsilon_{xc}[n] = \int d^3r n(\mathbf{r}) \epsilon_{xc}(n(\mathbf{r})) \quad (13.17)$$

where $\epsilon_{xc}(n)$ is the exchange-correlation energy (per particle) of a homogenous electron gas of density n (Jellium):

$$\epsilon_{xc}(n) = \frac{e^2}{a_0 r_s^2} (b r_s + c r_s \log r_s + \dots). \quad (13.18)$$

Other popular approximations include LSDA (low spin density approximation) and GGA, the “generalized gradient approximation”:

$$\epsilon_{xc}[n] = \int d^3r [g_{00}(n) + g_{22}(n)(\nabla n)^2 + g_{42}(n)(\nabla^2 n)^2 + \dots] \quad (13.19)$$

Next time, we will start discussing the semiclassical theory of conduction in metals. Please read A&M p214-218 in preparation.

14 Semiclassical Theory of Conduction I

14.1 Semiclassical QM in Crystals

We start with a wave packet with $\Delta x \Delta p \geq \hbar$, $\psi(\mathbf{r}, t) = \sum_{\mathbf{k}} g(\mathbf{k}) e^{i\mathbf{k} \cdot \mathbf{r} - \frac{\hbar^2 \mathbf{k}^2}{2m} t}$. By considering the “center of mass” of the wave packet, we are able to derive equations of motion from the Schrodinger Equation (non-trivially) - In free space, these read:

$$\begin{aligned} \frac{d\mathbf{r}}{dt} &= \frac{\hbar \mathbf{k}}{m} \\ \hbar \frac{d\mathbf{k}}{dt} &= -e(\mathbf{E} + \frac{1}{c} \mathbf{v} \times \mathbf{H}) \end{aligned} \quad (14.1)$$

In a periodic crystal, we need to revise this slightly; we have a “wave packet” described by Bloch wavefunctions, $\psi(\mathbf{r}, t) = \sum_{\mathbf{k}} g(\mathbf{k}) \psi_{n\mathbf{k}}(\mathbf{r}) e^{-i\epsilon_n(\mathbf{k})t/\hbar}$ where $\psi_{n\mathbf{k}}(\mathbf{r})$ are the Bloch eigenstates. In this setting, we obtain the equations of motion:

$$\begin{aligned} \mathbf{v}(\mathbf{k}) &= \frac{1}{\hbar} \nabla_{\mathbf{k}} \epsilon_n(\mathbf{k}) \\ \hbar \frac{d\mathbf{k}}{dt} &= -e(\mathbf{E} + \frac{1}{c} \mathbf{v} \times \mathbf{H}) \end{aligned} \quad (14.2)$$

Note that although crystal momentum is not equal to the actual momentum, we have:

$$\nabla_{\mathbf{k}} \epsilon_{\mathbf{k}} = \frac{\hbar^2}{2m} \int d\mathbf{r} \psi_{n\mathbf{k}}^* (-i\nabla) \psi_{n\mathbf{k}} \quad (14.3)$$

where the expectation value of velocity appears on the RHS. So, we can apply Eq. (14.2) if we suppose some localization in k -space, i.e. $\nabla k \ll \frac{1}{a}$. Note that in this limit that due to the uncertainty relation $\Delta x \Delta k \geq 1$, we have that $\Delta x \gg a$, i.e the particle is very spread out in position space.

14.2 Limits of Validity

We assume that electrons stay in one band, and that \mathbf{E}, \mathbf{B} vary slowly:

$$\begin{aligned} e|\mathbf{E}|a &\ll \frac{\Delta^2}{\epsilon_F} \\ \hbar\omega_c &\ll \frac{\Delta^2}{\epsilon_F} \\ \hbar\omega_{AC} &\ll \frac{\Delta^2}{\epsilon_F} \end{aligned} \quad (14.4)$$

here Δ is the band gap, a the lattice spacing, and $\omega_c = \frac{eB}{m^*c}$ is the cyclotron frequency.

Note that typically $|\mathbf{E}| \sim 1 \text{ N m}^{-1}$, so with $a \sim 10^{-10} \text{ m}$, in order for the condition to hold we require $\Delta \gg 10^{-5} \text{ eV}$. In contrast, $|\mathbf{B}| \sim 1 \text{ T}$ typically, where then $\hbar\omega_c \sim 10^{-4} \text{ eV}$ which requires $\Delta \gg 10^{-2} \text{ eV}$. So, the semiclassical approximation breaks down due to the magnetic field (known as magnetic breakdown).

14.3 Filled Bands are Inert

Let us define the current:

$$\mathbf{j} = (-e) \sum_{\mathbf{k} < k_F} \mathbf{v}(\mathbf{k}) = -e \int_{\mathbf{k} < k_F} \frac{d^3k}{(2\pi)^3} \frac{1}{\hbar} \nabla_{\mathbf{k}} \epsilon_{\mathbf{k}} \quad (14.5)$$

From this, we observe that filled bands are inert; looking for example at the one-dimensional case:

$$\mathbf{k} = -e \int_{k < k_F} \frac{dk}{2\pi} \frac{1}{\hbar} \frac{\partial}{\partial k} \epsilon_k = \frac{e}{2\pi\hbar} \epsilon_k \Big|_{-\pi/2}^{\pi/2} = 0 \quad (14.6)$$

where the last equality is obtained via periodicity. From this observation, we would conclude that only bands near the Fermi surface are relevant for transport properties (as far as the semiclassical approximation is concerned) - filled bands will never contribute, due to this periodicity.

14.4 Motion in a Uniform Electric Field

For uniform \mathbf{E} , we have:

$$\hbar \frac{d\mathbf{k}}{dt} = -e\mathbf{E} \quad (14.7)$$

which is immediately solved to be:

$$\mathbf{k}(t) = \mathbf{k}(0) - \frac{e\mathbf{E}t}{\hbar} \quad (14.8)$$

so in time we have a displacement of \mathbf{k} -values, and at a certain time, the filled states fully “pass through” - the below sketch illustrates this in the 1-dimensional case:

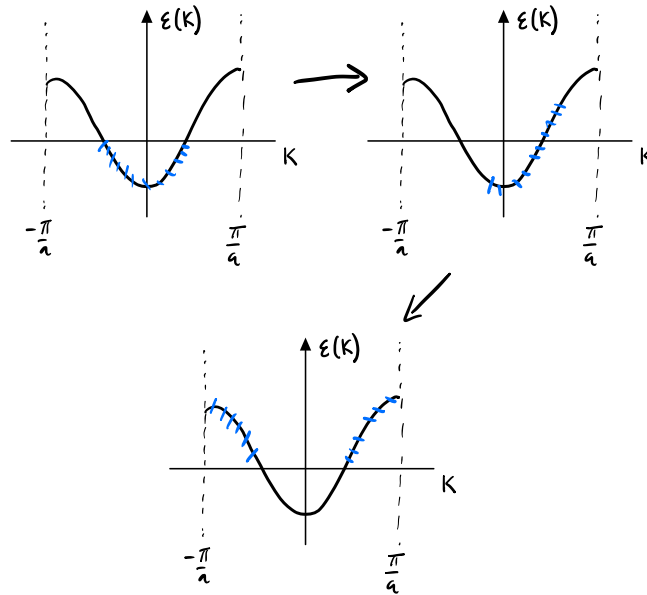


Figure 14.1: 1-dimensional cartoon of the displacement of \mathbf{k} values in time as a result of a uniform \mathbf{E} -field.

Hence, since $\mathbf{v}(t) = \mathbf{v}(\mathbf{k}(t)) = \mathbf{v}(\mathbf{k}(0) - \frac{e\mathbf{E}t}{\hbar})$ and at some time the filled levels “pass through”, we have oscillatory electrons/AC current!

Note in a real crystal, this phenomena does not actually occur due to electron collisions that relax $\mathbf{v}(t)$. These break the periodicity that the Bloch wavefunctions have - in a metal these symmetry-breaking components may be phonons, impurities etc.

14.5 The Hole Picture

Let us rearrange the (vanishing) integral over the full band we had from our previous calculation, by splitting it into two parts:

$$0 = -e \int_{\text{full band}} \frac{d^3k}{(2\pi)^3} \mathbf{v}(\mathbf{k}) = -e \int_{\text{occupied}} \frac{d^3k}{(2\pi)^3} \mathbf{v}(\mathbf{k}) + (-e) \int_{\text{unoccupied}} \frac{d^3k}{(2\pi)^3} \mathbf{v}(\mathbf{k}) \quad (14.9)$$

From this we obtain an equivalent way of describing the current:

$$\mathbf{j} = +e \int_{\text{unoccupied}} \frac{d^3k}{(2\pi)^3} \mathbf{v}(\mathbf{k}) \quad (14.10)$$

which describes the current in terms of positive charge carriers corresponding to the unoccupied parts of the band. This has practical use (e.g.) for a band which is mostly filled - where it is much easier to compute with the unoccupied states, rather than the occupied ones.

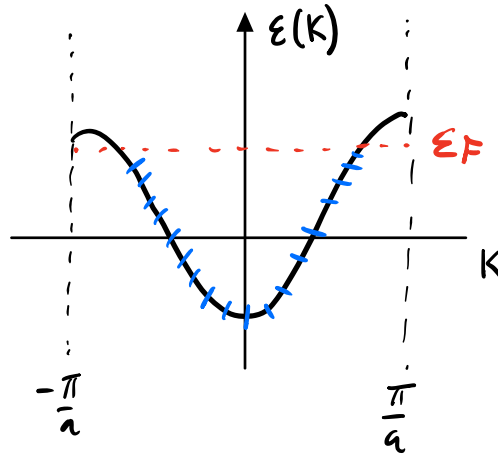


Figure 14.2: A sketch of a mostly filled band, for which the hole picture would be more convenient to do calculations in.

14.6 Motion in Uniform Magnetic Field

We now consider a scenario with a uniform \mathbf{H} -field (and zero electric field). We then have:

$$\hbar \frac{d\mathbf{k}}{dt} = -e(\mathbf{v} \times \mathbf{H}) \quad (14.11)$$

With $\epsilon_{\mathbf{k}} = \frac{\hbar^2(k_x^2 + k_y^2)}{2m}$, we have that $\mathbf{v} = \frac{1}{\hbar} \nabla_{\mathbf{k}} \epsilon_{\mathbf{k}} = \frac{1}{\hbar} \frac{\hbar^2}{m} (k_x, k_y) = \frac{\hbar}{m} (k_x, k_y)$ (Note: \mathbf{k} moves through state space, but \mathbf{v} is the real motion of velocity!). From this it follows that:

$$\frac{d\mathbf{k}}{dt} \sim \mathbf{v} \cdot \boldsymbol{\theta} \quad (14.12)$$

where $\boldsymbol{\theta}$ is some tangential vector. So, the electrons will move on intersections of constant $\epsilon_{\mathbf{k}}$ and surfaces perpendicular to \mathbf{H} !

15 Semiclassical Theory of Conduction II

We have our configurational space \mathbf{k}, \mathbf{r} . For some volume in this space, as we go infinitesimal we have:

$$g_n(\mathbf{k}, \mathbf{r}, t) \frac{d\mathbf{k}d\mathbf{r}}{4\pi^3} = dN \quad (15.1)$$

where g_n is the density of states. In equilibrium, we have that:

$$g_n \rightarrow g_0 = \frac{1}{e^{\beta(\epsilon_{\mathbf{k}} - \mu)} + 1}. \quad (15.2)$$

If we recall the Drude model, there we looked at a toy model of electrons travelling through a wire, with equation of motion:

$$\frac{d\mathbf{p}}{dt} = \mathbf{F} - \frac{\mathbf{p}}{\tau} \quad (15.3)$$

where the probability of a collision in an infinitesimal time interval dt given by $dp = \frac{dt}{\tau}$ with τ the relaxation time and $\int dp = 1$.

15.1 Non-equilibrium Distribution Function

With this picture in mind, we have that the number of electrons emerging from each collision after time dt is:

$$dg_n = \frac{dt}{\tau(\mathbf{k}, \mathbf{r})} g_0(\mathbf{k}, \mathbf{r}, t) \frac{d^3k d^3r}{4\pi^3} \quad (15.4)$$

where the $g_0(\mathbf{k}, \mathbf{r}, t) \frac{d^3k d^3r}{4\pi^3}$ are the number of electrons that appear in the volume.

Now, let us assume that $g_0(\mathbf{k}, \mathbf{r}, t)$ is unaltered by collisions, that g_0 does not depend on g (and as above, $g \rightarrow g_0$ in equilibrium). The trajectory of an electron is described by $\mathbf{r}(t), \mathbf{k}(t)$.

Define $P(t, t')$ as the number of electrons that make it to $\mathbf{k}(t), \mathbf{r}(t)$ without a collision from $\mathbf{k}(t'), \mathbf{r}(t')$. Note that some electrons will leave this trajectory due to collisions.

We then have that:

$$dN = g(\mathbf{k}, \mathbf{r}, t) \frac{d^3k d^3r}{4\pi^3} = \int_{-\infty}^t P(t, t') \frac{dt'}{\tau(\mathbf{k}(t'), \mathbf{r}(t'))} g_0(\mathbf{k}(t'), \mathbf{r}(t'), t) \frac{d^3k d^4r}{4\pi^3} \quad (15.5)$$

It is worth noting some terminology now - g is the non-equilibrium distribution function, while g_0 is the equilibrium distribution function.

Now, we have:

$$P(t, t') = P(t, t' + dt') \left(1 - \frac{dt'}{\tau(t')} \right) \quad (15.6)$$

so in the limit where $dt' \rightarrow 0$:

$$\frac{\partial P(t, t')}{\partial t'} = \frac{P(t, t')}{\tau(t')} \quad (15.7)$$

which has solution:

$$P(t, t') = e^{-\int_{t'}^t \frac{dt''}{\tau(t'')}} \quad (15.8)$$

and we note that $P(t, t) = 1$ (of course, as there are no collisions).

Using this, we can write the non-equilibrium distribution function as:

$$\begin{aligned}
g(t) &= \int_{-\infty}^t dt' g_0(t') \frac{P(t, t')}{\tau(t')} \\
&= \int_{-\infty}^t dt' g_0(t') \frac{\partial P(t, t')}{\partial t'} \\
&= g_0(t') P(t, t') \Big|_{-\infty}^t - \int_{-\infty}^t P(t, t') \frac{\partial}{\partial t'} g_0(t') \quad (\text{Integration by parts}) \\
&= g_0(t) P(t, t) - g_0(-\infty) P(t, -\infty) - \int_{-\infty}^t P(t, t') \frac{\partial g_0(t')}{\partial t'}
\end{aligned} \tag{15.9}$$

Now noting that $P(t, t) = 1$ and $P(t, -\infty) = 0$ (as no electrons from $-\infty$ make it to t):

$$g(t) = g_0(t) - \int_{-\infty}^t dt' P(t, t') \frac{dg_0(t')}{dt} \tag{15.10}$$

Now, with $g_0(\mathbf{k}, \mathbf{r}) = \frac{1}{e^{\beta(\epsilon_{\mathbf{k}} - \mu)} + 1}$, and so $g_0(\mathbf{k}(t), \mathbf{r}(t)) = \frac{1}{e^{\beta(\mathbf{r}(t))(\epsilon(\mathbf{k}(t)) - \mu(\mathbf{r}(t)))} + 1}$ (where this corresponds to “local” occupation assuming “local thermodynamic equilibrium”), we compute with the chain rule that:

$$\frac{dg_0}{dt} = \frac{\partial g_0}{\partial \epsilon_{\mathbf{k}}} \frac{\partial \epsilon_{\mathbf{k}}}{\partial \mathbf{k}} \frac{d\mathbf{k}}{dt} + \frac{\partial g_0}{\partial \mu} \frac{\partial \mu}{\partial \mathbf{r}} \frac{d\mathbf{r}}{dt} + \frac{\partial g_0}{\partial T} \frac{\partial T}{\partial \mathbf{r}} \frac{d\mathbf{r}}{dt} \tag{15.11}$$

Now, recalling that:

$$\begin{aligned}
\frac{d\mathbf{r}}{dt} &= \mathbf{v}_n(\mathbf{k}) = \frac{1}{\hbar} \nabla_{\mathbf{k}} \epsilon_{\mathbf{k}} \\
\hbar \frac{d\mathbf{k}}{dt} &= -e(\mathbf{E} + \frac{\mathbf{v}}{c} \times \mathbf{H})
\end{aligned} \tag{15.12}$$

we find:

$$\frac{\partial g_0}{\partial \epsilon_{\mathbf{k}}} \frac{\partial \epsilon_{\mathbf{k}}}{\partial \mathbf{k}} \frac{d\mathbf{k}}{dt} = \frac{\partial f}{\partial \epsilon} \mathbf{v} \cdot e(\mathbf{E} + \frac{\mathbf{v}}{c} \times \mathbf{H}) = \frac{\partial f}{\partial \epsilon} \mathbf{v} \cdot e\mathbf{E} \tag{15.13}$$

where the second term vanishes as $\mathbf{v} \times \mathbf{H}$ is perpendicular to \mathbf{v} . For the other terms:

$$\frac{\partial g_0}{\partial \mu} \frac{\partial \mu}{\partial \mathbf{r}} \frac{d\mathbf{r}}{dt} = -\frac{\partial f}{\partial \epsilon} \nabla \mu \cdot \mathbf{v} \tag{15.14}$$

$$\frac{\partial g_0}{\partial T} \frac{\partial T}{\partial \mathbf{r}} \frac{d\mathbf{r}}{dt} = -\frac{\partial f}{\partial \epsilon} \left(\frac{\epsilon - \mu}{T} \right) \nabla T \cdot \mathbf{v} \tag{15.15}$$

So therefore the non-equilibrium distribution function can be written as:

$$g(t) = g_0(t) + \int_{-\infty}^t dt' P(t, t') \left[-\frac{\partial f}{\partial \epsilon} \mathbf{v} \cdot \left(-e\mathbf{E} - \nabla \mu - \frac{\epsilon - \mu}{T} \nabla T \right) \right] \tag{15.16}$$

where $P(t, t') = e^{-(t-t')/\tau(\epsilon_{\mathbf{k}})}$ assuming $\tau = \tau(\epsilon_{\mathbf{k}})$.

Note that last lecture we wrote down a current as:

$$\mathbf{j} = -e \int_{\epsilon < \epsilon_F} \frac{dk}{4\pi^3} \mathbf{v}(\mathbf{k}) \tag{15.17}$$

with the equilibrium distribution (Fermi-Dirac distribution) we now generalize this to:

$$\mathbf{j} = -e \int \frac{d^3k}{4\pi^3} \mathbf{v}(\mathbf{k}) \cdot g(\mathbf{k}) \tag{15.18}$$

where $g(\mathbf{k})$ corresponds to some other, messier distribution.

16 Semiclassical Theory of Conduction III

16.1 Solving for Non-Equilibrium Distribution

We have:

$$g(t) = g^0(t) = \int_{-\infty}^t dt' P(t, t') \left[\left(-\frac{\partial f}{\partial \epsilon} \right) \mathbf{v} \cdot \left(-e\mathbf{E} - \nabla\mu - \frac{\epsilon - \mu}{T} \nabla T \right) \right] \quad (16.1)$$

$$P(t, t') = \exp\left(-\int_{t'}^t \frac{d\bar{t}}{\tau(\bar{t})}\right) \quad (16.2)$$

$$\dot{\mathbf{r}} = \mathbf{v}, \quad \hbar \dot{\mathbf{k}} = -e \left[\mathbf{E} + \frac{1}{c} (\mathbf{v} \times \mathbf{B}) \right] \quad (16.3)$$

But in this form solving for the non-equilibrium distribution is only possible numerically. However, there are three simplifying assumptions that allow for us to get a closed form:

1. Weak \mathbf{E} -field and ∇T . Usually this is a good approximation - most of the time it is impossible to apply a sufficiently strong \mathbf{E} field or a thermal gradient such that these terms matter significantly.
2. Spatially uniform \mathbf{E} and ∇T . This again is a good approximation; \mathbf{E} will likely be uniform on the length-scale of electron collisions, as will thermal gradients. Under these assumptions, the energy (which is the only place where t' enters) only depends on the magnetic field - however the magnetic field conserves energy, so t' enters only in the first piece of the integral!
3. The relaxation time depends \mathbf{k} only through $\epsilon_n(\mathbf{k})$. This is the energy-dependent relaxation approximation. It is the hardest approximation to justify, and there isn't a deep reason for it beyond making the calculations simple. This implies that $P(t, t') = e^{-(t-t')/\tau_n(\mathbf{k})}$.

With these assumptions, the non-equilibrium distribution takes the form (note: there is no more spatial dependence/translationally invariant situation here!):

$$g(\mathbf{k}, t) = g^0(\mathbf{k}) + \int_{-\infty}^t dt' e^{-(t-t')/\tau(\epsilon(\mathbf{k}))} \left[\left(-\frac{\partial f}{\partial \epsilon} \right) \mathbf{v}(\mathbf{k}(t')) \cdot [-e\mathbf{E}(t') - \nabla\mu(t') - \frac{\epsilon(\mathbf{k}) - \mu}{T} \nabla T(t')] \right] \quad (16.4)$$

16.2 DC electrical conductivity

We consider $\mathbf{B} = \mathbf{0}$, $\nabla T = \mathbf{0}$, $\mathbf{E} = E_0 \hat{\mathbf{n}} \neq 0$. In this scenario, \mathbf{v} is a constant, \mathbf{E} is a constant, and so we can take things out of the integral and perform the remaining integral, which is trivial:

$$g(\mathbf{k}) = g^0(\mathbf{k}) - e\mathbf{E} \cdot \mathbf{v}(\mathbf{k}) \tau(\epsilon(\mathbf{k})) \left(-\frac{\partial f}{\partial \epsilon} \right) \quad (16.5)$$

We can obtain the electrical current as:

$$\mathbf{j} = -e \int \frac{d^3k}{4\pi^3} \mathbf{v}(\mathbf{k}) g(\mathbf{k}) = \hat{\sigma} \mathbf{E} \quad (16.6)$$

Note the g^0 term will give us no contribution as we have no current at equilibrium. $\hat{\sigma}$ is the *conductivity tensor*. This is the generalized Ohm's law. We can get it by summing the per-band conductivity tensors (here expressed in their Cartesian components):

$$\hat{\sigma} = \sum_n \hat{\sigma}^{(n)}, \quad \hat{\sigma}_{\mu\nu}^{(n)} = e^2 \int \frac{d^3k}{4\pi^3} \tau_n(\epsilon_n(\mathbf{k})) v_n^\mu(\mathbf{k}) v_n^\nu(\mathbf{k}) \left(-\frac{\partial f}{\partial \epsilon} \right)_{\epsilon=\epsilon_n(\mathbf{k})} \quad (16.7)$$

Note that $v_n, \partial f$ are relatively well-known but $\tau(\epsilon_n(\mathbf{k}))$ is generally not. This is because relaxation can come from many sources. It is hard to theoretically calculate.

Some remarks:

- (i) Anisotropy - In general because $\hat{\sigma}$ is a tensor, it is possible to have a scenario where \mathbf{j} is not aligned with \mathbf{E} , which can occur when $\hat{\sigma}$ has off-diagonal elements. However, one can show that for free electrons, and for cubic crystals, $\sigma_{\mu\nu} = \sigma_0 \delta_{\mu\nu}$ by symmetry so $\mathbf{j} \parallel \mathbf{E}$. In less symmetric crystals (e.g. orthorhombic) we can see that they are not parallel.
- (ii) Filled bands do not contribute to $\hat{\sigma}$ at $T = 0$. Obviously empty bands would contribute nothing, but it is more surprising that a filled band does not contribute either. Formally, we can see it from the $\left(-\frac{\partial f}{\partial \epsilon}\right)_{\epsilon=\epsilon_n(\mathbf{k})}$ - at $T = 0$ the fermi function is a step function, with the step at ϵ_F , so if we take a derivative, it is just $\delta(\epsilon - \epsilon_F)$. But since we evaluate at $\epsilon = \epsilon_n(\mathbf{k})$ a band energy, the derivative must vanish. You will get to explore this in the HW, where you will study the conduction of a band insulator at finite temperature.
- (iii) Equivalence of electron and hole pictures. We consider a situation where the band structure looks as:
Now, consider:

$$\mathbf{v}(\epsilon) \left(-\frac{\partial f}{\partial \epsilon}\right)_{\epsilon=\epsilon_n(\mathbf{k})} = -\frac{1}{\hbar} \frac{\partial}{\partial \mathbf{k}} f(\epsilon(\mathbf{k})) \quad (16.8)$$

so:

$$\hat{\sigma}^n = e^2 \tau(\epsilon_F) \int \frac{d^3k}{4\pi^3} \frac{\partial \mathbf{v}_n(\mathbf{k})}{\partial \mathbf{k}} f(\epsilon(\mathbf{k})) = e^2 \tau(\epsilon_F) \int_{\text{occ. levels}} \frac{d^3k}{4\pi^3} M^{-1}(\mathbf{k}) \quad (16.9)$$

where we have used an integration by parts to rewrite $\hat{\sigma}^n$, and are assuming that $T = 0$. Note that

$M_{\mu\nu}^{-1}(\mathbf{k}) = \frac{1}{\hbar^2} \frac{\partial^2 \epsilon(\mathbf{k})}{\partial k_\mu \partial k_\nu} = \frac{1}{\hbar} \frac{\partial v_\mu(\epsilon)}{\partial k_\nu}$. We can then write:

$$\hat{\sigma} = e^2 \tau(\epsilon_F) \left[\int_{\text{all states}} - \int_{\text{empty states}} \right] \quad (16.10)$$

But if we integrate over the entire Brioullin zone, then $\int \frac{d^3k}{4\pi^3} M^{-1}(\mathbf{k}) = 0$ as we integrate over a periodic function. The bottom line is:

$$\hat{\sigma} = -e^2 \tau(\epsilon_F) \int_{\text{empty states}} \frac{d^3k}{4\pi^3} M^{-1}(\mathbf{k}) \quad (16.11)$$

which is the “hole picture”.

16.3 Thermal Conductivity

Everyone is familiar with electrical conductivity measurements. Thermal conductivity is similar; we apply some thermal gradient to our material, and then measure the thermal current. There is some subtlety here; not only does heat flow, but charge also flows - we produce both j^q (heat current) but also j (electrical current). Thermal conductivity is defined as the ratio of the heat current to the thermal gradient, under the condition which no electrical current flows.

Heat current is also a bit subtle in its definition. It's not a pure energy current, one also has to consider entropy. We write down the two relations (from our course in thermodynamics) and draw analogies:

$$\begin{aligned} dQ &= TdS \leftrightarrow \mathbf{j}^q = T\mathbf{j}^s \\ TdS &= dU - \mu dN \leftrightarrow T\mathbf{j}^s = \mathbf{j}^e - \mu\mathbf{j}^n \end{aligned} \quad (16.12)$$

so we suddenly have four currents to worry about. We do this because $\mathbf{j}^\epsilon, \mathbf{j}^n$ are easy to calculate from our semi-classical theory, but it is \mathbf{j}^q that we want. So, next we will calculate \mathbf{j}^ϵ and \mathbf{j}^n , and use these relations to reconstruct the electrical and heat current. We write:

$$\begin{pmatrix} \mathbf{j}^\epsilon \\ \mathbf{j}^n \end{pmatrix} = \sum_n \int \frac{d^3k}{4\pi} \begin{pmatrix} \epsilon_n(\mathbf{k}) \\ 1 \end{pmatrix} \mathbf{v}_n(\mathbf{k}) g_n(\mathbf{k}) \quad (16.13)$$

The heat current is then:

$$\mathbf{j}^q = \mathbf{j}^\epsilon - \mu \mathbf{j}^n = \sum_n \int \frac{d^3k}{4\pi^3} (\epsilon_n(\mathbf{k}) - \mu) v_n(\mathbf{k}) g_n(\mathbf{k}) \quad (16.14)$$

Now what we do is we take the non-equilibrium distribution function with all the simplifying assumptions, and stick it in. Now we have both potentially a nonzero electric field and thermal gradient. Once again the contribution from the equilibrium piece will vanish. What we will get is the following:

$$\begin{aligned} \mathbf{j} &= \hat{L}^{11} \epsilon + \hat{L}^{12} (-\nabla T) \\ \mathbf{j}^q &= \hat{L}^{21} \epsilon + \hat{L}^{22} (-\nabla T) \end{aligned} \quad (16.15)$$

where $\epsilon = \mathbf{E} + \frac{\nabla \mu}{e}$. Note we have the obvious diagonal pieces (electric field contributes to electrical current, thermal gradient contributes thermal gradient) but also off-diagonal; thermoelectric effects! E.g. $\hat{L}^{21} \epsilon$ explains refrigerators, where we use electricity to remove heat. $\hat{L}^{12} (-\nabla T)$ explains space probes, which use heat from the sun to generate electricity. The theory we have developed allows us to calculate these coefficients, and if we summarize how this looks:

$$\hat{L}^{11} = \mathcal{L}^{(0)}, \hat{L}^{22} = \frac{1}{e^2 T} \mathcal{L}^{(2)}, \hat{L}^{21} = T \hat{L}^{12} = -\frac{1}{e} \mathcal{L}^{(1)} \quad (16.16)$$

where:

$$\begin{aligned} \mathcal{L}^{(\alpha)} &= e^2 \int \frac{d^3k}{4\pi^3} \left(-\frac{\partial f}{\partial \epsilon} \right) \tau(\epsilon(\mathbf{k})) v_\mu(\mathbf{k}) v_\nu(\mathbf{k}) [\epsilon(\mathbf{k}) - \mu]^\alpha \\ &= \int d\epsilon \left(-\frac{\partial f}{\partial \epsilon} \right) (\epsilon - \mu)^\alpha \hat{\Pi}(\epsilon) \end{aligned} \quad (16.17)$$

$$\text{where } \hat{\Pi}_{\mu\nu}(\epsilon) = e^2 \tau(\epsilon) \int \frac{d^3k}{4\pi^3} \delta(\epsilon - \epsilon(\mathbf{k})) v_\mu(\mathbf{k}) v_\nu(\mathbf{k})$$

so therefore:

$$\begin{aligned} \hat{L}^{11} &= \hat{\Pi}(\epsilon_F) = \hat{\sigma} \\ \hat{L}^{21} &= T \hat{L}^{12} = -\frac{\pi^2}{3e} (k_B T)^2 \hat{\Pi}' \\ \hat{L}^{22} &= \frac{\pi^2}{3} \frac{k_B^2 T}{e^2} \hat{\sigma} \end{aligned} \quad (16.18)$$

where $\hat{\Pi}' = \left(\frac{\partial \hat{\Pi}(\epsilon)}{\partial \epsilon} \right)_{\epsilon=\epsilon_F}$. These formulas are valid at $k_B T \ll \epsilon_F$ (Sommerfield expansion). The diagonal components are simple and relate to the electrical conductivity tensor. But there are off-diagonal effects which underlie the thermoelectric properties of materials.

The last thing we will do is to extract the thermal conductivity - we want a situation where we apply a thermal gradient, and there is zero electrical current (only a thermal current). To this end, we must apply an electric field such that the electric field cancels out the thermoelectric term in \mathbf{j} :

$$\mathbf{j} = 0 \implies \epsilon = -(\hat{L}^{11})^{-1} \hat{L}^{12} (-\nabla T) \quad (16.19)$$

This then gives:

$$\mathbf{j}^q = \hat{K}(-\nabla T) \quad (16.20)$$

where:

$$\hat{K} = \hat{L}^{22} - \hat{L}^{21}(\hat{L}^{11})^{-1}\hat{L}^{12} \quad (16.21)$$

as expected, the dominant piece is \hat{L}^{22} , but there is a correction coming from the thermo-electric effects.

One can show that the thermoelectric piece is on the order of $\left(\frac{k_B T}{\epsilon_F}\right)^2$ which is small. So one finds that:

$$\hat{K} \approx \hat{L}^{22} + O\left(\left(\frac{k_B T}{\epsilon_F}\right)^2\right) \xrightarrow{T \rightarrow 0} \hat{L}^{22} \quad (16.22)$$

so recalling \hat{L}^{22} , we find at low T in metals that:

$$\hat{K} = \frac{\pi^2}{3} \left(\frac{k_B}{e}\right)^2 \hat{\sigma} \quad (16.23)$$

which is known as the “Wiedemann-Franz”⁶ law. It is reproduced in experiment. It is related to the fact that in some regimes, the heat is carried in electrons, which carry one unit of electric charge, and $k_B T$ of energy. In the assigned homework, we will see that this is *not* true in a semiconductor - it is strongly violated.

This concludes the semiclassical theory of conduction in metals. We will now switch gears and discuss electron-phonon interactions next class.

⁶Unfortunately, no relation to Marcel as far as we know

17 Electron-Phonon Interactions

We discussed tight-binding electrons which lead to bands, and lattice vibrations which lead to phonons - but both happen at the same time, so we now ask how do these two interact/intertwine?

Despite a fairly simple Hamiltonian that describes this, it's a surprisingly difficult problem, with unsolved questions remaining to this day.

17.1 Deriving the Electron-Phonon Interaction Hamiltonian

We proceed by considering electrons moving in the ionic potential, but we will not assume the ions are frozen in a perfectly periodic formation. We write:

$$H_1 = \sum_{\mathbf{k}, \mathbf{k}', l} \langle \mathbf{k} | U_0(\mathbf{r} - \mathbf{R}_l - \boldsymbol{\mu}_l) | \mathbf{k}' \rangle c_{\mathbf{k}}^\dagger c_{\mathbf{k}'} \quad (17.1)$$

where U_0 is a single-ion potential, \mathbf{R}_l is the ion equilibrium position, and $\boldsymbol{\mu}_l$ the ion displacement.

As usual, we work in a basis where $|\mathbf{k}\rangle \cong \frac{1}{\sqrt{V}} e^{i\mathbf{k} \cdot \mathbf{r}}$. So we calculate the matrix element, which amounts to fourier transforming the potential of these ions. What we get out is;

$$H_1 = \sum_{\mathbf{k}, \mathbf{k}', l} e^{i(\mathbf{k}' - \mathbf{k}) \cdot (\mathbf{R}_l + \boldsymbol{\mu}_l)} V_{\mathbf{k} - \mathbf{k}'} c_{\mathbf{k}}^\dagger c_{\mathbf{k}'} \quad (17.2)$$

Here, $V_{\mathbf{k}}$ is the FT of $U_0(\mathbf{r})$. We will now proceed to make the assumption that $|\boldsymbol{\mu}_l| \ll a$ and expand:

$$e^{i(\mathbf{k}' - \mathbf{k}) \cdot \boldsymbol{\mu}_l} \approx 1 + i(\mathbf{k}' - \mathbf{k}) \cdot \boldsymbol{\mu}_l = 1 + \frac{i}{\sqrt{N}} (\mathbf{k}' - \mathbf{k}) \sum_{\mathbf{q}} e^{i\mathbf{q} \cdot \mathbf{R}_l} \boldsymbol{\mu}_{\mathbf{q}} \quad (17.3)$$

We now substitute this back into the equation. The leading term (1) is just the Bloch Hamiltonian:

$$H_{\text{Bloch}} = \sum_{\mathbf{k}, \mathbf{k}'} \left(\sum_l e^{i(\mathbf{k} - \mathbf{k}') \cdot \mathbf{R}_l} \right) V_{\mathbf{k} - \mathbf{k}'} c_{\mathbf{k}}^\dagger c_{\mathbf{k}'} = N \sum_{\mathbf{k}, \mathbf{G}} V_{\mathbf{G}} c_{\mathbf{k} + \mathbf{G}}^\dagger c_{\mathbf{k}} \quad (17.4)$$

this we already studied. Going to the next term, we get electron-phonon interaction:

$$H_{e-p} = \frac{i}{\sqrt{N}} \sum_{\mathbf{k}, \mathbf{k}'} (\mathbf{k}' - \mathbf{k}) \cdot \boldsymbol{\mu}_{\mathbf{k} - \mathbf{k}'} V_{\mathbf{k} - \mathbf{k}'} c_{\mathbf{k}}^\dagger c_{\mathbf{k}'} \quad (17.5)$$

where we again have carried out the sum over lattice indices l . Now, let us express $\boldsymbol{\mu}_{\mathbf{q}}$ in terms of phonon operators:

$$H_{e-p} = i \sum_{\mathbf{k}, \mathbf{k}', s} \left(\frac{N\hbar}{2M\omega_{\mathbf{k} - \mathbf{k}', s}} \right)^{1/2} (\mathbf{k}' - \mathbf{k}) \cdot \mathbf{s} V_{\mathbf{k} - \mathbf{k}'} (a_{\mathbf{k}' - \mathbf{k}, s}^\dagger + a_{\mathbf{k} - \mathbf{k}', s}) c_{\mathbf{k}}^\dagger c_{\mathbf{k}'} \quad (17.6)$$

Looking at the creation/annihilation operators, we can see that an electron with momentum \mathbf{k}' gets destroyed and a new one with momentum \mathbf{k} is created. In the process, a phonon with that difference is either emitted or created. Diagrammatically:

Also, look at $(\mathbf{k}' - \mathbf{k}) \cdot \mathbf{s}$; this means that only longitudinal phonons are important. In a generic solid, one cannot classify longitudinal/transverse phonons, but let us assume a sufficient amount of isotropy in the solid such that this is possible.

In the following, we make two assumptions:

- (i) We only have longitudinal phonons.
- (ii) Neglect the effect of H_{Bloch} - study el-ph interactions in the limit of free-electrons.



Figure 17.1: Feynman Diagrams for the processes in H_{e-p} .

We therefore have the full Hamiltonian:

$$H = \sum_{\mathbf{k}} \epsilon_{\mathbf{k}} c_{\mathbf{k}}^{\dagger} c_{\mathbf{k}} + \sum_{\mathbf{q}} \omega_{\mathbf{q}} a_{\mathbf{q}}^{\dagger} a_{\mathbf{q}} + \sum_{\mathbf{k}, \mathbf{q}} M_{\mathbf{q}} (a_{-\mathbf{q}}^{\dagger} + a_{\mathbf{q}}) c_{\mathbf{k}+\mathbf{q}}^{\dagger} c_{\mathbf{k}} \quad (17.7)$$

where $M_{\mathbf{q}} = i\sqrt{\frac{N\hbar}{2M\omega_{\mathbf{q}}}} |\mathbf{q}| V_{\mathbf{q}}$. The first term we can call H_0 (this just describes “free” electrons and phonons) and the second term we can call H' (which describes the interactions). Even solving this simplified Hamiltonian we will find is nontrivial - the remainder of the lecture will be studying some of its consequences.

17.2 Kohn Anomaly and Peierls Instability

This describes the effect of the Fermi sea on phonons. The same Kohn who is responsible for DFT! We consider a single phonon propagating in the presence of the Fermi sea. Assume a weak el-ph interaction, and treat H' as a perturbation.

Consider unperturbed eigenstates $|\phi_i\rangle$, that satisfy $H_0|\phi_i\rangle = \epsilon_i|\phi_i\rangle$. Consider $|\phi_1\rangle = |FS\rangle a_{\mathbf{p}}^{\dagger} |0\rangle$ where $|FS\rangle$ describes the fermi sea of electrons and $a_{\mathbf{p}}^{\dagger} |0\rangle$ describes a single-phonon state.

Now, let's use second-order perturbation theory to calculate the perturbed energies.

$$E_1 \approx E_1^{(0)} + \langle \phi_1 | H' | \phi_1 \rangle + \langle \phi_1 | H' (E_1^{(0)} - H_0)^{-1} H' | \phi_1 \rangle \quad (17.8)$$

$E_1^{(0)} = \epsilon_1$ is just the unperturbed energy. $\langle \phi_1 | H' | \phi_1 \rangle = 0$ as there is an odd number of phonon operators in H' . Now, for the second order term - this might seem unfamiliar compared to what you are used to. If you insert the completeness relation $\mathbb{I} = \sum_i |\phi_i\rangle \langle \phi_i|$ twice (next to the H' 's) you will reproduce the usual expression. Importantly, it is only at second order where one finds interesting physics.

So, let's calculate this second order term:

$$E_1^{(2)} = \langle \phi_1 | \sum_{\mathbf{k}, \mathbf{q}} M_{\mathbf{q}} (a_{-\mathbf{q}}^{\dagger} + a_{\mathbf{q}}) c_{\mathbf{k}+\mathbf{q}}^{\dagger} c_{\mathbf{k}} (\epsilon_1 - H_0)^{-1} \sum_{\mathbf{k}', \mathbf{q}'} M_{\mathbf{q}'} (a_{-\mathbf{q}'}^{\dagger} + a_{\mathbf{q}'}) c_{\mathbf{k}'+\mathbf{q}'}^{\dagger} c_{\mathbf{k}'} | \phi_1 \rangle \quad (17.9)$$

we make our lives simpler by the following observation - if we annihilate a phonon then we must create another, or if we create a phonon we must annihilate another. So, there are two possible pairings which give a nonzero contribution, namely $a_{-\mathbf{q}}^{\dagger} a_{\mathbf{q}'}$ and $a_{\mathbf{q}} a_{-\mathbf{q}'}^{\dagger}$.

It turns out these two pairings correspond to different processes. The first term will be important for the question we are addressing at the moment - we put the other term on the backburner, as it explains how phonons effect the behaviour of the electrons.

First of all, in the initial state we only have one phonon with momentum \mathbf{p} , so for $a_{\mathbf{q}'} | \phi_1 \rangle$ to not vanish we require $\mathbf{q}' = \mathbf{p}$, and further we have to recreate the same phonon, so $\mathbf{q}' = \mathbf{p} = -\mathbf{q}$. This reduces two of the sums. We are then left with the \mathbf{k}, \mathbf{k}' sums. By considering that the same electron must be destroyed and recreated in the Fermi sea, we obtain $\mathbf{k}' = \mathbf{k} + \mathbf{q}$.

The other thing to consider; $(\epsilon_1 - H_0)^{-1}$ looks nasty, like the inverse of some operator. We will not be taking the inverse of anything here; we just consider that H_0 when acting on an eigenstate should yield the correct energy, so we replace:

$$(\epsilon_1 - H_0)^{-1} \rightarrow -(\epsilon_{\mathbf{k}'+\mathbf{p}} - \epsilon_{\mathbf{k}'} + \hbar\omega_{\mathbf{p}}) \quad (17.10)$$

so then:

$$E_1^{(2)} = -\sum_{\mathbf{k}} |M_{\mathbf{p}}|^2 \frac{\langle \phi_1 | a_{\mathbf{p}}^\dagger c_{\mathbf{k}-\mathbf{p}}^\dagger c_{\mathbf{k}} c_{\mathbf{k}}^\dagger c_{\mathbf{k}-\mathbf{p}} a_{\mathbf{p}} | \phi_1 \rangle}{\epsilon_{\mathbf{k}} - \epsilon_{\mathbf{k}-\mathbf{p}} + \hbar\omega_{\mathbf{p}}} = -\sum_{\mathbf{k}} |M_{\mathbf{p}}|^2 \frac{\langle \phi_1 | c_{\mathbf{k}-\mathbf{p}}^\dagger c_{\mathbf{k}-\mathbf{p}} (1 - c_{\mathbf{k}}^\dagger c_{\mathbf{k}}) a_{\mathbf{p}}^\dagger a_{\mathbf{p}} | \phi_1 \rangle}{\epsilon_{\mathbf{k}} - \epsilon_{\mathbf{k}-\mathbf{p}} + \hbar\omega_{\mathbf{p}}} \quad (17.11)$$

so if we beautify our expression with $\mathbf{k} \rightarrow \mathbf{k} + \mathbf{p}$:

$$E_1^{(2)} = \hbar\delta\omega_{\mathbf{p}} = -|M_{\mathbf{p}}|^2 \sum_{\mathbf{k}} \frac{n_{\mathbf{k}}(1 - n_{\mathbf{k}+\mathbf{p}})}{\epsilon_{\mathbf{k}+\mathbf{p}} - \epsilon_{\mathbf{k}} + \hbar\omega_{\mathbf{p}}} \quad (17.12)$$

where we note the definition of the number operator, and that $a_{\mathbf{p}}^\dagger a_{\mathbf{p}} = 1$ as we have one phonon with momentum \mathbf{p} .

17.2.1 3 dimensions - Kohn Anomaly

In 3-dimensions, note that $\delta\omega_{\mathbf{p}}$ is finite everywhere but has an infinite slope $\frac{\partial\omega_{\mathbf{p}}}{\partial\mathbf{p}} \rightarrow \infty$ as $|\mathbf{p}| \rightarrow 2k_F$. This is known as the *Kohn Anomaly*, and is often observed in metals (which have a Fermi momenta). This is interesting - you can use this technique to map out the Fermi surface of the metal, if you are able to measure the dispersion of a phonon in sufficient detail!

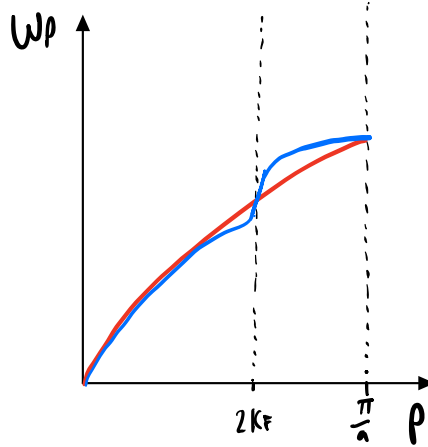


Figure 17.2: Dispersion Relation for 3-D electron-phonon interaction Hamiltonian. The derivative of the dispersion diverges at $p = 2k_F$.

What's the intuition behind this? Diagrammatically, we have a electron-hole bubble:

and why does this give us the intuition for the anomaly? Intuitively, the Kohn anomaly can be interpreted as phonon spending some fraction of time as a electron-hole pair. Because $v_F \gg v_{ph}$, the phonon speed of propagation increases dramatically. Why only near the fermi momentum? Kinematic constraints - you need the correct amount of momentum to create an electron-hole pair. At precisely $2k_F$, the fraction of time the phonon spends as an electron-hole pair approaches 100%. Note a more careful analysis of the behaviour near $2k_F$ would show that the derivative of the dispersion does not actually diverge but approaches the fermi velocity, as we might expect.

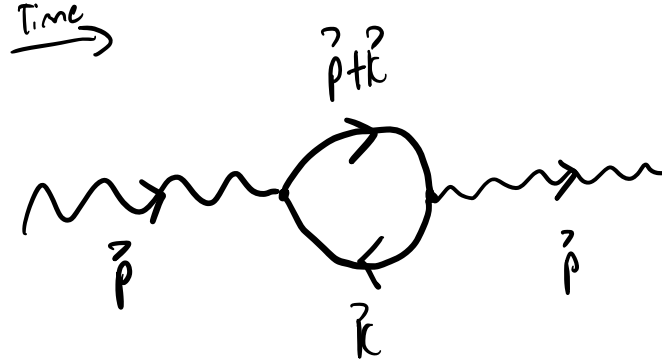


Figure 17.3: Feynman Diagram of Electron-Hole Bubble.

17.2.2 1 dimension - Peirels Instability

Assume $|\omega_q| \ll \epsilon_k, \epsilon_{k+p}$. We then evaluate:

$$\hbar\delta\omega_p = - \int_{-\infty}^{\infty} dk \frac{n_k(1 - n_{k-q})}{\epsilon_{k-q} - \epsilon_k} \quad (17.13)$$

but the integral vanishes everywhere where the product of $n_k, (1 - n_{k-q})$ is not supported, i.e. everywhere except $[-k_F, -k_F + q]$ (see sketch below)

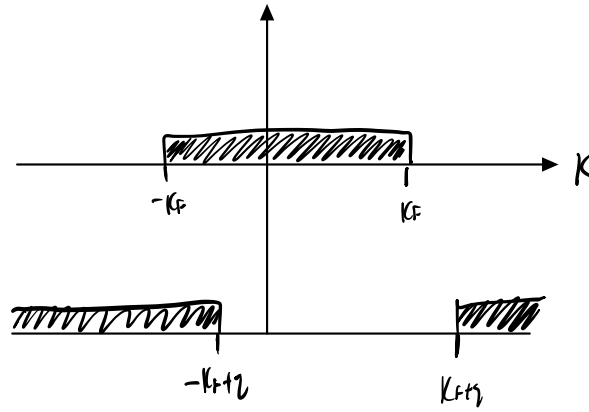


Figure 17.4: Sketch of the support of $n_k, 1 - n_{k-q}$.

so:

$$\hbar\delta\omega_p = - \int_{-k_F}^{-k_F+q} dk \frac{1}{(k-q)^2 - k^2} = \frac{1}{2q} \ln|q-2k| \Big|_{-k_F}^{q-k_F} = \frac{1}{2q} (\ln|q-2k_F| - \ln|q+2k_F|) \quad (17.14)$$

and so we conclude:

$$\delta\omega_q \sim \frac{1}{2q} (\ln|q-2k_F| - \ln|q+2k_F|) \quad (17.15)$$

Although we assume $q > 0$, the same occurs for $q < 0$, as the above is symmetric in the sign of q . But now notice - this correction is divergent when $q = \pm 2k_F$! This is a much stronger effect - in 3D the correction was finite but the derivative diverged, but here the correction itself diverges.

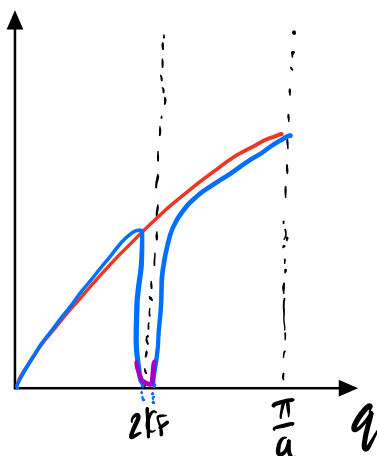


Figure 17.5: Dispersion Relation for 1-D electron-phonon interaction Hamiltonian. Naive analysis has the dispersion (blue) diverging at $q = 2k_F$ but otherwise similar to the original dispersion (red), but a more careful analysis shows that the dispersion goes to zero at $2k_F$ instead (purple).

Note: A more careful treatment where we do not neglect $\hbar\omega_q$ in the denominator shows that $\omega_q \rightarrow 0$ as $q \rightarrow 2k_F$ (rather than diverging). This leads to “Bose Condensation” of phonons at $q = \pm 2k_F$ - there is always a macroscopic number of phonons at this frequency, as there is no energy cost to be at that frequency. Physically, what does this mean? This corresponds to a static distortion of the lattice (we see a spontaneous change in the geometry), known as the *Peierls instability*. Next time we look at the canonical example of this, which is polyacetylene - it undergoes a dimerization transition which can be observed as this instability.

17.3 Polyacetylene - A Classic Peierls Example

Last time we saw the effect of electron-phonon interactions in 1D, resulting in Peierls instability. The canonical example of this phenomenon is in polyacetylene, which is a carbon polymer with 1 hydrogen atom per site.

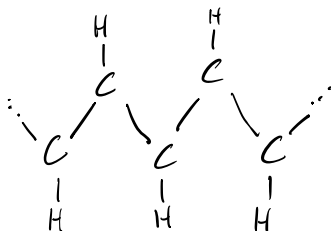


Figure 17.6: Symbolic depiction of the polyacetylene molecule.

TODO: Fix figure Explaining the chemistry: there are 4 valence electrons, 3 are hybridized (s, p_x, p_y) and form bonds and do not contribute to conduction. One is in the p_z orbital (which sticks out of the

plane) and this is the one electron that contributes to the conduction band.

So, we have the contribution of one electron per carbon atom, and therefore a simple prediction from a tight-binding model gives the dispersion: which seems to suggest that the molecule should be a conductor!

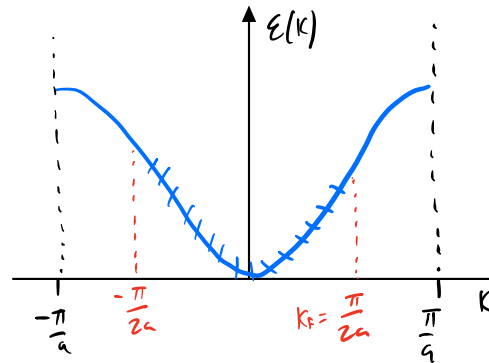


Figure 17.7: Dispersion relation for the polyacetylene molecule as predicted by the tight-binding molecule. This suggests the molecule is a conductor, but in fact it is insulating.

But when one prepares these molecules experimentally, we find that these molecules are insulating. We can understand this by studying the effects of electron-phonon interaction. Here we have $k_f = \frac{\pi}{2a}$ and so $2k_F = \frac{\pi}{a}$, i.e. at the edge of the Brillouin zone we see ω_q goes to zero.

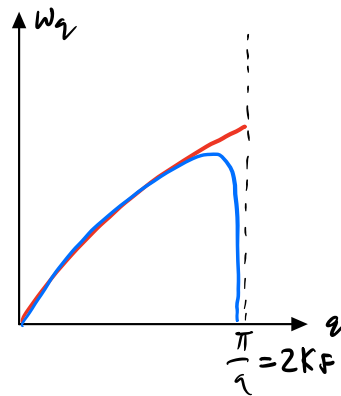


Figure 17.8: The electron-phonon interactions lead to the phonon dispersion going to zero at the edge of the Brillouin zone.

Now, since $\mu_q(\mathbf{r}) \sim e^{iq\mathbf{r}} = r^{il} a_{\mathbf{r}}^{\frac{\pi}{a}} = e^{il\pi} = \pm 1$, we will find that the Peierls instability results in a static distortion where odd sites get shifted to the left and even sites get shifted to the right - which results in a dimerized chain, with the bond lengths alternating:

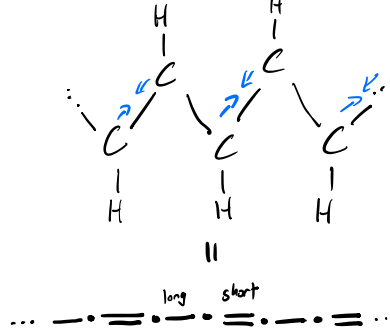


Figure 17.9: The phonon-electron interactions lead to a static deformation of the 1-D chain, leading to dimerization.

This is “dimerized polyacetylene”. Computing the band structure for this, we find:

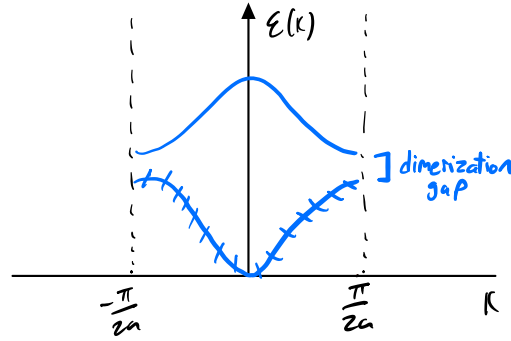


Figure 17.10: Dispersion relation for the dimerized polyacetylene chain.

where we have an insulator (the lower band is totally filled, and the upper band totally empty) and we have a “dimerization gap”.

17.4 Polarons and Mass Enhancement

Now, we ask how electron-phonon coupling affects the electron band structure. We saw an extreme example of this above in 1-D, but let us see how the (weaker) effect manifests in 3-D. This can also be understood based on the Frolich Hamiltonian which we derived last time, and follows from the second term in Eq. (17.9).

The contraction of this term yields:

$$E_0^{(2)} = - \sum_{\mathbf{k}, \mathbf{q}} |M_{\mathbf{q}}|^2 \frac{\langle \phi_0 | a_{\mathbf{q}} c_{\mathbf{k}}^{\dagger} c_{\mathbf{k}-\mathbf{q}} a_{\mathbf{q}}^{\dagger} c_{\mathbf{k}-\mathbf{q}}^{\dagger} c_{\mathbf{k}} | \phi_0 \rangle}{\epsilon_{\mathbf{k}} - \epsilon_{\mathbf{k}-\mathbf{q}} - \hbar\omega_{\mathbf{q}}} \quad (17.16)$$

Diagrammatically, the above process can be visualized as:

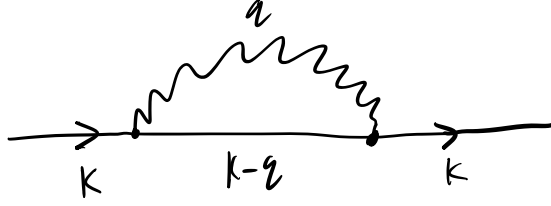


Figure 17.11: Feynman diagram for the process depicted in Eq. (17.16).

Note that Eq. (17.16) is called the electron “self-energy” due to phonons⁷. We calculate this expression exactly as we did before, and obtain:

$$E_0^{(2)} = \sum_{\mathbf{k}, \mathbf{q}} |M_{\mathbf{q}}|^2 \frac{n_{\mathbf{k}}(1 - n_{\mathbf{k}-\mathbf{q}})}{\epsilon_{\mathbf{k}} - \epsilon_{\mathbf{k}-\mathbf{q}} - \hbar\omega_{\mathbf{q}}} \quad (17.17)$$

How do we analyze this expression? The effective single electron Hamiltonian reads:

$$H_e = \sum_{\mathbf{k}} (\epsilon_{\mathbf{k}} + \delta\epsilon_{\mathbf{k}}) c_{\mathbf{k}}^\dagger c_{\mathbf{k}} = \sum_{\mathbf{k}} (\epsilon_{\mathbf{k}} + \delta\epsilon_{\mathbf{k}}) n_{\mathbf{k}} \quad (17.18)$$

Now, its clear that the correction $\delta\epsilon_{\mathbf{k}}$ can be obtained by taking the derivative of the correction w.r.t $n_{\mathbf{k}}$:

$$\begin{aligned} \delta\epsilon_{\mathbf{k}} &= \frac{\partial E_0^{(2)}}{\partial n_{\mathbf{k}}} = \sum_{\mathbf{q}} |M_{\mathbf{q}}|^2 \left(\frac{1 - n_{\mathbf{k}-\mathbf{q}}}{\epsilon_{\mathbf{k}} - \epsilon_{\mathbf{k}-\mathbf{q}} - \hbar\omega_{\mathbf{q}}} - \frac{n_{\mathbf{k}+\mathbf{q}}}{\epsilon_{\mathbf{k}+\mathbf{q}} - \epsilon_{\mathbf{k}} - \hbar\omega_{\mathbf{q}}} \right) \\ &= \sum_{\mathbf{q}} |M_{\mathbf{q}}|^2 \left(\frac{1}{\epsilon_{\mathbf{k}} - \epsilon_{\mathbf{k}-\mathbf{q}} - \hbar\omega_{\mathbf{q}}} - \frac{2\hbar\omega_{\mathbf{q}} n_{\mathbf{k}-\mathbf{q}}}{(\epsilon_{\mathbf{k}} - \epsilon_{\mathbf{k}-\mathbf{q}})^2 - (\hbar\omega_{\mathbf{q}})^2} \right) \end{aligned} \quad (17.19)$$

Note that the second term comes about because the fixed \mathbf{k} we take the derivative with respect to coincides with the \mathbf{k} in the sum.

The first term represents a correction independent of $n_{\mathbf{k}}$, and would be present even in an insulating crystal. It changes the effective mass of the electron near $\mathbf{k} = \mathbf{0}$. We say that the electron is “dressed” with a cloud of phonons and becomes the so called “polaron”. In some crystals, this enhancement can be very large - factor of 10, for example (corresponds to a large number of phonons attached to the electrons).

The second term leads to a velocity change near the Fermi level:

$$\hbar v_{\mathbf{k}} = \frac{\partial(\epsilon_{\mathbf{k}} + \delta\epsilon_{\mathbf{k}})}{\partial \mathbf{k}} = \frac{\partial \epsilon_{\mathbf{k}}}{\partial \mathbf{k}} \left(1 + \frac{d}{d\epsilon_{\mathbf{k}}} \sum_{\mathbf{q}} |M_{\mathbf{q}}|^2 \frac{2\hbar\omega_{\mathbf{q}} n_{\mathbf{k}-\mathbf{q}}}{(\epsilon_{\mathbf{k}} - \epsilon_{\mathbf{k}-\mathbf{q}})^2 - (\hbar\omega_{\mathbf{q}})^2} \right) \approx \hbar \mathbf{v}_{\mathbf{k}}^0 (1 - \alpha) \quad (17.20)$$

with $\alpha > 0$. The final approximation we make by expanding around k_F . Again we have a “mass enhancement” as we have a lower velocity. Diagrammatically, we can interpret this as the electron spends time as a slower electron + phonon before recombining, hence travels at a reduced speed. The dispersion looks like:

where we see near the Fermi surface that the velocity is reduced, attributed to the electron being slowed down by the phonon cloud.

Higher order terms have multiple phonons splitting off from the electron and recombining - but these much look much more complicated.

⁷There is an analogous process in quantum electrodynamics, with an electron emitting a photon and absorbing it back.

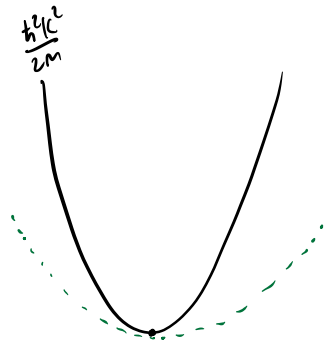


Figure 17.12: Sketch of the dispersion relation for the electrons - the phonons dress the electrons, enhancing their mass and changing the curvature of the dispersion as a result.

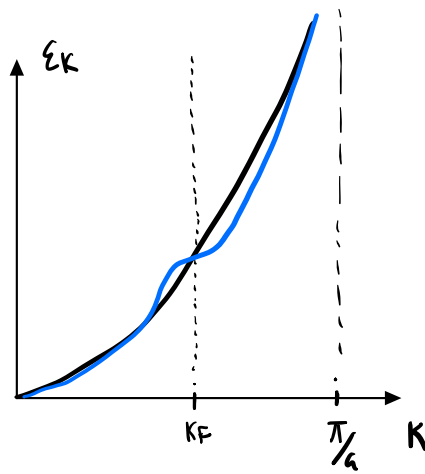


Figure 17.13: Dispersion relation for the mass-enhanced electrons. At k_F there is a small effect from the phonons, slowing them down and leading to a wiggle in the dispersion there.

18 Phonon-Mediated Attraction Between Electrons

We've looked at some effects of electron-phonon interaction, but we have saved the most spectacular for last - that is superconductivity. Phonons can create attractive interactions between electrons, which can give rise to superconductivity - an interesting historical note is that this phenomenon was known experimentally for many decades before a good theory to describe it was discovered!

Recall the electron-electron interaction (photon exchange) in quantum electrodynamics - this leads to Coloumb repulsion. Analogously, we have an exchange of phonons between electrons in solids. However, this can give rise to attractive interactions - why? Photons travel at $c \gg v_{el}$, but phonons travel at the speed of sound $c_s \ll v_{el}$; it is a matter of relative velocities.

18.1 Canonical Transformations

We sketch a derivation of this based on the Frolich Hamiltonian. Our goal is to show that some attractive interaction is possible - we will not analyze this in great detail however. We are doing this primarily to introduce a new, very useful technique - known as *canonical transformations*. We write:

$$H = H_0 + H' = \sum_{\mathbf{k}} \epsilon_{\mathbf{k}} c_{\mathbf{k}}^{\dagger} c_{\mathbf{k}} + \sum_{\mathbf{q}} \omega_{\mathbf{q}} a_{\mathbf{q}}^{\dagger} a_{\mathbf{q}} + M \sum_{\mathbf{k}, \mathbf{q}} c_{\mathbf{k}+\mathbf{q}}^{\dagger} c_{\mathbf{k}} (a_{-\mathbf{q}}^{\dagger} + a_{\mathbf{q}}). \quad (18.1)$$

From this we want to derive an effective "electron-only" Hamiltonian H_{eff} . Diagrammatically, the process we consider is:

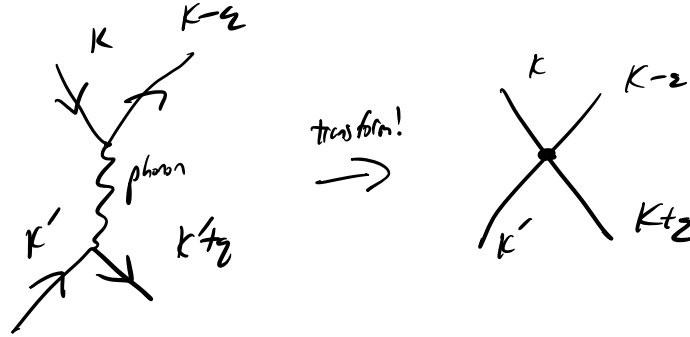


Figure 18.1: Feynman diagram for the electronic interaction via phonon exchange (left). Via a canonical transformation, we will recast this interaction into one that does not contain phonons (just electrons interacting - right).

where we want to ignore the phonons and just consider an interaction between electrons (analogous to QED where one ignores the photon field). This is where canonical transformations will come in handy. We consider:

$$H = H_0 + \lambda H'$$

where λ is some parameter. We transform it to:

$$H \rightarrow \tilde{H} = e^{-S} H e^S = H + [H, S] + \frac{1}{2} [[H, S], S] + \dots \quad (18.2)$$

usually in the literature, S is anti-hermitian (hence the *is* that you might expect are not written) but keep in mind that the above is a unitary transform. We have then expanded it using the Baker-Campbell-Hausdorf formula.

Now, we find an operator S such that \tilde{H} is independent of λ to linear order. It may not be immediately obvious why this achieves what we want, but it will.

We then have:

$$\tilde{H} = H_0 + \lambda H' + [H_0, S] + \lambda [H', S] + \dots \quad (18.3)$$

We assume that S must be linear in λ . So, $\lambda H' + [H_0, S]$ is linear in λ and we demand it vanishes ($\lambda [H', S]$ is quadratic in λ , and so on for higher order terms)

We therefore find S such that:

$$[H_0, S] = -\lambda H' \quad (18.4)$$

In some cases, if the operators are sufficiently simple one can find this directly. But there exists a way to find S for an arbitrary situation - we proceed as follows. We adopt a basis $H_0|\phi_m\rangle = \epsilon_m|\phi_m\rangle$. We then take matrix elements of both sides of the equation using these eigenstates:

$$\langle\phi_n|(H_0S - SH_0)|\phi_m\rangle = -\lambda\langle\phi_n|H'|\phi_m\rangle$$

Now, the trick is to act H_0 on the right and H_0 on the left, yielding the energy:

$$(\epsilon_n - \epsilon_m)\langle\phi_n|S|\phi_m\rangle = -\lambda\langle\phi_n|H'|\phi_m\rangle$$

So dividing by the energies, we can obtain arbitrary matrix elements of S (note: it might seem like we run into trouble with this if we look at diagonal matrix elements - however we get around this by absorbing any diagonal perturbations into H_0 , so the diagonal terms of H' are zero)! I.e. we have the solution:

$$\langle\phi_n|S|\phi_m\rangle = \lambda \frac{\langle\phi_n|H'|\phi_m\rangle}{\epsilon_m - \epsilon_n} \quad (18.5)$$

Finally, we have:

$$\tilde{H} = H_0 + \lambda[H', S] + \frac{1}{2}[[H_0, S], S] + \frac{\lambda}{2}[[H', S], S]$$

By assumption $[H_0, S] = -\lambda H'$, so $\lambda[H', S] + \frac{1}{2}[[H_0, S], S] = O(\lambda^2)$. But, $\frac{\lambda}{2}[[H', S], S] + \dots O(\lambda^3)$. And so the result after combining everything is:

$$\tilde{H} = H_0 + \frac{1}{2}\lambda[H', S] + O(\lambda^3). \quad (18.6)$$

which achieves what we set out to do - we have found S such that the transformed Hamiltonian is to leading order independent of λ ! Next class, we apply this technique to the Frolich Hamiltonian. And when we evaluate the $[H', S]$ term, we will see that it has no phonons, and corresponds to an attractive interaction.

18.2 Applying Canonical Transformations to the Frolich Hamiltonian

We consider $T = 0$, i.e. there are no thermal phonons.

As basis states, we take eigenstates of the phonon Hamiltonian $H_0^\phi = \sum_{\mathbf{q}} \omega_{\mathbf{q}} a_{\mathbf{q}}^\dagger a_{\mathbf{q}}$. Only two eigenstates here will be of consequence, namely $|0\rangle$ (no phonons) and $|\mathbf{q}\rangle$ (one phonon of momenta \mathbf{q}) - note that higher phonon states do not contribute as starting from the ground state we are not able to excite to a state with more than one phonon. We use (18.5) to calculate:

$$\langle\mathbf{q}|S|0\rangle = M \sum_{\mathbf{k}, \mathbf{p}} \frac{\langle\mathbf{q}|c_{\mathbf{k}+\mathbf{p}}^\dagger c_{\mathbf{k}}(a_{-\mathbf{p}}^\dagger + a_{\mathbf{p}})|0\rangle}{\epsilon_{\mathbf{k}} - \epsilon_{\mathbf{k}+\mathbf{p}} - \omega_{\mathbf{q}}} = M \sum_{\mathbf{k}} \frac{c_{\mathbf{k}-\mathbf{q}}^\dagger c_{\mathbf{k}}}{\epsilon_{\mathbf{k}} - \epsilon_{\mathbf{k}-\mathbf{q}} - \omega_{\mathbf{q}}} \quad (18.7)$$

where in the last equality we use that the matrix element vanishes unless $\mathbf{p} = -\mathbf{q}$. Analogously, we can obtain:

$$\langle 0|S|\mathbf{q}\rangle = M^* \sum_{\mathbf{k}} \frac{c_{\mathbf{k}+\mathbf{q}}^\dagger c_{\mathbf{k}}}{\epsilon_{\mathbf{k}} - \epsilon_{\mathbf{k}+\mathbf{q}} + \omega_{\mathbf{q}}} \quad (18.8)$$

So we can therefore find the leading $\sim \lambda^2$ correction in \tilde{H} :

$$\begin{aligned} \langle 0|[H', S]|0\rangle &= \langle 0|H'S|0\rangle - \langle 0|SH'|0\rangle \\ &= \sum_{\mathbf{q}} \left(\langle 0|H'|\mathbf{q}\rangle \langle \mathbf{q}|S|0\rangle - \langle 0|S|\mathbf{q}\rangle \langle \mathbf{q}|H'|0\rangle \right) \\ &= |M|^2 \sum_{\mathbf{q}, \mathbf{k}, \mathbf{k}'} c_{\mathbf{k}'+\mathbf{q}}^\dagger c_{\mathbf{k}'} c_{\mathbf{k}-\mathbf{q}}^\dagger c_{\mathbf{k}} \left(\frac{1}{\epsilon_{\mathbf{k}} - \epsilon_{\mathbf{k}-\mathbf{q}} - \omega_{\mathbf{q}}} - \frac{1}{\epsilon_{\mathbf{k}'} - \epsilon_{\mathbf{k}'+\mathbf{q}} + \omega_{\mathbf{q}}} \right) \end{aligned} \quad (18.9)$$

where in the second equality we have inserted the resolution of the identity, and in the third equality we evaluate the matrix elements. We have “integrated out the phonons” in this way. Now, in the last term let us rename $\mathbf{k} \leftrightarrow \mathbf{k}'$, reverse $\mathbf{q} \rightarrow -\mathbf{q}$ and so:

$$\langle 0|[H', S]|0\rangle = |M|^2 \sum_{\mathbf{q}, \mathbf{k}, \mathbf{k}'} \frac{\omega_{\mathbf{q}}}{(\epsilon_{\mathbf{k}} - \epsilon_{\mathbf{k}-\mathbf{q}})^2 - \omega_{\mathbf{q}}^2} c_{\mathbf{k}'+\mathbf{q}}^\dagger c_{\mathbf{k}'} c_{\mathbf{k}-\mathbf{q}}^\dagger c_{\mathbf{k}} \quad (18.10)$$

which (looking at the first term) can be negative! So, we conclude that the phonon-mediated interaction is attractive whenever this occurs, i.e.:

$$|\epsilon_{\mathbf{k}} - \epsilon_{\mathbf{k}-\mathbf{q}}| < \omega_{\mathbf{q}} \implies \frac{\hbar^2}{2m} |2\mathbf{k} \cdot \mathbf{q} - \mathbf{q}^2| < c_s |\mathbf{q}| \quad (18.11)$$

where in the last implication we put in the dispersion relation for free electrons $\epsilon_{\mathbf{k}} = \frac{\hbar^2 \mathbf{k}^2}{2m}$, and the dispersion relation for acoustic phonons (linear). This occurs whenever $|\mathbf{k}| \approx k_F$ and $\mathbf{q} \approx 2\mathbf{k}$.

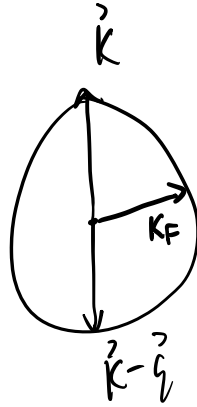


Figure 18.2: Electrons at different poles of the Fermi sphere feel an attractive interaction, due to phonons.

I.e. whenever two electrons are on opposite sides of the Fermi surface, their interaction will be attractive. And this we will see leads to superconductivity!

18.3 A physical picture for attractive electrons

We consider an ionic lattice. The electrons exert an attractive force on the ions in the lattice. But the electron moves very quickly. Another electron then would see the increased positive charge density due to

the ions - the second electron then feels attraction to this increased positive charge density. The important thing in the language that we use is that this is a retarded attraction - which is to say that the second electron is attracted to the previous position of the first electron. This is good, because when the second electron feels the attraction, the first electron is already faraway - so the Coloumb interaction does not matter in this consideration. So, even though this phonon mediated attraction is fairly weak compared to Coloumb, it can often win out - and this is all because of the relative velocities of the players in the interaction.

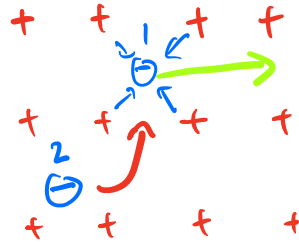


Figure 18.3: Cartoon sketch of the electron-electron attraction mechanism. Electron 1 attracts the positively charged ions in the lattice, and then quickly moves elsewhere in the lattice, creating an area of high positive charge density. Electron 2 is then attracted to the area that electron 1 once was.

19 Superconductivity

19.1 Phenomenology and Historical Background

We will not provide a large background of the phenomenology of superconductivity - the first 2-3 pages of any textbook section about it will explain it in depth. We just give a couple points and then explain the Cooper problem which is the microscopic foundations for the effect.

- In 1911, Kamerlingh-Onnes discovered that the low-temperature ($< 4\text{K}$) resistance of mercury goes to zero. Actually, his assistant discovered this and called Onnes at a conference, who doubted the result initially. The big contribution of Onnes was actually how to liquify Helium, as this is what was necessary to cool things down to low T ⁸.
- Meissner and Dehshfield found that $\mathbf{B} = \mathbf{0}$ (expels magnetic fields) for superconductors - this in some sense is more fundamental of a property, because a perfect metal (with no impurities) would have no resistance at $T = 0$, but such a perfect metal would not have this expelling of magnetic fields property.
- Flux quantization. This says that if we take a piece of superconductor, make a hole in it, and measure the magnetic flux through the hole, we would find that $\Phi = n\Phi_0$ where Φ_0 is the superconducting flux quantum $\Phi_0 = \frac{hc}{2e}$.

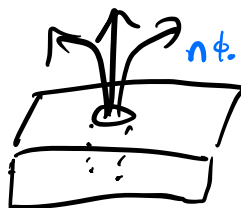


Figure 19.1: If we open a hole in a superconductor, the magnetic flux is quantized with $\Phi = n\Phi_0$.

How do we understand this effect? Note the theory of superconductivity was developed in 1956, so it took almost 50 years from the discovery of the phenomenon before we got an explanatory theory.

Note in 2011 (the 100 year anniversary of SC) a book was published with many famous papers (both theory and experiment), entertainingly many theory papers which were absolutely wrong.

19.2 Cooper Instability

Cooper instability⁹ is where for arbitrarily weak attractive interactions, electrons in a Fermi sea are unstable to formation of pairs, or electron-electron bound states. These are what are known as Cooper pairs. Superconductivity can be thought of as the condensation of e-e pairs, which act as bosons. Unlike Helium atoms which are charge neutral, Cooper pairs carry a charge and therefore can carry current. The first clue to this is the $2e$ (rather than e) appearing in the superconducting flux quantum.

To discuss this problem, we consider $T = 0$ and consider the filled Fermi sphere.

⁸Onnes had a streak of cooling many things with liquid Helium... like frogs, though nothing interesting happens to them at 4K - they just die. Fortunately mercury is more interesting

⁹An interesting sociological point - when Cooper came up with this, he was actually a graduate student. He got a Nobel prize a few years later, but unfortunately didn't come up with anything else notable after this (though the other two awardees did).

Cooper considered two electrons that are slightly above the Fermi sphere, and considered the Hamiltonian of these two electrons:

$$H = \frac{\mathbf{p}_1^2}{2m} + \frac{\mathbf{p}_2^2}{2m} + V(\mathbf{r}_1 - \mathbf{r}_2) \quad (19.1)$$

we assume that the total momentum of the two is zero, so $\mathbf{k}_1 + \mathbf{k}_2 = \mathbf{0}$. Further we assume that the electrons are in the spin singlet state (total spin is 0).

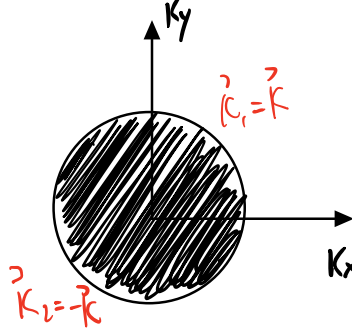


Figure 19.2: Setup of Cooper instability - We consider two electrons with total momentum zero, in a spin singlet state, sitting slightly above the Fermi sea.

Under these assumptions, one is able to solve this problem. We make an ansatz:

$$\psi(\mathbf{r}_1, \mathbf{r}_2) = \sum_{|\mathbf{k}| > k_F} g_{\mathbf{k}} e^{i\mathbf{k} \cdot \mathbf{r}_1} e^{-i\mathbf{k} \cdot \mathbf{r}_2} \frac{|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle}{\sqrt{2}} \quad (19.2)$$

Fermi statistics implies that $g_{\mathbf{k}} = g_{-\mathbf{k}}$. Because the spin part is already anti-symmetric, under interchange the g part must be symmetric so under the interchange we have a net negative sign. From this, we can conclude that the wavefunction is only a function of the difference of the two coordinates:

$$\psi(\mathbf{r}_1 - \mathbf{r}_2) = \sum_{|\mathbf{k}| > k_F} g_{\mathbf{k}} \cos(\mathbf{k} \cdot (\mathbf{r}_1 - \mathbf{r}_2)) \frac{|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle}{\sqrt{2}}$$

Now, we look for $g_{\mathbf{k}}$ by solving the Schrodinger equation with this Hamiltonian:

$$H\psi_0 = E\psi_0$$

What we get out is:

$$(E - 2\epsilon_{\mathbf{k}})g_{\mathbf{k}} = \sum_{|\mathbf{k}'| > k_F} V_{\mathbf{k}\mathbf{k}'} g_{\mathbf{k}'} \quad (19.3)$$

where $\epsilon_{\mathbf{k}} = \frac{\hbar^2 \mathbf{k}^2}{2m}$, and:

$$V_{\mathbf{k}\mathbf{k}'} = \frac{1}{\Omega} \int d^3r V(\mathbf{r}) e^{i(\mathbf{k}' - \mathbf{k}) \cdot \mathbf{r}} \quad (19.4)$$

now we make a Cooper ansatz for this interaction. If we take a generic interaction (such as that we found from the el-ph interaction) we cannot solve this analytically. However, with the ansatz:

$$V_{\mathbf{k}\mathbf{k}'} = \begin{cases} -V & |\epsilon_{\mathbf{k}} - \epsilon_F|, |\epsilon_{\mathbf{k}'} - \epsilon_F| < \hbar\omega_c \\ 0 & \text{otherwise} \end{cases} \quad (19.5)$$

With this ansatz:

$$(E - 2\epsilon_{\mathbf{k}}) = -V \sum'_{\mathbf{k}'} g_{\mathbf{k}'}$$

where the $'$ on the summation denotes the restriction of being sufficiently close to the Fermi energy. From this we obtain:

$$g_{\mathbf{k}} = V \frac{\sum'_{\mathbf{k}'} g_{\mathbf{k}'}}{2\epsilon_{\mathbf{k}} - E}$$

and so:

$$\frac{1}{V} = \sum'_{\mathbf{k}} \frac{1}{2\epsilon_{\mathbf{k}} - E} \quad (19.6)$$

we now solve for the energy eigenvalue E - we obtain this by going to the continuum approximation and replacing sum with integral:

$$\frac{1}{V} = \int_{\epsilon_F}^{\epsilon_F + \hbar\omega_c} d\epsilon N(\epsilon) \frac{1}{2\epsilon - E} = N(\epsilon_F) \int_{\epsilon_F}^{\epsilon_F + \hbar\omega_c} d\epsilon \frac{1}{2\epsilon - E} = \frac{1}{2} N(\epsilon_F) \ln \left(1 + \frac{2\hbar\omega_c}{2\epsilon_F - E} \right) \quad (19.7)$$

And we obtain the algebraic equation for E :

$$1 + \frac{2\hbar\omega_c}{2\epsilon_F - E} = e^{2/VN(\epsilon_F)} \quad (19.8)$$

If we make the weak coupling assumption of $VN(\epsilon_F) \ll 1$, then the argument of the exponential becomes very large. Therefore we obtain the condition that $\epsilon_F \approx E$. In this limit we can neglect the $+1$ on the LHS and solve for E to be:

$$E \approx 2\epsilon_F - 2\hbar\omega_c e^{-\frac{2}{VN(\epsilon_F)}} \quad (19.9)$$

Normally, we would expect that $E > 2\epsilon_F$ as the electrons should be above the Fermi surface (the Fermi sphere being full)! But with the attractive interaction, we find a state that has smaller energy than that - the attractive interaction lowers the energy. $E < 2\epsilon_F$, meaning we have a bound state. Another point is that it exists at arbitrarily small $V > 0$. A final point - the expression is not analytic in V . This implies that Cooper instability is a non-perturbative phenomena. This is an important insight, and may explain why it took so long to formulate this theory - perturbative approaches to probe superconductivity are destined to fail, as $e^{-\frac{1}{V}}$ has no Taylor expansion about $V = 0$ - there is an essential singularity at $V = 0$.

Now the important thing to ask - what happens when we try to describe all the electrons? It is quite contrived to have a problem of a full Fermi sea and two extra electrons. However, it will nevertheless lead to the whole BCS theory of superconductivity we will discuss next lecture.

20 The BCS Ground State

We discuss the 1956 result of Bardeen, Cooper and Schrieffer, who answered the question - what happens to many (as opposed to one, as we studied last lecture) electrons in the presence of attractive interaction?

Today, we go through a historical account of the story. Schrieffer wrote down the wavefunction via intuition, and then it was shown that this had minimal energy. This is a bit unwieldy and complex, but still interesting. On Wednesday, we discuss the more modern approach to SCs, which also will be more applicable to finite temperature (vs. today's lecture's result, which is only applicable at $T = 0$).

20.1 Ground State Wavefunction

We write:

$$|\psi_G\rangle = \prod_{\mathbf{k}} \left(u_{\mathbf{k}} + v_{\mathbf{k}} c_{\mathbf{k}\uparrow}^\dagger c_{-\mathbf{k}\downarrow}^\dagger \right) |0\rangle \quad (20.1)$$

and then solve the variational problem where $\mu_{\mathbf{k}}, v_{\mathbf{k}}$ are the variational parameters. We will find that the minimum energy variational state will have lower energy than if we just put all the electrons in the lowest energy states.

For normalization, we have the condition

$$|u_{\mathbf{k}}|^2 + |v_{\mathbf{k}}|^2 = 1. \quad (20.2)$$

$|v_{\mathbf{k}}|^2$ is the probability that the pair $(k \uparrow, -k \downarrow)$ is occupied.

An immediate possible objection - $|\psi_G\rangle$ is a superposition of states with different electron numbers. There are famous physicists who to this day do not accept the wavefunction. To clarify this point, we can write:

$$|\psi_G\rangle = \sum_N \lambda_N |\psi_N\rangle \quad (20.3)$$

where $|\psi_N\rangle$ is a state with a definite number N of electrons. Bardeen would answer to this objection - this is true, but λ_N is very sharply peaked around $N = \bar{N}$. This is a simple calculation:

$$\bar{N} = \langle \psi_G | \hat{N} | \psi_G \rangle = \sum_{\mathbf{k}} 2|v_{\mathbf{k}}|^2 \sim \Omega \quad (20.4)$$

the last expression follows from the fact that $|v_{\mathbf{k}}|$ are constants of order one, so taking $\sum_{\mathbf{k}} \rightarrow \Omega \int d^3k$, we see that the sum scales with volume. Similarly, we can calculate:

$$\langle (\hat{N} - \bar{N})^2 \rangle = 4 \sum_{\mathbf{k}} |u_{\mathbf{k}}|^2 |v_{\mathbf{k}}|^2 \sim \Omega \quad (20.5)$$

Therefore:

$$\frac{\delta N}{\bar{N}} = \frac{\sqrt{(\hat{N} - \bar{N})^2}}{\bar{N}} \sim \frac{\sqrt{\Omega}}{\bar{N}} \sim \frac{1}{\sqrt{\Omega}} \sim \frac{1}{\sqrt{\bar{N}}} \quad (20.6)$$

Now, since $\bar{N} \sim 10^{22}$, $\delta N \sim \sqrt{\bar{N}} \sim 10^{11}$ so the distribution $P(\lambda_N)$ is sharply peaked:

So, the BCS wavefunction describes a state with the number of electrons extremely sharply peaked around \bar{N} . Note we could calculate the distribution in principle by expanding out the product and then evaluating the products of $u_{\mathbf{k}}$ s and $v_{\mathbf{k}}$ s.

20.2 Calculation of $u_{\mathbf{k}}, v_{\mathbf{k}}$

We consider the Pairing Hamiltonian, or the BCS Hamiltonian as it is known today. Writing it down, we have:

$$\hat{H} - \mu \hat{N} = \sum_{\mathbf{k}\sigma} (\epsilon_{\mathbf{k}} - \mu) c_{\mathbf{k}\sigma}^\dagger c_{\mathbf{k}\sigma} + \sum_{\mathbf{k}\mathbf{l}} V_{\mathbf{k}\mathbf{l}} c_{\mathbf{k}\uparrow}^\dagger c_{-\mathbf{k}\downarrow}^\dagger c_{-\mathbf{l}\downarrow} c_{\mathbf{l}\uparrow} \quad (20.7)$$

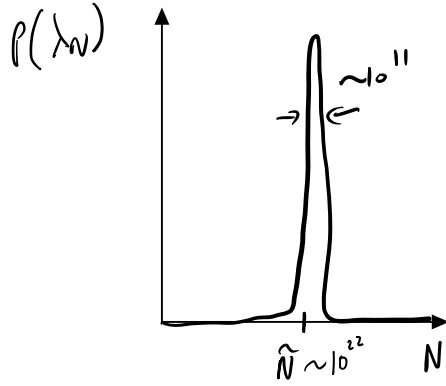


Figure 20.1: Plot of $P(\Delta_N)$ as a function of the electron number N . The distribution is sharply peaked around \bar{N} .

The first term is the kinetic energy, the second is the interaction. Where does this come from? We recall the general form of the two body interaction is:

$$\sum_{\mathbf{k}\mathbf{p}\mathbf{q}\sigma\sigma'} V_{\mathbf{q}} c_{\mathbf{k}-\mathbf{q}\sigma}^\dagger c_{\mathbf{p}+\mathbf{q}\sigma'}^\dagger c_{\mathbf{p}\sigma'} c_{\mathbf{k}\sigma} \quad (20.8)$$

And then motivated by our discussion last class of two electrons on opposite sides of the Fermi sphere having attractive interaction, we enforce $\mathbf{p} = -\mathbf{k}$ and $\mathbf{k} + \mathbf{q} = \mathbf{l}$. This yields the interaction term in the BCS Hamiltonian.

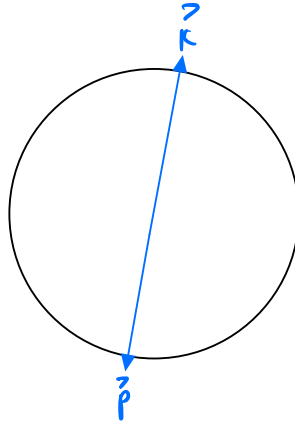


Figure 20.2: We restrict the sum of interactions to be electrons at opposing points above the Fermi sphere.

We now calculate the ground state energy. This is another simple, if lengthy calculation:

$$E_S = \langle \psi_G | \hat{H} - \mu \hat{N} | \psi_G \rangle = 2 \sum_{\mathbf{k}} \xi_{\mathbf{k}} |v_{\mathbf{k}}|^2 + \sum_{\mathbf{k}, \mathbf{l}} V_{\mathbf{k}\mathbf{l}} u_{\mathbf{k}} v_{\mathbf{k}}^* u_{\mathbf{l}}^* v_{\mathbf{l}} \quad (20.9)$$

where $\xi_{\mathbf{k}} = \epsilon_{\mathbf{k}} - \mu$. We minimize this with respect to $u_{\mathbf{k}}, v_{\mathbf{k}}$. Let us assume that $u_{\mathbf{k}}, v_{\mathbf{k}} \in \mathbb{R}$ (relaxing this assumption does not result in immediately significant changes). An immediate simplification is that

$u_{\mathbf{k}}, v_{\mathbf{k}}$ are not independent via the normalization condition in Eq. (20.2). So, let us parameterize $u_{\mathbf{k}} = \sin(\theta_{\mathbf{k}}), \cos(\theta_{\mathbf{k}})$. We then obtain:

$$\begin{aligned} E_s &= 2 \sum_{\mathbf{k}} \xi_{\mathbf{k}} \cos^2(\theta_{\mathbf{k}}) + \sum_{\mathbf{k}, \mathbf{l}} V_{\mathbf{k}\mathbf{l}} \cos(\theta_{\mathbf{k}}) \sin(\theta_{\mathbf{k}}) \cos(\theta_{\mathbf{l}}) \sin(\theta_{\mathbf{l}}) \\ &= \sum_{\mathbf{k}} \xi_{\mathbf{k}} (1 + \cos 2\theta_{\mathbf{k}}) + \frac{1}{4} \sum_{\mathbf{k}, \mathbf{l}} V_{\mathbf{k}\mathbf{l}} \sin 2\theta_{\mathbf{k}} \sin 2\theta_{\mathbf{l}} \end{aligned} \quad (20.10)$$

Now we minimize (for a given \mathbf{k})

$$\begin{aligned} 0 &= \frac{\partial E_s}{\partial \theta_{\mathbf{k}}} = -2\xi_{\mathbf{k}} \sin 2\theta_{\mathbf{k}} + \sum_{\mathbf{l}} V_{\mathbf{k}\mathbf{l}} \cos 2\theta_{\mathbf{k}} \sin 2\theta_{\mathbf{l}} \\ \Rightarrow \tan 2\theta_{\mathbf{k}} &= \frac{\sum_{\mathbf{l}} V_{\mathbf{k}\mathbf{l}} \sin 2\theta_{\mathbf{l}}}{2\xi_{\mathbf{k}}} \end{aligned} \quad (20.11)$$

Now, let us define:

$$\begin{aligned} \Delta_{\mathbf{k}} &= -\sum_{\mathbf{l}} V_{\mathbf{k}\mathbf{l}} u_{\mathbf{l}} v_{\mathbf{l}} = -\frac{1}{2} \sum_{\mathbf{l}} V_{\mathbf{k}\mathbf{l}} \sin 2\theta_{\mathbf{l}} \\ E_{\mathbf{k}} &= \sqrt{\xi_{\mathbf{k}}^2 + \Delta_{\mathbf{k}}^2} \end{aligned} \quad (20.12)$$

So with this, we have:

$$\tan 2\theta_{\mathbf{k}} = -\frac{\Delta_{\mathbf{k}}}{\xi_{\mathbf{k}}} \quad (20.13)$$

Now, writing:

$$\begin{aligned} 2u_{\mathbf{k}}v_{\mathbf{k}} &= \sin 2\theta_{\mathbf{k}} = \frac{1}{\sqrt{1 + \tan^2 2\theta_{\mathbf{k}}}} = \frac{\Delta_{\mathbf{k}}}{E_{\mathbf{k}}} \\ v_{\mathbf{k}}^2 - u_{\mathbf{k}}^2 &= \cos 2\theta_{\mathbf{k}} = \frac{-1}{\sqrt{2 + \tan^2 2\theta_{\mathbf{k}}}} = -\frac{\eta_{\mathbf{k}}}{E_{\mathbf{k}}} \end{aligned} \quad (20.14)$$

We can now solve:

$$\begin{aligned} v_{\mathbf{k}}^2 &= \frac{1}{2} \left(1 - \frac{\xi_{\mathbf{k}}}{E_{\mathbf{k}}} \right) \\ u_{\mathbf{k}}^2 &= \frac{1}{2} \left(1 + \frac{\xi_{\mathbf{k}}}{E_{\mathbf{k}}} \right) \end{aligned} \quad (20.15)$$

The last step is to calculate $\Delta_{\mathbf{k}}$ (from its definition) so we can make sense of $E_{\mathbf{k}}$. We find:

$$\Delta_{\mathbf{k}} = -\frac{1}{2} \sum_{\mathbf{l}} \frac{\Delta_{\mathbf{l}}}{E_{\mathbf{l}}} V_{\mathbf{k}\mathbf{l}} = -\frac{1}{2} \sum_{\mathbf{l}} \frac{\Delta_{\mathbf{l}}}{\sqrt{\xi_{\mathbf{l}}^2 + \Delta_{\mathbf{l}}^2}} V_{\mathbf{k}\mathbf{l}} \quad (20.16)$$

this is now a self-consistent equation for Δ . To solve, we use Cooper's Ansatz:

$$V_{\mathbf{k}\mathbf{l}} = \begin{cases} -V & \text{if } |\xi_{\mathbf{k}}|, |\xi_{\mathbf{l}}| \leq \hbar\omega_c \\ 0 & \text{otherwise} \end{cases} \quad (20.17)$$

So then:

$$\Delta_{\mathbf{k}} = \begin{cases} \Delta & \text{when } |\xi_{\mathbf{k}}| \leq \hbar\omega_c \\ 0 & \text{otherwise} \end{cases} \quad (20.18)$$

And so we obtain the equation:

$$\Delta = \frac{1}{2}V \sum_1' \frac{\Delta}{\sqrt{\Delta^2 + \xi_1^2}} \quad (20.19)$$

the trivial solution is $\Delta = 0$. If $\Delta \neq 0$, then we can divide out by Δ on both sides, and obtain:

$$\frac{2}{V} = \sum_1' \frac{1}{\sqrt{\Delta^2 + \xi_1^2}} = \int_{-\hbar\omega_c}^{\hbar\omega_c} d\xi \frac{N(\xi)}{\sqrt{\xi^2 + \Delta^2}} \approx N(0) \int_{-\hbar\omega_c}^{\hbar\omega_c} \frac{d\xi}{\sqrt{\xi^2 + \Delta^2}} = 2N(0) \operatorname{arcsinh}\left(\frac{\hbar\omega_c}{\Delta}\right) \quad (20.20)$$

Inverting this to find Δ , we find:

$$\Delta = \frac{\hbar\omega_c}{\sinh(1/VN(0))} \quad (20.21)$$

this starts to look like the Cooper problem. At weak coupling of $VN(0) \ll 1$, the argument of \sinh is large and we can approximate it as an exponential:

$$\Delta \rightarrow 2\hbar\omega_c e^{-1/VN(0)} \quad (20.22)$$

A few remarks on the physics behind this. This is a non-perturbative result, like Cooper instability. If you expand in powers of V , the expansion will fail as there is an essential singularity at $V = 0$. We also note that a non-trivial solution exists for any V attractive. So, the state exists self-consistently!

Let us now plot v_k^2 given this result.

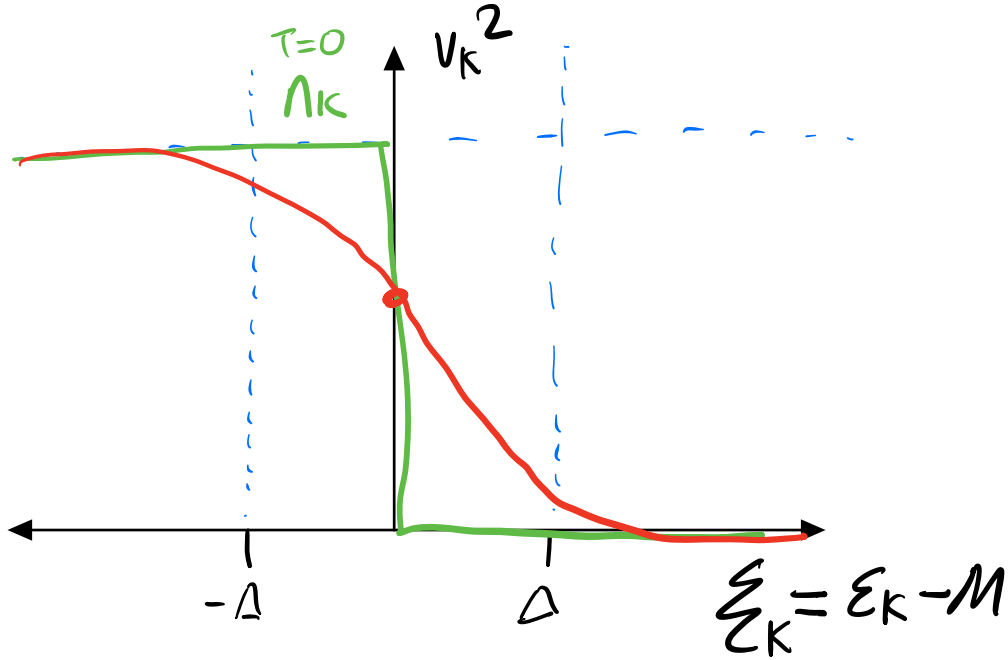


Figure 20.3: Plot of v_k^2 vs ξ_k , with the $T = 0$ Fermi function plotted for comparison.

When $\Delta \neq 0$, we have a “paired state”, different from the filled Fermi sphere.

20.3 BCS Condensation Energy

Finally, we show that this paired state has lower energy than the Fermi sphere:

$$\begin{aligned} E_s &= \langle \psi_G | \hat{H} - \mu \hat{N} | \psi_G \rangle \\ &= \sum_{\mathbf{k}} \left(\tilde{\xi}_{\mathbf{k}} - \frac{\tilde{\xi}_{\mathbf{k}}^2}{E_{\mathbf{k}}} \right) - \frac{\Delta^2}{V} \end{aligned} \quad (20.23)$$

compare this with the Fermi sphere energy:

$$E_n = \langle \psi_G | \hat{H} - \mu \hat{N} | \psi_G \rangle_{\Delta=0} = 2 \sum_{|\mathbf{k}| < k_F} \tilde{\xi}_{\mathbf{k}} \quad (20.24)$$

and so:

$$\boxed{\delta E = E_s - E_n = -\frac{1}{2} N(0) \Delta^2} \quad (20.25)$$

note the minus sign - the condensation energy is negative $\delta E \leq 0$, anytime Δ is nonzero. This implies that $|\Psi_G\rangle$ is a stable state. It has energy lower than the Fermi sphere.

If you read the original BCS paper, they go onto discuss the physical implications of this. But unfortunately because the calculation was at zero temperature, the calculations are lengthy and roundabout. Instead, next time we show the modern treatment of this problem via a standard mean-field theory, and do a finite temperature calculation. From those results we will be able to readily evaluate the properties of a superconducting state.