# Solid State Physics III

## F. Mila

Master course
Year 2021-2022

# Contents

# Chapter 1

# Introduction

## 1.1 Scope

Solid state physics deals with the properties of condensed matter in its solid state, that is in the phase in which the relative position of nuclei is fixed on average. This definition excludes liquids but includes amorphous systems such as glasses, or partially ordered systems such as liquid crystals. By properties we mean the way solids behave in the presence of external stimuli such as a mechanical action, a heat pulse, or an electromagnetic field in the broad sense (electric field, magnetic field, electromagnetic wave). This subject is extremely vast, and it is absolutely necessary to make a selection. The goal of this course is twofold:

1. Develop modern tools and concepts that allow one to deal with problems of current interest in fundamental and applied research in solid state physics (second quantization, degenerate perturbation theory, canonical transformations,. . . )

2. Use them to study a number of problems chosen for their intrinsic interest (magnetism, Quantum Hall Effect, superconductivity, etc.).

This is not the only way to proceed. Indeed, once they are well understood, essentially all phenomena discovered in solids can be described by elementary methods with appropriate ingredients. It is thus perfectly possible to teach advanced solid state physics without resorting to these sophisticated techniques. However, these tools are of current (not to say systematic) use in contemporary scientific literature. Besides, they are essential to convince oneself of the validity of the simple descriptions of the phenomena. This justifies the choice of developing them in detail so that they can be used in other contexts.

## 1.2 Independent electrons in a periodic potential

### 1.2.1 Bloch theorem

A large number of mysteries have been elucidated in solid physics by considering the electron gas as a gas of non-interacting fermions moving in a periodic potential. From that point of view, the fundamental theorem is *Bloch theorem.* Let us look for the solution of Schrödinger equation

$$\left(-\frac{\hbar^2}{2m}\nabla^2 + U(\mathbf{r})\right)\psi(\mathbf{r}) = \varepsilon\psi(\mathbf{r})$$

where $U(\mathbf{r})$ is a periodic potential ($U(\mathbf{r}+\mathbf{R}) = U(\mathbf{r})$ if $\mathbf{R} = \sum_j n_j \mathbf{a}_j$ with $\mathbf{a}_j$ 3 linearly independent vectors and $n_j$ three integers).

First formulation of Bloch theorem

The solutions are of the form

$$\psi_{\mathbf{k}}(\mathbf{r}) = e^{i\mathbf{k}.\mathbf{r}} u_{\mathbf{k}}(\mathbf{r})$$

where $u_{\mathbf{k}}(\mathbf{r})$ is a function that has the same periodicity as $U(\mathbf{r})$.

Second formulation of Bloch theorem

The solutions satisty

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

#### Equivalence of the two formulations

Let us first prove that the two formulations of Bloch theorem are indeed equivalent :

1. Let us suppose that $\psi_{\mathbf{k}}(\mathbf{r}) = e^{i\mathbf{k}.\mathbf{r}} u_{\mathbf{k}}(\mathbf{r})$, where $u_{\mathbf{k}}$ has the periodicity of the lattice. We have :

$$\begin{aligned} \psi_{\mathbf{k}}(\mathbf{r}+\mathbf{R}) &= e^{i\mathbf{k}.\mathbf{R}} e^{i\mathbf{k}.\mathbf{r}} u_{\mathbf{k}}(\mathbf{r}+\mathbf{R}) \\ &= e^{i\mathbf{k}.\mathbf{R}} e^{i\mathbf{k}.\mathbf{r}} u_{\mathbf{k}}(\mathbf{r}) = e^{i\mathbf{k}.\mathbf{R}}\psi_{\mathbf{k}}(\mathbf{r}) \end{aligned}$$

2. Conversely, let us suppose that :

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

   and let us define the function $u_{\mathbf{k}}(\mathbf{r})$ by :

$$u_{\mathbf{k}}(\mathbf{r}) = e^{-i\mathbf{k}.\mathbf{r}}\psi_{\mathbf{k}}(\mathbf{r})$$

   We have :

$$\begin{aligned} u_{\mathbf{k}}(\mathbf{r}+\mathbf{R}) &= e^{-i\mathbf{k}.\mathbf{r}} \underbrace{e^{-i\mathbf{k}.\mathbf{R}}\psi_{\mathbf{k}}(\mathbf{r}+\mathbf{R})}_{\psi_{\mathbf{k}}(\mathbf{r})} \\ &= u_{\mathbf{k}}(\mathbf{r}) \end{aligned}$$

   $\Rightarrow u_{\mathbf{k}}$ is indeed periodic.

### Proof of Bloch theorem

Let us now prove the second formulation of Bloch theorem. For this, let us first introduce the translation operator $T_{\mathbf{R}}$ defined by

$$T_{\mathbf{R}}\psi(\mathbf{r}) = \psi(\mathbf{r}+\mathbf{R}) \tag{1.1}$$

Now, if $\mathbf{R} = \sum_j n_j \mathbf{a}_j$, this operator commutes with $H$. Indeed,

$$T_a \frac{\partial^2 \psi}{\partial x^2}(x) = \frac{\partial^2 \psi}{\partial x^2}(x+a)$$

and

$$\begin{aligned}\frac{\partial^2}{\partial x^2} T_a \psi(x) = \frac{\partial^2}{\partial x^2}\left[\psi(x+a)\right] &= \frac{\partial}{\partial x}\left[\frac{\partial \psi}{\partial x}(x+a)\underbrace{\frac{d}{dx}(x+a)}_{=1}\right] \\ &= \frac{\partial^2 \psi}{\partial x^2}(x+a)\underbrace{\frac{d}{dx}(x+a)}_{=1}\end{aligned}$$

$$\Rightarrow T_{\mathbf{a}}\left(-\frac{\hbar^2}{2m}\nabla^2\psi(\mathbf{r})\right) = -\frac{\hbar^2}{2m}\nabla^2 T_{\mathbf{a}}\psi(\mathbf{r})$$

This is true for all $\mathbf{a}$, hence in particular for the lattice translations. So, $T_{\mathbf{R}}$ commutes with $-\frac{\hbar^2}{2m}\nabla^2$.

Besides,

$$\begin{aligned}T_{\mathbf{R}} U(\mathbf{r})\psi(\mathbf{r}) &= U(\mathbf{r}+\mathbf{R})\psi(\mathbf{r}+\mathbf{R}) \\ &= U(\mathbf{r})\psi(\mathbf{r}+\mathbf{R}) = U(\mathbf{r}) T_{\mathbf{R}}\psi(\mathbf{r})\end{aligned}$$

So $T_{\mathbf{R}}$ commutes with $H$.

Now, if some operators commute, one can diagonalize them in a common basis. So let us look for the eigenstates of $T_{\mathbf{a}_j}$. They must satisfy :

$$T_{\mathbf{a}_j}\psi(\mathbf{r}) = \lambda_j \psi(\mathbf{r})$$

With Born-Von Karman boundary conditions :

$$\psi(\mathbf{r}+N_j\mathbf{a}_j) = \psi(\mathbf{r})$$

for a system with $N = N_1 N_2 N_3$ unit cells, one must have :

$$T_{\mathbf{a}_j}^{N_j}\psi(\mathbf{r}) \equiv \psi(\mathbf{r}+N_j\mathbf{a}_j) = \psi(\mathbf{r})$$

$$\text{But} \qquad T_{\mathbf{a}_j}^{N_j}\psi(\mathbf{r}) = \lambda_j^{N_j}\psi(\mathbf{r}) \Rightarrow \lambda_j^{N_j} = 1 \Rightarrow \lambda_j = e^{i\frac{2\pi}{N_j}m_j}$$

$$\text{Finally,} \qquad T_{\mathbf{R}}\psi(\mathbf{r}) = e^{\sum_j i\frac{2\pi}{N_j}m_j n_j}\psi(\mathbf{r})$$

The expression $\frac{2i\pi}{N_j}m_j n_j$ can be written as $i\mathbf{k}.\mathbf{R}$ with $\mathbf{k} = \sum_j \frac{m_j}{N_j}\mathbf{a}_j^*$ where the vectors $\mathbf{a}_j^*$ are defined by :

$$\mathbf{a}_i.\mathbf{a}_j^* = 2\pi\delta_{ij}$$

They define the *reciprocal lattice.*

This leads us to the second formulation of *Bloch theorem* :

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

### Brillouin zones

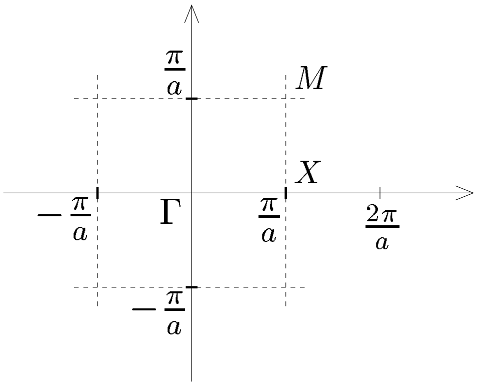
Let $\mathbf{G} = \sum_j m_j\mathbf{a}_j^*$ be a vector of the reciprocal lattice.

$$\mathbf{G}.\mathbf{R} = \sum_j 2\pi m_j n_j \Rightarrow e^{i\mathbf{G}.\mathbf{R}} = 1 \Rightarrow e^{i(\mathbf{k}+\mathbf{G}).\mathbf{R}} = e^{i\mathbf{k}.\mathbf{R}}$$

This implies that two wave-vectors $\mathbf{k}$ lead to the same solution if they only differ by a vector of the reciprocal lattice. It is thus sufficient to consider a unit cell of the reciprocal lattice, for instance the *1st Brillouin zone.*

Example : square lattice of parameter $a$.

$$\mathbf{a}_1 = a\hat{\mathbf{x}},\ \mathbf{a}_2 = a\hat{\mathbf{y}} \Rightarrow \mathbf{a}_1^* = \frac{2\pi}{a}\hat{\mathbf{x}},\ \mathbf{a}_2^* = \frac{2\pi}{a}\hat{\mathbf{y}}$$



## 1.2.2 Band theory

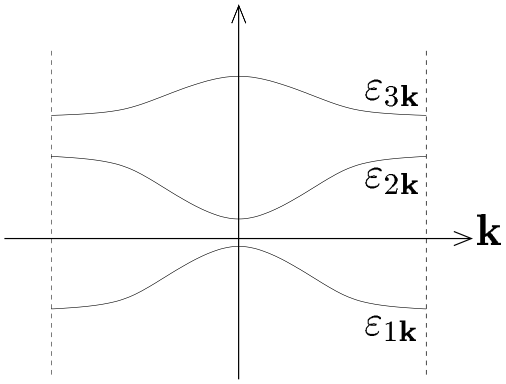
If one looks for the solutions of the Schrödinger equation of the form $\psi_{\mathbf{k}}(\mathbf{r}) = e^{i\mathbf{k}.\mathbf{r}}u_{\mathbf{k}}(\mathbf{r})$, the function $u_{\mathbf{k}}(\mathbf{r})$ must satisfy another differential equation that we can simply determine :

$$\nabla\left(e^{i\mathbf{k}.\mathbf{r}}u_{\mathbf{k}}(\mathbf{r})\right) = i\mathbf{k}e^{i\mathbf{k}.\mathbf{r}}u_{\mathbf{k}}(\mathbf{r}) + e^{i\mathbf{k}.\mathbf{r}}\nabla u_{\mathbf{k}}(\mathbf{r})$$

$$\begin{aligned}
\nabla^2\left(e^{i\mathbf{k}.\mathbf{r}}u_{\mathbf{k}}(\mathbf{r})\right) &= -\mathbf{k}^2 e^{i\mathbf{k}.\mathbf{r}}u_{\mathbf{k}}(\mathbf{r}) + i\mathbf{k}e^{i\mathbf{k}.\mathbf{r}}.\nabla u_{\mathbf{k}}(\mathbf{r}) + i\mathbf{k}e^{i\mathbf{k}.\mathbf{r}}.\nabla u_{\mathbf{k}}(\mathbf{r}) + e^{i\mathbf{k}.\mathbf{r}}\nabla^2 u_{\mathbf{k}}(\mathbf{r}) \\
&= e^{i\mathbf{k}.\mathbf{r}}\left(-\mathbf{k}^2 + 2i\mathbf{k}.\nabla + \nabla^2\right)u_{\mathbf{k}}(\mathbf{r})
\end{aligned}$$

$$\left[-\frac{\hbar^2}{2m}\left(-\mathbf{k}^2 + 2i\mathbf{k}.\nabla + \nabla^2\right) + U(\mathbf{r})\right]u_{\mathbf{k}}(\mathbf{r}) = \varepsilon_{\mathbf{k}}u_{\mathbf{k}}(\mathbf{r})$$

In general this equation has an infinite number of solutions that can classified according to an additional quantum number $n$ called the band index. The functions $\varepsilon_{n\mathbf{k}}$ constitute the *band structure.*



In the band theory, one assumes that the electrons do not interact with each other. The eigenstates with $N$ electrons can be obtained by putting one electron in each quantum state. At that stage, one must take into account the spin $\frac{1}{2}$ of the electrons. Each quantum state is characterized by its energy $\varepsilon_{n\mathbf{k}\sigma}$, $\sigma =\uparrow$ or $\downarrow$. In the absence of a magnetic field, and if the system has an inversion center, $\varepsilon_{n\mathbf{k}\uparrow} = \varepsilon_{n\mathbf{k}\downarrow}$.

In general, there are energies that do not correspond to $\varepsilon_{n\mathbf{k}\sigma}$ for any value of $(n\mathbf{k}\sigma)$. These energies build intervals called *gaps.*

Note : Strictly speaking, gaps only exist for infinite solids. When a surface is present, other solutions exist whose energy can fall into a gap. These solutions look like :

$$\psi(\mathbf{r}+\mathbf{R}) = z\psi(\mathbf{r})$$

where $z$ is a complex number of modulus not equal to 1. Such solutions are *a priori* possible since the translation operator can of course have eigenvalues of modulus $\neq 1$. Such solutions are not acceptable in the bulk because their norm diverges exponentially with the size of the system, but they are acceptable when there is a surface if the function *decreases* away from the surface. Such states are called *surface states.* They have been observed and play an important role in the electronic properties of surfaces (catalysis, ...). The Born-Von Karman solutions, which impose $|z| = 1$, do not allow one to describe surface states.

### Tight-binding approach

The band structure can also be seen as the result of hybridizing atomic levels. This is the tight-binding approach. The starting point is a set of atomic orbitals of energy $\epsilon_0$ located at the sites $i$ of a lattice. Let us denote by $|i\rangle$ the associated ket. In the absence of hybridization between the states, which will be true in the limit where the sites are very far apart, the energy level $\epsilon_0$ is $N$-fold degenerate, where $N$ is the number of sites. If the sites are brought closer to each other, a coupling will develop. It will take the form of a perturbation $V$ that lifts the degeneracy. Assuming that the coupling decays fast with distance, one can assume that only nearest-neighbor sites have a significant overlap:

$$\begin{cases} \langle i|V|j\rangle = -t \text{ if } i \text{ and } j \text{ are nearest neighbors} \\ \langle i|V|j\rangle = 0 \text{ otherwise} \end{cases}$$

where $-t$ is the hopping amplitude (the minus sign is conventional). Then, according to degenerate perturbation theory, the effective Hamiltonian is given by

$$\mathcal{H} = -t \sum_{\langle i,j\rangle} (|i\rangle\langle j| + |j\rangle\langle i|) + \epsilon_0 \sum_i |i\rangle\langle i|$$

where the sum over $\langle i, j\rangle$ runs over pairs of neighoring sites.

Let us assume for simplicity that the orbitals sit at the sites of a cubic lattice of parameter $a$ in dimension $D$, and let us define the Fourier transform of the states by

$$|\vec{k}> = \frac{1}{\sqrt{N}} \sum_i e^{-i\vec{k}\cdot\vec{R}_i} |i> .$$

Applying the Hamiltonian to this state leads to:

$$\begin{aligned} \mathcal{H}|\vec{k}> &= \epsilon_0|\vec{k}> + \frac{1}{\sqrt{N}} \sum_i e^{-i\vec{k}\cdot\vec{R}_i} (-t) \sum_{\vec{\tau}} |i+\vec{\tau}> \\ &= \epsilon_0|\vec{k}> + \frac{1}{\sqrt{N}} (-t) \sum_{\vec{\tau}} e^{i\vec{k}\cdot\vec{\tau}} \sum_i e^{-i\vec{k}\cdot(\vec{R}_i+\vec{\tau})} |i+\vec{\tau}> \\ &= \epsilon_0|\vec{k}> - t \sum_{\vec{\tau}} e^{i\vec{k}\cdot\vec{\tau}} |\vec{k}> \\ &= E_{\vec{k}}|\vec{k}> \end{aligned} \tag{1.2}$$

where the vectors $\vec{\tau}$ are the vectors connecting a site to its nearest neighbors, and where the dispersion $E_{\vec{k}}$ is given by

$$E_{\vec{k}} = \epsilon_0 - 2t \sum_\alpha \cos(k_\alpha a)$$

where $\alpha$ runs over the directions of space. This band is centered at $\epsilon_0$, and the band width is given by $W = 4Dt$. In the ground state, the particle goes to the bottom of the band and gains and energy equal to half the bandwidth $W/2$.

### Physical implications of band theory

The ground state of a bulk system is obtained by filling the energy levels starting from the lowest one. There are two possibilities :

1. All bands are either completely filled, or completely empty. The first excited state is separated from the ground state by the energy $\Delta$ that separates the last occupied state from the first empty one. As a consequence, it is impossible to move electrons by applying an arbitrarily small voltage $\Rightarrow$ the system is an *insulator*. Since there are $2N$ states per band, this is only possible if the number of electrons per unit cell is *even*.
2. One or several bands are partially filled. The energy of the last occupied state is called the *Fermi energy* and is denoted by $\varepsilon_F$. In that case, the separation between the last occupied state and the first empty one goes to 0 when $N \to \infty$. The system can react to infinitesimal excitations. Such systems are called *metals* (or conductors).

This classification is probably the greatest success of band theory. Before this theory, the existence of metals and insulators was a great mystery. Many consequences of this theory have been verified. For instance, this theory predicts the existence of a particular set of wave vectors called the *Fermi surface* defined by the equation :

$$\varepsilon_{n\mathbf{k}\sigma} = \varepsilon_F$$

The topology of this surface has truly remarkable consequences. In particular, it can lead to a conduction by holes, hence to a positive Hall coefficient, a result impossible to understand in the context of Drude or Sommerfeld theories of metals. It also has direct consequences for the de Haas-Van Alphen effect (oscillations of the magnetization of metals with the magnetic field) or for the Chubnikov-de Haas effect (oscillations of the magnetoresistance with the field) - see chapter on Electrons in a Magnetic Field.

Besides, the density of states at the Fermi level can be very different from its value in the Sommerfeld theory. Indeed, it is given by :

$$g_n(\varepsilon_F) = \int_{S_n(\varepsilon_F)} \frac{dS}{4\pi^3} \frac{1}{|\nabla \varepsilon_n(\mathbf{k})|}$$

where the integral runs over the Fermi surface. The density of states can have singularities if the gradient vanishes, or even divergences in dimension 2 or 1. This density of states directly enters quantities such as the specific heat :

$$c_v = \frac{\pi^2}{3} k_B^2 T g(\varepsilon_F)$$

or the Pauli suscpetibility :

$$\chi = \mu_B^2 g(\varepsilon_F)$$

Finally, the semiclassical theory of transport based on band theory allows one to explain many properties.

Nevertheless it is not possible to stop at the description in terms of independent electrons for two very different reasons :

1. Problem of principle : The Coulomb interaction between electrons is not negligible. It is not even small :

$$V(\mathbf{q}) = \frac{4\pi e^2}{\mathbf{q}^2} \text{ diverge pour } q \to 0$$

   Why does this approximation of independent electrons, that consists in replacing this interaction by an average effective potential, work so well in a large number of cases? To answer this question, and to understand the limits of validity of this theory, it is necessary to work with many-body wave functions. The only convenient representation goes under the name of second quantization (see Appendix 6).

2. The failures of band theory : in spite of considerable effort, including by Bloch himself, it has turned out to be impossible to explain superconductivity with band theory. One knows by now that superconductivity corresponds to a state described by a wave function that is radically different from the eigenfunctions of band theory (see Chapter 5 on superconductivity). This failure is not a complete one however since superconductivity is accurately described as an instability from a state well described by band theory.

   But there is a far more radical failure : certain systems with an *odd* number of electrons per unit cell are insulators, something rigorously impossible in the framework of band theory. This problem is the subject of the following section.

## 1.3 Mott insulators

In 1937, only a few years after the development of band theory (Bloch 1929, Peierls 1929, Wilson 1931), de Boer et Verwey noticed that NiO, an insulator with a gap of about $4\,eV$, should be a metal according to band theory. In the discussion that followed the presentation of the results, Peierls said that this could be a consequence of electron-electron interactions. It is only several years later that Mott managed to formulate this explanation in a precise way. Since then, many counter-examples have been found. The most famous is $\mathrm{La_2CuO_4}$ : doped with Sr ($\mathrm{La_{2-x}Sr_xCuO_4}$) this compound becomes a superconductor at $40\,K$, a discovery for which Bednorz and Müller have been awarded the 1987 Nobel prize.

The band structure of this compound has been calculated by Mattheiss in 1987 (see Fig. 1.1). Since there is an odd number of electrons per unit cell, there is at least one partially filled band. This is indeed the case : one band is half filled.

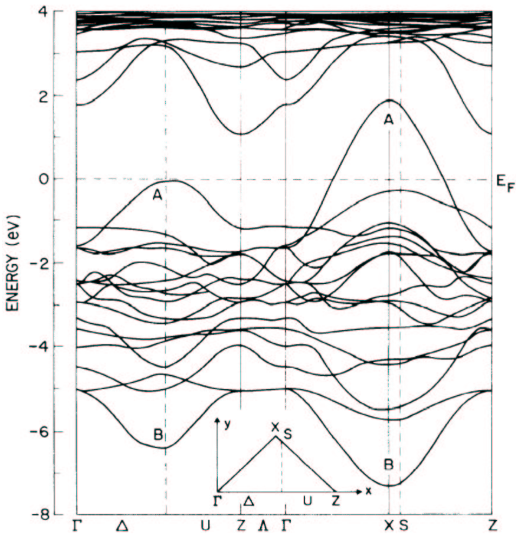


Figure 1.1: LAPW energy bands for $La_2CuO_4$ along symmetry lines in the bct Brillouin zone. L.F. Mattheiss, Phys. Rev. Lett. **58**, 1028 (1987).

To understand this band structure, the simplest thing to do is to see the solid as a collection of atoms. Each atom has orbitals that are more or less filled.

La : [Xe] $6s^2$ $5d^1$
Cu : [Ar] $3d^{10}$ $4s^1$
O : $1s^2$ $2s^2$ $2p^4$

The relative position in energy of the external orbitals of the atoms controls their filling in the solid. It turns out that the 2p orbitals of oxygen are lower than the external orbitals of Cu and La. As a consequence, the electronic configurations in the solid are as follows :

$La^{3+}$ : [Xe]
$Cu^{2+}$ : [Ar] $3d^9$
$O^{2-}$ : $1s^2$ $2s^2$ $2p^6$

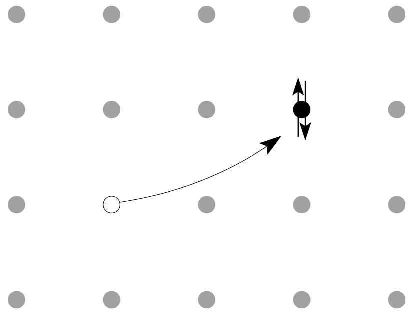
In total, $La_2^{3+}Cu^{2+}O_4^{2-}$ is neutral, as it should!

Besides, if we think in terms of the tight-binding approach, the hopping integrals between certain 3d orbitals of copper and certain 2p orbitals of oxygen are quite large ($\simeq 1\,eV$), and the separation between the atomic levels is not too large ($\simeq 2\,eV$). This leads to a significant hybridization between these orbitals, and the Fermi level lies in an energy range where there are 17 overlapping bands : 5 originating from the 3d levels of copper, and 12 (4×3) from the 2p levels of the four oxygen atoms of the unit cell.

This approach is always possible, and it describes very well the band structure of metals. Why does it fail to describe properly the ground state of several compounds like $La_2CuO_4$ ? The central argument put forward by Mott consists in studying the electronic excitations of a solid composed of hydrogen atoms as a function of the distance between the atoms.

To keep things as simple as possible, let us try to build an imaginary solid forming a square lattice with hydrogen atoms. The unit cell contains one hydrogen atom, and there is one electron per unit cell. According to band theory, the system must thus be a metal. However, if we denote by $a$ the distance between hydrogen atoms, the system cannot be metallic when $a$ gets very large. Indeed, let us try to induce a current into this system. To achieve this, we must take an electron and put it far away. For the system to be metallic, it must be possible to do this only providing the system with an infinitesimal energy. So let us try to evaluate the energy needed for this process.

If $a$ is very large, the system has essentially one electron per site to start with. When sending an electron far away, one creates one hole and one doubly occupied site :



The potential energy of that state is roughly given by :

$$E(H^-) + E(H^+) - 2E(H)$$

This energy is strictly positive because of the Coulomb repulsion between the electrons. It is usually denoted by $U$.

However, the system can gain kinetic energy: the hole can delocalize, as well as the doubly occupied site. Let us assume for simplicity that at all sites occupied by a single electron, the spin is the same, say $\uparrow$. Then indeed the motion of the hole can be described by a tight binding model of hopping amplitude $-t$ since letting the hole move to a neighbouring site is equivalent to letting the electron that occupied that site hop on the empty site. If we denote by $W$ the associated bandwidth, the hole will occupy the bottom of the band and gain $W/2$.

The story is very similar for the doubly occupied site. If we assume again for simplicity that all sites are occupied by a single electron with $\uparrow$ spin, the doubly occupied site corresponds to a site with an extra electron with spin $\downarrow$, and this electron can hop from one site to its neighbours with amplitude $-t$, leading again to an effective motion of the doubly occupied site with hopping amplitude $-t$. Still denoting by $W$ the associate bandwidth, the doubly occupied site will occupy the bottom of the band and gain $W/2$.

In reality, the spins do not need to be the same at all sites, and $W/2$ is an upper bound of the energy gain for both the hole and the doubly occupied sites, but the order of magnitude can be shown to be correct.

Now, the bandwidth has been estimated in the context of the tight-binding method. In 2D, the dispersion relation is given by:

$$E(\mathbf{k}) = -2t(\cos k_x + \cos k_y)$$

and the bandwidth is given by $W = 8t$.

Finally, since both the hole and the doubly occupied site gain approximately $W/2$, the energy of a configuration with one hole and one doubly occupied site is given by:

$$\Delta_c = U - W$$

In this expression, $U$ is essentially independent of the lattice parameter $a$ because it is an atomic quantity. By contrast, $W$ depends strongly on $a$. When $a$ is large as compared to the Bohr radius $a_0$, the overlap between wavefunctions that enters the calculation of the transfer integral $t$ involves the exponential tails of the wavefunctions, and $t \propto e^{-\frac{a}{a_0}}$ is exponentially small. In that limit, $\Delta_c \simeq U$ is a positive and large number $\Rightarrow$ the system is an *insulator*. Such insulators are called *Mott insulators*. $\Delta_c$ is then called the *charge gap*.

When $a$ decreases, $t$ increases, and there has to be a value of the order of $a_0$ for which $U = W$. Below this value, the calculation of $\Delta_c$ is not valid anymore. The gap cannot be negative! Mott's argument thus predicts the existence of a metal-insulator transition when $a$ increases.

In principle, this transition can be observed by applying pressure to a Mott insulator to reduce $a$ and increase $t$. Historically a transition of this type has been observed for the first time in $V_2O_3$ in the early seventies. There are other ingredients that make the transition in that compound more complicated, but the validity of Mott's argument is now generally accepted.

There is a fundamental difference between a Mott insulator and a band insulator. Indeed, if it is necessary to pay an energy $\Delta_c$ to make a charge excitation in a Mott insulator, it is not necessary to pay this energy to make a spin excitation: one can flip the spin of an electron without creating a doubly occupied site. The description of the magnetism of Mott insulators is treated in the next chapter. To discuss this physics in a precise way in a general framework that includes simultaneously charge and spin excitations, one must master tools that allow one to describe interacting electrons. The best adapted formalism, the second quantization, is presented in the last Chapter (Appendix).

# Chapter 2

# Magnetism of Mott insulators

## 2.1 Introduction

As we have seen in the previous chapter, Mott insulators are radically different from band insulators because they possess low-energy magnetic excitations. Indeed the insulating state correponds to a configuration where each atom carries an odd number of electrons, say 1 for simplicity, and the spin of this electron can be ↑ or ↓. In the purely atomic limit where atoms are infinitely far apart, the excitation spectrum is very simple. The ground state corresponds to all configurations with one electron per site. It is thus $2^N$ fold degenerate, and its energy is equal to $N\epsilon_0$, where $\epsilon_0$, the energy of an electron on a site, can be taken as the reference of energy ($\epsilon_0 = 0$).

The first excited state corresponds to a configuration where one site is empty and one site is doubly occupied. Its degeneracy is $N(N-1)2^{N-2}$, and its energy $U$. And so on $\cdots$

$$\begin{array}{rcll} E & = & 2U & \quad \cdots \\ E & = & U & \quad N(N-1)2^{N-2} \\ E & = & 0 & \quad 2^N. \end{array}$$

If one brings the atoms closer to each other, the first effect will be to lift the degeneracy in the subspace $E = 0$. The minimal model to describe this physics is called the *Hubbard model*. It is defined by the Hamiltonian:

$$\mathcal{H} = -t \sum_{<i,j>,\sigma} \left( c^+_{i,\sigma} c_{j,\sigma} + h.c. \right) + U \sum_i n_{i,\uparrow} n_{i,\downarrow}.$$

The first term is a hopping term betwenn nearest neighbours. This is the term that dominates the charge transfer processes when one brings atoms together from infinity. The second term measures the number of doubly occupied sites.

To determine the **effective Hamiltonian** that describes the lifting of degeneracy in the ground state manifold, it is thus necessary to carry on degenerate perturbation theory with respect to the kinetic term

$$\mathcal{H}_0 = -t \sum_{<i,j>,\sigma} \left( c^+_{i,\sigma} c_{j,\sigma} + \text{h.c.} \right) .$$

## 2.2 Elementary two-site calculation

To understand the effect of the kinetic hopping term, it is useful to start with a two-site calculation. Let us keep track of the two sites with indices 1 and 2. There are four possible quantum states for a particle defined by the creation operators $c^+_{1\uparrow}$, $c^+_{1\downarrow}$, $c^+_{2\uparrow}$ et $c^+_{2\downarrow}$. Let us build the two-particle states in the Fock space $|n_{1\uparrow}, n_{1\downarrow}, n_{2\uparrow}, n_{2\downarrow}\rangle$. There are six of them, given by:

$$\begin{aligned}
|1\rangle &\equiv |1,0,1,0\rangle = c^+_{1\uparrow} c^+_{2\uparrow} |0\rangle \\
|2\rangle &\equiv |0,1,0,1\rangle = c^+_{1\downarrow} c^+_{2\downarrow} |0\rangle \\
|3\rangle &\equiv |1,0,0,1\rangle = c^+_{1\uparrow} c^+_{2\downarrow} |0\rangle \\
|4\rangle &\equiv |0,1,1,0\rangle = c^+_{1\downarrow} c^+_{2\uparrow} |0\rangle \\
|5\rangle &\equiv |1,1,0,0\rangle = c^+_{1\uparrow} c^+_{1\downarrow} |0\rangle \\
|6\rangle &\equiv |0,0,1,1\rangle = c^+_{2\uparrow} c^+_{2\downarrow} |0\rangle
\end{aligned} \tag{2.1}$$

Applying the full two-site Hamiltonian to these states leads to:

$$\begin{aligned}
H|1\rangle &= 0 \\
H|2\rangle &= 0 \\
H|3\rangle &= -t(|5\rangle + |6\rangle) \\
H|4\rangle &= t(|5\rangle + |6\rangle) \\
H|5\rangle &= U|5\rangle - t(|3\rangle - |4\rangle) \\
H|6\rangle &= U|6\rangle - t(|3\rangle - |4\rangle)
\end{aligned} \tag{2.2}$$

These results are obvious given the form of the Hamiltonian, up to one subtlety, the sign, due to the fact that fermionic operators anticommute. Indeed, let us consider for example the effect of the term $c^+_{2\downarrow}c_{1\downarrow}$ on state $|4\rangle$. We get:

$$\begin{aligned}
c^+_{2\downarrow}c_{1\downarrow}|4\rangle &= c^+_{2\downarrow}c_{1\downarrow}c^+_{1\downarrow}c^+_{2\uparrow}|0\rangle \\
&= c^+_{2\downarrow}(1-c^+_{1\downarrow}c_{1\downarrow})c^+_{2\uparrow}|0\rangle \\
&= c^+_{2\downarrow}c^+_{2\uparrow}|0\rangle \\
&= -c^+_{2\uparrow}c^+_{2\downarrow}|0\rangle \\
&= -|6\rangle
\end{aligned}$$

Grouping states, we get:

$$\begin{aligned}
H\left(\frac{|3\rangle+|4\rangle}{\sqrt{2}}\right) &= 0 \\
H\left(\frac{|3\rangle-|4\rangle}{\sqrt{2}}\right) &= -2t\left(\frac{|5\rangle+|6\rangle}{\sqrt{2}}\right) \\
H\left(\frac{|5\rangle+|6\rangle}{\sqrt{2}}\right) &= -2t\left(\frac{|3\rangle-|4\rangle}{\sqrt{2}}\right)+U\left(\frac{|5\rangle+|6\rangle}{\sqrt{2}}\right) \\
H\left(\frac{|5\rangle-|6\rangle}{\sqrt{2}}\right) &= U\left(\frac{|5\rangle-|6\rangle}{\sqrt{2}}\right)
\end{aligned} \tag{2.3}$$

The states $\left(\frac{|3\rangle-|4\rangle}{\sqrt{2}}\right)$ and $\left(\frac{|5\rangle+|6\rangle}{\sqrt{2}}\right)$ lead to the eigenvalue equation

$$E(E-U)-4t^2=0$$

whose solutions are

$$E_\pm=\frac{U\pm\sqrt{U^2+16t^2}}{2}.$$

Up to second order in $t$, these energies are given by

$$E_-=-\frac{4t^2}{U},\quad E_+=U+\frac{4t^2}{U}$$

The ground state is thus a non-degenerate state of energy $-\frac{4t^2}{U}$, the first excited state has energy 0, and it is three-fold degenerate. The other excited states are at energies of order $U$, more precisely $U$ and $U+\frac{4t^2}{U}$. There are indeed four states whose energy would be equal to 0 if the hopping integral $t$ was equal to 0. They correspond to the four states with one an only one electron per site. The hopping term lifts the degeneracy and organizes the spectrum into a singlet of energy $-\frac{4t^2}{U}$ and a triplet of energy 0. This is exactly what one would get for two spins 1/2 coupled by the Hamiltonian

$$H=\frac{4t^2}{U}(\vec{S}_1\cdot\vec{S}_2-\frac{1}{4}) \tag{2.4}$$

We are now going to show that this is indeed the form of the effective Hamiltonian to second order in $t$.

## 2.3 Effective Hamiltonian

There are two methods to derive in a rigorous way the form of the effective Hamiltonian to second order: canonical transformations, and degenerate perturbation theory. We will use a canonical transformation later in this course. In the present case, one can directly apply degenerate perturbation theory.

To first order in $\mathcal{H}_0$, degenerate perturbation theory predicts that the effective Hamiltonian is given by

$$< m|\mathcal{H}_{eff}|n >=< m|\mathcal{H}_0|n >$$

where $|m >$ et $|n >$ are kets of the ground state manifold. But since we are at half-filling, both $|m >$ and $|n >$ must have one and only one electron per site. Since $\mathcal{H}_0$ lets one electron hop from one site to another, it necessarily creates an empty site and a doubly occupied site. Its matrix elements in the ground state manifold thus vanish identically, and one must push degenerate perturbation theory to next order.

To second order, the matrix elements are given by

$$< m|\mathcal{H}_{eff}|n >= \sum_{|k>}{}' \frac{< m|\mathcal{H}_0|k >< k|\mathcal{H}_0|n >}{E_0 - E_k}$$

where the sum $\sum'_{|k>}$ is restricted to states which do not belong to the ground state manifold.

Since $\mathcal{H}_0$ creates a doubly occupied site, $E_k = U$, which implies

$$< m|\mathcal{H}_{eff}|n >= -\frac{1}{U}\sum_{|k>}{}' < m|\mathcal{H}_0|k >< k|\mathcal{H}_0|n >$$

The operator $\sum'_{|k>} |k >< k|$ is the identity in the subspace of states with a doubly occupied site. It is convenient to replace it by another representation of the identity in this subspace, namely $\sum_j n_{j,\sigma}n_{j,-\sigma}$, leading to

$$\mathcal{H}_{eff} = \mathcal{H}_0 \frac{\sum_j n_{j,\sigma}n_{j,-\sigma}}{-U}\mathcal{H}_0$$

Let us consider the j-th term of this sum. Only terms of the type

$$\sum_{i(j),\sigma} c^+_{j,\sigma}c_{i,\sigma}$$

will contribute, where $i(j)$ means that one sums over the neighbors of $j$. The corresponding term of $\mathcal{H}_{eff}$ can be written

$$-\frac{t^2}{U}\sum_{i(j)\sigma'\sigma''} c^+_{i,\sigma'}c_{j,\sigma'}n_{j,\sigma}n_{j,-\sigma}c^+_{j,\sigma''}c_{i,\sigma''}$$

There are four cases to be considered:

$\boxed{\sigma' = \sigma'' = \sigma}$

$$c^+_{i,\sigma}c_{j,\sigma}n_{j,\sigma}n_{j,-\sigma}c^+_{j,\sigma}c_{i,\sigma} = n_{i,\sigma}n_{j,-\sigma}c_{j,\sigma}n_{j,\sigma}c^+_{j,\sigma}.$$

Now

$$c_{j,\sigma}n_{j,\sigma}c^+_{j,\sigma} = (1 - n_{j,\sigma})$$

since

$$c\,n = cc^+c = (1 - c^+c)c = c \Longrightarrow cc^+ = 1 - n.$$

The term of the Hamiltonian then takes the form:

$$c^+_{i,\sigma}c_{j,\sigma}n_{j,\sigma}n_{j,-\sigma}c^+_{i,\sigma}c_{j,\sigma} = n_{i,\sigma}n_{j,-\sigma}(1 - n_{j,\sigma})$$

But in the subspace in which $\mathcal{H}_{eff}$ operates, all sites are occupied by one and only one electron. Thus we have:

$$n_{j,\sigma} + n_{j,-\sigma} = 1 \Longrightarrow 1 - n_{j,\sigma} = n_{j,-\sigma}.$$

Besides, since $n_{j,-\sigma}n_{j,-\sigma} = n_{j,-\sigma}$, we get the following contribution:

$$c^+_{i,\sigma}c_{j,\sigma}n_{j,\sigma}n_{j,-\sigma}c^+_{i,\sigma}c_{j,\sigma} = n_{i,\sigma}n_{j,-\sigma}.$$

$\boxed{\sigma' = \sigma'' = -\sigma}$

In the same way,

$$c^+_{i,-\sigma}c_{j,-\sigma}n_{j,\sigma}n_{j,-\sigma}c^+_{i,-\sigma}c_{j,-\sigma} = n_{i,-\sigma}n_{j,\sigma}.$$

$\boxed{\sigma'' = -\sigma' = \sigma}$

$$\begin{aligned} c^+_{i,-\sigma}c_{j,-\sigma}n_{j,\sigma}n_{j,-\sigma}c^+_{j,\sigma}c_{i,\sigma} &= c^+_{i,-\sigma}c_{i,\sigma}\underbrace{c_{j,-\sigma}n_{j,-\sigma}}_{c_{j,-\sigma}}\underbrace{n_{j,\sigma}c^+_{j,\sigma}}_{c^+_{j,\sigma}} \\ &= -c^+_{i,-\sigma}c_{i,\sigma}c^+_{j,\sigma}c_{j,-\sigma}. \end{aligned}$$

$\boxed{\sigma'' = -\sigma' = -\sigma}$

Similarly,

$$c^+_{i,\sigma}c_{j,\sigma}n_{j,\sigma}n_{j,-\sigma}c^+_{j,-\sigma}c_{i,-\sigma} = -c^+_{i,-\sigma}c_{i,\sigma}c^+_{j,\sigma}c_{j,-\sigma}.$$

Collecting all terms, we get:

$$\mathcal{H}_{eff} = \frac{2t^2}{U}\sum_{<i,j>,\sigma}\left(-n_{i,\sigma}n_{j,-\sigma} + c^+_{i,-\sigma}c_{i,\sigma}c^+_{j,\sigma}c_{j,-\sigma}\right)$$

where the prefactor 2 comes from the fact that, for each pair $< i, j >$, one term comes from the term $n_{j,\sigma}n_{j,-\sigma}$ of the sum $\sum_j n_{j,\sigma}n_{j,-\sigma}$, and one term from $n_{i,\sigma}n_{i,-\sigma}$.

Now, in second quantization, the spin operators can be written (see exercises):

$$\begin{cases} S^z_i = \frac{1}{2}\left(n_{i,\uparrow} - n_{i,\downarrow}\right) \\ S^+_i = c^+_{i,\uparrow}c_{i,\downarrow} \\ S^-_i = c^+_{i,\downarrow}c_{i,\uparrow} \end{cases}$$

and the density operators $n_i = n_{i,\uparrow} + n_{i,\downarrow}$. Using the following identities:

$$S^z_iS^z_j - \frac{1}{4}n_in_j = -\frac{1}{2}\left(n_{i,\uparrow}n_{j,\downarrow} + n_{i,\downarrow}n_{j,\uparrow}\right)$$
$$S^+_iS^-_j + S^-_iS^+_j = c^+_{i,\uparrow}c_{i,\downarrow}c^+_{j,\downarrow}c_{j,\uparrow} + c^+_{i,\downarrow}c_{i,\uparrow}c^+_{j,\uparrow}c_{j,\downarrow}$$

on can rewrite the effective Hamiltonian in terms of spin operators as:

$$\mathcal{H}_{eff} = \frac{4t^2}{U}\sum_{<i,j>}\left(S^z_iS^z_j - \frac{1}{4}n_in_j + \frac{1}{2}\left(S^+_iS^-_j + S^-_iS^+_j\right)\right).$$

But $\vec{S}_i \cdot \vec{S}_j = S^z_iS^z_j + \frac{1}{2}\left(S^+_iS^-_j + S^-_iS^+_j\right)$ and $n_i = n_j = 1$ in the ground state manifold. Finally, the effective Hamiltonian takes the form:

$$\mathcal{H}_{eff} = \frac{4t^2}{U}\sum_{<i,j>}\left(\vec{S}_i \cdot \vec{S}_j - \frac{1}{4}\right),$$

Up to a constant, this is the *Heisenberg model* defined by:

$$\mathcal{H}_{Heis} = J\sum_{<i,j>}\vec{S}_i \cdot \vec{S}_j,$$

with the coupling constant $J = \frac{4t^2}{U}$. This constant is *positive*. So this Hamiltonian tends to favour configurations where nearest neighbour spins are antiparallel. The physical reason is very simple: if two electrons on neighbouring sites have parallel spins, none of them can hop on the other site because of Pauli principle. By contrast, if they spin are antiparallel, each electron can hop on the neighbouring site, and these processes allow the system to lower its energy.

## 2.4 Heisenberg model - general considerations

The Hamiltonian that we have obtained in the previous section is a particular case of the Heisenberg model which can be defined by the more general Hamiltonian:

$$H = \sum_{(i,j),\alpha=x,y,z} J_{ij}^{\alpha}\, S_i^{\alpha} S_j^{\alpha}.$$

The symbol $(i,j)$ means that one sums over pairs. The parameters that determine the properties of this model are:

1. The sign of the couplings.

   If the couplings are positive, as in the previous section, they tend to align the spins in opposite directions. One then speaks of **antiferromagnetic** couplings. If they are negative, they tend to align the spins in the same direction. In that case, one speaks of **ferromagnetic** couplings. (Beware! It is not unusual to write the Hamiltonian with a minus sign in front, in which case $J > 0$ corresponds to a ferromagnetic coupling and vice versa).

2. The anisotropy of the couplings. If the couplings $J^x$, $J^y$ and $J^z$ are equal, like for the model derived from the Hubbard model, one speaks of **isotropic** couplings. Otherwise the model is said to be **anisotropic**. The most important special cases are the **Ising model** ($J^x = J^y = 0$) and the **XY model** ($J^z = 0$).

3. The range of the couplings.

   One often considers couplings limited to first neighbours, but they are physical situations in which one must consider longer range couplings, with consequences that can be quite dramatic (see the last section of this chapter).

4. The value of the spin S.

   In the case of the Hubbard model, one gets a model with spin 1/2. But the spin of a magnetic system can a priori take any half-integer or integer value S. Some properties depend in a crucial way on the value of $S$. It is sometimes very useful to consider that the spins are vectors of length 1 and not operators. This is called the **classical limit**. It corresponds in a certain sense to the $S \to +\infty$ limit of the quantum case.

5. The dimensionality of space.

> Most of the important properties are radically different depending on whether one deals with a system of dimension 1, 2, or 3.

It is to a large extent the large number of parameters that makes magnetism a vast subject, and we will limit ourselves to the investigation of a few special situations. Besides, there are only a few cases where exact results could be obtained, usually with rather involved calculations, and in most cases one has to resort to approximate methods to study a given property. The method to choose depends on the energy scale, or equivalently on the temperature scale one is interested in. If one wants to study low-temperature properties, i.e. $T/J \ll 1$, one should try to determine the ground state and the low-energy excited states. The most commonly used approach is spin-wave theory. By contrast, if one is interested in temperatures of the order of $J$ or larger, all states play an important role, and a good understanding of the low-energy states is no longer sufficient. The relevant techniques are those of statistical physics: mean field, renormalization group, etc.

## 2.5 Spin algebra

The commutation relations between spin operators are:

$$[S^x, S^y] = iS^z.$$

with the convention $\hbar = 1$. The ladder operators defined by:

$$\begin{aligned} S^+ &= S^x + iS^y \\ S^- &= S^x - iS^y, \end{aligned}$$

satisfy the commutation relations:

$$\begin{cases} [S^+, S^-] = [S^x, -iS^y] + [iS^y, S^x] = 2S^z \\ [S^z, S^+] = S^+ \\ [S^z, S^-] = -S^- \end{cases}$$

The action of the spin operators is given by :

$$\begin{aligned} S^+|m> &= \sqrt{S(S+1) - m(m+1)}\ |m+1> \\ S^-|m> &= \sqrt{S(S+1) - m(m-1)}\ |m-1> \end{aligned}$$

$$\begin{aligned} S^+S^-|m> &= (S(S+1) - m(m-1))\ |m> \\ S^-S^+|m> &= (S(S+1) - m(m+1))\ |m> \end{aligned}$$

One gets back the commutation relation:

$$\begin{aligned}\left(S^+S^- - S^-S^+\right)|m> &= \left(-m(m-1) + m(m+1)\right)|m> \\ &= 2m|m>\end{aligned}$$

and so

$$[S^+, S^-] = 2S^z$$

To determine the action of the operator $\vec{S}^2$, we first note that:

$$\left(S^+S^- + S^-S^+\right) = 2S(S+1) - 2m^2$$

But $(S^z)^2 = m^2$, leading to:

$$\left((S^z)^2 + \frac{1}{2}(S^+S^- + S^-S^+)\right) = m^2 + S(S+1) - m^2 = S(S+1).$$

Symmetries: Let us determine the symmetries of the spin Hamiltonian

$$H = \sum_{(i,j)} J_{ij}\ \vec{S}_i \cdot \vec{S}_j.$$

Let us define $\vec{S}_{tot} = \sum_i \vec{S}_i$. One can prove that:

1) $[S^\alpha_{tot}, H] = 0, \quad \alpha = x, y, z, +, -$

2) $[\vec{S}^2_{tot}, H] = 0$

Indeed,

$$\begin{aligned}[S^z_i, \vec{S}_i \cdot \vec{S}_j] &= [S^z_i, S^x_i]S^x_j + [S^z_i, S^y_i]S^y_j \\ &= iS^y_i S^x_j - iS^x_i S^y_j\end{aligned}$$

$$\begin{aligned}[S^z_j, \vec{S}_i \cdot \vec{S}_j] &= S^x_i[S^z_j, S^x_j] + S^y_i[S^z_j, S^y_j] \\ &= iS^x_i S^y_i - iS^y_i S^x_j\end{aligned}$$

$$\begin{aligned}&\Rightarrow\ [S^z_i + S^z_j, \vec{S}_i \cdot \vec{S}_j] = 0 \\ &\Rightarrow\ [\sum_i S^z_i, \vec{S}_i \cdot \vec{S}_j] = 0 \\ &\Rightarrow\ [S^z_{tot}, \vec{S}_i \cdot \vec{S}_j] = 0 \\ &\Rightarrow\ [S^z_{tot}, H] = 0\end{aligned}$$

Since the Hamiltonian $H$ only involves scalar products, this relation must be true for any direction:

$$[S^\alpha_{tot}, H] = 0, \quad \alpha = x, y, z$$

In addition, since $S^+_{tot}$ et $S^-_{tot}$ are linear combinations of $S^x_{tot}$ et $S^y_{tot}$, one also has:

$$[S^+_{tot}, H] = [S^-_{tot}, H] = 0$$

Finally, since$\vec{S}^2_{tot} = (S^z_{tot})^2 + (S^x_{tot})^2 + (S^y_{tot})^2$, one sees immediately that:

$$[\vec{S}^2_{tot}, H] = 0$$

## 2.6 Ferromagnetic case

Let us consider the Hamiltonian

$$H = -J \sum_{<i,j>} \vec{S}_i.\vec{S}_j, \quad J > 0$$

where $\sum_{<i,j>}$ denotes the sum over nearest-neighbour pairs on a Bravais lattice.

### 2.6.1 Ground state

Classical case:

It is often very useful to start by studying the so-called *classical* limit in which spins are assumed to be classical, 3-component vectors of unit length. In the case of a ferromagnetic system, the energy is minimal for the configurations where all spins point in the same direction since such a configuration independently minimizes the energy of each term of the sum.

Quantum case:

By analogy with the classical case, one looks for the quantum equivalent of states where all spins are aligned. To achieve this, one proceeds in three steps:

1) $|m_1 = S, m_2 = S, ..., m_N = S >$ is an eigenstate.

$S^z_i S^z_j = S^2$, $S^+_i S^-_j$ gives 0. This state is thus an eigenstate of energy:

$$E = -J \sum_{<ij>} S^2 = -JS^2 \underbrace{\left(\frac{z}{2}N\right)}_{\text{number of pairs}}$$

2) This is the state of lowest energy. For this to be true, it is sufficient to prove that:

$$< \vec{S}_i \cdot \vec{S}_j > \leq S^2$$

But

$$\begin{aligned}\vec{S}_i \cdot \vec{S}_j &= \frac{1}{2}\left[\left(\vec{S}_i + \vec{S}_j\right)^2 - \vec{S}_i^2 - \vec{S}_j^2\right] \\ &= \frac{1}{2}\left(\vec{S}_i + \vec{S}_j\right)^2 - S(S+1)\end{aligned}$$

and the maximum of $\vec{S}_i \cdot \vec{S}_j$ is given by:

$$\begin{aligned}\max(\vec{S}_i \cdot \vec{S}_j) &= \frac{1}{2}\max\left(\vec{S}_i + \vec{S}_j\right)^2 - S(S+1) \\ &= \frac{1}{2}2S(2S+1) - S(S+1) = S^2\end{aligned}$$

3) Degeneracy: $2NS+1$. Indeed, since $S^-_{tot}$ commutes with the Hamiltonian, one gets new ground states from the previous states by applying successively $S^-_{tot}$ until one reaches the state $|m_1 = -S, m_2 = -S, ..., m_N = -S >$, which is annihilated by $S^-_{tot}$.

### 2.6.2 Exact low-energy states:

Let's consider the state

$$|i> = |S \ldots \underbrace{S-1}_{i} S \ldots >$$

It is related to $|F>$ by $|i> = \frac{1}{\sqrt{2S}} S_i^- |F>$ since

$$S_i^- |F> = \sqrt{S(S+1) - S(S-1)} |S \ldots \underbrace{S-1}_{i} S \ldots >$$

Moreover,

$$S_i^+ |i> = \sqrt{2S} |F>$$

Let us act on it with the components of the Hamiltonian:

$$\begin{aligned}H^z |i> &= -J \sum_{<kl>} S_k^z S_l^z |i> = -JS^2 \left(\frac{z}{2}N - z\right) |i> -JS(S-1)z|i> \\ &= (-JS^2 \frac{z}{2}N + JSz)|i> = (E_0 + JSz)|i>\end{aligned}$$

$$\begin{aligned}
H^{\perp}|i> &= -J\sum_{j(i)}\frac{1}{2}\left(S_i^+S_j^- + \underbrace{S_i^-S_j^+}_{=0}\right)|i> \\
&= -\frac{J}{2}\sum_{j(i)} S_j^- \underbrace{S_i^+|i>}_{=\sqrt{2S}|F>} \\
&= -\frac{J}{2}\sum_{j(i)} 2S|j> \\
&= -JS\sum_{j(i)}|j>
\end{aligned}$$

Acting on the state $|\vec{k}>= \frac{1}{\sqrt{N}}\sum_i e^{-i\vec{k}\cdot\vec{R}_i}|i>$, this Hamiltonian gives:

$$\begin{aligned}
H^{\perp}|\vec{k}> &= \frac{1}{\sqrt{N}}\sum_i e^{-i\vec{k}\cdot\vec{R}_i}(-JS)\sum_{\vec{\tau}}|i+\vec{\tau}> \\
&= \frac{1}{\sqrt{N}}(-JS)\sum_{\vec{\tau}} e^{i\vec{k}\cdot\vec{\tau}}\sum_i e^{-i\vec{k}\cdot(\vec{R}_i+\vec{\tau})}|i+\vec{\tau}> \\
&= -JS\sum_{\vec{\tau}} e^{i\vec{k}\cdot\vec{\tau}}|\vec{k}>
\end{aligned}$$

$\Longrightarrow$

$$\begin{aligned}
H|\vec{k}> &= E_0|\vec{k}> + JSz|\vec{k}> - JS\sum_{\vec{\tau}} e^{i\vec{k}\cdot\vec{\tau}}|\vec{k}> \\
&= E_{\vec{k}}|\vec{k}>
\end{aligned}$$

with

$$E_{\vec{k}} = E_0 + JS\sum_{\vec{\tau}}\left(1 - e^{i\vec{k}\cdot\vec{\tau}}\right)$$

Nota Bene 1: For $\vec{k} = 0$, one finds $E_{\vec{k}} = E_0$. Is this possible? Yes. This state can be obtained from $|F>$ by a rotation of the total spin, and the Hamiltonian commutes with the total spin.

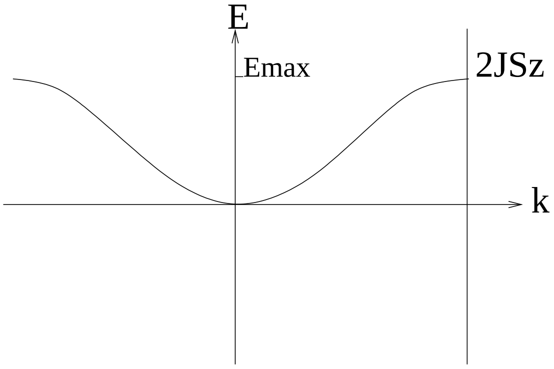
Nota Bene 2: This calculation is very similar to that done for the tight-binding model.

Dispersion:
$e^{i\vec{k}\cdot\vec{\tau}} + e^{-i\vec{k}\cdot\vec{\tau}} = 2\cos(\vec{k}\cdot\vec{\tau}), \quad \vec{\tau} = a\hat{x}, a\hat{y}, a\hat{z}$
After transforming $\sum_{\vec{\tau}}\left(1 - e^{i\vec{k}\cdot\vec{\tau}}\right)$:

$$\begin{aligned}
\sum_{\vec{\tau}}\left(1 - e^{i\vec{k}\cdot\vec{\tau}}\right) &= \sum_{\alpha} 2\left(1 - \cos(k_\alpha a)\right) \\
&= 4\sum_{\alpha}\sin^2\left(\frac{k_\alpha a}{2}\right)
\end{aligned}$$

$\Longrightarrow$

$$E_{\vec{k}} = E_0 + 4JS\sum_\alpha \sin^2\left(\frac{k_\alpha a}{2}\right)$$



For small $\vec{k}$: $\sin^2\left(\frac{k_\alpha a}{2}\right) \sim \frac{k_\alpha^2 a^2}{4} \Longrightarrow$

$$E_{\vec{k}} = E_0 + JS\vec{k}^2 a^2$$

The dispersion is quadratic for small $\vec{k}$.

**Exercise:** Calculate $<\vec{k}|S_i^x S_j^x + S_i^y S_j^y \vec{k}>$ et $<\vec{k}|S_i^z S_j^z|\vec{k}>$.

**Solution:** $S_i^x S_j^x + S_i^y S_j^y = \frac{1}{2}\left(S_i^+ S_j^- + S_i^- S_j^+\right), \quad |\vec{k}> = \frac{1}{\sqrt{N}}\sum_n e^{-i\vec{k}\cdot\vec{R}_n}|n>$
Let us first look at:

$$S_i^- S_j^+ |\vec{k}> = \frac{1}{\sqrt{N}} e^{-i\vec{k}\cdot\vec{R}_j} S_i^- S_j^+ |j>$$

Now

$$S_j^+|j> = \sqrt{S(S+1) - S(S-1)}|F> = \sqrt{2S}|F>$$

and

$$S_i^- S_j^+|j> = \sqrt{2S} S_i^-|F> = 2S|i>$$

fSo

$$S_i^- S_j^+|\vec{k}> = \frac{2S}{\sqrt{N}} e^{-i\vec{k}\cdot\vec{R}_j}|i> .$$

In the same way, one shows that:

$$S_i^+ S_j^-|\vec{k}> = \frac{2S}{\sqrt{N}} e^{-i\vec{k}\cdot\vec{R}_i}|j> .$$

We then get:

$$\begin{aligned}
<\vec{k}|\frac{1}{2}\left(S_i^+S_j^- + S_i^-S_j^+\right)|\vec{k}> &= \frac{2S}{2N}\sum_n e^{i\vec{k}\cdot\vec{R}_n}\left(e^{-i\vec{k}\cdot\vec{R}_j} <n|i> + e^{-i\vec{k}\cdot\vec{R}_i} <n|j>\right) \\
&= \frac{2S}{2N}\left(e^{i\vec{k}\cdot(\vec{R}_i-\vec{R}_j)} + e^{i\vec{k}\cdot(\vec{R}_j-\vec{R}_i)}\right) \\
&= \frac{2S}{N}\cos\vec{k}\cdot(\vec{R}_i-\vec{R}_j).
\end{aligned}$$

We now look for the expectation value of $<\vec{k}|S_i^zS_j^z|\vec{k}>$. Following the same procedure, we get:

$$S_j^z|\vec{k}>= \frac{1}{\sqrt{N}}\left(\sum_{n\neq j} Se^{-i\vec{k}\cdot\vec{R}_n}|n> + (S-1)e^{-i\vec{k}\cdot\vec{R}_j}|j>\right)$$

When $i \neq j$,

$$S_i^zS_j^z|\vec{k}>= \frac{1}{\sqrt{N}}\left(\sum_{n\neq i, n\neq j} S^2e^{-i\vec{k}\cdot\vec{R}_n}|n> + S(S-1)e^{-i\vec{k}\cdot\vec{R}_j}|j> + S(S-1)e^{-i\vec{k}\cdot\vec{R}_i}|i>\right)$$

Finally:

$$\begin{aligned}
<\vec{k}|S_i^zS_j^z|\vec{k}> &= \frac{1}{N}\sum_m e^{i\vec{k}\cdot\vec{R}_m}\left(\sum_{n\neq i, n\neq j} S^2e^{-i\vec{k}\cdot\vec{R}_n} <m|n>\right) \\
&+ \sum_m e^{i\vec{k}\cdot\vec{R}_m}\left(S(S-1)e^{-i\vec{k}\cdot\vec{R}_j} <m|j> + S(S-1)e^{-i\vec{k}\cdot\vec{R}_i} <m|i>\right) \\
&= \frac{1}{N}\left[\sum_{n\neq i, n\neq j} S^2 + S(S-1) + S(S-1)\right] \\
&= S^2 - \frac{2}{N}S
\end{aligned}$$

independent of the indices $i$ and $j$.

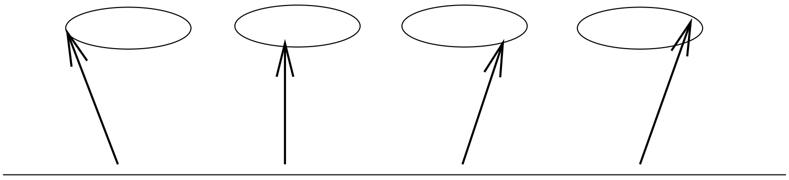


Figure 2.1: period $\sim \frac{1}{k}$

Low energy mode Goldstone Theorem (broken continuous symmetry $\Longrightarrow$ mode whose energy goes to 0 as $\vec{k}$ goes to $\vec{0}$).

Are there ther low-energy states? Yes, of course. One can overturn 2 spins, 3 spins, etc ..., but there is a problem. $|\vec{k},\vec{k'}>=\frac{1}{N}\sum_{i,j}e^{i\left(\vec{k}\vec{r_i}+\vec{k}\vec{r_j}\right)}S_i^-S_j^-|F>$ is not an eigenstate of $H$. Indeed, for the one magnon states, one looks for the solution as $|\Psi>=\sum_i A_iS_i^-|F>$ and we have seen that $A_i=e^{i\vec{k}\cdot\vec{R}_i}$ allows one to diagonalize $H$. For the two-magnon states, one must look for a solution of the form $|\Psi>=\sum_{i,j}A_{i,j}S_i^-S_j^-|F>$. But the coefficients $A_{i,j}$ must satisfy different equations depending on whether $i$ and $j$ are nearest neighbours or not. This problem does not have an exact solution except in dimension 1 for spins 1/2 (Bethe ansatz, 1931).

### 2.6.3 Holstein-Primakoff transformation

For a given $S$, and provided it is not too small, one can create several spin deviations on a given site → the operators $S_i^{\pm}$ are rather bosons than fermions.

Exact representation:

$$\begin{cases} S_i^z = S - a_i^+a_i \\ S_i^+ = \sqrt{2S}\,\sqrt{1-\frac{a_i^+a_i}{2S}}\;a_i \\ S_i^- = \sqrt{2S}\;a_i^+\sqrt{1-\frac{a_i^+a_i}{2S}} \end{cases}$$

Approximations:

- One forgets that the condition $a_i^+a_i \leq 2S$ should be fulfilled.
- One only keeps terms linear in S in the Hamiltonian. So in particular one can assume: $S_i^+=\sqrt{2S}a_i$ and $S_i^-=\sqrt{2S}a_i^+$.

Let us check the commutation relations:

$$\begin{aligned} [S^+,S^-] &= S^+S^- - S^-S^+ \\ &= 2S\left[\sqrt{1-\frac{n}{2S}}\,\underbrace{a_ia_i^+}_{1+a_i^+a_i}\,\sqrt{1-\frac{n}{2S}} - a_i^+(1-\frac{n}{2S})a_i\right] \\ &= 2S\left[(1-\frac{n}{2S})+\sqrt{1-\frac{n}{2S}}a_i^+a_i\sqrt{1-\frac{n}{2S}} - a_i^+(1-\frac{n}{2S})a_i\right] \end{aligned}$$

Using the relations:

$$a_i^+a_ia_i = (a_ia_i^+-1)a_i$$
$$n_ia_i = a_in_i - a_i$$

and

$$\begin{aligned} a_i^+\left(1-\frac{n_i}{2S}\right)a_i &= a_i^+a_i - \frac{a_i^+n_ia_i}{2S} \\ &= a_i^+a_i\left(1-\frac{n_i}{2S}\right)+\frac{n_i}{2S} \end{aligned}$$

leads to:

$$\begin{aligned}
[S^+, S^-] &= 2S\left[1-\frac{n_i}{2S}+\left(1-\frac{n_i}{2S}\right)n_i-\left(1-\frac{n_i}{2S}\right)n_i-\frac{n_i}{2S}\right] \\
&= 2S\left[1-\frac{n_i}{S}\right] = 2S-2n = 2S^z.
\end{aligned}$$

In the same spirit,

$$\begin{aligned}
\frac{1}{\sqrt{2S}}[S^z, S^+] &= [n, \frac{S^+}{\sqrt{2S}}] = -[n, \sqrt{1-\frac{n}{2S}}a_i] \\
&= -n\sqrt{1-\frac{n}{2S}}a_i + \sqrt{1-\frac{n}{2S}}a_i n \\
&= -\sqrt{1-\frac{n}{2S}}[n, a_i] = -\sqrt{1-\frac{n}{2S}}(-a_i) = \frac{S^+}{\sqrt{2S}}
\end{aligned}$$

Let us now calculate the Hamiltonian:

$$\begin{aligned}
H &= -J\left(\sum_{<i,j>} S^2 - Sa_i^+a_i - Sa_j^+a_j + \frac{1}{2}2S(a_i^+a_j + a_j^+a_i)\right) \\
&= -J\frac{Nz}{2}S^2 - JS\sum_{<i,j>}\left(-a_i^+a_i - a_j^+a_j + a_i^+a_j + a_j^+a_i\right) \\
&= -J\frac{Nz}{2}S^2 - \frac{JS}{2}\sum_{i,\vec{\tau}}\left(-a_i^+a_i - a_{i+\vec{\tau}}^+a_{i+\vec{\tau}} + a_i^+a_{i+\vec{\tau}} + a_{i+\vec{\tau}}^+a_i\right)
\end{aligned}$$

Let us define the Fourier transforms as:

$$\begin{aligned}
&a_{\vec{k}}^+ = \tfrac{1}{\sqrt{N}}\textstyle\sum_i e^{i\vec{k}\cdot\vec{r}_i}a_i^+, \quad a_{\vec{k}} = \tfrac{1}{\sqrt{N}}\textstyle\sum_i e^{-i\vec{k}\cdot\vec{r}_i}a_i \\
&a_i^+ = \tfrac{1}{\sqrt{N}}\textstyle\sum_{\vec{k}} e^{-i\vec{k}\cdot\vec{r}_i}a_{\vec{k}}^+, \quad a_i = \tfrac{1}{\sqrt{N}}\textstyle\sum_{\vec{k}} e^{i\vec{k}\cdot\vec{r}_i}a_{\vec{k}}
\end{aligned}$$

The Hamiltonian can be written:

$$\begin{aligned}
H &= -J\frac{Nz}{2}S^2 - \frac{JS}{2}\sum_{i,\vec{\tau}}\frac{1}{N}\sum_{\vec{k}_1,\vec{k}_2} \\
&\quad \left(-e^{-i\vec{k}_1\cdot\vec{r}_i+i\vec{k}_2\cdot\vec{r}_i} - e^{-i\vec{k}_1\cdot(\vec{r}_i+\vec{\tau})+i\vec{k}_2\cdot(\vec{r}_i+\vec{\tau})} + e^{-i\vec{k}_1\cdot\vec{r}_i+i\vec{k}_2\cdot(\vec{r}_i+\vec{\tau})} + e^{-i\vec{k}_1\cdot(\vec{r}_i+\vec{\tau})+i\vec{k}_2\cdot\vec{r}_i}\right) \\
&\times\ a_{\vec{k}_1}^+a_{\vec{k}_2} \\
&= -J\frac{Nz}{2}S^2 - \frac{JS}{2}\sum_{\vec{\tau}}\sum_{\vec{k}_1,\vec{k}_2}\left(-1 - e^{-i\vec{k}_1\cdot\vec{\tau}+i\vec{k}_2\cdot\vec{\tau}} + e^{i\vec{k}_2\cdot\vec{\tau}} + e^{-i\vec{k}_1\cdot\vec{\tau}}\right)\delta_{\vec{k}_1,\vec{k}_2}a_{\vec{k}_1}^+a_{\vec{k}_2} \\
&= -J\frac{Nz}{2}S^2 - \frac{JS}{2}\sum_{\vec{\tau}}\sum_{\vec{k}}\left(-1 - 1 + e^{i\vec{k}\cdot\vec{\tau}} + e^{-i\vec{k}\cdot\vec{\tau}}\right)a_{\vec{k}}^+a_{\vec{k}}
\end{aligned}$$

Finally

$$H = E_0 + \sum_{\vec{k}} \epsilon_{\vec{k}}\ a^+_{\vec{k}} a_{\vec{k}}$$

with

$$\epsilon_{\vec{k}} = JS \sum_{\vec{\tau}} \left(1 - e^{i\vec{k}\cdot\vec{\tau}}\right)$$

Interpretation: Let us adopt the representation in terms of the boson occupation numbers.

$|0, \cdots, 0> \quad \rightarrow \quad$ Ground state

$|0, \cdots, \underbrace{1}_{\vec{k}}, \cdots, 0> \quad \rightarrow \quad$ one magnon state

$|0, \cdots, \underbrace{1}_{\vec{k}_1}, \cdots, 0, \cdots, \underbrace{1}_{\vec{k}_2}, \cdots, 0> \quad \rightarrow \quad$ two-magnon state$\cdots$

Thermodynamics of spin waves:

$< S^z_i >= S - < a^+_i a_i >$ with $a^+_i a_i = \frac{1}{N} \sum_{\vec{k}_1, \vec{k}_2} e^{-i\vec{k}_1 \vec{r}_i + i\vec{k}_2 \vec{r}_i} a^+_{\vec{k}_1} a_{\vec{k}_2}$.

The magnetization is given by:

$$\begin{aligned} < m > &= \frac{1}{N} \sum_i < S^z_i >= S - \frac{1}{N^2} \sum_i \sum_{\vec{k}_1, \vec{k}_2} e^{-i\vec{k}_1 \vec{r}_i + i\vec{k}_2 \vec{r}_i} < a^+_{\vec{k}_1} a_{\vec{k}_2} > \\ &= S - \frac{1}{N} \sum_{\vec{k}} < a^+_{\vec{k}} a_{\vec{k}} > \\ &= S - \frac{1}{N} \sum_{\vec{k}} \frac{1}{e^{\beta \epsilon_{\vec{k}}} - 1} \\ &= S - \frac{1}{(2\pi)^D} \int \frac{d\vec{k}}{e^{\beta \epsilon_{\vec{k}}} - 1}, \end{aligned}$$

with $\beta = \frac{1}{k_B T}$.

Remember that $\epsilon_{\vec{k}} \sim JSa^2k^2$. Let us evaluate the contribution from small wave vectors:

$$< m >= S - \underbrace{\frac{a^D}{(2\pi)^D} \int_{|\vec{k}| \le k_0} \frac{d\vec{k}}{e^{\beta JSa^2k^2} - 1}}_{I} - \frac{a^D}{(2\pi)^D} \int_{|\vec{k}| > k_0} \frac{d\vec{k}}{e^{\beta \epsilon_{\vec{k}}} - 1}$$

Let us suppose that $\beta$ is fixed. For $k$ small enough, $e^{\beta JSa^2k^2} \sim 1 + \beta Ja^2k^2$. The integral $I$ becomes:

$$I = \frac{a^D}{(2\pi)^D} \int_{|\vec{k}| \leq k_0} \frac{d\vec{k}}{\beta JSa^2k^2}$$

- <u>3D:</u> This is a convergent integral because $d\vec{k} \propto k^2 dk$. To get the temperature dependence, we go back to the original integral and perform the change of variables $u = \sqrt{\beta}k$. This leads to:

  $$\int \frac{k^2 dk}{e^{\beta JSa^2k^2} - 1} = \frac{1}{\beta^{3/2}} \int \frac{u^2 du}{e^{JSa^2u^2} - 1} \propto T^{3/2} \tag{2.5}$$

  This is Bloch's law: the correction to the magnetization of a ferromagnet at low temperature is proportional to $T^{3/2}$.

- <u>2D:</u>

  $$\frac{a^2}{2\pi} \frac{1}{\beta JS} \int_{|\vec{k}|=0}^{k_0} \frac{k dk}{k^2} \qquad \text{Diverges!}$$

- <u>1D:</u>

  $$\frac{a}{(2\pi)} \frac{1}{\beta JS} \int_{|\vec{k}|=0}^{k_0} \frac{dk}{k^2} \qquad \text{Diverges even more!}$$

In $D = 1$ and $D = 2$, there is no magnetic order at finite temperature (Hohenberg-Mermin-Wagner theorem). By contrast, in dimension 3, there is a finite magnetization if the temperature is not too high.

## 2.7 Antiferromagnetic case

### 2.7.1 Classical case

The classical ground state is the Néel state with all spins ↑ on sublattice A and all spins ↓ on sublattice B:

$$\begin{array}{cccc} \uparrow & \downarrow & \uparrow & \downarrow \\ \downarrow & \uparrow & \downarrow & \uparrow \end{array}$$

### 2.7.2 Quantum formulation

By analogy, let us consider the state:

$$|m_1^A = S, \cdots, m_{\frac{N}{2}}^A = S, m_1^B = -S, \cdots, m_{\frac{N}{2}}^B = -S > .$$

Is this the ground state? Certainly not: it is not even an eigenstate!

To see this, let us consider two neighbouring sites ($i \in A$, $j \in B$):

$$\begin{cases} S_i^+ S_j^- \to 0 \\ S_i^- S_j^+ \to m_i^A = S - 1, \quad m_j^B = -S + 1 \end{cases}$$

This state is coupled to other states. How to proceed?

The situation is somehow similar to the ferromagnetic case at $T > 0$. One must include fluctuations.

Holstein-Primakoff transformation

Since spins point in different directions on the two sublattices, one must distinguish two cases:

Sublattice A

$$\begin{cases} S_i^z = S - a_i^+ a_i \\ S_i^+ = \sqrt{2S}\sqrt{1 - \frac{n_i}{2S}}\ a_i \\ S_i^- = \sqrt{2S}\ a_i^+ \sqrt{1 - \frac{n_i}{2S}} \end{cases}$$

Sublattice B

$$\begin{cases} S_j^z = -S + a_j^+ a_j \\ S_j^+ = \sqrt{2S}\ a_j^+ \sqrt{1 - \frac{n_j}{2S}} \\ S_j^- = \sqrt{2S}\sqrt{1 - \frac{n_j}{2S}}\ a_j \end{cases}$$

Linear approximation

Keeping only terms of order $S^2$ and $S$, the Hamiltonian becomes:

$$\begin{aligned} H &= J \sum_{i\in A} \sum_{j(i)} \left( -S^2 + S a_i^+ a_i + S a_j^+ a_j + \frac{2S}{2} \left( a_i a_j + a_i^+ a_j^+ \right) \right) \\ &= -JS^2 \frac{N}{2} z + SJ \sum_{i \in A} \sum_{\vec{\tau}} \left( a_i^+ a_i + a_{i+\vec{\tau}}^+ a_{i+\vec{\tau}} + a_i a_{i+\vec{\tau}} + a_{i+\vec{\tau}}^+ a_{i+\vec{\tau}}^+ \right) \\ &= -JS^2 \frac{N}{2} z + \frac{JS}{2} \sum_i \sum_{\vec{\tau}} \frac{1}{N} \sum_{\vec{k}_1,\vec{k}_2} \left[ \left( e^{-i\vec{k}_1\vec{r}_i + i\vec{k}_2\vec{r}_i} + e^{i(-\vec{k}_1+\vec{k}_2)(\vec{r}_i+\vec{\tau})} \right) a_{\vec{k}_1}^+ a_{\vec{k}_2} \right] \\ &+ \frac{JS}{2} \sum_i \sum_{\vec{\tau}} \frac{1}{N} \sum_{\vec{k}_1,\vec{k}_2} \left[ e^{i\vec{k}_1\vec{r}_i + i\vec{k}_2(\vec{r}_i+\vec{\tau})} a_{\vec{k}_1} a_{\vec{k}_2} + e^{-i\vec{k}_1\vec{r}_i - i\vec{k}_2(\vec{r}_i+\vec{\tau})} a_{\vec{k}_1}^+ a_{\vec{k}_2}^+ \right] \\ &= -JS^2 \frac{N}{2} z + \frac{JS}{2} \sum_{\vec{\tau}} \sum_{\vec{k}} \left( a_{\vec{k}}^+ a_{\vec{k}} + a_{\vec{k}}^+ a_{\vec{k}} + a_{\vec{k}} a_{-\vec{k}} e^{-i\vec{k}\vec{\tau}} + a_{\vec{k}}^+ a_{-\vec{k}}^+ e^{i\vec{k}\vec{\tau}} \right) \end{aligned}$$

Problem: The Hamiltonian is not the sum of harmonic oscillators because it contains terms of the form $a_{\vec{k}} a_{-\vec{k}}$ and $a_{\vec{k}}^+ a_{-\vec{k}}^+$.

### 2.7.3 Bogolioubov transformation

If the Hamiltonian was of the form:

$$H = \sum_{\vec{k}} \omega_{\vec{k}} \alpha^+_{\vec{k}} \alpha_{\vec{k}} + \mathrm{Cte}$$

one would have:

$$[H, \alpha_{\vec{k}}] = -\omega_{\vec{k}} \alpha_{\vec{k}}$$

Now, since

$$[a^+_{\vec{k}} a_{\vec{k}}, a_{\vec{k}}] = \underbrace{a^+_{\vec{k}} a_{\vec{k}}}_{-1+a_{\vec{k}}a^+_{\vec{k}}} \ a_{\vec{k}} - a_{\vec{k}} a^+_{\vec{k}} a_{\vec{k}} = -a_{\vec{k}},$$

$$[a^+_{\vec{k}} a^+_{-\vec{k}}, a_{\vec{k}}] = [a^+_{\vec{k}}, a_{\vec{k}}] a^+_{-\vec{k}} = -a^+_{-\vec{k}}.$$

on gets the following expression for the commutator:

$$\begin{aligned} [H, a_{\vec{k}}] &= -\frac{JS}{2} \sum_{\vec{\tau}} \left( a_{\vec{k}} + a_{\vec{k}} + a^+_{-\vec{k}} \left( e^{i\vec{k}\vec{\tau}} + e^{-i\vec{k}\vec{\tau}} \right) \right) \\ &= -JS \sum_{\vec{\tau}} \left( a_{\vec{k}} + a^+_{-\vec{k}} \cos \vec{k} \cdot \vec{\tau} \right). \end{aligned}$$

Similarly, since

$$[a^+_{\vec{k}} a_{\vec{k}}, a^+_{\vec{k}}] = a^+_{\vec{k}} \underbrace{a_{\vec{k}} a^+_{\vec{k}}}_{1+a^+_{\vec{k}} a_{\vec{k}}} - a^+_{\vec{k}} a^+_{\vec{k}} a_{\vec{k}} = a^+_{\vec{k}},$$

$$[a_{\vec{k}} a_{-\vec{k}}, a^+_{\vec{k}}] = +a_{-\vec{k}}$$

one gets:

$$[H, a^+_{-\vec{k}}] = JS \sum_{\vec{\tau}} \left( a^+_{-\vec{k}} + a_{\vec{k}} \cos \vec{k} \cdot \vec{\tau} \right).$$

Since $[H, a_{\vec{k}}]$ and $[H, a^+_{-\vec{k}}]$ are linear combinations of $a_{\vec{k}}$ and $a^+_{-\vec{k}}$, it must be possible to bring the Hamiltonian into the required form with bosonic operators $\alpha^+_{\vec{k}}$ and $\alpha_{\vec{k}}$ that are linear combinations of $a_{\vec{k}}$ et $a^+_{-\vec{k}}$. More precisely, let us look for operators of the form:

$$\alpha_{\vec{k}} = u_{\vec{k}} a_{\vec{k}} + v_{\vec{k}} a^+_{-\vec{k}}$$

The coefficients $u_{\vec{k}}$ and $v_{\vec{k}}$ are determined by the conditions:

$$\begin{aligned}
[H, \alpha_{\vec{k}}] &= JS \sum_{\vec{\tau}} \left[ \left( -u_{\vec{k}} + v_{\vec{k}} \cos \vec{k} \cdot \vec{\tau} \right) a_{\vec{k}} + \left( v_{\vec{k}} - u_{\vec{k}} \cos \vec{k} \cdot \vec{\tau} \right) a^+_{-\vec{k}} \right] \\
&= -\omega_{\vec{k}} \alpha_{\vec{k}} \quad , \quad \omega_{\vec{k}} > 0
\end{aligned}$$

which leads to:

$$\begin{cases}
-u_{\vec{k}} JSz + v_{\vec{k}} JSz_{\vec{k}} = -\omega_{\vec{k}} u_{\vec{k}} \\
-u_{\vec{k}} JSz_{\vec{k}} + v_{\vec{k}} JSz = -\omega_{\vec{k}} v_{\vec{k}}
\end{cases}$$

where we have introduced the notation:

$$z_{\vec{k}} = \sum_{\vec{\tau}} \cos \vec{k} \cdot \vec{\tau}$$

Solutions exist provided:

$$\left( -\omega_{\vec{k}} - JSz \right) \left( JSz - \omega_{\vec{k}} \right) + (JSz_{\vec{k}})^2 = 0$$

$$\Rightarrow \omega_{\vec{k}}^2 = (JSz)^2 - (JSz_{\vec{k}})^2$$

Besides, to ensure that $[\alpha_{\vec{k}}, \alpha^+_{\vec{k}}] = 1$, the coefficients $u_{\vec{k}}$ and $v_{\vec{k}}$ must satisfy the normalization condition $u^2_{\vec{k}} - v^2_{\vec{k}} = 1$. But from the above equation we get

$$u^2_{\vec{k}} - v^2_{\vec{k}} = u^2_{\vec{k}} \frac{2\omega_{\vec{k}}}{JSz + \omega_{\vec{k}}}.$$

Since $JSz + \omega_{\vec{k}} > 0$, this imposes to choose $\omega_{\vec{k}} > 0$:

$$\boxed{\omega_{\vec{k}} = JSz \sqrt{1 - \gamma^2_{\vec{k}}}}$$

with $\gamma_{\vec{k}} = \frac{z_{\vec{k}}}{z} = \cos k_x$ in1D, $\frac{1}{2}(\cos k_x + \cos k_y)$ in 2D $\cdots$ For small $k$, $\omega_{\vec{k}} \propto k$: In an antiferromagnet, the dispersion of spin waves is *linear* in $k$.

The eigenvalue equations imply:

$$\begin{aligned}
\frac{v_{\vec{k}}}{u_{\vec{k}}} &= \frac{JSz_{\vec{k}}}{JSz + \omega_{\vec{k}}} = \text{sign}(z_{\vec{k}}) \frac{\sqrt{(JSz - \omega_{\vec{k}})(JSz + \omega_{\vec{k}})}}{JSz + \omega_{\vec{k}}} \\
&= \text{sign}(z_{\vec{k}}) \sqrt{\frac{JSz - \omega_{\vec{k}}}{JSz + \omega_{\vec{k}}}}
\end{aligned}$$

Besides, the normalisation condition imposes

$$u^2_{\vec{k}} \left( \frac{2\omega_{\vec{k}}}{JSz + \omega_{\vec{k}}} \right) = 1$$

One can thus choose as solutions

$$u_{\vec{k}} = \sqrt{\frac{JSz + \omega_{\vec{k}}}{2\omega_{\vec{k}}}} \qquad v_{\vec{k}} = \mathrm{sign}(z_{\vec{k}})\sqrt{\frac{JSz - \omega_{\vec{k}}}{2\omega_{\vec{k}}}}$$

The Hamiltonian is thus of the form:

$$H = \sum_{\vec{k}} \omega_{\vec{k}} \alpha^+_{\vec{k}} \alpha_{\vec{k}} + \mathrm{Cte}$$

To determine the constant, one must invert these relations

$$\begin{cases} \alpha_{\vec{k}} = u_{\vec{k}} a_{\vec{k}} + v_{\vec{k}} a^+_{-\vec{k}} \\ \alpha^+_{-\vec{k}} = u_{\vec{k}} a^+_{-\vec{k}} + v_{\vec{k}} a_{\vec{k}} \end{cases}$$

which leads to:

$$\begin{cases} a_{\vec{k}} = u_{\vec{k}} \alpha_{\vec{k}} - v_{\vec{k}} \alpha^+_{-\vec{k}} \\ a^+_{-\vec{k}} = -v_{\vec{k}} \alpha_{\vec{k}} + u_{\vec{k}} \alpha^+_{-\vec{k}} \end{cases} \qquad \begin{aligned} a^+_{\vec{k}} &= u_{\vec{k}} \alpha^+_{\vec{k}} - v_{\vec{k}} \alpha_{-\vec{k}} \\ a_{-\vec{k}} &= -v_{\vec{k}} \alpha^+_{\vec{k}} + u_{\vec{k}} \alpha_{-\vec{k}} \end{aligned}$$

and introduce these expressions into the Hamiltonian written in terms of the $a^+_{\vec{k}}$ , $a_{\vec{k}}$ given by:

$$H = -JS^2 \frac{zN}{2} + \frac{JS}{2} \sum_{\vec{k}} \left( 2z a^+_{\vec{k}} a_{\vec{k}} + z_{\vec{k}} a_{\vec{k}} a_{-\vec{k}} + z_{\vec{k}} a^+_{\vec{k}} a^+_{-\vec{k}} \right)$$

Now,

$$\begin{aligned} a^+_{\vec{k}} a_{\vec{k}} &= \left( u_{\vec{k}} \alpha^+_{\vec{k}} - v_{\vec{k}} \alpha_{-\vec{k}} \right) \left( u_{\vec{k}} \alpha_{\vec{k}} - v_{\vec{k}} \alpha^+_{-\vec{k}} \right) \\ &= u^2_{\vec{k}}\ \alpha^+_{\vec{k}} \alpha_{\vec{k}} + v^2_{\vec{k}}\ \alpha_{-\vec{k}} \alpha^+_{-\vec{k}} - u_{\vec{k}} v_{\vec{k}} \left( \alpha^+_{\vec{k}} \alpha^+_{-\vec{k}} + \alpha_{\vec{k}} \alpha_{-\vec{k}} \right) \\ &= u^2_{\vec{k}}\ \alpha^+_{\vec{k}} \alpha_{\vec{k}} + v^2_{\vec{k}}\ \alpha^+_{-\vec{k}} \alpha_{-\vec{k}} + v^2_{\vec{k}} - u_{\vec{k}} v_{\vec{k}} \left( \alpha^+_{\vec{k}} \alpha^+_{-\vec{k}} + \alpha_{\vec{k}} \alpha_{-\vec{k}} \right) \end{aligned}$$

$$\begin{aligned} a_{\vec{k}} a_{-\vec{k}} &= \left( u_{\vec{k}} \alpha_{\vec{k}} - v_{\vec{k}} \alpha^+_{-\vec{k}} \right) \left( -v_{\vec{k}} \alpha^+_{\vec{k}} + u_{\vec{k}} \alpha_{-\vec{k}} \right) \\ &= -u_{\vec{k}} v_{\vec{k}} \left( \alpha_{\vec{k}} \alpha^+_{\vec{k}} + \alpha^+_{-\vec{k}} \alpha_{-\vec{k}} \right) + u^2_{\vec{k}}\ \alpha_{\vec{k}} \alpha_{-\vec{k}} + v^2_{\vec{k}}\ \alpha^+_{-\vec{k}} \alpha^+_{\vec{k}} \\ &= -u_{\vec{k}} v_{\vec{k}} \left( \alpha^+_{\vec{k}} \alpha_{\vec{k}} + \alpha^+_{-\vec{k}} \alpha_{-\vec{k}} \right) - u_{\vec{k}} v_{\vec{k}} + u^2_{\vec{k}}\ \alpha_{\vec{k}} \alpha_{-\vec{k}} + v^2_{\vec{k}}\ \alpha^+_{-\vec{k}} \alpha^+_{\vec{k}} \end{aligned}$$

$$\begin{aligned} a^+_{\vec{k}} a^+_{-\vec{k}} &= \left( u_{\vec{k}} \alpha^+_{\vec{k}} - v_{\vec{k}} \alpha_{-\vec{k}} \right) \left( -v_{\vec{k}} \alpha_{\vec{k}} + u_{\vec{k}} \alpha^+_{-\vec{k}} \right) \\ &= -u_{\vec{k}} v_{\vec{k}} \left( \alpha^+_{\vec{k}} \alpha_{\vec{k}} + \alpha_{-\vec{k}} \alpha^+_{-\vec{k}} \right) + u^2_{\vec{k}}\ \alpha^+_{\vec{k}} \alpha^+_{-\vec{k}} + v^2_{\vec{k}}\ \alpha_{-\vec{k}} \alpha_{\vec{k}} \\ &= -u_{\vec{k}} v_{\vec{k}} \left( \alpha^+_{\vec{k}} \alpha_{\vec{k}} + \alpha^+_{-\vec{k}} \alpha_{-\vec{k}} \right) - u_{\vec{k}} v_{\vec{k}} + u^2_{\vec{k}}\ \alpha^+_{\vec{k}} \alpha^+_{-\vec{k}} + v^2_{\vec{k}}\ \alpha_{-\vec{k}} \alpha_{\vec{k}} \end{aligned}$$

Let us regroup the terms:

The constant term is given by $2zv_{\vec{k}}^2 - 2z_{\vec{k}}u_{\vec{k}}v_{\vec{k}}$. But according to the eigenvalue equations, one has:

$$v_{\vec{k}}JSz - u_{\vec{k}}JSz_{\vec{k}} = -\omega_{\vec{k}}v_{\vec{k}}$$

and so:

$$\begin{aligned}
\frac{JS}{2}\sum_{\vec{k}}\left(2zv_{\vec{k}}^2 - 2z_{\vec{k}}u_{\vec{k}}v_{\vec{k}}\right) &= \sum_{\vec{k}}\left(JSzv_{\vec{k}}^2 - JSu_{\vec{k}}z_{\vec{k}}v_{\vec{k}}\right)\\
&= -\sum_{\vec{k}}\omega_{\vec{k}}v_{\vec{k}}^2\\
&= -\sum_{\vec{k}}\omega_{\vec{k}}\frac{JSz-\omega_{\vec{k}}}{2\omega_{\vec{k}}} = -\frac{JSz}{2}\underbrace{\sum_{\vec{k}}1}_{N}+\frac{1}{2}\sum_{\vec{k}}\omega_{\vec{k}}
\end{aligned}$$

The constant term of the Hamiltonian is thus equal to:

$$-JS(S+1)\frac{zN}{2}+\frac{1}{2}\sum_{\vec{k}}\omega_{\vec{k}}$$

The non-constant terms are given by:

$$\begin{aligned}
\alpha_{\vec{k}}^{+}\alpha_{\vec{k}} &\times \left(2zu_{\vec{k}}^2 + 2zv_{\vec{k}}^2 - 4z_{\vec{k}}u_{\vec{k}}v_{\vec{k}}\right)\\
&= \alpha_{\vec{k}}^{+}\alpha_{\vec{k}}\left(\left(2zu_{\vec{k}} - 2z_{\vec{k}}v_{\vec{k}}\right)u_{\vec{k}} + \left(2zv_{\vec{k}} - 2z_{\vec{k}}u_{\vec{k}}\right)v_{\vec{k}}\right)\\
&= \alpha_{\vec{k}}^{+}\alpha_{\vec{k}}\left(\frac{2}{JS}\omega_{\vec{k}}u_{\vec{k}}^2 - \frac{2}{JS}\omega_{\vec{k}}v_{\vec{k}}^2\right)\\
&= \alpha_{\vec{k}}^{+}\alpha_{\vec{k}}\frac{2}{JS}\omega_{\vec{k}}
\end{aligned}$$

which leads to a contribution to the Hamiltonian equal to $\sum_{\vec{k}}\omega_{\vec{k}}\alpha_{\vec{k}}^{+}\alpha_{\vec{k}}$. Finally,

$$\begin{aligned}
\alpha_{\vec{k}}\alpha_{-\vec{k}} &\times \left(-2zu_{\vec{k}}v_{\vec{k}} + z_{\vec{k}}u_{\vec{k}}^2 + z_{\vec{k}}v_{\vec{k}}^2\right)\\
&= \alpha_{\vec{k}}\alpha_{-\vec{k}}\left(\left(z_{\vec{k}}u_{\vec{k}} - zv_{\vec{k}}\right)u_{\vec{k}} + \left(z_{\vec{k}}v_{\vec{k}} - zu_{\vec{k}}\right)v_{\vec{k}}\right)\\
&= \alpha_{\vec{k}}\alpha_{-\vec{k}}\left(\omega_{\vec{k}}u_{\vec{k}}v_{\vec{k}} - \omega_{\vec{k}}u_{\vec{k}}v_{\vec{k}}\right)\\
&= 0
\end{aligned}$$

This leads to our final result:

$$\boxed{H = -JS(S+1)\tfrac{zN}{2} + \textstyle\sum_{\vec{k}}\left(\alpha_{\vec{k}}^{+}\alpha_{\vec{k}} + \tfrac{1}{2}\right)\omega_{\vec{k}}}$$

Staggered magnetization

The equivalent of the magnetization of the ferromagnetic case is the staggered magnetization defined by:

$$M_{\text{stag}} = \frac{1}{N} < \sum_{i\in A} S_z^i - \sum_{i\in B} S_z^j >$$

According to the Holstein-Primakoff transformation, we have:

$$M_{\text{stag}} = S - \frac{1}{N} \sum_{\vec{k}} < a_{\vec{k}}^+ a_{\vec{k}} >$$

Using the expression of $a_{\vec{k}}^+ a_{\vec{k}}$ in terms of the operators $\alpha_{\vec{k}}^+$, $\alpha_{\vec{k}}$, we get:

$$\begin{aligned} < a_{\vec{k}}^+ a_{\vec{k}} > &= u_{\vec{k}}^2\, n(\vec{k}) + v_{\vec{k}}^2 \left(1 + n(\vec{k})\right) \\ &= \left(u_{\vec{k}}^2 + v_{\vec{k}}^2\right) n(\vec{k}) + v_{\vec{k}}^2 \end{aligned}$$

- $\boxed{\mathrm{T} = 0}$:

  $$< a_{\vec{k}}^+ a_{\vec{k}} >= v_{\vec{k}}^2 = \frac{JSz - \omega_{\vec{k}}}{2\omega_{\vec{k}}} \neq 0 \; !$$

  There are quantum fluctuations.

  *D = 1 :*

  $$< m >= S - \frac{2}{2\pi} \int_0^{\frac{\pi}{a}} dk \frac{JSz - \omega_k}{2\omega_k}$$

  But for $k$ small, $\omega_k \sim k$. As a consequence, the integral diverges and there is no antiferromagnetic order.

  *D ≥ 2 :* There is antiferromagnetic order if the integral is smaller than $S$. It is often true. For instance, on the square lattice, the integral is approximately equal to 0.2. This calculation thus predicts that there must be antiferromagnetic order for all $S$.

- $\boxed{\mathrm{T} \neq 0}$:

  $$< a_{\vec{k}}^+ a_{\vec{k}} >= \underbrace{\frac{JSz}{\omega_{\vec{k}}} \frac{1}{e^{\beta\omega_{\vec{k}}} - 1}}_{\sim \frac{1}{k^2} \text{for k small}} + \frac{JSz - \omega_{\vec{k}}}{2\omega_{\vec{k}}}$$

  *D ≤ 2*: The integral diverges → no long-range antiferromagnetic order at $T > 0$ (Hohenberg-Mermin-Wagner theorem).

Spin-wave ground state:

The vacuum for the $a^+$ bosons is the Néel state. This is not the ground state! The ground state satisfies $\alpha_{\vec{k}}|F>=0$. But $\left(u_{\vec{k}}a_{\vec{k}}+v_{\vec{k}}a^+_{-\vec{k}}\right)|0>=v_{\vec{k}}a^+_{-\vec{k}}|0>\neq 0$.

One can show that the ground state is given by:

$$|F>=\prod_{\vec{k}}{}' \frac{1}{u_{\vec{k}}}\exp\left(-\frac{v_{\vec{k}}}{u_{\vec{k}}}a^+_{\vec{k}}a^+_{-\vec{k}}\right)|0>$$

where $\prod'_{\vec{k}}$ means that the product is restricted to wave vectors $\vec{k}$ and $\vec{k}'$ such that $\vec{k}\neq-\vec{k}'$. To this end, one much show that:

$$\alpha_{\vec{k}}|F>=0$$

with $\alpha_{\vec{k}}=u_{\vec{k}}a_{\vec{k}}+v_{\vec{k}}a^+_{-\vec{k}}$.

$$\begin{aligned} a_{\vec{k}}|F> &= a_{\vec{k}}\prod_{\vec{l}}{}'\frac{1}{u_{\vec{l}}}\exp\left(-\frac{v_{\vec{l}}}{u_{\vec{l}}}a^+_{\vec{l}}a^+_{-\vec{l}}\right)|0> \\ &= \prod_{\vec{l}\neq\vec{k},-\vec{k}}{}'\frac{1}{u_{\vec{l}}}\exp\left(-\frac{v_{\vec{l}}}{u_{\vec{l}}}a^+_{\vec{l}}a^+_{-\vec{l}}\right)\frac{1}{u_{\vec{k}}}a_{\vec{k}}\sum_{n=1}^{+\infty}\frac{\left[-\frac{v_{\vec{k}}}{u_{\vec{k}}}a^+_{\vec{k}}a^+_{-\vec{k}}\right]^n}{n!}|0> \end{aligned}$$

where the sum starts at $n=1$ because the term $n=0$ gives 0 since $a_{\vec{k}}|0>=0$. Now,

$$[a_{\vec{k}},(a^+_{\vec{k}})^n]=n(a^+_{\vec{k}})^{n-1}$$

Let us prove this by recurrence. This is true for $n=1$ since $[a_{\vec{k}},a^+_{\vec{k}}]=1$. Let us suppose this is true for $n-1$. Then:

$$\begin{aligned} [a_{\vec{k}},(a^+_{\vec{k}})^n] &= [a_{\vec{k}},(a^+_{\vec{k}})^{n-1}]a^+_{\vec{k}}+(a^+_{\vec{k}})^{n-1}[a_{\vec{k}},a^+_{\vec{k}}] \\ &= (n-1)(a^+_{\vec{k}})^{n-2}a^+_{\vec{k}}+(a^+_{\vec{k}})^{n-1} \\ &= n(a^+_{\vec{k}})^{n-1} \end{aligned}$$

Thus,

$$a_{\vec{k}}(a^+_{\vec{k}})^n=n(a^+_{\vec{k}})^{n-1}+\underbrace{(a^+_{\vec{k}})^n a_{\vec{k}}}_{\text{gives 0 on the vacuum}}$$

and one gets:

$$\begin{aligned} a_{\vec{k}}|F> &= \prod_{\vec{l}\neq\vec{k},-\vec{k}}{}'\frac{1}{u_{\vec{l}}}\exp\left(-\frac{v_{\vec{l}}}{u_{\vec{l}}}a^+_{\vec{l}}a^+_{-\vec{l}}\right)\frac{1}{u_{\vec{k}}}\sum_{n=1}^{+\infty}\frac{(-\frac{v_{\vec{k}}}{u_{\vec{k}}})^n(a^+_{\vec{k}})^{n-1}(a^+_{-\vec{k}})^n}{(n-1)!}|0> \\ &= \prod_{\vec{l}\neq\vec{k},-\vec{k}}{}'\frac{1}{u_{\vec{l}}}\exp\left(-\frac{v_{\vec{l}}}{u_{\vec{l}}}a^+_{\vec{l}}a^+_{-\vec{l}}\right)\frac{1}{u_{\vec{k}}}\frac{-v_{\vec{k}}}{u_{\vec{k}}}a^+_{-\vec{k}}\exp\left(-\frac{v_{\vec{k}}}{u_{\vec{k}}}a^+_{\vec{k}}a^+_{-\vec{k}}\right)|0> \end{aligned}$$

Finally:

$$u_{\vec{k}} a_{\vec{k}} |F> = -v_{\vec{k}} a^+_{-\vec{k}} |F>$$

so that

$$\alpha_{\vec{k}} |F> = 0.$$

Finally, let us check the normalization of $|F>$.

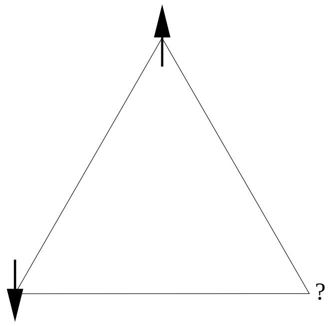
$$\begin{aligned} <F|F> &= <0| \prod_{\vec{k}}{}' \frac{1}{u^2_{\vec{k}}} \sum_n \left( -\frac{v_{\vec{k}}}{u_{\vec{k}}} \right)^{2n} \frac{(a_{-\vec{k}} a_{\vec{k}})^n (a^+_{\vec{k}} a^+_{-\vec{k}})^n}{n!^2} |0> \\ &= \prod_{\vec{k}}{}' \frac{1}{u^2_{\vec{k}}} \sum_n \left( -\frac{v_{\vec{k}}}{u_{\vec{k}}} \right)^{2n} \\ &= \prod_{\vec{k}}{}' \frac{1}{u^2_{\vec{k}}} \frac{1}{1 - (\frac{v_{\vec{k}}}{u_{\vec{k}}})^2} \\ &= 1. \end{aligned}$$

## 2.8 Frustrated magnets

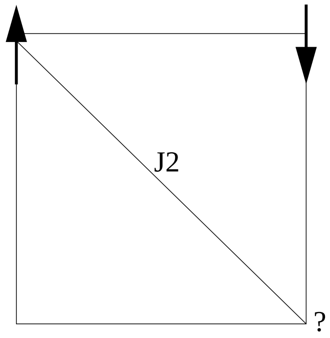
Definition: Frustrated magnets are magnets for which it is impossible to minimise simultaneously the energy of each bond in the classical case.

Examples:

- Triangular lattice → topological origin



- Square lattice with first and second neighbour interactions



### 2.8.1 Looking for the classical ground state

This is a problem that has no general solution. The solution is known however when the magnetic moments sit on a Bravais lattice (i.e. when there is one magnetic site per unit cell).

To see this, let us consider the energy of a configuration

$$E = \frac{1}{2} \sum_i \sum_{\vec{R}_n} J_{\vec{R}_n} \vec{S}_{\vec{R}_i} \cdot \vec{S}_{\vec{R}_i + \vec{R}_n}$$

There is a factor $\frac{1}{2}$ because each bond appears twice.

One must look for the minimum under the constraint $||\vec{S}_{\vec{R}_i}||^2 = 1$ for all sites.

Fourier transform:

$$\vec{S}_{\vec{k}} = \frac{1}{\sqrt{N}}\sum_i \vec{S}_i e^{-i\vec{k}\cdot\vec{R}_i} \Longrightarrow \vec{S}_i = \frac{1}{\sqrt{N}}\sum_{\vec{k}} \vec{S}_{\vec{k}} e^{i\vec{k}\cdot\vec{R}_i}$$

$$J_{\vec{k}} = \sum_{\vec{R}_n} J_{\vec{R}_n} e^{-i\vec{k}\cdot\vec{R}_n} \Longrightarrow J_{\vec{R}_n} = \frac{1}{N}\sum_{\vec{k}} J_{\vec{k}} e^{i\vec{k}\cdot\vec{R}_n}$$

$$E = \frac{1}{2}\sum_{\vec{R}_i}\sum_{\vec{R}_n}\frac{1}{N^2}\sum_{\vec{k}_1\vec{k}_2\vec{k}_3} J_{\vec{k}_1}\vec{S}_{\vec{k}_2}\cdot\vec{S}_{\vec{k}_3} e^{\vec{k}_1\cdot\vec{R}_n+\vec{k}_2\cdot\vec{R}_i+\vec{k}_3\cdot(\vec{R}_i+\vec{R}_n)}$$

$\Longrightarrow \vec{k}_2 = -\vec{k}_3 = \vec{k}_1$. The energy can be written as:

$$E = \frac{1}{2}\sum_{\vec{k}} J_{\vec{k}}\vec{S}_{\vec{k}}\cdot\vec{S}_{-\vec{k}}$$

On then proceeds in two steps:

- One solves the problem under the weaker constraint $\sum_i ||\vec{S}_i||^2 = N$.
- One checks if one can find a ground state that satisfies the stronger constraint $||\vec{S}_i||^2 = 1$ for all sites.

Now

$$\begin{aligned}\sum_i \vec{S}_i^2 &= \frac{1}{N}\sum_{\vec{k}_1\vec{k}_2}\sum_i e^{i(\vec{k}_1+\vec{k}_2)\cdot\vec{R}_i}\vec{S}_{\vec{k}_1}\cdot\vec{S}_{\vec{k}_1} \\ &= \sum_{\vec{k}}\vec{S}_{\vec{k}}\cdot\vec{S}_{-\vec{k}}\end{aligned}$$

One must thus minimise

$$E = \frac{1}{2}\sum_{\vec{k}} J_{\vec{k}}\vec{S}_{\vec{k}}\cdot\vec{S}_{-\vec{k}}$$

under the constraint $\sum_{\vec{k}}\vec{S}_{\vec{k}}\cdot\vec{S}_{-\vec{k}} = N = \text{cst}$.

Let's assume that $J_{\vec{k}}$ is minimal for $\vec{k} = \vec{k}_0$. The minimum of the energy is obtained iff $\vec{S}_{\vec{k}} = \vec{0}$, $\vec{k} \neq \vec{k}_0$, $-\vec{k}_0$ and $\vec{S}_{\vec{k}_0}\vec{S}_{-\vec{k}_0} + \vec{S}_{-\vec{k}_0}\vec{S}_{\vec{k}_0} = N$.

In real space, the spins are then given by

$$\vec{S}_i = \frac{1}{\sqrt{N}}\left(\vec{S}_{\vec{k}_0}e^{i\vec{k}_0\cdot\vec{R}} + \vec{S}_{-\vec{k}_0}e^{-i\vec{k}_0\cdot\vec{R}}\right)$$

Can one find $\vec{S}_{\vec{k}_0}$ and $\vec{S}_{-\vec{k}_0}$ such that:

- $\vec{S}_{\vec{k}_0} \cdot \vec{S}_{-\vec{k}_0} = \frac{N}{2}$

but also such that the local constraints:

- $\vec{S}_i$ réel
- $||\vec{S}_i||^2 = 1$

are satisfied for all $i$? Yes! For example:

$$\vec{S}_{\vec{k}_0} = \begin{pmatrix} \frac{\sqrt{N}}{2} \\ -i\frac{\sqrt{N}}{2} \\ 0 \end{pmatrix} \qquad \vec{S}_{-\vec{k}_0} = \begin{pmatrix} \frac{\sqrt{N}}{2} \\ i\frac{\sqrt{N}}{2} \\ 0 \end{pmatrix}$$

leads to:

$$\vec{S}_i = \begin{pmatrix} \cos\left(\vec{k}_0 \cdot \vec{R}\right) \\ \sin\left(\vec{k}_0 \cdot \vec{R}\right) \\ 0 \end{pmatrix}$$

One gest a helical structure of pitch vector $\vec{k}_0$.

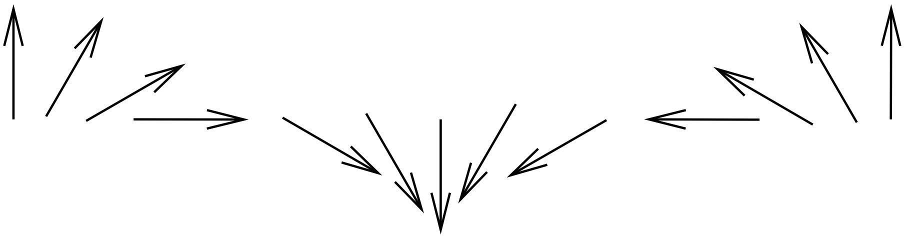


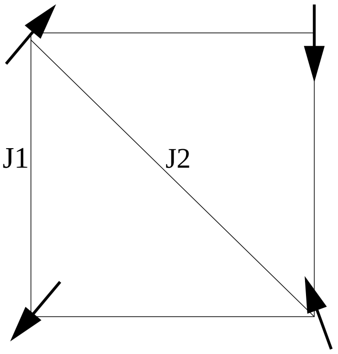
Figure 2.2: Example of a 1D helix of wave vector $\vec{k}_0 = \frac{2\pi}{12a}$.

Conclusions:

- The structure of the classical ground state is a helix whose wave vector is given by the minimum of $J_{\vec{k}}$.
- If there are several wave vectors that minimise $J_{\vec{k}}$, one can sometimes make linear combinations, and this can lead to a continuous degeneracy of the ground state.

Example: $J_1 - J_2$ model on the square lattice

$$\begin{aligned} H &= J_1 \sum_{<i,j>} \vec{S}_i \cdot \vec{S}_j + J_2 \sum_{<<i,j>>} \vec{S}_i \cdot \vec{S}_j \\ &= \sum_{\vec{k}} J(\vec{k}) \vec{S}_{\vec{k}} \cdot \vec{S}_{-\vec{k}} \end{aligned}$$



$$\begin{aligned} J(\vec{k}) &= J_1\left(e^{i\vec{k}\vec{x}}+e^{-i\vec{k}\vec{x}}+e^{i\vec{k}\vec{y}}+e^{-i\vec{k}\vec{y}}\right) \\ &+ J_2\left(e^{i\vec{k}(\vec{x}+\vec{y})}+e^{i\vec{k}(\vec{x}-\vec{y})}+e^{-i\vec{k}(\vec{x}+\vec{y})}+e^{i\vec{k}(-\vec{x}+\vec{y})}\right) \\ &= 2J_1\left(\cos(k_x)+\cos(k_y)\right)+2J_2\left(\cos(k_x+k_y)+\cos(k_x-k_y)\right) \\ &= 2J_1\left(\cos(k_x)+\cos(k_y)\right)+4J_2\cos(k_x)\cos(k_y) \end{aligned}$$

The minimization leads to:

$$\frac{\partial J}{\partial k_x} = -2J_1\sin(k_x)-4J_2\sin(k_x)\cos(k_y)=0$$
$$\frac{\partial J}{\partial k_y} = -2J_1\sin(k_y)-4J_2\sin(k_y)\cos(k_x)=0$$

$$\begin{pmatrix} k_x=0,\ k_x=\pi \\ k_y=0,\ k_y=\pi \end{pmatrix}$$

or

$$\cos k_y = \frac{-J_1}{2J_2} = \cos k_x$$

which is only possible if $\frac{J_1}{2J_2} < 1 \rightarrow \frac{J_1}{2} < J_2$.

$k_x = k_y = 0 \rightarrow J(\vec{k}) = 4J_1 + 4J_2$

$k_x = 0, k_y = \pi \rightarrow J(\vec{k}) = -4J_2$

$k_x = k_y = \pi \rightarrow J(\vec{k}) = -4J_1 + 4J_2$

$\cos k_x = \cos k_x = \frac{-J_1}{2J_2}$

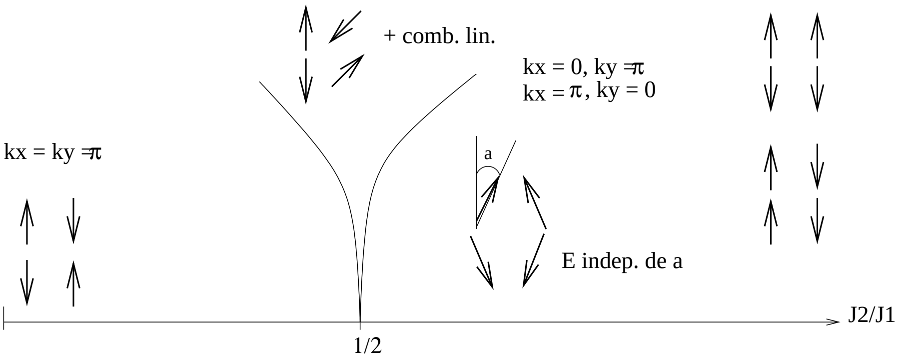
$$\begin{aligned} \rightarrow J(\vec{k}) &= 2J_1\left(\frac{-2J_1}{2J_2}\right)+4J_2\frac{J_1^2}{4J_2^2} \\ &= \frac{-2J_1^2}{J_2}+\frac{J_1^2}{J_2}=-\frac{J_1^2}{J_2} \end{aligned}$$

But if $J_1 < 2J_2$, $-4J_2 < \frac{-J_1^2}{J_2}$. The ground state is thus reached for $k_x = 0, k_y = \pi$ or $k_x = \pi, k_y = 0$.

For $\frac{J_2}{J_1} = \frac{1}{2}$, the minimum $J(\vec{k}) = -2J_1$ is reached for

$$k_x = \pi, \forall\ k_y \quad \text{ou} \quad k_y = \pi, \forall\ k_x$$

Finally, for $J_1 < 2J_2$, one can make a linear combination of two helices. This leads to a continuous family of degenerate classical ground states.



### 2.8.2 Spin waves in a helical system

Let us consider a system described by the Hamiltonian

$$\begin{aligned} H &= \frac{1}{2}\sum_i \sum_{\vec{R}_n} J_{\vec{R}_n} \vec{S}_i \cdot \vec{S}_{i+\vec{R}_n} \\ J(\vec{k}) &= \sum_{\vec{R}_n} J_{\vec{R}_n} e^{-i\vec{k}\cdot\vec{R}_n} \end{aligned}$$

and let us suppose that $J(\vec{k})$ is minimal for $\vec{k} = \vec{Q}$.

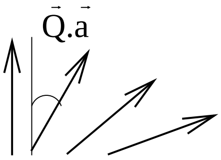
We are going to perform a rotation in spin space such that the classical ground state is described by $S_i^u = S_i^v = 0$, $S_i^w = 1$.

To achieve this, let us suppose that the ground state is described by

$$\vec{S}_i = \begin{pmatrix} 0 \\ \sin(\vec{Q}\cdot\vec{R}_i) \\ \cos(\vec{Q}\cdot\vec{R}_i) \end{pmatrix}$$

and let us perform a rotation of angle $\theta_i = \vec{Q}\cdot\vec{R}_i$ around $S^x$:

$S^u = S^x$, $S^v = \cos\theta_i S^y - \sin\theta_i S^z$ and $S^w = \sin\theta_i S^y + \cos\theta_i S^z$.



In the new basis, the Hamiltonian takes the form:

$$\begin{aligned} H &= \frac{1}{2}\sum_i \sum_{\vec{R}_n} J_{\vec{R}_n}[S_i^u S_{i+n}^u \\ &+ (\cos\theta_i S_i^v + \sin\theta_i S_i^w)(\cos\theta_{i+n} S_{i+n}^v + \sin\theta_{i+n} S_{i+n}^w) \\ &+ (-\sin\theta_i S_i^v + \cos\theta_i S_i^w)(-\sin\theta_{i+n} S_{i+n}^v + \cos\theta_{i+n} S_{i+n}^w)] \end{aligned}$$

since

$$\begin{aligned} S^y &= \cos\theta_i S_i^v + \sin\theta_i S_i^w \\ S^z &= -\sin\theta_i S_i^v + \cos\theta_i S_i^w \end{aligned}$$

$$\begin{aligned} H &= \frac{1}{2}\sum_i \sum_{\vec{R}_n} J_{\vec{R}_n}[S_i^u S_{i+n}^u \\ &+ S_i^v S_{i+n}^v(\cos\theta_i \cos\theta_{i+n} + \sin\theta_i \sin\theta_{i+n}) \\ &+ S_i^w S_{i+n}^w(\sin\theta_i \sin\theta_{i+n} + \cos\theta_i \cos\theta_{i+n}) \\ &+ S_i^v S_{i+n}^w(\cos\theta_i \sin\theta_{i+n} - \sin\theta_i \cos\theta_{i+n}) \\ &+ S_i^w S_{i+n}^v(-\cos\theta_i \sin\theta_{i+n} + \sin\theta_i \cos\theta_{i+n})] \\ &= \frac{1}{2}\sum_i \sum_{\vec{R}_n} J_{\vec{R}_n}\left[S_i^u S_{i+n}^u + \cos(\theta_i - \theta_{i+n})\left(S_i^v S_{i+n}^v + S_i^w S_{i+n}^w\right) + \sin(\theta_i - \theta_{i+n})\left(S_i^w S_{i+n}^v - S_i^v S_{i+}^w\right.\right. \end{aligned}$$

Next, let us perform the Holstein-Primakoff transformation:

$$\begin{aligned} S_i^w &= S - a_i^+ a_i \\ S_i^+ &\equiv S_i^u + iS_i^v \simeq \sqrt{2S} a_i \\ S_i^- &\equiv S_i^u - iS_i^v \simeq \sqrt{2S} a_i^+ \end{aligned}$$

$$\begin{aligned}
S_i^u S_{i+n}^u &= \frac{1}{4}\left(S_i^+ + S_i^-\right)\left(S_{i+n}^+ + S_{i+n}^-\right) \\
&= \frac{S}{2}\left(a_i a_{i+n} + a_i^+ a_{i+n}^+ + a_i a_{i+n}^+ + a_i^+ a_{i+n}\right) \\
S_i^v S_{i+n}^v &= -\frac{1}{4}\left(S_i^+ - S_i^-\right)\left(S_{i+n}^+ - S_{i+n}^-\right) \\
&= -\frac{S}{2}\left(a_i a_{i+n} + a_i^+ a_{i+n}^+ - a_i a_{i+n}^+ - a_i^+ a_{i+n}\right) \\
S_i^w S_{i+n}^w &= S^2 - S(a_i^+ a_i + a_{i+n}^+ a_{i+n}) \\
S_i^w S_{i+n}^v &= \frac{S\sqrt{2S}}{2i}\left(a_{i+n} + a_{i+n}^+\right) \\
S_i^v S_{i+n}^w &= \frac{S\sqrt{2S}}{2i}\left(a_i + a_i^+\right)
\end{aligned}$$

The Hamiltonian now takes the form:

$$\begin{aligned}
H &= \frac{1}{2}\sum_i \sum_{\vec{R}_n} J_{\vec{R}_n}[\frac{S}{2}(a_i a_{i+n} + a_i^+ a_{i+n}^+)(1 - \cos(\theta_i - \theta_{i+n})) \\
&+ \frac{S}{2}(a_i^+ a_{i+n} + a_{i+n}^+ a_i)(1 + \cos(\theta_i - \theta_{i+n})) \\
&+ [S^2 - S(a_i^+ a_i + a_{i+n}^+ a_{i+n})]\cos(\theta_i - \theta_{i+n}) \\
&+ \frac{S\sqrt{2S}}{2i}(a_i + a_i^+ + a_{i+n} + a_{i+n}^+)\sin(\theta_i - \theta_{i+n})]
\end{aligned}$$

$$J_{\vec{R}_n} = \frac{1}{N}\sum_{\vec{k}} J_{\vec{k}} e^{i\vec{k}\cdot\vec{R}_n} \quad , \quad a_i = \frac{1}{\sqrt{N}}\sum_{\vec{k}} e^{i\vec{k}\cdot\vec{r}_i} a_{\vec{k}} \quad , \quad a_i^+ = \frac{1}{\sqrt{N}}\sum_{\vec{k}} e^{-i\vec{k}\cdot\vec{r}_i} a_{\vec{k}}^+$$

*A priori* this Hamiltonian contains one-operator terms of the form $a_i$ et $a_i^+$. But the classical energy, which is given by:

$$E_{\text{class}} = \frac{1}{2}\sum_i \sum_{\vec{R}_n} J_{\vec{R}_n}\cos(\theta_i - \theta_{i+n})$$

must be minimal. If one groups the terms containing a given angle $\theta_i$, one gets

$$\sum_{\vec{R}_n} J_{\vec{R}_n}\cos(\theta_i - \theta_{i+n})$$

since each pair appears twice in the classical energy. Besides, the condition that the energy be minimal imposes

$$\begin{aligned}
\frac{\partial E_{\text{class}}}{\partial \theta_i} &= 0 \\
\Longrightarrow \sum_{\vec{R}_n} J_{\vec{R}_n}\sin(\theta_i - \theta_{i+n}) &= 0
\end{aligned}$$

As a consequence, the one-operator terms linear in $a_i$ et $a_i^+$ drop.

$$
\begin{aligned}
H &= \frac{1}{2}\sum_i \sum_{\vec{R}_n} \frac{1}{N^2} \sum_{\vec{k}_1,\vec{k}_2,\vec{k}_3} J_{\vec{k}_1} e^{i\vec{k}_1\cdot\vec{R}_n} \Bigg[ \frac{S}{2} \Big( a_{\vec{k}_2} a_{\vec{k}_3} e^{i\vec{k}_2\cdot\vec{r}_i + i\vec{k}_3\cdot(\vec{r}_i+\vec{R}_n)} \\
&+ a^+_{\vec{k}_2} a^+_{\vec{k}_3} e^{-i\vec{k}_2\cdot\vec{r}_i - i\vec{k}_3\cdot(\vec{r}_i+\vec{R}_n)} \Big) \left( 1 - \frac{e^{i\vec{Q}\cdot\vec{R}_n} + e^{-i\vec{Q}\cdot\vec{R}_n}}{2} \right) \\
&+ \frac{S}{2} \Big( a^+_{\vec{k}_2} a_{\vec{k}_3} e^{-i\vec{k}_2\cdot\vec{r}_i + i\vec{k}_3\cdot(\vec{r}_i+\vec{R}_n)} + a^+_{\vec{k}_3} a_{\vec{k}_2} e^{i\vec{k}_2\cdot\vec{r}_i - i\vec{k}_3\cdot(\vec{r}_i+\vec{R}_n)} \Big) \\
&\times \left( 1 + \frac{e^{i\vec{Q}\cdot\vec{R}_n} + e^{-i\vec{Q}\cdot\vec{R}_n}}{2} \right) \\
&+ S^2 \frac{e^{i\vec{Q}\cdot\vec{R}_n} + e^{-i\vec{Q}\cdot\vec{R}_n}}{2} \\
&- S \Big( a^+_{\vec{k}_2} a_{\vec{k}_3} e^{-i\vec{k}_2\cdot\vec{r}_i + i\vec{k}_3\cdot\vec{r}_i} + a^+_{\vec{k}_2} a_{\vec{k}_3} e^{-i\vec{k}_2\cdot(\vec{r}_i+\vec{R}_n) + i\vec{k}_3\cdot(\vec{r}_i+\vec{R}_n)} \Big) \\
&\times \left( \frac{e^{i\vec{Q}\cdot\vec{R}_n} + e^{-i\vec{Q}\cdot\vec{R}_n}}{2} \right) \Bigg]
\end{aligned}
$$

Let us look at the terms:

- $e^{i\vec{R}_n(\vec{k}_1+\vec{k}_3)} e^{i\vec{r}_i(\vec{k}_2+\vec{k}_3)}$ $\qquad \vec{k}_1 = -\vec{k}_3 \ , \ \vec{k}_2 = -\vec{k}_3$

$$\rightarrow \frac{1}{2}\sum_{\vec{k}} J_{\vec{k}} \frac{S}{2} a_{\vec{k}} a_{-\vec{k}}$$

- $e^{i\vec{R}_n(\vec{k}_1+\vec{k}_3+\vec{Q})} e^{i\vec{r}_i(\vec{k}_2+\vec{k}_3)}$ $\qquad \vec{k}_3 = -\vec{k}_1 - \vec{Q} \ , \ \vec{k}_2 = \vec{k}_1 + \vec{Q}$

$$\rightarrow -\frac{1}{4}\sum_{\vec{k}} J_{\vec{k}-\vec{Q}} \frac{S}{2} a_{\vec{k}} a_{-\vec{k}}$$

- $e^{i\vec{R}_n(\vec{k}_1+\vec{k}_3-\vec{Q})} e^{i\vec{r}_i(\vec{k}_2+\vec{k}_3)}$

$$\rightarrow -\frac{1}{4}\sum_{\vec{k}} J_{\vec{k}+\vec{Q}} \frac{S}{2} a_{\vec{k}} a_{-\vec{k}}$$

- $e^{i\vec{R}_n(\vec{k}_1-\vec{k}_3)} e^{-i\vec{r}_i(\vec{k}_2+\vec{k}_3)}$

$$\rightarrow \frac{1}{2}\sum_{\vec{k}} J_{\vec{k}} \frac{S}{2} a^+_{\vec{k}} a^+_{-\vec{k}}$$

- $e^{i\vec{R}_n(\vec{k}_1-\vec{k}_3+\vec{Q})}e^{-i\vec{r}_i(\vec{k}_2+\vec{k}_3)}$

$$\rightarrow -\frac{1}{4}\sum_{\vec{k}} J_{\vec{k}-\vec{Q}}\frac{S}{2}a^+_{\vec{k}}a^+_{-\vec{k}}$$

- $e^{i\vec{R}_n(\vec{k}_1-\vec{k}_3-\vec{Q})}e^{-i\vec{r}_i(\vec{k}_2+\vec{k}_3)}$

$$\rightarrow -\frac{1}{4}\sum_{\vec{k}} J_{\vec{k}+\vec{Q}}\frac{S}{2}a^+_{\vec{k}}a^+_{-\vec{k}}$$

- $e^{i\vec{R}_n(\vec{k}_1+\vec{k}_3)}e^{i\vec{r}_i(-\vec{k}_2+\vec{k}_3)}$

$$\rightarrow \quad \frac{1}{2}\sum_{\vec{k}} J_{-\vec{k}}\frac{S}{2}a^+_{\vec{k}}a_{\vec{k}}$$

- and all the other terms:

$$\begin{aligned}
&\rightarrow \quad \frac{1}{4}\sum_{\vec{k}} J_{-\vec{k}-\vec{Q}}\frac{S}{2}a^+_{\vec{k}}a_{\vec{k}}\\
&\rightarrow \quad \frac{1}{4}\sum_{\vec{k}} J_{-\vec{k}+\vec{Q}}\frac{S}{2}a^+_{\vec{k}}a_{\vec{k}}\\
&\rightarrow \quad \frac{1}{2}\sum_{\vec{k}} J_{\vec{k}}\frac{S}{2}a^+_{\vec{k}}a^+_{\vec{k}}\\
&\rightarrow \quad \frac{1}{4}\sum_{\vec{k}} J_{\vec{k}+\vec{Q}}\frac{S}{2}a^+_{\vec{k}}a_{\vec{k}}\\
&\rightarrow \quad \frac{1}{4}\sum_{\vec{k}} J_{\vec{k}-\vec{Q}}\frac{S}{2}a^+_{\vec{k}}a_{\vec{k}}\\
&\rightarrow \quad \frac{1}{2}\sum_{i}\sum_{\vec{R}_n} S^2 J_{\vec{R}_n}\frac{e^{i\vec{Q}\vec{R}_n}+e^{-i\vec{Q}\vec{R}_n}}{2} = \frac{S^2}{2}N\frac{1}{2}\left(J_{\vec{Q}}+J_{-\vec{Q}}\right)\\
&\rightarrow \quad -\frac{S}{4}J_{\vec{Q}}a^+_{\vec{k}}a_{\vec{k}}\\
&\rightarrow \quad -\frac{S}{4}J_{-\vec{Q}}a^+_{\vec{k}}a_{\vec{k}}\\
&\rightarrow \quad -\frac{S}{4}J_{-\vec{Q}}a^+_{\vec{k}}a_{\vec{k}}\\
&\rightarrow \quad -\frac{S}{4}J_{\vec{Q}}a^+_{\vec{k}}a_{\vec{k}}
\end{aligned}$$

Let us group the terms:

$$\begin{aligned}
&\left(a^+_{\vec{k}}a^+_{-\vec{k}} + a_{\vec{k}}a_{-\vec{k}}\right)\frac{S}{4}\sum_{\vec{k}}\left(J_{\vec{k}} - \frac{1}{2}\left(J_{\vec{k}+\vec{Q}} + J_{\vec{k}-\vec{Q}}\right)\right) \\
&+a^+_{\vec{k}}a_{\vec{k}}\left[\frac{S}{4}\sum_{\vec{k}}\left(J_{-\vec{k}} + \frac{1}{2}\left(J_{-\vec{k}+\vec{Q}} + J_{-\vec{k}-\vec{Q}}\right) + J_{\vec{k}}\right.\right. \\
&\qquad\left.\left.+\frac{1}{2}\left(J_{\vec{k}+\vec{Q}} + J_{\vec{k}-\vec{Q}}\right)\right) - \frac{S}{2}(J_{\vec{Q}} + J_{-\vec{Q}})\right] \\
&\qquad\qquad+\frac{S^2}{2}N\frac{1}{2}\left(J_{\vec{Q}} + J_{-\vec{Q}}\right)
\end{aligned}$$

or, since $J_{\vec{k}} = J_{-\vec{k}}$:

$$\begin{aligned}
H &= \frac{NS^2}{2}J_{\vec{Q}} + \frac{S}{2}\sum_{\vec{k}}\left[J_{\vec{k}} + \frac{1}{2}\left(J_{\vec{k}+\vec{Q}} + J_{\vec{k}-\vec{Q}}\right) - 2J_{\vec{Q}}\right]a^+_{\vec{k}}a_{\vec{k}} \\
&+ \frac{S}{4}\sum_{\vec{k}}\left[J_{\vec{k}} - \frac{1}{2}\left(J_{\vec{k}+\vec{Q}} + J_{\vec{k}-\vec{Q}}\right)\right]\left(a_{\vec{k}}a_{-\vec{k}} + a^+_{\vec{k}}a^+_{-\vec{k}}\right)
\end{aligned}$$

This Hamiltonian has the same for as in the antiferromagnetic case, and a Bogoliubov transformation leads to:

$$\begin{aligned}
\omega_{\vec{k}} &= S\sqrt{\left(J_{\vec{k}} - J_{\vec{Q}}\right)\left(\frac{1}{2}\left(J_{\vec{k}+\vec{Q}} + J_{\vec{k}-\vec{Q}}\right) - J_{\vec{Q}}\right)} \\
\omega_{\vec{k}} &= 0 \quad \text{si} \quad \begin{cases} \vec{k} = 0 \\ \vec{k} = \pm\vec{Q} \end{cases}
\end{aligned}$$

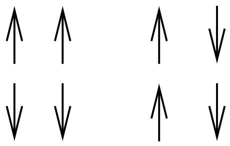
### 2.8.3 Complement: $J_1 - J_2$ model

Let us describe some effects of quantum fluctuations in the case of the $J_1 - J_2$ model.

- $J_2 > \frac{J_1}{2}$: lifting of the classical degeneracy by zero-point fluctuations.

$$\begin{aligned}
E &= \text{cst} + \sum_{\vec{k}}\left(\alpha^+_{\vec{k}}\alpha_{\vec{k}} + \frac{1}{2}\right)\omega_{\vec{k}}(\theta) \\
\rightarrow E_0 &= \text{cst} + \sum_{\vec{k}}\omega_{\vec{k}}(0)
\end{aligned}$$

The collinear structures are selected:



- $J_2 < \frac{J_1}{2}$:

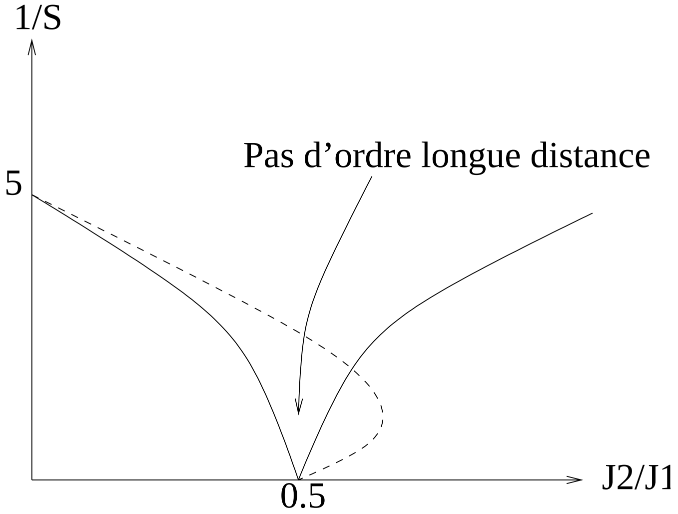
$$< m >= S - \underbrace{\int \frac{d\vec{k} f(\vec{k})}{\omega_{\vec{k}}}}_{\Delta<m>}$$

$f(\vec{k}) \rightarrow$ cst when $\vec{k} \rightarrow \vec{0}$.

The correction will thus be large if the spin-wave velocity is small. But

$$\begin{aligned} \omega_{\vec{k}} &= \left\{ \left[ SJ_1 z + SJ_2 z(\beta_{\vec{k}} - 1) \right]^2 - \left( SJ_1 z \gamma_{\vec{k}} \right)^2 \right\}^{\frac{1}{2}} \\ \gamma_{\vec{k}} &= \frac{1}{z} \sum_{\vec{\tau}_1} e^{i\vec{k}\cdot\vec{\tau}_1} \quad , \quad \beta_{\vec{k}} = \frac{1}{z} \sum_{\vec{\tau}_2} e^{i\vec{k}\cdot\vec{\tau}_2} \end{aligned}$$

Expanding at small $k$ leads to $\omega_{\vec{k}} \sim ck$ with $c \rightarrow 0$ when $J_2 \rightarrow \frac{J_1}{2}$.
$\longrightarrow$ the correction diverges:



Is this true for large $S$? Probably not. The correction is infinite for $S$ infinite, but for $S < +\infty$, one must take the higher order corrections into account. It turns out that $\Delta < m >$ diverges as $\ln S$:

$\rightarrow \; < m >= S - cst \times \ln S > 0$ when $S \rightarrow +\infty$

What happens for $S = \frac{1}{2}$? Another kind of ground state, a product of dimer singlets, has been suggested, but this is still an open problem.

Simplest example: $1D$, $J_1 - J_2$ model with $J_1 = 2J_2$.

Before we study this problem for $S = 1/2$, let us look at the 1D classical $J_1 - J_2$ model:

$$\begin{aligned} J(k) &= 2J_1 \cos k + 2J_2 \cos 2k \\ \frac{\partial J}{\partial k} &= -2J_1 \sin k - 4J_2 \sin 2k \\ &= -2 \sin k \, (J_1 + 4J_2 \cos k) \end{aligned}$$

1. $k = 0 \quad \Rightarrow \quad J(k) = 2J_1 + 2J_2$

2. $k = \pi \quad \Rightarrow \quad J(k) = -2J_1 + 2J_2$

3. $\cos k = \frac{-J_1}{4J_2}$ (si $J_2 \geq \frac{J_1}{4}$)

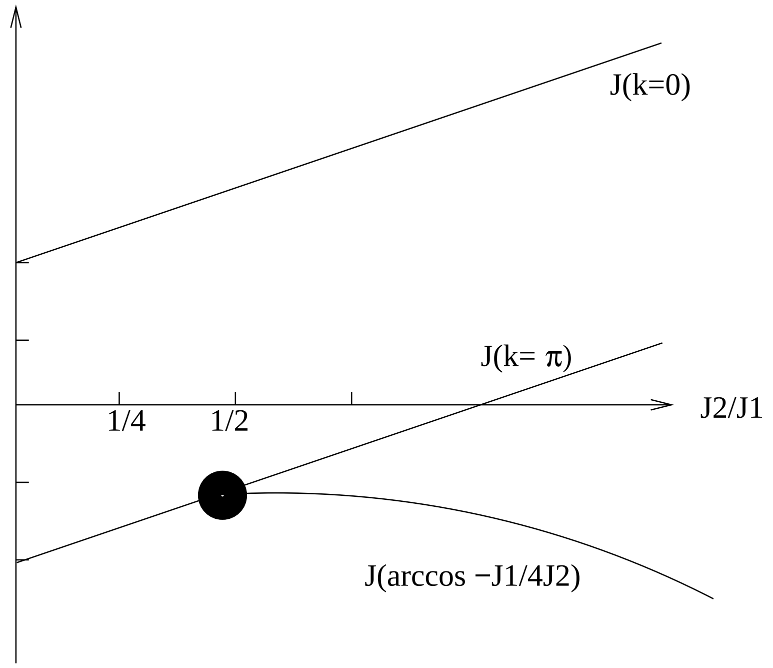
$$\begin{aligned} \Rightarrow J(k) &= \frac{-2J_1^2}{4J_2} + 2J_2 \left( 2\frac{J_1^2}{16J_2^2} - 1 \right) \\ &= -\frac{1}{4}\frac{J_1^2}{J_2} - 2J_2 \end{aligned}$$

$$\begin{aligned} & J \left( \arccos \frac{-J_1}{4J_2} \right) \leq J(\pi) \\ \Leftrightarrow \; & -\frac{1}{4}\frac{J_1^2}{J_2} - 2J_2 \leq -2J_1 + 2J_2 \\ \Leftrightarrow \; & -\frac{1}{4}\frac{J_1^2}{J_2} + 2J_1 \leq 4J_2 \\ \Leftrightarrow \; & 0 \leq J_1^2 - 8J_1J_2 + 16J_2^2 \\ \Leftrightarrow \; & 0 \leq (4J_2 - J_1)^2 \qquad \text{OK} \end{aligned}$$

Let us come back to the quantum case and consider the model:

$$H = J \sum_i \left( 2\vec{S}_i \cdot \vec{S}_{i+1} + \vec{S}_i \cdot \vec{S}_{i+2} \right) \quad , \quad N \;\; \text{sites}$$

with $(J_1 = 2J_2,\, J_2 = J)$ and $(S = \frac{1}{2})$.

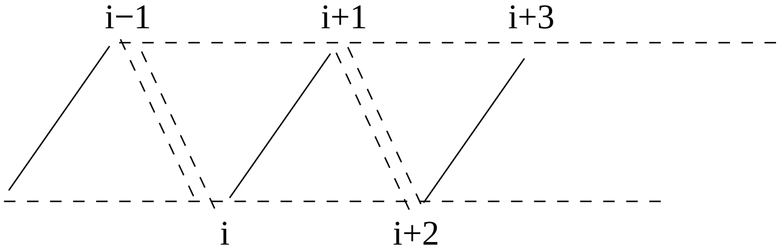


The Hamiltonian can be rewritten as:

$$H = J\left(2\vec{S}_1 \cdot \vec{S}_2 + \left(\vec{S}_1 + \vec{S}_2\right) \cdot \vec{S}_3 + \vec{S}_2 \cdot \left(\vec{S}_3 + \vec{S}_4\right) + \cdots\right)$$

or

$$H = J \sum_{i\ odd} \left[2\vec{S}_i \cdot \vec{S}_{i+1} + \left(\vec{S}_i + \vec{S}_{i+1}\right) \cdot \vec{S}_{i+2} + \vec{S}_{i+1} \cdot \left(\vec{S}_{i+2} + \vec{S}_{i+3}\right)\right]$$



Let us consider the state $|\phi>= [1,2] \otimes [3,4] \otimes \cdots$ where $[j,k] = \frac{1}{\sqrt{2}}(|\uparrow_j \downarrow_k> - |\downarrow_j \uparrow_k>)$ is the singlet built out of the pair of spins $(j,k)$.

What is $\left(\vec{S}_1 \cdot \vec{S}_2\right)[1,2]$ equal to?

$$\vec{S}_1 \cdot \vec{S}_2 = \frac{1}{2}\left[\left(\vec{S}_1 + \vec{S}_2\right)^2 - \vec{S}_1^2 - \vec{S}_2^2\right] = \frac{1}{2}\left[\left(\vec{S}_1 + \vec{S}_2\right)^2 - 2 \times \frac{3}{4}\right]$$

and $\left(\vec{S}_1+\vec{S}_2\right)^2[1,2]=0$ since $[1,2]$ is the singlet.

One thus has $\left(\vec{S}_1\cdot\vec{S}_2\right)[1,2]=-\frac{3}{4}[1,2]$.

Calculation of $\left(\vec{S}_1+\vec{S}_2\right)\cdot\vec{S}_3[1,2]$:

$$\left(\vec{S}_1+\vec{S}_2\right)\cdot\vec{S}_3 \;=\; \left(S_1^z+S_2^z\right)S_3^z+\frac{1}{2}\left(S_1^++S_2^+\right)S_3^-+\frac{1}{2}\left(S_1^-+S_2^-\right)S_3^+,$$

and the state is given by $[1,2]\sigma_3=\frac{1}{\sqrt{2}}\left(|\uparrow\downarrow>-|\downarrow\uparrow>\right)\otimes|\sigma_3>$.

Let us then look at the action of each of the three terms:

$$\begin{aligned}
\left(S_1^z+S_2^z\right)S_3^z &= \frac{1}{\sqrt{2}}\left(-\frac{1}{4}\sigma_3+\frac{1}{4}\sigma_3\right)[1,2]\sigma_3=0\\
\frac{1}{2}\left(S_1^++S_2^+\right)S_3^- &= \begin{cases}\frac{1}{\sqrt{2}}\left(|\uparrow\uparrow>-|\uparrow\uparrow>\right)|\downarrow>=0 & \text{si } \sigma_3=\frac{1}{2}\\ 0 \quad \text{otherwise}\end{cases}\\
\frac{1}{2}\left(S_1^++S_2^+\right)S_3^- &= \begin{cases}0 \quad \text{if } \sigma_3=\frac{1}{2}\\ \frac{1}{\sqrt{2}}\left(|\downarrow\downarrow>-|\downarrow\downarrow>\right)|\uparrow>=0 & \text{si } \sigma_3=-\frac{1}{2}\end{cases}
\end{aligned}$$

One is left with $\left(\vec{S}_1+\vec{S}_2\right)\cdot\vec{S}_3[1,2]=0$.

Finally, acting with the Hamiltonian $H$ on the state $|\phi>$ gives

$$\boxed{H|\phi>=\left(\frac{N}{2}\times 2J\times\frac{-3}{4}\right)|\phi>=-\frac{3}{4}JN|\phi>}$$

So, $|\phi>$ is an eigenstate of energy $-\frac{3}{4}JN$.

Let us now try to show that this is the ground state. One can rewrite $H$ as: $H=\sum_i h_i$ with $h_i=J\left(\vec{S}_i\vec{S}_{i+1}+\vec{S}_{i+1}\vec{S}_{i+2}+\vec{S}_i\vec{S}_{i+2}\right)$. But

$$\begin{aligned}
h_i &= \frac{J}{2}\left[\left(\vec{S}_i+\vec{S}_{i+1}+\vec{S}_{i+2}\right)^2-\vec{S}_i^2-\vec{S}_{i+1}^2-\vec{S}_{i+2}^2\right]\\
&= \frac{J}{2}\left[\left(\vec{S}_i+\vec{S}_{i+1}+\vec{S}_{i+2}\right)^2-3\times\frac{3}{4}\right]
\end{aligned}$$

The sum of 3 spins $\frac{1}{2}$ is a spin $\frac{1}{2}$ or a spin $\frac{3}{2}$.

$$\Longrightarrow\left(\vec{S}_i+\vec{S}_{i+1}+\vec{S}_{i+2}\right)^2=\begin{cases}\frac{1}{2}\left(\frac{1}{2}+1\right)=\frac{3}{4}\\ \text{ou}\frac{3}{2}\left(\frac{3}{2}+1\right)=\frac{15}{4}\end{cases} \tag{2.6}$$

The smallest eigenvalue of $h_i$ is $-\frac{3}{4}J$. Now

$$\begin{aligned}
H=\sum_i h_i \quad &\Longrightarrow\quad E_{\text{fond}}(H)\geq\sum_i E_{\text{fond}}h_i\\
&\Longrightarrow\quad E_{\text{fond}}(H)\geq N\times\frac{-3}{4}J
\end{aligned}$$

since $E(\phi) = -\frac{3}{4}NJ$.

So, $|\phi>$ is indeed the ground state.

# Chapter 3

# Electrons in a magnetic field

This chapter is devoted to the properties of itinerant electrons in a magnetic field. This is a very vast subject that one cannot fully cover in a few hours. We will concentrate on three remarkable effects: the de Haas-van Alphen effect, the integer Quantum Hall Effect, and the fractional Quantum Hall Effect. All these effects rely in an essential way on the quantification of the motion of a charged particle in the presence of a uniform magnetic field $\vec{B} = B\hat{z}$.

## 3.1 Landau levels

### 3.1.1 Classical problem

In the presence of a magnetic field, a particle of charge $q$ is subject to the Lorentz force

$$\vec{F} = \frac{q}{c}(\vec{v} \wedge \vec{B}) = \frac{q}{c}B \begin{pmatrix} v_y \\ -v_x \\ 0 \end{pmatrix} \tag{3.1}$$

with $v_x = \dot{x}$ and $v_y = \dot{y}$, so that

$$\begin{cases} m\ddot{x} = \dfrac{qB}{c}\dot{y} \\ m\ddot{y} = -\dfrac{qB}{c}\dot{x} \\ m\ddot{z} = 0 \end{cases} \tag{3.2}$$

In the $z$ direction, this is the motion of a free particle.

In the plane perpendicular to the field, one gets:

$$\begin{cases} \ddot{x} = \omega_c \dot{y} \\ \ddot{y} = -\omega_c \dot{x} \end{cases} \tag{3.3}$$

with $\omega_c = \frac{qB}{mc}$. These differential equations can be integrated:

$$\begin{cases} \dot{x} = \omega_c (y - y_0) \\ \dot{y} = -\omega_c (x - x_0) \end{cases} \tag{3.4}$$

leading to

$$\begin{cases} \ddot{x} = -\omega_c^2 (x - x_0) \\ \ddot{y} = -\omega_c^2 (y - y_0) \end{cases} \tag{3.5}$$

Let us define $\eta_x = x - x_0$ and $\eta_y = y - y_0$. We get:

$$\begin{cases} \ddot{\eta}_x = -\omega_c^2 \eta_x \\ \ddot{\eta}_y = -\omega_c^2 \eta_y \end{cases} \Rightarrow \begin{cases} \eta_x = r_x \sin(\omega_c t + \varphi_x) \\ \eta_y = r_y \cos(\omega_c t + \varphi_y) \end{cases} \tag{3.6}$$

But $\dot{\eta}_x = r_x \omega_c \cos(\omega_c t + \varphi_x)$ must be equal to $\omega_c \eta_y$, so

$$\begin{cases} r_x = r_y = r \\ \varphi_x = \varphi_y = \varphi \end{cases} \tag{3.7}$$

$$\Rightarrow \begin{cases} x = x_0 + r \sin(\omega_c t + \varphi) \\ y = y_0 + r \cos(\omega_c t + \varphi) \end{cases} \tag{3.8}$$

The motion of the particle projected on the plane corresponds to a uniform motion along a circle of center $(x_0, y_0)$ and of radius $r$ at a frequency $\omega_c = \frac{qB}{mc}$.

### 3.1.2 Lagrangian approach

The equations of motion can be rederived from the Lagrangian

$$\mathcal{L} = \mathcal{L}_{xy} + \tfrac{1}{2} m \dot{z}^2, \tag{3.9}$$

with

$$\mathcal{L}_{xy}(x, y; \dot{x}, \dot{y}) = \frac{1}{2} m(\dot{x}^2 + \dot{y}^2) + \frac{q}{c} \left[ A_x \dot{x} + A_y \dot{y} \right], \tag{3.10}$$

where $A_x(x, y)$ and $A_y(x, y)$ are the components of a vector $\vec{A}$ such that $\vec{B} = \mathrm{rot}\vec{A}$, or $B = \partial_x A_y - \partial_y A_x$. Indeed the Euler-Lagrange equations

$$\frac{\mathrm{d}}{\mathrm{d}t} \frac{\partial \mathcal{L}}{\partial \dot{q}_i} - \frac{\partial \mathcal{L}}{\partial q_i} = 0, \tag{3.11}$$

lead to:

$$\Rightarrow \begin{cases} \dfrac{\mathrm{d}}{\mathrm{d}t}\left(m\dot{x}+\dfrac{q}{c}A_x\right)-\dfrac{q}{c}\left[\dot{x}\partial_x A_x+\dot{y}\partial_x A_y\right]=0, \\ \dfrac{\mathrm{d}}{\mathrm{d}t}\left(m\dot{y}+\dfrac{q}{c}A_y\right)-\dfrac{q}{c}\left[\dot{x}\partial_y A_x+\dot{y}\partial_y A_y\right]=0. \end{cases} \tag{3.12}$$

$$\Rightarrow \begin{cases} m\ddot{x}+\dfrac{q}{c}\partial_t A_x+\dfrac{q}{c}\left(\dot{x}\partial_x A_x+\dot{y}\partial_y A_x\right)-\dfrac{q}{c}\left(\dot{x}\partial_x A_x+\dot{y}\partial_x A_y\right)=0 \\ m\ddot{y}+\dfrac{q}{c}\partial_t A_y+\dfrac{q}{c}\left(\dot{x}\partial_x A_y+\dot{y}\partial_y A_y\right)-\dfrac{q}{c}\left(\dot{x}\partial_y A_x+\dot{y}\partial_y A_y\right)=0 \end{cases} \tag{3.13}$$

If the magnetic does not depend on time, $\partial_t A_x = \partial_t A_y = 0$.

$$\Rightarrow \begin{cases} m\ddot{x}=\dfrac{q}{c}\dot{y}\left(\partial_x A_y-\partial_y A_x\right) \\ m\ddot{y}=\dfrac{q}{c}\dot{x}\left(\partial_y A_x-\partial_x A_y\right), \end{cases} \tag{3.14}$$

or

$$\Rightarrow \begin{cases} \ddot{x}=\dfrac{qB}{mc}\dot{y} \\ \ddot{y}=\dfrac{qB}{mc}\dot{x}. \end{cases} \tag{3.15}$$

### 3.1.3 Hamiltonian formulation

One must express the velocities in terms of the momenta.

$$\begin{cases} p_x=\dfrac{\partial\mathcal{L}}{\partial\dot{x}}=m\dot{x}+\dfrac{q}{c}A_x \\ p_y=\dfrac{\partial\mathcal{L}}{\partial\dot{y}}=m\dot{y}+\dfrac{q}{c}A_y \end{cases} \quad \Rightarrow \quad \begin{cases} \dot{x}=\dfrac{1}{m}\left(p_x-\dfrac{q}{c}A_x\right) \\ \dot{y}=\dfrac{1}{m}\left(p_y-\dfrac{q}{c}A_y\right) \end{cases} \tag{3.16}$$

$$\Rightarrow \mathcal{H}_{xy}=\frac{1}{2}m\left(\dot{x}^2+\dot{y}^2\right)=\frac{1}{2m}\left[\left(p_x-\frac{q}{c}A_x\right)^2+\left(p_y-\frac{q}{c}A_y\right)^2\right] \tag{3.17}$$

Finally,

$$\boxed{\mathcal{H}=\frac{1}{2m}\left(\vec{p}-\frac{q}{c}\vec{A}\right)^2.} \tag{3.18}$$

### 3.1.4 Quantum treatment

Let us start from the Hamiltonian formulation, and let us suppose that $\hat{x}, \hat{y}, \hat{p}_x, \hat{p}_y, ...$ are operators satisfying the commutation relations

$$[\hat{x}_i, \hat{p}_j] = i\hbar\delta_{ij}. \tag{3.19}$$

$$\to \mathcal{H} = \frac{1}{2m}\left(p_x - \frac{q}{c}A_x\right)^2 + \frac{1}{2m}\left(p_y - \frac{q}{c}A_y\right)^2 + \frac{1}{2m}p_z^2. \tag{3.20}$$

Let us define $\Pi_x = p_x - \frac{q}{c}A_x$ and $\Pi_y = p_y - \frac{q}{c}A_y$. Since $A_x$ and $A_y$ do not depend on $z$,

$$[p_z, \Pi_x] = [p_z, \Pi_y] = 0. \tag{3.21}$$

Let us calculate the commutator $[\Pi_x, \Pi_y]$,

$$[\Pi_x, \Pi_y] = \left[p_x - \frac{q}{c}A_x, p_y - \frac{q}{c}A_y\right] = -\frac{q}{c}\left[p_x, A_y\right] - \frac{q}{c}\left[A_x, p_y\right]. \tag{3.22}$$

But

$$[p_x, f(x)]\varphi(x) = -i\hbar\partial_x\left(f(x)\varphi(x)\right) + f(x)i\hbar\partial_x\varphi(x) = -i\hbar(\partial_x f(x))\varphi(x), \tag{3.23}$$

$$\to [p_x, f(x)] = -i\hbar\partial_x f(x). \tag{3.24}$$

So

$$[p_x, A_y] = -i\hbar\partial_x A_y, \quad [A_x, p_y] = +i\hbar\partial_y A_x, \tag{3.25}$$

leading to

$$[\Pi_x, \Pi_y] = -\frac{q}{c}(-i\hbar)\left(\partial_x A_y - \partial_y A_x\right), \tag{3.26}$$

or

$$\boxed{[\Pi_x, \Pi_y] = i\hbar\frac{qB}{c}.} \tag{3.27}$$

Let us define $\hat{P} = \hat{\Pi}_x$, $\hat{Q} = -\frac{c}{Bq}\Pi_y$. We have $[\hat{Q}, \hat{P}] = i\hbar$ and

$$\mathcal{H} = \frac{1}{2m}\hat{P}^2 + \frac{1}{2}m\omega^2\hat{Q}^2 + \frac{1}{2m}p_z^2, \tag{3.28}$$

with

$$\frac{1}{2m}\Pi_y^2 = \frac{1}{2}m\omega^2\frac{c^2}{q^2B^2}\Pi_y^2, \tag{3.29}$$

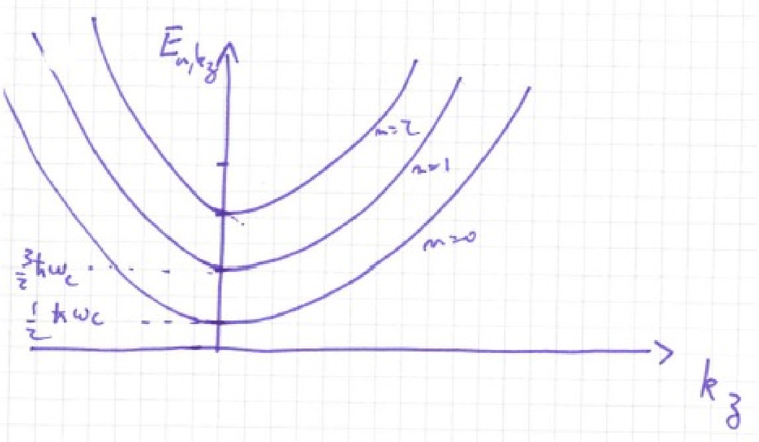
$$\Rightarrow \omega^2 = \frac{q^2B^2}{m^2c^2}, \text{ soit } \omega = \frac{qB}{mc}. \tag{3.30}$$

With these variables, the Hamiltonian reads:

$$\boxed{\mathcal{H} = \frac{1}{2m}\hat{P}^2 + \frac{1}{2}m\omega_c^2\hat{Q}^2 + \frac{1}{2m}p_z^2.} \tag{3.31}$$

The Hamiltonian can thus be written as the sum of a harmonic oscillator of frequency $\omega_c$ and of the Hamiltonian of a free particle in the $z$ direction. The eigen energies are thus of the form:

$$\boxed{E_{n,k_z} = \hbar\omega_c(n + \tfrac{1}{2}) + \frac{\hbar^2k_z^2}{2m}} \tag{3.32}$$



Each of these curves corresponds to a Landau level.

## Degeneracy

For a given value of $k_z$, how many levels are there per level? To get a first idea, let us consider the number $N$ of quantum states whose energy is between $E$ et $E+\hbar\omega_c$. To calculate this number, we write that it is given by the volume in $\vec{k}$ space of the points such that

$$E \leq \frac{\hbar^2 k^2}{2m} \leq E + \hbar\omega_c \tag{3.33}$$

divided by the elementary volume per $\vec{k}$ point. In 2D, in a box of dimension $L_x \times L_y$, this elementary volume is given by $\frac{(2\pi)^2}{L_x L_y}$. The volume of $\vec{k}$ points satisfying (3.33) is a surface in 2D given by:

$$2\pi \int_{k_1}^{k_2} k\mathrm{d}k = 2\pi \frac{k_2^2 - k_1^2}{2}, \tag{3.34}$$

with

$$E = \frac{\hbar^2 k_1^2}{2m}, \quad E + \hbar\omega_c = \frac{\hbar^2 k_2^2}{2m}, \tag{3.35}$$

$$\rightarrow 2\pi \times \frac{1}{2} \times \frac{2m}{\hbar^2}\hbar\omega_c = \frac{2\pi}{\hbar}\omega_c m. \tag{3.36}$$

The number of quantum states we are looking for is thus given by:

$$N = \frac{2\pi}{\hbar}\omega_c m \frac{L_x L_y}{2\pi} = L_x L_y \frac{qB}{2\pi\hbar c}. \tag{3.37}$$

This number is independent of $E$ (the density of states is a constant in 2D). So this a good candidate for the degeneracy of a Landau level.

To confirm this result more precisely, one should calculate the exact spectrum of a system with periodic boundary conditions, a problem which unfortunately does not have a simple analytical solution.

Nevertheless, one can convince oneself that this result is correct by going back to the solution of the problem. The motion in the $x, y$ plane is described by 4 operators $\hat{x}, \hat{p}_x, \hat{y}, \hat{p}_y$. Now, we have shown that the Hamiltonian only involves two operators $\Pi_x = p_x - \frac{q}{c}A_x$, $\Pi_y = p_y - \frac{q}{c}A_y$. So there must be two degrees of freedom that do not enter the Hamiltonian. From the classical solution, these degrees of freeedom are $y_0 = y - \frac{1}{\omega_c}\dot{x}$ and $x_0 = x + \frac{1}{\omega_c}\dot{y}$, or, in the Hamiltonian formulation,

$$\begin{cases} Y = y - \dfrac{1}{m\omega_c}\Pi_x \\ X = x + \dfrac{1}{m\omega_c}\Pi_y \end{cases} \tag{3.38}$$

These operators commute with the Hamiltonian. Besides, since $[x, \Pi_x] = [y, \Pi_y] = i\hbar$, one gets:

$$\begin{aligned} [X, Y] &= [x + \frac{1}{m\omega_c}\Pi_y, y - \frac{1}{m\omega_c}\Pi_x] \\ &= -\frac{1}{m\omega_c}[x, \Pi_x] + \frac{1}{m\omega_c}[\Pi_y, y] - \frac{1}{m^2\omega_c^2}[\Pi_y, \Pi_x] \\ &= -\frac{2}{m\omega_c}i\hbar + \frac{1}{m^2\omega_c^2}i\hbar\frac{qB}{c} \\ &= -i\frac{\hbar c}{qB}. \end{aligned} \tag{3.39}$$

The elementary volume in phase space associated to these variables is given by $2\pi\frac{\hbar c}{qB}$ (for $\hat{x}$ and $\hat{p}$, it is given by $2\pi\hbar$, i.e. $2\pi|[\hat{x}, \hat{p}]|$). In a sample of dimension $L_x, L_y$, the total volume is $L_xL_y$,

$$\rightarrow \boxed{\text{deg} = \frac{L_xL_y}{2\pi\hbar c}qB.} \tag{3.40}$$

## Density of states

Let us consider the density of states of one band $n$. It is given by:

$$n(E)\mathrm{d}E = 2n(k_z)\mathrm{d}k_z = 2 \times \text{deg} \times \frac{\mathrm{d}k_z}{2\pi/L_z} = 2 \times \frac{L_xL_y}{2\pi} \times \frac{qB}{\hbar c}\frac{L_z}{2\pi}\mathrm{d}k_z. \tag{3.41}$$

But $E = \hbar\omega_c(n + \frac{1}{2}) + \frac{\hbar^2k_z^2}{2m}$,

$$\mathrm{d}E = 2\frac{\hbar^2k_z}{2m}\mathrm{d}k_z \quad \rightarrow \mathrm{d}k_z = \frac{m}{\hbar^2k_z}\mathrm{d}E \tag{3.42}$$

and

$$k_z^2 = \frac{2m}{\hbar^2}\left(E - \hbar\omega_c(n+\frac{1}{2})\right), \tag{3.43}$$

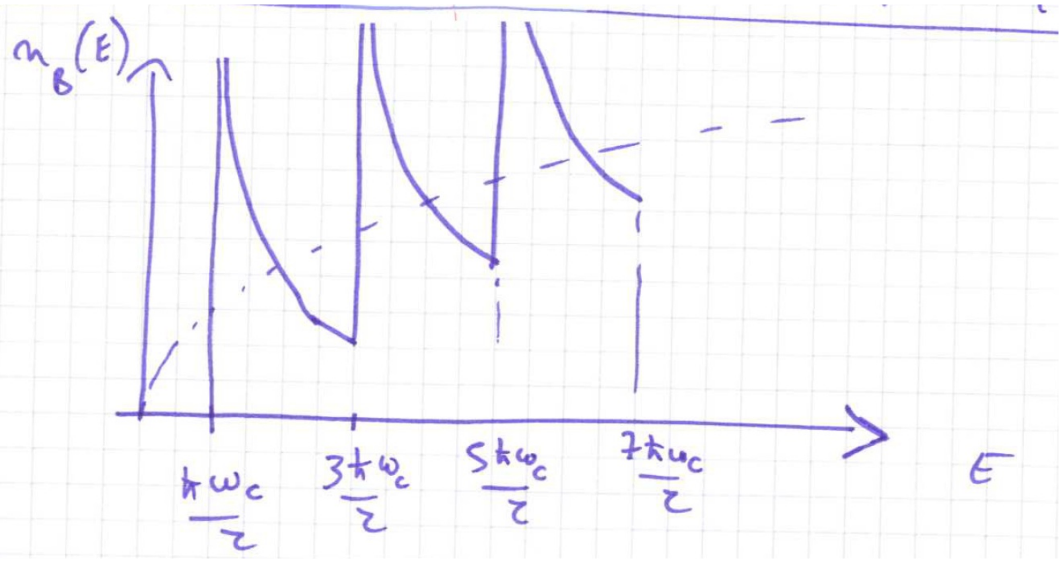
$$\rightarrow \mathrm{d}k_z = \frac{m}{\hbar^2}\frac{\mathrm{d}E}{\sqrt{\frac{2m}{\hbar^2}}\sqrt{E-\hbar\omega_c(n+\frac{1}{2})}}, \text{ si } E \geq \hbar\omega_c(n+\tfrac{1}{2}) \tag{3.44}$$

so that

$$\begin{aligned} \rightarrow n(E) &= \frac{V}{4\pi^2}\frac{qB}{\hbar c}\sqrt{\frac{2m}{\hbar^2}}\frac{\Theta\left(E-\hbar\omega_c(n+\frac{1}{2})\right)}{\sqrt{E-\hbar\omega_c(n+\frac{1}{2})}} \\ &= \frac{V}{8\pi^2}\hbar\omega_c\left(\frac{2m}{\hbar^2}\right)^{3/2}\frac{\Theta\left(E-\hbar\omega_c(n+\frac{1}{2})\right)}{\sqrt{E-\hbar\omega_c(n+\frac{1}{2})}} \end{aligned} \tag{3.45}$$

Finally, the total density of states, including a factor 2 for the spin, is given by:

$$n(E) = \frac{V}{8\pi^2}\hbar\omega_c\left(\frac{2m}{\hbar^2}\right)^{3/2}\sum_{n=0}^{\infty}\frac{\Theta\left(E-\hbar\omega_c(n+\frac{1}{2})\right)}{\sqrt{E-\hbar\omega_c(n+\frac{1}{2})}}. \tag{3.46}$$



The dashed line represents the density of states for $B = 0$. It is given by

$$n(E) = 2\frac{V}{4\pi^2}\left(\frac{2m}{\hbar^2}\right)^{3/2}\sqrt{E}\mathrm{d}E. \tag{3.47}$$

The curve in the presence of a magnetic field is very different from a simple square root. It exhibits singularities at energies given by $\hbar\omega_c(n+\frac{1}{2})$.

This property of the density of states is responsible for a remarkable phenomenon: many quantities such as the magnetization or the resistance exhibit oscillations as a function of $\frac{1}{B}$.

Indeed, when $B$ varies, one gets a new singularity when

$$E_F = \frac{\hbar e B}{mc}\left(n + \tfrac{1}{2}\right),$$

or

$$\frac{1}{B} = \frac{\hbar e}{mcE_F}\left(n + \tfrac{1}{2}\right). \tag{3.48}$$

If one calculates any thermodynamic quantity as a function of $\frac{1}{B}$, it will have regularly spaced anomalies. For the small values of $B$, these anomalies lead to a regular behaviour that shows up as oscillations in $\frac{1}{B}$. These oscillations are called the de Haas - van Alphen effect for the magnetization and the Shubnikov - de Haas effect for the resistance.

## Landau gauge

In many situations, it is useful to know the eigenfunctions of the problem. For a value of $k_z$, each Landau level has a degeneracy deg $= L_x L_y \frac{qB}{2\pi\hbar c}$. This degeneracy leads to different choices of basis. A convenient way to construct a basis of eigenfunctions of $\mathcal{H}_{xy}$ consists in choosing a gauge. Two choices are particularly useful:

- Landau gauge

$$A_x = -By, \quad A_y = 0, \quad A_z = 0. \tag{3.49}$$

In this gauge, the directions $x$ and $y$ are treated differently. It is useful for problems in rectangular geometry.

- Symmetric gauge

$$A_x = -\frac{By}{2}, \quad A_y = \frac{Bx}{2}, \quad A_z = 0. \tag{3.50}$$

This gauge leads to eigenfunctions that turned out to be essential in the construction of variational wave functions for the fractional Quantum Hall Effect.

### Landau gauge

$$\begin{aligned}\mathcal{H} &= \frac{1}{2m}\left(p_x - \frac{q}{c}A_x\right)^2 + \frac{1}{2m}p_y^2 + \frac{1}{2m}p_z^2 \\ &= \frac{1}{2m}\left(p_x + \frac{qB}{c}y\right)^2 + \frac{1}{2m}p_y^2 + \frac{1}{2m}p_z^2.\end{aligned} \tag{3.51}$$

$p_x$ and $p_z$ commute with $\mathcal{H}$. Thus one can look for eigenfunctions of the form

$$\varphi(x,y,z) = \mathrm{e}^{ik_x x}\mathrm{e}^{ik_z z}\varphi(y). \tag{3.52}$$

$$\rightarrow \frac{1}{2m}\left[\left(\hbar k_x + \frac{qB}{c}y\right)^2\right]\varphi(y) - \frac{\hbar^2}{2m}\frac{\partial^2\varphi}{\partial y^2} + \frac{\hbar^2 k_z^2}{2m}\varphi(y) = E\varphi(y) \tag{3.53}$$

Let us define

$$y_{k_x} = -\frac{c\hbar k_x}{qB} \text{ et } \omega_c = \frac{qB}{mc}. \tag{3.54}$$

We get:

$$\frac{1}{2}m\omega_c^2\left(y - y_{k_x}\right)^2\varphi(y) - \frac{\hbar^2}{2m}\frac{\partial^2\varphi}{\partial y^2} + \frac{\hbar^2 k_z^2}{2m}\varphi(y) = E\varphi(y). \tag{3.55}$$

This is a shifted harmonic oscillator of frequency $\omega_c$,

$$\rightarrow E_{n,k_x,k_z} = \hbar\omega_c\left(n + \tfrac{1}{2}\right) + \frac{\hbar^2 k_z^2}{2m}. \tag{3.56}$$

The energy is independent of $k_x$. This leads again to the degeneracy already calculated for the Landau levels. Indeed, $k_x$ must be quantized in units of $\frac{2\pi}{L_x}$, i.e. it must be of the form $k_x = p\frac{2\pi}{L_x}$, $p$ integer, and the coordinate $y_{k_x}$ must lie between 0 et $L_y$

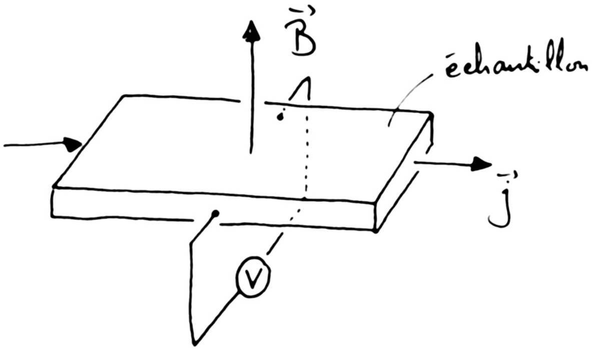
$$\Rightarrow 0 \leq c\hbar p\frac{2\pi}{L_x}\frac{1}{qB} \leq L_y. \tag{3.57}$$

$$\Rightarrow \boxed{p \leq \frac{L_x L_y qB}{2\pi\hbar c}} \tag{3.58}$$

## 3.2 The integer Quantum Hall Effect

### 3.2.1 The classical Hall effect

The Hall effect, discovered by Hall in 1879, consists in the fact that, in the presence of an applied magnetic field $\vec{B} = B\hat{z}$, a current in the $x$ direction $\vec{j} = j\hat{x}$ induces in a conductor a voltage in the $\hat{y}$ direction.



This effect can be understood in the context of the Drude theory:

$$\frac{\mathrm{d}\vec{p}}{\mathrm{d}t} = -e\left(\vec{E} + \frac{\vec{p}}{m} \wedge \frac{\vec{B}}{c}\right) - \frac{\vec{p}}{\tau} \tag{3.59}$$

where $\tau$ is the relaxation time. The stationary solution is given by

$$\frac{\mathrm{d}\vec{p}}{\mathrm{d}t} = 0 \tag{3.60}$$

so that

$$\begin{cases} -e\left(E_x + \dfrac{p_y}{mc}B\right) - \dfrac{p_x}{\tau} = 0 \\ -e\left(E_y - \dfrac{p_x}{mc}B\right) - \dfrac{p_y}{\tau} = 0 \end{cases} \tag{3.61}$$

Let us rewrite these equations with the help of the cyclotron frequency $\omega_c = \frac{eB}{mc}$ and of the Drude conductivity $\sigma_0 = \frac{ne^2\tau}{m}$. One gets:

$$\begin{cases} \sigma_0 E_x = -ne\dfrac{p_x}{m} - ne\dfrac{p_y}{m}\omega_c\tau \\ \sigma_0 E_y = ne\dfrac{p_x}{m}\omega_c\tau - ne\dfrac{p_y}{m} \end{cases} \tag{3.62}$$

The current density $\vec{j}$ is defined by:

$$\vec{j} = -ne\frac{\vec{p}}{m} \tag{3.63}$$

This system can be written:

$$\vec{E} = \rho\vec{j} \tag{3.64}$$

$$\text{with} \qquad \rho = \frac{1}{\sigma_0}\begin{pmatrix} 1 & \omega_c\tau \\ -\omega_c\tau & 1 \end{pmatrix}, \qquad \text{and} \qquad R_H = \rho_{yx} = \frac{-B}{nec} \tag{3.65}$$

where $R_H$ is the Hall resistance. The conductivity tensor is obtained by a matrix inversion: $\sigma = \rho^{-1}$. Now

$$\begin{pmatrix} 1 & \omega_c\tau \\ -\omega_c\tau & 1 \end{pmatrix}\begin{pmatrix} 1 & -\omega_c\tau \\ \omega_c\tau & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\left(1 + \omega_c^2\tau^2\right) \tag{3.66}$$

so that

$$\sigma = \begin{pmatrix} \sigma_L & -\sigma_H \\ \sigma_H & \sigma_L \end{pmatrix} \tag{3.67}$$

$$\text{with} \qquad \sigma_L = \frac{\sigma_0}{1 + \omega_c^2\tau^2} \qquad \text{and} \qquad \sigma_H = \sigma_0\frac{\omega_c\tau}{1 + \omega_c^2\tau^2} \tag{3.68}$$

If $\tau \to \infty$ (no dissipation),

$$\rho = \begin{pmatrix} 0 & -R_H \\ R_H & 0 \end{pmatrix} \qquad \text{and} \qquad \sigma = \begin{pmatrix} 0 & R_H^{-1} \\ -R_H^{-1} & 0 \end{pmatrix} \tag{3.69}$$

In 1980, Von Klitzing has investigated the Hall effect in a 2D electron gas. He observed plateaus in the function $\rho_{x,y}(B)$ for values of $R_H$ given by:

$$|R_H| = \left(\frac{h}{e^2}\right)\frac{1}{n}, \quad n \text{ integer.} \tag{3.70}$$

Inside these plateaus, $\rho_{xx} = 0$

$$\boxed{\sigma_H = n\frac{e^2}{h}} \tag{3.71}$$

This is the **Integer Quantum Hall Effect**. In 1983, Tsui, Störmer and Gossard discovered additional plateaus at rational values:

$$\sigma_H = \nu\frac{e^2}{h}, \qquad \nu = \frac{p}{q}. \tag{3.72}$$

This is the **Fractional Quantum Hall Effect**. The integer Quantum Hall Effect is a consequence of Landau levels and of disorder. The fractional Quantum Hall Effect is a consequence of the repulsion between electrons.

### 3.2.2 Hall effect and Landau levels

To discuss the integer Quantum Hall Effect, let us redo the calculation in the Landau gauge with an electric field in the $y$ direction. If the field is uniform, the potential is given by $V(y) = -Ey$, and the Hamiltonian becomes:

$$\begin{aligned} \mathcal{H} &\to \mathcal{H} + eEy \\ &= \frac{1}{2m}\left(p_x - \frac{eB}{c}y\right)^2 + \frac{1}{2m}p_y^2 + eEy \end{aligned} \tag{3.73}$$

Let us consider again $\phi(x,y) = e^{ik_x x}\phi(y)$. The eigenvalue equation is given by

$$\left[\frac{1}{2m}\left(\hbar k_x - \frac{eB}{c}y\right)^2 + \frac{1}{2m}p_y^2 + eEy\right]\phi(y) = \varepsilon\phi(y) \tag{3.74}$$

It is again a shifted harmonic oscillator, but this time the center of the harmonic oscillator depends on the electric field:

$$\begin{aligned} & \frac{1}{2m}\left(\hbar k_x - \frac{eB}{c}y\right)^2 + eEy = \frac{1}{2}m\omega_c^2(y - y(k_x))^2 + A \\ \Rightarrow \quad & \frac{\hbar^2 k_x^2}{2m} - \frac{eB}{mc}\hbar k_x y + \frac{1}{2m}\frac{e^2B^2}{c^2}y^2 + eEy = \frac{1}{2}m\omega_c^2 y^2 - m\omega_c^2 y y(k_x) + \frac{1}{2}m\omega_c^2 y(k_x)^2 + A \\ \Rightarrow \quad & \begin{cases} -\dfrac{eB}{mc}\hbar k_x + eE = -m\omega_c^2 y(k_x) \\ \dfrac{\hbar^2 k_x^2}{2m} = \dfrac{1}{2}m\omega_c^2 y(k_x)^2 + A \end{cases} \end{aligned} \tag{3.75}$$

$$\begin{aligned}
\Rightarrow \qquad y(k_x) &= -\frac{eE}{m\omega_c^2} + \frac{c\hbar k_x}{eB} \\
A &= \frac{\hbar^2 k_x^2}{2m} - \frac{1}{2}m\omega_c^2 y(k_x)^2 \\
&= \frac{\hbar^2 k_x^2}{2m} - \frac{1}{2}m\frac{e^2B^2}{m^2c^2}\left[\frac{e^2E^2}{m^2\omega_c^4} + \frac{c^2\hbar^2k_x^2}{e^2B^2} - \frac{2eE}{m\omega_c^2}\frac{c\hbar k_x}{eB}\right] \\
&= -\frac{1}{2}m\frac{e^2E^2}{m^2\omega_c^2} + \frac{1}{2}m\omega_c^2\frac{2eE}{m\omega_c^2}\frac{c\hbar k_x}{eB} \\
&= -\frac{1}{2}m\frac{E^2c^2}{B^2} + E\frac{c\hbar k_x}{B} \\
&= -\frac{1}{2}mc^2\frac{E^2}{B^2} + c\hbar k_x\frac{E}{B}
\end{aligned} \tag{3.76}$$

$$\begin{aligned}
\text{But} \qquad eEy(k_x) &= -\frac{e^2E^2}{me^2B^2}m^2c^2 + c\hbar k_x\frac{eE}{eB} \\
&= -mc^2\frac{E^2}{B^2} + c\hbar k_x\frac{E}{B}
\end{aligned} \tag{3.77}$$

Finally,

$$A = eEy(k_x) + \frac{1}{2}mc^2\frac{E^2}{B^2} \tag{3.78}$$

The eigenenergies are thus given by

$$\boxed{\varepsilon_{n,k_x} = \hbar\omega_c\left(n + \frac{1}{2}\right) + eEy(k_x) + \frac{1}{2}mc^2\frac{E^2}{B^2}} \tag{3.79}$$

with

$$\boxed{y(k_x) = \frac{c\hbar}{eB}k_x - \frac{eE}{m\omega_c^2}} \tag{3.80}$$

This time, the degeneracy is lifted: the energy depends on $k_x$.

Let us compute the current carried by each eigenstate:

$$\begin{aligned}
\vec{j} &= -e\vec{v} = -e\frac{\vec{p} + \frac{e}{c}\vec{A}}{m} \\
\hat{j}_x &= -\frac{e}{m}\left(p_x + \frac{e}{c}A_x\right) = -\frac{e}{m}\left(p_x - \frac{eBy}{c}\right)
\end{aligned} \tag{3.81}$$

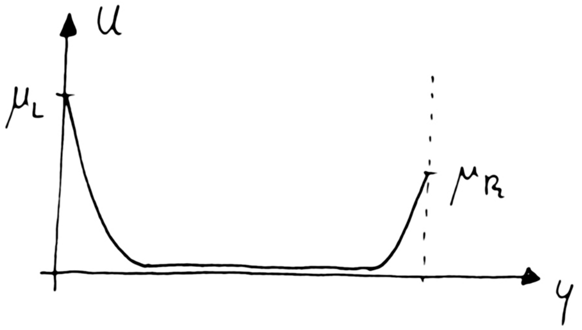
$$\begin{aligned}
\left\langle n, k_x \middle| \hat{j}_x \middle| n, k_x \right\rangle &= -\frac{e}{m}\hbar k_x + \frac{e^2}{mc}By_{k_x} \\
&= -\frac{e}{m}\hbar k_x + \frac{e^2}{mc}B\frac{c\hbar}{eB}k_x - \frac{e^2}{mc}B\frac{eE}{m\omega_c^2} \\
&= -\frac{e^2}{mc}B\frac{eEm^2c^2}{me^2B^2} = -ec\frac{E}{B}
\end{aligned} \tag{3.82}$$

or

$$\langle j_x \rangle = -ec\frac{E}{B} \tag{3.83}$$

Besides, $\langle j_y \rangle = 0$ because the ground state corresponds to a harmonic oscillator in the $y$ direction. In this model where the electric field is uniform, the total current is thus simply proportional to the number of electrons since all eigenstates carry the same current. This is the classical result when there is no dissipation: $\sigma_H = R_H^{-1} = -n\frac{ec}{B} \to j_x = -nec\frac{E}{B}$.

Let us now suppose that the potential is not simply linear, but that it confines the electron gas:



In the center, the potential varies very little, the field practically vanishes, and the states do not carry current. So the total current is no longer simply a linear function of the filling. Nevertheless, one can convince oneself that, if a Landau level is full, the intensity of the current depends only on the voltage difference, regardless of how the electric field varies inside the sample. Indeed, assuming that the electric field does not vary too fast, one can use the result above to estimate the current by simply taking into account the $y$ dependence of the electric field:

$$\langle j_x \rangle(k_x) = -\frac{ec}{B}E(y(k_x)) \tag{3.84}$$

The voltage bias in the $y$ direction is given by:

$$\Delta V = -\int E(y)dy = -\sum_{k_x} E[y(k_x)]\Delta y \tag{3.85}$$

where $\Delta y$ is the spacing between values of $y$ due to the quantification of $k_x$. But $k_x$ is quantized in units of $2\pi/L_x$, and $y(k_x) = \frac{c\hbar}{eB}k_x - \frac{eE}{m\omega_c^2}$, so that

$$\Delta y = \frac{c\hbar}{eB}\Delta k_x = \frac{c\hbar}{eB}\frac{2\pi}{L_x} \tag{3.86}$$

The expectation value of the total current is thus given by:

$$\begin{aligned} J_{tot} &= \sum_{k_x} \langle j_x \rangle (k_x) && (3.87) \\ &= -\frac{ec}{B}\frac{1}{\Delta y}\sum_{k_x} E(y(k_x))\Delta y && (3.88) \\ &= -\frac{ec}{B}\frac{eB}{c\hbar}\frac{L_x}{2\pi}(-\Delta V) && (3.89) \\ &= \frac{e^2}{h}\Delta V L_x && (3.90) \end{aligned}$$

The current density is equal to $J_{tot}/L_xL_y$, and its intensity is equal to the density times the section, here the width $L_y$. It follows that:

$$\boxed{I = \frac{e^2}{h}\Delta V} \tag{3.91}$$

If $n$ Landau levels are filled, the Hall conductance is thus given by:

$$G_{xy} = n\frac{e^2}{h} \tag{3.92}$$

Besides, there is no current in the $y$ direction: $G_{yy} = 0$. Finally:

$$G = \begin{pmatrix} 0 & \frac{ne^2}{h} \\ -\frac{ne^2}{h} & 0 \end{pmatrix} \quad \text{and} \quad R = \begin{pmatrix} 0 & -\frac{h}{e^2n} \\ \frac{h}{e^2n} & 0 \end{pmatrix} \tag{3.93}$$

It is very instructive to compare this result to the classical result:

$$R_H = \frac{-B}{n_{\text{el}}ec} \tag{3.94}$$

where $n_{\text{el}}$ denotes the number of electrons. Since there are $\frac{L_xL_yeB}{2\pi\hbar c}$ states per Landau level, if $n$ levels are filled, the electronic density is given by:

$$n_{\text{el}} = n\ \text{deg}\frac{1}{L_xL_y} = n\frac{L_xL_yeB}{2\pi\hbar c}\frac{1}{L_xL_y} = n\frac{eB}{2\pi\hbar c} = \frac{neB}{hc} \tag{3.95}$$

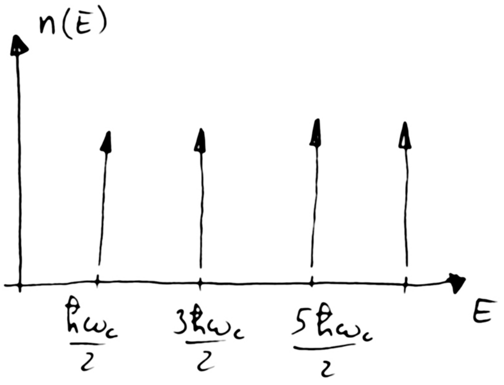
$$\Rightarrow R_H = \frac{-Bhc}{neBec} = -\frac{h}{ne^2} \tag{3.96}$$

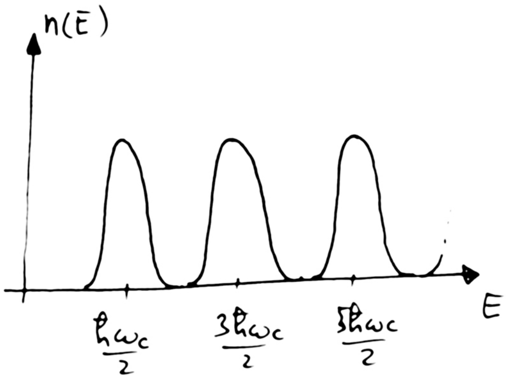
Thus, for a filling corresponding to filled Landau levels, the Hall resistance just takes its classical value. What is remarkable is that this value is stabilized over a field range. This can only be understood as a consequence of disorder.

### 3.2.3 Effect of disorder

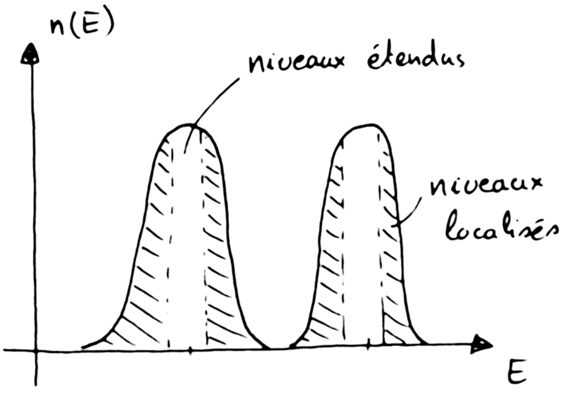
Until now, we have considered that the potential was periodic in the $x$ direction. In a real sample, the potential is of course not periodic in any direction. What is the impact of disorder on the density of states?

In the absence of disorder, the density of states consists of a series of $\delta$ peaks of weight $\frac{L_x L_y}{2\pi\hbar c} qB$. In the presence of disorder, the peaks are broadened:





The presence of plateaus in $R_H$ can only be explained if the density of states of a Landau level consists of "localized" states on one hand, and "extended" states on the other hand:



Then, if one starts from a magnetic field such that $n$ Landau levels are filled, to increase or decrease the magnetic field will change the structure of Landau levels in such a way that localized levels will be emptied or filled. But these localized states do not carry any current, and the Hall resistance will not change until extended levels are reached.

Why does the density of states have this structure? There is a simple hand-waving argument to understand this. In the classical case, since the Lorentz force is perpendicular to the velocity, it does not work, and the particles follow equipotential curves. For a Landau level that corresponds to a given velocity, the lowest-energy trajectories follow equipotential curves at the bottom of valleys, the highest-energy ones at the top of the peaks. They are localized and give rise to quantum states that do not carry any current. In between, there must exist trajectories that go from one end of the sample to the other. These are the trajectories that contribute to the Hall resistance.

## 3.3 The Fractional Quantum Hall Effect

### 3.3.1 Introduction

Let us consider again the case of a system without disorder and without electric field. If a Landau level is partially filled, the ground state is massively degenerate. Indeed to build it one can pick any levels among the $\frac{L_x L_y}{2\pi\hbar c} eB$ single particle states of the single particle ground state manifold since they are degenerate. However electrons interact via Coulomb repulsion. This perturbation must a priori lift the degeneracy. This is an extremely difficult problem that does not possess any simple solution, even approximate. The picture that has imposed itself is based on variational wave functions proposed by Laughlin in 1983. These variational wave functions rely in an essential way on the solution of Landau levels in the symmetric gauge. Let us thus start by reviewing the solution of this problem, using a formulation in which the position of the particles is represented by a complex variable in the plane, opening the way to Laughlin's generalization to the many-body case.

### 3.3.2 Landau levels and symmetric gauge

Let us consider the gauge

$$A_x = -\frac{By}{2}, \quad A_y = \frac{Bx}{2}, \quad A_z = 0.$$

We have shown that, quite generally, the Hamiltonian takes the form

$$H = \frac{1}{2m}\hat{P}^2 + \frac{1}{2}m\omega_c^2\hat{Q}^2$$

with $\hat{P} = \hat{\Pi}_x$, $\hat{Q} = \frac{c}{eB}\hat{\Pi}_y \quad (e = -q)$ and

$$\begin{cases} \hat{\Pi}_x = \hat{p}_x + \dfrac{e}{c}A_x \\ \hat{\Pi}_y = \hat{p}_y + \dfrac{e}{c}A_y \end{cases} \qquad [\hat{\Pi}_x, \hat{\Pi}_y] = -i\hbar\frac{e}{c}B.$$

Besides, there is an additional degeneracy due to the fact that the operators

$$\begin{cases} \hat{X} = \hat{x} - \dfrac{1}{m\omega_c}\hat{\Pi}_y \\ \hat{Y} = \hat{y} + \dfrac{1}{m\omega_c}\hat{\Pi}_x \end{cases}$$

commute with the Hamiltonian. (NB: the sign has changed because $\omega_c = \frac{eB}{mc} = -\frac{qB}{mc}$.) Indeed,

$$[\hat{X}, \hat{\Pi}_x] = [\hat{x} - \frac{1}{m\omega_c}\hat{\Pi}_y, \hat{\Pi}_x] = i\hbar + \frac{1}{m\omega_c}[\hat{\Pi}_y, \hat{\Pi}_x] = i\hbar - i\hbar\frac{eB}{c}\frac{1}{m\omega_c} = 0.$$

In the same way, $[\hat{X}, \hat{\Pi}_y] = [\hat{Y}, \hat{\Pi}_x] = [\hat{Y}, \hat{\Pi}_y] = 0$. In addition,

$$\begin{aligned} [\hat{X}, \hat{Y}] &= [\hat{x} - \frac{1}{m\omega_c}\hat{\Pi}_y, \hat{y} + \frac{1}{m\omega_c}\hat{\Pi}_x] \\ &= \frac{1}{m\omega_c}[\hat{x}, \hat{\Pi}_x] + \frac{1}{m\omega_c}[\hat{y}, \hat{\Pi}_y] + \frac{1}{m^2\omega_c^2}[\hat{\Pi}_x, \hat{\Pi}_y] \\ &= \frac{1}{m\omega_c}(i\hbar + i\hbar - i\hbar) = \frac{i\hbar}{m\omega_c} = \frac{i\hbar c}{eB}. \end{aligned}$$

It is convenient to formulate the problem in terms of the magnetic length $l_B$ defined by

$$l_B^2 = \frac{\hbar c}{eB} \Rightarrow l_B = \sqrt{\frac{\hbar c}{eB}},$$

and of the operators $\hat{\eta}_x$, $\hat{\eta}_y$ defined by

$$\begin{cases} m\omega_c\hat{\eta}_x = \hat{\Pi}_y \\ -m\omega_c\hat{\eta}_y = \hat{\Pi}_x \end{cases}$$

(these definitions come from the classical treatment: $\eta_x = x - x_0$, $\eta_y = y - y_0$ and $\dot{x} = -\omega_c(y - y_0)$, $\dot{y} = \omega_c(x - x_0)$).

The problem can then be rewritten as:

$$H = \frac{1}{2}m\omega_c^2(\hat{\eta}_x^2 + \hat{\eta}_y^2)$$

with $[\hat{\eta}_x, \hat{\eta}_y] = -il_B^2$ and $[\hat{X}, \hat{Y}] = il_B^2$.

Let us introduce the operators

$$\begin{cases} a = \frac{1}{\sqrt{2}l_B}(\hat{\eta}_x - i\hat{\eta}_y) & a^\dagger = \frac{1}{\sqrt{2}l_B}(\hat{\eta}_x + i\hat{\eta}_y) \\ \Rightarrow \hat{\eta}_x = \frac{l_B}{\sqrt{2}}(a^\dagger + a) & \hat{\eta}_y = -i\frac{l_B}{\sqrt{2}}(a^\dagger - a) \end{cases}$$

$$\begin{cases} b = \frac{1}{\sqrt{2}l_B}(\hat{X} + i\hat{Y}) & b^\dagger = \frac{1}{\sqrt{2}l_B}(\hat{X} - i\hat{Y}) \\ \Rightarrow \hat{X} = \frac{l_B}{\sqrt{2}}(b^\dagger + b) & \hat{Y} = i\frac{l_B}{\sqrt{2}}(b^\dagger - b) \end{cases}$$

On can easily check that $[a, a^\dagger] = [b, b^\dagger] = 1$, et $[a, b] = [a, b^\dagger] = 0$.

The Hamiltonian can be rewritten:

$$H = \hbar\omega_c \left(a^\dagger a + \frac{1}{2}\right) + \hbar\omega' \left(b^\dagger b + \frac{1}{2}\right)$$

with $\omega' = 0$.

This allows one to define the basis $|n, m\rangle$ by:

$$|n, m\rangle = \frac{(a^\dagger)^n}{\sqrt{n!}} \frac{(b^\dagger)^m}{\sqrt{m!}} |0, 0\rangle.$$

Of course:

$$a^\dagger |n, m\rangle = \sqrt{n+1}|n+1, m\rangle, a|n, m\rangle = \sqrt{n}|n-1, m\rangle,$$

$$b^\dagger |n, m\rangle = \sqrt{m+1}|n, m+1\rangle, b|n, m\rangle = \sqrt{m}|n, m-1\rangle.$$

NB1: So far, the argument does not depend on the gauge.

NB2: The operators $a^\dagger$ and $b^\dagger$ are equal to the operators $a_d^\dagger$ and $a_g^\dagger$ seen in the exercise if ones picks the symmetric gauge.

To find the eigenfunctions in a given gauge, one just needs to write the operators $a^\dagger, a, b^\dagger, b$ in terms of $x, p_x, y, p_y$ in this gauge, and to determine the ground state wave function from the condition $a|0, 0\rangle = b|0, 0\rangle = 0$.

With the conventional notation for negatively charged particles

$$z = x - iy, \bar{\partial} = (\partial_x - i\partial_y)/2, z^* = x + iy, \partial = (\partial_x + i\partial_y)/2,$$

one gets:

$$\begin{cases} a = \sqrt{2}(\frac{z}{4l_B} + l_B\bar{\partial}), & a^\dagger = \sqrt{2}(\frac{z^*}{4l_B} - l_B\partial) \\ b = \sqrt{2}(\frac{z^*}{4l_B} + l_B\partial), & b^\dagger = \sqrt{2}(\frac{z}{4l_B} - l_B\bar{\partial}) \end{cases}.$$

Intermediate steps: in the symmetric gauge,

$$\eta_x = \frac{1}{2}x + \frac{l_B^2}{\hbar}p_y, \quad \eta_y = \frac{1}{2}y - \frac{l_B^2}{\hbar}p_x,$$

$$X = x - \eta_x = \frac{1}{2}x - \frac{l_B^2}{\hbar}p_y, \quad Y = y - \eta_y = \frac{1}{2}y + \frac{l_B^2}{\hbar}p_x.$$

Moreover, $m\omega_c = \frac{\hbar}{l_B^2} = \frac{eB}{c}$.

A state of the first Landau level satisfies:

$$a|n=0,m\rangle = 0 \Rightarrow (z + 4l_B^2\bar{\partial})\Phi_{n=0}(z,z^*) = 0.$$

But

$$\begin{aligned} \bar{\partial}f(z,z^*) &= \frac{1}{2}\left[\partial_x f(z,z^*) - i\partial_y f(z,z^*)\right] \\ &= \frac{1}{2}\left[\frac{\partial f}{\partial z} + \frac{\partial f}{\partial z^*} - i\frac{\partial f}{\partial z}(-i) - i\frac{\partial f}{\partial z^*}(i)\right] \end{aligned}$$

$$\boxed{\bar{\partial}f(z,z^*) = \frac{\partial f}{\partial z^*}.}$$

Similarly,

$$\begin{aligned} \partial f(z,z^*) &= \frac{1}{2}\left[\partial_x f(z,z^*) + i\partial_y f(z,z^*)\right] \\ &= \frac{1}{2}\left[\frac{\partial f}{\partial z} + \frac{\partial f}{\partial z^*} + i\frac{\partial f}{\partial z}(-i) + i\frac{\partial f}{\partial z^*}(i)\right] \end{aligned}$$

$$\boxed{\partial f(z,z^*) = \frac{\partial f}{\partial z}.}$$

Let us consider the gaussian $\varphi(z,z^*) = \mathrm{e}^{-|z|^2/4l_B^2}$. Its derivative $\bar{\partial}\varphi(z,z^*)$ is given by:

$$\bar{\partial}\varphi(z,z^*) = \frac{\partial}{\partial z^*}\left(\mathrm{e}^{-|z|^2/4l_B^2}\right) = \frac{-z}{4l_B^2}\varphi(z,z^*),$$

$$\Rightarrow (z + 4l_B^2\bar{\partial})\varphi(z,z^*) = 0$$

$\varphi(z,z^*)$ is thus in the first Landau level.

But for any function that depends only on $z$, $\bar{\partial}f(z) = 0$. The general form of a state of the first Landau level is thus

$$f(z)\mathrm{e}^{-|z|^2/4l_B^2}$$

Similarly, the $m = 0$ states satisfy

$$(z^* + 4l_B^2\partial)\Phi_{m=0}(z, z^*) = 0,$$

whose general solution is

$$g(z^*)\mathrm{e}^{-|z|^2/4l_B^2}.$$

The wave function $\Phi_{n=0,m=0}$ is thus a gaussian. After normalization, it is given by:

$$\Phi_{n=0,m=0}(z, z^*) = \frac{1}{\sqrt{2\pi l_B^2}}\mathrm{e}^{-|z|^2/4l_B^2}.$$

The other states can be obtained from this one by successive applications of $a^\dagger, b^\dagger$. In particular,

$$\begin{aligned}\Phi_{n=0,m}(z, z^*) &= \frac{(\sqrt{2})^m}{\sqrt{2\pi l_B^2 m!}}\left(\frac{z}{4l_B} - l_B\bar{\partial}\right)^m \mathrm{e}^{-|z|^2/4l_B^2} \\ &= \frac{1}{\sqrt{2\pi l_B^2 m!}}\left(\frac{z}{\sqrt{2}l_B}\right)^m \mathrm{e}^{-|z|^2/4l_B^2},\end{aligned}$$

since

$$\left(\frac{z}{4l_B} - l_B\bar{\partial}\right) f(z)\mathrm{e}^{-|z|^2/4l_B^2} = \frac{z}{2l_B} f(z)\mathrm{e}^{-|z|^2/4l_B^2}.$$

Similarly,

$$\Phi_{n,m=0}(z, z^*) = \frac{1}{\sqrt{2\pi l_B^2 n!}}\left(\frac{z^*}{\sqrt{2}l_B}\right)^n \mathrm{e}^{-|z|^2/4l_B^2}.$$

The states corresponding to different values of $m$ for $n = 0$ satisfy

$$\langle\hat{R}\rangle \equiv \langle n = 0, m|\hat{R}|n = 0, m\rangle = 0$$

since $\hat{X}$ and $\hat{Y}$ are linear combinations of $b$ and $b^\dagger$, and $\langle m|b, b^\dagger|m\rangle = 0$, and

$$\langle|\hat{R}|\rangle \equiv \langle\sqrt{X^2 + Y^2}\rangle = l_B\langle\sqrt{2b^\dagger b + 1}\rangle = l_B\sqrt{2m + 1}.$$

When $m$ increases, the particle gets away from the origin in a symmetric way.

### 3.3.3 The wave function of the $\nu = 1$ state:

Let us first consider the non-interacting case. A completely filled Landau level corresponds to a totally polarized spin wavefunction. Indeed, the Zeeman coupling of the spin to the magnetic field lifts the degeneracy, and each Landau level corresponds to two Landau levels separated by $g\mu_B H$.

Since the spin component of the wave function is totally symmetric, the orbital component must be totally antisymmetric. Besides, to be a solution of the one-particle problem, the functions must be of the form

$$f(z)\mathrm{e}^{-|z|^2/4l_B^2}.$$

So one looks for a function of the form

$$\psi_{\nu=1}(\{z_i, z_i^*\}) = f_N(\{z_i\})\mathrm{e}^{-\sum_j |z_j|^2/4l_B^2}$$

where $f_N$ is totally antisymmetric in the exchange of the variables $z_i$.

Besides, a convenient basis of the one particle states is given by

$$\Phi_{n=0,m}(z, z^*) = \frac{1}{\sqrt{2\pi l_B^2 m!}}\left(\frac{z}{\sqrt{2}l_B}\right)^m \mathrm{e}^{-|z|^2/4l_B^2},$$

where $m$ can take all the values from 0 to $N-1$ ($N$ = degeneracy = number of electrons when $\nu = 1$).

The function $f_N(\{z_i\})$ that can be written simply in this basis and that is totally antisymmetric is a Slater determinant built out of the polynomials $z_i^0, z_i^1, ..., z_i^N$,

$$f_N(\{z_i\}) = \begin{vmatrix} z_1^0 & z_1^1 & \dots & z_1^{N-1} \\ z_2^0 & z_2^1 & \dots & z_2^{N-1} \\ \vdots & \vdots & \ddots & \vdots \\ z_N^0 & z_N^1 & \dots & z_N^{N-1} \end{vmatrix}.$$

This is a Vandermonde determinant. It can be written as:

$$f_N(\{z_i\}) = \prod_{i<j}(z_i - z_j).$$

This wave function is unique. It is thus an eigenstate of any interaction Hamiltonian that does not mix Landau levels.

It is instructive to calculate the density. Since this is a Slater determinant, it is the sum of the densities of each function. Indeed,

$$\hat{\rho}(z) = \sum_i \delta(z - z_i),$$

and thus,

$$\begin{aligned}
\rho(z) &= \frac{1}{N!}\sum_P (\mathrm{sgn}(P))^2 \sum_i \int \mathrm{d}z_1 \ldots \mathrm{d}z_N \delta(z-z_i)|\varphi_{P(1)}(z_1)|^2 \ldots |\varphi_{P(n)}(z_N)|^2 \\
&= \frac{1}{N!}\sum_P \sum_i |\varphi_{P(i)}(z)|^2 \\
&= \frac{1}{N!}\sum_P \sum_i |\varphi_i(z)|^2 \\
&= \sum_i |\varphi_i(z)|^2 .
\end{aligned}$$

But

$$\varphi_m(z) = \frac{1}{\sqrt{2\pi l_B^2 m!}} \left(\frac{z}{\sqrt{2} l_B}\right)^m \mathrm{e}^{-|z|^2/4l_B^2}.$$

So,

$$\begin{aligned}
\rho(z) &= \left(\sum_m \frac{1}{2\pi l_B^2} \frac{|z|^{2m}}{m!(2l_B^2)^m}\right) \mathrm{e}^{-|z|^2/2l_B^2} \\
&= \frac{1}{2\pi l_B^2} \mathrm{e}^{|z|^2/2l_B^2}\, \mathrm{e}^{-|z|^2/2l_B^2} \\
\to \rho(z) &= \frac{1}{2\pi l_B^2}.
\end{aligned}$$

The density is uniform, and of course it corresponds to one electron per elementary area $2\pi l_B^2$.

### 3.3.4 $\nu < 1$: Laughlin's wave functions

When the filling is smaller than 1, Laughlin has proposed to generalize the wave function of the $\nu = 1$ case in the following way:

$$\psi(z_1, ..., z_N) = \prod_{i<j} f(z_i - z_j) \mathrm{e}^{-\sum_i |z_i|^2/4l_B^2}.$$

The prefactor is called a Jastrow factor. The wave function is constrainded by the following conditions:

- It only involves states of the the first Landau level
  $\to f$ does not depend on the $z_i^*$.
- $\psi$ is totally antisymmetric
  $\to f$ is odd.

- $\psi$ is an eigenstate of the angular momentum since, in the symmetric gauge, the angular momentum commutes with the Hamiltonian. But for an eigenstate of the first Landau level, the angular momentum is proportional to $m$ ($L_z/\hbar = a^\dagger a - b^\dagger b$). $\prod_{i<j} f(z_i - z_j)$ must thus be a polynomial of fixed degree $\Rightarrow f(z_i - z_j) = (z_i - z_j)^m$.

Finally, $f(z_i - z_j)$ must be of the form $(z_i - z_j)^m$, $m$ odd,

$$\Rightarrow \boxed{\psi_m(z_1, ..., z_N) = \prod_{i<j} (z_i - z_j)^m \mathrm{e}^{-\sum_i |z_i|^2/4l_B^2}}.$$

Let us explore the physics of this wave function.

Density: Since this is not a Slater determinant, there is no simple way to calculate the density. But there is a simple argument to estimate it. For a given particle, the highest power is $m(N-1)$. To build this function, one must thus use states whose radius is of the order of $l_B\sqrt{2m(N-1)+1}$. One must thus take a sample of size

$$\pi l_B^2 (2m(N-1)+1) = \pi l_B^2 2mN + O(1).$$

The density is given by

$$\frac{N}{2\pi l_B^2 mN} = \frac{1}{2\pi l_B^2 m}$$

$$\Rightarrow \boxed{\nu = \frac{1}{m}}.$$

So Laughlin wave functions allow one to describe states of filling $\nu = \frac{1}{m}$, $m$ odd.

Energy: These wave functions turn out to be remarkably good. To get an idea why, let us consider the other states of density $\frac{1}{m}$:

$$\psi(z_1, ..., z_N) = \prod_{i<j} (z_i + z_j)^{m-p} (z_i - z_j)^p \mathrm{e}^{-\sum_i |z_i|^2/4l_B^2}$$

with $p$ odd. These states are much less competitive energetically as soon as $p < m$ because the probability that two particles are close to each other is larger. Since the Coulomb repulsion is repulsive, large values of $p$ are favoured. In fact, the only serious contender is the Wigner crystal, a wave function that breaks the translational symmetry and localizes electrons to form a triangular lattice. Comparing the Coulomb energy in both states, one can show that the Wigner crystal is stabilized for $\nu < \frac{1}{7}$, while Laughlin states are stabilized for $\frac{1}{7} \leq \nu \leq 1$.

Excitations: If one adds one electron to the system, this costs a finite energy. To convince oneself of this property, we start by noting that, in a Laughlin state, the relative angular momentum of two particles is equal to $m$. Let us write the interaction potential energy between two particles as

$$V(i,j) = \sum_{m'=0}^{\infty} v_{m'} P_{m'}(ij),$$

where $P_{m'}(ij)$ is the projector on the subspace of angular momentum $m'$. If one assumes that $v_{m'} = 0$ if $m' \geq m$, then the function $\psi_{1/m}$ is an eigenstate of $\hat{V} = \sum_{i<j} V(i,j)$ with

$$\hat{V}\psi_{1/m} = 0.$$

If one tries to add one particle, one cannot stay in a state where all pairs have a relative angular momentum $m$ (this would force one out of the sample). One can convince oneself that one needs to create pairs of angular momentum $< m$. Indeed, this is necessary for the total polynomial to keep a degree in each variable less than or equal to $m(N-1)$. This costs a finite energy. There is thus a charge gap. This property remains true if one includes the terms of $V(i,j)$ with $m' \geq m$ in perturbation.

The presence of a gap is absotutely necessary to understand the presence of a plateau: it is because it is impossible to add a particle for an infinitesimal cost that $\sigma_{xx} = \rho_{xx} = 0$, a necessary condition for the appearance of a plateau (like in the case of completely filled Landau level). This condition is however not sufficient, and one must again invoke disorder to explain the constant value of $R_H$ around $\nu = \frac{1}{m}$.

**Fractional charge:** Let us now think about the excitations that one can induce without changing the number of particles. Simple electron-hole excitations are very deep modifications of the wave function. One can implement simpler modifications as follows:

$$\psi_{1/m}^{(1)} = \prod_j z_j \psi_{1/m}.$$

In this function, the polynomial is of degree $\geq 1$ in each variable. Thus the level $m = 0$ is not used. Similarly, one can create $\psi_{1/m}^{(j)}$ for $j \leq m$ by creating $m$ excitations of the same type. For $j = m$, the wave function can be written:

$$\psi_{1/m}^{(m)} = \prod_j z_j^m \prod_{i<j} (z_i - z_j)^m \mathrm{e}^{-\sum_i |z_i|^2/4l_B^2} = \psi_{1/m}(z_0 = 0, z_1, ..., z_N).$$

This is the $\nu = 1/m$ Laughlin wave function with a particle sitting at $z_0 = 0$. The wave function $\psi_{1/m}^{(m)}(z_0 = 0, z_1, ..., z_N)$ thus corresponds to a state with one hole of charge $e$ at the center, and one electron of charge $-e$ at the periphery. But this excitation is composed of $m$ elementary excitations that each consists in multiplying the wave function by $\prod_j z_j$. Each of these excitations carries a charge $e/m$. This is one of the most remarkable predictions of Laughlin. It has been checked in noise measurements.

**Generalizations** Laughlin's wave functions allow one to understand the plateaus $\nu = 1/m$, $m$ odd. However, other plateaux have been observed, noticeably at $\frac{2}{5}$. Various approaches have been suggested to generalize Laughlin's theory. The most promising is due to Jain. It consists in building Laughlin's states using higher

Landau levels, and to project back onto the first level. A detailed discussion of this contruction would lead us too far. We just note that, within this approach, it is plausible that there are plateaus for

$$\boxed{\nu = \frac{n}{2np \pm 1}} \quad n, p \, \text{integers} \geq 1.$$

# Appendix

### Angular momentum

$$l_B^2 = \frac{\hbar c}{eB} = \frac{mc}{eb}\frac{\hbar}{m} = \frac{\hbar}{m\omega_c}$$

$$\hat{L}_z = \hat{x}\hat{p}_y - \hat{y}\hat{p}_x$$

$$\begin{cases} \hat{\Pi}_x = \hat{p}_x - \dfrac{e}{2c}B\hat{y} \\ \hat{\Pi}_y = \hat{p}_y + \dfrac{e}{2c}B\hat{x} \end{cases} \qquad \begin{cases} \hat{X} = \hat{x} - \dfrac{1}{m\omega_c}\hat{\Pi}_y \\ \hat{Y} = \hat{y} + \dfrac{1}{m\omega_c}\hat{\Pi}_x \end{cases}$$

$$\begin{cases} m\omega_c\hat{X} = m\omega_c\hat{x} - \hat{\Pi}_y \\ m\omega_c\hat{Y} = m\omega_c\hat{y} + \hat{\Pi}_x \end{cases}$$

$$\begin{cases} \hat{x} = \hat{X} + \dfrac{1}{m\omega_c}\hat{\Pi}_y \\ \hat{y} = \hat{Y} - \dfrac{1}{m\omega_c}\hat{\Pi}_x \end{cases}$$

$$\begin{aligned} p_x &= \Pi_x + \frac{eB}{2c}y = \Pi_x + \frac{1}{2}m\omega_c y \\ &= \Pi_x + \frac{1}{2}m\omega_c Y - \frac{1}{2}\Pi_x \\ &= \frac{1}{2}m\omega_c Y + \frac{1}{2}\Pi_x \end{aligned}$$

$$\begin{aligned} p_y &= \Pi_y - \frac{eB}{2c}x = \Pi_y - \frac{1}{2}m\omega_c x \\ &= \Pi_y - \frac{1}{2}m\omega_c X - \frac{1}{2}\Pi_y \\ &= -\frac{1}{2}m\omega_c X + \frac{1}{2}\Pi_y \end{aligned}$$

$$
\begin{aligned}
L_z &= xp_y - yp_x \\
&= \left(X + \frac{1}{m\omega_c}\Pi_y\right)\left(-\frac{1}{2}m\omega_c X + \frac{1}{2}\Pi_y\right) - \left(Y - \frac{1}{m\omega_c}\Pi_x\right)\left(\frac{1}{2}m\omega_c Y + \frac{1}{2}\Pi_x\right) \\
&= -\frac{1}{2}m\omega_c X^2 + \frac{1}{2m\omega_c}\Pi_y^2 - \frac{1}{2}m\omega_c Y^2 + \frac{1}{2m\omega_c}\Pi_x^2 \\
&= \frac{1}{2}m\omega_c\left[-(X^2+Y^2) + (\eta_x^2 + \eta_y^2)\right] \\
&= \frac{\hbar}{2l_B^2}\left[-\frac{l_B^2}{2}(b^\dagger + b)^2 + \frac{l_B^2}{2}(b^\dagger - b)^2 + \frac{l_B^2}{2}(a^\dagger + a)^2 - \frac{l_B^2}{2}(a^\dagger - a)^2\right] \\
&= \frac{\hbar}{2}\left[-b^\dagger b - bb^\dagger + a^\dagger a + aa^\dagger\right] \\
&= \hbar\left[a^\dagger a + \frac{1}{2} - b^\dagger b - \frac{1}{2}\right] \\
&= \hbar\left[a^\dagger a - b^\dagger b\right]
\end{aligned}
$$

$$
\begin{cases}
\vec{R}_{CM} = \dfrac{\vec{r}_1 + \vec{r}_2}{2} \\
\vec{P}_{CM} = \vec{p}_1 + \vec{p}_2
\end{cases}
\qquad
\begin{cases}
\vec{R}_{rel} = \vec{r}_1 - \vec{r}_2 \\
\vec{P}_{rel} = \dfrac{\vec{p}_1 - \vec{p}_2}{2}
\end{cases}
$$

$$\rightarrow \vec{l}_1 + \vec{l}_2 = \vec{l}_{CM} + \vec{l}_{rel}$$

$$\vec{l}_{CM} = \vec{R}_{CM} \wedge \vec{P}_{CM}, \qquad \vec{l}_{rel} = \vec{R}_{rel} \wedge \vec{P}_{rel}$$

$$
\begin{aligned}
l_{rel}^z &= -\frac{m\omega_c}{4}\left\{(X_1 - X_2)^2 + (Y_1 - Y_2)^2 - (\eta_1^x - \eta_2^x)^2 - (\eta_1^y - \eta_2^y)^2\right\} \\
&= \frac{\hbar}{2}\left\{a_1^\dagger a_1 + a_2^\dagger a_2 - a_1^\dagger a_2 - a_1 a_2^\dagger - b_1^\dagger b_1 - b_2^\dagger b_2 + b_1^\dagger b_2 + b_1 b_2^\dagger\right\}
\end{aligned}
$$

$$
\begin{aligned}
l_{CM}^z &= -\frac{m\omega_c}{4}\left\{(X_1 + X_2)^2 + (Y_1 + Y_2)^2 - (\eta_1^x + \eta_2^x)^2 - (\eta_1^y + \eta_2^y)^2\right\} \\
&= \frac{\hbar}{2}\left\{a_1^\dagger a_1 + a_2^\dagger a_2 + a_1^\dagger a_2 + a_1 a_2^\dagger - b_1^\dagger b_1 - b_2^\dagger b_2 - b_1^\dagger b_2 - b_1 b_2^\dagger\right\}
\end{aligned}
$$

$$l_{CM}^z + l_{rel}^z = l_1^z + l_2^z$$

$$l_1^z = \hbar\left(a_1^\dagger a_1 - b_1^\dagger b_1\right), \qquad l_2^z = \hbar\left(a_2^\dagger a_2 - b_2^\dagger b_2\right)$$

$$l^z_{rel} = \frac{\hbar}{2}\left[(a_1^\dagger - a_2^\dagger)(a_1 - a_2) - (b_1^\dagger - b_2^\dagger)(b_1 - b_2)\right]$$

$$l^z_{CM} = \frac{\hbar}{2}\left[(a_1^\dagger + a_2^\dagger)(a_1 + a_2) - (b_1^\dagger + b_2^\dagger)(b_1 + b_2)\right]$$

$$\rightarrow l^z_{rel} = \hbar\left[\alpha_-^\dagger\alpha_- - \beta_-^\dagger\beta_-\right] \qquad l^z_{CM} = \hbar\left[\alpha_+^\dagger\alpha_+ - \beta_+^\dagger\beta_+\right]$$

$$\begin{aligned} \alpha_-^\dagger &= \frac{1}{\sqrt{2}}(a_1^\dagger - a_2^\dagger), \qquad & \beta_-^\dagger &= \frac{1}{\sqrt{2}}(b_1^\dagger - b_2^\dagger) \\ \alpha_+^\dagger &= \frac{1}{\sqrt{2}}(a_1^\dagger + a_2^\dagger), \qquad & \beta_+^\dagger &= \frac{1}{\sqrt{2}}(b_1^\dagger + b_2^\dagger) \end{aligned}$$

$$\begin{aligned} \psi &= (z_1 + z_2)^M (z_1 - z_2)^m \mathrm{e}^{-\sum_i |z_i|^2/4l_B^2} \\ &\propto \left(\beta_+^\dagger\right)^M \left(\beta_-^\dagger\right)^m |0\rangle \end{aligned}$$

If one considers the angular momentum in the $-z$ direction,

$$\begin{aligned} \tilde{l}^z_{rel}|\psi\rangle &= m \quad (\hbar = 1) \\ \tilde{l}^z_{CM}|\psi\rangle &= M \quad (\hbar = 1) \\ \tilde{l}^z_{tot}|\psi\rangle &= M + m \quad (\hbar = 1) \end{aligned}$$

# Chapter 4

# Electron-phonon interaction

## 4.1 Introduction

In band theory, the lattice is assumed to be static. In reality, ions vibrate around their equilibrium position. These vibrations have direct consequences on the properties of solids (specific heat, fusion, thermal conductivity) and indirect consequences due to their influence on electrons. In particular, they contribute to the resistivity, and they are responsible for the effective attractive interaction between electrons that lies at the root of superconductivity.

## 4.2 Phonons - Reminder

Let us denote by $\vec{R}_{n,\alpha} = \vec{R}_n + \vec{R}_\alpha$ the equilibrium position of an atom, where $\vec{R}_n$ is the position of the unit cell and $\vec{R}_\alpha$ the position inside the unit cell, and let us denote by $\vec{s}_{n,\alpha}(t)$ the displacement of this atom with respect to its equilibrium position. The kinetic energy can be written:

$$T = \sum_{n,\alpha,i} \frac{M_\alpha}{2} \dot{s}^2_{n,\alpha,i} \ , \quad n = 1, \ldots, N \ ; \quad \alpha = 1, \ldots, r \ ; \quad i = x,\ y,\ z$$

where $M_\alpha$ is the mass of the atom $\alpha$ in the unit cell. Furthermore let us denote by $V(\{\vec{s}_{n,\alpha}\})$ the potential energy of a system considered as a function of the positions of the ions. If one is interested in the small displacements with respect to the equilibrium position, one can expand this function around the equilibrium position. Up to second order, one gets:

$$V(\{\vec{s}_{n,\alpha}\}) - V(\{\vec{o}\}) = \frac{1}{2} \sum_{n,\alpha,i} \sum_{n',\alpha',i'} \phi^{n',\alpha',i'}_{n,\alpha,i} s_{n,\alpha,i} s_{n',\alpha',i'}$$

with

$$\phi^{n',\alpha',i}_{n,\alpha,i} = \left. \frac{\partial^2 V}{\partial s_{n,\alpha,i} \partial s_{n',\alpha',i'}} \right|_{\vec{s}_{n,\alpha} = \vec{o}}$$

The vibrations of the solid thus appear as a particular case of the general problem of small oscillations treated in analytical mechanics. In particular, with the help of a change of variables, one can map the problem onto that of a collection of harmonic oscillators. These variables are called the normal coordinates. In the case of the vibrations of a solid, the periodicity of the problem allows one to simplify the problem by performing a Fourier transform. The price to pay is to introduce complex coordinates.

Up to a constant, the Lagrangian of the system is thus given by:

$$L = T - V = \sum_{n,\alpha,i} \frac{M_\alpha}{2} \dot{s}^2_{n,\alpha,i} \; - \frac{1}{2} \sum_{n,\alpha,i} \sum_{n',\alpha',i'} \phi^{n',\alpha',i'}_{n,\alpha,i} \, s_{n,\alpha,i} s_{n',\alpha',i'}$$

The Lagrange equations for the variable $s_{n,\alpha,i}$ take the form:

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{s}_{n,\alpha,i}}\right) - \frac{\partial L}{\partial s_{n,\alpha,i}} = 0$$

$$\Rightarrow M_\alpha \ddot{s}_{n,\alpha,i} = - \sum_{n',\alpha',i'} \phi^{n',\alpha',i'}_{n,\alpha,i} \, s_{n',\alpha',i'}$$

We are interested in the eigenmodes of this system, i.e. in solutions of the type

$$s_{n,\alpha,i}(t) = \frac{1}{\sqrt{M_\alpha}} \; u_{n,\alpha,i} \; e^{-i\omega t}$$

Since this is a linear system, both the real part and the imaginary part are solutions. One gets:

$$\omega^2 u_{n,\alpha,i} = \sum_{n',\alpha',i'} D^{n',\alpha',i'}_{n,\alpha,i} \, u_{n',\alpha',i'}$$

with

$$D^{n',\alpha',i'}_{n,\alpha,i} = \frac{\phi^{n',\alpha',i'}_{n,\alpha,i}}{\sqrt{M_\alpha M_{\alpha'}}}$$

The squares of the eigenfrequencies are thus the eigenvalues of the matrix $D$. Now, this matrix is symmetric since $\phi^{n',\alpha',i'}_{n,\alpha,i}$ are second order partial derivatives. Its eigenvalues are thus real. Besides, the condition that the energy is minimal (and not just extremal) implies that the eigenvalues are positive. Let us denote by $\omega_j > 0$ the square root of $\omega_j^2$. We thus have:

$$3 \times r \times N \quad \text{eigenfrequencies}$$

where the factor 3 comes from $i = x, y, z$ , $r$ is the number of atoms in the unit cell, and $N$ is the number of unit cells in the crystal. The eigenvector associated to $\omega_j^2$ is denoted by $u^{(j)}_{n,\alpha,i}$.

So far, the calculation is completely general. The translational invariance of the system implies that $\phi_{n,\alpha,i}^{n',\alpha',i'}$ depends only on $n-n'$. In such a case, one can look for the solutions as

$$u_{n,\alpha,i} = c_{\alpha,i} e^{i\vec{q}\cdot\vec{R}_n}$$

$$\begin{aligned}\Rightarrow \omega^2 c_{\alpha,i} &= \sum_{\alpha',i'}\sum_{n'} \frac{1}{\sqrt{M_\alpha M_{\alpha'}}}\phi_{\alpha,i,n}^{\alpha',i',n'} e^{i\vec{q}\cdot(\vec{R}_{n'}-\vec{R}_n)} c_{\alpha',i'} \\ &= \sum_{\alpha',i'} D_{\alpha,i}^{\alpha',i'}(\vec{q}) c_{\alpha',i'}\end{aligned}$$

For each value of $\vec{q}$, we thus have a system of $3r$ equations, which leads to $3r$ eigenfrequencies:

$$\omega_j(\vec{q}), \quad j = 1, \ldots, 3r$$

The eigenvector associated to $\omega_j(\vec{q})$ has $3r$ components $e_{\alpha,i}^{(j)}(\vec{q})$. It is often written as $r$ vectors with three components $\vec{e}_\alpha^{(j)}(\vec{q})$. These vectors are called the polarization vectors.

The final solution for the displacement is thus a linear combination of the displacements $\vec{e}_\alpha^{(j)}(\vec{q}) e^{i(\vec{q}\cdot\vec{R}_n - \omega_j(\vec{q})t)}$. It is convenient to include the time dependence into the coefficient, and write this linear combination as:

$$\Rightarrow s_{n,\alpha,i}(t) = \frac{1}{\sqrt{NM_\alpha}} \sum_{j,\vec{q}} Q_j(\vec{q},t) e_{\alpha,i}^{(j)}(\vec{q}) e^{i\vec{q}\cdot\vec{R}_n}$$

There are in total $3rN$ quantities $Q_j(\vec{q},t)$ ($3r$ for $j$, $N$ for $\vec{q}$). One can consider them as a new set of variables. In order for the $s_{n,\alpha,i}(t)$ to be real, the $Q_j(\vec{q},t)$ must be complex and satisfy $Q_j^*(\vec{q},t) = Q_j(-\vec{q},t)$ (one can show that $\omega_j(-\vec{q}) = \omega_j(\vec{q})$ and choose $\vec{e}_\alpha^{(j)}(\overrightarrow{-q}) = \vec{e}_\alpha^{(j)}(\vec{q})$).

Besides, the eigenvectors of a symmetric matrix are orthogonal:

$$\sum_{\alpha,i} e_{\alpha,i}^{(j)}(\vec{q}) e_{\alpha,i}^{(j')}(\vec{q}') = \delta_{j,j'}\delta_{\vec{q},\vec{q}'}$$

With the help of these relations one can show that the Lagrangian can be written:

$$L = \frac{1}{2}\sum_{j,\vec{q}} \left[\dot{Q}_j^*(\vec{q},t)\dot{Q}_j(\vec{q},t) - \omega_j^2(\vec{q}) Q_j^*(\vec{q},t) Q_j(\vec{q},t)\right]$$

The Hamiltonian can then be obtained with the help of a Legendre transformation:

$$P_j(\vec{q},t) = \frac{\partial L}{\partial \dot{Q}_j(\vec{q},t)} = \dot{Q}_j^*(\vec{q},t)$$

$$\Rightarrow H = \frac{1}{2}\sum_{j,\vec{q}} \left[P_j^*(\vec{q},t) P_j(\vec{q},t) + \omega_j^2 Q_j^*(\vec{q},t) Q_j(\vec{q},t)\right]$$

Quantization :

One postulates, as for the harmonic oscillator, the usual commutation relations between the positions and the momenta of the ions. This leads to the following commutation relations for the normal coordinates:

$$[Q_j(\vec{q}), P_{j'}(\vec{q}')] = i\hbar \delta_{\vec{q},\vec{q}'} \delta_{j,j'}$$

If, still by analogy with the harmonic oscillator, one defines creation and annihilation operators by:

$$\begin{cases} a_j^+(\vec{q}) = \frac{1}{\sqrt{2\hbar\omega_j(\vec{q})}}(\omega_j Q_j^*(\vec{q}) - iP_j(\vec{q})) \\ a_j(\vec{q}) = \frac{1}{\sqrt{2\hbar\omega_j(\vec{q})}}(\omega_j Q_j(\vec{q}) - iP_j^*(\vec{q})) \end{cases}$$

the Hamiltonian takes the form:

$$H = \sum_{j,\vec{q}} \hbar\omega_j(\vec{q})\left(a_j^+(\vec{q})a_j(\vec{q}) + \frac{1}{2}\right)$$

The operators $a_j^+(\vec{q})$, $a_j(\vec{q})$ satisfy bosonic commutation rules:

$$[a_j(\vec{q}), a_{j'}^+(\vec{q}')] = \delta_{j,j'}\delta_{\vec{q},\vec{q}'}$$

The inverse relations giving the positions and momenta in terms of the $a_j(\vec{q})$ and $a_j^+(\vec{q})$ are:

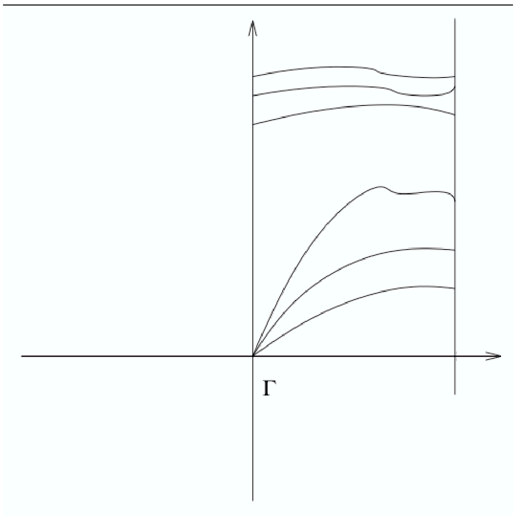
$$\begin{cases} Q_j(\vec{q}) = \sqrt{\frac{\hbar}{2\omega_j(\vec{q})}}(a_j^+(-\vec{q}) + a_j(\vec{q})) \\ P_j(\vec{q}) = i\sqrt{\frac{\hbar\omega_j(\vec{q})}{2}}(a_j^+(\vec{q}) - a_j(-\vec{q})) \end{cases}$$

The form of the Hamiltonian shows that the system behaves as a gas of bosons.

General discussion:

The curves $\omega_j(\vec{q})$ are called the dispersion relations. There are $3r$ of them. When $\vec{q} \to 0$, all the unit cells vibrate in phase. If all atoms inside the unit cells vibrate in phase, $\omega_j(\vec{q})$ must go to 0. There are 3 modes that satisfy this condition (global displacement in the three directions of space). These modes are called acoustic modes. The remaining $3r - 3$ modes do not go to 0 as $\vec{q} \to 0$. They are called optical modes.

The general form of the dispersion is thus:



Thermodynamics

The calculation of the specific heat at low temperture is completely equivalent to that of antiferromagnets since acoustic phonons have a linear dispersion, like magnons in an antiferromagnet. It leads to

$$C_v \sim T^3$$

A high temperature, the specific heat saturates when the temperature becomes larger than the largest phonon frequency.

Other properties:

For successful that it is, this harmonic model suffers from a number of failures. In particular, it is impossible to account for thermal expansion and thermal conductivity without including anharmonic effects. In second quantization, they take the form of terms with 3 phonons, 4 phonons, etc...

## 4.3 Electron-phonon interaction

The electron-phonon interaction comes from the Coulomb interaction between the electrons and the ions, that can be written:

$$H_{el-ion} = \sum_{l,i} V(\vec{r}_l - \vec{R}_i)$$

where $\vec{r}_l$ stands for the position of the electron and $\vec{R}_i$ for that of the ion, and from the fact that $\vec{R}$ is not fixed at its equilibrium position but can make small oscillations around it. With the help of the previous notations, one can write:

$$\vec{R}_{n,\alpha}(t) = \underbrace{\vec{R}_n + \vec{R}_\alpha}_{\vec{R}_{n,\alpha}} + \vec{s}_{n,\alpha}(t)$$

If $\vec{s}_{n,\alpha}(t)$ is small, one can expand $V_\alpha(\vec{r}_l - \vec{R}_i)$, where $V$ depends on the site in the unit cell:

$$V_\alpha(\vec{r}_l - \vec{R}_{n,\alpha} - \vec{s}_{n,\alpha}) = V_\alpha(\vec{r}_l - \vec{R}_{n,\alpha}) - \vec{s}_{n,\alpha} \cdot \vec{\nabla} V_\alpha(\vec{r}_l - \vec{R}_{n,\alpha})$$

The first term describes the motion of electrons in the periodic potential of the crystal. It is taken care of by Bloch theorem.

The second term describes the interaction with the lattice vibrations. If one puts the expression of $\vec{s}_{n,\alpha}$ in terms of the normal coordinates in the expression of the Hamiltonian, one gets:

$$H_{el-ph} = -\sum_{\alpha,n,l} \frac{1}{\sqrt{NM_\alpha}} \sum_{j,\vec{q}} Q_j(\vec{q}) \vec{e}_\alpha^{(j)}(\vec{q}) \cdot \vec{\nabla} V_\alpha(\vec{r}_l - \vec{R}_{n,\alpha}) e^{i\vec{q}\cdot\vec{R}_n}$$

To express this Hamiltonian in second quantization, one proceeds in two steps:

- For the phonon part, one just has to replace $Q_j(\vec{q})$ by its expression in terms of $a_j$ and $a_j^+$:

$$Q_j(\vec{q}) = \sqrt{\frac{\hbar}{2\omega_j(\vec{q})}} \Big( a_j^+(-\vec{q}) + a_j(\vec{q}) \Big)$$

- For the electron part, one must use the rules of second quantizaiton, i.e. one must calculate the matrix element of the part of the interaction that depends on the electron coordinates, namely $\vec{\nabla} V_\alpha(\vec{r}_l - \vec{R}_{n,\alpha})$. For this, it is useful to expand $v_\alpha$ in Fourier series :

$V_\alpha(\vec{r}) = \sum_{\vec{K}} e^{i\vec{K}\cdot\vec{r}} V_{\alpha,\vec{K}}$ $\quad$ $\vec{K}$ arbitrary because $V_\alpha(\vec{r})$ is not periodic

$$\Rightarrow \vec{\nabla} V_\alpha(\vec{r}_l - \vec{R}_{n,\alpha}) = \sum_{\vec{K}} e^{i\vec{K}(\vec{r}_l - \vec{R}_{n,\alpha})} V_{\alpha,\vec{K}} i\vec{K}$$

The matrix elements of this potential between Bloch waves are given by:

$$< \vec{k}', \sigma' | \vec{\nabla} V_\alpha | \vec{k}, \sigma > = \sum_{\vec{K}} e^{-i\vec{K}\vec{R}_{n,\alpha}} V_{\alpha,\vec{K}} i\vec{K} < \vec{k}', \sigma' | e^{i\vec{K}\vec{r}} | \vec{k}, \sigma >$$

with

$$< \vec{k}', \sigma' | e^{i\vec{K}\vec{r}} | \vec{k}, \sigma > = \int d\vec{r} u^*_{\vec{k}'}(\vec{r}) u_{\vec{k}}(\vec{r}) e^{i(-\vec{k}' + \vec{K} + \vec{k})\vec{r}} \delta_{\sigma,\sigma'}$$

This integral vanishes unless

$$\vec{k}' = \vec{K} + \vec{k} + \vec{G}$$

where $\vec{G}$ is a vector of the reciprocal lattice. Indeed, $u^*_{\vec{k}'}(\vec{r})u_{\vec{k}}(\vec{r})$ is a periodic function. It can thus be expanded in Fourier series:

$$u^*_{\vec{k}'}(\vec{r})u_{\vec{k}}(\vec{r}) = \sum_{\vec{G}} f(\vec{G})e^{i\vec{G}.\vec{r}}$$

where the vectors $\vec{G}$ are vectors of the reciprocal lattice, which leads to

$$< \vec{k}', \sigma'|e^{i\vec{K}\vec{r}}|\vec{k}, \sigma >= \delta_{\sigma,\sigma'} \sum_{\vec{G}} f(\vec{G}) \int d\vec{r} e^{i(-\vec{k}'+\vec{K}+\vec{k}+\vec{G})\vec{r}}$$

Now, the integral $\int d\vec{r} e^{i(-\vec{k}'+\vec{K}+\vec{k}+\vec{G}).\vec{r}}$ vanishes unless $-\vec{k}' + \vec{K} + \vec{k} + \vec{G} = \vec{0}$.

Moreover, the sum over $\vec{R}_n$ can be written:

$$\sum_{\vec{R}_n} e^{i\vec{q}.\vec{R}_n} e^{-i\vec{K}.\vec{R}_n} = \sum_{\vec{R}_n} e^{i(\vec{q}-\vec{K}).\vec{R}_n}$$

This sum vanishes unless $\vec{K} = \vec{q} + \vec{G}$ where $\vec{G}$ is a vector of the reciprocal lattice.

As a consequence, $\vec{k}$ and $\vec{k}'$ must be related by

$$\vec{k}' = \vec{q} + \vec{k} + \vec{G}$$

But $\vec{k}$ and $\vec{k}'$ must both belong to the first Brillouin zone. So, for a given pair of vectors $\vec{k}$ and $\vec{q}$, there is one and only one vector $\vec{G}$ that brings $\vec{k} + \vec{q}$ back into the first Brillouin zone.

One thus distinguishes two types of processes:

- The "normal" processes such that $\vec{k} + \vec{q} \in$ BZ ($\vec{G} = \vec{0}$).
- The "Umklapp" processes such that $\vec{k} + \vec{q} \notin$ BZ $\Rightarrow \vec{G} \neq \vec{0}$ and $\vec{k}' \neq \vec{k} + \vec{q}$.

Finally, the interaction can be written

$$\begin{aligned} H &= - \sum_{\alpha,\sigma,\vec{k},\vec{q},j} \sqrt{\frac{N\hbar}{2M_\alpha \omega_j(\vec{q})}} \\ &\times \sum_{\vec{G}'} \left[ \vec{e}^{(j)}_{\alpha,\vec{q}} \cdot i(\vec{q} + \vec{G}') \right] \int V_\alpha(\vec{r}')e^{i(\vec{q}+\vec{G}')\cdot(\vec{r}'-\vec{R}_\alpha)} d\vec{r}' \\ &\times \int u^*_{\vec{k}+\vec{q}+\vec{G}}(\vec{r})u_{\vec{k}}(\vec{r})d\vec{r} \\ &\times \left[ a^+_{-\vec{q}}(j) + a_{\vec{q}}(j) \right] c^+_{\vec{k}+\vec{q}+\vec{G},\sigma} c_{\vec{k},\sigma} \end{aligned}$$

where $\vec{G}$ is such that $\vec{k}+\vec{q}+\vec{G} \in 1^{\text{st}}$ Brillouin zone. This Hamiltonian can be written in a more compact way as:

$$H = -\sum_{\vec{k},\vec{k}',\vec{q},j,\sigma} M(\vec{k},\vec{k}',\vec{q},j)\left[a^+_{-\vec{q}}(j) + a_{\vec{q}}(j)\right] c^+_{\vec{k}',\sigma} c_{\vec{k},\sigma}$$

with $\vec{q} = \vec{k}' - \vec{k} + \vec{G}$

$$\begin{cases} \vec{G} = \vec{0} & \text{for normal processes} \\ \vec{G} \neq \vec{0} & \text{for Umklapp processes} \end{cases}$$

## 4.4 The effective electron-electron interaction

One of the consequences of the electron-phonon interaction is its important contribution to resistivity. This assumes that the electrons can be describes by a Fermi-Dirac distribution, or equivalently that one essentially deals with free electrons.

But the ions are charged particles, and the electron-electron interaction must be modified by their vibrations. The calculation of the effective electron-electron interaction in the presence of phonons can be done from two points of view:

- The phonons contribute to the dielectric function, which induces a modification of the electron-electron interaction.
- The electrons interact via the exchange of virtual phonons.

We will adopt the second point of view to do the calculation.

The starting point is the Hamiltonian describing the electron-phonon interaction:

$$H_{el-ph} = \sum_{\vec{k},\vec{q},\sigma} M_{\vec{q}}\left[a^+_{-\vec{q}} + a_{\vec{q}}\right] c^+_{\vec{k}+\vec{q},\sigma} c_{\vec{k},\sigma}$$

where for simplicity we only consider one branch of phonons, and where we limit ourselves to normal processes (can be justified a posteriori). Furthermore, we assume that the electrons are described by a Hamiltonian of the type :

$$H_{el} = \sum_{\vec{k},\sigma} \epsilon_{\vec{k}} c^+_{\vec{k},\sigma} c_{\vec{k},\sigma}$$

and that phonons are described by the Hamiltonian

$$H_{ph} = \sum_{\vec{q}} \hbar\omega_{\vec{q}}(a^+_{\vec{q}} a_{\vec{q}} + \frac{1}{2})$$

If the matrix elements $M_{\vec{q}}$ are small, we can expect the electron gas to be described by $H_0 = H_{el}$ plus a small correction. But if $|\psi\rangle$ is an eigenstate of $H_{el} + H_{ph}$, then

$\langle\psi|H_{el-ph}|\psi\rangle = 0$. So if we want to treat $H_{el-ph}$ as a perturbation, we must go to second order. This is reminiscent of the calculation done for Mott insulators, where degenerate perturbation theory has been used to obtain an effective Hamiltonian, the Heisenberg model with a coupling constant $J = \frac{4t^2}{U}$ of order 2 in the kinetic term. In the present case however, the starting point is not degenerate, and one must resort to another technique to derive an effective Hamiltonian for the electron gas that takes into account the electron-phonon interaction.

By definition, an effective Hamiltonian is a Hamiltonian that must have the same spectrum of eigenvalues as the original one. Now there is a systematic method to generate Hamiltonians having the same eigenvalues as a given Hamiltonian known as canonical transformation.

Canonical transformation

A canonical transformation is a unitary transformation of the Hamiltonian and of the vectors of the Hilbert space that does not modify the spectrum:

$$\begin{array}{ccc} \hat{H} & \to & U\hat{H}U^{-1} \\ |\Psi> & \to & U|\Psi> \end{array}$$

with $U^+ = U^{-1}$.

The following properties are then satisfied:

1. The new Hamiltonian is hermitian. Indeed,
   $(U\hat{H}U^{-1})^+ = (U^{-1})^+\hat{H}U^+ = U\hat{H}U^{-1}$ QED
2. If $|\Psi>$ is an eigenstate of $\hat{H}$ with eigenvalue $E$, then $U|\Psi>$ is an eigenstate of $U\hat{H}U^{-1}$ with the same eigenvalue. Indeed,
   $U\hat{H}U^{-1}U|\Psi> = U\hat{H}|\Psi> = UE|\Psi> = EU|\Psi>$.

The property $U^+ = U^{-1}$ is satisfied by operators of the form:

$$U = e^{iS}$$

provided $S$ is hermitian. Indeed,

$$U^+ = (e^{iS})^+ = e^{-iS^+} = e^{-iS} = (e^{iS})^{-1}$$

The transformed Hamiltonian can be written as an expansion in terms of commutators:

$$\begin{aligned} e^{iS}\hat{H}e^{-iS} &= \hat{H} + i[S,\hat{H}] + \frac{i^2}{2!}[S,[S,\hat{H}]] + \ldots \\ &= \hat{H} + \sum_{n=1}^{\infty}\frac{i^n}{n!}[S,[S,\ldots[S,\hat{H}]]] \end{aligned}$$

Proof :

Let us define the operator $\hat{H}(\lambda)$ by :

$$\hat{H}(\lambda) \equiv e^{i\lambda S}\hat{H}e^{-i\lambda S}$$

Then

$$\begin{aligned}
\frac{d\hat{H}(\lambda)}{d\lambda} &= iSe^{i\lambda S}\hat{H}e^{-i\lambda S} + e^{i\lambda S}\hat{H}(-iS)e^{-i\lambda S} \\
&= i[S, \hat{H}(\lambda)] \\
\frac{d^2\hat{H}(\lambda)}{d\lambda^2} &= i[S, \frac{d\hat{H}(\lambda)}{d\lambda}] = i^2[S, [S, \hat{H}(\lambda)]]
\end{aligned}$$

If $\hat{H}(\lambda)$ has an expansion in $\lambda$, it can thus be written :

$$\hat{H}(\lambda) = \hat{H} + i\lambda[S, \hat{H}(0)] + \frac{i^2}{2!}\lambda^2[S, [S, \hat{H}(0)]] + \ldots$$

For $\lambda = 1$, this leads to the expansion we are looking for since $\hat{H}(0) = \hat{H}$.

Let us come back to the original problem. One looks for a description of the electron gas interacting with phonons that does not contain terms of order 1 in $M_{\vec{q}}$ any more, and that contains purely electronic terms of order 2.

Looking at the expansion up to order 2,

$$\tilde{H} = H + i[S, H] - \frac{1}{2!}[S, [S, H]]$$

$S$ must satisfy two conditions:

- It is or order 1 in $M_{\vec{q}}$
- It is such that $[S, H]$ cancels $H_{el-ph}$ to first order.

To find the solution of

$$i[S, H_{el} + H_{ph}] = -H_{el-ph}$$

let us start by calculating the commutator of a term of $H_{el-ph}$ with $H_{el} + H_{ph}$.

$$\begin{aligned}
[a^+_{-\vec{q}}c^+_{\vec{k}+\vec{q}}c_{\vec{k}}, c^+_l c_l] &= a^+_{-\vec{q}}[c^+_{\vec{k}+\vec{q}}c_{\vec{k}}, c^+_l c_l] \\
&= a^+_{-\vec{q}}\Big\{\delta_{\vec{k},l}\Big(c^+_{\vec{k}+\vec{q}}c_{\vec{k}}c^+_{\vec{k}}c_{\vec{k}} - c^+_{\vec{k}}c_{\vec{k}}c^+_{\vec{k}+\vec{q}}c_{\vec{k}}\Big) \\
&+ \delta_{\vec{k}+\vec{q},l}\Big(c^+_{\vec{k}+\vec{q}}c_{\vec{k}}c^+_{\vec{k}+\vec{q}}c_{\vec{k}+\vec{q}} - c^+_{\vec{k}+\vec{q}}c_{\vec{k}+\vec{q}}c^+_{\vec{k}+\vec{q}}c_{\vec{k}}\Big)\Big\} \\
&= a^+_{-\vec{q}}\Big\{\delta_{\vec{k},l}c^+_{\vec{k}+\vec{q}}c_{\vec{k}} - \delta_{\vec{k}+\vec{q},l}c^+_{\vec{k}+\vec{q}}c_{\vec{k}}\Big\}
\end{aligned}$$

$$\Rightarrow [a^+_{-\vec{q}} c^+_{\vec{k}+\vec{q}} c_{\vec{k}}, H_{el}] = a^+_{-\vec{q}} (\epsilon_{\vec{k}} - \epsilon_{\vec{k}+\vec{q}}) c^+_{\vec{k}+\vec{q}} c_{\vec{k}}$$

Similarly

$$[a_{\vec{q}} c^+_{\vec{k}+\vec{q}} c_{\vec{k}}, H_{el}] = a_{\vec{q}} (\epsilon_{\vec{k}} - \epsilon_{\vec{k}+\vec{q}}) c^+_{\vec{k}+\vec{q}} c_{\vec{k}}$$

Moreover,

$$\begin{aligned} [a^+_{-\vec{q}}, a^+_{\vec{k}} a_{\vec{k}}] &= -\delta_{\vec{k},-\vec{q}} a^+_{-\vec{q}} \\ \Rightarrow [a^+_{-\vec{q}}, H_{ph}] &= -\hbar\omega_{-\vec{q}} a^+_{-\vec{q}} \end{aligned}$$

Similarly,

$$\begin{aligned} [a_{\vec{q}}, a^+_{\vec{k}} a_{\vec{k}}] &= \delta_{\vec{k},\vec{q}} a_{\vec{q}} \\ \Rightarrow [a_{\vec{q}}, H_{ph}] &= \hbar\omega_{\vec{q}} a_{\vec{q}} \end{aligned}$$

Finally,

$$[a^+_{-\vec{q}} c^+_{\vec{k}+\vec{q}} c_{\vec{k}}, H_{el} + H_{ph}] = (\epsilon_{\vec{k}} - \epsilon_{\vec{k}+\vec{q}} - \hbar\omega_{\vec{q}}) a^+_{-\vec{q}} c^+_{\vec{k}+\vec{q}} c_{\vec{k}}$$

$$[a_{\vec{q}} c^+_{\vec{k}+\vec{q}} c_{\vec{k}}, H_{el} + H_{ph}] = (\epsilon_{\vec{k}} - \epsilon_{\vec{k}+\vec{q}} + \hbar\omega_{\vec{q}}) a_{\vec{q}} c^+_{\vec{k}+\vec{q}} c_{\vec{k}}$$

The solution of the equation

$$[S, H_{el} + H_{ph}] = iH_{el-ph}$$

can thus be written

$$\begin{aligned} S &= i \sum_{\vec{k},\vec{q},\sigma} \frac{M_{\vec{q}}}{\epsilon_{\vec{k}} - \epsilon_{\vec{k}+\vec{q}} - \hbar\omega_{\vec{q}}} a^+_{-\vec{q}} c^+_{\vec{k}+\vec{q},\sigma} c_{\vec{k},\sigma} \\ &+ i \sum_{\vec{k},\vec{q},\sigma} \frac{M_{\vec{q}}}{\epsilon_{\vec{k}} - \epsilon_{\vec{k}+\vec{q}} + \hbar\omega_{\vec{q}}} a_{\vec{q}} c^+_{\vec{k}+\vec{q},\sigma} c_{\vec{k},\sigma} \end{aligned}$$

(the spin has been added at the end of the calculation).

The operator $S$ satisfies the two conditions. The first terms in powers of $M_{\vec{q}}$ are thus of order 2. They are given by :

$$\begin{aligned} i & \quad [S, H_{el-ph}] - \frac{1}{2}[S, \underbrace{[S, H_{el} + H_{ph}]}_{iH_{el-ph}}] \\ &= \frac{i}{2}[S, H_{el-ph}] = \frac{1}{2}[H_{el-ph}, \frac{S}{i}] \end{aligned}$$

The only thing that remains to be done is to calculate the commutator $[S, H_{el-ph}]$. From the expression of $S$ and of $H_{el-ph}$, it is clear that the generic term of this commutator is of the form:

$$a^{(+)}_{\pm\vec{q}} a^{(+)}_{\pm\vec{q}'} c^+_{\vec{k}'+\vec{q}',\sigma'} c_{\vec{k}',\sigma'} c^+_{\vec{k}+\vec{q},\sigma} c_{\vec{k},\sigma}$$

Since our goal is to derive an effective interaction between electrons, we will only keep terms that only involve fermionic operators. The other terms have a vanishing expectation value in the vacuum and do not contribute in the limit $T \to 0$.

Let us write:

$$\begin{aligned} H_{el-ph} &= \sum_{\vec{k}_1,\vec{q}_1,\sigma_1} M_{\vec{q}_1} a^+_{-\vec{q}_1} c^+_{\vec{k}_1+\vec{q}_1,\sigma_1} c_{\vec{k}_1,\sigma_1} \\ &+ \sum_{\vec{k}_1,\vec{q}_1,\sigma_1} M_{\vec{q}_1} a_{\vec{q}_1} c^+_{\vec{k}_1+\vec{q}_1,\sigma_1} c_{\vec{k}_1,\sigma_1} \end{aligned}$$

$$\begin{aligned} \frac{S}{i} &= \sum_{\vec{k}_2,\vec{q}_2,\sigma_2} \frac{M_{\vec{q}_2}}{\epsilon_{\vec{k}_2} - \epsilon_{\vec{k}_2+\vec{q}_2} - \hbar\omega_{\vec{q}_2}} a^+_{-\vec{q}_2} c^+_{\vec{k}_2+\vec{q}_2,\sigma_2} c_{\vec{k}_2,\sigma_2} \\ &+ \sum_{\vec{k}_2,\vec{q}_2,\sigma_2} \frac{M_{\vec{q}_2}}{\epsilon_{\vec{k}_2} - \epsilon_{\vec{k}_2+\vec{q}_2} + \hbar\omega_{\vec{q}_2}} a_{\vec{q}_2} c^+_{\vec{k}_2+\vec{q}_2,\sigma_2} c_{\vec{k}_2,\sigma_2} \end{aligned}$$

Two terms lead to an expression that does not depend on $(a^+, a)$ :

$$\begin{aligned} & [a^+_{-\vec{q}_1} c^+_{\vec{k}_1+\vec{q}_1,\sigma_1} c_{\vec{k}_1,\sigma_1}, a_{-\vec{q}_1} c^+_{\vec{k}_2-\vec{q}_1,\sigma_2} c_{\vec{k}_2,\sigma_2}] \\ =\ & a^+_{-\vec{q}_1} a_{-\vec{q}_1} c^+_{\vec{k}_1+\vec{q}_1,\sigma_1} c_{\vec{k}_1,\sigma_1} c^+_{\vec{k}_2-\vec{q}_1,\sigma_2} c_{\vec{k}_2,\sigma_2} \\ -\ & \underbrace{a_{-\vec{q}_1} a^+_{-\vec{q}_1}}_{1+a^+_{-\vec{q}_1} a_{-\vec{q}_1}} c^+_{\vec{k}_2-\vec{q}_1,\sigma_2} c_{\vec{k}_2,\sigma_2} c^+_{\vec{k}_1+\vec{q}_1,\sigma_1} c_{\vec{k}_1,\sigma_1} \\ =\ & a^+_{-\vec{q}_1} a_{-\vec{q}_1} [c^+_{\vec{k}_1+\vec{q}_1,\sigma_1} c_{\vec{k}_1,\sigma_1}, c^+_{\vec{k}_2-\vec{q}_1,\sigma_2} c_{\vec{k}_2,\sigma_2}] - c^+_{\vec{k}_2-\vec{q}_1,\sigma_2} c_{\vec{k}_2,\sigma_2} c^+_{\vec{k}_1+\vec{q}_1,\sigma_1} c_{\vec{k}_1,\sigma_1} \end{aligned}$$

The first term can be neglected at zero temperature, and the second one is purely electronic. Similarly,

$$\begin{aligned} & [a_{\vec{q}_1} c^+_{\vec{k}_1+\vec{q}_1,\sigma_1} c_{\vec{k}_1,\sigma_1}, a^+_{\vec{q}_1} c^+_{\vec{k}_2-\vec{q}_1,\sigma_2} c_{\vec{k}_2,\sigma_2}] \\ =\ & a^+_{\vec{q}_1} a_{\vec{q}_1} [c^+_{\vec{k}_1+\vec{q}_1,\sigma_1} c_{\vec{k}_1,\sigma_1}, c^+_{\vec{k}_2-\vec{q}_1,\sigma_2} c_{\vec{k}_2,\sigma_2}] + c^+_{\vec{k}_1+\vec{q}_1,\sigma_1} c_{\vec{k}_1,\sigma_1} c^+_{\vec{k}_2-\vec{q}_1,\sigma_2} c_{\vec{k}_2,\sigma_2} \end{aligned}$$

$$\begin{aligned} \Rightarrow H^{(2)}_{eff} &= \frac{1}{2} \sum_{\vec{k}_1,\vec{k}_2,\vec{q},\sigma_1,\sigma_2} \left( \frac{M^2_{\vec{q}}}{\epsilon_{\vec{k}_2} - \epsilon_{\vec{k}_2-\vec{q}} + \hbar\omega_{\vec{q}}} (-c^+_{\vec{k}_2-\vec{q},\sigma_2} c_{\vec{k}_2,\sigma_2} c^+_{\vec{k}_1+\vec{q},\sigma_1} c_{\vec{k}_1,\sigma_1}) \right. \\ &+ \left. \frac{M^2_{\vec{q}}}{\epsilon_{\vec{k}_2} - \epsilon_{\vec{k}_2-\vec{q}} - \hbar\omega_{\vec{q}}} (c^+_{\vec{k}_1+\vec{q},\sigma_1} c_{\vec{k}_1,\sigma_1} c^+_{\vec{k}_2-\vec{q},\sigma_2} c_{\vec{k}_2,\sigma_2}) \right) \end{aligned}$$

If one rearranges the creation and annihilation operators to put them in the same order (the bilinear terms generated on the way can be included in $\epsilon_{\vec{k}}$), one gets:

$$H^{(2)}_{eff} = \frac{1}{2} \sum_{\vec{k}_1,\vec{k}_2,\vec{q},\sigma_1,\sigma_2} V_{\vec{k}_2,\vec{q}} c^+_{\vec{k}_1+\vec{q},\sigma_1} c_{\vec{k}_1,\sigma_1} c^+_{\vec{k}_2-\vec{q},\sigma_2} c_{\vec{k}_2,\sigma_2}$$

with

$$\begin{aligned} V_{\vec{k}_2,\vec{q}} &= M_{\vec{q}}^2 \left( \frac{1}{\epsilon_{\vec{k}_2} - \epsilon_{\vec{k}_2-\vec{q}} - \hbar\omega_{\vec{q}}} - \frac{1}{\epsilon_{\vec{k}_2} - \epsilon_{\vec{k}_2-\vec{q}} + \hbar\omega_{\vec{q}}} \right) \\ &= M_{\vec{q}}^2 \frac{2\hbar\omega_{\vec{q}}}{(\epsilon_{\vec{k}_2} - \epsilon_{\vec{k}_2-\vec{q}})^2 - (\hbar\omega_{\vec{q}})^2} \end{aligned}$$

Finally, the effective Hamiltonian takes the form:

$$H_{eff}^{(2)} = \frac{1}{2} \sum_{\vec{k},\vec{k}',\vec{q},\sigma,\sigma'} V_{\vec{k},\vec{q}} c_{\vec{k}+\vec{q},\sigma}^{+} c_{\vec{k}'-\vec{q},\sigma'}^{+} c_{\vec{k}',\sigma'} c_{\vec{k},\sigma}$$

with

$$V_{\vec{k},\vec{q}} = M_{\vec{q}}^2 \frac{2\hbar\omega_{\vec{q}}}{(\epsilon_{\vec{k}} - \epsilon_{\vec{k}+\vec{q}})^2 - (\hbar\omega_{\vec{q}})^2}$$

With this ordering of the fermionic operators, the interaction is repulsive if the coefficient if positive, and it is attractive if the coefficient is negative (see the case of the Coulomb repulsion in Chapter 6 - Appendix). But $V_{\vec{k},\vec{q}} < 0$ if $|\epsilon_{\vec{k}} - \epsilon_{\vec{k}+\vec{q}}| < \hbar\omega_{\vec{q}}$ ! So the electron-phonon interaction induces an effective attraction between electrons with similar energies, thus leading to a qualitative modification of the electron-electron interaction. This attraction becomes significant as soon as the energies are close, not the momenta. As a consequence, it is larger than the Coulomb interaction as soon as $q$ is not too small since the Coulomb interaction decays as $1/q^2$ and is only important for small momentum transfer.

The simple physical picture of this attraction is the following: When an electron travels through the lattice, it attracts the ions. After its passage, there is an excess of positive charge that acts as an attractive potential for the other electrons.

# Chapter 5

# Superconductivity

## 5.1 Introduction

Superconductivity is probably the most remarkable phenomenon ever observed in solids. The most spectacular aspect is the *rigorous* absence of resistance below a certain temperature. In view of the effect of phonons on transport, the absence of $T^5$ contribution to resistivity implies that one of the hypotheses is no longer fulfilled below a certain temperature. The idea that has finally imposed itself is that the electron gas undergoes a phase transition into another state no longer described by a Fermi sea, but by a Bose-Einstein condensate of electron pairs that form because of the effective attraction between electrons due to the electron-phonon interaction.

The starting point is thus the effective Hamiltonian derived in the previous chapter. For simplicity, one writes it:

$$\begin{aligned}
H^{(2)}_{eff} &= -\frac{V}{2N}\sum_{\vec{k},\vec{k}',\vec{q},\sigma,\sigma'} c^{+}_{\vec{k}+\vec{q},\sigma} c^{+}_{\vec{k}'-\vec{q},\sigma'} c_{\vec{k}',\sigma'} c_{\vec{k},\sigma} \quad , \quad V > 0 \\
&\text{if} \quad |\epsilon_{\vec{k}+\vec{q}} - \epsilon_F|, |\epsilon_{\vec{k}} - \epsilon_F|, |\epsilon_{\vec{k}'-\vec{q}} - \epsilon_F|, |\epsilon_{\vec{k}'} - \epsilon_F| < \hbar\omega_D \\
H^{(2)}_{eff} &= 0 \quad \text{otherwise}
\end{aligned}$$

In other words, one assumes that there is an attraction between all electrons that are close to the Fermi surface, and that there is no electron-electron interaction otherwise.

If one considers the case of two electrons in the presence of a Fermi sea (see exercises), one can show that attraction induces a bound state that corresponds to a pair of electrons. But a pair of electrons behaves more or less as a boson. In particular, the antisymmetrization of the wave function, which imposes to work with wave functions that are all different if one wants to build an $N$ electron wave function out of one electron wave functions, does not prevent one from using $N/2$ times the same pair wave-function: the antisymmetrization does not make the resulting wave function vanish.

The analogy between the superconducting transition and the Bose-Einstein condensation is clear: the new ground state must correspond to a condensate of pairs (called Cooper pairs). Once this idea has been accepted, the difficulty is technical: How to work with a wave function that is not a Slater determinant or a simple combination of Slater determinants? By analogy with the Bose-Einstein condensation, the only simple method is based on wave functions with a number of particles that is not fixed. In their original paper, Bardeen, Cooper, and Schrieffer have used a variational approach based on a wave function that mixes states with different number of pairs. This wave function, known as the BCS wave function, can also be found as the ground state of a mean-field theory. This approach relies on a Bogoliubov transformation analogous to that introduced for antiferromagnets. So this is the presentation we have adopted in these lectures. The variational calculation will be done in the exercises.

## 5.2 The mean-field approach

Since the idea is to describe the ground state in terms of pairs, it is natural to use expectation values of pair creation operators $< c^+_{\vec{k},\uparrow} c^+_{-\vec{k},\downarrow} >$ (and their hermitian conjugate, pair annihilation operators) to decouple the Hamiltonian. Now, one can show that, in the absence of current, the energy is minimal if only pairs of zero momentum and zero total spin are used. In other words, one only needs to keep the terms $k' = -k$, $\sigma' = -\sigma$ in the sum. This means that one can work with the reduced Hamiltonian (also called BCS Hamiltonian) :

$$H_{\text{red}} = \sum_{\vec{k},\sigma} \epsilon_{\vec{k}} c^+_{\vec{k},\sigma} c_{\vec{k},\sigma} - \frac{V}{N} \sum_{\vec{k},\vec{k}'} c^+_{\vec{k},\uparrow} c^+_{-\vec{k},\downarrow} c_{-\vec{k}',\downarrow} c_{\vec{k}',\uparrow}$$

where one has replaced the sum over $\sigma$ by a factor 2.

The mean field decoupling can be written:

$$\begin{aligned} c^+_{\vec{k},\uparrow} c^+_{-\vec{k},\downarrow} c_{-\vec{k}',\downarrow} c_{\vec{k}',\uparrow} &\simeq < c^+_{\vec{k},\uparrow} c^+_{-\vec{k},\downarrow} > c_{-\vec{k}',\downarrow} c_{\vec{k}',\uparrow} + c^+_{\vec{k},\uparrow} c^+_{-\vec{k},\downarrow} < c_{-\vec{k}',\downarrow} c_{\vec{k}',\uparrow} > \\ &- < c^+_{\vec{k},\uparrow} c^+_{-\vec{k},\downarrow} >< c_{-\vec{k}',\downarrow} c_{\vec{k}',\uparrow} > \end{aligned} \tag{5.1}$$

Introducing an order parameter $\Delta$ assumed to be real by

$$\frac{\Delta}{V} = \frac{1}{N} \sum_{\vec{k}} < c^+_{\vec{k},\uparrow} c^+_{-\vec{k},\downarrow} >= \frac{1}{N} \sum_{\vec{k}} < c_{-\vec{k}',\downarrow} c_{\vec{k}',\uparrow} >$$

leads to the mean-field Hamiltonian

$$H_{MF} = \sum_{\vec{k},\sigma} \epsilon_{\vec{k}} c^+_{\vec{k},\sigma} c_{\vec{k},\sigma} - \Delta \sum_{\vec{k}} (c^+_{\vec{k},\uparrow} c^+_{-\vec{k},\downarrow} + c_{-\vec{k},\downarrow} c_{\vec{k},\uparrow}) + \frac{N\Delta^2}{V}$$

This Hamiltonian is quadratic, but it contains products of creation operators and products of annihilation operators. So, to diagonalize it, one must use a Bogoliubov transformation. The full mean-field solution of the problem thus relies on three steps:

a) Bogolioubov transformation to diagonalize the Hamiltonian;

b) Calculation of the ground state of this Hamiltonian;

c) Calculation of the expectation values $\langle c^+_{k,\sigma} c^+_{-k,-\sigma}\rangle \Rightarrow$ self-consistent equations for the parameters $\Delta$.

a) Bogolioubov transformation :

We look for fermionic operators defined by

$$\begin{cases} \alpha^+_{\vec{k},\uparrow} = u_{\vec{k}} c^+_{\vec{k}} - v_{\vec{k}} c_{-\vec{k},\downarrow} \\ \alpha_{-\vec{k},\downarrow} = u_{\vec{k}} c_{-\vec{k},\downarrow} + v_{\vec{k}} c^+_{\vec{k},\uparrow} \end{cases} \tag{5.2}$$

If the Hamiltonian is quadratic in these operators with energy $E_{\vec{k}}$, then the commutators of $H$ with these operators must be given by

$$\begin{cases} [H, \alpha^+_{\vec{k},\uparrow}] = E_{\vec{k}} \alpha^+_{\vec{k},\uparrow} \\ [H, \alpha_{-\vec{k},\downarrow}] = -E_{\vec{k}} \alpha_{-\vec{k},\downarrow} \end{cases} \tag{5.3}$$

These equations can be rewritten in terms of the original fermionic operators. A straightforward calculation of the commutators then leads to

$$\begin{cases} -\epsilon_{\vec{k}} u_{\vec{k}} - \Delta v_{\vec{k}} = -E_{\vec{k}} u_{\vec{k}} \\ \Delta u_{\vec{k}} - \epsilon_{\vec{k}} v_{\vec{k}} = E_{\vec{k}} v_{\vec{k}} \end{cases} \tag{5.4}$$

or

$$\begin{cases} (E_{\vec{k}} - \epsilon_{\vec{k}}) u_{\vec{k}} - \Delta v_{\vec{k}} = 0 \\ \Delta u_{\vec{k}} - (E_{\vec{k}} + \epsilon_{\vec{k}}) v_{\vec{k}} = 0 \end{cases} \tag{5.5}$$

$$\begin{aligned} &\Rightarrow \quad -(E^2_{\vec{k}} + \epsilon^2_{\vec{k}}) + \Delta^2 = 0 \\ &\Rightarrow \quad E_{\vec{k}} = \sqrt{\epsilon^2_{\vec{k}} + \Delta^2} \end{aligned}$$

We have chosen the positive solution because the negative one would just correspond to a different ordering of the creation and annihilation operators in the Hamiltonian. For the operators $\alpha^+_{\vec{k},\uparrow}$ and $\alpha_{-\vec{k},\downarrow}$ to be fermionic, the coefficients $u_{\vec{k}}$ and $v_{\vec{k}}$ must satisfy the normalization condition

$$u^2_{\vec{k}} + v^2_{\vec{k}} = 1$$

But from the above system of equations they are also related by

$$v_{\vec{k}} = \frac{\Delta}{E_{\vec{k}} + \epsilon_{\vec{k}}} u_{\vec{k}}$$

$$\begin{aligned}
\Rightarrow \qquad & u_{\vec{k}}^2\Big(1+\frac{\Delta^2}{(E_{\vec{k}}+\epsilon_{\vec{k}})^2}\Big)=1 \\
\Rightarrow \qquad & u_{\vec{k}}^2=\frac{1}{1+\frac{\Delta^2}{(E_{\vec{k}}+\epsilon_{\vec{k}})^2}} \\
= \quad & \frac{(E_{\vec{k}}+\epsilon_{\vec{k}})^2}{(E_{\vec{k}}+\epsilon_{\vec{k}})^2+\Delta^2} \\
= \quad & \frac{(E_{\vec{k}}+\epsilon_{\vec{k}})^2}{2E_{\vec{k}}^2+2E_{\vec{k}}\epsilon_{\vec{k}}} \\
= \quad & \frac{1}{2}\Big(1+\frac{\epsilon_{\vec{k}}}{E_{\vec{k}}}\Big)
\end{aligned}$$

Similarly

$$v_{\vec{k}}^2=1-u_{\vec{k}}^2=\frac{1}{2}\Big(1-\frac{\epsilon_{\vec{k}}}{E_{\vec{k}}}\Big)$$

Finally, the Hamiltonian takes the form

$$H=\sum_{\vec{k}}E_{\vec{k}}(\alpha^+_{\vec{k},\uparrow}\alpha_{\vec{k},\uparrow}+\alpha^+_{-\vec{k},\downarrow}\alpha_{-\vec{k},\downarrow})+Cst$$

In the following, unless there is an ambiguity, we will always associate $\vec{k}$ with a spin ↑ and $-\vec{k}$ with a spin ↓, and we will adopt the compact notations :

$$\begin{aligned}
\alpha^+_{\vec{k}} &\equiv \alpha^+_{\vec{k},\uparrow} \\
\alpha_{\vec{k}} &\equiv \alpha_{\vec{k},\uparrow} \\
\alpha^+_{-\vec{k}} &\equiv \alpha^+_{-\vec{k},\downarrow} \\
\alpha_{-\vec{k}} &\equiv \alpha_{-\vec{k},\downarrow}
\end{aligned}$$

and similarly for the $c^+$ and $c$ operators.

b) Ground state

With these notations, the effective Hamiltonian after decoupling reads:

$$H=\sum_{\vec{k}}E_{\vec{k}}(\alpha^+_{\vec{k}}\alpha_{\vec{k}}+\alpha^+_{-\vec{k}}\alpha_{-\vec{k}})+Cst$$

Since $E_{\vec{k}}>0$, the ground state corresponds to the vacuum of the $\alpha^+_{\vec{k}}$ particles, i.e. it must satisfy:

$$\begin{aligned}
\alpha_{\vec{k}}|\phi_0> &= 0 \\
\alpha_{-\vec{k}}|\phi_0> &= 0
\end{aligned}$$

for all $\vec{k}$.

If one denotes by $|0>$ the vacuum of the $c^+_{\vec{k},\sigma}$ particles, the state

$$\prod_{\vec{k}} \alpha_{\vec{k}}\alpha_{-\vec{k}}|0>$$

satisfies these conditions since the $\alpha_{\vec{k}}$ operators are fermions.

But

$$\begin{aligned}
&\prod_{\vec{k}} \alpha_{\vec{k}}\alpha_{-\vec{k}}|0> \\
&= \prod_{\vec{k}}(u_{\vec{k}}c_{\vec{k}} - v_{\vec{k}}c^+_{-\vec{k}})(u_{\vec{k}}c_{-\vec{k}} + v_{\vec{k}}c^+_{\vec{k}})|0> \\
&= \prod_{\vec{k}}(u^2_{\vec{k}}c_{\vec{k}}c_{-\vec{k}} + u_{\vec{k}}v_{\vec{k}}c_{\vec{k}}c^+_{\vec{k}} - v_{\vec{k}}u_{\vec{k}}c^+_{-\vec{k}}c_{-\vec{k}} - v^2_{\vec{k}}c^+_{-\vec{k}}c^+_{\vec{k}})|0> \\
&= \prod_{\vec{k}}(u_{\vec{k}}v_{\vec{k}} + v^2_{\vec{k}}c^+_{\vec{k}}c^+_{-\vec{k}})|0> \\
&= \prod_{\vec{k}} v_{\vec{k}}(u_{\vec{k}} + v_{\vec{k}}c^+_{\vec{k}}c^+_{-\vec{k}})|0>
\end{aligned}$$

Besides, this wave fucntion must be normalized :

$$\begin{aligned}
\Rightarrow\ & <0|\prod_{\vec{k}} \alpha^+_{-\vec{k}}\alpha^+_{\vec{k}}\alpha_{\vec{k}}\alpha_{-\vec{k}}|0> \\
= \ & <0|\prod_{\vec{k}} v^2_{\vec{k}}(u_{\vec{k}} + v_{\vec{k}}c_{-\vec{k}}c_{\vec{k}})(u_{\vec{k}} + v_{\vec{k}}c^+_{\vec{k}}c^+_{-\vec{k}})|0> \\
= \ & <0|\prod_{\vec{k}} v^2_{\vec{k}}(u^2_{\vec{k}} + v^2_{\vec{k}}c_{-\vec{k}}c_{\vec{k}}c^+_{\vec{k}}c^+_{-\vec{k}})|0> \\
= \ & \prod_{\vec{k}} v^2_{\vec{k}}(u^2_{\vec{k}} + v^2_{\vec{k}}) = v^2_{\vec{k}}
\end{aligned}$$

because $c_{-\vec{k}}c_{\vec{k}}c^+_{\vec{k}}c^+_{-\vec{k}})|0>=|0>$. Finally, the normalized ground state is given by:

$$\boxed{|\phi_0>=\prod_{\vec{k}}(u_{\vec{k}} + v_{\vec{k}}c^+_{\vec{k}}c^+_{-\vec{k}})|0>}$$

c) Self-consistent equations

It is clear that this ground state is very different from a Fermi sea. In particular, it mixes states with different numbers of particles. That's why expectation values such as $< c^+_{\vec{k}}c^+_{-\vec{k}} >$ can be non vanishing.

Let us now calculate

$$\Delta = \frac{V}{N}\sum_{\vec{k}} < c_{-\vec{k}}c_{\vec{k}} >$$

$$\begin{aligned}
< c_{-\vec{k}} c_{\vec{k}} > &= < 0| \prod_l (u_l + v_l c_{-l} c_l) c_{-\vec{k}} c_{\vec{k}} \prod_{l'} (u_{l'} + v_{l'} c^+_{l'} c^+_{-l'})|0 > \\
&= < 0| u_k c_{-\vec{k}} c_{\vec{k}} \prod_{l \neq k} (u_l + v_l c_{-l} c_l) \prod_{l'} (u_{l'} + v_{l'} c^+_{l'} c^+_{-l'})|0 > \\
&= < 0| u_k c_{-\vec{k}} c_{\vec{k}} (u_k + v_k c^+_k c^+_{-k}) \prod_{l \neq k} (u_l + v_l c_{-l} c_l) \prod_{l' \neq k} (u_{l'} + v_{l'} c^+_{l'} c^+_{-l'})|0 > \\
&= < 0|(u_{\vec{k}} c_{-\vec{k}} c_{\vec{k}} v_{\vec{k}} c^+_{\vec{k}} c^+_{-\vec{k}})|0 > \\
&= u_{\vec{k}} v_{\vec{k}}
\end{aligned}$$

$$\Rightarrow \Delta = \frac{V}{N} \sum_{\vec{k}} u_{\vec{k}} v_{\vec{k}}$$

But

$$u_{\vec{k}} v_{\vec{k}} = \frac{1}{2} \sqrt{1 - \frac{\epsilon^2_{\vec{k}}}{E^2_{\vec{k}}}} = \frac{1}{2} \frac{\Delta}{\sqrt{\epsilon^2_{\vec{k}} + \Delta^2}}$$

$$\Rightarrow \Delta = \frac{V}{2N} \Delta \sum_{\vec{k}} \frac{1}{\sqrt{\epsilon^2_{\vec{k}} + \Delta^2}}$$

There is of course the solution $\Delta = 0$ that corresponds to the normal state. But this equation possesses another non-zero solution:

$$\begin{aligned}
1 &= \frac{V}{2N} \sum_{\vec{k}} \frac{1}{\sqrt{\epsilon^2_{\vec{k}} + \Delta^2}} \\
1 &= \frac{V}{2} \int g(\epsilon) \frac{1}{\sqrt{\epsilon^2 + \Delta^2}} d\epsilon
\end{aligned}$$

This integral is limited to energies close to $\epsilon_F$ (within the Debye frequency $\omega_D$) for which $g(\epsilon) \simeq g(\epsilon_F)$

$$\begin{aligned}
1 &\simeq \frac{V}{2} g(\epsilon_F) \int_{-\omega_D}^{+\omega_D} \frac{1}{\sqrt{\epsilon^2 + \Delta^2}} d\epsilon \\
1 &\simeq \frac{V}{2} g(\epsilon_F) \Big[ \ln(\epsilon + \sqrt{\epsilon^2 + \Delta^2}) \Big]_{-\omega_D}^{+\omega_D} \\
1 &\simeq \frac{V}{2} g(\epsilon_F) \ln \left[ \frac{\omega_D + \sqrt{\omega_D^2 + \Delta^2}}{-\omega_D + \sqrt{\omega_D^2 + \Delta^2}} \right]
\end{aligned}$$

Let us suppose that $\Delta$ is small as compared to $\omega_D$ (to be checked a posteriori)

$$\begin{aligned}
1 &\simeq \frac{V}{2} g(\epsilon_F) \ln \left[ \frac{\omega_D + \omega_D(1 + \frac{\Delta^2}{2\omega_D^2})}{\frac{\Delta^2}{2\omega_D}} \right] \\
1 &\simeq \frac{V}{2} g(\epsilon_F) \ln \left( \frac{4\omega_D^2}{\Delta^2} \right) \\
1 &\simeq \frac{V}{2} g(\epsilon_F) 2 \ln \left( \frac{2\omega_D}{\Delta} \right) \\
\frac{1}{V g(\epsilon_F)} &= \ln \left( \frac{2\omega_D}{\Delta} \right) \\
&\Rightarrow \Delta = 2\omega_D e^{\frac{-1}{V g(\epsilon_F)}}
\end{aligned}$$

Note: The parameter $\Delta$ does not have an expansion in powers of $V$. So perturbative approaches cannot be used.

The original method of BCS consists in assuming that the ground state is of the form $\prod_{\vec{k}} (u_{\vec{k}} + v_{\vec{k}} c^+_{\vec{k}} c^+_{-\vec{k}}) |0>$, and to minimiuze the energy with respect to $u_{\vec{k}}$ and $v_{\vec{k}}$. This variational method leads to the same results.

## 5.3 Interpretation of the results

To understand the physical implications of the existence of the solution with $\Delta \neq 0$, we must:

a) Compare its energy to the solution with $\Delta = 0$.

b) Come up with a physical interpretation of the wave function.

c) Extract from the solution the elementary excitations and the behaviour at finite temperature.

Condensation energy :

To calculate the energy of both solutions, let us go back to the mean-field Hamiltonian

$$\Rightarrow H_{MF} = \sum_{\vec{k},\sigma} \epsilon_{\vec{k}} c^+_{\vec{k},\sigma} c_{\vec{k},\sigma} - \Delta \sum_{\vec{k}} (c^+_{\vec{k},\uparrow} c^+_{-\vec{k},\downarrow} + c_{-\vec{k},\downarrow} c_{\vec{k},\uparrow}) + \frac{N\Delta^2}{V}$$

and let us invert the Bogolioubov transform to express $H_{MF}$ in terms of $\alpha^+_{\vec{k},\sigma}$:

$$\left\{ \begin{array}{ll} \alpha^+_{\vec{k},\uparrow} &= u_{\vec{k}} c^+_{\vec{k},\uparrow} - v_{\vec{k}} c_{-\vec{k},\downarrow} \\ \alpha_{-\vec{k},\downarrow} &= u_{\vec{k}} c_{-\vec{k},\downarrow} + v_{\vec{k}} c^+_{\vec{k},\uparrow} \end{array} \right. \Rightarrow \left\{ \begin{array}{ll} c^+_{\vec{k},\uparrow} &= u_{\vec{k}} \alpha^+_{\vec{k},\uparrow} + v_{\vec{k}} \alpha_{-\vec{k},\downarrow} \\ c_{\vec{k},\uparrow} &= u_{\vec{k}} \alpha_{\vec{k},\uparrow} + v_{\vec{k}} \alpha^+_{-\vec{k},\downarrow} \\ c^+_{-\vec{k},\downarrow} &= u_{\vec{k}} \alpha^+_{-\vec{k},\downarrow} - v_{\vec{k}} \alpha_{\vec{k},\uparrow} \\ c_{-\vec{k},\downarrow} &= u_{\vec{k}} \alpha_{-\vec{k},\downarrow} - v_{\vec{k}} \alpha^+_{\vec{k},\uparrow} \end{array} \right.$$

taking into account the fact that $v_{\vec{k}}$ changes sign between $\vec{k},\uparrow$ and $-\vec{k},\downarrow$.

$$\begin{aligned}
c^+_{\vec{k},\uparrow}c_{\vec{k},\uparrow} &= (u_{\vec{k}}\alpha^+_{\vec{k},\uparrow} + v_{\vec{k}}\alpha_{-\vec{k},\downarrow})(u_{\vec{k}}\alpha_{\vec{k},\uparrow} + v_{\vec{k}}\alpha^+_{-\vec{k},\downarrow}) \\
&= u^2_{\vec{k}}\alpha^+_{\vec{k},\uparrow}\alpha_{\vec{k},\uparrow} + v^2_{\vec{k}}(1-\alpha^+_{-\vec{k},\downarrow}\alpha_{-\vec{k},\downarrow}) \\
&+ u_{\vec{k}}v_{\vec{k}}(\alpha^+_{\vec{k},\uparrow}\alpha^+_{-\vec{k},\downarrow} + \alpha_{-\vec{k},\downarrow}\alpha_{\vec{k},\uparrow})
\end{aligned}$$

$$\begin{aligned}
c^+_{\vec{k},\downarrow}c_{\vec{k},\downarrow} &= (u_{\vec{k}}\alpha^+_{\vec{k},\downarrow} - v_{\vec{k}}\alpha_{-\vec{k},\uparrow})(u_{\vec{k}}\alpha_{\vec{k},\downarrow} - v_{\vec{k}}\alpha^+_{-\vec{k},\uparrow}) \\
&= u^2_{\vec{k}}\alpha^+_{\vec{k},\downarrow}\alpha_{\vec{k},\downarrow} + v^2_{\vec{k}}(1-\alpha^+_{-\vec{k},\uparrow}\alpha_{-\vec{k},\uparrow}) \\
&- u_{\vec{k}}v_{\vec{k}}(\alpha_{-\vec{k},\uparrow}\alpha_{\vec{k},\downarrow} + \alpha^+_{\vec{k},\downarrow}\alpha^+_{-\vec{k},\uparrow})
\end{aligned}$$

$$\begin{aligned}
c^+_{\vec{k},\uparrow}c^+_{-\vec{k},\downarrow} &= (u_{\vec{k}}\alpha^+_{\vec{k},\uparrow} + v_{\vec{k}}\alpha_{-\vec{k},\downarrow})(u_{\vec{k}}\alpha^+_{-\vec{k},\downarrow} - v_{\vec{k}}\alpha_{\vec{k},\uparrow}) \\
&= u^2_{\vec{k}}\alpha^+_{\vec{k},\uparrow}\alpha^+_{-\vec{k},\downarrow} - v^2_{\vec{k}}\alpha_{-\vec{k},\downarrow}\alpha_{\vec{k},\uparrow} \\
&- u_{\vec{k}}v_{\vec{k}}\alpha^+_{\vec{k},\uparrow}\alpha_{\vec{k},\uparrow} + u_{\vec{k}}v_{\vec{k}}(1-\alpha^+_{-\vec{k},\downarrow}\alpha_{-\vec{k},\downarrow})
\end{aligned}$$

$$\begin{aligned}
c_{-\vec{k},\downarrow}c_{\vec{k},\uparrow} &= (u_{\vec{k}}\alpha_{-\vec{k},\downarrow} - v_{\vec{k}}\alpha^+_{\vec{k},\uparrow})(u_{\vec{k}}\alpha_{\vec{k},\uparrow} + v_{\vec{k}}\alpha^+_{-\vec{k},\downarrow}) \\
&= u^2_{\vec{k}}\alpha_{-\vec{k},\downarrow}\alpha_{\vec{k},\uparrow} - v^2_{\vec{k}}\alpha^+_{\vec{k},\uparrow}\alpha^+_{-\vec{k},\downarrow} \\
&+ u_{\vec{k}}v_{\vec{k}}(1-\alpha^+_{-\vec{k},\downarrow}\alpha_{-\vec{k},\downarrow}) - u_{\vec{k}}v_{\vec{k}}\alpha^+_{\vec{k},\uparrow}\alpha_{\vec{k},\uparrow}
\end{aligned}$$

<u>Coefficient of $\alpha^+_{\vec{k},\uparrow}\alpha_{\vec{k},\uparrow}$:</u>

$$\begin{aligned}
&\epsilon_{\vec{k}}u^2_{\vec{k}} - \epsilon_{\vec{k}}v^2_{\vec{k}} + \Delta u_{\vec{k}}v_{\vec{k}} + \Delta u_{\vec{k}}v_{\vec{k}} \\
&= \epsilon_{\vec{k}}\frac{\epsilon_{\vec{k}}}{E_{\vec{k}}} + 2\Delta\sqrt{\frac{1}{4}\left(1-\frac{\epsilon^2_{\vec{k}}}{E^2_{\vec{k}}}\right)} \\
&= \frac{\epsilon^2_{\vec{k}}}{E_{\vec{k}}} + \Delta\sqrt{\frac{\Delta^2}{E^2_{\vec{k}}}} = E_{\vec{k}}
\end{aligned}$$

<u>Coefficient of $\alpha^+_{\vec{k},\uparrow}\alpha^+_{-\vec{k},\downarrow}$:</u>

$$\begin{aligned}
&\epsilon_{\vec{k}}u_{\vec{k}}v_{\vec{k}} + \epsilon_{\vec{k}}u_{\vec{k}}v_{\vec{k}} - \Delta u^2_{\vec{k}} + \Delta v^2_{\vec{k}} \\
&= 2\epsilon_{\vec{k}}u_{\vec{k}}v_{\vec{k}} + \Delta(v^2_{\vec{k}} - u^2_{\vec{k}}) \\
&= \epsilon_{\vec{k}}\frac{\Delta}{E_{\vec{k}}} + \Delta(-\frac{\epsilon_{\vec{k}}}{E_{\vec{k}}}) = 0
\end{aligned}$$

<u>Constant term :</u>

$$\begin{aligned}
&\sum_{\vec{k}}(\epsilon_{\vec{k}}v^2_{\vec{k}} + \epsilon_{\vec{k}}v^2_{\vec{k}} - \Delta u_{\vec{k}}v_{\vec{k}} - \Delta u_{\vec{k}}v_{\vec{k}}) + \frac{N\Delta^2}{V} \\
&= \sum_{\vec{k}}(\epsilon_{\vec{k}} - \frac{\epsilon^2_{\vec{k}}}{E_{\vec{k}}} - \frac{\Delta^2}{E_{\vec{k}}}) + \frac{N\Delta^2}{V}
\end{aligned}$$

But from the gap equation:

$$\sum_{\vec{k}} \frac{1}{E_{\vec{k}}} = \frac{2N}{V}$$

$$\Rightarrow \sum_{\vec{k}} (\epsilon_{\vec{k}} - \frac{\epsilon_{\vec{k}}^2}{E_{\vec{k}}}) - \frac{N\Delta^2}{V}$$

Altogether, the Hamiltonian can be written :

$$\begin{aligned} H_{CM} &= \sum_{\vec{k}} E_{\vec{k}} (\alpha^+_{\vec{k},\uparrow}\alpha_{\vec{k},\uparrow} + \alpha^+_{-\vec{k},\downarrow}\alpha_{-\vec{k},\downarrow}) \\ &+ \sum_{\vec{k}} (\epsilon_{\vec{k}} - \frac{\epsilon_{\vec{k}}^2}{E_{\vec{k}}}) - \frac{N\Delta^2}{V} \end{aligned}$$

In the ground state, since $E_{\vec{k}} > 0$, there is no particle $\alpha^+_{\vec{k}}$. Let us compare this energy with that of the non-superconducting case ($\Delta = 0$).

$\underline{\Delta = 0}$

$$\begin{aligned} E_{\vec{k}} = \sqrt{\epsilon_{\vec{k}}^2} &= \epsilon_{\vec{k}} \quad \text{si} \quad \epsilon_{\vec{k}} > 0 \quad (\vec{k} \text{ empty}) \\ &= -\epsilon_{\vec{k}} \quad \text{si} \quad \epsilon_{\vec{k}} < 0 \quad (\vec{k} \text{ occ.}) \end{aligned}$$

$$\Rightarrow E_{GS} = \sum_{\vec{k}_{occ}} 2\epsilon_{\vec{k}}$$

$\underline{\Delta \neq 0}$

$$\begin{aligned} & E_{GS}(\Delta \neq 0) - E_{GS}(\Delta = 0) \\ = & \sum_{\vec{k}_{occ}} (-\epsilon_{\vec{k}} - \frac{\epsilon_{\vec{k}}^2}{E_{\vec{k}}}) + \sum_{\vec{k}_{\text{empty}}} (\epsilon_{\vec{k}} - \frac{\epsilon_{\vec{k}}^2}{E_{\vec{k}}}) - \frac{N\Delta^2}{V} \\ = & \; 2 \sum_{\vec{k}_{\text{empty}}} (\epsilon_{\vec{k}} - \frac{\epsilon_{\vec{k}}^2}{E_{\vec{k}}}) - \frac{N\Delta^2}{V} \end{aligned}$$

$$\frac{\Delta E}{N} \simeq 2g(\epsilon_F) \int_0^{\omega_D} d\epsilon (\epsilon - \frac{\epsilon^2}{\sqrt{\epsilon^2 + \Delta^2}}) - \frac{\Delta^2}{V}$$

Let us calculate

$$2g(\epsilon_F) \int_0^{\omega_D} \epsilon d\epsilon (1 - \frac{\epsilon}{\sqrt{\epsilon^2 + \Delta^2}})$$

To this end, it is useful to perform the change of variables $\epsilon = \Delta \sinh u$, $d\epsilon = \cosh u \, du$

$$\begin{aligned} \Rightarrow \quad & 2g(\epsilon_F) \int_0^{u_D} \left(1 - \frac{\Delta \sinh u}{\Delta \cosh u}\right) \Delta^2 \sinh u \cosh u \, du \\ = \quad & 2g(\epsilon_F)\Delta^2 \int_0^{u_D} (1 - \tanh u) \sinh u \cosh u \, du \end{aligned}$$

But

$$(1-\tanh u)\sinh u\cosh u = (\cosh u - \sinh u)\sinh u = e^{-u}\sinh u = \frac{1-e^{-2u}}{2}$$

so that

$$\begin{aligned} &= \Delta^2 g(\epsilon_F)\int_0^{u_D}(1-e^{-2u})\,du \\ &= \Delta^2 g(\epsilon_F)(u_D + \frac{e^{-2u_D}-1}{2}) \end{aligned}$$

<u>Proposition:</u> $u_D = \frac{1}{Vg(\epsilon_F)}$.
Indeed, $\sinh u_D = \frac{\omega_D}{\Delta}$.
But the gap is given by :

$$\begin{aligned} 1 &= Vg(\epsilon_F)\int_0^{u_D}\frac{d\epsilon}{\sqrt{\epsilon^2+\Delta^2}} \\ &= Vg(\epsilon_F)\text{Argsh}\left(\frac{\omega_D}{\Delta}\right) \end{aligned}$$

$$\Rightarrow \frac{\omega_D}{\Delta} = \sinh\frac{1}{Vg(\epsilon_F)} \Rightarrow u_D = \frac{1}{Vg(\epsilon_F)}$$

So, $u_D \gg 1$. One can thus neglect $e^{-2u_D}$ in the previous result. Finally, the condensation energy is given by:

$$\Delta^2 g(\epsilon_F)\left(\frac{1}{Vg(\epsilon_F)} - \frac{1}{2}\right) - \frac{\Delta^2}{V}$$

leading to

$$E_{cond} = -\frac{1}{2}g(\epsilon_F)\Delta^2$$

As soon as there is a solution $\Delta \neq 0$, its energy is lower. Bus as soon as $V$ is positive, i.e. as soon as the interaction is attractive, there is a solution $\Delta \neq 0$. Finally,

$$\boxed{\text{Attraction} \Rightarrow \text{superconductivity}}$$

<u>b) Interpretation of the wave function :</u>

So there is a mean-field solution whose energy is lower than that of the Fermi sea as soon as $V > 0$. Let us try to understand the physics of this solution.

The ground state wave function reads:

$$|\phi_0> = \prod_{\vec{k}}(u_{\vec{k}} + v_{\vec{k}}c^+_{\vec{k},\uparrow}c^+_{-\vec{k},\downarrow})|0>$$

The probability for the pair $c^+_{\vec{k},\uparrow}c^+_{-\vec{k},\downarrow}$ to be occupied is equal to $v^2_{\vec{k}}$, the probability for it not to be occupied is equal to $u^2_{\vec{k}}$. It is clear that this wave function contains states with different particle numbers. This is not a problem as long as, in the limit of very large systems, the distribution of particles is "peaked" around a well defined mean value.

Let us thus calculate $\bar{N} =< N >$ and $\delta N = (< (N-\bar{N})^2 >)^{\frac{1}{2}}$

$$< N >=< \phi_0 | \sum_{\vec{k},\sigma} c^+_{\vec{k},\sigma} c_{\vec{k},\sigma} | \phi_0 >$$

Let us consider a term $c^+_{\vec{k}_0,\uparrow}c_{\vec{k}_0,\uparrow}$. All the operators $u_{\vec{k}} + v_{\vec{k}} c^+_{\vec{k},\uparrow} c^+_{-\vec{k},\downarrow}$ commute with it except $u_{\vec{k}_0} + v_{\vec{k}_0} c^+_{\vec{k}_0,\uparrow}c^+_{-\vec{k}_0,\downarrow}$.

$$< N >= \underbrace{2}_{\text{spin}} \sum_{\vec{k}} < 0 | (u_{\vec{k}} + v_{\vec{k}} c_{-\vec{k},\downarrow} c_{\vec{k},\uparrow}) c^+_{\vec{k},\uparrow} c_{\vec{k},\uparrow} (u_{\vec{k}} + v_{\vec{k}} c^+_{\vec{k},\uparrow} c^+_{-\vec{k},\downarrow}) |0>$$

$$\begin{aligned}
& c^+_{\vec{k},\uparrow} c_{\vec{k},\uparrow} (u_{\vec{k}} + v_{\vec{k}} c^+_{\vec{k},\uparrow} c^+_{-\vec{k},\downarrow}) |0> \\
= \; & v_{\vec{k}} c^+_{\vec{k},\uparrow} (1 - c^+_{\vec{k},\uparrow} c_{\vec{k},\uparrow}) c^+_{-\vec{k},\downarrow} |0> \\
= \; & v_{\vec{k}} c^+_{\vec{k},\uparrow} c^+_{-\vec{k},\downarrow} |0> \\
& \Rightarrow < N >= 2 \sum_{\vec{k}} v^2_{\vec{k}}
\end{aligned}$$

$$\begin{aligned}
\delta N^2 &= < N^2 - 2\bar{N}N + \bar{N}^2 > = < N^2 > - \bar{N}^2 \\
< N^2 > &= < \phi_0 | \sum_{\vec{k},\vec{k}',\sigma,\sigma'} c^+_{\vec{k},\sigma} c_{\vec{k},\sigma} c^+_{\vec{k}',\sigma'} c_{\vec{k}',\sigma'} | \phi_0 >
\end{aligned}$$

If $(\vec{k}',\sigma') \neq (\vec{k},\sigma)$ and $\neq (-\vec{k},-\sigma)$, the terms commute $\to v^2_{\vec{k}} v^2_{\vec{k}'}$

If $(\vec{k}',\sigma') = (\vec{k},\sigma) \Rightarrow c^+_{\vec{k},\sigma} \underbrace{c_{\vec{k},\sigma} c^+_{\vec{k},\sigma}}_{1 - c^+_{\vec{k},\sigma} c_{\vec{k},\sigma}} c_{\vec{k},\sigma} \to v^2_{\vec{k}}$

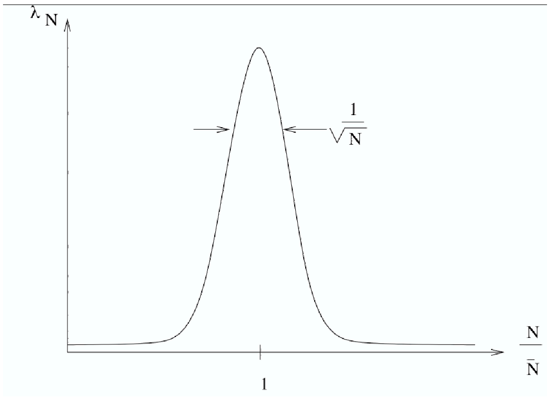
If $(\vec{k}',\sigma') = (-\vec{k},-\sigma) \Rightarrow c^+_{\vec{k},\uparrow} c_{\vec{k},\uparrow} c^+_{-\vec{k},\downarrow} c_{-\vec{k},\downarrow} (u_{\vec{k}} + v_{\vec{k}} c^+_{\vec{k},\uparrow} c^+_{-\vec{k},\downarrow}) \to v^2_{\vec{k}}$

$$\begin{aligned}
\Rightarrow < N^2 > &= \left( \sum_{\vec{k},\sigma} v^2_{\vec{k}} \right)^2 - \underbrace{\sum_{\vec{k},\sigma} \left( v^4_{\vec{k}} - v^2_{\vec{k}} \right)}_{(\vec{k}',\sigma') = (\vec{k},\sigma)} - \underbrace{\sum_{\vec{k},\sigma} \left( v^4_{\vec{k}} - v^2_{\vec{k}} \right)}_{(\vec{k}',\sigma') = (-\vec{k},-\sigma)} \\
&= \left( \sum_{\vec{k}} 2 v^2_{\vec{k}} \right)^2 - 2 \sum_{\vec{k}} 2 v^4_{\vec{k}} + 2 \sum_{\vec{k}} 2 v^2_{\vec{k}} \\
&= \bar{N}^2 + 4 \sum_{\vec{k}} v^2_{\vec{k}} (1 - v^2_{\vec{k}})
\end{aligned}$$

$$\Rightarrow \ \delta N^2 = 4 \sum_{\vec{k}} u_{\vec{k}}^2 v_{\vec{k}}^2$$

Conclusion :

$\bar{N}$ grows as $N$, but $\delta N$ only grows as $\sqrt{N}$. So, if one writes $\phi_0 = \sum_N \lambda_N \phi_N$, where $< \phi_N|\hat{N}|\phi_N >= N$, the $\lambda_N$ are distributed according to



So the expectation value of any observable calculated in $|\phi_0>$ is the same as that calculated in $|\phi_{\bar{N}}>$ in the limit $N \to +\infty$. But this wave function is much easier to manipulate.

It is in fact easy to understand the close relationship between this wave function and wave functions with a fixed number of particles. Let us consider the case where $2N$ electrons are concerned by the instability (i.e. $2N$ electrons whose energy $\epsilon_{\vec{k}}$ lies between $\epsilon_F - \omega_D$ and $\epsilon_F$). There are thus $2N$ values of $\vec{k}$ such that $\epsilon_F - \omega_D < \epsilon_{\vec{k}} < \epsilon_F + \omega_D$, and the Fermi sea consists in filling up the $N$ values such that $\epsilon_F - \omega_D < \epsilon_{\vec{k}} < \epsilon_F$ with an $\uparrow$ spin or a $\downarrow$ spin.

If one tries to construct a wave function only built out of pairs to gain the potential energy of $H_{\text{red}}$, and having a well defined number of particles, a natural choice would be:

$$|\phi_N>= \frac{1}{N!} \left( \sum_{\vec{k}} g_{\vec{k}} c^+_{\vec{k},\uparrow} c^+_{-\vec{k},\downarrow} \right)^N |0>$$

If

$$g_{\vec{k}} \begin{cases} 1 & \text{when } \epsilon_{\vec{k}} < \epsilon_F \\ 0 & \text{otherwise} \end{cases}$$

one gets back the Fermi sea, and since there are $N!$ terms, one must divide by $N!$.

Starting from this wave function, one could look for the $g_{\vec{k}}$ that minimize the energy. But this wave function is <u>very</u> complicated. It contains of the order of $(2N)^N$ terms!

Now, if we add various $|\phi_N>$ to build

$$\begin{aligned}|\phi> &= \sum_N |\phi_N> \\ &= \sum_N \frac{1}{N!}\left(\sum_{\vec{k}} g_{\vec{k}} c^+_{\vec{k},\uparrow} c^+_{-\vec{k},\downarrow}\right)^N |0>\end{aligned}$$

this function can be rewritten:

$$|\phi>= e^{\sum_{\vec{k}} g_{\vec{k}} c^+_{\vec{k},\uparrow} c^+_{-\vec{k},\downarrow}}|0>$$

But for fermions this coherent state takes a simple form: in the expansion of the exponential, the only terms that survive are those of order 0 or 1 beacuse of the Pauli principle

$$\Rightarrow |\phi> \quad = \quad \prod_{\vec{k}}(1 + g_{\vec{k}} c^+_{\vec{k},\uparrow} c^+_{-\vec{k},\downarrow})|0>$$

If one writes $g_{\vec{k}} = \frac{v_{\vec{k}}}{u_{\vec{k}}}$, one gets back the BCS wave function. This wave function is much easier to manipulate. It is thanks to it that Bardeen, Cooper and Schrieffer got the Nobel prize!

c) Elementary excitations :

It is clear from the form of $H_{CM}$ that an elementary excitation consists in adding one particle $\alpha^+_{\vec{k},\sigma}$. This costs an energy

$$E_{\vec{k}} = \sqrt{\epsilon^2_{\vec{k}} + \Delta^2}$$

This energy is greater than or equal to $\Delta$.

There is thus an energy gap $\Delta$ in the single particle density of states of the system. The very basis of the theory of transport, with diffusion processes of quasi particles at the Fermi level, is thus not appropriate any more. One needs to investigate the properties of the electron gas in this state.

## 5.4 Finite temperatures

At finite temperature, one can still use mean-field theory, but now the gap involves the thermal mean value of $< c^+_{\vec{k},\uparrow} c^+_{-\vec{k},\downarrow} >$.

Now,

$$\begin{aligned}c^+_{\vec{k},\uparrow} c^+_{-\vec{k},\downarrow} &= (u_{\vec{k}}\alpha^+_{\vec{k},\uparrow} + v_{\vec{k}}\alpha_{-\vec{k},\downarrow})(u_{\vec{k}}\alpha^+_{-\vec{k},\downarrow} - v_{\vec{k}}\alpha_{\vec{k},\uparrow}) \\ &= -u_{\vec{k}}v_{\vec{k}}\alpha^+_{\vec{k},\uparrow}\alpha_{\vec{k},\uparrow} + u_{\vec{k}}v_{\vec{k}}\alpha_{-\vec{k},\downarrow}\alpha^+_{-\vec{k},\downarrow} + u^2_{\vec{k}}\alpha^+_{\vec{k},\uparrow}\alpha^+_{-\vec{k},\downarrow} - v^2_{\vec{k}}\alpha_{-\vec{k},\downarrow}\alpha_{\vec{k},\uparrow} \\ &= u_{\vec{k}}v_{\vec{k}}(1 - \alpha^+_{\vec{k},\uparrow}\alpha_{\vec{k},\uparrow} - \alpha^+_{-\vec{k},\downarrow}\alpha_{-\vec{k},\downarrow}) + u^2_{\vec{k}}\alpha^+_{\vec{k},\uparrow}\alpha^+_{-\vec{k},\downarrow} - v^2_{\vec{k}}\alpha_{-\vec{k},\downarrow}\alpha_{\vec{k},\uparrow}\end{aligned}$$

$$\Rightarrow < c^+_{\vec{k},\uparrow} c^+_{-\vec{k},\downarrow} > = \ u_{\vec{k}} v_{\vec{k}} (1 - 2f(E_{\vec{k}}))$$

since $< \alpha^+_{\vec{k},\uparrow}\alpha_{\vec{k},\uparrow} >=< \alpha^+_{-\vec{k},\downarrow}\alpha_{-\vec{k},\downarrow} >= \frac{1}{e^{\beta E_{\vec{k}}+1}}$ et $< \alpha^+_{\vec{k},\uparrow}\alpha^+_{-\vec{k},\downarrow} >=< \alpha_{-\vec{k},\downarrow}\alpha_{\vec{k},\uparrow} >= 0$

So the gap equation is given by :

$$\Delta = \frac{V}{N}\sum_{\vec{k}} \frac{\Delta}{2E_{\vec{k}}}(1 - 2f(E_{\vec{k}}))$$

$$1 - 2f(E_{\vec{k}}) = \frac{e^{\beta E_{\vec{k}}} + 1 - 2}{e^{\beta E_{\vec{k}}} + 1} = \frac{e^{\beta \frac{E_{\vec{k}}}{2}} - e^{-\beta \frac{E_{\vec{k}}}{2}}}{e^{\beta \frac{E_{\vec{k}}}{2}} + e^{-\beta \frac{E_{\vec{k}}}{2}}} = \tanh \frac{\beta E_{\vec{k}}}{2}$$

The non-vanishing solution thus satisfies:

$$1 = \frac{V}{N}\sum_{\vec{k}} \frac{1}{2E_{\vec{k}}} \tanh \frac{\beta E_{\vec{k}}}{2}$$

When $T$ increases, $\beta \to 0$ and this equation is satisfied for smaller and smaller values of $\Delta$. In particular, there is a temperature $T_c$ where $\Delta(T_c) = 0$ defined by

$$1 = \frac{V}{N}\sum_{\vec{k}} \frac{1}{2|\epsilon_{\vec{k}}|} \tanh \frac{\beta \epsilon_{\vec{k}}}{2} \quad (\beta = \frac{1}{T_c})$$

$$1 = Vg(\epsilon_F)\int_0^{\omega_D} \tanh \frac{\beta\epsilon}{2}\frac{d\epsilon}{\epsilon}$$

$$\frac{1}{Vg(\epsilon_F)} = \int_0^{\frac{\omega_D}{2T_c}} \frac{\tanh u}{u} du \quad (u = \frac{\beta\epsilon}{2})$$

Let us suppose that $T_c \ll \omega_D$ (this is reasonable since one expects that $T_c \simeq \Delta(0) \ll \omega_D$ because thermal fluctuations destroy the gap if $T \gg \Delta(0)$).

$$\begin{aligned}
\int_0^{\frac{\omega_D}{2T_c}} \frac{\tanh u}{u} du &= \int_0^{\eta} \frac{\tanh u}{u} du + \int_\eta^{\frac{\omega_D}{2T_c}} \frac{\tanh u}{u} du \\
&\simeq \int_0^{\eta} \frac{\tanh u}{u} du + \int_\eta^{\frac{\omega_D}{2T_c}} \frac{du}{u} \quad (\eta \gg 1 \Rightarrow \tanh u \simeq 1) \\
&= \ln \frac{\omega_D}{T_c} + \underbrace{\int_0^{\eta} \frac{\tanh u}{u} du - \ln \eta - \ln 2}_{\text{depends very little on } \eta, \to \ln \frac{2\gamma}{\pi} = \ln 1.13}
\end{aligned}$$

where $\gamma$ is the Euler constant.

$$\Rightarrow T_c \simeq 1.13\, \omega_D e^{\frac{-1}{Vg(\epsilon_F)}}$$

This fomula cannot be used directly to predict $T_c$ since $\omega_D$, $V$ and $g(\epsilon_F)$ are adjustable parameters.

However, if we compare it to the formula for $\Delta$, we get

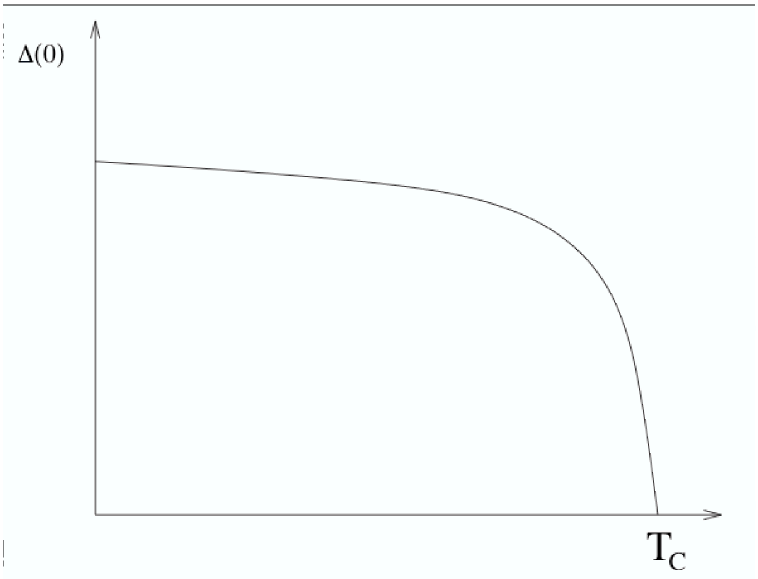
$$\frac{\Delta}{k_B T_c} = \frac{2}{1.13} = 1.76$$

$\Delta$ and $T_c$ can be measured independently, and this relation is well verified in many superconductors. This is one of the confirmations of BCS theory.

Besides, this theory predicts that $\Delta(T)$ behaves as:

$$\Delta(T) \simeq 3.2\, T_c \sqrt{1 - \frac{T}{T_c}}$$

The exponent $\frac{1}{2}$ is typical of mean field theory.



## 5.5 Electrodynamics of superconductors

The London equation

Before we embark on the calculation of the response of a superconductor to an external field, it is useful to start with some remarks on the constitutive equations which, together with Maxwell's equations, could describe the basic properties of superconductors.

There are two such properties:

1) Vanishing resistivity (1911)

2) Meissner(-Ochsenfeld) effect (1933) : the magnetic field vanishes inside a superconductor. It goes from the external value to 0 over a characteristic length $\lambda$ called the "penetration length".

Let us look in more detail at the consequences of these properties on the constitutive equations:

1) Vanishing resistivity:

In a metal, one generally assumes that the motion of an electron is governed by the equation:

$$m\frac{d\vec{v}}{dt} + m\frac{\vec{v}}{\tau} = e\vec{E}$$

where $\tau$ is a phenomenological relaxation time that one can for instance calculate in the context of the electron-phonon interaction. The general solution is the sum of a particular solution, for example $\vec{v} = \frac{e\tau}{m}\vec{E}$, and of the general solution of the equation without the right hand side:

$$\vec{v} = \vec{v}_0 \, e^{-\frac{t}{\tau}}$$
$$\Rightarrow \vec{v} = \vec{v}_0 \, e^{-\frac{t}{\tau}} + \frac{e\tau}{m}\vec{E}$$

After a time $t \gg \tau$, one thus gets:

$$\vec{v} = \frac{e\tau}{m}\vec{E}.$$

But $\vec{j} = ne\vec{v}$

$$\Rightarrow \vec{j} = \frac{ne^2\tau}{m}\vec{E}$$

or $\vec{j} = \sigma\vec{E}$, with $\sigma = \frac{ne^2\tau}{m}$.

This is the constitutive relation of metals.

In superconductors, $\sigma = \infty$ $(R = 0) \Rightarrow \tau = +\infty$

$$\Rightarrow m\frac{d\vec{v}}{dt} = e\vec{E}$$

But $\vec{j} = n_s e\vec{v}$, where $n_s$ is the superfluid density, i.e. the density of electrons that take part into superconductivity, so that

$$m\left(\frac{1}{n_s e}\right)\frac{d\vec{j}}{dt} = e\vec{E}$$

or

$$\vec{E} = \frac{m}{n_s e^2}\frac{d\vec{j}}{dt}$$

This relation only fixes the current up to a constant. It is thus not sufficient to describe the response of superconductors.

2) Meissner effect:

If one combines this equation with Maxwell equation

$$\vec{E} = -\frac{1}{c}\frac{\partial \vec{A}}{\partial t}$$

one finds:

$$\begin{aligned} \frac{m}{n_s e^2}\frac{d}{dt}\vec{j} = -\frac{1}{c}\frac{\partial \vec{A}}{\partial t} \\ \Rightarrow \frac{m}{n_s e^2}\vec{j} + \frac{1}{c}\vec{A} = \vec{c} \end{aligned}$$

where $\vec{c}$ is a constant of integration independent of the time. The Meissner effect is a consequence of this equation if $\vec{\text{rot}}\ \vec{c} = \vec{0}$. Indeed, Maxwell's equations imply

$$\begin{aligned} \vec{\text{rot}}\vec{B} &= \frac{4\pi}{c}\vec{j} \\ \Rightarrow \vec{\text{rot}}\ \vec{\text{rot}}\vec{B} &= \frac{4\pi}{c}\vec{\text{rot}}\vec{j} \end{aligned}$$

But $\vec{\text{rot}}\ \vec{\text{rot}}\vec{B} = -\Delta\vec{B} + \underbrace{\vec{\text{grad}}\ \text{div}\vec{B}}_{=0}$ and $\vec{\text{rot}}\vec{j} = -\frac{n_s e^2}{mc}\underbrace{\vec{\text{rot}}\vec{A}}_{\vec{B}}$. So the equation relating $\vec{j}$ to $\vec{A}$ implies

$$\Delta\vec{B} = \frac{4\pi n_s e^2}{mc^2}\vec{B} + \vec{\text{rot}}\ \vec{c}$$

If $\vec{\text{rot}}\ \vec{c} = \vec{0}$, we get

$$\Rightarrow \Delta\vec{B} = \frac{n_s e^2 4\pi}{mc^2}\vec{B}$$

Let us define

$$\frac{1}{\lambda^2} = \frac{n_s e^2 4\pi}{mc^2}$$

We thus have:

$$\Delta\vec{B} = \frac{1}{\lambda^2}\vec{B}$$

Let us adopt a geometry where $\vec{B} = \vec{B}(z)$. Integrating this equation leads to:

$$\Rightarrow \vec{B} = \vec{B}(0)e^{-\frac{z}{\lambda}}$$

The magnetic field $\vec{B}$ decreases exponentially over a characteristic length given by $\lambda$.

Since $\rho = 0$ in a metal, $\text{div}\vec{j} = 0$. Thus, $\frac{1}{c}\text{div}\vec{A} = \text{div}\ \vec{c}$. If one chooses a gauge in which $\text{div}\vec{A} = 0$ and $\vec{A} \to \vec{0}$ in the bulk, one can choose $\vec{c} = \vec{0}$. Under these conditions, $\vec{j}$ and $\vec{A}$ are related by:

$$\vec{j} = -\frac{n_s e^2}{mc}\vec{A} = -\frac{c}{\lambda^2 4\pi}\vec{A}$$

This equation is known as London equation (F. et H. London, 1935). This equation is valid in the London gauge:

- $\mathrm{div}\vec{A} = 0$.
- $\vec{A} \to \vec{0}$ in the bulk.

Note : the Meissner effect "implies" in a certain sense the absence of resistivity. Indeed, the field vanishes inside a superconductor because the external field is compensated by a field created at the surface of the sample. But a magnetic field does not produce work ($\vec{F} = e\vec{v} \wedge \vec{B} \perp \vec{v}$). For the current to be able to stay, it has to be able to circulate without dissipation. We will thus, without loss of generality, concentrate on the response to a static magnetic field.

Response to a static magnetic field:

Let us consider a static field in the gauge $\mathrm{div}\vec{A} = 0$.

$$\vec{p} = m\vec{v} + \frac{e}{c}\vec{A}$$

$$\begin{aligned}
\Rightarrow H_{cin} &= \sum_i \left(\frac{1}{2}mv_i^2\right) = \frac{1}{2m}\sum_i \left(\vec{p}_i - \frac{e}{c}\vec{A}_i\right)^2 \\
&= \frac{1}{2m}\sum_i \left(-i\hbar\vec{\nabla}_i - \frac{e}{c}\vec{A}(\vec{r}_i)\right)^2 \\
&= \underbrace{H^0_{cin}}_{\frac{1}{2m}\sum_i \vec{p}_i^2} \underbrace{+\frac{ie\hbar}{2mc}\sum_i \left(\vec{\nabla}_i \cdot \vec{A}(\vec{r}_i) + \vec{A}(\vec{r}_i)\cdot\vec{\nabla}_i\right)}_{H_1} + O(A^2)
\end{aligned}$$

The term $O(A^2)$ can be ignored for the calculation of $\vec{j}$ to first order.

Let us introduce the Fourier transform of the vector potential $\vec{A}(\vec{r}) = \sum_{\vec{q}}\vec{a}(\vec{q})e^{i\vec{q}\cdot\vec{r}}$, and let us determine the form of $H_1$ in second quantization. For the term $\vec{\nabla}_i \cdot \vec{A}(\vec{r}_i)$, we have :

$$\begin{aligned}
&\textstyle\sum_{\vec{q}} \int \frac{e^{-i\vec{k}'\cdot\vec{r}}}{\sqrt{V}}\vec{\nabla}\cdot\vec{a}(\vec{q})e^{i\vec{q}\cdot\vec{r}}\frac{e^{i\vec{k}\cdot\vec{r}}}{\sqrt{V}}d\vec{r} \\
= &\textstyle\sum_{\vec{q}} \frac{1}{V}\int e^{-i\vec{k}'\cdot\vec{r}}i(\vec{k}+\vec{q})\cdot\vec{a}(\vec{q})e^{i(\vec{k}+\vec{q})\cdot\vec{r}}d\vec{r}
\end{aligned}$$

Since the field is assumed to be transverse ($\mathrm{div}\vec{A} = 0$), we have $\vec{q}\cdot\vec{a}(\vec{q}) = 0$, and we get :

$$\sum_{\vec{q}} i\vec{k}\cdot\vec{a}(\vec{q})\delta_{\vec{k}',\vec{k}+\vec{q}}$$

The contribution to $H_1$ is thus given by :

$$-\frac{e\hbar}{2mc}\sum_{\vec{k},\vec{q},\sigma}\vec{k}\cdot\vec{a}(\vec{q})c^+_{\vec{k}+\vec{q},\sigma}c_{\vec{k},\sigma}$$

For the term $\vec{A}(\vec{r_i})\cdot\vec{\nabla}_i$, we have :

$$\begin{aligned} & \textstyle\sum_{\vec{q}} \quad \int e^{-i\vec{k}'\cdot\vec{r}}\vec{a}(\vec{q})\int e^{i\vec{q}\cdot\vec{r}}\cdot\vec{\nabla}e^{i\vec{k}\cdot\vec{r}}d\vec{r} \\ = & \textstyle\sum_{\vec{q}} \quad \int e^{-i\vec{k}'\cdot\vec{r}}i\vec{k}\cdot\vec{a}(\vec{q})e^{i(\vec{k}+\vec{q})\cdot\vec{r}}d\vec{r} \end{aligned}$$

It gives the same contribution, and one finally gets:

$$H_1=\frac{-e\hbar}{mc}\sum_{\vec{k},\vec{q},\sigma}\vec{k}\cdot\vec{a}(\vec{q})\;c^{+}_{\vec{k}+\vec{q},\sigma}c_{\vec{k},\sigma}$$

The current operator can be written :

$$\vec{j}(\vec{r})=\frac{e}{2m}\sum_{\sigma}\left\{\Psi^{+}(\vec{r},\sigma)\left(-i\hbar\vec{\nabla}-\frac{e}{c}\vec{A}\right)\Psi(\vec{r},\sigma)+\left[\left(i\hbar\vec{\nabla}-\frac{e}{c}\vec{A}\right)\Psi^{+}(\vec{r},\sigma)\right]\Psi(\vec{r},\sigma)\right\}$$

where $\Psi(\vec{r},\sigma)$, $\Psi^{+}(\vec{r},\sigma)$ are field operators (see Appendix on second quantization and exercises). They can be exanded in Fourier series:

$$\Psi^{+}(\vec{r},\sigma)=\sum_{\vec{k}_1}e^{-i\vec{k}_1\cdot\vec{r}}c^{+}_{\vec{k}_1,\sigma};\quad \Psi(\vec{r},\sigma)=\sum_{\vec{k}_2}e^{i\vec{k}_2\cdot\vec{r}}c_{\vec{k}_2,\sigma}$$

as well as their gradient:

$$\nabla\Psi^{+}(\vec{r},\sigma)=\sum_{\vec{k}_1}(-i\vec{k}_1)e^{-i\vec{k}_1\cdot\vec{r}}c^{+}_{\vec{k}_1,\sigma};\quad \nabla\Psi(\vec{r},\sigma)=\sum_{\vec{k}_2}i\vec{k}_2e^{i\vec{k}_2\cdot\vec{r}}c_{\vec{k}_2,\sigma}$$

It is convenient to decompose the current into two contributions:

$$\vec{j}(\vec{r})=\vec{J}_1(\vec{r})+\vec{J}_2(\vec{r})$$

with

$$\vec{J}_1(\vec{r})=\frac{-ie\hbar}{2m}\sum_{\sigma}\left\{\Psi^{+}(\vec{r},\sigma)\vec{\nabla}\Psi(\vec{r},\sigma)-(\vec{\nabla}\Psi^{+}(\vec{r},\sigma))\Psi(\vec{r},\sigma)\right\}$$

and

$$\vec{J}_2(\vec{r})=\frac{-e^2}{mc}\vec{A}\sum_{\sigma}\Psi^{+}(\vec{r},\sigma)\Psi(\vec{r},\sigma)$$

The Fourier transform of $\vec{J}_1(\vec{r})$ is given by:

$$\begin{aligned} \vec{J}_1(\vec{q}) &= \int d\vec{r}\;e^{-i\vec{q}\cdot\vec{r}}J_1(\vec{r}) \\ &= \frac{-ie\hbar}{2m}\sum_{\sigma}\left(\int d\vec{r}e^{-i\vec{q}\cdot\vec{r}}\sum_{\vec{k}_1,\vec{k}_2}\left[e^{-i\vec{k}_1\cdot\vec{r}+i\vec{k}_2\cdot\vec{r}}(i\vec{k}_2)\;c^{+}_{\vec{k}_1,\sigma}c_{\vec{k}_2,\sigma}-e^{-i\vec{k}_1\cdot\vec{r}+i\vec{k}_2\cdot\vec{r}}(-i\vec{k}_1)\;c^{+}_{\vec{k}_1,\sigma}c_{\vec{k}_2,\sigma}\right]\right) \\ &= \frac{-ie\hbar}{2m}\sum_{\vec{k},\sigma}\left[i(\vec{k}+\vec{q})\;c^{+}_{\vec{k},\sigma}c_{\vec{k}+\vec{q},\sigma}+i\vec{k}\;c^{+}_{\vec{k},\sigma}c_{\vec{k}+\vec{q},\sigma}\right] \\ &= \frac{e\hbar}{m}\sum_{\vec{k},\sigma}(\vec{k}+\frac{1}{2}\vec{q})\;c^{+}_{\vec{k},\sigma}c_{\vec{k}+\vec{q},\sigma} \end{aligned}$$

Besides,

$$\begin{aligned}\vec{J}_2(\vec{r}) &= \frac{-e^2}{mc}\vec{A}\underbrace{\sum_\sigma \Psi^+(\vec{r},\sigma)\Psi(\vec{r},\sigma)}_{\text{density}}\\ &= -\frac{e^2 n(\vec{r})}{mc}\vec{A}\end{aligned}$$

$\vec{J}_2$ is proportional to $\vec{A}$, and this relation is similar to the London equation, but it involves the <u>total</u> density and not the superconducting density. The term $\vec{J}_1$ must significantly reduce this contribution and even cancel it altogether in a normal metal!

If one wishes to determine the full response, one can express $H_1$ with the help of $\vec{J}_1$ and calculate the response in the context of linear response theory. This calculation is complicated (although not difficult), and here we will limit ourselves to the calculation of the expectation value of $\vec{J}_1(\vec{q})$ in the limit $\vec{q} \to 0$ (large wavelength). In this case, a simpler calculation is possible.

Let us start be writing the operators $H_1$ and $\vec{J}_1(\vec{q})$ in terms of the Bogolioubov operators:

$$\begin{aligned}H_1 &= \frac{-e\hbar}{mc}\sum_{\vec{k},\vec{q}}\vec{k}\cdot\vec{a}(\vec{q})\Big[(u_{\vec{k}+\vec{q}}\alpha^+_{\vec{k}+\vec{q},\uparrow} + v_{\vec{k}+\vec{q}}\alpha_{-\vec{k}-\vec{q},\downarrow})\\ &\times\ (u_{\vec{k}}\alpha_{\vec{k},\uparrow} + v_{\vec{k}}\alpha^+_{-\vec{k},\downarrow}) \underbrace{-}_{\vec{k}\to-\vec{k}-\vec{q}} (u_{\vec{k}}\alpha^+_{-\vec{k},\downarrow} - v_{\vec{k}}\alpha_{\vec{k},\uparrow})\\ &\times\ (u_{\vec{k}+\vec{q}}\alpha_{-\vec{k}-\vec{q},\downarrow} + v_{\vec{k}+\vec{q}}\alpha^+_{\vec{k}+\vec{k},\uparrow})\Big]\\ &= \frac{-e\hbar}{mc}\sum_{\vec{k},\vec{q}}\vec{k}\cdot\vec{a}(\vec{q})\Big[(u_{\vec{k}}u_{\vec{k}+\vec{q}} + v_{\vec{k}}v_{\vec{k}+\vec{q}})\\ &\times\ (\alpha^+_{\vec{k}+\vec{q},\uparrow}\alpha_{\vec{k},\uparrow} - \alpha^+_{-\vec{k},\downarrow}\alpha_{-\vec{k}-\vec{q},\downarrow})\\ &+\ (v_{\vec{k}}u_{\vec{k}+\vec{q}} - u_{\vec{k}}v_{\vec{k}+\vec{q}})\\ &\times\ (\alpha^+_{\vec{k}+\vec{q},\uparrow}\alpha^+_{-\vec{k},\downarrow} - \alpha_{-\vec{k}-\vec{q},\downarrow}\alpha_{\vec{k},\uparrow})\Big]\end{aligned}$$

In the limit $\vec{q}\to\vec{0}$, the second term disappears, and the first term gives:

$$\begin{aligned}H_1 &= \frac{-e\hbar}{mc}\sum_{\vec{k}}\vec{k}\cdot\vec{a}(\vec{0})(\alpha^+_{\vec{k},\uparrow}\alpha_{\vec{k},\uparrow} - \alpha^+_{-\vec{k},\downarrow}\alpha_{-\vec{k},\downarrow})\\ &\Rightarrow \left\{\begin{array}{l} E_{\vec{k},\uparrow} = E_{\vec{k}} - \frac{e\hbar}{mc}\vec{k}\cdot\vec{a}(\vec{0})\\ E_{\vec{k},\downarrow} = E_{\vec{k}} + \frac{e\hbar}{mc}\vec{k}\cdot\vec{a}(\vec{0})\end{array}\right.\end{aligned}$$

Similarly, in the limit $\vec{q}\to\vec{0}$, we have:

$$\vec{J}_1(\vec{q}\to\vec{0}) = \frac{e\hbar}{m}\sum_{\vec{k}}\vec{k}(\alpha^+_{\vec{k},\uparrow}\alpha_{\vec{k},\uparrow} - \alpha^+_{-\vec{k},\downarrow}\alpha_{-\vec{k},\downarrow})$$

But $< \alpha^+_{\vec{k},\sigma}\alpha_{\vec{k},\sigma} >= f(E_{\vec{k},\sigma})$

$$\Rightarrow < \vec{J}_1 >= \frac{e\hbar}{m}\sum_{\vec{k}}\vec{k}\left[f(E_{\vec{k},\uparrow}) - f(E_{\vec{k},\downarrow})\right]$$

Besides,

$$\begin{aligned} f(E_{\vec{k},\uparrow}) - f(E_{\vec{k},\downarrow}) &\simeq \frac{2e\hbar}{mc}\vec{k}\cdot\vec{a}(\vec{0})\left(-\frac{\partial f}{\partial E_{\vec{k}}}\right) \\ \Rightarrow < \vec{J}_1(\vec{q}\rightarrow\vec{0}) > &= \frac{2e^2\hbar^2}{m^2c}\sum_{\vec{k}}(\vec{k}\cdot\vec{a}(\vec{0}))\vec{k}\left(-\frac{\partial f}{\partial E_{\vec{k}}}\right) \end{aligned}$$

Let us suppose that $\vec{J}_1 \parallel \vec{a}(\vec{0})$, and that $k \simeq k_F$

$$\Rightarrow < \vec{J}_1(\vec{q}\rightarrow\vec{0}) >\simeq \frac{2e^2\hbar^2}{m^2c}\frac{k_F^2}{3}\sum_{\vec{k}}\left(-\frac{\partial f}{\partial E_{\vec{k}}}\right)\vec{a}(\vec{0})$$

where the factor $\frac{1}{3}$ comes from the average of the cosines. Indeed, let us assume that $\vec{N} = \sum_{\vec{k}}(\vec{k}\cdot\vec{a}(\vec{0}))\vec{k}h(k)$, where $h$ is an arbitrary function:

$$\begin{aligned} \vec{N}\cdot\vec{a}(\vec{0}) &= \sum_{\vec{k}}(\vec{a}(\vec{0})\cdot\vec{k})^2h(k) \\ &= \int dk\int 2\pi\sin\theta(\vec{a}(\vec{0})k\cos\theta)^2d\theta k^2h(k) \\ &= \int 2\pi\sin\theta\cos^2\theta d\theta\int dka^2(0)k^2k^2h(k) \\ &= 2\pi\frac{2}{3}\int dka^2(0)k^4h(k) \\ &= \frac{1}{3}2\pi\int_0^\pi\sin\theta d\theta\int dka^2(0)k^4h(k) \\ &= \frac{1}{3}\int d\vec{k}a^2(0)k^2h(k) \end{aligned}$$

Inserting $\frac{\hbar^2k_F^2}{2m} = \epsilon_F$, $g(\epsilon_F) = \frac{3n}{2\epsilon_F}$ and $\sum_{\vec{k}} = \frac{1}{2}g(\epsilon_F)\int d\epsilon$, one gets:

$$\begin{aligned} < \vec{J}_1(\vec{q}\rightarrow\vec{0}) > &\simeq \frac{2e^2}{3mc}\epsilon_Fg(\epsilon_F)\int_{-\infty}^{+\infty}d\epsilon\left(-\frac{\partial f}{\partial E_{\vec{k}}}\right)\vec{a}(\vec{0}) \\ &\simeq \frac{e^2}{mc}n\int_{-\infty}^{+\infty}d\epsilon\left(-\frac{\partial f}{\partial E_{\vec{k}}}\right)\vec{a}(\vec{0}) \end{aligned}$$

The total current can thus be written:

$$\begin{aligned} < \vec{j} > &= < \vec{J}_1 + \vec{J}_2 > \\ &= -\frac{e^2n}{mc}\vec{a}(\vec{0})\left[1 - \int_{-\infty}^{+\infty}d\epsilon\left(-\frac{\partial f}{\partial E_{\vec{k}}}\right)\right] \end{aligned}$$

with $E = \sqrt{\epsilon^2 + \Delta^2}$.

If $\Delta = 0$

$$\begin{aligned} \int_{-\infty}^{+\infty} d\epsilon \left(\frac{\partial f}{\partial E_{\vec{k}}}\right) &= \int_{-\infty}^{+\infty} dE \left(\frac{\partial f}{\partial E}\right) = -1 \\ &\Rightarrow \; <\vec{J}>=\vec{0} \end{aligned}$$

But if $\Delta \neq 0$,

$$-\int_{-\infty}^{+\infty} d\epsilon \left(\frac{\partial f}{\partial E_{\vec{k}}}\right) = 2\int_{\Delta}^{+\infty} dE \left(-\frac{\partial f}{\partial E}\frac{E}{\sqrt{E^2-\Delta^2}}\right) < 1$$

If one defines the density of normal fluid by

$$n_n = n\int_{-\infty}^{+\infty} d\epsilon \left(\frac{\partial f}{\partial E_{\vec{k}}}\right)$$

and the superfluid density by

$$n_s = n - n_n$$

one recovers London equation

$$\vec{j} \; = \; -\frac{e^2 n_s}{mc}\vec{a}(\vec{0})$$

N.B.:

$$f(E) \; = \; \frac{1}{e^{\beta E}+1}$$

At $T = 0$, $\frac{-\partial f}{\partial E} = \delta(E)$, and if there is a gap, $n_n = 0$ and $n_s(T = 0) = n$. This is true, but in this case it is important to distinguish between the wavevectors that participate to superconductivity and those who don't. The term $\vec{J}_1$ cancels anyway the contribution of $\vec{J}_2$ coming from the electrons that do not participate to superconductivity. $n$ is thus the density of electrons which are at $\omega_D$ of $\epsilon_F$.

Critical field and field penetration :

Let us now discuss in more detail the way a magnetic field is expelled from a superconducting sample. As we shall see, this is in general a very rich phenomenology that goes far beyond the mere decay of the field over the penetration depth.

Let us start by showing that superconductivity must be destroyed over a characteristic length $\xi_0$ that is *not* related to the penetration depth $\lambda$. According to the Heisenberg uncertainty relation,

$$\Delta p \Delta x \geq \hbar$$

But to build a wave-packet for a superconducting electron, we must have

$$\epsilon_F - \Delta < \frac{p^2}{2m} < \epsilon_F + \Delta \quad \text{with} \quad \epsilon_F = \frac{p_F^2}{2m}$$

Let us define

$$\frac{p_2^2}{2m} = \epsilon_F + \Delta$$

and

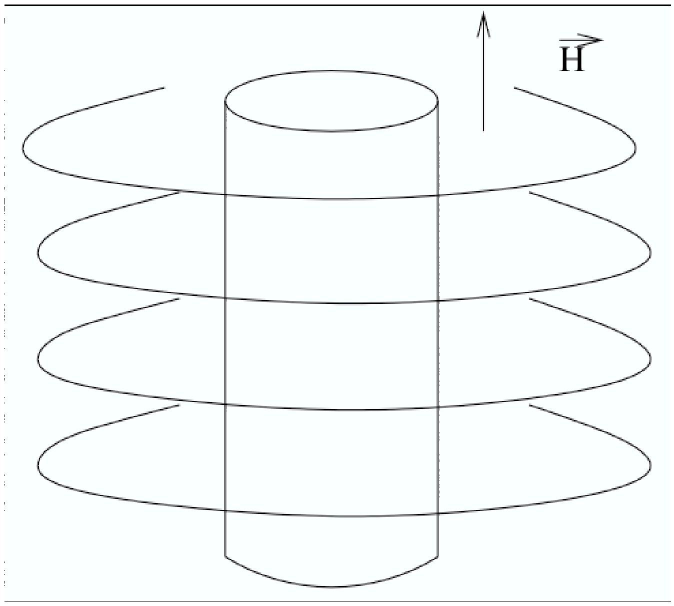
$$\frac{p_1^2}{2m} = \epsilon_F - \Delta$$

$$\begin{aligned} \frac{1}{2m}(p_2^2 - p_1^2) &= \frac{1}{2m}\underbrace{(p_2 - p_1)}_{\Delta p}\underbrace{(p_2 + p_1)}_{2p_F} = 2\Delta \\ &\Rightarrow \Delta p = 2\frac{\Delta m}{p_F} = \frac{2\Delta}{v_F} \\ &\Rightarrow \Delta x \geq \frac{\hbar v_F}{2\Delta} \simeq \xi_0 \end{aligned}$$

It is thus impossible to localize the particle over a distance smaller than $\xi_0$, and the superconducting state must be affected over lengths of order $\xi_0$.

There is another, maybe more intuitive way to derive this length. Indeed, it can also be seen as the size of a Cooper pair (see exercises).

The existence of two characteristic lengths is responsible for the very rich phenomenology of superconductors in a magnetic field. Let us first note that expelling the magnetic field is only possible as long as the external field is not too strong. To be concrete, let us consider a cylinder of metal inside a coil that creates a uniform field $\vec{H}$.



If one cools down the metal below its critical temperature, say for simplicity at $T = 0$, it can become superconducting. Let's look at the energy balance. If it becomes superconducting, its energy with respect to the normal state is equal to $E_{\text{cond}}$/volume unit. For this to be possible however, one has to expel the magnetic field, i.e. to create a surface current that opposes the external field. This costs an energy $\int \vec{j} \cdot \vec{A}_{ext} d\tau$ whose density is $\frac{H^2}{8\pi}$. So the energy per volume unit in the superconducting state is

$$\frac{H^2}{8\pi} - |E_{\text{cond}}|$$

It is advantageous to transit into the superconducting state as long as $H < H_c$, with

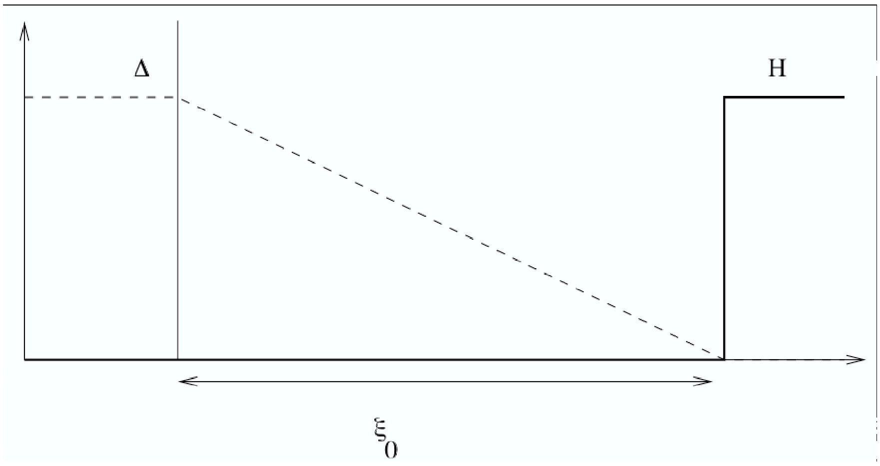
$$H_c^2 = 8\pi |E_{\text{cond}}|$$

However, the penetration of the magnetic field inside the sample is in general neither uniform nor progressive. This depends on the shape of the sample, but more importantly on the characteristics of the metal. In particular, the relative value of $\xi_0$ and $\lambda_L$ is crucial. It turns out that this varies considerably from one compound to the next. The description of this phenomenology is more compact in the context of the Landau-Ginsburg theory. But the basic idea can be explained in simple terms. To create an interface between a normal region and a superconducting one allows one to gain magnetic energy but has a cost in superconducting energy. If $\xi_0 \gg \lambda$, the cost of the interface is large, and the system will try to avoid creating them as much as possible. If, by contrast, $\xi_0 \ll \lambda$, an interface allows one to gain energy and, if the normal domain is small, the magnetic field can penetrate for values of the field much smaller than $H_c$. To increase the surface while keeping a volume as small as possible, the magnetic field penetrates as small cylinders known as vortices.

Estimation of the interface energy:

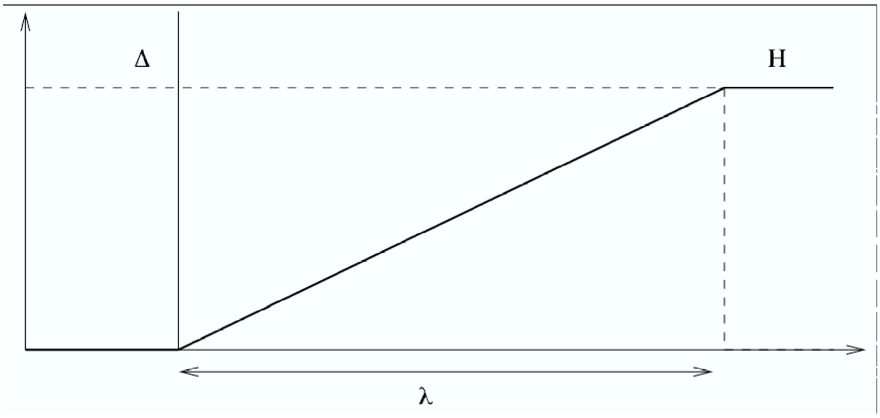
Let us now perform a rough estimation of the energy of an interface.

$\boxed{\xi_0 \gg \lambda}$



$$\frac{E}{S} = \frac{H^2}{8\pi}\xi_0 - |E_{\text{cond}}|\frac{\xi_0}{2} > 0 \quad \text{for} \quad H = H_c$$

$\boxed{\xi_0 \ll \lambda}$



$$\frac{E}{S} = \frac{H^2}{8\pi}\frac{\lambda}{2} - |E_{\text{cond}}|\lambda < 0 \quad \text{for} \quad H = H_c$$

The field can then penetrate as vortices for fields $H$ much smaller than $H_c$.

## 5.6 The Landau-Ginsburg theory

The most efficient description of the macroscopic properties of superconductors has been developed before the BCS theory. It is due to Landau and Ginsburg who postulated in 1951 the form of the free energy as a function of a complex order parameter that describes:

- The superconducting density $|\Psi|^2$;
- The coupling to the electromagnetic field through its phase

  $$\vec{A}' = \vec{A} + \vec{\nabla}\chi \Rightarrow \Psi' = \Psi e^{ie^* \frac{\chi}{hc}}$$

  under a gauge transformation.

The free-energy density can thus be written :

$$f = f_n + \alpha|\Psi|^2 + \frac{\beta}{2}|\Psi|^4 + \frac{1}{2m}\left|\left(-i\hbar\vec{\nabla} - \frac{e^*\vec{A}}{c}\right)\Psi\right|^2 + \frac{B^2}{8\pi}$$

The first two terms are the usual ones in a Landau expansion, the third one is the form of the gradient that is invariant under a gauge transformation for a complex order parameter.

Since this functional was introduced by Landau and Ginsburg, one has shown using Green functions (Gorkov, 1959) that this form of the free energy was a consequence of BCS, with $e^* = 2e$. This proof goes beyond the scope of these lectures, but the identification of $\Psi$ with $\Delta$ naturally leads to this form. Indeed, we $\Delta = \frac{V}{N}\sum_{\vec{k}}\langle c^+_{\vec{k},\uparrow} c^+_{-\vec{k},\downarrow}\rangle$ could *a priori* take complex values. For a homogeneous system without a magnetic field, the phase does not play any role. It drops from all physical properties, and it is constant. That is why we have chosen the real solution. However, when there is a magnetic field, the phase of the wave function is related to the choice of gauge. Indeed, the Schrödinger equation of a free particle in a field is given by.:

$$H\Psi = \frac{1}{2m}(\vec{p} - \frac{e}{c}\vec{A})^2\Psi = E\Psi$$

Let us perform the gauge transformation

$$\vec{A}' = \vec{A} + \vec{\nabla}\chi(\vec{r})$$

The Schrödinger equation

$$\frac{1}{2m}(\vec{p} - \frac{e}{c}\vec{A}')^2\Phi = E\Phi$$

is satisfied by $\Phi(\vec{r}) = e^{\frac{ie}{\hbar c}\chi(\vec{r})}\Psi(\vec{r})$.

Proof:

$$\begin{aligned}
(-i\hbar\vec{\nabla} - \frac{e}{c}\vec{A}')\Phi(\vec{r}) &= (-i\hbar\vec{\nabla} - \frac{e}{c}\vec{A}')e^{\frac{ie}{\hbar c}\chi(\vec{r})}\Psi(\vec{r}) \\
&= -\frac{e}{c}\vec{A}'\Phi(\vec{r}) - i\hbar\vec{\nabla}(e^{\frac{ie}{\hbar c}\chi(\vec{r})}\Psi(\vec{r})) \\
&= -\frac{e}{c}\vec{A}'\Phi(\vec{r}) - (i\hbar\frac{ie}{\hbar c}\vec{\nabla}\chi(\vec{r}))\Phi(\vec{r}) + e^{\frac{ie}{\hbar c}\chi(\vec{r})}(-i\hbar\vec{\nabla})\Psi(\vec{r}) \\
&= -\frac{e}{c}\vec{A}\Phi(\vec{r}) + e^{\frac{ie}{\hbar c}\chi(\vec{r})}(-i\hbar\vec{\nabla})\Psi(\vec{r}) \\
&= e^{\frac{ie}{\hbar c}\chi(\vec{r})}(\vec{p} - \frac{e}{c}\vec{A})\Psi(\vec{r})
\end{aligned}$$

To get the Hamiltonian, on must again apply $(-i\hbar\vec{\nabla} - \frac{e}{c}\vec{A}')$, performing the scalar product. Let us suppose that we have a vector, and not a scalar:

$$(-i\hbar\vec{\nabla} - \frac{e}{c}\vec{A}')e^{\frac{ie}{\hbar c}\chi(\vec{r})}\vec{F}(\vec{r}) = -\frac{e}{c}\vec{A}'e^{\frac{ie}{\hbar c}\chi(\vec{r})}\vec{F}(\vec{r}) - i\hbar\vec{\nabla}(e^{\frac{ie}{\hbar c}\chi(\vec{r})}\vec{F}(\vec{r}))$$

But $\vec{\nabla}\cdot(p\vec{A}) = p\vec{\nabla}\cdot\vec{A} + \vec{A}\cdot\vec{\nabla}p$

$$\begin{aligned}
\Rightarrow \vec{\nabla}\left(e^{\frac{ie}{\hbar c}\chi(\vec{r})}\vec{F}(\vec{r})\right) &= e^{\frac{ie}{\hbar c}\chi(\vec{r})}\vec{\nabla}\cdot\vec{F}(\vec{r}) + \vec{F}(\vec{r})\cdot\vec{\nabla}e^{\frac{ie}{\hbar c}\chi(\vec{r})} \\
&= e^{\frac{ie}{\hbar c}\chi(\vec{r})}\vec{\nabla}\cdot\vec{F}(\vec{r}) + e^{\frac{ie}{\hbar c}\chi(\vec{r})}\frac{ie}{\hbar c}\vec{\nabla}\chi(\vec{r})\cdot\vec{F}(\vec{r}) \\
\Rightarrow (-i\hbar\vec{\nabla} - \frac{e}{c}\vec{A}')e^{\frac{ie}{\hbar c}\chi(\vec{r})}\vec{F}(\vec{r}) &= e^{\frac{ie}{\hbar c}}(\vec{p} - \frac{e}{c}\vec{A})\cdot\vec{F}(\vec{r}) \\
\\
\Rightarrow H'\Phi &= e^{\frac{ie}{\hbar c}\chi(\vec{r})}H\Psi = e^{\frac{ie}{\hbar c}\chi(\vec{r})}E\Psi = E\Phi \qquad CQFD
\end{aligned}$$

This property implies that the kinetic energy

$$\int \Psi^*(\vec{r})\frac{1}{2m}(\vec{p} - \frac{e}{c}\vec{A})^2\Psi(\vec{r})d\vec{r}$$

is invariant under a gauge transformation. Besides, since $\vec{p} - \frac{e}{c}\vec{A}$ is hermitian, one can also write

$$\begin{aligned}
E_{\text{cin}} &= \frac{1}{2m}\int \Psi^*(\vec{r})(\vec{p} - \frac{e}{c}\vec{A})^2\Psi(\vec{r})d\vec{r} \\
&= \frac{1}{2m}\langle\Psi|(\vec{p} - \frac{e}{c}\vec{A})^2|\Psi\rangle \\
&= \frac{1}{2m}\langle\Psi|(\vec{p} - \frac{e}{c}\vec{A})^+(\vec{p} - \frac{e}{c}\vec{A})|\Psi\rangle \\
&= \frac{1}{2m}||(\vec{p} - \frac{e}{c}\vec{A})|\Psi\rangle||^2 \\
&= \frac{1}{2m}\int \left|(\vec{p} - \frac{e\vec{A}}{c})\Psi(\vec{r})\right|^2 d\vec{r}
\end{aligned}$$

Similarly, one can show that, for the gap $\Delta$, the form of the energy that is invariant under a gauge transformation is given in term of the gradient by :

$$\frac{1}{2m}\left|(-i\hbar\vec{\nabla}-\frac{2e\vec{A}}{c})\Delta\right|^2$$

The Landau-Ginsburg free energy is thus a functional of the fields $\vec{A}(\vec{r})$ and $\Psi(\vec{r})$ given by :

$$F[\vec{A},\Psi]=\int d\vec{r}\left(\alpha|\Psi|^2+\frac{\beta}{2}|\Psi|^4+\frac{1}{2m}\left|\left(-i\hbar\vec{\nabla}-\frac{2e\vec{A}}{c}\right)\Psi\right|^2+\frac{B^2}{8\pi}\right)$$

Since the vector potential and the order parameter can both vary in a superconductor, it is natural to minimize the free energy with respect to both fields. To get the equations that they must satisty, one just needs to extend the method used to derive the Lagrange equations from the principle of least action.

Let us start with a variation $\delta\vec{A}$ of the vector potential :

$$\begin{aligned}
F[\vec{A}+\delta\vec{A}] &= \int\frac{\left(\vec{\text{rot}}(\vec{A}+\delta\vec{A})\right)^2}{8\pi}d\vec{r}+\int d\vec{r}\left(\alpha|\Psi|^2+\frac{\beta}{2}|\Psi|^4\right)\\
&+ \int\frac{d\vec{r}}{2m}\left(-i\hbar\vec{\nabla}\Psi-2e\frac{\vec{A}+\delta\vec{A}}{c}\Psi\right)\left(i\hbar\vec{\nabla}\Psi^*-2e\frac{\vec{A}+\delta\vec{A}}{c}\Psi^*\right)\\
&\simeq F[\vec{A}]+\int\frac{\delta\vec{A}}{2m}\Bigg[-\frac{2e}{c}\Psi\left(i\hbar\vec{\nabla}\Psi^*-2e\frac{\vec{A}}{c}\Psi^*\right)\\
&- \frac{2e}{c}\Psi^*\left(-i\hbar\vec{\nabla}\Psi-2e\frac{\vec{A}}{c}\Psi\right)\Bigg]d\vec{r}\\
&+ \int\frac{(\vec{\text{rot}}\vec{A})\cdot\vec{\text{rot}}\delta\vec{A}}{4\pi}
\end{aligned}$$

But

$$\begin{aligned}
\text{div}(\vec{A}\times\vec{B}) &= \vec{B}\cdot\vec{\text{rot}}\vec{A}-\vec{A}\cdot\vec{\text{rot}}\vec{B}\\
&\Rightarrow \vec{B}\cdot\vec{\text{rot}}\delta\vec{A}=\vec{\text{rot}}\vec{B}\cdot\delta\vec{A}+\text{div}(\delta\vec{A}\times\vec{B})
\end{aligned}$$

and

$$\int\text{div}(\delta\vec{A}\times\vec{B})=\int\int_S\vec{n}\cdot(\delta\vec{A}\times\vec{B})d\sigma$$

One can limit oneself to variations such that $\delta\vec{A}=\vec{0}$ at the surface of the superconductor. Then,

$$\int\frac{1}{2m}\frac{\vec{B}\cdot\vec{\text{rot}}\delta\vec{A}}{4\pi}=\int\frac{1}{2m}\frac{\vec{\text{rot}}\vec{B}\cdot\delta\vec{A}}{4\pi}$$

In order for $F[\vec{A}+\delta\vec{A}]-F[\vec{A}]$ to vanish for all $\delta\vec{A}$, one must have $\vec{\mathrm{rot}}\vec{B}=\frac{4\pi}{c}\vec{j}$ with

$$\vec{j}=-\frac{ei\hbar}{m}\left(\Psi^*\vec{\nabla}\Psi-\Psi\vec{\nabla}\Psi^*\right)-\frac{4e^2\vec{A}}{mc}\Psi^*\Psi$$

Let us now make a variation around $\Psi$. For a function of a complex variable $f(z,\bar{z})$, the statement that the variations vanish is equivalent to saying that if $z=re^{i\theta}$,

$$\left\{\begin{array}{l}\frac{\partial f}{\partial r}=0\\ \frac{\partial f}{\partial \theta}=0\end{array}\right. \Rightarrow \left\{\begin{array}{l}\frac{\partial f}{\partial z}e^{i\theta}+\frac{\partial f}{\partial \bar{z}}e^{-i\theta}=0\\ \frac{\partial f}{\partial z}ire^{i\theta}+\frac{\partial f}{\partial \bar{z}}(-ir)e^{-i\theta}=0\end{array}\right. \Rightarrow \left\{\begin{array}{l}\frac{\partial f}{\partial z}=0\\ \frac{\partial f}{\partial \bar{z}}=0\end{array}\right.$$

if $r\neq 0$.

Let us thus express $F$ as a function of $(\Psi,\Psi^*)$, and let us write $\frac{\partial F}{\partial \Psi^*}=0$

$$F=\int d\vec{r}\left(\alpha\Psi\Psi^*+\frac{\beta}{2}\Psi^2\Psi^{*2}+\frac{1}{2m}(-i\hbar\vec{\nabla}-\frac{2e}{c}\vec{A})\Psi(i\hbar\vec{\nabla}-\frac{2e}{c}\vec{A})\Psi^*+\frac{B^2}{8\pi}\right)$$

$$\begin{aligned}F(\Psi,\Psi^*+\delta\Psi^*) &- F(\Psi,\Psi^*)\\ &= \int d\vec{r}\Big[\alpha\Psi+\beta|\Psi|^2\Psi+\frac{1}{2m}(-i\hbar\vec{\nabla}\Psi-\frac{2e}{c}\vec{A}\Psi)\\ &\times\ (i\hbar\vec{\nabla}-\frac{2e}{c}\vec{A})\Big]\delta\Psi^*=0\end{aligned}$$

An integration by part of the last term leads to

$$\vec{\nabla}\Psi\cdot\vec{\nabla}\delta\Psi^*=\mathrm{div}(\delta\Psi^*\vec{\nabla}\Psi)-\delta\Psi^*\underbrace{\mathrm{div}\vec{\nabla}\Psi}_{\Delta\Psi}$$

since $\vec{\nabla}(p\vec{A})=p\vec{\nabla}\cdot\vec{A}+\vec{A}\cdot\vec{\nabla}p$. So,

$$\begin{aligned}\int\vec{\nabla}\Psi\cdot\vec{\nabla}\delta\Psi^* &= \int\mathrm{div}(\delta\Psi^*\vec{\nabla}\Psi)-\int\delta\Psi^*\Delta\Psi\\ &= \int_S\delta\Psi^*\vec{n}\cdot\vec{\nabla}\Psi-\int\delta\Psi^*\Delta\Psi\end{aligned}$$

But $\int_S\delta\Psi^*\vec{n}\vec{\nabla}\Psi=0$ if one limits oneself to $\delta\Psi^*$ that vanish at the surface. So,

$$\int -i\hbar\vec{\nabla}\Psi\cdot i\hbar\vec{\nabla}\delta\Psi^*=\int\left[(-i\hbar)^2\vec{\nabla}^2\Psi\right]\delta\Psi^*$$

Besides,

$$\begin{aligned}\int\Psi\vec{\nabla}\delta\Psi^* &= -\int\delta\Psi^*\vec{\nabla}\Psi+\int\vec{\nabla}(\Psi\delta\Psi^*)\\ &= -\int\delta\Psi^*\vec{\nabla}\Psi+\int_S\vec{n}(\Psi\delta\Psi^*)d\sigma\end{aligned}$$

But $\int_S \vec{n}(\Psi\delta\Psi^*)d\sigma = 0$ if $\delta\Psi^* = 0$ at the surface. Thus,

$$\begin{aligned}
&\int d\vec{r}\left(-i\hbar\vec{\nabla}\Psi - \frac{2e\vec{A}}{c}\Psi\right)\left(i\hbar\vec{\nabla} - \frac{2e\vec{A}}{c}\right)\delta\Psi^* \\
= &\int d\vec{r}\left[(-i\hbar\vec{\nabla})^2\Psi + \left(-\frac{2e\vec{A}}{c}\right)^2\Psi - i\hbar\left(-\frac{2e\vec{A}}{c}\right)\vec{\nabla}\Psi - \frac{2e\vec{A}}{c}(-i\hbar)\vec{\nabla}\Psi\right]\delta\Psi^* \\
= &\int d\vec{r}\delta\Psi^*\left(-i\hbar\vec{\nabla} - \frac{2e\vec{A}}{c}\right)^2\Psi
\end{aligned}$$

In order to ensure that $F(\Psi, \Psi^* + \delta\Psi^*) - F(\Psi, \Psi^*) = 0$ at first order in $\delta\Psi^*$, one must have

$$\alpha\Psi + \beta|\Psi|^2\Psi + \frac{1}{2m}\left(-i\hbar\vec{\nabla} - \frac{2e}{c}\vec{A}\right)^2\Psi = 0$$

Finally, the minimization of the functional with respect to the order parameter $\Psi$ and to the vector potential $\vec{A}$ leads to the following pair of equations, known as the Landau-Ginsburg equations:

$$\boxed{\vec{j} = -\frac{ei\hbar}{m}\left(\Psi^*\vec{\nabla}\Psi - \Psi\vec{\nabla}\Psi^*\right) - \frac{4e^2\vec{A}}{mc}\Psi^*\Psi}$$

$$\boxed{\alpha\Psi + \beta|\Psi|^2\Psi + \frac{1}{2m}\left(-i\hbar\vec{\nabla} - \frac{2e}{c}\vec{A}\right)^2\Psi = 0}$$

These equations constitute the basis of the description of the macroscopic properties of superconductors. They contain two characteristic lengths that depend on $\alpha$ and $\beta$. To see this, let us consider two limiting cases :

1) Let us assume that the external magnetic field vanishes. In the bulk of the material, $\Psi$ satisfies the equation

$$\alpha\Psi + \beta|\Psi|^2\Psi - \frac{\hbar^2}{2m}\Delta\Psi = 0$$

According to Landau's theory of phase transitions, $\alpha = \alpha_0(T - T_c)$, and so $\alpha < 0$ if $T < T_c$ and $\alpha > 0$ if $T > T_c$.

The equation above has in general two obvious solutions :

- $\Psi = 0 \quad (\forall T)$
- $\Psi^2 = -\frac{\alpha}{\beta}$ if $\alpha < 0$, i.e. if $T < T_c$ (one has assumed that $\Psi$ is real for simplicity).

In the superconducting phase, there is thus a solution $\Psi_0 = \sqrt{\frac{-\alpha}{\beta}} \neq 0$.

Let us now consider an interface $z = 0$ between a normal region $(z < 0)$ and a superconducting region $(z > 0)$, and let us still assume that $\Psi$ is real. The differential equation takes the form :

$$-\frac{\hbar^2}{2m}\Psi'' + \alpha\Psi + \beta\Psi^3 = 0$$

where $\Psi'' = \frac{\partial^2\Psi}{\partial z^2} = \frac{d^2\Psi}{dz^2}$ since $\Psi$ only depends on $z$. If one multiplies by $\Psi'$ and integrates, one gets

$$\begin{aligned} & - \frac{\hbar^2}{2m}\Psi''\Psi' + \alpha\Psi\Psi' + \beta\Psi^3\Psi' = 0 \\ \Rightarrow & -\frac{\hbar^2}{4m}\Psi'^2 + \frac{\alpha}{2}\Psi^2 + \frac{\beta}{4}\Psi^4 = cst \end{aligned}$$

For this equation to have a solution $\Psi^2 = \frac{-\alpha}{\beta}$ for $z \to \infty$, one must have :

$$\frac{-\alpha^2}{2\beta} + \frac{\beta}{4}\frac{\alpha^2}{\beta^2} = cst$$

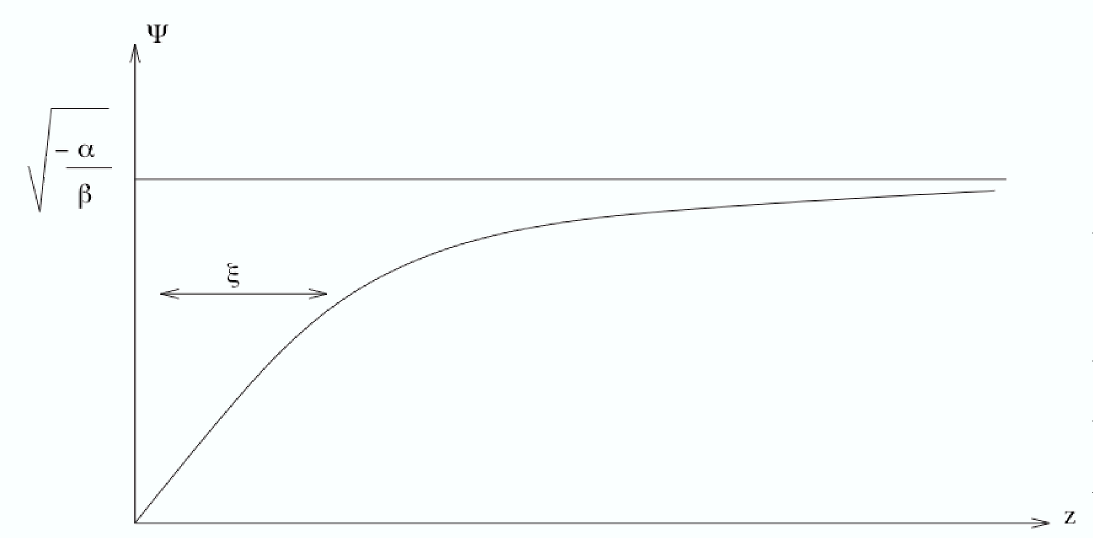
or $cst = -\frac{\alpha^2}{4\beta}$

$$\begin{aligned} \Rightarrow \frac{\hbar^2}{4m}\Psi'^2 &= \frac{\beta}{4}\Psi^4 + \frac{\alpha}{2}\Psi^2 + \frac{\alpha^2}{4\beta} \\ &= \frac{\beta}{4}\left(\Psi^4 + \frac{2\alpha}{\beta}\Psi^2 + \frac{\alpha^2}{\beta^2}\right) \\ &= \frac{\beta}{4}\left(\Psi^2 + \frac{\alpha}{\beta}\right)^2 \end{aligned}$$

$$\begin{aligned} \Rightarrow \Psi' &= -\frac{\sqrt{\beta m}}{\hbar}\left(\Psi^2 + \frac{\alpha}{\beta}\right) \\ \Psi &= \sqrt{\frac{-\alpha}{\beta}}\tanh\frac{z}{\sqrt{2}\xi} \end{aligned}$$

with $\xi^2 = \frac{\hbar^2}{2m|\alpha|}$ or $\xi = \frac{\hbar}{\sqrt{2m(-\alpha)}}$ (the factor $\sqrt{2}$ is conventional). Indeed,

$$\begin{aligned} \Psi' &= \sqrt{\frac{-\alpha}{\beta}}\frac{1}{\sqrt{2}\xi}\left(1 - \tanh^2\frac{z}{\sqrt{2}\xi}\right) \\ &= \sqrt{\frac{-\alpha}{\beta}}\frac{\sqrt{m(-\alpha)}}{\hbar}\left(1 + \frac{\beta}{\alpha}\Psi^2\right) \\ &= -\frac{\sqrt{\beta m}}{\hbar}\left(\Psi^2 + \frac{\alpha}{\beta}\right) \end{aligned}$$



This length is the analog of the coherence length. In the Landau-Ginsburg theory, its critical behaviour is given by $\xi(T) \sim \frac{1}{\sqrt{T_c - T}}$. This agrees with the microscopic expression $\xi = \frac{\hbar v_F}{\Delta}$ with $\Delta(T) \sim \sqrt{T_c - T}$.

2) Let us assume that the external magnetic field is very small, hence that the vector potential varies very slowly with distance. $\Psi$ is nearly constant, and $\vec{j}$ is approximately given by

$$\vec{j} = -\frac{4e^2}{mc}|\Psi_0|^2\vec{A}$$

To see this, one writes locally

$$\vec{A}(\vec{r}) = \underbrace{\vec{A}_0}_{\text{uniform}} + \delta\vec{A}(\vec{r})$$

and one expands in $\delta\vec{A}(\vec{r})$. To zeroth order, $\Psi$ is uniform according to the second Landau-Ginsburg equation, and the first equation then leads to the result.

Now this equation is nothing but London equation, with

$$\begin{aligned} -\frac{c}{4\pi\lambda^2} &= \frac{-4e^2}{mc}\frac{-\alpha}{\beta} \\ \Rightarrow \frac{1}{\lambda^2} &= \frac{16\pi e^2}{mc^2}\frac{-\alpha}{\beta} \end{aligned}$$

Like $\xi$, $\lambda \sim \frac{1}{\sqrt{T_c - T}}$ when $T \to T_c$.

This description thus agrees qualitatively with the discussion of the electrodynamics of superconductors of the previous section.

Critical field

As before, the critical field is obtained by equating the energy density of the field that one has to create to oppose the external field with the condensation energy:

$$\begin{aligned}\frac{H_c^2}{8\pi} &= -\left(\alpha|\Psi|^2 + \frac{\beta}{2}|\Psi|^4\right)\\ &= -\left(\alpha\frac{-\alpha}{\beta} + \frac{\beta}{2}\frac{\alpha^2}{\beta^2}\right) = \frac{\alpha^2}{\beta}\\ \Rightarrow H_c^2 &= \frac{4\pi\alpha^2}{\beta}\end{aligned}$$

Type I and type II superconductors:

We are now going to show that the ratio $\kappa = \frac{\lambda}{\xi}$ plays a crucial role in the properties of a superconductor. Let us assume that one decreases the magnetic field starting from a very strong magnetic field where the solution is given by $\Psi = 0$, and let us study at which field there exists a solution of the Landau-Ginsburg equations with $\Psi \neq 0$. More precisely, let us investigate the stability of the solution $\Psi = 0$ by studying if there are other solutions with $\Psi$ very small.

To first order in $\Psi$, the equation reads

$$\frac{1}{2m}\left(-i\hbar\vec{\nabla} - \frac{2e}{c}\vec{A}\right)^2\Psi = -\alpha\Psi$$

This equation is similar to the Schrödinger equation of a particle of charge $2e$ in a magnetic field $H$ described by $\vec{A}$. The energies are well known. They are given by

$$\left(n + \frac{1}{2}\right)\hbar\omega_c + \frac{p_z^2}{2m}$$

where $\omega_c$ is the cyclotron frequency

$$\omega_c = \frac{2eH}{mc}$$

When $H$ is very large, $\frac{\hbar\omega_c}{2} > -\alpha$, and the only solution is given by $\Psi = 0$. By when $H$ decreases, there is a solution as soon as $\frac{\hbar\omega_c}{2} = -\alpha$. There is thus another solution than $\Psi = 0$ as soon as $H < H_{c2}$ with

$$\begin{aligned}\frac{1}{2}\hbar\frac{2eH_{c2}}{mc} &= -\alpha\\ \Rightarrow H_{c2} &= -\frac{\alpha mc}{\hbar e}\end{aligned}$$

This field is not directly related to $H_c$, and one must distinguish between two cases:

$H_{c2} > H_c$ : For $H_c < H < H_{c2}$, it is not advantageous for the system to be everywhere in the normal state. It is better to take advantage of the non-zero solutions. However, the solution cannot correspond to a global expulsion of the magnetic field since $H > H_c$. Only part of the sample becomes superconducting. To determine the actual form of the global solution requires to go beyond the linear approximation. Indeed the Landau levels are degenerate, and the actual form of the lowest energy solution is governed by the next order terms that lift the degeneracy. One can show including such terms that these solutions build a lattice perpendicular to the surface of the sample ($H \perp$ surface) that encloses cylinders of normal metal. These cylinders are called vortices, and in general they form a triangular lattice.

The condition $H_{c2} > H_c$ can also be written

$$-\frac{\alpha mc}{\hbar e} > \sqrt{\frac{4\pi}{\beta}}(-\alpha)$$

or

$$\sqrt{\beta} > \frac{\sqrt{4\pi}\hbar e}{mc}$$

But

$$\begin{aligned}\kappa = \frac{\lambda}{\xi} &= \frac{\sqrt{m}c\sqrt{\beta}}{4\sqrt{\pi}e\sqrt{-\alpha}}\frac{\sqrt{2m(-\alpha)}}{\hbar} \\ &= \frac{mc\sqrt{\beta}}{\hbar e 2\sqrt{2}\sqrt{\pi}} = \sqrt{\beta}\frac{mc}{\sqrt{4\pi}\hbar e}\frac{1}{\sqrt{2}}\end{aligned}$$

This inequality can thus be written

$$\kappa > \frac{1}{\sqrt{2}}$$

In other words, $\frac{H_{c2}}{H_c} = \sqrt{2}\kappa$.

So, if $\xi < \sqrt{2}\lambda$, the system remains partially superconducting between $H_c$ and $H_{c2}$. Such superconductors are called type II superconductors.

$H_{c2} < H_c$ : The system undergoes a direct transition into a global superconductor. Such superconductors are called type I superconductors.

## 5.7 Flux quantization

Let us first show that the current only depends on the variation of the phase of the order parameter, and not on the variation of its amplitude. Indeed, if one writes

$$\Psi = |\Psi|e^{i\varphi}$$

one gets:

$$\vec{\nabla}\Psi = (\vec{\nabla}|\Psi|)e^{i\varphi} + |\Psi|i(\vec{\nabla}\varphi)e^{i\varphi} = (\vec{\nabla}|\Psi|)e^{i\varphi} + \Psi i\vec{\nabla}\varphi$$

so that

$$\begin{aligned}\Psi^*\vec{\nabla}\Psi - \Psi\vec{\nabla}\Psi^* &= \Psi^*(\vec{\nabla}|\Psi|)e^{i\varphi} + |\Psi|^2 i\vec{\nabla}\varphi - [\Psi(\vec{\nabla}|\Psi|)e^{-i\varphi} + \Psi|\Psi|(-i)(\vec{\nabla}\varphi)e^{-i\varphi}] \\ &= \vec{\nabla}|\Psi|(\Psi^* e^{i\varphi} - \Psi e^{-i\varphi}) + |\Psi|^2 2i\vec{\nabla}\varphi \\ &= |\Psi|^2 2i\vec{\nabla}\varphi\end{aligned}$$

$$\Rightarrow \vec{j} = \frac{2e\hbar}{m}|\Psi|^2\left(\vec{\nabla}\varphi - \frac{2e}{\hbar c}\vec{A}\right)$$

But deep inside a superconductor, $\vec{j} = \vec{0}$. Indeed, $\vec{B} = \vec{0} \Rightarrow \vec{\text{rot}}\vec{B} = \vec{0} \Rightarrow \vec{j} = \vec{0}$. So

$$\begin{aligned}&\Rightarrow \quad \oint \vec{j}\cdot d\vec{l} = 0 \\ &\Rightarrow \quad \underbrace{\oint \vec{\nabla}\varphi \cdot d\vec{l}}_{2n\pi} = \frac{2e}{\hbar c}\underbrace{\oint \vec{A}\cdot d\vec{l}}_{\int \vec{\text{rot}}\vec{A}\cdot d\vec{S}}\end{aligned}$$

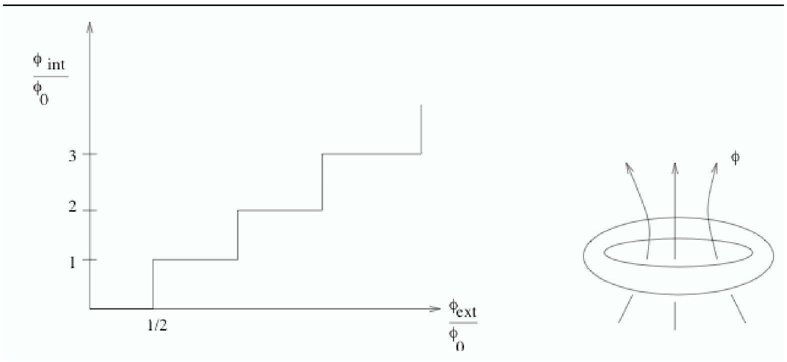
The first term must be equal to $2\pi n$ to ensure that the order parameter is well defined. Besides, $\int \vec{\text{rot}}\vec{A}\cdot d\vec{S} = \Phi$, where $\Phi$ is the flux of the magnetic field though the contour since $\vec{\text{rot}}\vec{A} = \vec{B}$. This leads to :

$$\Phi = n\frac{\pi\hbar c}{e} = n\phi_0$$

with $\phi_0 = \frac{\pi\hbar c}{e} = \frac{hc}{2e}$.

Consequences :

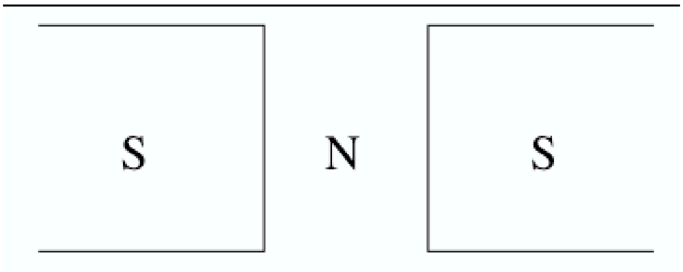
- Penetration of the field in type II superconductors : below $H_{c2}$, the sample is pierced by tubes of normal state. According to this argument, the flux of the magnetic field through such a tube must be a multiple of $\phi_0$. Now, we know that it is advantageous for the system to maximize its ratio surface/normal volume. So to get a flux $2\phi_0$ through the sample, it is better to create two vortices of flux $\phi_0$ than one of flux $2\phi_0$. So the magnetic flux penetrates as vortices of flux $\phi_0 = \frac{hc}{2e}$.
- If the charge of the effective charge carriers was $e$, as first assumed by Landau and Ginsburg, and not $2e$, the flux would be quantized in units of $\frac{hc}{e}$ and not $\frac{hc}{2e}$. The periodicity of various quantities as a function of the flux has been measured, and these results have clearly shown that the elementary flux is equal to $\frac{hc}{2e}$. In particular, if a superconducting ring is put into a field parallel to its axis, currents appear at the surface of the sample to adjust the field to a multiple of $\phi_0$.



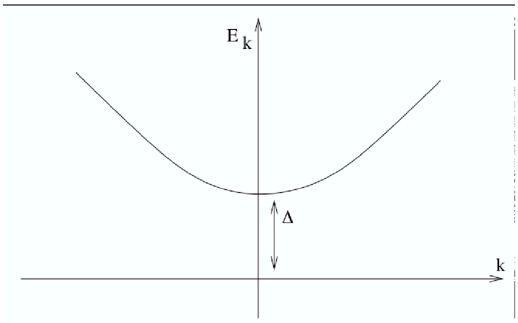
- Persistent currents : if one induces a current in a superconducting ring, this current must create a magnetic field whose flux is of the form $n\phi_0$. To decay, this current must overcome a potential barrier to reach the states $(n-1)\phi_0$, $(n-2)\phi_0$, . . . Now, this barrier is very high because the system must be in an intermediate state which is not globally superconducting. The probability is thus very weak, and the current can persist without decaying in any significant way for several years.

## 5.8 Josephson effect:

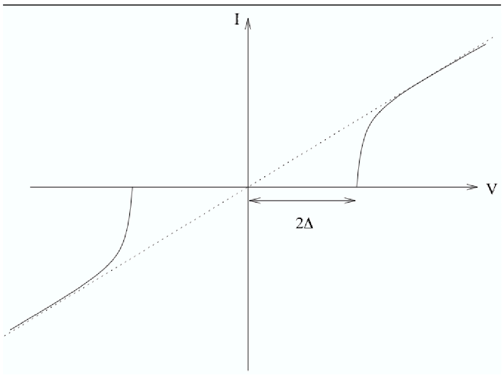
Let us consider two superconductors connected by a normal region



and let us try to let a current flow across the junction. Since the current goes through a normal region, the most natural thing to do is to assume that the current consists of single electrons. But to get an unpaired electron, one must break a pair and create two Bogoliubov particles. This costs at least $2\Delta$ in total ($\Delta$ for each particle).



To make an excitation in the system, one must thus apply a voltage $V = \frac{2\Delta}{e}$ to excite an electron. One thus expects the current-voltage characteristic to look like:



In other words, if one induces a current through the junction, even a very small one, one expects to induce a voltage drop at least equal to $2\Delta$.

In 1962, Josephson has produced a small revolution in the superconductivity community when he suggested that another type of current could flow through the junction, namely a pair current, and that this current could flow *without inducing a voltage drop*.

Josephson's original calculation has been done in the context of the BCS theory, looking at the effect of second order terms of the tunneling Hamiltonian through an insulating junction. But this effect can be understood more easily in the context of the Landau-Ginsburg theory.

The basic idea is very simple. Let us come back to the expression of the superconducting current as a function of the order parameter. The calculation of the previous section shows that, if $\vec{A}$ is equal to $\vec{0}$, and without any voltage, a current can flow if the phase of the order parameter is not uniform:

$$\vec{j} = \frac{2e\hbar}{m}|\Psi|^2\vec{\nabla}\phi$$

At thermal equilibrium, and in a singly-connected superconductor, this phase modulation is much too costly, and the current vanishes.

But if one considers a junction, the phase can take a different value on each side of the junction without having to pay this energy. If the order parameter does not vanish completely inside the junction, as will be the case if the junction is not too wide so that the electrons can tunnel through the junction, a superconducting current can a priori go through.

To find the expression of this current, it is of course impossible to ignore the variation of the amplitude of $\Psi$. The simplest calculation consists in going back to the Landau-Ginsburg equation for $\Psi$:

$$-\frac{\hbar^2}{2m}\Delta\Psi + \alpha\Psi + \beta|\Psi|^2\Psi = 0$$

Inside the junction, $\Psi$ changes very fast, and the term $\Delta\Psi$ dominates

$$\Rightarrow \Delta\Psi = 0$$

Let us suppose that the phase is equal to $\phi_1$ in the left superconductor, and to $\phi_2$ in the right superconductor, and that these two superconductors are the same so that the amplitude is the same in the bulk far from the junction, so that $\Psi(\vec{r}) = e^{i\phi_1}|\Psi_0|$ far on the left and $\Psi(\vec{r}) = e^{i\phi_2}|\Psi_0|$ far on the right.

Let us look for $\Psi$ as

$$\Psi(\vec{r}) = |\Psi_0|[f(\vec{r})e^{i\phi_1} + (1 - f(\vec{r}))e^{i\phi_2}]$$

where $f(\vec{r})$ is a real function equal to 1 in the left superconductor and 0 in the right one. This leads to

$$|\Psi_0|(e^{i\phi_1} - e^{i\phi_2})\Delta f = 0$$

Now, $\Delta\Psi = 0 \Rightarrow \Delta f = 0$ if $\phi_1 \neq \phi_2$.

In that case, the current is given by

$$\begin{aligned}
\vec{j} &= -\frac{ei\hbar}{m}\left(\Psi^*\vec{\nabla}\Psi - \Psi\vec{\nabla}\Psi^*\right) \\
\Psi^* &= |\Psi_0|\left(f(\vec{r})e^{-i\phi_1} + (1 - f(\vec{r}))e^{-i\phi_2}\right) \\
\vec{\nabla}\Psi &= |\Psi_0|\vec{\nabla}f(\vec{r})(e^{i\phi_1} - e^{i\phi_2})
\end{aligned}$$

$$\begin{aligned}
\Psi^*\vec{\nabla}\Psi &= |\Psi_0|^2\vec{\nabla}f(\vec{r})[f(\vec{r})e^{-i\phi_1} + (1 - f(\vec{r}))e^{-i\phi_2}](e^{i\phi_1} - e^{i\phi_2}) \\
&= |\Psi_0|^2\vec{\nabla}f(\vec{r})[f - fe^{-i\phi_1+i\phi_2} - (1 - f) + (1 - f)e^{-i\phi_2+i\phi_1}]
\end{aligned}$$

$$\Psi\vec{\nabla}\Psi^* = (\Psi^*\vec{\nabla}\Psi)^*$$

$$\begin{aligned}\Rightarrow \Psi^*\vec{\nabla}\Psi - \Psi\vec{\nabla}\Psi^* &= |\Psi_0|^2\vec{\nabla}f(\vec{r})[-f(e^{-i\phi_1+i\phi_2} - e^{i\phi_1-i\phi_2}) \\ &+ (1-f)(e^{-i\phi_2+i\phi_1} - e^{i\phi_2-i\phi_1})] \\ &= |\Psi_0|^2\vec{\nabla}f(\vec{r})(e^{-i\phi_2+i\phi_1} - e^{i\phi_2-i\phi_1}) \\ &= 2i\sin(\phi_1-\phi_2)|\Psi_0|^2\vec{\nabla}f(\vec{r})\end{aligned}$$

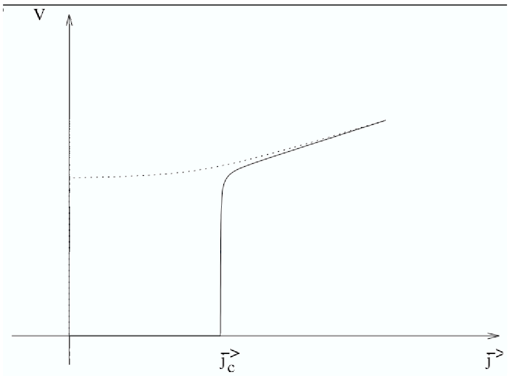
$$\Rightarrow \vec{j} = \frac{2e\hbar}{m}|\Psi_0|^2\vec{\nabla}f\sin(\phi_1-\phi_2)$$

The maximum current is reached when $\phi_1 - \phi_2 = \frac{\pi}{2}$, and it is of the order of

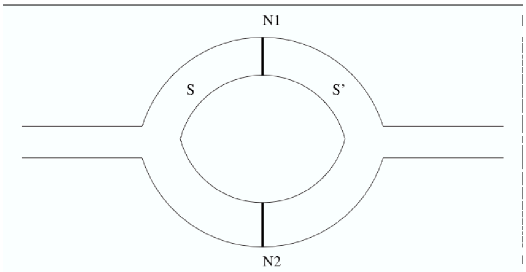
$$j_c = \frac{2e\hbar}{m}|\Psi_0|^2\frac{1}{L}$$

where $L$ is the width of the junction ($\vec{\nabla}f$ is of the order of $1/L$ since $f$ changes mostly inside the junction).

If one drives a current through an SNS' junction and measures the voltage drop, one expects to find:



The Josephson effect has considerable applications. In particular, the current also depends on the magnetic field. Then, if one builds two Josephon junctions according to the following geometry,



one can show that the maximal current is a periodic function of the flux $\phi$ through the ring, of periodicity $\phi_0$. But $\phi_0 = 10^{-10}G/cm^2$. One can thus detect field changes of the order of $10^{-9}$ Gauss. This is by far the most sensitive magnetometer.

# Chapter 6

# Appendix - Second quantization

## 6.1 Introduction

When dealing with systems of interacting particles, one has to work with many-particle wave functions. According to the principle of indistinguishability of identical particles, one must work with wave function which are symmetric (bosons) or antisymmetric (fermions) under the exchange of the coordinates of two particles. Indeed, let us consider a system of two particles. Since they cannot be distinguished, exchanging their coordinate cannot lead to different physical properties. The two wave functions must thus differ at most by a phase factor :

$$\psi(x_1, x_2) = e^{i\alpha}\psi(x_2, x_1)$$

But if one carries out this operation twice, one recovers the same wave function

$$\begin{aligned} &\Rightarrow e^{2i\alpha} = 1 \\ &\Rightarrow e^{i\alpha} = \pm 1 \quad \begin{aligned} &\rightarrow +1 \text{ bosons} \\ &\rightarrow -1 \text{ fermions} \end{aligned} \end{aligned}$$

So, if one denotes by $\{\psi_i(x)\}$ a basis of one-particle wave functions, one cannot simply use the set of functions

$$\Psi(x_1, \ldots, x_N) = \psi_{i_1}(x_1) \ldots \psi_{i_N}(x_N)$$

as a basis of the $N$-particle Hilbert space. One must build $N$ particle wave functions that are totally symmetric or totally antisymmetric when exchanging the coordinates of particles.

Remark :

One can construct other statistics than that of fermions and bosons. For instance, in 2D, exchanging the coordinates of two particles can be done in two topologically inequivalent ways (see Fig. 6.1). One comes back to the initial state only if one

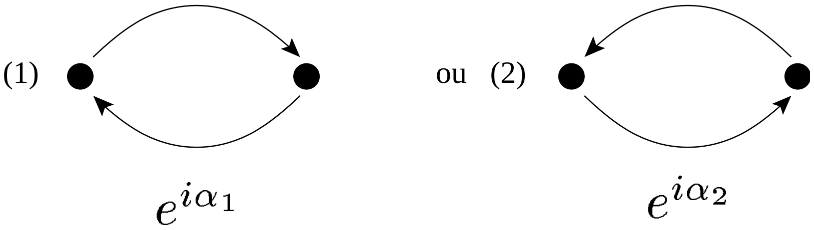


Figure 6.1:

successively applies (1) and (2).

$$\Rightarrow e^{i(\alpha_1+\alpha_2)} = 1 \Rightarrow \alpha_2 = -\alpha_1$$

But one can have $\alpha_1 \neq 0, \pi \Rightarrow$ *anyons.* This statistics approximately describes the excitations of a 2D electron gas in the presence of a strong magnetic field (Fractional Quantum Hall Effect).

## 6.2 Bosons

### 6.2.1 Fock space

One considers a basis of the one-particle Hilbert space $\{\psi_i(x)\}$, and one tries to construct $N$-particle wave functions that are totally symmetric under the exchange of the coordinates of any pair of particles. Let us start with the case of two particles. The wave function

$$\Psi(x_1, x_2) = \psi_i(x_1)\psi_i(x_2)$$

is symmetric and is thus acceptable. For two different states $i$ and $j$, one must build the wave function

$$\Psi(x_1, x_2) = \frac{1}{\sqrt{2}} \left[\psi_i(x_1)\psi_j(x_2) + \psi_j(x_1)\psi_i(x_2)\right]$$

to get a symmetric wave function. More generally, an $N$-particle wave function is defined by the number of times a given one-particle wave function $i$ appears. The wave function can then be written

$$\Psi(x_1, \ldots, x_N) = \frac{1}{\sqrt{\text{Norm}}} \sum_P \psi_{i_{P(1)}}(x_1) \ldots \psi_{i_{P(N)}}(x_N) \tag{6.1}$$

where the sum over $P$ runs over all the permutations of the integers $i_1 \ldots i_N$ that lead to a different term. In other words, if two permutations only differ by permutations inside families of equal indices, they only lead to a single term in the definition of the wave function. The number of such permutations (= number of terms in 6.1) is equal to :

$$\frac{N!}{N_1!N_2!\dots}$$

with the convention $N_i! = 1$ if $N_i = 0$, and of course $\sum_i N_i = N$. Besides, since the $\psi_i$ are mutually orthogonal, each term gives 1 with itself and 0 with the other terms when calculating $\langle\Psi|\Psi\rangle$. Thus,

$$\frac{1}{\sqrt{\mathrm{Norm}}} = \left(\frac{N_1!N_2!\dots}{N!}\right)^{\frac{1}{2}}$$

Example :

3 particles using $\psi_1$ twice and $\psi_2$ once.

$$\psi_1(x_1)\psi_1(x_2)\psi_2(x_3) + \psi_1(x_1)\psi_1(x_3)\psi_2(x_2) + \psi_1(x_2)\psi_1(x_3)\psi_2(x_1)$$

$\Rightarrow$ Norm = 3. General formula : $\frac{3!}{2!1!} = 3$

Since the wave function is entirely defined by the integers $N_i$, one can represent it by the ket :

$$|N_1, N_2, \dots\rangle \quad \text{(Fock space)}$$

### 6.2.2 Harmonic oscillator

The operators that allow one to conveniently travel inside a bosonic Fock space can be seen as a generalization of the creation and annihilation operators of the algebraic solution of the one-dimensional harmonic oscillator. Let us briefly recall how they are defined, and how they act on eigenstates.

$$H = \frac{p^2}{2m} + \frac{1}{2}m\omega^2x^2 \qquad [x,p] = i\hbar$$

1. Change of variables : $\left\{\begin{array}{l} X = \sqrt{\frac{m\omega}{\hbar}}x \\ P = \frac{1}{\sqrt{m\hbar\omega}}p \end{array}\right.$

$$\Rightarrow H = \frac{\hbar\omega}{2}(X^2 + P^2) \qquad [X,P] = i$$

2.

$$\left\{\begin{array}{lcl} a & = & \frac{1}{\sqrt{2}}(X + iP) \\ a^+ & = & \frac{1}{\sqrt{2}}(X - iP) \end{array}\right. \Rightarrow \left\{\begin{array}{lcl} X & = & \frac{1}{\sqrt{2}}(a^+ + a) \\ P & = & \frac{i}{\sqrt{2}}(a^+ - a) \end{array}\right.$$

3. $$[a,a^+]=1$$
$$a^+a=\frac{1}{2}(X^2+P^2-1)\Rightarrow H=\left(a^+a+\frac{1}{2}\right)\hbar\omega$$

4. Let us define $N=a^+a$, and let $|\varphi_\nu\rangle$ be a function such that $N|\varphi_\nu\rangle=\nu|\varphi_\nu\rangle$. Let us derive some properties of these operators :

   - If $\nu$ is an eigenvalue of $N$, then $\nu\geqslant 0$.
     Indeed, $\langle\varphi|a^+a|\varphi\rangle=\|a|\varphi\rangle\|^2\geqslant 0$
   - If $\nu=0$, then $a|\varphi_\nu\rangle=0$
     Indeed, $\nu=0\Rightarrow\|a|\varphi_\nu\rangle\|^2=0\Rightarrow a|\varphi_\nu\rangle=0$
   - If $\nu>0$, then $a|\varphi_\nu\rangle$ is an eigenvector of $N$ with eigenvalue $\nu-1$. Indeed,
     $$\begin{aligned}a^+aa|\varphi_\nu\rangle=(aa^+-1)a|\varphi_\nu\rangle &= a\,\underbrace{a^+a|\varphi_\nu\rangle}_{\nu|\varphi_\nu\rangle}-a|\varphi_\nu\rangle\\ &= (\nu-1)a|\varphi_\nu\rangle\end{aligned}$$
     Since the eigenvalues are positive or equal to zero, they must be integer. Otherwise, applying several times the operator $a$ would lead to states with negative eigenvalues.
   - $a^+|\varphi_\nu\rangle$ is never equal to zero.
     Indeed, $\|a^+|\varphi_\nu\rangle\|^2=\langle\varphi_\nu|aa^+|\varphi_\nu\rangle=(\nu+1)\langle\varphi_\nu|\varphi_\nu\rangle$.
     Now $\nu\geqslant 0\Rightarrow\nu+1>0$.
   - $a^+|\varphi_\nu\rangle$ is an eigenvector of $N$ with eigenvalue $\nu+1$.
     Indeed, $a^+aa^+|\varphi_\nu\rangle=a^+(a^+a+1)|\varphi_\nu\rangle=(\nu+1)a^+|\varphi_\nu\rangle$

   Finallly, $\nu$ can take all positive integer values, including 0.

5. The ground state is non degenerate. This is a property of the differential equation, whose lowest-energy solution is a gaussian. Thus, all eigenstates are non degenerate. The usual notation consists in replacing $|\varphi_n\rangle$ by $|n\rangle$.

6. Let $|n\rangle$ and $|n+1\rangle$ be two normalized states. One knows from the above that $a^+|n\rangle\propto|n+1\rangle$. To find the proportionality coefficient, let us write :
   $$a^+|n\rangle=c_{n+1}|n+1\rangle$$
   Let us take the norm of each member :
   $$|c_{n+1}|^2\underbrace{\langle n+1|n+1\rangle}_{1}=\langle n|aa^+|n\rangle=\langle n|1+a^+a|n\rangle=n+1$$

$$\Rightarrow c_{n+1} = \sqrt{n+1}$$

Similarly, writing $a|n\rangle = c_{n-1}|n-1\rangle$, one finds :

$$|c_{n-1}|^2 = \langle n|a^+a|n\rangle = n \quad \Rightarrow \quad c_{n-1} = \sqrt{n}$$

Thus,

$$\left\{ \begin{array}{rcr} a^+|n\rangle & = & \sqrt{n+1}\ |n+1\rangle \\ a|n\rangle & = & \sqrt{n}\ |n-1\rangle \end{array} \right.$$

### 6.2.3 Relationship between bosons and harmonic oscillators

**Creation and annihilation operators**

By analogy with the harmonic oscillator, it is convenient to introduce the elementary operators $a_i$ and $a_i^+$ defined by :

$$\begin{array}{rcl} a_i|N_1 \ldots N_i \ldots\rangle & = & \left\{ \begin{array}{l} \sqrt{N_i}\ |N_1 \ldots N_i - 1 \ldots\rangle \quad \text{if } N_i \geqslant 1 \\ 0 \quad \text{otherwise} \end{array} \right. \\ a_i^+|N_1 \ldots N_i \ldots\rangle & = & \sqrt{N_i + 1}\ |N_1 \ldots N_i + 1 \ldots\rangle \end{array}$$

Let us investigate the properties of these operators :

$$\begin{array}{rcl} a_i^+ a_i|N_1 \ldots N_i \ldots\rangle & = & \sqrt{N_i}\ a_i^+\ |N_1 \ldots N_i - 1 \ldots\rangle \\ & = & N_i\ |N_1 \ldots N_i \ldots\rangle \end{array}$$

$$\Rightarrow a_i^+ a_i = N_i \text{ diagonal}$$

$$\begin{array}{rcl} a_i a_i^+|N_1 \ldots N_i \ldots\rangle & = & \sqrt{N_i + 1}\ a_i\ |N_1 \ldots N_i + 1 \ldots\rangle \\ & = & (N_i + 1)\ |N_1 \ldots N_i \ldots\rangle \end{array}$$

$$\Rightarrow a_i a_i^+ = N_i + 1$$

This implies that :

$$\boxed{a_i a_i^+ - a_i^+ a_i = 1}$$

$$\left. \begin{array}{rcl} a_i a_j & = & a_j a_i \\ a_i^+ a_j & = & a_j a_i^+ \\ a_i^+ a_j^+ & = & a_j^+ a_i^+ \end{array} \right\} \text{if } i \neq j$$

$$\text{or finally} \left\{ \begin{array}{l} [a_i, a_j^+] = \delta_{ij} \\ [a_i, a_j] = [a_i^+, a_j^+] = 0 \end{array} \right.$$

### One-body operators

Let us consider a symmetrized one-body operator :

$$\hat{F}^{(1)} = \sum_a \hat{f}_a^{(1)}$$

where $\hat{f}_a^{(1)}$ only acts on the variable $x_a$. The matrix elements of the operator $\hat{f}^{(1)}$ are given by :

$$f_{ij}^{(1)} = \int \psi_i^*(x)\hat{f}^{(1)}\psi_j(x)dx$$

Let us calculate the expectation value of $\hat{F}^{(1)}$ between two states. Since $\hat{f}_i^{(1)}$ only acts on one variable at a time, two states are connected by $\hat{F}^{(1)}$ only if they at most differ by the function acting on one variable. We thus have two types of non vanishing matrix elements :

$$\left\{\begin{array}{ccc} \langle N_1 \dots N_i \dots |\hat{F}^{(1)}|N_1 \dots N_i \dots\rangle & = & \sum_i f_{ii}^{(1)} N_i \\ \langle \dots N_i \dots N_k - 1 \dots |\hat{F}^{(1)}| \dots N_i - 1 \dots N_k \dots\rangle & = & f_{ik}^{(1)}\sqrt{N_i N_k} \end{array}\right.$$

<u>Proof</u>

$$\langle N_1 N_2 \dots |\hat{F}^{(1)}|N_1 N_2 \dots\rangle = \sum_i f_{ii}^{(1)} N_i$$

$|N_1 N_2 \dots\rangle$ contains $\frac{N!}{N_1!N_2!\dots}$ terms. For each term, $\langle \hat{f}_a^{(1)}\rangle = f_{i_a i_a}^{(1)}$, where $i_a$ if the index of the function $\phi_i$ of the variable $x_a$. Since the variables $x_a$ enter all the functions $\phi_i$ of $|N_1 \dots\rangle$, we get :

$$\langle \sum_a \hat{f}_a^{(1)}\rangle = \sum_i f_{ii}^{(1)} N_i$$

where the left hand site is the expectation value for one term. Since this expectation value does not depend on the term in the sum, we get the result we wanted to prove:

$$\langle \dots N_i \dots N_k - 1 \dots |\hat{F}^{(1)}| \dots N_i - 1 \dots N_k \dots\rangle = f_{ik}^{(1)}\sqrt{N_i N_k}$$

The ket $| \dots N_i - 1 \dots N_k \dots\rangle$ contains $\frac{N!}{\dots N_k!(N_i-1)!}$ terms. For each term, there are $N_k$ terms in the sum $\sum_a \hat{f}_a^{(1)}$ that contribute to a given matrix element $f_{ik}^{(1)}$ . Besides, the normalization coefficients of the bras and the kets are given by :

$$\begin{array}{ccc} | \dots N_i - 1 \dots N_k \dots\rangle & \rightarrow & \sqrt{\frac{N_k!(N_i-1)!}{N!}} \\ \langle \dots N_i \dots N_k - 1 \dots | & \rightarrow & \sqrt{\frac{N_i!(N_k-1)!}{N!}} \end{array}$$

So the coefficient of $f_{ik}^{(1)}$ is given by :

$$\underbrace{\frac{N!}{\dots N_k!(N_i-1)!\dots}}_{\text{number of terms in the ket}} \quad \underbrace{N_k}_{\substack{\text{number of operators in} \\ \hat{F}^{(1)}\text{that give } f_{ik}^{(1)}}} \quad \underbrace{\sqrt{\frac{\dots N_k!(N_i-1)!\dots}{N!}}\sqrt{\frac{\dots (N_k-1)!N_i!\dots}{N!}}}_{\text{normalization coefficients}}$$

which finally leads to $\sqrt{N_i N_k}$.

The operator $\hat{F}^{(1)}$ can thus be written :

$$\boxed{\hat{F}^{(1)} = \sum_{ik} f_{ik}^{(1)} a_i^+ a_k}$$

Proof

- Diagonal elements

$$\langle N_1 N_2 \ldots | a_i^+ a_k | N_1 N_2 \ldots \rangle = 0 \quad \text{if } i \neq k$$

$$\begin{aligned} \Rightarrow \langle N_1 N_2 \ldots | \hat{F}^{(1)} | N_1 N_2 \ldots \rangle &= \sum_i f_{ii}^{(1)} \langle N_1 \ldots | a_i^+ a_i | N_1 \ldots \rangle \\ &= \sum_i f_{ii}^{(1)} N_i \quad \text{Q.E.D.} \end{aligned}$$

- Off-diagonal elements

$$\begin{aligned} \langle \ldots N_i N_k - 1 \ldots | a_j^+ a_l | \ldots N_i - 1 N_k \ldots \rangle &= 0 \quad \text{except if } l = k \text{ and } i = j \\ \langle \ldots N_i N_k - 1 \ldots | a_i^+ a_k | \ldots N_i - 1 N_k \ldots \rangle &= \sqrt{N_i N_k} \quad \text{Q.E.D.} \end{aligned}$$

### Two-body operators

Similarly, a symmetrized operator acting on the coordinates of two particles can be written :

$$\hat{F}^{(2)} = \sum_{a<b} \hat{f}_{ab}^{(2)}$$

and, in second quantization :

$$\hat{F}^{(2)} = \frac{1}{2} \sum_{iklm} \left( \hat{f}^{(2)} \right)_{lm}^{ik} a_i^+ a_k^+ a_l a_m$$

$$\text{with} \quad \left( \hat{f}^{(2)} \right)_{lm}^{ik} = \int \int \psi_i^*(x_1) \psi_k^*(x_2) \hat{f}^{(2)} \psi_l(x_2) \psi_m(x_1) dx_1 dx_2$$

Finally, the Fock space itself can be described simply in terms of creation and annihilation operators. Indeed, using repeatedly the identity :

$$|N_1 \ldots N_i + 1 \ldots \rangle = \frac{a_i^+}{\sqrt{N_i + 1}} |N_1 \ldots N_i \rangle$$

one can write :

$$|N_1 \ldots N_i \ldots \rangle = \prod_i \frac{\left( a_i^+ \right)^{N_i}}{\sqrt{N_i!}} |0\rangle$$

where $|0\rangle$ is the vacuum, the state without any particle.

This leads to a very compact formulation of the many-body problem in which the operators have the same form independent of the number of particles, and in which everything can be formulated in terms of operators satisfying very simple commutation rules.

## 6.3 Fermions

In the case of fermions, the wave function must be antisymmetric. Such a wave function can only be built out of different one-paerticle wave functions as Slater determinants :

$$\Psi(x_1,\ldots,x_N)=\frac{1}{\sqrt{N!}}\begin{vmatrix}\psi_{p_1}(x_1) & \psi_{p_1}(x_2) & \ldots & \psi_{p_1}(x_N)\\ \psi_{p_2}(x_1) & \psi_{p_2}(x_2) & \ldots & \psi_{p_2}(x_N)\\ \vdots & \vdots & & \vdots\\ \psi_{p_N}(x_1) & \psi_{p_N}(x_2) & \ldots & \psi_{p_N}(x_N)\end{vmatrix}$$

In the Fock space, each state is defined by an occupation number equal to 0 or 1 :

$$|N_1\ldots N_i\ldots\rangle$$

There is a subtlety with respect to bosons. Indeed, the order of the lines controls the final sign of the determinant. One must thus order the wave functions once for all and keep this ordering. The determinant can then be calculated according to the convention :

$$p_1<p_2<\ldots<p_N$$

If one calculates the matrix elements of a one-body operator, one can follow the calculation made for bosons with $N_i=0$ or 1, up to the fact that when calculating off-diagonal term, a sign appears because of the determinant :

$$\langle 1_i 0_k|F^{(1)}|0_i 1_k\rangle = f^{(1)}_{ik}(-1)^{\sum(i+1,k-1)}$$

$$\text{with}\qquad \sum(k,l)=\sum_{n=k}^{l}N_n$$

So, if one wants to represent this operator with creation and annihilation operators, by analogy to bosons, they must be defined by :

$$\left\{\begin{array}{lllll} c_i^+ & |N_1,\ldots,N_i,\ldots\rangle & =(-1)^{\sum(1,i-1)} & (1-N_i) & |N_1,\ldots,N_i+1,\ldots\rangle\\ c_i & |N_1,\ldots,N_i,\ldots\rangle & =(-1)^{\sum(1,i-1)} & N_i & |N_1,\ldots,N_i-1,\ldots\rangle \end{array}\right.$$

This leads to the following anticommutation relations :

$$\left\{\begin{array}{lll} \{c_i^+,c_k\} & \equiv & c_i^+c_k+c_kc_i^+=\delta_{ik}\\ \{c_i^+,c_k^+\} & = & \{c_i,c_k\}=0 \end{array}\right.$$

Proof

$\boxed{i = k}$

$$\begin{aligned}
c_i^+ c_i |0_i\rangle &= 0 \\
c_i c_i^+ |0_i\rangle &= c_i (-1)^{\sum(1,i-1)} |1_i\rangle = |0_i\rangle \\
c_i^+ c_i |1_i\rangle &= c_i^+ (-1)^{\sum(1,i-1)} |0_i\rangle = |1_i\rangle \\
c_i c_i^+ |1_i\rangle &= 0 \\
&\Rightarrow \boxed{c_i^+ c_i + c_i c_i^+ = 1}
\end{aligned}$$

$\boxed{i < k}$

$$\begin{aligned}
c_i^+ c_k^+ |N_1 \ldots\rangle &= (-1)^{N_1+\ldots+N_{k-1}} c_i^+ (1-N_k) |N_1 \ldots N_k + 1 \ldots\rangle \\
&= (-1)^{N_1+\ldots+N_{k-1}} (-1)^{N_1+\ldots+N_{i-1}} (1-N_i)(1-N_k) |\ldots N_i + 1 \ldots N_k + 1 \ldots\rangle \\
c_k^+ c_i^+ |N_1 \ldots\rangle &= (-1)^{N_1+\ldots+N_{i-1}} c_k^+ (1-N_i) |N_1 \ldots N_i + 1 \ldots N_k\rangle \\
&= (-1)^{N_1+\ldots+N_{i-1}} (-1)^{N_1+\ldots+N_i+1+N_{i+1}+\ldots+N_{k-1}} \\
&\quad (1-N_i)(1-N_k) |\ldots N_i + 1 \ldots N_k + 1 \ldots\rangle
\end{aligned}$$

$$\Rightarrow c_i^+ c_k^+ = -c_k^+ c_i^+ \Rightarrow c_i^+ c_k^+ + c_k^+ c_i^+ = 0$$

$$\text{or} \quad \{c_i^+, c_k^+\} = 0$$

Similarly for $\{c_i, c_k^+\}$ or $\{c_i, c_k\}$

Some remarks on anticommutation relations :

- $c_i^+ c_i^+ = c_i c_i = 0$. This is Pauli's principle : two fermions cannot occupy the same quantum state.
- Besides, one indeed has :

$$\begin{aligned}
c_i^+ c_k |0_i 1_k\rangle &= (-1)^{N_1+\ldots+N_{i-1}+N_{i+1}+\ldots+N_{k-1}} c_i^+ |0_i 0_k\rangle \quad (N_i = 0) \\
&= (-1)^{N_1+\ldots+N_{i-1}+N_{i+1}+\ldots+N_{k-1}} (-1)^{N_1+\ldots+N_{i-1}} |1_i 0_k\rangle \\
&= (-1)^{\sum(i+1,k-1)} |1_i 0_k\rangle
\end{aligned}$$

- Finally, $N = \sum_i c_i^+ c_i$ satisfies :

$$N|N_1 \ldots\rangle = \sum_i N_i |N_1 \ldots\rangle \Rightarrow N = \text{number of particles}$$

With these definitions, the expressions that we have derived for bosons still apply:

$$\Rightarrow \hat{F}^{(1)} = \sum_{ik} f_{ik}^{(1)} c_i^+ c_k$$

with

$$f_{ij}^{(1)} = \int \psi_i^*(x) \hat{f}^{(1)} \psi_j(x) dx$$

and

$$\hat{F}^{(2)} = \frac{1}{2}\sum_{iklm}\left(\hat{f}^{(2)}\right)^{ik}_{lm} a_i^+ a_k^+ a_l a_m$$

$$\text{with} \qquad \left(\hat{f}^{(2)}\right)^{ik}_{lm} = \int\int \psi_i^*(x_1)\psi_k^*(x_2)\hat{f}^{(2)}\psi_l(x_2)\psi_m(x_1)dx_1dx_2$$

For the two-body term, the order of the indices now matters since the operators anticommute.

## 6.4 Field operator

It is very useful to use linear combinations of creation and annihilation operators called *field operators* and defined by :

$$\begin{cases} \Psi(x) &= \sum_i \varphi_i(x)a_i \\ \Psi^+(x) &= \sum_i \varphi_i^*(x)a_i^+ \end{cases}$$

Since the operator $a_i^+$ creates one particle in the state of wave function $\varphi_i(x)$, the operator $\Psi^+(x_0)$ creates a particle in the state of wave function $\delta(x-x_0)$. In other words, if $x_0$ denotes the position in space, the field operator creates a particle at point $x_0$. This definition can be easily extended to other degrees of freedom, for instance the spin.

Proof

$$\begin{aligned} \langle x|a_i^+|0\rangle &= \varphi_i(x) \\ \Rightarrow \langle x|\Psi^+(x_0)|0\rangle &= \sum_i \varphi_i^*(x_0)\langle x|a_i^+|0\rangle \\ \sum_i \varphi_i^*(x_0)\varphi_i(x) &= \sum_i \varphi_i^*(x_0)\varphi_i(x) \\ &= \delta(x-x_0) \end{aligned}$$

Indeed, according to the definition of the "function" $\delta(x-x_0)$, one must have for any trial wave function $\varphi_j^*(x)$

$$\int \delta(x-x_0)\varphi_j^*(x)dx = \varphi_j^*(x_0)$$

But,

$$\int \sum_i \varphi_i^*(x_0)\varphi_i(x)\varphi_j^*(x)dx = \sum_i \varphi_i^*(x_0)\delta_{ij} = \varphi_j^*(x_0)$$

The formulae that allow one to write the one- and two-body operators take a very simple form :

$$\begin{aligned} \hat{F}^{(1)} &= \int \Psi^+(x)\hat{f}^{(1)}\Psi(x)dx \\ \hat{F}^{(2)} &= \frac{1}{2}\int\int \Psi^+(x)\Psi^+(x')\hat{f}^{(2)}\Psi(x')\Psi(x)dxdx' \end{aligned}$$

where $\hat{f}^{(1)}$ and $\hat{f}^{(2)}$ act on the functions of $x$ and $x'$. Indeed,

$$\hat{F}^{(1)} = \sum_{ij} a_i^+ a_j \int \varphi_i^*(x) \hat{f}^{(1)} \varphi_j(x) dx$$

or

$$\hat{F}^{(1)} = \sum_{ij} f_{ij}^{(1)} a_i^+ a_j$$

The same applies to $\hat{F}^{(2)}$.

The commutation relations of the field operators can be deduced from those of the operators $a_i^+$ and $a_i$. They are of course different for fermions and bosons :

- Bosons :

  $$\left[\Psi(x), \Psi^+(x')\right] = \delta(x - x')$$

  Indeed :

  $$\begin{aligned} \left[\Psi(x), \Psi^+(x')\right] &= \sum_{ij} \varphi_i(x) \varphi_j^*(x') \underbrace{[a_i, a_j^+]}_{\delta_{ij}} \\ &= \sum_i \varphi_i(x) \varphi_i^*(x') \\ &= \delta(x - x') \end{aligned}$$

  $$[\Psi(x), \Psi(x')] = \left[\Psi^+(x), \Psi^+(x')\right] = 0$$

- Fermions :

  $$\{\Psi(x), \Psi^+(x')\} = \delta(x - x')$$

  $$\{\Psi(x), \Psi(x')\} = \{\Psi^+(x), \Psi^+(x')\} = 0$$

## 6.5 Application : interacting electrons

Let us consider the problem of an electron gas in a potential $U(\mathbf{r})$ interacting through the Coulomb interaction $V(\mathbf{r}_1, \mathbf{r}_2) = \frac{e^2}{|\mathbf{r}_1 - \mathbf{r}_2|}$. Let us try to determine the form of the Hamiltonian in second quantization.

### 6.5.1 Reference eigenfunctions

When we built our totally symmetric or totally antisymmetric states, we built wave functions for the variable $x$ without specifying what it stands for. For electrons, the variable $x = (\mathbf{r}, \sigma)$ is a composite variable that stands for the position $\mathbf{r}$ and for a spin variable $\sigma$ that can take the values $+\frac{1}{2}$ and $-\frac{1}{2}$.

For the spatial part, the most convenient choice is to use plane waves :

$$\varphi_{\mathbf{k}}(\mathbf{r}) = \frac{1}{\sqrt{\Omega}} e^{i\mathbf{k}.\mathbf{r}}$$

where one considers a box of volume $\Omega$. The momentum $\mathbf{k}$ takes discrete values, and one has :

$$\int_\Omega \varphi^*_{\mathbf{k}_1}(\mathbf{r})\varphi_{\mathbf{k}_2}(\mathbf{r})d\mathbf{r} = \delta_{\mathbf{k}_1,\mathbf{k}_2}$$

For the spin, one must also use wave functions. Since this is not the standard way to proceed, let us write down explicitly the form of the wave functions for a spin $\frac{1}{2}$.

### Standard description

Two states $|\uparrow\rangle$ and $|\downarrow\rangle$

$$\begin{array}{lll} S^z|\uparrow\rangle = \frac{\hbar}{2}|\uparrow\rangle & S^+|\uparrow\rangle = 0 & S^-|\uparrow\rangle = |\downarrow\rangle \\ S^z|\downarrow\rangle = -\frac{\hbar}{2}|\downarrow\rangle & S^+|\downarrow\rangle = |\uparrow\rangle & S^-|\downarrow\rangle = 0 \end{array}$$

$$S^x = \frac{S^+ + S^-}{2} \quad S^y = \frac{S^+ - S^-}{2i}$$

### Description with vectors and matrices

$$|\uparrow\rangle \rightarrow \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad |\downarrow\rangle \rightarrow \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

$$S^z = \frac{\hbar}{2}\underbrace{\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}}_{\sigma_z} \quad S^x = \frac{\hbar}{2}\underbrace{\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}}_{\sigma_x} \quad S^y = \frac{\hbar}{2}\underbrace{\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}}_{\sigma_y}$$

$\sigma_x, \sigma_y, \sigma_z$ : Pauli matrices.

### Description in terms of wave functions

One can also represent the kets by wave functions. For a spin $\frac{1}{2}$, the variable $\sigma$ can take two values, say $\frac{1}{2}$ et $-\frac{1}{2}$. One can then define two wave functions by :

$$\left\{ \begin{array}{ll} \eta_{+\frac{1}{2}}(\frac{1}{2}) = 1 & \eta_{+\frac{1}{2}}(-\frac{1}{2}) = 0 \\ \eta_{-\frac{1}{2}}(\frac{1}{2}) = 0 & \eta_{-\frac{1}{2}}(-\frac{1}{2}) = 1 \end{array} \right.$$

or, more compactly,

$$\eta_{\sigma_1}(\sigma) = \delta_{\sigma\sigma_1}$$

These functions build an orthonormal basis :

$$\sum_\sigma \eta^*_{\sigma_1}(\sigma)\eta_{\sigma_2}(\sigma) = \delta_{\sigma_1\sigma_2} \quad \text{(to be checked)}$$

In this basis, the operators $S^x$, $S^y$ et $S^z$ act according to :

$$\begin{cases} S^z\eta_\sigma = \hbar\sigma\eta_\sigma \\ S^x\eta_\sigma = \frac{\hbar}{2}\eta_{-\sigma} \\ S^y\eta_\sigma = i\hbar\sigma\eta_{-\sigma} \end{cases}$$

The eigenfuntions are thus labeled by two indices :

$$\varphi_{\mathbf{k}\sigma}(\mathbf{r},\sigma') = \frac{1}{\sqrt{\Omega}} e^{i\mathbf{k}.\mathbf{r}} \eta_\sigma(\sigma')$$

The creation and annihilation operators are thus given by :

$$c^+_{\mathbf{k}\sigma} \quad \text{and} \quad c_{\mathbf{k}\sigma}$$

Similarly, the field operators are defined by :

$$\begin{aligned} \Psi(\mathbf{r},\sigma) &= \sum_{\mathbf{k},\sigma'} \varphi_{\mathbf{k}\sigma'}(\mathbf{r},\sigma) c_{\mathbf{k}\sigma'} \\ &= \sum_{\mathbf{k},\sigma'} \frac{1}{\sqrt{\Omega}} e^{i\mathbf{k}.\mathbf{r}} \delta_{\sigma\sigma'} c_{\mathbf{k}\sigma'} \\ &= \sum_{\mathbf{k}} \frac{1}{\sqrt{\Omega}} e^{i\mathbf{k}.\mathbf{r}} c_{\mathbf{k}\sigma} \\ &= \sum_{\mathbf{k}} \varphi_{\mathbf{k}}(\mathbf{r}) c_{\mathbf{k}\sigma} \end{aligned}$$

and

$$\Psi^+(\mathbf{r},\sigma) = \sum_{\mathbf{k}} \varphi^*_{\mathbf{k}}(\mathbf{r}) c^+_{\mathbf{k}\sigma}$$

### 6.5.2 Operators

There are two one-body operators :

#### Kinetic energy

$$H_{\text{kin}} = \sum_i -\frac{\hbar^2\nabla_i^2}{2m} \equiv \sum_i -\frac{\hbar^2\Delta_i}{2m}$$

$$\Delta_i\varphi = \frac{\partial^2\varphi}{\partial x_i^2} + \frac{\partial^2\varphi}{\partial y_i^2} + \frac{\partial^2\varphi}{\partial z_i^2}$$

**Potential energy**

$$H_{\text{pot}} = \sum_i U(\mathbf{r}_i)$$

$U(\mathbf{r}_i)\varphi$ is just the product by the function $U(\mathbf{r})$ evaluated at the point $\mathbf{r}_i$.

The Coulomb repulsion leads to a two-body term :

$$H_{\text{Coulomb}} = \frac{1}{2}\sum_{ij} V(\mathbf{r}_i - \mathbf{r}_j)$$

The factor $\frac{1}{2}$ has been introduced to count each pair only once.

### 6.5.3 Hamiltonian in second quantization

$$H_{\text{kin}} = \sum_{\mathbf{k}_1\mathbf{k}_2\sigma_1\sigma_2} c^+_{\mathbf{k}_1\sigma_1} c_{\mathbf{k}_2\sigma_2} \int d\mathbf{r} \sum_\sigma \varphi^*_{\mathbf{k}_1\sigma 1}(\mathbf{r},\sigma)\left(-\frac{\hbar^2\Delta}{2m}\right)\varphi_{\mathbf{k}_2\sigma_2}(\mathbf{r},\sigma)$$

$$\text{or} \quad \begin{array}{r} \varphi^*_{\mathbf{k}_1\sigma_1}(\mathbf{r},\sigma) = \frac{1}{\sqrt{\Omega}} e^{-i\mathbf{k}_1.\mathbf{r}} \delta_{\sigma\sigma_1} \\ \varphi_{\mathbf{k}_2\sigma_2}(\mathbf{r},\sigma) = \frac{1}{\sqrt{\Omega}} e^{i\mathbf{k}_2.\mathbf{r}} \delta_{\sigma\sigma_2} \end{array}$$

$$\begin{aligned} \Rightarrow H_{\text{kin}} &= \sum_{\mathbf{k}_1\mathbf{k}_2\sigma} c^+_{\mathbf{k}_1\sigma} c_{\mathbf{k}_2\sigma} \int \frac{d\mathbf{r}}{\Omega} e^{-i\mathbf{k}_1.\mathbf{r}} \frac{\hbar^2\mathbf{k}_2^2}{2m} e^{i\mathbf{k}_2.\mathbf{r}} \\ &= \sum_{\mathbf{k}_1\mathbf{k}_2\sigma} c^+_{\mathbf{k}_1\sigma} c_{\mathbf{k}_2\sigma} \frac{\hbar^2\mathbf{k}_2^2}{2m} \delta_{\mathbf{k}_1\mathbf{k}_2} \end{aligned}$$

$$\boxed{H_{\text{kin}} = \sum_{\mathbf{k}\sigma} \frac{\hbar^2\mathbf{k}^2}{2m} c^+_{\mathbf{k}\sigma} c_{\mathbf{k}\sigma}}$$

$$\begin{aligned} H_{\text{pot}} &= \sum_{\mathbf{k}_1\mathbf{k}_2\sigma_1\sigma_2} c^+_{\mathbf{k}_1\sigma_1} c_{\mathbf{k}_2\sigma_2} \int \frac{d\mathbf{r}}{\Omega} \sum_\sigma e^{-i\mathbf{k}_1.\mathbf{r}} U(\mathbf{r}) e^{i\mathbf{k}_2.\mathbf{r}} \delta_{\sigma\sigma_1}\delta_{\sigma\sigma_2} \\ &= \sum_{\mathbf{k}_1\mathbf{k}_2\sigma} c^+_{\mathbf{k}_1\sigma} c_{\mathbf{k}_2\sigma} \int \frac{d\mathbf{r}}{\Omega} e^{-i\mathbf{k}_1.\mathbf{r}} U(\mathbf{r}) e^{i\mathbf{k}_2.\mathbf{r}} \end{aligned}$$

Let us define the Fourier transform of the potential by :

$$\tilde{U}(\mathbf{k}) = \int \frac{d\mathbf{r}}{\Omega} e^{-i\mathbf{k}.\mathbf{r}} U(\mathbf{r})$$

This leads to :

$$H_{\text{pot}} = \sum_{\mathbf{k}_1\mathbf{k}_2\sigma} c^+_{\mathbf{k}_1\sigma} c_{\mathbf{k}_2\sigma} \tilde{U}(\mathbf{k}_1 - \mathbf{k}_2)$$

or :

$$\boxed{H_{\text{pot}} = \sum_{\mathbf{k}\mathbf{q}\sigma} \tilde{U}(\mathbf{q}) c^+_{\mathbf{k}+\mathbf{q}\sigma} c_{\mathbf{k}\sigma}}$$

The operator :

$$\rho_{-\mathbf{q}} = \sum_{\mathbf{k}\sigma} c^{+}_{\mathbf{k}+\mathbf{q}\sigma} c_{\mathbf{k}\sigma}$$

is the Fourier transform of the density operator. Indeed,

$$\left\{ \begin{array}{lcl} \rho_{\mathbf{q}} & = & \int d\mathbf{r} \rho(\mathbf{r}) e^{-i\mathbf{q}.\mathbf{r}} \\ \rho(\mathbf{r}) & = & \sum_{\sigma} \Psi^{+}(\mathbf{r},\sigma)\Psi(\mathbf{r},\sigma) \end{array} \right.$$

$$\begin{aligned} \Rightarrow \rho_{\mathbf{q}} &= \int \frac{d\mathbf{r}}{\Omega} e^{-i\mathbf{q}.\mathbf{r}} \sum_{\mathbf{k}_1\mathbf{k}_2\sigma} e^{-i\mathbf{k}_1.\mathbf{r}} e^{i\mathbf{k}_2.\mathbf{r}} c^{+}_{\mathbf{k}_1\sigma} c_{\mathbf{k}_2\sigma} \\ &= \sum_{\mathbf{k}\sigma} c^{+}_{k\sigma} c_{\mathbf{k}+\mathbf{q}\sigma} \end{aligned}$$

One can thus write :

$$H_{\text{pot}} = \sum_{\mathbf{q}} \tilde{U}(\mathbf{q}) \rho_{-\mathbf{q}}$$

$$\begin{aligned} H_{\text{Coul}} &= \frac{1}{2} \sum_{\substack{\mathbf{k}_1\mathbf{k}_2\mathbf{k}_3\mathbf{k}_4 \\ \sigma_1\sigma_2\sigma_3\sigma_4}} c^{+}_{\mathbf{k}_1\sigma_1} c^{+}_{\mathbf{k}_2\sigma_2} c_{\mathbf{k}_3\sigma_3} c_{\mathbf{k}_4\sigma_4} \\ &\quad \times \int d\mathbf{r} d\mathbf{r}' \sum_{\sigma\sigma'} \varphi^{*}_{\mathbf{k}_1\sigma_1}(\mathbf{r},\sigma) \varphi^{*}_{\mathbf{k}_2\sigma_2}(\mathbf{r}',\sigma') V(\mathbf{r}-\mathbf{r}') \varphi_{\mathbf{k}_3\sigma_3}(\mathbf{r}',\sigma') \varphi_{\mathbf{k}_4\sigma_4}(\mathbf{r},\sigma) \end{aligned}$$

The integral is given by :

$$\int \frac{d\mathbf{r} d\mathbf{r}'}{\Omega^2} V(\mathbf{r}-\mathbf{r}') e^{-i\mathbf{k}_1.\mathbf{r}} e^{-i\mathbf{k}_2.\mathbf{r}'} e^{i\mathbf{k}_3.\mathbf{r}'} e^{i\mathbf{k}_4.\mathbf{r}} \underbrace{\sum_{\sigma\sigma'} \delta_{\sigma_1\sigma} \delta_{\sigma_2\sigma'} \delta_{\sigma_3\sigma'} \delta_{\sigma_4\sigma}}_{\delta_{\sigma_1\sigma_4}\delta_{\sigma_2\sigma_3}}$$

or, defining $\mathbf{R} = \mathbf{r} - \mathbf{r}'$ :

$$\begin{aligned} & \int \frac{d\mathbf{R} d\mathbf{r}'}{\Omega^2} V(\mathbf{R}) e^{-i\mathbf{k}_1.(\mathbf{R}+\mathbf{r}')} e^{-i\mathbf{k}_2.\mathbf{r}'} e^{i\mathbf{k}_3.\mathbf{r}'} e^{i\mathbf{k}_4.(\mathbf{R}+\mathbf{r}')} \\ = & \underbrace{\int \frac{d\mathbf{R}}{\Omega} V(\mathbf{R}) e^{-i(\mathbf{k}_1-\mathbf{k}_4).\mathbf{R}}}_{\tilde{V}(\mathbf{k}_1-\mathbf{k}_4)} \underbrace{\int \frac{d\mathbf{r}'}{\Omega} e^{-i(\mathbf{k}_1+\mathbf{k}_2-\mathbf{k}_3-\mathbf{k}_4).\mathbf{r}'}}_{\delta_{\mathbf{k}_1+\mathbf{k}_2,\mathbf{k}_3+\mathbf{k}_4}} \end{aligned}$$

$$\boxed{H_{\text{Coul}} = \frac{1}{2} \sum_{\mathbf{k}\mathbf{k}'\mathbf{q}\sigma\sigma'} \tilde{V}(\mathbf{q}) c^{+}_{\mathbf{k}+\mathbf{q}\sigma} c^{+}_{\mathbf{k}'-\mathbf{q}\sigma'} c_{\mathbf{k}'\sigma'} c_{\mathbf{k}\sigma}}$$

With the help of the density operators, this can also be written :

$$H_{\text{Coul}} = \frac{1}{2} \sum_{\mathbf{q}} \tilde{V}(\mathbf{q}) \left( \rho_{\mathbf{q}} \rho_{-\mathbf{q}} - N \right)$$

Nota Bene : This expression is valid with the definition $\tilde{V}(\mathbf{q}) = \int \frac{d\mathbf{r}}{\Omega} e^{-i\mathbf{q}.\mathbf{r}} V(\mathbf{r})$.

If one defines :

$$V(\mathbf{q}) = \int d\mathbf{r} e^{-i\mathbf{q}.\mathbf{r}} V(\mathbf{r})$$

which, for Coulomb potential, leads to $V(\mathbf{q}) = \frac{4\pi e^2}{\mathbf{q}^2}$, then the Hamiltonian in second quantization takes the form :

$$H_1 = \frac{1}{2\Omega} \sum_{\mathbf{k}\mathbf{k}'\mathbf{q}\sigma\sigma'} V(\mathbf{q}) c^+_{\mathbf{k}+\mathbf{q}\sigma} c^+_{\mathbf{k}'-\mathbf{q}\sigma'} c_{\mathbf{k}'\sigma'} c_{\mathbf{k}\sigma}$$

This the form that is generally used. The inverse Fourier transform is then given by :

$$V(\mathbf{r}) = \frac{1}{\Omega} \sum_{\mathbf{q}} V(\mathbf{q}) e^{i\mathbf{q}.\mathbf{r}}$$

or, for Coulomb potential :

$$\frac{e^2}{r} = \frac{1}{\Omega} \sum_{\mathbf{q}} \frac{4\pi e^2}{\mathbf{q}^2} e^{i\mathbf{q}.\mathbf{r}}.$$