

EPFL

Solid State Physics III

Main author: **F. Mila**

Modifications by: **L. Herviou**

Master course
Year 2025-2026

Contents

1	Preamble	1
1.1	Scope	1
1.2	Organization of the manuscript	2
2	Introduction	3
2.1	A universal model with emergent properties	3
2.2	Independent electrons in a periodic potential	7
2.2.1	Bloch theorem	7
2.2.2	Band theory	10
2.3	Mott insulators	16
3	Magnetism of Mott insulators	19
3.1	Low-energy models in the Mott regime	19
3.1.1	Introduction	19
3.1.2	Elementary two-site calculation	20
3.1.3	Effective Hamiltonian	23
3.2	The Heisenberg model	26
3.2.1	General considerations	26
3.2.2	Spin algebra	27
3.2.3	Symmetries	28
3.3	Ferromagnetic case	29
3.3.1	Ground state	29
3.3.2	Exact low-energy states:	31
3.3.3	Holstein-Primakoff transformation	35
3.4	Antiferromagnetic case	39
3.4.1	Classical case	39
3.4.2	Quantum formulation	39
3.4.3	Bogoliubov transformation	41

3.5	Frustrated magnets	48
3.5.1	Looking for the classical ground state	48
3.5.2	Spin waves in a helical system	52
3.5.3	Bonus: $J_1 - J_2$ model	58
4	Electrons in a magnetic field	63
4.1	Landau levels	63
4.1.1	Classical problem	63
4.1.2	Lagrangian approach	64
4.1.3	Hamiltonian formulation	65
4.1.4	Quantum treatment	66
4.1.5	Degeneracy	67
4.1.6	Density of states	69
4.1.7	Exact solutions of the Landau problem	71
4.2	The integer Quantum Hall Effect	72
4.2.1	The classical Hall effect	72
4.2.2	Hall effect and Landau levels	74
4.2.3	Effect of disorder	79
4.3	The Fractional Quantum Hall Effect	80
4.3.1	Introduction	80
4.3.2	Landau levels and symmetric gauge	80
4.3.3	The wave function of the $\nu = 1$ state:	85
4.3.4	$\nu < 1$: Laughlin's wave functions	87
5	Electron-phonon interaction	94
5.1	Introduction	94
5.2	Phonons - Reminder	94
5.3	Quantization	97
5.3.1	Canonical quantization	97
5.3.2	Energetics	97
5.3.3	Thermodynamics	98
5.3.4	Limitations	98
5.4	Electron-phonon interaction	99
5.5	The effective electron-electron interaction	101
6	Superconductivity	107
6.1	Introduction	107

6.2	The mean-field approach	108
6.2.1	Bogoliubov transformation	109
6.2.2	Ground state	112
6.2.3	Self-consistent equations	113
6.3	Interpretation of the results	115
6.3.1	Condensation energy	115
6.3.2	Energy comparison	118
6.3.3	Interpretation of the wave function	119
6.3.4	Elementary excitations	122
6.4	Finite temperatures	122
6.5	Electrodynamics of superconductors	124
6.5.1	The London equation	124
6.5.2	Response to a static magnetic field:	127
6.5.3	Critical field and field penetration:	131
6.6	The Landau-Ginzburg theory	134
6.7	Flux quantization	143
6.8	Josephson effect:	144
7	Appendix - Second quantization	149
7.1	Introduction	149
7.2	Bosons	150
7.2.1	Fock space	150
7.2.2	Harmonic oscillator	153
7.2.3	Relationship between bosons and harmonic oscillators	154
7.3	Fermions	158
7.4	Field operator	160
7.5	Application: interacting electrons	162
7.5.1	Reference eigenfunctions	162
7.5.2	Operators	164
7.5.3	Hamiltonian in second quantization	165

Chapter 1

Preamble

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 Organization of the manuscript

These lecture notes are still a work in progress and will be updated throughout the semester. The main differences from the original lecture notes by Prof. Mila lie in the presentation and the organization of the sections. We will partially reshuffle the manuscript to match the order used in the course. Additionally, we will try to visually separate the main content of the course from the proofs and computations (shown in dark gray). We emphasize that these computations are crucial: as physicists, we must be able to calculate. While a good conceptual understanding remains most important, it is not possible to neglect the process of computing, which will be a large part of your work should you continue in academia.

Chapter 2

Introduction

2.1 A universal model with emergent properties

In this course, we explore the physics of materials. These are composed of a tremendously large set of atoms and electrons interacting primarily through the Coulomb force, as other fundamental forces are negligible at the energy scales we consider. Consequently, the "theory of everything" in solid-state physics reduces to a deceptively simple, universal Hamiltonian:

$$H = \sum_{j=1}^{N_A} -\frac{\hbar^2}{2M_j} \Delta_{\mathbf{R}_j} + \sum_{j=1}^{N_e} -\frac{\hbar^2}{2m_e} \Delta_{\mathbf{r}_j} + \sum_{i,j} \frac{(-e)q_j}{|\mathbf{r}_i - \mathbf{R}_j|} + \frac{q_i q_j}{|\mathbf{R}_i - \mathbf{R}_j|} + \frac{ee}{|\mathbf{r}_i - \mathbf{r}_j|}.$$

(2.1)

Here, \mathbf{R}_j corresponds to the position of an atom of mass M_j and charge q_j , while \mathbf{r}_j is the position of an electron of mass m_e and charge $-e$.

A central theme of our study will be how this conceptually straightforward equation gives rise to an astonishing array of complex and diverse behaviours. Many of these are *emergent* phenomena, arising from global, collective particle motions that cannot be deduced from the properties of any single constituent. This semester, we will focus on three principal topics that exemplify this emergence:

1. (Frustrated) Magnetism: In certain materials, electrons and atoms are largely immobile, and the key excitations are spin waves—large-scale collective oscillations of the valence electrons' spins. The resulting magnetic regime (ferro-, antiferro-, or paramagnetic) depends critically on the geometry and properties of the atomic lattice (see Fig. 2.1). In some geometries, frustration can produce an exponential number of classical ground states. In the quantum regime,

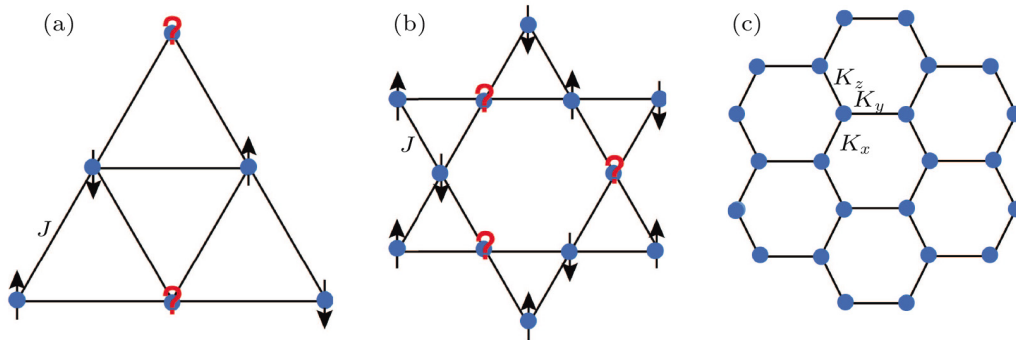


Figure 2.1: Coulomb interactions can freeze electrons at a density of one valence electrons per (relevant) atomic site. The remaining degree of freedom is then the electronic spin. If the effective interaction is antiferromagnetic, certain lattices (e.g., (a) triangular, (b) Kagome) can be frustrated: no classical solution minimize the energy. On the other hand, bipartite lattices (such as (c) honeycomb) admit solutions close to the classical limit. Fig. taken from [Chinese Physics B, 2018, 27\(10\): 106101](#).

this often leads to novel exotic states known as quantum spin liquids, whose properties and conditions for existence are subjects of very active research.

2. Quantum Hall effect: Classically, a two-dimensional electron gas in a perpendicular magnetic field exhibits a non-zero transverse conductance; applying an electric field in one direction induces a current in the perpendicular direction. The quantum mechanical discovery, made in 1980, was far more remarkable: at low temperatures in clean samples, the longitudinal resistance vanishes while the transverse conductivity becomes quantized in integer multiples of the conductance quantum,

$$\sigma_{xy} = \frac{e^2}{h} \times \nu, \quad (2.2)$$

where $\nu \in \mathbb{N}$ is related to the electron density. An experimental plot is shown in Fig. 2.2. This quantization is incredibly robust against impurities (stable up to 10^{-9}), leading to its use in metrology to define the ohm. The subsequent 1982 discovery of the fractional quantum Hall effect, with stable plateaus at non-integer ν , revealed a complex emergent phenomenon driven by electron interactions. Both effects are deeply connected to topology, and the nature of the fractional states remains a vibrant area of research.

3. Superconductivity: First observed in 1911, superconductivity is characterized by a complete disappearance of electrical resistance, usually at low temperatures, and the expulsion of magnetic fields—the Meissner effect. A theoretical understanding of conventional superconductivity was not achieved until the 1950s. The pivotal BCS theory (Bardeen, Cooper, and Schrieffer, 1957) explains it through the formation of bound electron pairs (Cooper pairs), whose

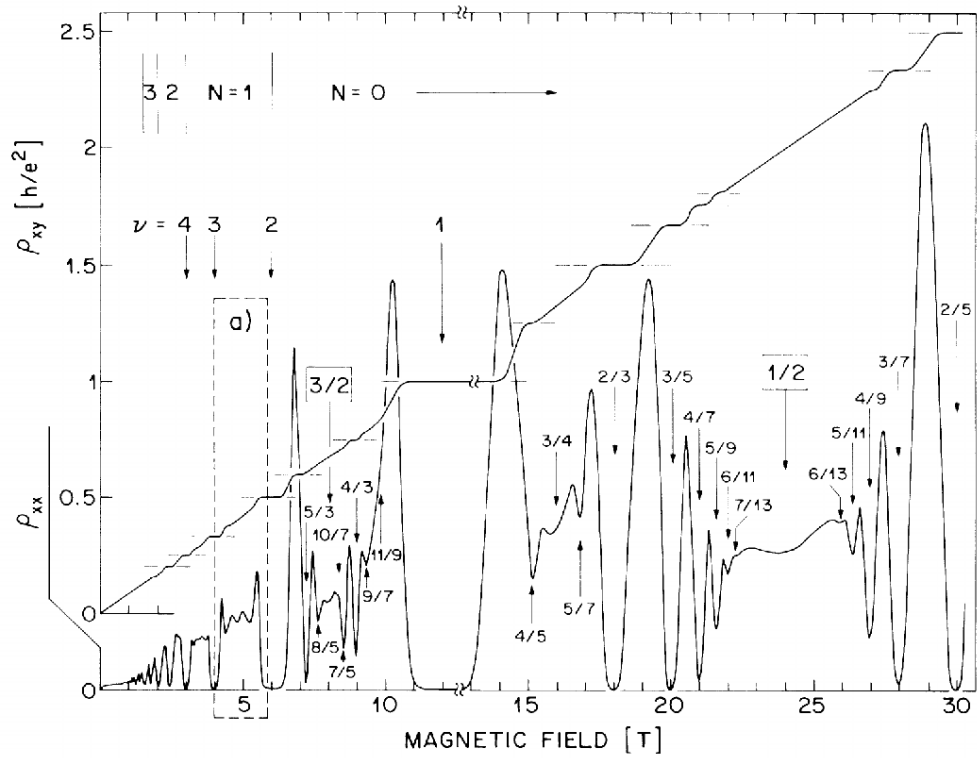


Figure 2.2: Experimental measurement of the diagonal resistivity ρ_{xx} and transverse resistance ρ_{xy} . Stable plateaus of quantized transverse resistance are observed at integer filling (see for example the plateau at $\nu = 1$ or $\nu = 2$) accompanied by a vanishing of the longitudinal resistivity. Secondary plateaus at fractional filling (see $\nu = 2/5$ at $\approx 30T$) are also visible: the fractional Quantum Hall effect. Fig. taken from [Phys. Rev. Lett. 59, 1776 \(1987\)](#).

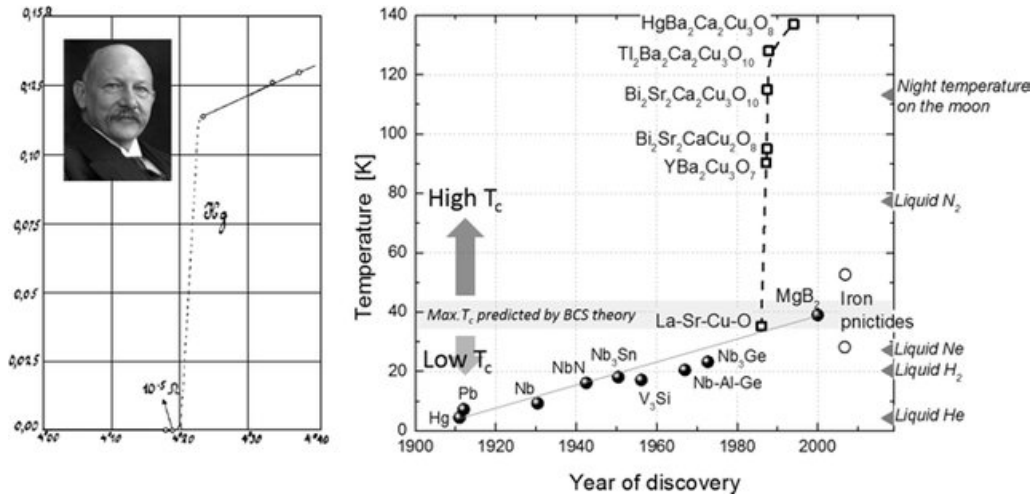


Figure 2.3: Left: original resistance measurements from Kamerlingh Onnes in mercury. At a critical temperature, the resistance abruptly vanishes, marking the opening of the superconducting phase. Right: a partial list of discovered superconductors with their critical temperature. Fig. from [Superconductors at the Nanoscale](#) by Wördenweber and Vanacken.

condensation into a superfluid state yields superconductivity. Although electrons repel each other via the Coulomb force, phonons (collective atomic vibrations) mediate an effective attraction that enables pairing. The late 1980s saw the discovery of unconventional superconductors like the cuprates, which superconduct at temperatures above 150 K and defy explanation by BCS theory (see Fig. 2.3). Understanding these materials remains one of the most significant open challenges in condensed matter physics.

Of course, the Hamiltonian in Eq. (2.1) is utterly unsolvable due to the many-body Coulomb interactions that couple all particles. Indeed, the system's dynamics are chaotic in the classical limit with just three particles, a famously difficult problem illustrated by the three-body problem in celestial mechanics [here](#).¹ This intractability defines the core task of a condensed matter physicist: to derive and study simpler effective models. By judiciously freezing degrees of freedom or cutting off interaction ranges, we aim to isolate physically relevant regimes that can be studied analytically or numerically. We operate on the well-justified assumption that the results from these simplified models remain qualitatively, and often quantitatively, valid in the real world.

A prime example of this approach is the Born-Oppenheimer approximation, which leverages the vast mass disparity between nuclei and electrons (a factor of ≈ 1800).

¹If a student is motivated to realize a similar simulation replacing the three stars and gravity by a mix of positive and negative particles, let me know!

This difference means atoms move on a much slower timescale than electrons. Thus, for low-temperature electronic properties, we can often treat the atoms as frozen, simplifying the Hamiltonian (2.1) into

$$H = \sum_{j=1}^{N_e} -\frac{\hbar^2}{2m_e} \Delta_{\mathbf{r}_j} + \sum_{i,j} \frac{ee}{|\mathbf{r}_i - \mathbf{r}_j|} + \underbrace{\sum_{i,j} \frac{(-e)q_j}{|\mathbf{r}_i - \mathbf{R}_j|}}_{\sum_j V(\mathbf{r}_j)}. \quad (2.3)$$

Here, the atoms generate a static external potential V in which the electrons move. Determining the stable configurations of the atoms themselves — the domain of ab initio computations — is a separate crucial challenge. We will employ this approximation in our studies of frustrated magnetism and the quantum Hall effect. Interestingly, understanding the origins of superconductivity will require us to partially relax this very assumption and reconsider the dynamics of the atoms.

2.2 Independent electrons in a periodic potential

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. In this section, we will focus on these non-interacting models. It is largely a reminder of what you have seen in previous solid state classes.

2.2.1 Bloch theorem

We consider a single electron moving in a three-dimensional periodic potential $U(\mathbf{r})$, for example generated by the atomic centres in the Born-Oppenheimer approximation. The basic Hamiltonian will be

$$\left(-\frac{\hbar^2}{2m} \nabla^2 + U(\mathbf{r}) \right) \psi(\mathbf{r}) = \varepsilon \psi(\mathbf{r}) \quad (2.4)$$

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). The fundamental theorem is *Bloch theorem*.

1: Bloch theorem (first version)

The solutions of Eq. (2.4) are of the form

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

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

2: Bloch theorem (second version)

The solutions of Eq. (2.4) satisfy

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

Comment. Bloch theorem is in fact a simple consequence of the translation invariance. It can be understood as a decomposition of the space of solutions in terms of irreducible representations of the discrete translation group.

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

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

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

2. Conversely, let us suppose that :

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

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

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

We have :

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

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

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

$$T_{\mathbf{R}}\psi(\mathbf{r}) = \psi(\mathbf{r} + \mathbf{R}) \tag{2.5}$$

We first prove that it commutes with H . Without lack of generality, we take $\mathbf{R} = a\hat{\mathbf{x}}$. Trivially,

$$\partial_y T_{\mathbf{R}}\psi(\mathbf{r}) = \partial_y \psi(x + a, y, z) = (\partial_y \psi)(x + a, y, z) = T_{\mathbf{R}}\partial_y \psi(\mathbf{r})$$

and

$$\partial_x T_{\mathbf{R}}\psi(x) = \partial_x \psi(x+a) = \frac{\partial \psi(x+a)}{\partial(x+a)} \underbrace{\frac{d}{dx}(x+a)}_{=1} = (\partial_x \psi)(x+a) = T_{\mathbf{R}}\partial_x \psi(x)$$

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 (i.e., periodic) boundary conditions :

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

for a system with $N = N_1N_2N_3$ unit cells, one must have :

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

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

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

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

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

They define the *reciprocal lattice*.

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

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

Brillouin zones

The reciprocal lattice is defined as

$$\mathbf{G} = \sum_j m_j \mathbf{a}_j^* \text{ with } \mathbf{a}_i \cdot \mathbf{a}_j^* = 2\pi \delta_{ij} \text{ and } m_j \in \mathbb{Z} \quad (2.6)$$

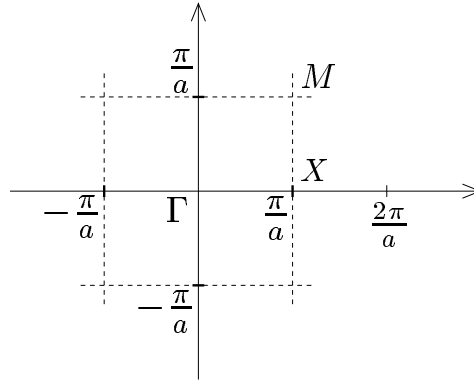
such that

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

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*. It is important to emphasize here that the Brillouin zone is a (n -dimensional) torus, i.e., it is periodic. Two points \mathbf{k} and \mathbf{q} are connected by any vector of the form $\mathbf{q} - \mathbf{k} + \mathbf{G}$, and therefore the distance separating them is $\min_{\mathbf{G}} \|\mathbf{q} - \mathbf{k} + \mathbf{G}\|$. Finally the vectors \mathbf{G} (or equivalently, \mathbf{a}_i^*) defines the lattice *dual* to the lattice generated by \mathbf{R} .

Example : For a square lattice of parameter a .

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



For a triangular lattice of parameter a .

$$\mathbf{a}_1 = a\hat{\mathbf{x}}, \quad \mathbf{a}_2 = a\left(\frac{1}{2}\hat{\mathbf{x}} + \frac{\sqrt{3}}{2}\hat{\mathbf{y}}\right) \Rightarrow \mathbf{a}_1^* = \frac{4\pi}{a\sqrt{3}}\left(\frac{\sqrt{3}}{2}\hat{\mathbf{x}} - \frac{1}{2}\hat{\mathbf{y}}\right), \quad \mathbf{a}_2^* = \frac{4\pi}{a\sqrt{3}}\hat{\mathbf{y}}$$

2.2.2 Band theory

Electronic bands and orbitals.

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

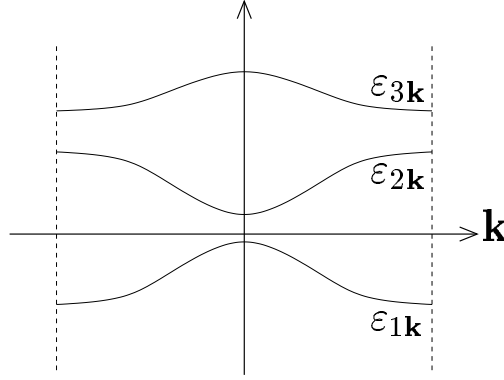
simply determine :

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

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

$$\left[-\frac{\hbar^2}{2m} \left(-\mathbf{k}^2 + 2i\mathbf{k} \cdot \nabla + \nabla^2 \right) + U(\mathbf{r}) \right] u_{\mathbf{k}}(\mathbf{r}) = \varepsilon_{\mathbf{k}} u_{\mathbf{k}}(\mathbf{r}) \quad (2.7)$$

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



At this stage, it is important to take into account the spin $\frac{1}{2}$ of the electrons that we have neglected so far. In general, each quantum state (or orbital) 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 band theory, one assumes that the electrons do not interact with each other. The many-body states with N electrons are obtained by choosing a set of N quantum states which will be occupied by an electron and will have total energy

$$E = \sum_{\text{occupied}} \varepsilon_{n,k,\sigma}$$

The groundstate with N electrons is therefore obtained by filling the N lowest-energy quantum 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 $|n\rangle$ of

energy ϵ_n located at the sites i of a lattice (typically different (Wyckoff) positions in the crystal generated by the atoms). Let us denote by $|i, n\rangle$ the associated ket. In the absence of hybridization between the states, i.e., in the limit where the sites are very far apart, each energy level ϵ_n is N -fold degenerate, where N is the number of sites. If the sites are brought closer to each other, a coupling will develop as electrons will be able to tunnel from one site to the next. It will take the form of a perturbation V that lifts this degeneracy. Assuming that the coupling decays fast with distance, and considering a single orbital per site for simplicity, one can assume that only nearest-neighbor sites have a significant overlap:

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

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

$$\mathcal{H} = -t \sum_{\langle i, j \rangle} (|i\rangle\langle j| + |j\rangle\langle i|) + \epsilon_0 \sum_i |i\rangle\langle i| \quad (2.8)$$

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

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

$$|\mathbf{k}\rangle = \frac{1}{\sqrt{N}} \sum_j e^{-i\mathbf{k}\cdot\mathbf{R}_j} |j\rangle \Leftrightarrow |j\rangle = \frac{1}{\sqrt{N}} \sum_{\mathbf{k}} e^{i\mathbf{k}\cdot\mathbf{R}_j} |\mathbf{k}\rangle. \quad (2.9)$$

The Fourier transform diagonalize translation-invariant system (this is the analogue of Bloch theorem for tight-binding Hamiltonians). Indeed, $|\mathbf{k}\rangle$ is an eigenstate of \mathcal{H} with energy

$$E_{\mathbf{k}} = \epsilon_0 - 2t \sum_{\alpha=1}^D \cos(k_{\alpha}a), \quad (2.10)$$

where α runs over the directions of space.

Proof: application of the state. Applying the Hamiltonian to this state leads to:

$$\begin{aligned}
\mathcal{H}|\mathbf{k}\rangle &= \epsilon_0|\mathbf{k}\rangle + \frac{1}{\sqrt{N}} \sum_j e^{-i\mathbf{k}\cdot\mathbf{R}_j} (-t) \sum_\tau |j+\tau\rangle \\
&= \epsilon_0|\mathbf{k}\rangle + \frac{1}{\sqrt{N}} (-t) \sum_\tau e^{i\mathbf{k}\cdot\tau} \sum_j e^{-i\mathbf{k}\cdot(\mathbf{R}_j+\tau)} |j+\tau\rangle \\
&= \epsilon_0|\mathbf{k}\rangle - t \sum_\tau e^{i\mathbf{k}\cdot\tau} |\mathbf{k}\rangle \\
&= (\epsilon_0 - 2t \sum_\alpha \cos(k_\alpha a)) |\mathbf{k}\rangle
\end{aligned}$$

where the vectors τ are the vectors connecting a site to its nearest neighbors.

Proof: transformation of the Hamiltonian.

$$\begin{aligned}
\mathcal{H} &= -\frac{t}{N} \sum_{j,\tau} \sum_{k,q} e^{i\mathbf{k}\cdot\mathbf{R}_j} e^{-i\mathbf{q}\cdot(\mathbf{R}_j+\tau)} |\mathbf{k}\rangle \langle \mathbf{q}| + h.c. + \frac{\epsilon_0}{N} \sum_j \sum_{k,q} e^{i(\mathbf{k}-\mathbf{q})\cdot\mathbf{R}_j} |\mathbf{k}\rangle \langle \mathbf{q}| \\
&= -t \sum_\tau \sum_{k,q} \delta_{\mathbf{k},\mathbf{q}} e^{-i\mathbf{q}\cdot\tau} |\mathbf{k}\rangle \langle \mathbf{q}| + h.c. + \epsilon_0 \sum_{k,q} \delta_{\mathbf{k},\mathbf{q}} |\mathbf{k}\rangle \langle \mathbf{q}| \\
&= \sum_{\mathbf{k}} (\epsilon_0 - 2t \sum_\tau \cos \mathbf{k}\cdot\tau) |\mathbf{k}\rangle \langle \mathbf{k}|,
\end{aligned}$$

where we used

$$\sum_j e^{i\mathbf{k}\cdot\mathbf{R}_j} = N\delta_{\mathbf{k},0}$$

if \mathbf{k} is a vector of the reciprocal lattice.

This band is centered at ϵ_0 , and the band width is given by $W = 4Dt$. In the lowest-energy state, the particle goes to the bottom of the band. Its energy is lowered by half the bandwidth $W/2$ as it delocalizes through the whole system.

Physical implications of band theory

Placing ourselves in the thermodynamic limit, there will in general be energies that do not correspond to any single-particle energy $\epsilon_{n\mathbf{k}\sigma}$ for any quantum numbers (n, \mathbf{k}, σ) . These energies form intervals known as gaps. Conversely, two bands sharing exactly the same energy are said to be degenerate; this often indicates the presence of a (possibly non-Abelian) symmetry.

As a reminder, the ground state of a bulk system is obtained by filling the energy levels starting from the lowest one. There are three possibilities:

1. All bands are either completely filled, or completely empty. The first excited state is separated from the ground state by the energy Δ that separates the last occupied state from the first empty one. The system is said to be *gapped*.

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 (due to the spin degeneracy, assuming no spin splitting), this is only possible if the number of electrons per unit cell is *even*.

2. One or several bands are partially filled. The system is said to be gapless. The energy of the last occupied state is called the *Fermi energy* and is denoted by ε_F . In that case, the separation between the last occupied state and the first empty one goes to 0 when $N \rightarrow \infty$. The system can react to infinitesimal excitations. Such systems are called *metals* (or conductors).
3. Bands are either completely filled, or completely empty, but two bands are touching at the Fermi energy. The state is therefore gapless, and will typically correspond to a semi-metal or a semi-conductor.

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 can look like:

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

where z is a complex number of modulus not equal to 1. Note that the Bloch theorem do not a priori apply as the system is not fully translation invariant. 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.

Strengths and limits of band theory

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 susceptibility:

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

The semi-classical theory of transport based on band theory in general explain many properties of materials.

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:

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

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 7).

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 6 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.

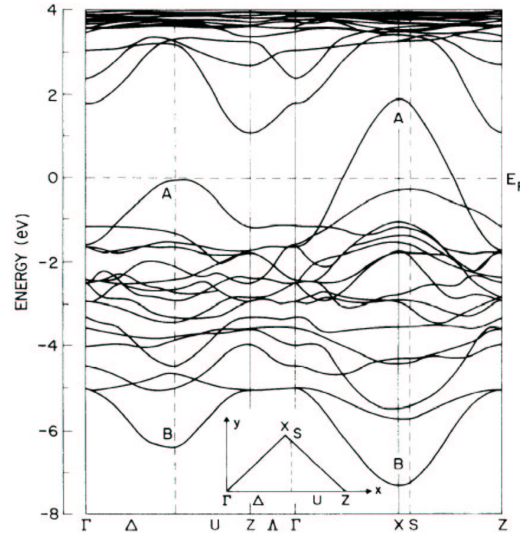


Figure 2.4: LAPW energy bands for La_2CuO_4 along symmetry lines in the bct Brillouin zone. L.F. Mattheiss, Phys. Rev. Lett. **58**, 1028 (1987).

2.3 Mott insulators

In 1937, only a few years after the development of band theory (Bloch 1928, Peierls 1929, Wilson 1931), de Boer and Verwey observed that NiO, an insulator with a band gap of approximately 4eV, should, according to band theory, be a metal. During the discussion following their presentation, Peierls suggested that this discrepancy might be a consequence of electron–electron interactions. It was only several years later that Mott formulated this explanation in a precise way. Since then, many counterexamples have been discovered. The most famous is La_2CuO_4 : when doped with Sr (forming $\text{La}_{2-x}\text{Sr}_x\text{CuO}_4$), this compound becomes superconducting below 40K, a discovery for which Bednorz and Müller were awarded the 1987 Nobel Prize.

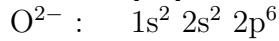
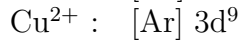
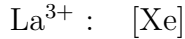
The band structure of La_2CuO_4 was calculated by Mattheiss in 1987 (see Fig. 2.4). Since there is an odd number of electrons per unit cell, at least one band must be partially filled. This is indeed the case : one band is half filled.

To understand this band structure, it is useful to begin by viewing the solid as a collection of atoms, each with specific atomic orbitals that are more or less filled. The atomic electron configurations are

$$\begin{aligned} \text{La} &: [\text{Xe}] 6s^2 5d^1 \\ \text{Cu} &: [\text{Ar}] 3d^{10} 4s^1 \\ \text{O} &: 1s^2 2s^2 2p^4 \end{aligned}$$

The relative energy positions of the valence orbitals control their filling in the solid. It turns out that the $2p$ orbitals of the oxygen lie lower in energy than the valence orbitals of Cu and La. As a result, the electronic configurations in the solid

become

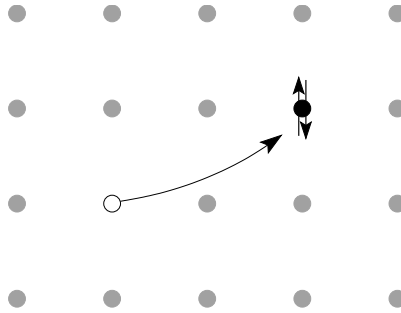


Thus, the compound $\text{La}_2^{3+}\text{Cu}^{2+}\text{O}_4^{2-}$ is charge-neutral, as expected.

Within a tight-binding approach, the hopping integrals between Cu-3*d* and O-2*p* orbitals are substantial ($\simeq 1\text{eV}$), and the energy separation between their atomic levels is modest ($\simeq 2\text{eV}$). This leads to significant hybridization between these orbitals. The Fermi level consequently lies within an energy region where 17 bands overlap: 5 from Cu-3*d* orbitals and 12 (from 4 O atoms \times 3 *p*-orbitals each) from O-2*p* orbitals.

Although this approach is generally applicable and successfully describes metallic band structures, it fails to accurately capture the ground state of this compound. The central argument advanced by Mott involves studying the electronic excitations in a solid composed of hydrogen atoms as a function of interatomic distance. For simplicity, consider an imaginary solid consisting of hydrogen atoms arranged on a square lattice. With one hydrogen atom per unit cell and one electron per atom, band theory predicts metallic behavior. However, if a denotes the distance between atoms, the system cannot remain metallic as a becomes very large. To induce a current, an electron must be moved far away, which in a metal should require only infinitesimal energy. Let us estimate the energy cost of this process.

For large a , the system begins with one electron per site. Moving an electron to a distant site creates one hole and one doubly occupied site:



The potential energy of this excited state is approximately

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

where H^- and H^+ represent the local holon and doublon. Without hopping, this energy is strictly positive because of the Coulomb repulsion between the electrons which we denote by U .

However, the system can gain kinetic energy through the delocalization of the hole and the doubly occupied site. Assuming for simplicity that all singly occupied sites have the same spin (e.g., \uparrow), both the hole and doublon can be described by a

tight-binding model with hopping amplitude $-t$. If W is the associated bandwidth, the hole and doublon each gain at most $W/2$ by occupying the band bottom. In reality, spins may vary, so $W/2$ is an upper bound, but the order of magnitude remains valid. In two dimensions, the dispersion is

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

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

Finally, 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 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*. Δ_c is then called the *charge gap*.

When a decreases, t increases, and there is a value of the order of a_0 for which $U = W$. Below this value, the calculation of Δ_c is not valid anymore as a 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 Δ_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 overturn 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 Appendix.

Chapter 3

Magnetism of Mott insulators

3.1 Low-energy models in the Mott regime

3.1.1 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 corresponds to a configuration where each atom carries an odd number of electrons, say 1 for simplicity, and the spin of this electron can be \uparrow or \downarrow . In the purely atomic limit where atoms are infinitely far apart, the excitation spectrum is very simple. We assume the presence of a single orbital per site ($\times 2$ with the spin degeneracy). 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 ϵ_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 \dots

$$\begin{aligned} \text{Groundstate manifold: } E &= 0 && 2^N. \\ \text{1st excited manifold: } E &= U && N(N-1)2^{N-2} \\ \text{2nd excited manifold: } E &= 2U && \dots \end{aligned}$$

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

$$\mathcal{H} = -t \sum_{\langle i,j \rangle, \sigma} (c_{i,\sigma}^\dagger c_{j,\sigma} + h.c.) + U \sum_i n_{i,\uparrow} n_{i,\downarrow}. \quad (3.1)$$

The first term is a hopping term between 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. It is a minimal representation of the Coulomb repulsion.

To determine the **effective Hamiltonian** that describes the lifting of degeneracy in the ground state manifold — when $U \gg |t|$ — it is thus necessary to carry on degenerate perturbation theory with respect to the kinetic term

$$\mathcal{H}_{\text{kin}} = -t \sum_{\langle i,j \rangle, \sigma} \left(c_{i,\sigma}^{\dagger} c_{j,\sigma} + \text{h.c.} \right). \quad (3.2)$$

Remark: Here, the particle density will be fixed, but it is possible to introduce a chemical potential

$$-\mu \sum_{j,\sigma} n_{j,\sigma}$$

if we want to study the system in other regimes. Similarly, one could introduce a Zeeman spin coupling with a magnetic field, with

$$-g \sum_{j,\alpha,\beta} c_{j,\alpha}^{\dagger} \left(\sum_{\gamma=x,y,z} B_{\gamma} \sigma^{\gamma} \right)_{\alpha,\beta} c_{j,\beta},$$

where σ^{γ} are the Pauli matrices and \mathbf{B} is the magnetic field.

3.1.2 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 $c_{1\uparrow}^{\dagger}$, $c_{1\downarrow}^{\dagger}$, $c_{2\uparrow}^{\dagger}$ et $c_{2\downarrow}^{\dagger}$. 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 $\binom{4}{2} = 6$ of them, given by:

$$\begin{aligned} |1\rangle &\equiv |1, 0, 1, 0\rangle = c_{1\uparrow}^{\dagger} c_{2\uparrow}^{\dagger} |0\rangle \\ |2\rangle &\equiv |0, 1, 0, 1\rangle = c_{1\downarrow}^{\dagger} c_{2\downarrow}^{\dagger} |0\rangle \\ |3\rangle &\equiv |1, 0, 0, 1\rangle = c_{1\uparrow}^{\dagger} c_{2\downarrow}^{\dagger} |0\rangle \\ |4\rangle &\equiv |0, 1, 1, 0\rangle = c_{1\downarrow}^{\dagger} c_{2\uparrow}^{\dagger} |0\rangle \\ |5\rangle &\equiv |1, 1, 0, 0\rangle = c_{1\uparrow}^{\dagger} c_{1\downarrow}^{\dagger} |0\rangle \\ |6\rangle &\equiv |0, 0, 1, 1\rangle = c_{2\uparrow}^{\dagger} c_{2\downarrow}^{\dagger} |0\rangle \end{aligned}$$

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

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

These results are straightforward given the form of the Hamiltonian, up to one subtlety, the sign, due to the fact fermionic anticommutation.

Indeed, let us consider for example the effect of the term $c_{2\downarrow}^+ c_{1\downarrow}$ on state $|4\rangle$. We get:

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

The eigenvalues of the system are simply given by 0 (three times degenerate), U and

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

We can notice that

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

so that we have only a 2-level system to solve:

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

i.e. we need to diagonalize

$$\begin{pmatrix} 0 & -2t \\ -2t & U \end{pmatrix} = \frac{U}{2} \text{Id} + \begin{pmatrix} -\frac{U}{2} & -2t \\ -2t & \frac{U}{2} \end{pmatrix}$$

It is now convenient to use the fact that the eigenvalues of

$$H = \sum_{\gamma=x,y,z} n_{\gamma} \sigma^{\gamma}$$

are simply given by $\|\mathbf{n}\|$. The proof is straightforward:

$$H^2 = \sum_{\gamma,\alpha} n_{\gamma} n_{\alpha} \sigma^{\gamma} \sigma^{\alpha} = \sum_{\gamma} n_{\gamma}^2 (\sigma^{\gamma})^2 + \frac{1}{2} \sum_{\gamma \neq \alpha} n_{\gamma} n_{\alpha} (\sigma^{\gamma} \sigma^{\alpha} + \sigma^{\alpha} \sigma^{\gamma})$$

We now use $\{\sigma^{\alpha}, \sigma^{\beta}\} = 2\delta_{\alpha,\beta}$ to obtain

$$H^2 = \sum_{\gamma} n_{\gamma}^2$$

Using $\text{Tr}(H) = 0$, we conclude.

In our case, the corresponding eigenvalues 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 and only one electron on each site (2^N states). The hopping term partially 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

$$H = \frac{4t^2}{U} (\mathbf{S}_1 \cdot \mathbf{S}_2 - \frac{1}{4}) \quad (3.3)$$

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

Remark: the existence of a triplet of degenerate states should alert you. In the absence of symmetries, generic Hamiltonian are never degenerate (the $t = 0$ limit of the Hubbard Hamiltonian is far from generic). Stable degeneracies are typically the symptom of a (non-Abelian) symmetry. Here, in particular, the existence of a singlet and a triplet of states strongly suggest the existence of a SU(2) rotational, with a spin-0 and spin-1 state. And indeed, the Hamiltonian in Eq. (3.3) is invariant under a global rotation of the spins.

3.1.3 Effective Hamiltonian

There are two primary methods to rigorously derive the form of the effective Hamiltonian to second order: canonical transformations and degenerate perturbation theory. Note that this distinction is somewhat arbitrary and is largely a matter of perspective. We will use a canonical transformation later in this course. In the present case, one can directly apply degenerate perturbation theory. We recall in the following the expressions for degenerate perturbation theory (which is in fact generic).

3: Degenerate perturbation theory (1st order)

We consider an Hamiltonian $\mathcal{H} = \mathcal{H}_0 + \mathcal{V}$. Let E_0 be the groundstate energy of \mathcal{H}_0 and E_1 the first excited energy, such that $V = \|\mathcal{V}\| \ll E_1 - E_0$. Let P be the orthogonal projector on the groundstate manifold. To first order in $V/(E_1 - E_0)$, the effective Hamiltonian describing the perturbation induced by \mathcal{V} on the groundstate manifold is

$$\mathcal{H}_{\text{eff}} = E_0 + P\mathcal{V}P.$$

4: Degenerate perturbation theory (2nd order)

Let Q be the orthogonal projector on the excited manifold, such that $P + Q = \text{Id}$. The second-order correction to the effective Hamiltonian is

$$\mathcal{H}_{\text{eff}}^{(2)} = -P\mathcal{V}Q \frac{1}{\mathcal{H}_0 - E_0} Q\mathcal{V}P.$$

Let us apply these formula to our example.

The low-energy subspace include all states with a single electron per site, and arbitrary spins (2^N states). To first order, following Th. 3, since \mathcal{H}_{kin} 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, denoting by $|m\rangle$ and $|n\rangle$ two elements of the low-energy subspace, the matrix elements are given by

$$\langle m | \mathcal{H}_{\text{eff}} | n \rangle = \sum'_{|k\rangle} \frac{\langle m | \mathcal{H}_{\text{kin}} | k \rangle \langle k | \mathcal{H}_{\text{kin}} | n \rangle}{E_0 - E_k}$$

where the sum $\sum'_{|k\rangle}$ is restricted to states which do not belong to the groundstate manifold. We can additionally restrain them to have a single doubly occupied state (and a single empty site), given the groundstate manifold. Consequently, $E_k = U$

and we obtain

$$\langle m | \mathcal{H}_{\text{eff}} | n \rangle = -\frac{1}{U} \sum'_{|k\rangle} \langle m | \mathcal{H}_{\text{kin}} | k \rangle \langle k | \mathcal{H}_{\text{kin}} | n \rangle$$

The operator $\sum'_{|k\rangle} |k\rangle \langle k|$ is the identity in the subspace of states with a single 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

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

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

$$c n = c c^+ c = (1 - c^+ c) c = c \implies c c^+ = 1 - n.$$

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:

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

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:

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

where the prefactor 2 comes from the fact that, for each pair $\langle i, j \rangle$, 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):

$$\begin{cases} S_i^z = \frac{1}{2} (n_{i,\uparrow} - n_{i,\downarrow}) \\ S_i^+ = c_{i,\uparrow}^+ c_{i,\downarrow} \\ S_i^- = c_{i,\downarrow}^+ c_{i,\uparrow} \end{cases}$$

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

$$\begin{aligned} S_i^z S_j^z - \frac{1}{4} n_i n_j &= -\frac{1}{2} (n_{i,\uparrow} n_{j,\downarrow} + n_{i,\downarrow} n_{j,\uparrow}) \\ S_i^+ S_j^- + S_i^- S_j^+ &= c_{i,\uparrow}^+ c_{i,\downarrow} c_{j,\downarrow}^+ c_{j,\uparrow} + c_{i,\downarrow}^+ c_{i,\uparrow} c_{j,\uparrow}^+ c_{j,\downarrow} \end{aligned}$$

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

$$\mathcal{H}_{eff} = \frac{4t^2}{U} \sum_{\langle i,j \rangle} \left(S_i^z S_j^z - \frac{1}{4} n_i n_j + \frac{1}{2} (S_i^+ S_j^- + S_i^- S_j^+) \right).$$

But $\mathbf{S}_i \cdot \mathbf{S}_j = S_i^z S_j^z + \frac{1}{2}(S_i^+ S_j^- + S_i^- S_j^+)$ and $n_i = n_j = 1$ in the ground state manifold. Finally, the effective Hamiltonian takes the form:

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

Importantly, this Hamiltonian does preserve the groundstate manifold: we do not add constraints. Up to a constant, this is the *Heisenberg model* defined by:

$$\mathcal{H}_{Heis} = J \sum_{\langle i,j \rangle} \mathbf{S}_i \cdot \mathbf{S}_j, \quad (3.4)$$

with the coupling constant $J = \frac{4t^2}{U}$. This constant is *positive*. 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.

3.2 The Heisenberg model

3.2.1 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 there 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). This will depend on the nature and size of the valence orbitals, and of the lattice spacing.
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 $S \rightarrow +\infty$ limit of the quantum case.
5. **The dimensionality of space.** Most of the properties are radically different depending on whether one deals with a system (and a lattice) 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 relevant methods depend on the energy scale, or equivalently on the temperature scale, that we study. If we study low-temperature properties, i.e. $T/J \ll 1$, one should typically try to determine the groundstate and the low-energy excited states. The most commonly used analytical 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. In both regimes, a large variety of numerical methods (Monte Carlo, exact diagonalization, tensor networks, machine learning) have been proposed.

3.2.2 Spin algebra

The commutation relations between spin operators are

$$[S^x, S^y] = iS^z \quad (3.5)$$

with the convention $\hbar = 1$. These commutation relations define the generators of the complex rotation group $SU(2)$ (up to some global factor). The ladder operators defined by

$$S^+ = S^x + iS^y \quad S^- = S^x - iS^y$$

satisfy the commutation relations:

$$\begin{cases} [S^+, S^-] = [S^x, -iS^y] + [iS^y, S^x] = 2S^z \\ [S^z, S^+] = S^+ \\ [S^z, S^-] = -S^- \end{cases} \quad (3.6)$$

The action of the spin operators on a state of total spin S is given by

$$\begin{cases} S^+ |m\rangle = \sqrt{S(S+1) - m(m+1)} |m+1\rangle \\ S^- |m\rangle = \sqrt{S(S+1) - m(m-1)} |m-1\rangle \end{cases} \quad (3.7)$$

With some minor algebra, we recover, for a spin S state,

$$(S^x)^2 + (S^y)^2 + (S^z)^2 = (S^z)^2 + \frac{1}{2}(S^+S^- + S^-S^+) = S(S+1) \quad (3.8)$$

3.2.3 Symmetries

Let us determine the symmetries of the spin Hamiltonian

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

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

- $[S_{\text{tot}}^\alpha, H] = 0$, $\alpha = x, y, z, +, -$
- $[\mathbf{S}_{\text{tot}}^2, H] = 0$

Proof:

$$\begin{aligned} [S_i^z, \mathbf{S}_i \cdot \mathbf{S}_j] &= [S_i^z, S_i^x] S_j^x + [S_i^z, S_i^y] S_j^y \\ &= iS_i^y S_j^x - iS_i^x S_j^y \end{aligned}$$

$$\begin{aligned} [S_j^z, \mathbf{S}_i \cdot \mathbf{S}_j] &= S_i^x [S_j^z, S_j^x] + S_i^y [S_j^z, S_j^y] \\ &= iS_i^x S_j^y - iS_i^y S_j^x \end{aligned}$$

$$\Rightarrow [S_i^z + S_j^z, \mathbf{S}_i \cdot \mathbf{S}_j] = 0$$

$$\Rightarrow \left[\sum_i S_i^z, \mathbf{S}_i \cdot \mathbf{S}_j \right] = 0$$

$$\Rightarrow [S_{\text{tot}}^z, \mathbf{S}_i \cdot \mathbf{S}_j] = 0$$

$$\Rightarrow [S_{\text{tot}}^z, H] = 0$$

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

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

In addition, since S_{tot}^+ et S_{tot}^- are linear combinations of S_{tot}^x et S_{tot}^y , one also has:

$$[S_{\text{tot}}^+, H] = [S_{\text{tot}}^-, H] = 0$$

Finally, since $\mathbf{S}_{\text{tot}}^2 = (S_{\text{tot}}^z)^2 + (S_{\text{tot}}^x)^2 + (S_{\text{tot}}^y)^2$, one sees immediately that:

$$[\mathbf{S}_{\text{tot}}^2, H] = 0.$$

As the total spin commutes with the Hamiltonian (with isotropic), it is a conserved quantity. Physically, it means that the Hamiltonian is invariant under rotation of all spins, that is to say that there is no preferred direction. It is rather straightforward given the structure of the model. We can therefore classify the eigenstates depending on their total spin: the Hamiltonian do not couple states with different $\mathbf{S}_{\text{tot}}^2$ or S^z . Additionally, states of total spin S are necessarily $2S + 1$ degenerate. Indeed, given an eigenstate $|\psi\rangle$, $\mathbf{S}_{\text{tot}}^\pm |\psi\rangle$ is either 0 or an eigenstate of same energy. This is a consequence of the *non-Abelian* nature of $\text{SU}(2)$.

3.3 Ferromagnetic case

Let us consider the Hamiltonian

$$H = -J \sum_{\langle i,j \rangle} \mathbf{S}_i \cdot \mathbf{S}_j, \quad J > 0$$

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

3.3.1 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. We denote N the total number of sites.

- $|F\rangle = |m_1 = S, m_2 = S, \dots, m_N = S\rangle$ is an eigenstate. $S_i^z S_j^z = S^2$, $S_i^+ S_j^-$ gives 0. This state is thus an eigenstate of energy:

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

- This is the state of lowest energy.

For this to be true, it is sufficient to prove that:

$$\langle \mathbf{S}_i \cdot \mathbf{S}_j \rangle \leq S^2$$

But

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

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

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

- Degeneracy: $2NS + 1$.

The state has total spin NS , as

$$\begin{aligned} S_{\text{tot}}^z |F\rangle &= 2NS, \\ S_{\text{tot}}^+ |F\rangle &= 0. \end{aligned}$$

Due to the $SU(2)$ invariance, the degeneracy is $2NS + 1$. The proof can also be constructive. 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 $|m_1 = -S, m_2 = -S, \dots, m_N = -S\rangle$, which is annihilated by S_{tot}^- .

Note that the groundstate multiplet is actually the only spin- NS multiplet.

3.3.2 Exact low-energy states:

Let's consider the state

$$|i\rangle = |S \dots \underbrace{S-1}_i S \dots\rangle$$

In other words, it is a state with a single lowered spin at position i . They are simply related to the polarized state $|F\rangle$:

$$|i\rangle = \frac{1}{\sqrt{2S}} S_i^- |F\rangle \quad \text{and} \quad S_i^+ |i\rangle = \sqrt{2S} |F\rangle.$$

We introduce its Fourier transform, the spin-wave $|\mathbf{k}\rangle$

$$|\mathbf{k}\rangle = \frac{1}{\sqrt{N}} \sum_i e^{-i\mathbf{k}\cdot\mathbf{R}_i} |i\rangle,$$

where \mathbf{R}_i are the sites positions. They are eigenstates of the Hamiltonian with energy

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

where τ labels the hopping directions.

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

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

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

Acting on the state $|\mathbf{k}\rangle$, this Hamiltonian gives:

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

\Rightarrow

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

with

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

Remarks.

- For $\mathbf{k} = 0$, one finds $E_{\mathbf{k}} = E_0$. Is this normal? Yes, it is part of the groundstate multiplet as $|\mathbf{k} = 0\rangle_{\alpha} S_{\text{tot}}^{-}|F\rangle$.
- The excitations are slowly oscillating spins as shown in Fig. 3.3.2
- This calculation is very similar to that done for the tight-binding model. Let us compute the dispersion on a cubic lattice. Using $e^{i\mathbf{k}\cdot\tau} + e^{-i\mathbf{k}\cdot\tau} = 2 \cos(\mathbf{k}\cdot\tau)$, and $\tau = \pm a\hat{x}, \pm a\hat{y}, \pm a\hat{z}$,

$$\sum_{\tau} (1 - e^{i\mathbf{k}\cdot\tau}) = 4 \sum_{\alpha} \sin^2\left(\frac{k_{\alpha}a}{2}\right)$$

The dispersion, represented in Fig. 3.1, is

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

- For small \mathbf{k} : $\sin^2\left(\frac{k_{\alpha}a}{2}\right) \sim \frac{k_{\alpha}^2 a^2}{4}$. The dispersion is therefore quadratic for small \mathbf{k} .

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

- The excitations $|\mathbf{k}\rangle$ have total spin $NS - 1$, except for $|\mathbf{k} = 0\rangle$. They are also macroscopically degenerate ($2NS - 1$ times).

In principle, we can conclude from $S_{\text{tot}}^z|\mathbf{k}\rangle = (NS - 1)|\mathbf{k}\rangle$ and $S_{\text{tot}}^+|\mathbf{k} \neq 0\rangle = 0$. We nonetheless show an explicit computation below.

We need to evaluate $\langle \mathbf{k} | \mathbf{S}_{\text{tot}}^2 | \mathbf{k} \rangle = \langle \mathbf{k} | \sum_{i,j} \sum_{\alpha} S_i^{\alpha} S_j^{\alpha} | \mathbf{k} \rangle$

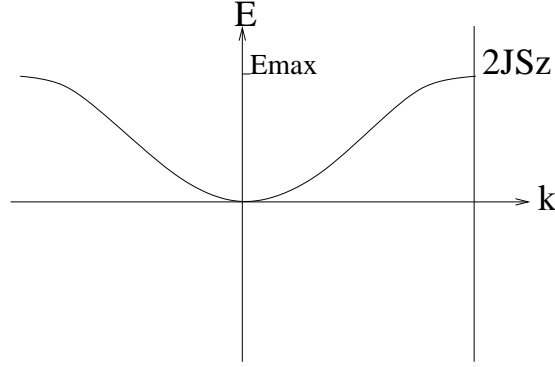


Figure 3.1: Schematic dispersion of the low-energy band in the ferromagnetic Heisenberg model.

Let us first look at

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

Now

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

and

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

So

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

In the same way, using $[S^+, S^-] = 2S^z$, one shows that:

$$S_i^+ S_j^- |\mathbf{k}\rangle = \frac{2S}{\sqrt{N}} e^{-i\mathbf{k}\cdot\mathbf{R}_i} |j\rangle + 2\delta_{i,j} S_i^z |\mathbf{k}\rangle.$$

We then get:

$$\langle \mathbf{k} | \frac{1}{2} (S_i^+ S_j^- + S_i^- S_j^+) | \mathbf{k} \rangle = \frac{2S}{N} \cos \mathbf{k} \cdot (\mathbf{R}_i - \mathbf{R}_j) + \delta_{i,j} \langle \mathbf{k} | S_i^z | \mathbf{k} \rangle.$$

Summing over all positions cancel the cosine if $\mathbf{k} \neq 0$, such that

$$\langle \mathbf{k} | \frac{1}{2} \sum_{i,j} (S_i^+ S_j^- + S_i^- S_j^+) | \mathbf{k} \rangle = \sum_j \langle \mathbf{k} | S_j^z | \mathbf{k} \rangle = NS - 1.$$

We now look for the expectation value of $\langle \mathbf{k} | S_i^z S_j^z | \mathbf{k} \rangle$. Following the same procedure, we get:

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

When $i \neq j$,

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

independent of the indices i and j .

When $i = j$,

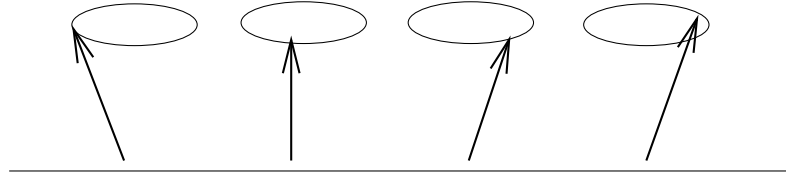
$$\begin{aligned}
(S_j^z)^2 |\mathbf{k}\rangle &= \frac{1}{\sqrt{N}} \left(\sum_{n \neq j} S^2 e^{-i\mathbf{k} \cdot \mathbf{R}_n} |n\rangle + (S-1)^2 e^{-i\mathbf{k} \cdot \mathbf{R}_j} |j\rangle \right) \\
\langle \mathbf{k} | (S_j^z)^2 | \mathbf{k} \rangle &= \frac{1}{N} \sum_{m, n \neq j} e^{i\mathbf{k} \cdot \mathbf{R}_m} S^2 e^{-i\mathbf{k} \cdot \mathbf{R}_n} \langle m | n \rangle \\
&+ \frac{1}{N} \sum_m e^{i\mathbf{k} \cdot \mathbf{R}_m} (S-1)^2 e^{-i\mathbf{k} \cdot \mathbf{R}_j} \langle m | j \rangle \\
&= \frac{1}{N} \left[\sum_{n \neq j} S^2 + (S-1)^2 \right] \\
&= S^2 - \frac{2}{N} S + \frac{1}{N}
\end{aligned}$$

Consequently,

$$\langle \mathbf{k} | \mathbf{S}_{\text{tot}}^2 | \mathbf{k} \rangle = N(N-1) \left(S^2 - \frac{2}{N} S \right) + N \left(S^2 - \frac{2}{N} S + \frac{1}{N} \right) + NS - 1 = NS \times (NS - 1)$$

Technically, we should still prove that $|\mathbf{k}\rangle$ are eigenstates, but it is straightforward.

- The existence of a low-energy branch whose dispersion goes to 0 as \mathbf{k} goes to 0 is a consequence of the (*Nambu*-)Goldstone theorem (see Th. 5).

Figure 3.2: The spinwaves have period $\sim \frac{1}{k}$

5: Nambu-Goldstone theorem

Consider a system invariant under a continuous symmetry. If the groundstate is not invariant under the symmetry, i.e. it belongs to an irrep of dimension > 1 , then there exists a massless mode of vanishing dispersion.

Note that this theorem also put constraints on the number of Goldstone modes depending on the generators of the symmetry group.

- Are there other low-energy states? Yes, of course. One can flip 2 spins, 3 spins, etc \dots , but there are some difficulties. $|\mathbf{k}, \mathbf{k}' \rangle = \frac{1}{N} \sum_{i,j} e^{i(\mathbf{k}\mathbf{r}_i + \mathbf{k}'\mathbf{r}_j)} S_i^- S_j^- |F \rangle$ is not an eigenstate of H . Indeed, for the one magnon states, one looks for the solution as $|\Psi \rangle = \sum_i A_i S_i^- |F \rangle$ and we have seen that $A_i = e^{i\mathbf{k}\cdot\mathbf{R}_i}$ allows one to diagonalize H . For the two-magnon states, one must look for a solution of the form $|\Psi \rangle = \sum_{i,j} A_{i,j} S_i^- S_j^- |F \rangle$. But the coefficients $A_{i,j}$ must satisfy different equations depending on whether i and j are nearest neighbours or not. In other words, the effective models for more than a single flip are strongly-interacting. This problem does not have an exact solution except in dimension 1 for spins 1/2 (Bethe ansatz, 1931).

3.3.3 Holstein-Primakoff transformation

We now want to study excitations around the groundstate $|F \rangle$. As we discussed, direct study is complex for more than a single spin flip. Consequently, we introduce a convenient transformation where we will map the spin degrees of freedom to bosonic degrees of freedom.

Exact representation:

$$\begin{cases} S_i^z = S - a_i^\dagger a_i \\ S_i^+ = \sqrt{2S} \sqrt{1 - \frac{a_i^\dagger a_i}{2S}} a_i \\ S_i^- = \sqrt{2S} a_i^\dagger \sqrt{1 - \frac{a_i^\dagger a_i}{2S}} \end{cases} \quad (3.9)$$

Let us check the commutation relations:

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

Using the relations:

$$\begin{aligned}
a_i^\dagger a_i a_i &= (a_i a_i^\dagger - 1) a_i \\
n_i a_i &= a_i n_i - a_i
\end{aligned}$$

and

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

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,

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

Approximations: Using the exact representation defined above does not typically simplify our Hamiltonian. The Hamiltonian remains strongly-interacting, and we have to deal with the square root of operators. The bosonic density also is bounded from above. It is instead typical to proceed to a series of approximations which will be valid when $S \gg 1$ and $S \gg \langle a^\dagger a \rangle$.

- One forgets that the condition $a_i^\dagger a_i \leq 2S$ should be fulfilled.

- We only keep terms linear in S in the Hamiltonian. In particular we can assume $S_i^+ = \sqrt{2S}a_i$ and $S_i^- = \sqrt{2S}a_i^\dagger$.

Under these approximations, the effective Hamiltonian simply becomes

$$H = E_0 + \sum_{\mathbf{k}} \epsilon_{\mathbf{k}} a_{\mathbf{k}}^\dagger a_{\mathbf{k}} \text{ with } \epsilon_{\mathbf{k}} = JS \sum_{\tau} (1 - e^{i\mathbf{k}\cdot\tau}). \quad (3.10)$$

Let us now calculate the Hamiltonian:

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

Let us define the Fourier transforms as:

$$\begin{aligned} a_{\mathbf{k}}^\dagger &= \frac{1}{\sqrt{N}} \sum_i e^{i\mathbf{k}\cdot\mathbf{r}_i} a_i^\dagger, & a_{\mathbf{k}} &= \frac{1}{\sqrt{N}} \sum_i e^{-i\mathbf{k}\cdot\mathbf{r}_i} a_i \\ a_i^\dagger &= \frac{1}{\sqrt{N}} \sum_{\mathbf{k}} e^{-i\mathbf{k}\cdot\mathbf{r}_i} a_{\mathbf{k}}^\dagger, & a_i &= \frac{1}{\sqrt{N}} \sum_{\mathbf{k}} e^{i\mathbf{k}\cdot\mathbf{r}_i} a_{\mathbf{k}} \end{aligned}$$

The Hamiltonian can be written:

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

Interpretation: We can represent our states $|F\rangle$ and $|\mathbf{k}\rangle$ in terms of their bosonic occupations.

$|0, \dots, 0\rangle \rightarrow$ Ground state

$|0, \dots, \underbrace{1}_{\mathbf{k}}, \dots, 0\rangle \rightarrow$ one magnon state

$|0, \dots, \underbrace{1}_{\mathbf{k}_1}, \dots, 0, \dots, \underbrace{1}_{\mathbf{k}_2}, \dots, 0\rangle \rightarrow$ two-magnon state...

Thermodynamics of spin waves: Due to the simple nature of the Hamiltonian (3.10), we can straightforwardly study it at finite temperature. Of course our results will only be valid at low enough temperature. We remind that for a boson of energy ε at inverse temperature $\beta = \frac{1}{k_B T}$, the occupation number is

$$\langle n \rangle = \frac{1}{e^{\beta\varepsilon} - 1}. \quad (3.11)$$

The magnetization is given by:

$$\begin{aligned} \langle m \rangle &= \frac{1}{N} \sum_i \langle S_i^z \rangle = S - \frac{1}{N} \sum_i \langle a_i^\dagger a_i \rangle \\ &= S - \frac{1}{N} \sum_{\mathbf{k}} \langle a_{\mathbf{k}}^\dagger a_{\mathbf{k}} \rangle \\ &= S - \frac{1}{N} \sum_{\mathbf{k}} \frac{1}{e^{\beta\varepsilon_{\mathbf{k}}} - 1} \\ &= S - \frac{1}{(2\pi)^D} \int \frac{d\mathbf{k}}{e^{\beta\varepsilon_{\mathbf{k}}} - 1}. \end{aligned}$$

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

$$\langle m \rangle = S - \underbrace{\frac{a^D}{(2\pi)^D} \int_{|\mathbf{k}| \leq k_0} \frac{d\mathbf{k}}{e^{\beta JSa^2k^2} - 1}}_I - \frac{a^D}{(2\pi)^D} \int_{|\mathbf{k}| > k_0} \frac{d\mathbf{k}}{e^{\beta\varepsilon_{\mathbf{k}}} - 1}$$

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

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

- **3D:** This is a convergent integral because $d\mathbf{k} \propto k^2 dk$. 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:

$$\int \frac{k^2 dk}{e^{\beta JSa^2k^2} - 1} = \frac{1}{\beta^{3/2}} \int \frac{u^2 du}{e^{JSa^2u^2} - 1} \propto T^{3/2} \quad (3.12)$$

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

- **2D:**

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

- 1D:

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

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.

3.4 Antiferromagnetic case

Let us now consider the Hamiltonian

$$H = J \sum_{\langle i,j \rangle} \mathbf{S}_i \cdot \mathbf{S}_j, \quad J > 0.$$

We will limit ourselves in this section to bipartite lattice. In a bipartite lattice, we can split sites into two independent sublattices A and B , such that A sites (resp. B) are never directly coupled. A cubic or a honeycomb lattice with nearest-neighbour couplings are examples of bipartite lattices/models. The triangular lattice is not.

3.4.1 Classical case

The classical ground state is the Néel state with all spins \uparrow on sublattice A and all spins \downarrow on sublattice B :

$$\begin{array}{c} \uparrow \downarrow \uparrow \downarrow \\ \downarrow \uparrow \downarrow \uparrow \end{array}$$

3.4.2 Quantum formulation

By analogy with the ferromagnetic case, we can consider the state

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

Is this the ground state? Certainly not: it is not even an eigenstate (and not an eigenstate of $\mathbf{S}_{\text{tot}}^2$)!

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

$$\begin{cases} S_i^+ S_j^- |m^A = S, m^B = -S \rangle = 0 \\ S_i^- S_j^+ |m^A = S, m^B = -S \rangle = 2\sqrt{2S} |m^A = S - 1, m^B = -S + 1 \rangle \\ S_i^z S_j^z |m^A = S, m^B = -S \rangle = -S^2 \end{cases}$$

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

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

Sublattice B

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

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

$$\boxed{-JS^2 \frac{N}{2} z + \frac{JS}{2} \sum_{\tau} \sum_{\mathbf{k}} \left(2a_{\mathbf{k}}^\dagger a_{\mathbf{k}} + a_{\mathbf{k}} a_{-\mathbf{k}} e^{-i\mathbf{k}\tau} + a_{\mathbf{k}}^\dagger a_{-\mathbf{k}}^\dagger e^{i\mathbf{k}\tau} \right)}. \quad (3.13)$$

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

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

Problem: The Hamiltonian is not the sum of harmonic oscillators because it contains terms of the form $a_{\mathbf{k}} a_{-\mathbf{k}}$ and $a_{\mathbf{k}}^\dagger a_{-\mathbf{k}}^\dagger$. In the next section, we will see how to solve generically this family of models. As they remain bilinears in terms of bosonic operators, these models are still *Gaussian* and therefore can be solved.

3.4.3 Bogoliubov transformation

The systematic way to solve the Hamiltonian (3.13) is called the Bogoliubov(-de-Gennes) transformation. The core idea is to work with new bosonic operators $\alpha_{\mathbf{k}} = u_{\mathbf{k}}a_{\mathbf{k}} + v_{\mathbf{k}}a_{-\mathbf{k}}^{\dagger}$. A boson $\alpha_{\mathbf{k}}$ is a superposition of a boson a propagating at momentum \mathbf{k} and a hole at momentum $-\mathbf{k}$. The Hamiltonian then takes the simple form

$$H = -JS(S+1)\frac{zN}{2} + \sum_{\mathbf{k}} \omega_{\mathbf{k}} \left(\alpha_{\mathbf{k}}^{\dagger} \alpha_{\mathbf{k}} + \frac{1}{2} \right), \quad (3.14)$$

where

$$\begin{aligned} \omega_{\mathbf{k}} &= JSz\sqrt{1 - \gamma_{\mathbf{k}}^2} \\ \gamma_{\mathbf{k}} &= \frac{1}{z} \sum_{\tau} \cos \mathbf{k} \cdot \tau. \\ u_{\mathbf{k}} &= \sqrt{\frac{JSz + \omega_{\mathbf{k}}}{2\omega_{\mathbf{k}}}} \quad v_{\mathbf{k}} = \text{sign}(\gamma_{\mathbf{k}}) \sqrt{\frac{JSz - \omega_{\mathbf{k}}}{2\omega_{\mathbf{k}}}} \end{aligned}$$

For small k , $\omega_{\mathbf{k}}$ is linear in k . In an antiferromagnet, the dispersion of spin waves is *linear* in k , and not quadratic. The groundstate of the Hamiltonian (3.14) is simply the empty state for the bosons $\alpha_{\mathbf{k}}$, i.e. the state where

$$\langle \alpha_{\mathbf{k}}^{\dagger} \alpha_{\mathbf{k}} \rangle = 0.$$

Importantly, as α 's are superposition of a 's and a^{\dagger} 's, the groundstate will generally have

$$\langle a_{\mathbf{k}}^{\dagger} a_{\mathbf{k}} \rangle \neq 0.$$

Proof: If the Hamiltonian was of the form:

$$H = \sum_{\mathbf{k}} \omega_{\mathbf{k}} \alpha_{\mathbf{k}}^{\dagger} \alpha_{\mathbf{k}} + \text{Cte}$$

one would have:

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

Now, since

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

$$[a_{\mathbf{k}}^{\dagger} a_{-\mathbf{k}}^{\dagger}, a_{\mathbf{k}}] = [a_{\mathbf{k}}^{\dagger}, a_{\mathbf{k}}] a_{-\mathbf{k}}^{\dagger} = -a_{-\mathbf{k}}^{\dagger}.$$

on gets the following expression for the commutator:

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

Similarly, since

$$\begin{aligned} [a_{\mathbf{k}}^{\dagger} a_{\mathbf{k}}, a_{\mathbf{k}}^{\dagger}] &= a_{\mathbf{k}}^{\dagger} \underbrace{a_{\mathbf{k}} a_{\mathbf{k}}^{\dagger}}_{1+a_{\mathbf{k}}^{\dagger} a_{\mathbf{k}}} - a_{\mathbf{k}}^{\dagger} a_{\mathbf{k}}^{\dagger} a_{\mathbf{k}} = a_{\mathbf{k}}^{\dagger}, \\ [a_{\mathbf{k}} a_{-\mathbf{k}}, a_{\mathbf{k}}^{\dagger}] &= +a_{-\mathbf{k}} \end{aligned}$$

one gets:

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

Since $[H, a_{\mathbf{k}}]$ and $[H, a_{-\mathbf{k}}^{\dagger}]$ are linear combinations of $a_{\mathbf{k}}$ and $a_{-\mathbf{k}}^{\dagger}$, it must be possible to bring the Hamiltonian into the required form with bosonic operators $\alpha_{\mathbf{k}}^{\dagger}$ and $\alpha_{\mathbf{k}}$ that are linear combinations of $a_{\mathbf{k}}$ et $a_{-\mathbf{k}}^{\dagger}$. We define

$$\alpha_{\mathbf{k}} = u_{\mathbf{k}} a_{\mathbf{k}} + v_{\mathbf{k}} a_{-\mathbf{k}}^{\dagger}$$

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

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

which leads to:

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

where we have introduced the notation:

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

Solutions exist provided:

$$\begin{aligned} (-\omega_{\mathbf{k}} - JSz)(JSz - \omega_{\mathbf{k}}) + (JSz_{\mathbf{k}})^2 &= 0 \\ \Rightarrow \omega_{\mathbf{k}}^2 &= (JSz)^2 - (JSz_{\mathbf{k}})^2 \end{aligned}$$

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

$$u_{\mathbf{k}}^2 - v_{\mathbf{k}}^2 = u_{\mathbf{k}}^2 \frac{2\omega_{\mathbf{k}}}{JSz + \omega_{\mathbf{k}}}.$$

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

$$\omega_{\mathbf{k}} = JSz \sqrt{1 - \gamma_{\mathbf{k}}^2}$$

with $\gamma_{\mathbf{k}} = \frac{z_{\mathbf{k}}}{z} = \cos k_x$ in 1D, $\frac{1}{2}(\cos k_x + \cos k_y)$ in 2D \dots

The eigenvalue equations imply:

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

Besides, the normalisation condition imposes

$$u_{\mathbf{k}}^2 \left(\frac{2\omega_{\mathbf{k}}}{JSz + \omega_{\mathbf{k}}} \right) = 1$$

One can thus choose as solutions

$$u_{\mathbf{k}} = \sqrt{\frac{JSz + \omega_{\mathbf{k}}}{2\omega_{\mathbf{k}}}} \quad v_{\mathbf{k}} = \text{sign}(z_{\mathbf{k}}) \sqrt{\frac{JSz - \omega_{\mathbf{k}}}{2\omega_{\mathbf{k}}}}$$

The Hamiltonian is thus of the form:

$$H = \sum_{\mathbf{k}} \omega_{\mathbf{k}} \alpha_{\mathbf{k}}^\dagger \alpha_{\mathbf{k}} + \text{Cte}$$

To determine the constant, one must invert these relations

$$\begin{cases} \alpha_{\mathbf{k}} = u_{\mathbf{k}} a_{\mathbf{k}} + v_{\mathbf{k}} a_{-\mathbf{k}}^\dagger \\ \alpha_{-\mathbf{k}}^\dagger = u_{\mathbf{k}} a_{-\mathbf{k}}^\dagger + v_{\mathbf{k}} a_{\mathbf{k}} \end{cases}$$

which leads to:

$$\begin{cases} a_{\mathbf{k}} = u_{\mathbf{k}} \alpha_{\mathbf{k}} - v_{\mathbf{k}} \alpha_{-\mathbf{k}}^\dagger & a_{\mathbf{k}}^\dagger = u_{\mathbf{k}} \alpha_{\mathbf{k}}^\dagger - v_{\mathbf{k}} \alpha_{-\mathbf{k}} \\ a_{-\mathbf{k}}^\dagger = -v_{\mathbf{k}} \alpha_{\mathbf{k}} + u_{\mathbf{k}} \alpha_{-\mathbf{k}}^\dagger & a_{-\mathbf{k}} = -v_{\mathbf{k}} \alpha_{\mathbf{k}}^\dagger + u_{\mathbf{k}} \alpha_{-\mathbf{k}} \end{cases}$$

and introduce these expressions into the Hamiltonian written in terms of the $a_{\mathbf{k}}^\dagger$, $a_{\mathbf{k}}$ given by:

$$H = -JS^2 \frac{zN}{2} + \frac{JS}{2} \sum_{\mathbf{k}} \left(2za_{\mathbf{k}}^\dagger a_{\mathbf{k}} + z_{\mathbf{k}} a_{\mathbf{k}} a_{-\mathbf{k}} + z_{\mathbf{k}} a_{\mathbf{k}}^\dagger a_{-\mathbf{k}}^\dagger \right)$$

Now,

$$\begin{aligned}
a_{\mathbf{k}}^\dagger a_{\mathbf{k}} &= (u_{\mathbf{k}}\alpha_{\mathbf{k}}^\dagger - v_{\mathbf{k}}\alpha_{-\mathbf{k}}) (u_{\mathbf{k}}\alpha_{\mathbf{k}} - v_{\mathbf{k}}\alpha_{-\mathbf{k}}^\dagger) \\
&= u_{\mathbf{k}}^2 \alpha_{\mathbf{k}}^\dagger \alpha_{\mathbf{k}} + v_{\mathbf{k}}^2 \alpha_{-\mathbf{k}} \alpha_{-\mathbf{k}}^\dagger - u_{\mathbf{k}}v_{\mathbf{k}} (\alpha_{\mathbf{k}}^\dagger \alpha_{-\mathbf{k}}^\dagger + \alpha_{\mathbf{k}} \alpha_{-\mathbf{k}}) \\
&= u_{\mathbf{k}}^2 \alpha_{\mathbf{k}}^\dagger \alpha_{\mathbf{k}} + v_{\mathbf{k}}^2 \alpha_{-\mathbf{k}}^\dagger \alpha_{-\mathbf{k}} + v_{\mathbf{k}}^2 - u_{\mathbf{k}}v_{\mathbf{k}} (\alpha_{\mathbf{k}}^\dagger \alpha_{-\mathbf{k}}^\dagger + \alpha_{\mathbf{k}} \alpha_{-\mathbf{k}})
\end{aligned}$$

$$\begin{aligned}
a_{\mathbf{k}} a_{-\mathbf{k}} &= (u_{\mathbf{k}}\alpha_{\mathbf{k}} - v_{\mathbf{k}}\alpha_{-\mathbf{k}}^\dagger) (-v_{\mathbf{k}}\alpha_{\mathbf{k}}^\dagger + u_{\mathbf{k}}\alpha_{-\mathbf{k}}) \\
&= -u_{\mathbf{k}}v_{\mathbf{k}} (\alpha_{\mathbf{k}}\alpha_{\mathbf{k}}^\dagger + \alpha_{-\mathbf{k}}^\dagger\alpha_{-\mathbf{k}}) + u_{\mathbf{k}}^2 \alpha_{\mathbf{k}}\alpha_{-\mathbf{k}} + v_{\mathbf{k}}^2 \alpha_{-\mathbf{k}}^\dagger\alpha_{\mathbf{k}}^\dagger \\
&= -u_{\mathbf{k}}v_{\mathbf{k}} (\alpha_{\mathbf{k}}^\dagger\alpha_{\mathbf{k}} + \alpha_{-\mathbf{k}}^\dagger\alpha_{-\mathbf{k}}) - u_{\mathbf{k}}v_{\mathbf{k}} + u_{\mathbf{k}}^2 \alpha_{\mathbf{k}}\alpha_{-\mathbf{k}} + v_{\mathbf{k}}^2 \alpha_{-\mathbf{k}}^\dagger\alpha_{\mathbf{k}}^\dagger
\end{aligned}$$

$$\begin{aligned}
a_{\mathbf{k}}^\dagger a_{-\mathbf{k}}^\dagger &= (u_{\mathbf{k}}\alpha_{\mathbf{k}}^\dagger - v_{\mathbf{k}}\alpha_{-\mathbf{k}}) (-v_{\mathbf{k}}\alpha_{\mathbf{k}} + u_{\mathbf{k}}\alpha_{-\mathbf{k}}^\dagger) \\
&= -u_{\mathbf{k}}v_{\mathbf{k}} (\alpha_{\mathbf{k}}^\dagger\alpha_{\mathbf{k}} + \alpha_{-\mathbf{k}}\alpha_{-\mathbf{k}}^\dagger) + u_{\mathbf{k}}^2 \alpha_{\mathbf{k}}^\dagger\alpha_{-\mathbf{k}}^\dagger + v_{\mathbf{k}}^2 \alpha_{-\mathbf{k}}\alpha_{\mathbf{k}} \\
&= -u_{\mathbf{k}}v_{\mathbf{k}} (\alpha_{\mathbf{k}}^\dagger\alpha_{\mathbf{k}} + \alpha_{-\mathbf{k}}^\dagger\alpha_{-\mathbf{k}}) - u_{\mathbf{k}}v_{\mathbf{k}} + u_{\mathbf{k}}^2 \alpha_{\mathbf{k}}^\dagger\alpha_{-\mathbf{k}}^\dagger + v_{\mathbf{k}}^2 \alpha_{-\mathbf{k}}\alpha_{\mathbf{k}}
\end{aligned}$$

Let us regroup the terms:

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

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

and so:

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

The constant term of the Hamiltonian is thus equal to:

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

The non-constant terms are given by:

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

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

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

This leads to our final result:

$$H = -JS(S+1)\frac{zN}{2} + \sum_{\mathbf{k}} \left(\alpha_{\mathbf{k}}^\dagger \alpha_{\mathbf{k}} + \frac{1}{2} \right) \omega_{\mathbf{k}}$$

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

$$M_{\text{alt}} = \frac{1}{N} \left\langle \sum_{i \in A} S_z^i - \sum_{i \in B} S_z^j \right\rangle$$

According to the Holstein-Primakoff transformation, we have:

$$M_{\text{alt}} = S - \frac{1}{N} \sum_{\mathbf{k}} \langle a_{\mathbf{k}}^\dagger a_{\mathbf{k}} \rangle$$

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

$$\begin{aligned}
a_{\mathbf{k}}^\dagger a_{\mathbf{k}} &= u_{\mathbf{k}}^2 \alpha_{\mathbf{k}}^\dagger \alpha_{\mathbf{k}} + v_{\mathbf{k}}^2 (1 + \alpha_{\mathbf{k}}^\dagger \alpha_{\mathbf{k}}) \\
&= (u_{\mathbf{k}}^2 + v_{\mathbf{k}}^2) \alpha_{\mathbf{k}}^\dagger \alpha_{\mathbf{k}} + v_{\mathbf{k}}^2
\end{aligned}$$

- **T = 0:** we compute expectation values in the groundstate of Eq. (3.14), i.e., where $\alpha_{\mathbf{k}}^\dagger \alpha_{\mathbf{k}} >$.

$$\langle a_{\mathbf{k}}^\dagger a_{\mathbf{k}} \rangle = v_{\mathbf{k}}^2 = \frac{JSz - \omega_{\mathbf{k}}}{2\omega_{\mathbf{k}}} \neq 0 !$$

These are the so-called *quantum fluctuations*: the groundstate is not the classical groundstate, but has non-zero vacuum fluctuations.

– $\mathbf{D} = \mathbf{1}$

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

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

– $\mathbf{D} \geq \mathbf{2}$ 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 .

• $\mathbf{T} \neq \mathbf{0}$:

$$\langle a_{\mathbf{k}}^\dagger a_{\mathbf{k}} \rangle = \frac{JSz}{\omega_{\mathbf{k}}} \underbrace{\frac{1}{e^{\beta\omega_{\mathbf{k}}} - 1}}_{\sim \frac{1}{k^2} \text{ for } k \text{ small}} + \frac{JSz - \omega_{\mathbf{k}}}{2\omega_{\mathbf{k}}}$$

The integral diverges if $D \leq 2 \rightarrow$ no long-range antiferromagnetic order at $T > 0$ (Hohenberg-Mermin-Wagner theorem).

Spin-wave ground state: The vacuum for the a^\dagger bosons is the Néel state. This is not the ground state! The ground state satisfies $\alpha_{\mathbf{k}}|AF\rangle = 0$. Indeed, denoting $|0_a\rangle$ the empty state for the bosons a :

$$(u_{\mathbf{k}}a_{\mathbf{k}} + v_{\mathbf{k}}a_{-\mathbf{k}}^\dagger)|0_a\rangle = v_{\mathbf{k}}a_{-\mathbf{k}}^\dagger|0_a\rangle \neq 0.$$

One can show that the ground state is given by:

$$|AF\rangle = |0_\alpha\rangle = \prod_{\mathbf{k}}' \frac{1}{u_{\mathbf{k}}} \exp\left(-\frac{v_{\mathbf{k}}}{u_{\mathbf{k}}} a_{\mathbf{k}}^\dagger a_{-\mathbf{k}}^\dagger\right) |0_a\rangle \quad (3.15)$$

where $\prod_{\mathbf{k}}'$ means that the product is restricted to wave vectors \mathbf{k} such that $\mathbf{k} \neq -\mathbf{k}$. Note that this groundstate is technically not unique, as $\omega_{\mathbf{0}} = \omega_{\boldsymbol{\pi}} = 0$ (where $\mathbf{k} = \boldsymbol{\pi}$ is a short hand for the vector (π, π, \dots)). This technically means that we can also chose an arbitrary value for the values of $a_{\mathbf{0}}$ or $a_{\boldsymbol{\pi}}$. We also have not carefully reviewed the occupation numbers of other symmetric points in the Brillouin zone — this is left as an exercise.

To this end, one much show that:

$$\alpha_{\mathbf{k}}|AF\rangle = 0$$

with $\alpha_{\mathbf{k}} = u_{\mathbf{k}}a_{\mathbf{k}} + v_{\mathbf{k}}a_{-\mathbf{k}}^{\dagger}$.

$$\begin{aligned} a_{\mathbf{k}}|AF\rangle &= a_{\mathbf{k}}\prod_1' \frac{1}{u_1} \exp\left(-\frac{v_1}{u_1}a_1^{\dagger}a_{-1}^{\dagger}\right)|0_a\rangle \\ &= \prod_{\mathbf{l}\neq\mathbf{k},-\mathbf{k}}' \frac{1}{u_1} \exp\left(-\frac{v_1}{u_1}a_1^{\dagger}a_{-1}^{\dagger}\right) \frac{1}{u_{\mathbf{k}}} a_{\mathbf{k}} \sum_{n=1}^{+\infty} \frac{\left[-\frac{v_{\mathbf{k}}}{u_{\mathbf{k}}}a_{\mathbf{k}}^{\dagger}a_{-\mathbf{k}}^{\dagger}\right]^n}{n!} |0_a\rangle \end{aligned}$$

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

$$[a_{\mathbf{k}}, (a_{\mathbf{k}}^{\dagger})^n] = n(a_{\mathbf{k}}^{\dagger})^{n-1}$$

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

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

Thus,

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

and one gets:

$$\begin{aligned} a_{\mathbf{k}}|AF\rangle &= \prod_{\mathbf{l}\neq\mathbf{k},-\mathbf{k}}' \frac{1}{u_1} \exp\left(-\frac{v_1}{u_1}a_1^{\dagger}a_{-1}^{\dagger}\right) \frac{1}{u_{\mathbf{k}}} \sum_{n=1}^{+\infty} \frac{\left(-\frac{v_{\mathbf{k}}}{u_{\mathbf{k}}}\right)^n (a_{\mathbf{k}}^{\dagger})^{n-1} (a_{-\mathbf{k}}^{\dagger})^n}{(n-1)!} |0_a\rangle \\ &= \prod_{\mathbf{l}\neq\mathbf{k},-\mathbf{k}}' \frac{1}{u_1} \exp\left(-\frac{v_1}{u_1}a_1^{\dagger}a_{-1}^{\dagger}\right) \frac{1}{u_{\mathbf{k}}} \frac{-v_{\mathbf{k}}}{u_{\mathbf{k}}} a_{-\mathbf{k}}^{\dagger} \exp\left(-\frac{v_{\mathbf{k}}}{u_{\mathbf{k}}}a_{\mathbf{k}}^{\dagger}a_{-\mathbf{k}}^{\dagger}\right) |0_a\rangle \end{aligned}$$

Finally:

$$u_{\mathbf{k}}a_{\mathbf{k}}|AF\rangle = -v_{\mathbf{k}}a_{-\mathbf{k}}^{\dagger}|AF\rangle$$

so that

$$\alpha_{\mathbf{k}}|AF\rangle = 0.$$

Finally, let us check the normalization of $|AF\rangle$.

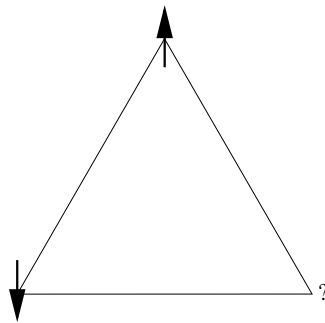
$$\begin{aligned} \langle AF|AF\rangle &= \langle 0_a | \prod_{\mathbf{k}} \frac{1}{u_{\mathbf{k}}^2} \sum_n \left(-\frac{v_{\mathbf{k}}}{u_{\mathbf{k}}}\right)^{2n} \frac{(a_{-\mathbf{k}}a_{\mathbf{k}})^n (a_{\mathbf{k}}^{\dagger}a_{-\mathbf{k}}^{\dagger})^n}{n!^2} |0_a\rangle \\ &= \prod_{\mathbf{k}} \frac{1}{u_{\mathbf{k}}^2} \sum_n \left(-\frac{v_{\mathbf{k}}}{u_{\mathbf{k}}}\right)^{2n} \\ &= \prod_{\mathbf{k}} \frac{1}{u_{\mathbf{k}}^2} \frac{1}{1 - \left(\frac{v_{\mathbf{k}}}{u_{\mathbf{k}}}\right)^2} \\ &= 1. \end{aligned}$$

3.5 Frustrated magnets

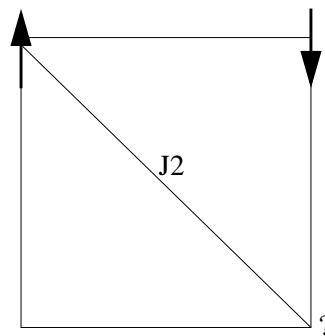
Definition: Frustrated magnets are magnets for which it is impossible to minimise simultaneously the energy of each bond in the classical case.

Examples:

- Triangular lattice \rightarrow topological origin



- Square lattice with first and second neighbour interactions



3.5.1 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

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

There is a factor $\frac{1}{2}$ because each bond appears twice in this formulation. Importantly, the sum over \mathbf{R}_n is a sum over all sites: the restriction in the form of the couplings is entirely contained in the function $J_{\mathbf{R}_n}$. We look for the minimum under the constraint $\|\mathbf{S}_{\mathbf{R}_i}\|^2 = 1$ for all sites.

Fourier transform:

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

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

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

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

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

One then proceeds in two steps:

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

Now

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

One must thus minimise

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

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

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

In real space, the spins are then given by

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

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

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

but also such that the local constraints:

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

are satisfied for all i ? Yes! For example:

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

leads to:

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

One get a helical structure of pitch vector \mathbf{k}_0 .

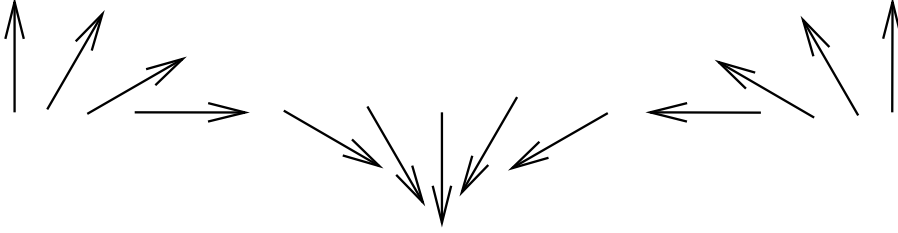


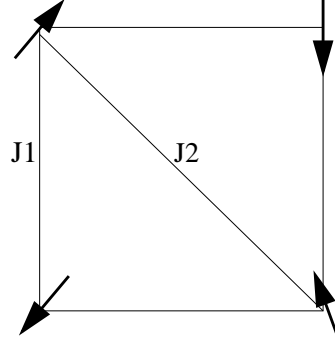
Figure 3.3: Example of a 1D helix of wave vector $\mathbf{k}_0 = \frac{2\pi}{12a}$.

Conclusions:

- The structure of the classical ground state is a helix whose wave vector is given by the minimum of $J_{\mathbf{k}}$.
- If there are several wave vectors that minimise $J_{\mathbf{k}}$, one can often make linear combinations of the solutions, and therefore there is a continuous (infinite) degeneracy of the ground state.

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

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



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

The minimization leads to:

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

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

or

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

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

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

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

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

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

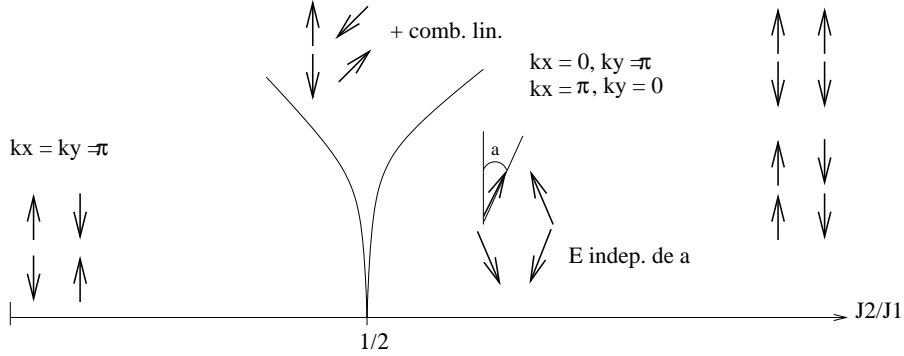
$$\begin{aligned}
\rightarrow J(\mathbf{k}) &= 2J_1 \left(\frac{-2J_1}{2J_2} \right) + 4J_2 \frac{J_1^2}{4J_2^2} \\
&= \frac{-2J_1^2}{J_2} + \frac{J_1^2}{J_2} = -\frac{J_1^2}{J_2}
\end{aligned}$$

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 $J(\mathbf{k}) = -2J_1$ 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.



3.5.2 Spin waves in a helical system

Let us consider a system described by the Hamiltonian

$$H = \frac{1}{2} \sum_i \sum_{\mathbf{R}_n} J_{\mathbf{R}_n} \mathbf{S}_i \cdot \mathbf{S}_{i+\mathbf{R}_n}$$

$$J(\mathbf{k}) = \sum_{\mathbf{R}_n} J_{\mathbf{R}_n} e^{-i\mathbf{k} \cdot \mathbf{R}_n}$$

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

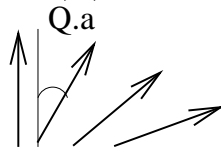
We are going to perform a rotation in spin space such that the classical ground state is described by $S_i^u = S_i^v = 0, S_i^w = 1$.

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

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

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

$$\begin{cases} S^u = S^x \\ S^v = \cos \theta_i S^y - \sin \theta_i S^z \\ S^w = \sin \theta_i S^y + \cos \theta_i S^z. \end{cases}$$



In the new basis, the Hamiltonian takes the form

$$H = \frac{1}{2} \sum_i \sum_{\mathbf{R}_n} J_{\mathbf{R}_n} \left[S_i^u S_{i+n}^u + \cos(\theta_i - \theta_{i+n}) (S_i^v S_{i+n}^v + S_i^w S_{i+n}^w) + \sin(\theta_i - \theta_{i+n}) (S_i^w S_{i+n}^v - S_i^v S_{i+n}^w) \right] \quad (3.16)$$

Proof:

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

since

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

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

Next, we perform the Holstein-Primakoff transformation:

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

The Hamiltonian now takes the form:

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

At this point, we have not specified the transformation θ_i . It needs to minimize the classical energy

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

which get rid of terms linear in a and a^\dagger . We obtain the final Hamiltonian

$$\begin{aligned}
H &= \frac{NS^2}{2} J_{\mathbf{Q}} + \frac{S}{2} \sum_{\mathbf{k}} \left[J_{\mathbf{k}} + \frac{1}{2} (J_{\mathbf{k}+\mathbf{Q}} + J_{\mathbf{k}-\mathbf{Q}}) - 2J_{\mathbf{Q}} \right] a_{\mathbf{k}}^\dagger a_{\mathbf{k}} \\
&+ \frac{S}{4} \sum_{\mathbf{k}} \left[J_{\mathbf{k}} - \frac{1}{2} (J_{\mathbf{k}+\mathbf{Q}} + J_{\mathbf{k}-\mathbf{Q}}) \right] (a_{\mathbf{k}} a_{-\mathbf{k}} + a_{\mathbf{k}}^\dagger a_{-\mathbf{k}}^\dagger)
\end{aligned} \tag{3.17}$$

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

$$\begin{cases} \omega_{\mathbf{k}} = S\sqrt{(J_{\mathbf{k}} - J_{\mathbf{Q}}) \left(\frac{1}{2} (J_{\mathbf{k}+\mathbf{Q}} + J_{\mathbf{k}-\mathbf{Q}}) - J_{\mathbf{Q}} \right)} \\ \omega_{\mathbf{k}} = 0 \quad \text{si} \quad \begin{cases} \mathbf{k} = 0 \\ \mathbf{k} = \pm\mathbf{Q} \end{cases} \end{cases} \tag{3.18}$$

Detailed computations:

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

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

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

must be minimal. If one groups the terms containing a given angle θ_i , one gets

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

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

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

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

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

Let us look at the terms:

$$\bullet e^{i\mathbf{R}_n \cdot (\mathbf{k}_1 + \mathbf{k}_3)} e^{i\mathbf{r}_i \cdot (\mathbf{k}_2 + \mathbf{k}_3)} \quad \mathbf{k}_1 = -\mathbf{k}_3, \mathbf{k}_2 = -\mathbf{k}_3$$

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

$$\bullet e^{i\mathbf{R}_n \cdot (\mathbf{k}_1 + \mathbf{k}_3 + \mathbf{Q})} e^{i\mathbf{r}_i \cdot (\mathbf{k}_2 + \mathbf{k}_3)} \quad \mathbf{k}_3 = -\mathbf{k}_1 - \mathbf{Q}, \mathbf{k}_2 = \mathbf{k}_1 + \mathbf{Q}$$

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

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

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

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

$$\rightarrow \frac{1}{2} \sum_{\mathbf{k}} J_{\mathbf{k}} \frac{S}{2} a_{\mathbf{k}}^\dagger a_{-\mathbf{k}}^\dagger$$

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

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

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

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

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

$$\rightarrow \frac{1}{2} \sum_{\mathbf{k}} J_{-\mathbf{k}} \frac{S}{2} a_{\mathbf{k}}^\dagger a_{\mathbf{k}}$$

- and all the other terms:

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

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

$$\rightarrow \frac{1}{2} \sum_{\mathbf{k}} J_{\mathbf{k}} \frac{S}{2} a_{\mathbf{k}}^\dagger a_{\mathbf{k}}$$

$$\rightarrow \frac{1}{4} \sum_{\mathbf{k}} J_{\mathbf{k}+\mathbf{Q}} \frac{S}{2} a_{\mathbf{k}}^\dagger a_{\mathbf{k}}$$

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

$$\rightarrow \frac{1}{2} \sum_i \sum_{\mathbf{R}_n} S^2 J_{\mathbf{R}_n} \frac{e^{i\mathbf{Q}\mathbf{R}_n} + e^{-i\mathbf{Q}\mathbf{R}_n}}{2} = \frac{S^2}{2} N \frac{1}{2} (J_{\mathbf{Q}} + J_{-\mathbf{Q}})$$

$$\rightarrow -\frac{S}{4} J_{\mathbf{Q}} a_{\mathbf{k}}^\dagger a_{\mathbf{k}}$$

$$\rightarrow -\frac{S}{4} J_{-\mathbf{Q}} a_{\mathbf{k}}^\dagger a_{\mathbf{k}}$$

$$\rightarrow -\frac{S}{4} J_{-\mathbf{Q}} a_{\mathbf{k}}^\dagger a_{\mathbf{k}}$$

$$\rightarrow -\frac{S}{4} J_{\mathbf{Q}} a_{\mathbf{k}}^\dagger a_{\mathbf{k}}$$

Let us group the terms:

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

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

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

3.5.3 Bonus: $J_1 - J_2$ 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.

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

The collinear structures are selected:

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

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

$$\langle m \rangle = S - \underbrace{\int \frac{d\mathbf{k} f(\mathbf{k})}{\omega_{\mathbf{k}}}}_{\Delta \langle m \rangle}$$

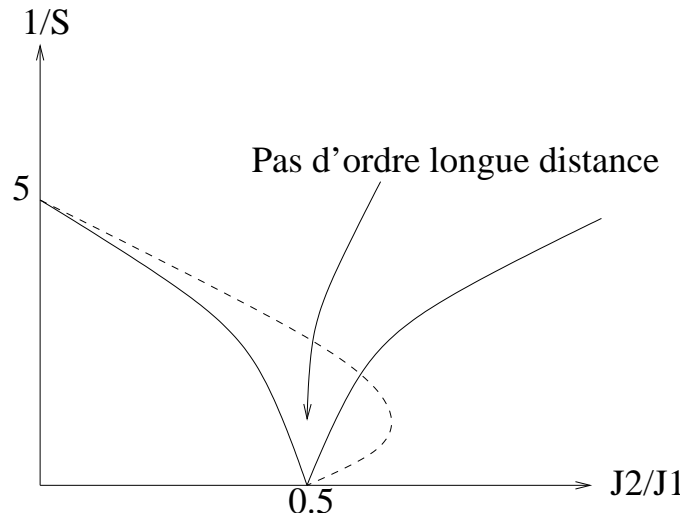
$f(\mathbf{k}) \rightarrow \text{cst}$ when $\mathbf{k} \rightarrow \mathbf{0}$.

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

$$\begin{aligned}\omega_{\mathbf{k}} &= \left\{ [SJ_1z + SJ_2z(\beta_{\mathbf{k}} - 1)]^2 - (SJ_1z\gamma_{\mathbf{k}})^2 \right\}^{\frac{1}{2}} \\ \gamma_{\mathbf{k}} &= \frac{1}{z} \sum_{\tau_1} e^{i\mathbf{k}\cdot\tau_1} \quad , \quad \beta_{\mathbf{k}} = \frac{1}{z} \sum_{\tau_2} e^{i\mathbf{k}\cdot\tau_2}\end{aligned}$$

Expanding at small k leads to $\omega_{\mathbf{k}} \sim ck$ with $c \rightarrow 0$ when $J_2 \rightarrow \frac{J_1}{2}$.

→ 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 $\ln S$:

→ $\langle m \rangle = S - \text{cst} \times \ln S > 0$ 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:

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

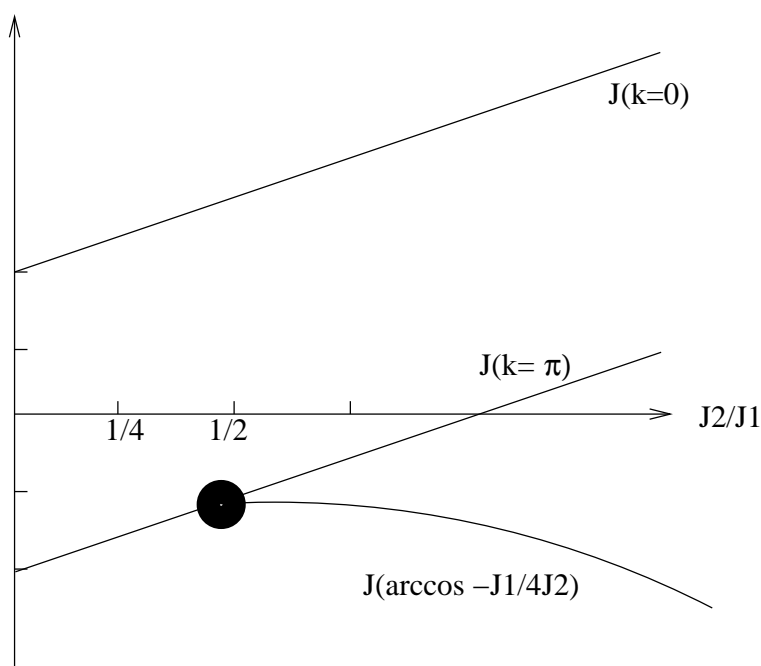
$$1. \quad k = 0 \quad \Rightarrow \quad J(k) = 2J_1 + 2J_2$$

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

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

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

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



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

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

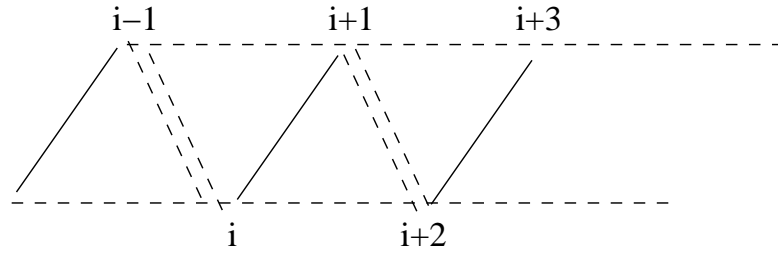
with ($J_1 = 2J_2$, $J_2 = J$) and ($S = \frac{1}{2}$).

The Hamiltonian can be rewritten as:

$$H = J (2\mathbf{S}_1 \cdot \mathbf{S}_2 + (\mathbf{S}_1 + \mathbf{S}_2) \cdot \mathbf{S}_3 + \mathbf{S}_2 \cdot (\mathbf{S}_3 + \mathbf{S}_4) + \dots)$$

or

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



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

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

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

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

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

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

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

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

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

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

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

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

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

So, $|\phi\rangle$ 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 $h_i = J(\mathbf{S}_i\mathbf{S}_{i+1} + \mathbf{S}_{i+1}\mathbf{S}_{i+2} + \mathbf{S}_i\mathbf{S}_{i+2})$. But

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

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

$$\implies (\mathbf{S}_i + \mathbf{S}_{i+1} + \mathbf{S}_{i+2})^2 = \begin{cases} \frac{1}{2} \left(\frac{1}{2} + 1\right) = \frac{3}{4} \\ \text{ou} \frac{3}{2} \left(\frac{3}{2} + 1\right) = \frac{15}{4} \end{cases} \quad (3.19)$$

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

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

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

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

Chapter 4

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 (studied as an exercise, but extremely important), the integer Quantum Hall Effect, and the fractional Quantum Hall Effect. All these effects rely in an essential way on the quantization of the motion of a charged particle in the presence of a uniform magnetic field $\mathbf{B} = B\hat{z}$.

Loic Herviou: I changed the units used in this chapter to the standard SI units (instead of CGS). If you catch a remaining line where this change has not been done, please contact me.

4.1 Landau levels

4.1.1 Classical problem

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

$$\mathbf{F} = q(\mathbf{v} \wedge \mathbf{B}) = qB \begin{pmatrix} v_y \\ -v_x \\ 0 \end{pmatrix}$$

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

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

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

In the plane perpendicular to the field, one gets:

$$\boxed{\begin{cases} \ddot{x} = \omega_c \dot{y} \\ \ddot{y} = -\omega_c \dot{x} \end{cases}} \quad (4.1)$$

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

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

leading to

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

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

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

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

$$\begin{cases} r_x = r_y = r \\ \varphi_x = \varphi_y = \varphi \end{cases} \\ \Rightarrow \begin{cases} x = x_0 + r \sin(\omega_c t + \varphi) \\ y = y_0 + r \cos(\omega_c t + \varphi) \end{cases}$$

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 the *cyclotron frequency*

$$\boxed{\omega_c = \frac{qB}{m}} \quad (4.2)$$

4.1.2 Lagrangian approach

The equations of motion can be rederived from the Lagrangian

$$\mathcal{L} = \mathcal{L}_{xy} + \frac{1}{2} m \dot{z}^2, \quad (4.3)$$

with

$$\mathcal{L}_{xy}(x, y; \dot{x}, \dot{y}) = \frac{1}{2}m(\dot{x}^2 + \dot{y}^2) + q[A_x\dot{x} + A_y\dot{y}], \quad (4.4)$$

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

$$\frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{q}_i} - \frac{\partial \mathcal{L}}{\partial q_i} = 0, \quad (4.5)$$

lead to:

$$\begin{aligned} &\Rightarrow \begin{cases} \frac{d}{dt} (m\dot{x} + qA_x) - q[\dot{x}\partial_x A_x + \dot{y}\partial_x A_y] = 0, \\ \frac{d}{dt} (m\dot{y} + qA_y) - q[\dot{x}\partial_y A_x + \dot{y}\partial_y A_y] = 0. \end{cases} \\ &\Rightarrow \begin{cases} m\ddot{x} + q\partial_t A_x + q(\dot{x}\partial_x A_x + \dot{y}\partial_y A_x) - q(\dot{x}\partial_x A_x + \dot{y}\partial_x A_y) = 0 \\ m\ddot{y} + q\partial_t A_y + q(\dot{x}\partial_x A_y + \dot{y}\partial_y A_y) - q(\dot{x}\partial_y A_x + \dot{y}\partial_y A_y) = 0 \end{cases} \end{aligned}$$

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

$$\Rightarrow \begin{cases} m\ddot{x} = q\dot{y}(\partial_x A_y - \partial_y A_x) \\ m\ddot{y} = q\dot{x}(\partial_y A_x - \partial_x A_y), \end{cases}$$

or

$$\Rightarrow \begin{cases} \ddot{x} = \frac{qB}{m}\dot{y} \\ \ddot{y} = \frac{qB}{m}\dot{x}. \end{cases}$$

4.1.3 Hamiltonian formulation

One must express the velocities in terms of the momenta.

$$\begin{aligned} &\begin{cases} p_x = \frac{\partial \mathcal{L}}{\partial \dot{x}} = m\dot{x} + qA_x \\ p_y = \frac{\partial \mathcal{L}}{\partial \dot{y}} = m\dot{y} + qA_y \end{cases} \Rightarrow \begin{cases} \dot{x} = \frac{1}{m}(p_x - qA_x) \\ \dot{y} = \frac{1}{m}(p_y - qA_y) \end{cases} \\ &\Rightarrow \mathcal{H}_{xy} = \frac{1}{2}m(\dot{x}^2 + \dot{y}^2) = \frac{1}{2m}[(p_x - qA_x)^2 + (p_y - qA_y)^2] \end{aligned}$$

Finally,

$$\mathcal{H} = \frac{1}{2m}(\mathbf{p} - q\mathbf{A})^2. \quad (4.6)$$

4.1.4 Quantum treatment

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

$$[\hat{x}_i, \hat{p}_j] = i\hbar\delta_{ij}. \quad (4.7)$$

We remind that this canonical approach to quantization is to relate non-trivial commutator for quantum variables to non-trivial Poisson bracket. The resulting Hamiltonian is simply

$$\mathcal{H} = \frac{1}{2m} (p_x - qA_x)^2 + \frac{1}{2m} (p_y - qA_y)^2 + \frac{1}{2m} p_z^2. \quad (4.8)$$

Let us define $\Pi_x = p_x - qA_x$ and $\Pi_y = p_y - qA_y$. Since A_x and A_y do not depend on z ,

$$[p_z, \Pi_x] = [p_z, \Pi_y] = 0.$$

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

$$[\Pi_x, \Pi_y] = [p_x - qA_x, p_y - qA_y] = -q[p_x, A_y] - q[A_x, p_y].$$

But

$$\begin{aligned} [p_x, f(x)]\varphi(x) &= -i\hbar\partial_x(f(x)\varphi(x)) + f(x)i\hbar\partial_x\varphi(x) = -i\hbar(\partial_x f(x))\varphi(x), \\ \rightarrow [p_x, f(x)] &= -i\hbar\partial_x f(x). \end{aligned}$$

So

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

leading to

$$[\Pi_x, \Pi_y] = -q(-i\hbar)(\partial_x A_y - \partial_y A_x),$$

or

$$[\Pi_x, \Pi_y] = i\hbar q B. \quad (4.9)$$

From dimensional analysis, it is interesting to introduce the *magnetic length*

$$l_B^2 = \frac{\hbar}{|q|B} \quad (4.10)$$

such that (the negative sign accounts for the negative charges of the electrons)

$$[\Pi_x, \Pi_y] = -i\frac{\hbar^2}{l_B^2}. \quad (4.11)$$

Let us define $\hat{P} = \hat{\Pi}_x$, $\hat{Q} = -\frac{1}{Bq}\Pi_y$. We have $[\hat{Q}, \hat{P}] = i\hbar$ and

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

with

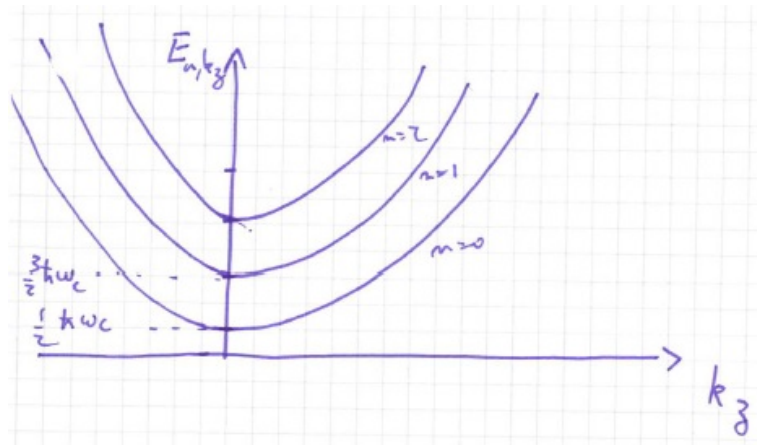
$$\begin{aligned} \frac{1}{2m}\Pi_y^2 &= \frac{1}{2}m\omega^2\frac{1}{q^2B^2}\Pi_y^2, \\ \omega^2 &= \frac{q^2B^2}{m^2}, \text{ soit } \boxed{\omega = \frac{qB}{m}}. \end{aligned} \quad (4.12)$$

With these variables, the Hamiltonian reads:

$$\boxed{\mathcal{H} = \frac{1}{2m}\hat{P}^2 + \frac{1}{2}m\omega_c^2\hat{Q}^2 + \frac{1}{2m}p_z^2.} \quad (4.13)$$

The Hamiltonian can thus be written as the sum of a harmonic oscillator of frequency ω_c and of the Hamiltonian of a free particle in the z direction. The eigenenergies are thus of the form:

$$\boxed{E_{n,k_z} = \hbar\omega_c\left(n + \frac{1}{2}\right) + \frac{\hbar^2k_z^2}{2m}} \quad (4.14)$$



Each of these curves corresponds to a *Landau level*.

4.1.5 Degeneracy

For a given value of k_z , how many levels are there per level? Importantly, we obtained only two quantum numbers for a three-dimensional system! To get a first

rough idea, let us consider the number N of quantum states whose energy is between E et $E + \hbar\omega_c$ in a free electronic model without magnetic field. To calculate this number, we write that it is the volume in \mathbf{k} space of the points such that

$$E \leq \frac{\hbar^2 k^2}{2m} \leq E + \hbar\omega_c \quad (4.15)$$

divided by the elementary volume per \mathbf{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 \mathbf{k} points satisfying (4.15) is a surface defined by:

$$2\pi \int_{k_1}^{k_2} k dk = 2\pi \frac{k_2^2 - k_1^2}{2}, \quad (4.16)$$

with

$$E = \frac{\hbar^2 k_1^2}{2m}, \quad E + \hbar\omega_c = \frac{\hbar^2 k_2^2}{2m}, \quad (4.17)$$

$$\rightarrow 2\pi \times \frac{1}{2} \times \frac{2m}{\hbar^2} \hbar\omega_c = \frac{2\pi}{\hbar} \omega_c m. \quad (4.18)$$

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

$$N = \frac{2\pi}{\hbar} \omega_c m \frac{L_x L_y}{2\pi} = L_x L_y \frac{qB}{2\pi\hbar}. \quad (4.19)$$

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. This is a *macroscopic* number: in the thermodynamic limit, it is infinitely degenerate. 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, we can convince ourselves 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 freedom 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,

$$\begin{cases} Y = y - \frac{1}{m\omega_c} \Pi_x \\ X = x + \frac{1}{m\omega_c} \Pi_y \end{cases} \quad (4.20)$$

Since $[x, \Pi_x] = [y, \Pi_y] = i\hbar$, one gets:

$$\begin{aligned}
[X, Y] &= \left[x + \frac{1}{m\omega_c} \Pi_y, y - \frac{1}{m\omega_c} \Pi_x \right] \\
&= -\frac{1}{m\omega_c} [x, \Pi_x] + \frac{1}{m\omega_c} [\Pi_y, y] - \frac{1}{m^2\omega_c^2} [\Pi_y, \Pi_x] \\
&= -\frac{2}{m\omega_c} i\hbar + \frac{1}{m^2\omega_c^2} i\hbar qB \\
&= -i \frac{\hbar}{qB}.
\end{aligned} \tag{4.21}$$

The elementary volume in phase space associated to these variables is given by $2\pi \frac{\hbar}{qB}$ (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 $L_x L_y$,

$$\text{deg} = \frac{L_x L_y}{2\pi\hbar} qB = \frac{\Phi}{\phi_0}. \tag{4.22}$$

In this expression, Φ is the total magnetic flux going through the sample, and ϕ_0 is the quantum of flux for particles of charge q .

4.1.6 Density of states

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

$$n(E) dE = 2n(k_z) dk_z = 2 \times \text{deg} \times \frac{dk_z}{2\pi/L_z} = 2 \times \frac{L_x L_y}{2\pi} \times \frac{qB L_z}{\hbar} \frac{dk_z}{2\pi}. \tag{4.23}$$

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

$$dE = 2 \frac{\hbar^2 k_z}{2m} dk_z \Rightarrow dk_z = \frac{m}{\hbar^2 k_z} dE \tag{4.24}$$

and

$$k_z^2 = \frac{2m}{\hbar^2} \left(E - \hbar\omega_c(n + \frac{1}{2}) \right), \tag{4.25}$$

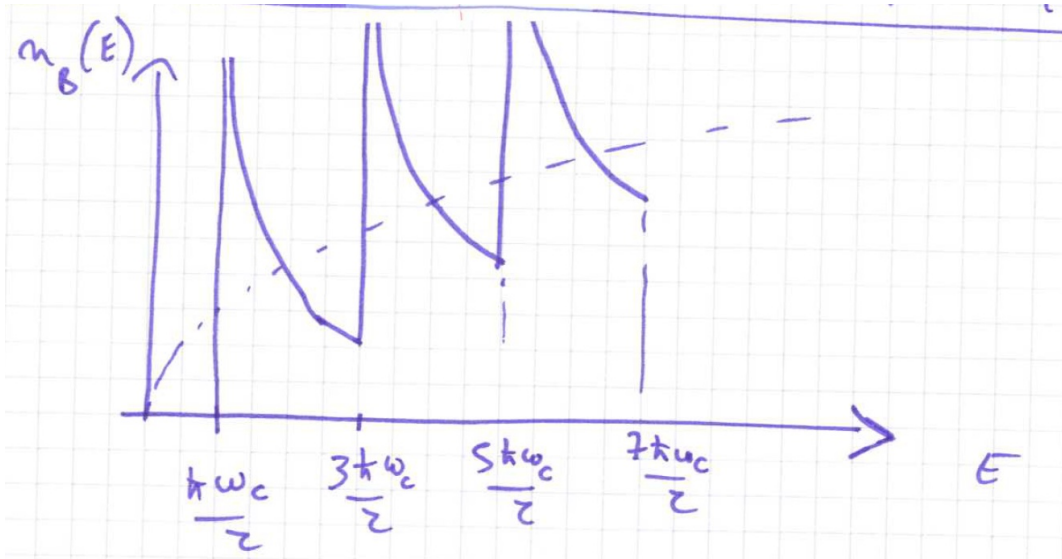
$$dk_z = \frac{m}{\hbar^2} \frac{dE}{\sqrt{\frac{2m}{\hbar^2} \left(E - \hbar\omega_c(n + \frac{1}{2}) \right)}}, \text{ si } E \geq \hbar\omega_c(n + \frac{1}{2}) \tag{4.26}$$

so that

$$\begin{aligned}
n(E) &= \frac{V}{4\pi^2} \frac{qB}{\hbar} \sqrt{\frac{2m}{\hbar^2}} \frac{\Theta \left(E - \hbar\omega_c(n + \frac{1}{2}) \right)}{\sqrt{E - \hbar\omega_c(n + \frac{1}{2})}} \\
&= \frac{V}{8\pi^2} \hbar\omega_c \left(\frac{2m}{\hbar^2} \right)^{3/2} \frac{\Theta \left(E - \hbar\omega_c(n + \frac{1}{2}) \right)}{\sqrt{E - \hbar\omega_c(n + \frac{1}{2})}}
\end{aligned} \tag{4.27}$$

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

$$n(E) = \frac{V}{8\pi^2} \hbar\omega_c \left(\frac{2m}{\hbar^2}\right)^{3/2} \sum_{n=0}^{\infty} \frac{\Theta\left(E - \hbar\omega_c\left(n + \frac{1}{2}\right)\right)}{\sqrt{E - \hbar\omega_c\left(n + \frac{1}{2}\right)}}. \quad (4.28)$$



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

$$n(E) = 2 \frac{V}{4\pi^2} \left(\frac{2m}{\hbar^2}\right)^{3/2} \sqrt{E} dE. \quad (4.29)$$

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

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

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

or

$$\frac{1}{B} = \frac{\hbar e}{m E_F} \left(n + \frac{1}{2}\right). \quad (4.30)$$

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.

4.1.7 Exact solutions of the Landau problem

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 $\text{deg} = L_x L_y \frac{qB}{2\pi\hbar}$. 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**

$$A_x = -By, \quad A_y = 0, \quad A_z = 0. \quad (4.31)$$

In this gauge, the directions x and y are treated differently. It is useful for problems in rectangular geometry. We can of course exchange the roles of x and y :

$$A_x = 0, \quad A_y = Bx, \quad A_z = 0. \quad (4.32)$$

- **Symmetric gauge**

$$A_x = -\frac{By}{2}, \quad A_y = \frac{Bx}{2}, \quad A_z = 0. \quad (4.33)$$

This gauge leads to eigenfunctions that turned out to be essential in the construction of variational wave functions for the fractional Quantum Hall Effect. It is equivalent, in cylinder coordinates, to

$$A_r = 0, \quad A_\theta = \frac{Br}{2}, \quad A_z = 0. \quad (4.34)$$

A case study: the Landau gauge

Let us assume that we are in the Landau gauge. We will now rederive the degeneracy of the groundstate, and compute exactly the eigenfunctions of the Landau Hamiltonian. We will assume an infinite plane, in order to avoid having to deal with boundary conditions. Our choice of gauge is also compatible with periodic boundary conditions in the x and z directions. It will correspond to a quantization of the respective momenta.

$$\begin{aligned} \mathcal{H} &= \frac{1}{2m} (p_x - qA_x)^2 + \frac{1}{2m} p_y^2 + \frac{1}{2m} p_z^2 \\ &= \frac{1}{2m} (p_x + qBy)^2 + \frac{1}{2m} p_y^2 + \frac{1}{2m} p_z^2. \end{aligned} \quad (4.35)$$

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

$$\varphi(x, y, z) = e^{ik_x x} e^{ik_z z} \varphi(y). \quad (4.36)$$

We then need to solve

$$\frac{1}{2m} [(\hbar k_x + qBy)^2] \varphi(y) - \frac{\hbar^2}{2m} \frac{\partial^2 \varphi}{\partial y^2} + \frac{\hbar^2 k_z^2}{2m} \varphi(y) = E\varphi(y) \quad (4.37)$$

Let us define

$$y_{k_x} = -\frac{\hbar k_x}{qB} \text{ et } \omega_c = \frac{qB}{m}. \quad (4.38)$$

We get:

$$\frac{1}{2} m \omega_c^2 (y - y_{k_x})^2 \varphi(y) - \frac{\hbar^2}{2m} \frac{\partial^2 \varphi}{\partial y^2} + \frac{\hbar^2 k_z^2}{2m} \varphi(y) = E\varphi(y). \quad (4.39)$$

This is a shifted harmonic oscillator of frequency ω_c ,

$$E_{n,k_x,k_z} = \hbar \omega_c \left(n + \frac{1}{2} \right) + \frac{\hbar^2 k_z^2}{2m}. \quad (4.40)$$

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

$$0 \leq \hbar p \frac{2\pi}{L_x} \frac{1}{qB} \leq L_y \Rightarrow p \leq \frac{L_x L_y qB}{2\pi \hbar} \quad (4.41)$$

Finally, we can easily deduce the eigenfunctions using the standard formula for the harmonic oscillators:

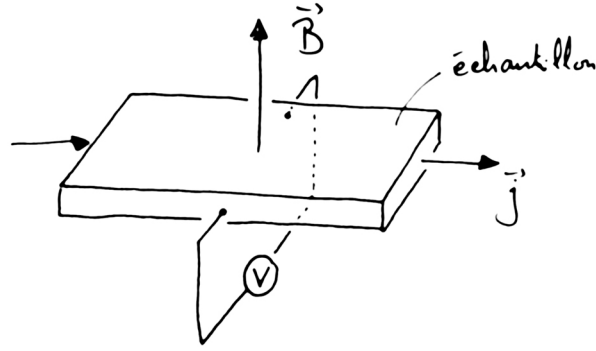
$$\varphi_{n,k_x,k_z}(x, y, z) = e^{ik_x x} e^{ik_z z} \times \frac{1}{\sqrt{2^n n! l_B \sqrt{\pi}}} e^{-\frac{(y-y_{k_x})^2}{2l_B^2}} H_n\left(\frac{y-y_{k_x}}{l_B}\right), \quad (4.42)$$

where $l_B = \hbar/|q|B$ is the magnetic length. The position of the center of the orbital (y_{k_x}) depends on the momentum of the particle in the orthogonal direction. This is the so-called *position-momentum coupling* characteristic of the movement of charged particles in the magnetic field. While the exact form of the wavefunctions depends of the choice of gauge (unlike any observable), this coupling remains true in all gauge. It is one of the core components of the phenomena we will study in the rest of the chapter.

4.2 The integer Quantum Hall Effect

4.2.1 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 $\mathbf{B} = B\hat{z}$, a current in the x direction $\mathbf{j} = j\hat{x}$ induces in a conductor a voltage in the \hat{y} direction (or conversely, a voltage induces a transverse current).



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

$$\frac{d\mathbf{p}}{dt} = -e \left(\mathbf{E} + \frac{\mathbf{p}}{m} \wedge \mathbf{B} \right) - \frac{\mathbf{p}}{\tau} \quad (4.43)$$

where τ is the relaxation time. The stationary solution is given by

$$\frac{d\mathbf{p}}{dt} = 0 \quad (4.44)$$

so that

$$\begin{cases} -e \left(E_x + \frac{p_y}{m} B \right) - \frac{p_x}{\tau} = 0 \\ -e \left(E_y - \frac{p_x}{m} B \right) - \frac{p_y}{\tau} = 0 \end{cases} \quad (4.45)$$

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

$$\begin{cases} \sigma_0 E_x = -ne \frac{p_x}{m} - ne \frac{p_y}{m} \omega_c \tau \\ \sigma_0 E_y = ne \frac{p_x}{m} \omega_c \tau - ne \frac{p_y}{m} \end{cases} \quad (4.46)$$

The current density \mathbf{j} is defined by:

$$\mathbf{j} = -ne \frac{\mathbf{p}}{m} \quad (4.47)$$

This system can be written:

$$\mathbf{E} = \rho \mathbf{j} \quad (4.48)$$

$$\text{with } \rho = \frac{1}{\sigma_0} \begin{pmatrix} 1 & \omega_c \tau \\ -\omega_c \tau & 1 \end{pmatrix}, \quad \text{and } R_H = \rho_{yx} = \frac{-B}{ne} \quad (4.49)$$

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

$$\begin{pmatrix} 1 & \omega_c \tau \\ -\omega_c \tau & 1 \end{pmatrix} \begin{pmatrix} 1 & -\omega_c \tau \\ \omega_c \tau & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} (1 + \omega_c^2 \tau^2) \quad (4.50)$$

so that

$$\sigma = \begin{pmatrix} \sigma_L & -\sigma_H \\ \sigma_H & \sigma_L \end{pmatrix} \quad (4.51)$$

$$\text{with } \sigma_L = \frac{\sigma_0}{1 + \omega_c^2 \tau^2} \quad \text{and} \quad \sigma_H = \sigma_0 \frac{\omega_c \tau}{1 + \omega_c^2 \tau^2} \quad (4.52)$$

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

$$\rho = \begin{pmatrix} 0 & -R_H \\ R_H & 0 \end{pmatrix} \quad \text{and} \quad \sigma = \begin{pmatrix} 0 & R_H^{-1} \\ -R_H^{-1} & 0 \end{pmatrix} \quad (4.53)$$

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:

$$|R_H| = \left(\frac{h}{e^2} \right) \frac{1}{n}, \quad n \text{ integer.} \quad (4.54)$$

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

$$\boxed{\sigma_H = n \frac{e^2}{h}} \quad (4.55)$$

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

$$\sigma_H = \nu \frac{e^2}{h}, \quad \nu = \frac{p}{q}. \quad (4.56)$$

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. Fig. 4.1 shows the experimental results by Willet and al, in 1987, as an illustration.

4.2.2 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. As we consider electrons strictly confined in two-dimensions, we drop the variable z . If the field is uniform, the potential is given by $V(y) = -Ey$, and the Hamiltonian becomes:

$$\begin{aligned} \mathcal{H} &\rightarrow \mathcal{H} + eEy \\ &= \frac{1}{2m} (p_x - eBy)^2 + \frac{1}{2m} p_y^2 + eEy \end{aligned} \quad (4.57)$$

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

$$\left[\frac{1}{2m} (\hbar k_x - eBy)^2 + \frac{1}{2m} p_y^2 + eEy \right] \phi(y) = \varepsilon \phi(y) \quad (4.58)$$

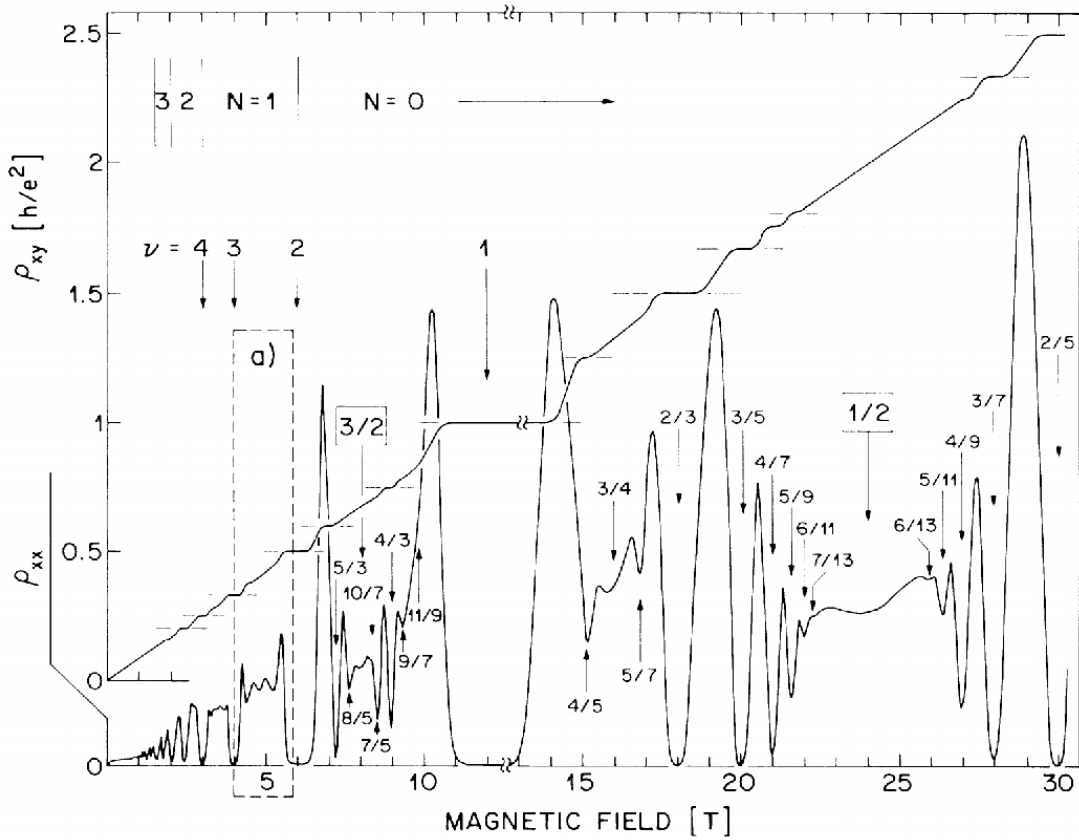


Figure 4.1: Experimental measurement of the diagonal resistivity ρ_{xx} and transverse resistance ρ_{xy} . Note the presence of very stable plateaus of ρ_{xy} at filling $\nu = 1$ (or $\nu = 2$). This is the first integer quantum Hall plateau, accompanied by a vanishing of the longitudinal conductance. Additional plateaus can be seen at fractional levels. Fig. taken from [Phys. Rev. Lett. 59, 1776 \(1987\)](#).

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

$$\begin{aligned}
 & \frac{1}{2m} (\hbar k_x - eBy)^2 + eEy = \frac{1}{2} m \omega_c^2 (y - y(k_x))^2 + A \\
 \Rightarrow & \frac{\hbar^2 k_x^2}{2m} - \frac{eB}{m} \hbar k_x y + \frac{1}{2m} e^2 B^2 y^2 + eEy = \frac{1}{2} m \omega_c^2 y^2 - m \omega_c^2 y y(k_x) + \frac{1}{2} m \omega_c^2 y(k_x)^2 + A \\
 \Rightarrow & \begin{cases} -\frac{eB}{m} \hbar k_x + eE = -m \omega_c^2 y(k_x) \\ \frac{\hbar^2 k_x^2}{2m} = \frac{1}{2} m \omega_c^2 y(k_x)^2 + A \end{cases}
 \end{aligned} \tag{4.59}$$

$$\begin{aligned}
\Rightarrow y(k_x) &= -\frac{eE}{m\omega_c^2} + \frac{\hbar k_x}{eB} \\
A &= \frac{\hbar^2 k_x^2}{2m} - \frac{1}{2}m\omega_c^2 y(k_x)^2 \\
&= \frac{\hbar^2 k_x^2}{2m} - \frac{1}{2}m \frac{e^2 B^2}{m^2} \left[\frac{e^2 E^2}{m^2 \omega_c^4} + \frac{\hbar^2 k_x^2}{e^2 B^2} - \frac{2eE}{m\omega_c^2} \frac{\hbar k_x}{eB} \right] \\
&= -\frac{1}{2}m \frac{e^2 E^2}{m^2 \omega_c^2} + \frac{1}{2}m\omega_c^2 \frac{2eE}{m\omega_c^2} \frac{\hbar k_x}{eB} \\
&= -\frac{1}{2}m \frac{E^2}{B^2} + E \frac{\hbar k_x}{B} \\
&= -\frac{1}{2}m \frac{E^2}{B^2} + \hbar k_x \frac{E}{B}
\end{aligned} \tag{4.60}$$

$$\begin{aligned}
\text{But } eEy(k_x) &= -\frac{e^2 E^2}{m e^2 B^2} m^2 + \hbar k_x \frac{eE}{eB} \\
&= -m \frac{E^2}{B^2} + \hbar k_x \frac{E}{B}
\end{aligned} \tag{4.61}$$

Finally,

$$A = eEy(k_x) + \frac{1}{2}m \frac{E^2}{B^2} \tag{4.62}$$

The eigenenergies are thus given by

$$\boxed{\varepsilon_{n,k_x} = \hbar\omega_c \left(n + \frac{1}{2} \right) + eEy(k_x) + \frac{1}{2}m \frac{E^2}{B^2}} \tag{4.63}$$

with

$$\boxed{y(k_x) = \frac{\hbar}{eB} k_x - \frac{eE}{m\omega_c^2}} \tag{4.64}$$

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

Let us compute the current carried by each eigenstate:

$$\begin{aligned}
\mathbf{j} &= -e\mathbf{v} = -e \frac{\mathbf{p} + e\mathbf{A}}{m} \\
\hat{j}_x &= -\frac{e}{m} (p_x + eA_x) = -\frac{e}{m} (p_x - eBy)
\end{aligned} \tag{4.65}$$

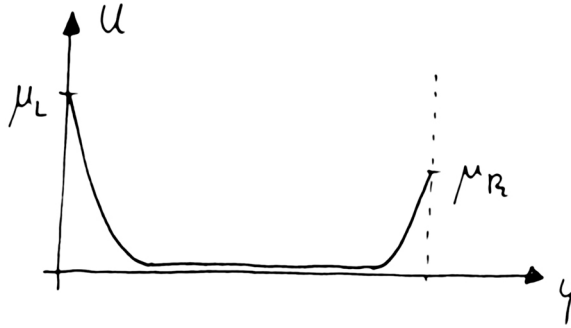
$$\begin{aligned}
\langle n, k_x | \hat{j}_x | n, k_x \rangle &= -\frac{e}{m} \hbar k_x + \frac{e^2}{m} By_{k_x} \\
&= -\frac{e}{m} \hbar k_x + \frac{e^2}{m} B \frac{\hbar}{eB} k_x - \frac{e^2}{m} B \frac{eE}{m\omega_c^2} \text{ same current. This is the classical result when there is} \\
&= -\frac{e^2}{m} B \frac{eEm^2}{me^2 B^2} = -e \frac{E}{B}
\end{aligned} \tag{4.66}$$

or

$$\langle j_x \rangle = -e \frac{E}{B} \quad (4.67)$$

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: $\sigma_H = R_H^{-1} = -n \frac{e}{B} \rightarrow j_x = -ne \frac{E}{B}$.

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:

$$\langle j_x \rangle(k_x) = -\frac{e}{B} E(y(k_x)) \quad (4.68)$$

The voltage bias in the y direction is given by:

$$\Delta V = -\int E(y) dy = -\sum_{k_x} E[y(k_x)] \Delta y \quad (4.69)$$

where Δy 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 $y(k_x) = \frac{\hbar}{eB} k_x - \frac{eE}{m\omega_c^2}$, so that

$$\Delta y = \frac{\hbar}{eB} \Delta k_x = \frac{\hbar}{eB} \frac{2\pi}{L_x} \quad (4.70)$$

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

$$J_{tot} = \sum_{k_x} \langle j_x \rangle(k_x) \quad (4.71)$$

$$= -\frac{e}{B} \frac{1}{\Delta y} \sum_{k_x} E(y(k_x)) \Delta y \quad (4.72)$$

$$= -\frac{e}{B} \frac{eB}{\hbar} \frac{L_x}{2\pi} (-\Delta V) \quad (4.73)$$

$$= \frac{e^2}{h} \Delta V L_x \quad (4.74)$$

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

$$I_n = \frac{e^2}{h} \Delta V \quad (4.75)$$

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

$$G_{xy} = n \frac{e^2}{h} \quad (4.76)$$

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

$$G = \begin{pmatrix} 0 & \frac{ne^2}{h} \\ -\frac{ne^2}{h} & 0 \end{pmatrix} \quad \text{and} \quad R = \begin{pmatrix} 0 & -\frac{h}{e^2 n} \\ \frac{h}{e^2 n} & 0 \end{pmatrix} \quad (4.77)$$

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

$$R_H = \frac{-B}{n_{el} e} \quad (4.78)$$

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

$$n_{el} = n \deg \frac{1}{L_x L_y} = n \frac{L_x L_y e B}{2\pi \hbar} \frac{1}{L_x L_y} = n \frac{e B}{2\pi \hbar} = \frac{n e B}{h} \quad (4.79)$$

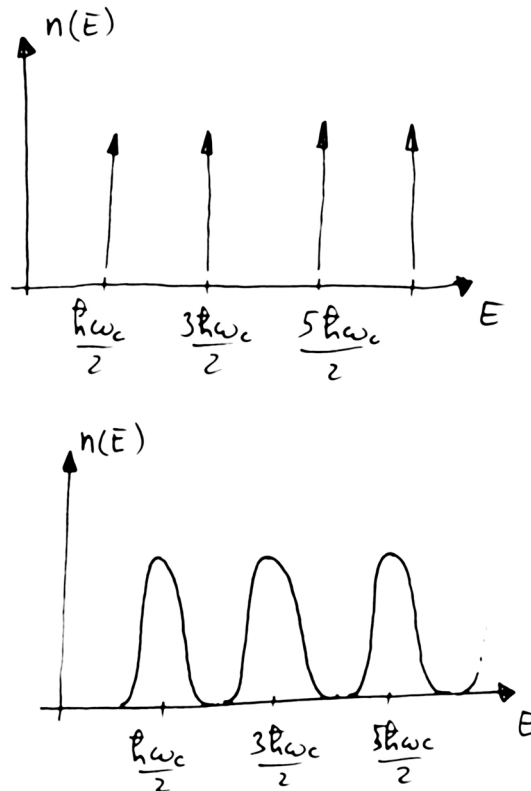
$$\Rightarrow R_H = \frac{-B h}{n e B e} = -\frac{h}{n e^2} \quad (4.80)$$

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. Another remarkable property is that it corresponds to the current transported by n independent one-dimensional fermionic modes, despite the macroscopic width of the sample.

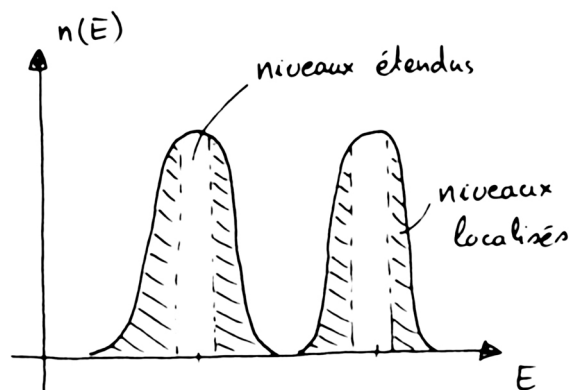
4.2.3 Effect of disorder

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 δ peaks of weight $\frac{L_x L_y}{2\pi\hbar} qB$. 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.

4.3 The Fractional Quantum Hall Effect

4.3.1 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 $\frac{L_x L_y}{2\pi\hbar} eB$ 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. Said otherwise, there is no energy competing with the Coulomb interaction and we are always in a strongly-interacting limit. This is an extremely difficult problem that does not possess any simple solution, even approximate. The picture that has imposed itself is based on a set of variational wavefunctions proposed by Laughlin in 1983. These variational wavefunctions 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.

4.3.2 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 can show, independently of the gauge, that the Hamiltonian takes the form

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

with $\hat{P} = \hat{\Pi}_x$, $\hat{Q} = \frac{1}{eB}\hat{\Pi}_y$ ($e = -q$) and

$$\begin{cases} \hat{\Pi}_x = \hat{p}_x + eA_x \\ \hat{\Pi}_y = \hat{p}_y + eA_y \end{cases} \quad [\hat{\Pi}_x, \hat{\Pi}_y] = -i\hbar eB.$$

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

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

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

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

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

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

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

$$l_B^2 = \frac{\hbar}{eB}$$

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: $\eta_x = x - x_0$, $\eta_y = y - y_0$ et $\dot{x} = -\omega_c(y - y_0)$, $\dot{y} = \omega_c(x - x_0)$).

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

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

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

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

The Hamiltonian can be rewritten:

$$H = \hbar\omega_c \left(a^\dagger a + \frac{1}{2} \right) + \hbar\omega' \left(b^\dagger b + \frac{1}{2} \right) \quad (4.81)$$

with $\omega' = 0$.

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

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

Of course:

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

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

Comments:

- So far, the argument does not depend on the gauge.
- The operators a^\dagger and b^\dagger are equal to the operators a_d^\dagger and a_g^\dagger seen in the exercise if one 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:

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

Intermediate steps: in the symmetric gauge,

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

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

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

$$\boxed{\bar{\partial}f(z, z^*) = \frac{\partial f}{\partial z^*}.} \quad (4.82)$$

Similarly,

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

$$\boxed{\partial f(z, z^*) = \frac{\partial f}{\partial z}.} \quad (4.83)$$

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

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

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

$\varphi(z, z^*)$ 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

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

Similarly, the $m = 0$ states satisfy

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

whose general solution is

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

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

$$\Phi_{n=0, m=0}(z, z^*) = \frac{1}{\sqrt{2\pi l_B^2}} e^{-|z|^2/4l_B^2}. \quad (4.84)$$

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

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

$$\Phi_{n=0, m}(z, z^*) = \frac{1}{\sqrt{2\pi l_B^2 m!}} \left(\frac{z}{\sqrt{2}l_B} \right)^m e^{-|z|^2/4l_B^2}, \quad (4.85)$$

since

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

Similarly,

$$\Phi_{n, m=0}(z, z^*) = \frac{1}{\sqrt{2\pi l_B^2 n!}} \left(\frac{z^*}{\sqrt{2}l_B} \right)^n e^{-|z|^2/4l_B^2}. \quad (4.86)$$

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. Said otherwise, the weights of the orbitals are located around a circle of increasing radius.

4.3.3 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 $g\mu_B H$. A quick comment: this argument is in fact not entirely correct: the Zeeman energy is generally extremely small and does not dominate compared to the Coulomb energy. In the absence of a Zeeman field, due to the structure of the Coulomb interaction, the groundstate of the interacting Hamiltonian would be an integer quantum Hall state for the charge degrees of freedom, and the spin degrees would form a ferromagnet. The weak Zeeman field is then enough to break the ferromagnetic phase and polarize (maximally) the state in the direction of the magnetic field. The distinction between the two approaches is largely irrelevant to the rest of the chapter.

Since the spin component of the wavefunction 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

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

So one looks for a function of the form

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

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

In the lowest Landau level, a convenient basis of the one particle states is given by

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

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

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 $z_i^0, z_i^1, \dots, z_i^N$,

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

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

$$f_N(\{z_i\}) = \prod_{i < j} (z_i - z_j). \quad (4.87)$$

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 electronic 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,

$$\begin{aligned} \rho(z) &= \frac{1}{N!} \sum_P (\text{sgn}(P))^2 \sum_i \int dz_1 \dots dz_N \delta(z - z_i) |\varphi_{P(1)}(z_1)|^2 \dots |\varphi_{P(n)}(z_N)|^2 \\ &= \frac{1}{N!} \sum_P \sum_i |\varphi_{P(i)}(z)|^2 \\ &= \frac{1}{N!} \sum_P \sum_i |\varphi_i(z)|^2 \\ &= \sum_i |\varphi_i(z)|^2. \end{aligned}$$

But

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

So,

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

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

4.3.4 $\nu < 1$: 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:

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

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

- It only involves states of the the first Landau level
 $\rightarrow f$ does not depend on the z_i^* .
- ψ is totally antisymmetric
 $\rightarrow f$ is odd.
- ψ 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 $\Rightarrow f(z_i - z_j) = (z_i - z_j)^m$.

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

$$\psi_m(z_1, \dots, z_N) = \prod_{i < j} (z_i - z_j)^m e^{-\sum_i |z_i|^2 / 4l_B^2}. \quad (4.88)$$

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

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

$$\nu = \frac{1}{m}. \quad (4.89)$$

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

Energy: These wavefunctions turn out to be remarkably accurate approximation of the groundstates of the Coulomb interaction. To get an idea why, let us consider the other states of density $\frac{1}{m}$:

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

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. Let us 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' . The P_m operator are also called *Haldane pseudopotentials*. They form a basis of central two-body interactions in two dimensions. Assuming $v_{m'} = 0$ if $m' \geq m$, and $v_{m'} > 0$ if $m' < m$, the function $\psi_{1/m}$ is an eigenstate of $\hat{V} = \sum_{i < j} V(i, j)$ with

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

Now we want to add one particle without changing the size of the samples. Within our approximation, this means that we want the total polynomial to keep a degree less or equal than $m(N-1)$ in each variables. With this constraint, adding a particle requires creating of pairs of angular momentum $< m$. 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 $m' \geq m$ as perturbations.

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

It is nonetheless possible to create 0 energy states starting from the Laughlin states when $v_{m' \geq m} = 0$. Indeed,

$$\psi_{1/m}^P = P(z_1, \dots) \psi_{1/m}$$

is a valid fermionic wavefunction for any totally symmetric polynomial P . It is also a state of energy 0 as P can only increase the relative momenta of particles.

Gapless edge modes: A crucial property of the quantum Hall effect is the existence of edge modes. A simple way to visualize their appearance is to introduce a trapping potential. In a circular geometry, we could for example add a quadratic potential

$$V_{\text{trap}} = V \frac{r^2}{l_B^2}.$$

While we could estimate its action on the Laughlin state, it remains challenging, so instead, we will use a simpler ersatz:

$$V_{\text{trap}} = 2V L_z.$$

Indeed, we have

$$V_{\text{trap}} \Phi_{n=0,m} = mV \Phi_{n=0,m},$$

while the radius of the orbital $\Phi_{n=0,m}$ is approximately $l_b \sqrt{2m+1}$.

If we suppose $V \ll E_{\text{gap}}$, the gap above the Laughlin state, we can assume that $\psi_{1/m}$ is still the ground state of the system at filling $\frac{1}{m}$. Now let's turn to $\psi_{1/m}^P$. We take P to be of constant degree that we denote d_P . Clearly, the energy of $\psi_{1/m}^P$ is simply $2V d_P + E_0$. These states form a low-energy branch, linear in momentum. These states acquire a finite energy due to an extension at the edge of the system (we also technically create a hole at the core of the sample, which we will neglect: it is an artifact of working with a constant number of particles): we have linear energy edge states. Now we can count these states by listing the linearly independent polynomials of each degree:

$$\begin{aligned} d_P = 1 : P_1 &= \sum_j z_j \\ d_P = 2 : &\begin{cases} P_1 = \sum_j z_j^2 \\ P_2 = \sum_{i \neq j} z_i z_j \end{cases} \\ d_P = 3 : &\begin{cases} P_1 = \sum_j z_j^3 \\ P_2 = \sum_{i \neq j} z_i^2 z_j \\ P_3 = \sum_{i \neq j \neq k} z_i z_j z_k \end{cases} \\ d_P = 4 : &\begin{cases} P_1 = \sum_j z_j^4 \\ P_2 = \sum_{i \neq j} z_i^3 z_j \\ P_3 = \sum_{i \neq j} z_i^2 z_j^2 \\ P_4 = \sum_{i \neq j \neq k} z_i^2 z_j z_k \\ P_5 = \sum_{i \neq j \neq k \neq l} z_i z_j z_k z_l \end{cases} \end{aligned}$$

We have therefore 1 groundstate, 1 state at $\delta L = 1$, 2 states at $\delta L = 2$, 3 states at $\delta L = 3$ and 5 states at $\delta L = 4$. This reproduces the counting of states for chiral

one-dimensional bosons (or fermions):

$$H = \sum_{m \geq 1} m b_m^\dagger b_m$$

$$H = \sum_{k \in \mathbb{Z}} \left(k + \frac{1}{2}\right) c_k^\dagger c_k.$$

This suggests, and it can be proven rigorously, that the edge states behave as emergent chiral fermions (or bosons, it is equivalent in one-dimension).

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 ≥ 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:

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

This is the $\nu = 1/m$ Laughlin wave function with a particle sitting at $z_0 = 0$. The wave function $\psi_{1/m}^{(m)}(z_0 = 0, z_1, \dots, z_N)$ 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 construction would lead us too far. We just note that, within this approach, it is plausible that there are plateaus for

$$\nu = \frac{n}{2np \pm 1} \quad n, p \text{ integers } \geq 1. \quad (4.90)$$

Appendix

Angular momentum

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

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

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

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

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

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

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

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

$$\begin{cases} \mathbf{R}_{CM} = \frac{\mathbf{r}_1 + \mathbf{r}_2}{2} \\ \mathbf{P}_{CM} = \mathbf{p}_1 + \mathbf{p}_2 \end{cases} \quad \begin{cases} \mathbf{R}_{rel} = \mathbf{r}_1 - \mathbf{r}_2 \\ \mathbf{P}_{rel} = \frac{\mathbf{p}_1 - \mathbf{p}_2}{2} \end{cases}$$

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

$$\mathbf{l}_{CM} = \mathbf{R}_{CM} \wedge \mathbf{P}_{CM}, \quad \mathbf{l}_{rel} = \mathbf{R}_{rel} \wedge \mathbf{P}_{rel}$$

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

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

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

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

$$\begin{aligned}
l_{rel}^z &= \frac{\hbar}{2} [(a_1^\dagger - a_2^\dagger)(a_1 - a_2) - (b_1^\dagger - b_2^\dagger)(b_1 - b_2)] \\
l_{CM}^z &= \frac{\hbar}{2} [(a_1^\dagger + a_2^\dagger)(a_1 + a_2) - (b_1^\dagger + b_2^\dagger)(b_1 + b_2)] \\
\rightarrow l_{rel}^z &= \hbar [\alpha_-^\dagger \alpha_- - \beta_-^\dagger \beta_-] \quad l_{CM}^z = \hbar [\alpha_+^\dagger \alpha_+ - \beta_+^\dagger \beta_+]
\end{aligned}$$

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

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

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

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

Chapter 5

Electron-phonon interaction

5.1 Introduction

In band theory, the lattice is assumed to be static. In reality, ions (can) 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.

5.2 Phonons - Reminder

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

$$T = \sum_{n,\alpha,i} \frac{M_\alpha}{2} \dot{s}_{n,\alpha,i}^2, \quad n = 1, \dots, N; \quad \alpha = 1, \dots, r; \quad i = x, y, z \quad (5.1)$$

where M_α is the mass of the atom α in the unit cell. Furthermore let us denote by $V(\{\mathbf{s}_{n,\alpha}\})$ 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:

$$V(\{\mathbf{s}_{n,\alpha}\}) - V(\{\mathbf{0}\}) = \frac{1}{2} \sum_{n,\alpha,i} \sum_{n',\alpha',i'} \phi_{n,\alpha,i}^{n',\alpha',i'} s_{n,\alpha,i} s_{n',\alpha',i'} \quad (5.2)$$

with

$$\phi_{n,\alpha,i}^{n',\alpha',i'} = \left. \frac{\partial^2 V}{\partial s_{n,\alpha,i} \partial s_{n',\alpha',i'}} \right|_{\mathbf{s}_{n,\alpha}=\mathbf{0}}$$

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.

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

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

The Lagrange equations for the variable $s_{n,\alpha,i}$ take the form:

$$\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{s}_{n,\alpha,i}} \right) - \frac{\partial L}{\partial s_{n,\alpha,i}} = 0$$

$$M_\alpha \ddot{s}_{n,\alpha,i} = - \sum_{n',\alpha',i'} \phi_{n,\alpha,i}^{n',\alpha',i'} s_{n',\alpha',i'} \quad (5.3)$$

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 (the Jacobian must be positive). Let us denote by $\omega_j > 0$ the square root of ω_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 ω_j^2 is denoted by $u_{n,\alpha,i}^{(j)}$.

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

$$\begin{aligned} u_{n,\alpha,i} &= c_{\alpha,i} e^{i\mathbf{q}\cdot\mathbf{R}_n} \\ \Rightarrow \omega^2 c_{\alpha,i} &= \sum_{\alpha',i'} \sum_{n'} \frac{1}{\sqrt{M_\alpha M_{\alpha'}}} \phi_{\alpha,i,n}^{\alpha',i',n'} e^{i\mathbf{q}\cdot(\mathbf{R}_{n'} - \mathbf{R}_n)} c_{\alpha',i'} \\ &= \sum_{\alpha',i'} D_{\alpha,i}^{\alpha',i'}(\mathbf{q}) c_{\alpha',i'} \end{aligned}$$

For each value of \mathbf{q} , we thus have a system of $3r$ equations, which leads to $3r$ eigenfrequencies:

$$\omega_j(\mathbf{q}), \quad j = 1, \dots, 3r$$

The eigenvector associated to $\omega_j(\mathbf{q})$ has $3r$ components $e_{\alpha,i}^{(j)}(\mathbf{q})$. It is often written as r vectors with three components $\mathbf{e}_\alpha^{(j)}(\mathbf{q})$. These vectors are called the *polarization* vectors.

The final solution for the displacement is thus a linear combination of the displacements $\mathbf{e}_\alpha^{(j)}(\mathbf{q}) e^{i(\mathbf{q}\cdot\mathbf{R}_n - \omega_j(\mathbf{q})t)}$. It is convenient to include the time dependence into the coefficient, and write this linear combination as:

$$s_{n,\alpha,i}(t) = \frac{1}{\sqrt{NM_\alpha}} \sum_{j,\mathbf{q}} Q_j(\mathbf{q}, t) e_{\alpha,i}^{(j)}(\mathbf{q}) e^{i\mathbf{q}\cdot\mathbf{R}_n} \quad (5.4)$$

There are in total $3rN$ quantities $Q_j(\mathbf{q}, t)$ ($3r$ for j , N for \mathbf{q}). One can consider them as a new set of variables. In order for the $s_{n,\alpha,i}(t)$ to be real, the $Q_j(\mathbf{q}, t)$ must be complex and satisfy $Q_j^*(\mathbf{q}, t) = Q_j(-\mathbf{q}, t)$ (one can show that $\omega_j(-\mathbf{q}) = \omega_j(\mathbf{q})$ and choose $\mathbf{e}_\alpha^{(j)}(-\mathbf{q}) = \mathbf{e}_\alpha^{(j)}(\mathbf{q})$).

Besides, the eigenvectors of a symmetric matrix are orthogonal:

$$\sum_{\alpha,i} e_{\alpha,i}^{(j)}(\mathbf{q}) e_{\alpha,i}^{(j')}(\mathbf{q}') = \delta_{j,j'} \delta_{\mathbf{q},\mathbf{q}'}$$

With the help of these relations one can show that the Lagrangian can be written:

$$L = \frac{1}{2} \sum_{j,\mathbf{q}} \left[\dot{Q}_j^*(\mathbf{q}, t) \dot{Q}_j(\mathbf{q}, t) - \omega_j^2(\mathbf{q}) Q_j^*(\mathbf{q}, t) Q_j(\mathbf{q}, t) \right]$$

The Hamiltonian can then be obtained with the help of a Legendre transformation:

$$P_j(\mathbf{q}, t) = \frac{\partial L}{\partial \dot{Q}_j(\mathbf{q}, t)} = \dot{Q}_j^*(\mathbf{q}, t)$$

$$H = \frac{1}{2} \sum_{j, \mathbf{q}} \left[P_j^*(\mathbf{q}, t) P_j(\mathbf{q}, t) + \omega_j^2 Q_j^*(\mathbf{q}, t) Q_j(\mathbf{q}, t) \right] \quad (5.5)$$

We recognize a set of uncoupled harmonic oscillators, a direct consequence of expanding to second-order only around the equilibrium position.

5.3 Quantization

5.3.1 Canonical quantization

Following canonical quantization, we postulate the usual commutation relations between the positions and the momenta of the ions. This leads to the following commutation relations for the normal coordinates:

$$[Q_j(\mathbf{q}), P_{j'}(\mathbf{q}')] = i\hbar \delta_{\mathbf{q}, \mathbf{q}'} \delta_{j, j'}$$

If, still by analogy with the harmonic oscillator, one defines creation and annihilation operators by:

$$\begin{cases} a_j^\dagger(\mathbf{q}) = \frac{1}{\sqrt{2\hbar\omega_j(\mathbf{q})}} (\omega_j Q_j^*(\mathbf{q}) - iP_j(\mathbf{q})) \\ a_j(\mathbf{q}) = \frac{1}{\sqrt{2\hbar\omega_j(\mathbf{q})}} (\omega_j Q_j(\mathbf{q}) - iP_j^*(\mathbf{q})) \end{cases}$$

the Hamiltonian takes the form:

$$H = \sum_{j, \mathbf{q}} \hbar\omega_j(\mathbf{q}) \left(a_j^\dagger(\mathbf{q}) a_j(\mathbf{q}) + \frac{1}{2} \right)$$

The operators $a_j^\dagger(\mathbf{q})$, $a_j(\mathbf{q})$ satisfy bosonic commutation rules:

$$[a_j(\mathbf{q}), a_{j'}^\dagger(\mathbf{q}')] = \delta_{j, j'} \delta_{\mathbf{q}, \mathbf{q}'}$$

The inverse relations giving the positions and momenta in terms of the $a_j(\mathbf{q})$ and $a_j^\dagger(\mathbf{q})$ are:

$$\begin{cases} Q_j(\mathbf{q}) = \sqrt{\frac{\hbar}{2\omega_j(\mathbf{q})}} (a_j^\dagger(-\mathbf{q}) + a_j(\mathbf{q})) \\ P_j(\mathbf{q}) = i\sqrt{\frac{\hbar\omega_j(\mathbf{q})}{2}} (a_j^\dagger(\mathbf{q}) - a_j(-\mathbf{q})) \end{cases}$$

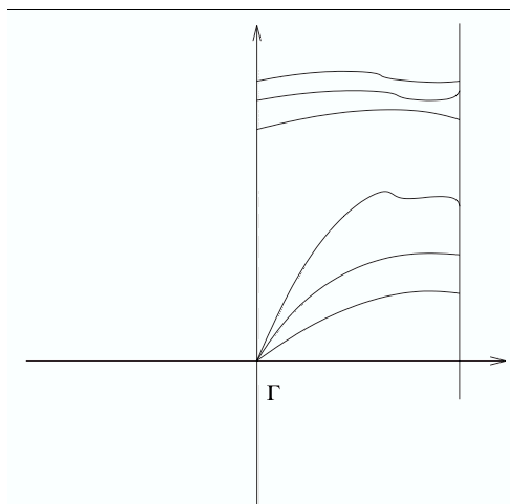
The form of the Hamiltonian shows that the system behaves as a gas of *bosons*. It is again typical from a second-order expansion around an equilibrium position.

5.3.2 Energetics

The $3r$ functions $\omega_j(\mathbf{q})$ are called dispersion relations. When $\mathbf{q} \rightarrow 0$, all the unit cells vibrate in phase. If all atoms inside the unit cells vibrate in phase, $\omega_j(\mathbf{q})$ 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 $\mathbf{q} \rightarrow 0$. They are called *optical modes*.

The general form of the dispersion is thus:



5.3.3 Thermodynamics

The calculation of the specific heat at low temperature is completely equivalent to that of antiferromagnets since acoustic phonons have a linear dispersion, like magnons in an antiferromagnet. It leads to

$$C_v \sim T^3 \quad (5.6)$$

A high temperature, the specific heat saturates when the temperature becomes larger than the largest phonon frequency.

5.3.4 Limitations

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. . . , that is to say, interactions.

5.4 Electron-phonon interaction

The electron-phonon interaction comes from the Coulomb interaction between the electrons and the ions, that can be written:

$$H_{el-ion} = \sum_{l,i} V(\mathbf{r}_l - \mathbf{R}_i)$$

where \mathbf{r}_l stands for the position of the electron and \mathbf{R}_i for that of the ion, and from the fact that \mathbf{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:

$$\mathbf{R}_{n,\alpha}(t) = \underbrace{\mathbf{R}_n + \mathbf{R}_\alpha}_{\mathbf{R}_{n,\alpha}} + \mathbf{s}_{n,\alpha}(t),$$

where we took $\mathbf{R}_{n,\alpha}$ to be the equilibrium position. If $\mathbf{s}_{n,\alpha}(t)$ is small, we can again expand $V_\alpha(\mathbf{r}_l - \mathbf{R}_i)$, where V depends on the site in the unit cell:

$$V_\alpha(\mathbf{r}_l - \mathbf{R}_{n,\alpha} - \mathbf{s}_{n,\alpha}) = V_\alpha(\mathbf{r}_l - \mathbf{R}_{n,\alpha}) - \mathbf{s}_{n,\alpha} \cdot \nabla V_\alpha(\mathbf{r}_l - \mathbf{R}_{n,\alpha})$$

The first term describes the motion of electrons in the periodic potential of the crystal at equilibrium. It is a simple static, periodic potential which is included in the Bloch theorem.

The second term describes the coupling between electrons and the lattice vibrations. If we replace $\mathbf{s}_{n,\alpha}$ by the solutions we obtained in the previous section, we obtain

$$H_{el-ph} = - \sum_{\alpha,n,l} \frac{1}{\sqrt{NM_\alpha}} \sum_{j,\mathbf{q}} Q_j(\mathbf{q}) \mathbf{e}_\alpha^{(j)}(\mathbf{q}) \cdot \nabla V_\alpha(\mathbf{r}_l - \mathbf{R}_{n,\alpha}) e^{i\mathbf{q} \cdot \mathbf{R}_n}$$

To express this Hamiltonian in second quantization, we proceed in two steps:

- For the phonon part, we just have to replace $Q_j(\mathbf{q})$ by its expression in terms of a_j and a_j^\dagger :

$$Q_j(\mathbf{q}) = \sqrt{\frac{\hbar}{2\omega_j(\mathbf{q})}} (a_j^\dagger(-\mathbf{q}) + a_j(\mathbf{q}))$$

- For the electron part, we calculate the matrix element of $\nabla V_\alpha(\mathbf{r}_l - \mathbf{R}_{n,\alpha})$ to obtain its second quantized form.

For this, we expand V_α in Fourier series: $V_\alpha(\mathbf{r}) = \sum_{\mathbf{K}} e^{i\mathbf{K} \cdot \mathbf{r}} V_{\alpha,\mathbf{K}}$ where \mathbf{K} is arbitrary because $V_\alpha(\mathbf{r})$ is not necessarily periodic. We naturally obtain

$$\nabla V_\alpha(\mathbf{r}_l - \mathbf{R}_{n,\alpha}) = \sum_{\mathbf{K}} e^{i\mathbf{K}(\mathbf{r}_l - \mathbf{R}_{n,\alpha})} V_{\alpha,\mathbf{K}} i\mathbf{K}$$

The matrix elements of this potential between Bloch waves are given by:

$$\langle \mathbf{k}', \sigma' | \nabla V_\alpha | \mathbf{k}, \sigma \rangle = \sum_{\mathbf{K}} e^{-i\mathbf{K}\mathbf{R}_{n,\alpha}} V_{\alpha,\mathbf{K}} i\mathbf{K} \langle \mathbf{k}', n', \sigma' | e^{i\mathbf{K}\mathbf{r}} | \mathbf{k}, n, \sigma \rangle$$

with

$$\langle \mathbf{k}', \sigma' | e^{i\mathbf{K}\mathbf{r}} | \mathbf{k}, \sigma \rangle = \int d\mathbf{r} u_{\mathbf{k}',n'}^*(\mathbf{r}) u_{\mathbf{k},n}(\mathbf{r}) e^{i(-\mathbf{k}'+\mathbf{K}+\mathbf{k})\mathbf{r}} \delta_{\sigma,\sigma'}$$

This integral vanishes unless

$$\mathbf{k}' = \mathbf{K} + \mathbf{k} + \mathbf{G}$$

where \mathbf{G} is a vector of the reciprocal lattice. Indeed, $u_{\mathbf{k}'}^*(\mathbf{r})u_{\mathbf{k}}(\mathbf{r})$ is a periodic function. It can thus be expanded in Fourier series:

$$u_{\mathbf{k}',n'}^*(\mathbf{r})u_{\mathbf{k},n}(\mathbf{r}) = \sum_{\mathbf{G}} f_{n',n}(\mathbf{G}) e^{i\mathbf{G}\cdot\mathbf{r}}$$

where the vectors \mathbf{G} are vectors of the reciprocal lattice, which leads to

$$\langle \mathbf{k}', n', \sigma' | e^{i\mathbf{K}\mathbf{r}} | \mathbf{k}, n, \sigma \rangle = \delta_{\sigma,\sigma'} \sum_{\mathbf{G}} f_{n',n}(\mathbf{G}) \int d\mathbf{r} e^{i(-\mathbf{k}'+\mathbf{K}+\mathbf{k}+\mathbf{G})\mathbf{r}}$$

Now, the integral $\int d\mathbf{r} e^{i(-\mathbf{k}'+\mathbf{K}+\mathbf{k}+\mathbf{G})\mathbf{r}}$ vanishes unless $-\mathbf{k}' + \mathbf{K} + \mathbf{k} + \mathbf{G} = \mathbf{0}$.

Moreover, the sum over \mathbf{R}_n can be written:

$$\sum_{\mathbf{R}_n} e^{i\mathbf{q}\cdot\mathbf{R}_n} e^{-i\mathbf{K}\cdot\mathbf{R}_n} = \sum_{\mathbf{R}_n} e^{i(\mathbf{q}-\mathbf{K})\cdot\mathbf{R}_n}$$

This sum vanishes unless $\mathbf{K} = \mathbf{q} + \mathbf{G}$ where \mathbf{G} is a vector of the reciprocal lattice.

As a consequence, \mathbf{k} and \mathbf{k}' must be related by

$$\mathbf{k}' = \mathbf{q} + \mathbf{k} + \mathbf{G}$$

But \mathbf{k} and \mathbf{k}' must both belong to the first Brillouin zone. So, for a given pair of vectors \mathbf{k} and \mathbf{q} , there is one and only one vector \mathbf{G} that brings $\mathbf{k} + \mathbf{q}$ back into the first Brillouin zone.

One thus distinguishes two types of processes:

- The "normal" processes such that $\mathbf{k} + \mathbf{q} \in \text{BZ}$ ($\mathbf{G} = \mathbf{0}$).
- The "Umklapp" processes such that $\mathbf{k} + \mathbf{q} \notin \text{BZ} \Rightarrow \mathbf{G} \neq \mathbf{0}$ and $\mathbf{k}' \neq \mathbf{k} + \mathbf{q}$.

Finally, the interaction can be written

$$\begin{aligned} H &= - \sum_{\alpha,\sigma,\mathbf{k},\mathbf{q},j,n,n'} \sqrt{\frac{N\hbar}{2M_\alpha\omega_j(\mathbf{q})}} \\ &\times \sum_{\mathbf{G}} \left[\mathbf{e}_{\alpha,\mathbf{q}}^{(j)} \cdot i(\mathbf{q} + \mathbf{G}) \right] \int V_\alpha(\mathbf{r}') e^{i(\mathbf{q}+\mathbf{G})\cdot(\mathbf{r}'-\mathbf{R}_\alpha)} d\mathbf{r}' \\ &\times \int u_{\mathbf{k}+\mathbf{q}+\mathbf{G},n'}^*(\mathbf{r}) u_{\mathbf{k},n}(\mathbf{r}) d\mathbf{r} \\ &\times \left[a_{-\mathbf{q}}^\dagger(j) + a_{\mathbf{q}}(j) \right] c_{\mathbf{k}+\mathbf{q}+\mathbf{G},n',\sigma}^\dagger c_{\mathbf{k},n,\sigma} \end{aligned}$$

where \mathbf{G} is such that $\mathbf{k} + \mathbf{q} + \mathbf{G} \in 1^{\text{st}}$ Brillouin zone.

The Hamiltonian can be finally written in a more compact way as:

$$H = - \sum_{\mathbf{k}, \mathbf{k}', \mathbf{q}, j, \sigma} M_{n', n}(\mathbf{k}, \mathbf{k}', \mathbf{q}, j) \left[a_{-\mathbf{q}}^\dagger(j) + a_{\mathbf{q}}(j) \right] c_{\mathbf{k}', n', \sigma}^\dagger c_{\mathbf{k}, n, \sigma} \quad (5.7)$$

with $\mathbf{q} = \mathbf{k}' - \mathbf{k} + \mathbf{G}$

$$\begin{cases} \mathbf{G} = \mathbf{0} & \text{for normal processes} \\ \mathbf{G} \neq \mathbf{0} & \text{for Umklapp processes.} \end{cases}$$

Note that σ here is both a spin and orbital index.

5.5 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 described 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:

$$H_{el-ph} = \sum_{\mathbf{k}, \mathbf{q}, \sigma} M_{\mathbf{q}} \left[a_{-\mathbf{q}}^\dagger + a_{\mathbf{q}} \right] c_{\mathbf{k}+\mathbf{q}, \sigma}^\dagger c_{\mathbf{k}, \sigma}$$

where for simplicity we only consider one branch of phonons, and a single electronic orbital. We also limit ourselves to normal processes, which can be justified a posteriori. Furthermore, we assume that the electrons and phonons are described by

$$H_{el} = \sum_{\mathbf{k}, \sigma} \epsilon_{\mathbf{k}} c_{\mathbf{k}, \sigma}^\dagger c_{\mathbf{k}, \sigma}, \quad H_{ph} = \sum_{\mathbf{q}} \hbar \omega_{\mathbf{q}} \left(a_{\mathbf{q}}^\dagger a_{\mathbf{q}} + \frac{1}{2} \right).$$

If the matrix elements $M_{\mathbf{q}}$ 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. We will resort to an equivalent 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*. A canonical transformation is a unitary transformation of the Hamiltonian and of the vectors of the Hilbert space that does not modify the spectrum:

$$\begin{aligned}\hat{H} &\rightarrow U\hat{H}U^{-1} \\ |\Psi\rangle &\rightarrow U|\Psi\rangle\end{aligned}$$

with $U^\dagger = U^{-1}$. The following properties are then satisfied:

1. The new Hamiltonian is hermitian.
2. If $|\Psi\rangle$ is an eigenstate of \hat{H} with eigenvalue E , then $U|\Psi\rangle$ is an eigenstate of $U\hat{H}U^{-1}$ with the same eigenvalue.

The property $U^\dagger = U^{-1}$ is satisfied by operators of the form $U = e^{iS}$ provided S is hermitian. The transformed Hamiltonian can be written as an expansion in terms of commutators:

$$\begin{aligned}e^{iS}\hat{H}e^{-iS} &= \hat{H} + i[S, \hat{H}] + \frac{i^2}{2!}[S, [S, \hat{H}]] + \dots \\ &= \hat{H} + \sum_{n=1}^{\infty} \frac{i^n}{n!}[S, [S, \dots [S, \hat{H}]]]\end{aligned}$$

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 λ , it can thus be written :

$$\hat{H}(\lambda) = \hat{H} + i\lambda[S, \hat{H}(0)] + \frac{i^2}{2!}\lambda^2[S, [S, \hat{H}(0)]] + \dots$$

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 $M_{\mathbf{q}}$ 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 of order 1 in $M_{\mathbf{q}}$
- 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 $H_{el} + H_{ph}$.

$$\begin{aligned} [a_{-\mathbf{q}}^\dagger c_{\mathbf{k}+\mathbf{q}}^\dagger c_{\mathbf{k}}, c_l^\dagger c_l] &= a_{-\mathbf{q}}^\dagger [c_{\mathbf{k}+\mathbf{q}}^\dagger c_{\mathbf{k}}, c_l^\dagger c_l] \\ &= a_{-\mathbf{q}}^\dagger \left\{ \delta_{\mathbf{k},l} (c_{\mathbf{k}+\mathbf{q}}^\dagger c_{\mathbf{k}} c_{\mathbf{k}}^\dagger c_{\mathbf{k}} - c_{\mathbf{k}}^\dagger c_{\mathbf{k}} c_{\mathbf{k}+\mathbf{q}}^\dagger c_{\mathbf{k}}) \right. \\ &\quad \left. + \delta_{\mathbf{k}+\mathbf{q},l} (c_{\mathbf{k}+\mathbf{q}}^\dagger c_{\mathbf{k}} c_{\mathbf{k}+\mathbf{q}}^\dagger c_{\mathbf{k}+\mathbf{q}} - c_{\mathbf{k}+\mathbf{q}}^\dagger c_{\mathbf{k}+\mathbf{q}} c_{\mathbf{k}+\mathbf{q}}^\dagger c_{\mathbf{k}}) \right\} \\ &= a_{-\mathbf{q}}^\dagger \left\{ \delta_{\mathbf{k},l} c_{\mathbf{k}+\mathbf{q}}^\dagger c_{\mathbf{k}} - \delta_{\mathbf{k}+\mathbf{q},l} c_{\mathbf{k}+\mathbf{q}}^\dagger c_{\mathbf{k}} \right\} \end{aligned}$$

$$\Rightarrow [a_{-\mathbf{q}}^\dagger c_{\mathbf{k}+\mathbf{q}}^\dagger c_{\mathbf{k}}, H_{el}] = a_{-\mathbf{q}}^\dagger (\epsilon_{\mathbf{k}} - \epsilon_{\mathbf{k}+\mathbf{q}}) c_{\mathbf{k}+\mathbf{q}}^\dagger c_{\mathbf{k}}$$

Similarly

$$[a_{\mathbf{q}} c_{\mathbf{k}+\mathbf{q}}^\dagger c_{\mathbf{k}}, H_{el}] = a_{\mathbf{q}} (\epsilon_{\mathbf{k}} - \epsilon_{\mathbf{k}+\mathbf{q}}) c_{\mathbf{k}+\mathbf{q}}^\dagger c_{\mathbf{k}}$$

Moreover,

$$\begin{aligned} [a_{-\mathbf{q}}^\dagger, a_{\mathbf{k}}^\dagger a_{\mathbf{k}}] &= -\delta_{\mathbf{k},-\mathbf{q}} a_{-\mathbf{q}}^\dagger \\ \Rightarrow [a_{-\mathbf{q}}^\dagger, H_{ph}] &= -\hbar\omega_{-\mathbf{q}} a_{-\mathbf{q}}^\dagger \end{aligned}$$

Similarly,

$$\begin{aligned} [a_{\mathbf{q}}, a_{\mathbf{k}}^\dagger a_{\mathbf{k}}] &= \delta_{\mathbf{k},\mathbf{q}} a_{\mathbf{q}} \\ \Rightarrow [a_{\mathbf{q}}, H_{ph}] &= \hbar\omega_{\mathbf{q}} a_{\mathbf{q}} \end{aligned}$$

Finally,

$$[a_{-\mathbf{q}}^\dagger c_{\mathbf{k}+\mathbf{q}}^\dagger c_{\mathbf{k}}, H_{el} + H_{ph}] = (\epsilon_{\mathbf{k}} - \epsilon_{\mathbf{k}+\mathbf{q}} - \hbar\omega_{\mathbf{q}}) a_{-\mathbf{q}}^\dagger c_{\mathbf{k}+\mathbf{q}}^\dagger c_{\mathbf{k}}$$

$$[a_{\mathbf{q}} c_{\mathbf{k}+\mathbf{q}}^\dagger c_{\mathbf{k}}, H_{el} + H_{ph}] = (\epsilon_{\mathbf{k}} - \epsilon_{\mathbf{k}+\mathbf{q}} + \hbar\omega_{\mathbf{q}}) a_{\mathbf{q}} c_{\mathbf{k}+\mathbf{q}}^\dagger c_{\mathbf{k}}$$

The solution of the equation

$$[S, H_{el} + H_{ph}] = iH_{el-ph}$$

can thus be written

$$S = i \sum_{\mathbf{k}, \mathbf{q}, \sigma} \frac{M_{\mathbf{q}}}{\epsilon_{\mathbf{k}} - \epsilon_{\mathbf{k}+\mathbf{q}} - \hbar\omega_{\mathbf{q}}} a_{-\mathbf{q}}^\dagger c_{\mathbf{k}+\mathbf{q}, \sigma}^\dagger c_{\mathbf{k}, \sigma} + i \sum_{\mathbf{k}, \mathbf{q}, \sigma} \frac{M_{\mathbf{q}}}{\epsilon_{\mathbf{k}} - \epsilon_{\mathbf{k}+\mathbf{q}} + \hbar\omega_{\mathbf{q}}} a_{\mathbf{q}} c_{\mathbf{k}+\mathbf{q}, \sigma}^\dagger c_{\mathbf{k}, \sigma}$$

(5.8)

(the spin has been added at the end of the calculation).

The operator S satisfies the two conditions. The first terms in powers of $M_{\mathbf{q}}$ are thus of order 2. They are given by :

$$\begin{aligned} i [S, H_{el-ph}] &= \frac{1}{2} [S, \underbrace{[S, H_{el} + H_{ph}]}_{iH_{el-ph}}] \\ &= \frac{i}{2} [S, H_{el-ph}] = \frac{1}{2} [H_{el-ph}, \frac{S}{i}] \end{aligned}$$

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:

$$a_{\pm\mathbf{q}}^{(\dagger)} a_{\pm\mathbf{q}'}^{(\dagger)} c_{\mathbf{k}'+\mathbf{q}', \sigma'}^\dagger c_{\mathbf{k}', \sigma'} c_{\mathbf{k}+\mathbf{q}, \sigma}^\dagger c_{\mathbf{k}, \sigma}$$

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 $T \rightarrow 0$.

Let us write:

$$\begin{aligned} H_{el-ph} &= \sum_{\mathbf{k}_1, \mathbf{q}_1, \sigma_1} M_{\mathbf{q}_1} a_{-\mathbf{q}_1}^\dagger c_{\mathbf{k}_1+\mathbf{q}_1, \sigma_1}^\dagger c_{\mathbf{k}_1, \sigma_1} \\ &\quad + \sum_{\mathbf{k}_1, \mathbf{q}_1, \sigma_1} M_{\mathbf{q}_1} a_{\mathbf{q}_1} c_{\mathbf{k}_1+\mathbf{q}_1, \sigma_1}^\dagger c_{\mathbf{k}_1, \sigma_1} \\ \frac{S}{i} &= \sum_{\mathbf{k}_2, \mathbf{q}_2, \sigma_2} \frac{M_{\mathbf{q}_2}}{\epsilon_{\mathbf{k}_2} - \epsilon_{\mathbf{k}_2+\mathbf{q}_2} - \hbar\omega_{\mathbf{q}_2}} a_{-\mathbf{q}_2}^\dagger c_{\mathbf{k}_2+\mathbf{q}_2, \sigma_2}^\dagger c_{\mathbf{k}_2, \sigma_2} \\ &\quad + \sum_{\mathbf{k}_2, \mathbf{q}_2, \sigma_2} \frac{M_{\mathbf{q}_2}}{\epsilon_{\mathbf{k}_2} - \epsilon_{\mathbf{k}_2+\mathbf{q}_2} + \hbar\omega_{\mathbf{q}_2}} a_{\mathbf{q}_2} c_{\mathbf{k}_2+\mathbf{q}_2, \sigma_2}^\dagger c_{\mathbf{k}_2, \sigma_2} \end{aligned}$$

Two terms lead to an expression that does not depend on (a^\dagger, a) :

$$\begin{aligned}
& [a_{-\mathbf{q}_1}^\dagger c_{\mathbf{k}_1+\mathbf{q}_1, \sigma_1}^\dagger c_{\mathbf{k}_1, \sigma_1}, a_{-\mathbf{q}_1} c_{\mathbf{k}_2-\mathbf{q}_1, \sigma_2}^\dagger c_{\mathbf{k}_2, \sigma_2}] \\
= & a_{-\mathbf{q}_1}^\dagger a_{-\mathbf{q}_1} c_{\mathbf{k}_1+\mathbf{q}_1, \sigma_1}^\dagger c_{\mathbf{k}_1, \sigma_1} c_{\mathbf{k}_2-\mathbf{q}_1, \sigma_2}^\dagger c_{\mathbf{k}_2, \sigma_2} \\
- & \underbrace{a_{-\mathbf{q}_1}^\dagger a_{-\mathbf{q}_1}^\dagger}_{1+a_{-\mathbf{q}_1}^\dagger a_{-\mathbf{q}_1}} c_{\mathbf{k}_2-\mathbf{q}_1, \sigma_2}^\dagger c_{\mathbf{k}_2, \sigma_2} c_{\mathbf{k}_1+\mathbf{q}_1, \sigma_1}^\dagger c_{\mathbf{k}_1, \sigma_1} \\
= & a_{-\mathbf{q}_1}^\dagger a_{-\mathbf{q}_1} [c_{\mathbf{k}_1+\mathbf{q}_1, \sigma_1}^\dagger c_{\mathbf{k}_1, \sigma_1}, c_{\mathbf{k}_2-\mathbf{q}_1, \sigma_2}^\dagger c_{\mathbf{k}_2, \sigma_2}] - c_{\mathbf{k}_2-\mathbf{q}_1, \sigma_2}^\dagger c_{\mathbf{k}_2, \sigma_2} c_{\mathbf{k}_1+\mathbf{q}_1, \sigma_1}^\dagger c_{\mathbf{k}_1, \sigma_1}
\end{aligned}$$

Similarly,

$$\begin{aligned}
& [a_{\mathbf{q}_1} c_{\mathbf{k}_1+\mathbf{q}_1, \sigma_1}^\dagger c_{\mathbf{k}_1, \sigma_1}, a_{\mathbf{q}_1}^\dagger c_{\mathbf{k}_2-\mathbf{q}_1, \sigma_2}^\dagger c_{\mathbf{k}_2, \sigma_2}] \\
= & a_{\mathbf{q}_1}^\dagger a_{\mathbf{q}_1} [c_{\mathbf{k}_1+\mathbf{q}_1, \sigma_1}^\dagger c_{\mathbf{k}_1, \sigma_1}, c_{\mathbf{k}_2-\mathbf{q}_1, \sigma_2}^\dagger c_{\mathbf{k}_2, \sigma_2}] + c_{\mathbf{k}_1+\mathbf{q}_1, \sigma_1}^\dagger c_{\mathbf{k}_1, \sigma_1} c_{\mathbf{k}_2-\mathbf{q}_1, \sigma_2}^\dagger c_{\mathbf{k}_2, \sigma_2} \\
\Rightarrow H_{eff}^{(2)} = & \frac{1}{2} \sum_{\mathbf{k}_1, \mathbf{k}_2, \mathbf{q}, \sigma_1, \sigma_2} \left(\frac{M_{\mathbf{q}}^2}{\epsilon_{\mathbf{k}_2} - \epsilon_{\mathbf{k}_2-\mathbf{q}} + \hbar\omega_{\mathbf{q}}} (-c_{\mathbf{k}_2-\mathbf{q}, \sigma_2}^\dagger c_{\mathbf{k}_2, \sigma_2} c_{\mathbf{k}_1+\mathbf{q}, \sigma_1}^\dagger c_{\mathbf{k}_1, \sigma_1}) \right. \\
& \left. + \frac{M_{\mathbf{q}}^2}{\epsilon_{\mathbf{k}_2} - \epsilon_{\mathbf{k}_2-\mathbf{q}} - \hbar\omega_{\mathbf{q}}} (c_{\mathbf{k}_1+\mathbf{q}, \sigma_1}^\dagger c_{\mathbf{k}_1, \sigma_1} c_{\mathbf{k}_2-\mathbf{q}, \sigma_2}^\dagger c_{\mathbf{k}_2, \sigma_2}) \right)
\end{aligned}$$

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 $\epsilon_{\mathbf{k}}$), one gets:

$$H_{eff}^{(2)} = \frac{1}{2} \sum_{\mathbf{k}_1, \mathbf{k}_2, \mathbf{q}, \sigma_1, \sigma_2} V_{\mathbf{k}_2, \mathbf{q}} c_{\mathbf{k}_1+\mathbf{q}, \sigma_1}^\dagger c_{\mathbf{k}_1, \sigma_1} c_{\mathbf{k}_2-\mathbf{q}, \sigma_2}^\dagger c_{\mathbf{k}_2, \sigma_2}$$

with

$$\begin{aligned}
V_{\mathbf{k}_2, \mathbf{q}} &= M_{\mathbf{q}}^2 \left(\frac{1}{\epsilon_{\mathbf{k}_2} - \epsilon_{\mathbf{k}_2-\mathbf{q}} - \hbar\omega_{\mathbf{q}}} - \frac{1}{\epsilon_{\mathbf{k}_2} - \epsilon_{\mathbf{k}_2-\mathbf{q}} + \hbar\omega_{\mathbf{q}}} \right) \\
&= M_{\mathbf{q}}^2 \frac{2\hbar\omega_{\mathbf{q}}}{(\epsilon_{\mathbf{k}_2} - \epsilon_{\mathbf{k}_2-\mathbf{q}})^2 - (\hbar\omega_{\mathbf{q}})^2}
\end{aligned}$$

Finally, the effective Hamiltonian takes the form:

$$H_{eff}^{(2)} = \frac{1}{2} \sum_{\mathbf{k}, \mathbf{k}', \mathbf{q}, \sigma, \sigma'} V_{\mathbf{k}, \mathbf{q}} c_{\mathbf{k}+\mathbf{q}, \sigma}^\dagger c_{\mathbf{k}', \sigma'}^\dagger c_{\mathbf{k}', \sigma'} c_{\mathbf{k}, \sigma} \quad (5.9)$$

with

$$V_{\mathbf{k}, \mathbf{q}} = M_{\mathbf{q}}^2 \frac{2\hbar\omega_{\mathbf{q}}}{(\epsilon_{\mathbf{k}} - \epsilon_{\mathbf{k}+\mathbf{q}})^2 - (\hbar\omega_{\mathbf{q}})^2}$$

With this ordering of the fermionic operators, the interaction is repulsive if the coefficient is positive, and it is attractive if the coefficient is negative (see the case of

the Coulomb repulsion in the appendix). But $V_{\mathbf{k},\mathbf{q}} < 0$ if $|\epsilon_{\mathbf{k}} - \epsilon_{\mathbf{k}+\mathbf{q}}| < \hbar\omega_{\mathbf{q}}$! 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 6

Superconductivity

6.1 Introduction

Superconductivity is probably the most remarkable phenomenon ever observed in solids. The most spectacular aspect is the *rigorous* absence of resistance below a critical temperature. In view of the effect of phonons on transport, the absence of T^3 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, we write it as

$$H_{eff}^{(2)} = -\frac{V}{2N} \sum_{\mathbf{k}, \mathbf{k}', \mathbf{q}, \sigma, \sigma'} f(\mathbf{k}, \mathbf{k}', \mathbf{q}) c_{\mathbf{k}+\mathbf{q}, \sigma}^\dagger c_{\mathbf{k}'-\mathbf{q}, \sigma'}^\dagger c_{\mathbf{k}', \sigma'} c_{\mathbf{k}, \sigma} \quad (6.1)$$

with $V > 0$, and $f(\mathbf{k}, \mathbf{k}', \mathbf{q}) = 1$ if $|\epsilon_{\mathbf{k}+\mathbf{q}} - \epsilon_F|, |\epsilon_{\mathbf{k}} - \epsilon_F|, |\epsilon_{\mathbf{k}'-\mathbf{q}} - \epsilon_F|, |\epsilon_{\mathbf{k}'} - \epsilon_F| < \hbar\omega_D$ and 0 otherwise. In other words, we assume 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 exercise), one can show that attraction induces a bound state that corresponds to a pair of electrons. 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 wavefunction: the antisymmetrization does not necessarily make the resulting wavefunction vanish.

The analogy between the superconducting transition and the Bose-Einstein condensation is clear: the new ground state will 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. It is the latter approach we have adopted in these lectures. The variational calculation will be done in the exercises.

6.2 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 $\langle c_{\mathbf{k},\uparrow}^\dagger c_{-\mathbf{k},\downarrow}^\dagger \rangle$ (and their hermitian conjugate, pair annihilation operators) to decouple the Hamiltonian. Note that it does explicitly break the $U(1)$ charge conservation of the Hamiltonian. It is nonetheless required in order to simplify computations. Now, we 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 approximation is not universally valid: exotic superconducting materials breaking any or both of these assumptions exist.

In any case, we work with the reduced Hamiltonian (also called BCS Hamiltonian)

:

$$H_{\text{red}} = \sum_{\mathbf{k},\sigma} \epsilon_{\mathbf{k}} c_{\mathbf{k},\sigma}^\dagger c_{\mathbf{k},\sigma} - \frac{V}{N} \sum_{\mathbf{k},\mathbf{k}'} c_{\mathbf{k},\uparrow}^\dagger c_{-\mathbf{k},\downarrow}^\dagger c_{-\mathbf{k}',\downarrow} c_{\mathbf{k}',\uparrow} \quad (6.2)$$

where one has replaced the sum over σ by a factor 2. We have dropped the function f for simplicity. It could in principle be included, as $f(\mathbf{k})$ is compatible with our hypothesis.

We use a minimal mean field decoupling:

$$\begin{aligned} c_{\mathbf{k},\uparrow}^\dagger c_{-\mathbf{k},\downarrow}^\dagger c_{-\mathbf{k}',\downarrow} c_{\mathbf{k}',\uparrow} &\simeq \langle c_{\mathbf{k},\uparrow}^\dagger c_{-\mathbf{k},\downarrow}^\dagger \rangle c_{-\mathbf{k}',\downarrow} c_{\mathbf{k}',\uparrow} + c_{\mathbf{k},\uparrow}^\dagger c_{-\mathbf{k},\downarrow}^\dagger \langle c_{-\mathbf{k}',\downarrow} c_{\mathbf{k}',\uparrow} \rangle \\ &- \langle c_{\mathbf{k},\uparrow}^\dagger c_{-\mathbf{k},\downarrow}^\dagger \rangle \langle c_{-\mathbf{k}',\downarrow} c_{\mathbf{k}',\uparrow} \rangle \end{aligned}$$

Note that we could be more complete, and also introduce also a density-density decoupling, without significantly changing the results. It would mainly contribute

as an effective self-consistent chemical potential. We introduce the order parameter Δ , assumed to be real for simplicity (a gauge transform allows us to proceed to this simplification):

$$\Delta = \frac{V}{N} \sum_{\mathbf{k}} \langle c_{\mathbf{k},\uparrow}^\dagger c_{-\mathbf{k},\downarrow}^\dagger \rangle = \frac{V}{N} \sum_{\mathbf{k}} \langle c_{-\mathbf{k},\downarrow} c_{\mathbf{k},\uparrow} \rangle \quad (6.3)$$

leads to the mean-field Hamiltonian

$$H_{MF} = \sum_{\mathbf{k},\sigma} \epsilon_{\mathbf{k}} c_{\mathbf{k},\sigma}^\dagger c_{\mathbf{k},\sigma} - \Delta \sum_{\mathbf{k}} (c_{\mathbf{k},\uparrow}^\dagger c_{-\mathbf{k},\downarrow}^\dagger + c_{-\mathbf{k},\downarrow} c_{\mathbf{k},\uparrow}) + \frac{N\Delta^2}{V}. \quad (6.4)$$

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:

1. Bogoliubov transformation to diagonalize the Hamiltonian,
2. Calculation of the ground state of this Hamiltonian,
3. Calculation of the expectation values $\langle c_{\mathbf{k},\sigma}^\dagger c_{-\mathbf{k},-\sigma}^\dagger \rangle$ and resolution the self-consistent equations for the parameters Δ .

6.2.1 Bogoliubov transformation

We provide in this section two equivalent derivations of the Bogoliubov transformation. The first one follows the method we used for bosonic models. The second one takes advantage of the simplified fermionic algebra.

First derivation

We look for fermionic operators defined by

$$\begin{cases} \alpha_{\mathbf{k},\uparrow}^\dagger = u_{\mathbf{k}} c_{\mathbf{k},\uparrow}^\dagger - v_{\mathbf{k}} c_{-\mathbf{k},\downarrow} \\ \alpha_{-\mathbf{k},\downarrow} = u_{\mathbf{k}} c_{-\mathbf{k},\downarrow} + v_{\mathbf{k}} c_{\mathbf{k},\uparrow}^\dagger \end{cases}$$

If the Hamiltonian is quadratic in these operators with energy $E_{\mathbf{k}}$, then the commutators of H with these operators must be given by

$$\begin{cases} [H, \alpha_{\mathbf{k},\uparrow}^\dagger] = E_{\mathbf{k}} \alpha_{\mathbf{k},\uparrow}^\dagger \\ [H, \alpha_{-\mathbf{k},\downarrow}] = -E_{\mathbf{k}} \alpha_{-\mathbf{k},\downarrow} \end{cases}$$

These equations can be rewritten in terms of the original fermionic operators. A straightforward calculation of the commutators then leads to

$$\begin{cases} -\epsilon_{\mathbf{k}}u_{\mathbf{k}} - \Delta v_{\mathbf{k}} = -E_{\mathbf{k}}u_{\mathbf{k}} \\ \Delta u_{\mathbf{k}} - \epsilon_{\mathbf{k}}v_{\mathbf{k}} = E_{\mathbf{k}}v_{\mathbf{k}} \end{cases}$$

or

$$\begin{cases} (E_{\mathbf{k}} - \epsilon_{\mathbf{k}})u_{\mathbf{k}} - \Delta v_{\mathbf{k}} = 0 \\ \Delta u_{\mathbf{k}} - (E_{\mathbf{k}} + \epsilon_{\mathbf{k}})v_{\mathbf{k}} = 0 \end{cases}$$

$$\Rightarrow -(E_{\mathbf{k}}^2 + \epsilon_{\mathbf{k}}^2) + \Delta^2 = 0$$

$$\Rightarrow E_{\mathbf{k}} = \sqrt{\epsilon_{\mathbf{k}}^2 + \Delta^2}$$

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 $\alpha_{\mathbf{k},\uparrow}^\dagger$ and $\alpha_{-\mathbf{k},\downarrow}$ to be fermionic, the coefficients $u_{\mathbf{k}}$ and $v_{\mathbf{k}}$ must satisfy the normalization condition

$$u_{\mathbf{k}}^2 + v_{\mathbf{k}}^2 = 1$$

But from the above system of equations they are also related by

$$v_{\mathbf{k}} = \frac{\Delta}{E_{\mathbf{k}} + \epsilon_{\mathbf{k}}}u_{\mathbf{k}}$$

$$\Rightarrow u_{\mathbf{k}}^2 \left(1 + \frac{\Delta^2}{(E_{\mathbf{k}} + \epsilon_{\mathbf{k}})^2} \right) = 1$$

$$\Rightarrow u_{\mathbf{k}}^2 = \frac{1}{1 + \frac{\Delta^2}{(E_{\mathbf{k}} + \epsilon_{\mathbf{k}})^2}}$$

$$= \frac{(E_{\mathbf{k}} + \epsilon_{\mathbf{k}})^2}{(E_{\mathbf{k}} + \epsilon_{\mathbf{k}})^2 + \Delta^2}$$

$$= \frac{(E_{\mathbf{k}} + \epsilon_{\mathbf{k}})^2}{2E_{\mathbf{k}}^2 + 2E_{\mathbf{k}}\epsilon_{\mathbf{k}}}$$

$$= \frac{1}{2} \left(1 + \frac{\epsilon_{\mathbf{k}}}{E_{\mathbf{k}}} \right)$$

Similarly

$$v_{\mathbf{k}}^2 = 1 - u_{\mathbf{k}}^2 = \frac{1}{2} \left(1 - \frac{\epsilon_{\mathbf{k}}}{E_{\mathbf{k}}} \right)$$

Finally, the Hamiltonian takes the form

$$H = \sum_{\mathbf{k}} E_{\mathbf{k}} (\alpha_{\mathbf{k},\uparrow}^\dagger \alpha_{\mathbf{k},\uparrow} + \alpha_{-\mathbf{k},\downarrow}^\dagger \alpha_{-\mathbf{k},\downarrow}) + Cst$$

In the following, unless there is an ambiguity, we will always associate \mathbf{k} with a spin \uparrow and $-\mathbf{k}$ with a spin \downarrow , and we will adopt the compact notations :

$$\begin{aligned}\alpha_{\mathbf{k}}^{\dagger} &\equiv \alpha_{\mathbf{k},\uparrow}^{\dagger} \\ \alpha_{\mathbf{k}} &\equiv \alpha_{\mathbf{k},\uparrow} \\ \alpha_{-\mathbf{k}}^{\dagger} &\equiv \alpha_{-\mathbf{k},\downarrow}^{\dagger} \\ \alpha_{-\mathbf{k}} &\equiv \alpha_{-\mathbf{k},\downarrow}\end{aligned}$$

and similarly the c^{\dagger} and c operators.

Second derivation

This derivation is more straightforward to implement numerically. First, we note that the Hamiltonian couples fermions with quantum number (\mathbf{k}, \uparrow) and $(-\mathbf{k}, \downarrow)$. We therefore can rewrite the Hamiltonian as

$$H_{MF} = \frac{N\Delta^2}{V} + \sum_{\mathbf{k}} \left(\epsilon_{\mathbf{k}} c_{\mathbf{k},\uparrow}^{\dagger} c_{\mathbf{k},\uparrow} + \epsilon_{-\mathbf{k}} c_{-\mathbf{k},\downarrow}^{\dagger} c_{-\mathbf{k},\downarrow} - \Delta (c_{\mathbf{k},\uparrow}^{\dagger} c_{-\mathbf{k},\downarrow}^{\dagger} + c_{-\mathbf{k},\downarrow} c_{\mathbf{k},\uparrow}) \right). \quad (6.5)$$

With this resummation, each term in the sum is entirely independent: operators at a given (\mathbf{k}, σ) only occurs in a single element of the sum. This transformation is *crucial* to avoid errors, as it automatically enforces the effective particle-hole symmetry required in the Bogoliubov formalism.

We can then define the Nambu spinors $\Psi_{\mathbf{k}}^{\dagger} = (c_{\mathbf{k},\uparrow}^{\dagger}, c_{-\mathbf{k},\downarrow})$, and the effective Hamiltonian

$$h(k) = \begin{pmatrix} \epsilon_{\mathbf{k}} & -\Delta \\ -\Delta & -\epsilon_{-\mathbf{k}} \end{pmatrix},$$

such that

$$H_{MF} = \frac{N\Delta^2}{V} + \sum_{\mathbf{k}} \epsilon_{-\mathbf{k}} + \Psi_{\mathbf{k}}^{\dagger} h(\mathbf{k}) \Psi_{\mathbf{k}}. \quad (6.6)$$

We assume inversion symmetry for the kinetic energy, and therefore $\epsilon_{\mathbf{k}} = \epsilon_{-\mathbf{k}}$ (again this is not fundamental, but simplify computations). We can then rewrite

$$h(\mathbf{k}) = \epsilon_{\mathbf{k}} \sigma^z - \Delta \sigma^x,$$

whose spectrum is $\pm E_{\mathbf{k}}$ with $E_{\mathbf{k}} = \sqrt{\epsilon_{\mathbf{k}}^2 + \Delta^2}$. To obtain the eigenvalues, we can note

$$h(\mathbf{k}) = E_{\mathbf{k}} \begin{pmatrix} \cos \theta_{\mathbf{k}} & -\sin \theta_{\mathbf{k}} \\ -\sin \theta_{\mathbf{k}} & -\cos \theta_{\mathbf{k}} \end{pmatrix}, \text{ with } \cos \theta_{\mathbf{k}} = \frac{\epsilon_{\mathbf{k}}}{E_{\mathbf{k}}} \text{ and } \sin \theta_{\mathbf{k}} = \frac{\Delta}{E_{\mathbf{k}}}.$$

We know that the two eigenvectors are simply $(\cos \theta_{\mathbf{k}}/2, -\sin \theta_{\mathbf{k}}/2)$ and $(\sin \theta_{\mathbf{k}}/2, \cos \theta_{\mathbf{k}}/2)$, and the matrix

$$U_{\mathbf{k}} = \begin{pmatrix} \cos \theta_{\mathbf{k}}/2 & \sin \theta_{\mathbf{k}}/2 \\ -\sin \theta_{\mathbf{k}}/2 & \cos \theta_{\mathbf{k}}/2 \end{pmatrix},$$

such that

$$H_{MF} = \frac{N\Delta^2}{V} + \sum_{\mathbf{k}} \varepsilon_{-\mathbf{k}} + \Psi_{\mathbf{k}}^\dagger U_{\mathbf{k}} \begin{pmatrix} E_{\mathbf{k}} & 0 \\ 0 & E_{-\mathbf{k}} \end{pmatrix} U_{\mathbf{k}}^\dagger \Psi_{\mathbf{k}}.$$

We can therefore define $(\alpha_{\mathbf{k},\uparrow}^\dagger, \alpha_{-\mathbf{k},\downarrow}) = \Psi_{\mathbf{k}} U_{\mathbf{k}}$, and obtain

$$H_{MF} = \frac{N\Delta^2}{V} + \sum_{\mathbf{k}} (\varepsilon_{-\mathbf{k}} - E_{\mathbf{k}}) + E_{\mathbf{k}} (\alpha_{\mathbf{k},\uparrow}^\dagger \alpha_{\mathbf{k},\uparrow} + \alpha_{-\mathbf{k},\downarrow}^\dagger \alpha_{-\mathbf{k},\downarrow}).$$

We also recover

$$u_{\mathbf{k}} = \cos \theta_{\mathbf{k}}/2 = \frac{1}{\sqrt{2}} \sqrt{1 + \frac{\varepsilon_{\mathbf{k}}}{E_{\mathbf{k}}}}$$

$$v_{\mathbf{k}} = \sin \theta_{\mathbf{k}}/2 = \frac{1}{\sqrt{2}} \sqrt{1 - \frac{\varepsilon_{\mathbf{k}}}{E_{\mathbf{k}}}}.$$

6.2.2 Ground state

In the following, unless there is an ambiguity, we will always associate \mathbf{k} with a spin \uparrow and $-\mathbf{k}$ with a spin \downarrow , and we will adopt the compact notations :

$$\alpha_{\mathbf{k}} \equiv \alpha_{\mathbf{k},\uparrow}$$

$$\alpha_{-\mathbf{k}} \equiv \alpha_{-\mathbf{k},\downarrow}$$

and similarly with the c operators.

With these notations, the effective Hamiltonian after decoupling reads:

$$H = \sum_{\mathbf{k}} E_{\mathbf{k}} (\alpha_{\mathbf{k}}^\dagger \alpha_{\mathbf{k}} + \alpha_{-\mathbf{k}}^\dagger \alpha_{-\mathbf{k}}) + Cst$$

Since $E_{\mathbf{k}} > 0$, the ground state corresponds to the vacuum of the $\alpha_{\mathbf{k}}^\dagger$ particles, i.e. it must satisfy for all \mathbf{k} :

$$\alpha_{\mathbf{k}} |\phi_0\rangle = 0$$

$$\alpha_{-\mathbf{k}} |\phi_0\rangle = 0.$$

If one denotes by $|0_c\rangle$ the vacuum of the $c_{\mathbf{k},\sigma}^\dagger$ particles, the state

$$\prod_{\mathbf{k}} \alpha_{\mathbf{k}} \alpha_{-\mathbf{k}} |0\rangle$$

satisfies these conditions since the $\alpha_{\mathbf{k}}$ operators are fermions. In the original fermion basis, the normalized ground state is given by:

$$|\phi_0\rangle = |0_\alpha\rangle = \prod_{\mathbf{k}} (u_{\mathbf{k}} + v_{\mathbf{k}} c_{\mathbf{k}}^\dagger c_{-\mathbf{k}}^\dagger) |0_c\rangle \quad (6.7)$$

Proof: we develop the formula as:

$$\begin{aligned}
\prod_{\mathbf{k}} \alpha_{\mathbf{k}} \alpha_{-\mathbf{k}} |0\rangle &= \prod_{\mathbf{k}} (u_{\mathbf{k}} c_{\mathbf{k}} - v_{\mathbf{k}} c_{-\mathbf{k}}^{\dagger}) (u_{\mathbf{k}} c_{-\mathbf{k}} + v_{\mathbf{k}} c_{\mathbf{k}}^{\dagger}) |0\rangle \\
&= \prod_{\mathbf{k}} (u_{\mathbf{k}}^2 c_{\mathbf{k}} c_{-\mathbf{k}} + u_{\mathbf{k}} v_{\mathbf{k}} c_{\mathbf{k}} c_{\mathbf{k}}^{\dagger} - v_{\mathbf{k}} u_{\mathbf{k}} c_{-\mathbf{k}}^{\dagger} c_{-\mathbf{k}} - v_{\mathbf{k}}^2 c_{-\mathbf{k}}^{\dagger} c_{\mathbf{k}}^{\dagger}) |0\rangle \\
&= \prod_{\mathbf{k}} (u_{\mathbf{k}} v_{\mathbf{k}} + v_{\mathbf{k}}^2 c_{\mathbf{k}}^{\dagger} c_{-\mathbf{k}}^{\dagger}) |0\rangle \\
&= \prod_{\mathbf{k}} v_{\mathbf{k}} (u_{\mathbf{k}} + v_{\mathbf{k}} c_{\mathbf{k}}^{\dagger} c_{-\mathbf{k}}^{\dagger}) |0\rangle
\end{aligned}$$

Besides, this wave function must be normalized :

$$\begin{aligned}
\langle 0 | \prod_{\mathbf{k}} \alpha_{-\mathbf{k}}^{\dagger} \alpha_{\mathbf{k}}^{\dagger} \alpha_{\mathbf{k}} \alpha_{-\mathbf{k}} |0\rangle &= \langle 0 | \prod_{\mathbf{k}} v_{\mathbf{k}}^2 (u_{\mathbf{k}} + v_{\mathbf{k}} c_{-\mathbf{k}} c_{\mathbf{k}}) (u_{\mathbf{k}} + v_{\mathbf{k}} c_{\mathbf{k}}^{\dagger} c_{-\mathbf{k}}^{\dagger}) |0\rangle \\
&= \langle 0 | \prod_{\mathbf{k}} v_{\mathbf{k}}^2 (u_{\mathbf{k}}^2 + v_{\mathbf{k}}^2 c_{-\mathbf{k}} c_{\mathbf{k}} c_{\mathbf{k}}^{\dagger} c_{-\mathbf{k}}^{\dagger}) |0\rangle \\
&= \prod_{\mathbf{k}} v_{\mathbf{k}}^2 (u_{\mathbf{k}}^2 + v_{\mathbf{k}}^2) = \prod_{\mathbf{k}} v_{\mathbf{k}}^2
\end{aligned}$$

because $c_{-\mathbf{k}} c_{\mathbf{k}} c_{\mathbf{k}}^{\dagger} c_{-\mathbf{k}}^{\dagger} |0\rangle = |0\rangle$.

6.2.3 Self-consistent equations

It is clear that this ground states is very different from a Fermi sea. In particular, it mixes states with different numbers of particles. That's why expectation values such as $\langle c_{\mathbf{k}}^{\dagger} c_{-\mathbf{k}}^{\dagger} \rangle$ can be non-vanishing.

Let us now show that

$$\Delta = \frac{V}{N} \sum_{\mathbf{k}} \langle c_{-\mathbf{k}} c_{\mathbf{k}} \rangle = \frac{V}{N} \sum_{\mathbf{k}} u_{\mathbf{k}} v_{\mathbf{k}} = \frac{V}{2N} \Delta \sum_{\mathbf{k}} \frac{1}{E_{\mathbf{k}}} \quad (6.8)$$

Proof:

$$\begin{aligned}
\langle c_{-\mathbf{k}} c_{\mathbf{k}} \rangle &= \langle 0 | \prod_l (u_l + v_l c_{-l} c_l) c_{-\mathbf{k}} c_{\mathbf{k}} \prod_{l'} (u_{l'} + v_{l'} c_{l'}^{\dagger} c_{-l'}^{\dagger}) |0\rangle \\
&= \langle 0 | u_{\mathbf{k}} c_{-\mathbf{k}} c_{\mathbf{k}} \prod_{l \neq \mathbf{k}} (u_l + v_l c_{-l} c_l) \prod_{l'} (u_{l'} + v_{l'} c_{l'}^{\dagger} c_{-l'}^{\dagger}) |0\rangle \\
&= \langle 0 | u_{\mathbf{k}} c_{-\mathbf{k}} c_{\mathbf{k}} (u_{\mathbf{k}} + v_{\mathbf{k}} c_{\mathbf{k}}^{\dagger} c_{-\mathbf{k}}^{\dagger}) \prod_{l \neq \mathbf{k}} (u_l + v_l c_{-l} c_l) \prod_{l' \neq \mathbf{k}} (u_{l'} + v_{l'} c_{l'}^{\dagger} c_{-l'}^{\dagger}) |0\rangle \\
&= \langle 0 | (u_{\mathbf{k}} c_{-\mathbf{k}} c_{\mathbf{k}} v_{\mathbf{k}} c_{\mathbf{k}}^{\dagger} c_{-\mathbf{k}}^{\dagger}) |0\rangle \\
&= u_{\mathbf{k}} v_{\mathbf{k}}
\end{aligned}$$

We obtain

$$\Delta = \frac{V}{N} \sum_{\mathbf{k}} u_{\mathbf{k}} v_{\mathbf{k}}$$

But

$$u_{\mathbf{k}} v_{\mathbf{k}} = \cos \frac{\theta_{\mathbf{k}}}{2} \sin \frac{\theta_{\mathbf{k}}}{2} = \frac{1}{2} \sin \theta_{\mathbf{k}} = \frac{1}{2} \frac{\Delta}{E_{\mathbf{k}}}$$

Therefore

$$\Delta = \frac{V}{2N} \Delta \sum_{\mathbf{k}} \frac{1}{\sqrt{\epsilon_{\mathbf{k}}^2 + \Delta^2}}$$

The solution $\Delta = 0$ of course corresponds to the normal state. This equation possesses another non-zero solution:

$$\begin{aligned} 1 &= \frac{V}{2N} \sum_{\mathbf{k}} \frac{1}{\sqrt{\epsilon_{\mathbf{k}}^2 + \Delta^2}} \\ 1 &= \frac{V}{2} \int g(\epsilon) \frac{1}{\sqrt{\epsilon^2 + \Delta^2}} d\epsilon \end{aligned}$$

In the limit of weak Δ , this implies that

$$\boxed{\Delta \approx 2\omega_D e^{\frac{-1}{Vg(\epsilon_F)}},} \quad (6.9)$$

where ω_D is the Debye frequency.

The integral is limited to energies close to ϵ_F (within the Debye frequency ω_D) for which $g(\epsilon) \simeq g(\epsilon_F)$, if we remember the introduction of this chapter.

$$\begin{aligned} 1 &\simeq \frac{V}{2} g(\epsilon_F) \int_{-\omega_D}^{+\omega_D} \frac{1}{\sqrt{\epsilon^2 + \Delta^2}} d\epsilon \\ 1 &\simeq \frac{V}{2} g(\epsilon_F) \left[\ln(\epsilon + \sqrt{\epsilon^2 + \Delta^2}) \right]_{-\omega_D}^{+\omega_D} \\ 1 &\simeq \frac{V}{2} g(\epsilon_F) \ln \left[\frac{\omega_D + \sqrt{\omega_D^2 + \Delta^2}}{-\omega_D + \sqrt{\omega_D^2 + \Delta^2}} \right] \end{aligned}$$

Let us suppose that Δ is small as compared to ω_D (to be checked a posteriori)

$$\begin{aligned}
1 &\simeq \frac{V}{2}g(\epsilon_F) \ln \left[\frac{\omega_D + \omega_D \left(1 + \frac{\Delta^2}{2\omega_D^2}\right)}{\frac{\Delta^2}{2\omega_D}} \right] \\
1 &\simeq \frac{V}{2}g(\epsilon_F) \ln \left(\frac{4\omega_D^2}{\Delta^2} \right) \\
1 &\simeq \frac{V}{2}g(\epsilon_F) 2 \ln \left(\frac{2\omega_D}{\Delta} \right) \\
\frac{1}{Vg(\epsilon_F)} &= \ln \left(\frac{2\omega_D}{\Delta} \right) \\
\Delta &= 2\omega_D e^{\frac{-1}{Vg(\epsilon_F)}}
\end{aligned}$$

Importantly, the parameter Δ 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 $\prod_{\mathbf{k}}(u_{\mathbf{k}} + v_{\mathbf{k}}c_{\mathbf{k}}^{\dagger}c_{-\mathbf{k}}^{\dagger})|0\rangle$, and to minimize the energy with respect to $u_{\mathbf{k}}$ and $v_{\mathbf{k}}$. This variational method leads to the same results.

6.3 Interpretation of the results

To understand the physical implications of the existence of the solution with $\Delta \neq 0$, we must:

1. Compare its energy to the solution with $\Delta = 0$.
2. Come up with a physical interpretation of the wave function.
3. Extract from the solution the elementary excitations and the behaviour at finite temperature.

6.3.1 Condensation energy

To calculate the energy of both solutions, let us go back to the mean-field Hamiltonian

$$H_{MF} = \sum_{\mathbf{k},\sigma} \epsilon_{\mathbf{k}} c_{\mathbf{k},\sigma}^{\dagger} c_{\mathbf{k},\sigma} - \Delta \sum_{\mathbf{k}} (c_{\mathbf{k},\uparrow}^{\dagger} c_{-\mathbf{k},\downarrow}^{\dagger} + c_{-\mathbf{k},\downarrow} c_{\mathbf{k},\uparrow}) + \frac{N\Delta^2}{V}.$$

The ground state energy can be computed as

$$\langle 0_\alpha | H_{MF} | 0_\alpha \rangle = \sum_{\mathbf{k}, \sigma} \epsilon_{\mathbf{k}} \langle 0_\alpha | c_{\mathbf{k}, \sigma}^\dagger c_{\mathbf{k}, \sigma} | 0_\alpha \rangle - \Delta \sum_{\mathbf{k}} \langle 0_\alpha | c_{\mathbf{k}, \uparrow}^\dagger c_{-\mathbf{k}, \downarrow}^\dagger | 0_\alpha \rangle + \langle 0_\alpha | c_{-\mathbf{k}, \downarrow} c_{\mathbf{k}, \uparrow} | 0_\alpha \rangle + \frac{N\Delta^2}{V},$$

where $|0_\alpha\rangle$ is the empty state of the α fermions.

To compute the expectation values, we can either use the explicit expression of the groundstate we obtained, or inverse the Bogoliubov transform. Additionally, we could have used directly the expression of the Hamiltonian we obtained.

Direct computation:

$$\begin{aligned} \langle 0_\alpha | c_{\mathbf{k}, \uparrow}^\dagger c_{\mathbf{k}, \uparrow} | 0_\alpha \rangle &= \langle 0_c | \prod_{\mathbf{q}'} (u_{\mathbf{q}'}^* + v_{\mathbf{q}'}^* c_{-\mathbf{q}', \downarrow} c_{\mathbf{q}', \uparrow}) c_{\mathbf{k}, \uparrow}^\dagger c_{\mathbf{k}, \uparrow} \prod_{\mathbf{q}} (u_{\mathbf{q}} + v_{\mathbf{q}} c_{\mathbf{q}, \uparrow}^\dagger c_{-\mathbf{q}, \downarrow}^\dagger) | 0_c \rangle \\ &= \langle 0_c | (u_{\mathbf{k}}^* + v_{\mathbf{k}}^* c_{-\mathbf{k}, \downarrow} c_{\mathbf{k}, \uparrow}) c_{\mathbf{k}, \uparrow}^\dagger c_{\mathbf{k}, \uparrow} (u_{\mathbf{k}} + v_{\mathbf{k}} c_{\mathbf{k}, \uparrow}^\dagger c_{-\mathbf{k}, \downarrow}^\dagger) | 0_c \rangle \\ &= |v_{\mathbf{k}}|^2 \end{aligned}$$

Similarly,

$$\langle 0_\alpha | c_{\mathbf{k}, \downarrow}^\dagger c_{\mathbf{k}, \downarrow} | 0_\alpha \rangle = |v_{-\mathbf{k}}|^2$$

$$\begin{aligned} \langle 0_\alpha | c_{\mathbf{k}, \uparrow}^\dagger c_{-\mathbf{k}, \downarrow}^\dagger | 0_\alpha \rangle &= \langle 0_c | (u_{\mathbf{k}}^* + v_{\mathbf{k}}^* c_{-\mathbf{k}, \downarrow} c_{\mathbf{k}, \uparrow}) c_{\mathbf{k}, \uparrow}^\dagger c_{-\mathbf{k}, \downarrow}^\dagger (u_{\mathbf{k}} + v_{\mathbf{k}} c_{\mathbf{k}, \uparrow}^\dagger c_{-\mathbf{k}, \downarrow}^\dagger) | 0_c \rangle \\ &= u_{\mathbf{k}} v_{\mathbf{k}}^* \end{aligned}$$

$$\langle 0_\alpha | c_{-\mathbf{k}, \downarrow} c_{\mathbf{k}, \uparrow} | 0_\alpha \rangle = u_{\mathbf{k}}^* v_{\mathbf{k}}$$

The groundstate energy, taking into account that u and v are reals, is

$$E = \frac{N\Delta^2}{V} + 2 \sum_{\mathbf{k}} \epsilon_{\mathbf{k}} v_{\mathbf{k}}^2 - 2\Delta \sum_{\mathbf{k}} u_{\mathbf{k}} v_{\mathbf{k}} = \frac{N\Delta^2}{V} + \sum_{\mathbf{k}} \epsilon_{\mathbf{k}} - E_{\mathbf{k}} \quad (6.10)$$

We recover the expression for the Hamiltonian.

Inverse Bogoliubov We invert the Bogoliubov transform to express H_{MF} in terms of $\alpha_{\mathbf{k}, \sigma}^\dagger$:

$$\begin{cases} \alpha_{\mathbf{k}, \uparrow}^\dagger &= u_{\mathbf{k}} c_{\mathbf{k}, \uparrow}^\dagger - v_{\mathbf{k}} c_{-\mathbf{k}, \downarrow} \\ \alpha_{-\mathbf{k}, \downarrow} &= u_{\mathbf{k}} c_{-\mathbf{k}, \downarrow} + v_{\mathbf{k}} c_{\mathbf{k}, \uparrow}^\dagger \end{cases} \Rightarrow \begin{cases} c_{\mathbf{k}, \uparrow}^\dagger &= u_{\mathbf{k}} \alpha_{\mathbf{k}, \uparrow}^\dagger + v_{\mathbf{k}} \alpha_{-\mathbf{k}, \downarrow} \\ c_{\mathbf{k}, \uparrow} &= u_{\mathbf{k}} \alpha_{\mathbf{k}, \uparrow} + v_{\mathbf{k}} \alpha_{-\mathbf{k}, \downarrow}^\dagger \\ c_{-\mathbf{k}, \downarrow}^\dagger &= u_{\mathbf{k}} \alpha_{-\mathbf{k}, \downarrow}^\dagger - v_{\mathbf{k}} \alpha_{\mathbf{k}, \uparrow} \\ c_{-\mathbf{k}, \downarrow} &= u_{\mathbf{k}} \alpha_{-\mathbf{k}, \downarrow} - v_{\mathbf{k}} \alpha_{\mathbf{k}, \uparrow}^\dagger \end{cases}$$

taking into account the fact that $v_{\mathbf{k}}$ changes sign between \mathbf{k}, \uparrow and $-\mathbf{k}, \downarrow$.

$$\begin{aligned} c_{\mathbf{k},\uparrow}^\dagger c_{\mathbf{k},\uparrow} &= (u_{\mathbf{k}}\alpha_{\mathbf{k},\uparrow}^\dagger + v_{\mathbf{k}}\alpha_{-\mathbf{k},\downarrow})(u_{\mathbf{k}}\alpha_{\mathbf{k},\uparrow} + v_{\mathbf{k}}\alpha_{-\mathbf{k},\downarrow}^\dagger) \\ &= u_{\mathbf{k}}^2\alpha_{\mathbf{k},\uparrow}^\dagger\alpha_{\mathbf{k},\uparrow} + v_{\mathbf{k}}^2(1 - \alpha_{-\mathbf{k},\downarrow}^\dagger\alpha_{-\mathbf{k},\downarrow}) \\ &\quad + u_{\mathbf{k}}v_{\mathbf{k}}(\alpha_{\mathbf{k},\uparrow}^\dagger\alpha_{-\mathbf{k},\downarrow}^\dagger + \alpha_{-\mathbf{k},\downarrow}\alpha_{\mathbf{k},\uparrow}) \end{aligned}$$

$$\begin{aligned} c_{\mathbf{k},\downarrow}^\dagger c_{\mathbf{k},\downarrow} &= (u_{\mathbf{k}}\alpha_{\mathbf{k},\downarrow}^\dagger - v_{\mathbf{k}}\alpha_{-\mathbf{k},\uparrow})(u_{\mathbf{k}}\alpha_{\mathbf{k},\downarrow} - v_{\mathbf{k}}\alpha_{-\mathbf{k},\uparrow}^\dagger) \\ &= u_{\mathbf{k}}^2\alpha_{\mathbf{k},\downarrow}^\dagger\alpha_{\mathbf{k},\downarrow} + v_{\mathbf{k}}^2(1 - \alpha_{-\mathbf{k},\uparrow}^\dagger\alpha_{-\mathbf{k},\uparrow}) \\ &\quad - u_{\mathbf{k}}v_{\mathbf{k}}(\alpha_{-\mathbf{k},\uparrow}\alpha_{\mathbf{k},\downarrow} + \alpha_{\mathbf{k},\downarrow}^\dagger\alpha_{-\mathbf{k},\uparrow}^\dagger) \end{aligned}$$

$$\begin{aligned} c_{\mathbf{k},\uparrow}^\dagger c_{-\mathbf{k},\downarrow}^\dagger &= (u_{\mathbf{k}}\alpha_{\mathbf{k},\uparrow}^\dagger + v_{\mathbf{k}}\alpha_{-\mathbf{k},\downarrow})(u_{\mathbf{k}}\alpha_{-\mathbf{k},\downarrow}^\dagger - v_{\mathbf{k}}\alpha_{\mathbf{k},\uparrow}) \\ &= u_{\mathbf{k}}^2\alpha_{\mathbf{k},\uparrow}^\dagger\alpha_{-\mathbf{k},\downarrow}^\dagger - v_{\mathbf{k}}^2\alpha_{-\mathbf{k},\downarrow}\alpha_{\mathbf{k},\uparrow} \\ &\quad - u_{\mathbf{k}}v_{\mathbf{k}}\alpha_{\mathbf{k},\uparrow}^\dagger\alpha_{\mathbf{k},\uparrow} + u_{\mathbf{k}}v_{\mathbf{k}}(1 - \alpha_{-\mathbf{k},\downarrow}^\dagger\alpha_{-\mathbf{k},\downarrow}) \end{aligned}$$

$$\begin{aligned} c_{-\mathbf{k},\downarrow} c_{\mathbf{k},\uparrow} &= (u_{\mathbf{k}}\alpha_{-\mathbf{k},\downarrow} - v_{\mathbf{k}}\alpha_{\mathbf{k},\uparrow}^\dagger)(u_{\mathbf{k}}\alpha_{\mathbf{k},\uparrow} + v_{\mathbf{k}}\alpha_{-\mathbf{k},\downarrow}^\dagger) \\ &= u_{\mathbf{k}}^2\alpha_{-\mathbf{k},\downarrow}\alpha_{\mathbf{k},\uparrow} - v_{\mathbf{k}}^2\alpha_{\mathbf{k},\uparrow}^\dagger\alpha_{-\mathbf{k},\downarrow}^\dagger \\ &\quad + u_{\mathbf{k}}v_{\mathbf{k}}(1 - \alpha_{-\mathbf{k},\downarrow}^\dagger\alpha_{-\mathbf{k},\downarrow}) - u_{\mathbf{k}}v_{\mathbf{k}}\alpha_{\mathbf{k},\uparrow}^\dagger\alpha_{\mathbf{k},\uparrow} \end{aligned}$$

Coefficient of $\alpha_{\mathbf{k},\uparrow}^\dagger\alpha_{\mathbf{k},\uparrow}$:

$$\begin{aligned} &\epsilon_{\mathbf{k}}u_{\mathbf{k}}^2 - \epsilon_{\mathbf{k}}v_{\mathbf{k}}^2 + \Delta u_{\mathbf{k}}v_{\mathbf{k}} + \Delta u_{\mathbf{k}}v_{\mathbf{k}} \\ &= \epsilon_{\mathbf{k}}\frac{\epsilon_{\mathbf{k}}}{E_{\mathbf{k}}} + 2\Delta\sqrt{\frac{1}{4}\left(1 - \frac{\epsilon_{\mathbf{k}}^2}{E_{\mathbf{k}}^2}\right)} \\ &= \frac{\epsilon_{\mathbf{k}}^2}{E_{\mathbf{k}}} + \Delta\sqrt{\frac{\Delta^2}{E_{\mathbf{k}}^2}} = E_{\mathbf{k}} \end{aligned}$$

Coefficient of $\alpha_{\mathbf{k},\uparrow}^\dagger\alpha_{-\mathbf{k},\downarrow}^\dagger$:

$$\begin{aligned} &\epsilon_{\mathbf{k}}u_{\mathbf{k}}v_{\mathbf{k}} + \epsilon_{\mathbf{k}}u_{\mathbf{k}}v_{\mathbf{k}} - \Delta u_{\mathbf{k}}^2 + \Delta v_{\mathbf{k}}^2 \\ &= 2\epsilon_{\mathbf{k}}u_{\mathbf{k}}v_{\mathbf{k}} + \Delta(v_{\mathbf{k}}^2 - u_{\mathbf{k}}^2) \\ &= \epsilon_{\mathbf{k}}\frac{\Delta}{E_{\mathbf{k}}} + \Delta\left(-\frac{\epsilon_{\mathbf{k}}}{E_{\mathbf{k}}}\right) = 0 \end{aligned}$$

Constant term :

$$\begin{aligned} &\sum_{\mathbf{k}}(\epsilon_{\mathbf{k}}v_{\mathbf{k}}^2 + \epsilon_{\mathbf{k}}v_{\mathbf{k}}^2 - \Delta u_{\mathbf{k}}v_{\mathbf{k}} - \Delta u_{\mathbf{k}}v_{\mathbf{k}}) + \frac{N\Delta^2}{V} \\ &= \sum_{\mathbf{k}}\left(\epsilon_{\mathbf{k}} - \frac{\epsilon_{\mathbf{k}}^2}{E_{\mathbf{k}}} - \frac{\Delta^2}{E_{\mathbf{k}}}\right) + \frac{N\Delta^2}{V} \end{aligned}$$

But from the gap equation:

$$\begin{aligned} \sum_{\mathbf{k}} \frac{1}{E_{\mathbf{k}}} &= \frac{2N}{V} \\ \Rightarrow \sum_{\mathbf{k}} \left(\epsilon_{\mathbf{k}} - \frac{\epsilon_{\mathbf{k}}^2}{E_{\mathbf{k}}} \right) &= \frac{N\Delta^2}{V} \end{aligned}$$

Altogether, the Hamiltonian can be written :

$$\begin{aligned} H_{CM} &= \sum_{\mathbf{k}} E_{\mathbf{k}} (\alpha_{\mathbf{k},\uparrow}^\dagger \alpha_{\mathbf{k},\uparrow} + \alpha_{-\mathbf{k},\downarrow}^\dagger \alpha_{-\mathbf{k},\downarrow}) \\ &+ \sum_{\mathbf{k}} \left(\epsilon_{\mathbf{k}} - \frac{\epsilon_{\mathbf{k}}^2}{E_{\mathbf{k}}} \right) - \frac{N\Delta^2}{V} \end{aligned}$$

In the ground state, since $E_{\mathbf{k}} > 0$, there is no particle $\alpha_{\mathbf{k}}^\dagger$.

6.3.2 Energy comparison

Let us compare this energy with that of the non-superconducting case ($\Delta = 0$).

Case $\Delta = 0$

$$\begin{aligned} E_{\mathbf{k}} = \sqrt{\epsilon_{\mathbf{k}}^2} &= \epsilon_{\mathbf{k}} \quad \text{si } \epsilon_{\mathbf{k}} > 0 \quad (\mathbf{k} \text{ empty}) \\ &= -\epsilon_{\mathbf{k}} \quad \text{si } \epsilon_{\mathbf{k}} < 0 \quad (\mathbf{k} \text{ occ.}) \end{aligned}$$

Consequently, the groundstate energy is simply

$$E_{GS} = \sum_{\mathbf{k}_{occ}} 2\epsilon_{\mathbf{k}}$$

Case $\Delta \neq 0$ We find that

$$\Delta E = E_{GS}(\Delta \neq 0) - E_{GS}(\Delta = 0) = -g(\varepsilon_F) \frac{\Delta^2}{2} \quad (6.11)$$

Derivation:

$$\begin{aligned} \Delta E &= \sum_{\mathbf{k}_{occ}} (-\epsilon_{\mathbf{k}} - E_{\mathbf{k}}) + \sum_{\mathbf{k}_{empty}} (\epsilon_{\mathbf{k}} - E_{\mathbf{k}}) + \frac{N\Delta^2}{V} \\ &= 2 \sum_{\mathbf{k}_{empty}} (\epsilon_{\mathbf{k}} - E_{\mathbf{k}}) + \frac{N\Delta^2}{V} \end{aligned}$$

Going through the integral formulation

$$\begin{aligned}\frac{\Delta E}{N} &\approx 2g(\varepsilon_F) \int_0^{\omega_d} (\varepsilon - \sqrt{\varepsilon^2 + \Delta^2}) d\varepsilon + \frac{\Delta^2}{V} \\ &= \frac{\Delta^2}{V} + g(\varepsilon_F) \left(\omega_d^2 - \omega_d \sqrt{\omega_d^2 + \Delta^2} - \Delta^2 \operatorname{atanh} \left(\frac{\omega_d}{\sqrt{\Delta^2 + \omega_d^2}} \right) \right)\end{aligned}$$

Now we can use that $\Delta = 2\omega_D e^{-\frac{1}{Vg(\varepsilon_F)}} \ll \omega_d$

$$\begin{aligned}\frac{\Delta E}{N} &\approx \frac{\Delta^2}{V} - g(\varepsilon_F) \frac{\Delta^2}{2} (1 + \log 2 - \log \frac{\Delta^2}{2\omega_d^2}) \\ &\approx \frac{\Delta^2}{V} - g(\varepsilon_F) \frac{\Delta^2}{2} (1 + \log 2 - \log(2e^{-\frac{2}{Vg(\varepsilon_F)}})) \\ &= -g(\varepsilon_F) \frac{\Delta^2}{2}.\end{aligned}$$

As soon as there is a solution $\Delta \neq 0$, its energy is lower than the ground state energy. On the other hand, as soon as V is positive, i.e. as soon as the interaction is attractive, there is a solution $\Delta \neq 0$. Consequently, superconductivity seem to generically appear for all attractive interactions.

Attraction \Rightarrow superconductivity

6.3.3 Interpretation of the wave function

The mean-field solution is lower in energy than the Fermi sea as soon as $V > 0$. Let us try to understand the physics of this solution.

The ground state wave function reads:

$$|\phi_0\rangle = \prod_{\mathbf{k}} (u_{\mathbf{k}} + v_{\mathbf{k}} c_{\mathbf{k},\uparrow}^\dagger c_{-\mathbf{k},\downarrow}^\dagger) |0_c\rangle$$

The probability for the pair $c_{\mathbf{k},\uparrow}^\dagger c_{-\mathbf{k},\downarrow}^\dagger$ to be occupied is equal to $v_{\mathbf{k}}^2$, the probability for it not to be occupied is equal to $u_{\mathbf{k}}^2$. 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 $\bar{N} = \langle N \rangle$ and $\delta N = (\langle (n - \bar{N})^2 \rangle)^{\frac{1}{2}}$

$$\langle N \rangle = \langle \phi_0 | \sum_{\mathbf{k},\sigma} c_{\mathbf{k},\sigma}^\dagger c_{\mathbf{k},\sigma} | \phi_0 \rangle$$

From previous computations, we see that

$$\langle N \rangle = 2 \sum_{\mathbf{k}} v_{\mathbf{k}}^2$$

Similarly,

$$\begin{aligned} \delta N^2 &= \langle \phi_0 | N^2 | \phi_0 \rangle - \langle N \rangle^2 \\ \langle N^2 \rangle &= \langle \phi_0 | \sum_{\mathbf{k}, \mathbf{k}', \sigma, \sigma'} c_{\mathbf{k}, \sigma}^\dagger c_{\mathbf{k}, \sigma} c_{\mathbf{k}', \sigma'}^\dagger c_{\mathbf{k}', \sigma'} | \phi_0 \rangle \end{aligned}$$

We can use Wick theorem here, as $|\phi_0\rangle$ is the groundstate of a non-interacting Hamiltonian.

$$\begin{aligned} \langle N^2 \rangle &= \sum_{\mathbf{k}, \mathbf{k}', \sigma, \sigma'} \langle c_{\mathbf{k}, \sigma}^\dagger c_{\mathbf{k}, \sigma} \rangle \langle c_{\mathbf{k}', \sigma'}^\dagger c_{\mathbf{k}', \sigma'} \rangle - \langle c_{\mathbf{k}, \sigma}^\dagger c_{\mathbf{k}', \sigma'}^\dagger \rangle \langle c_{\mathbf{k}, \sigma} c_{\mathbf{k}', \sigma'} \rangle + \langle c_{\mathbf{k}, \sigma}^\dagger c_{\mathbf{k}', \sigma'} \rangle \langle c_{\mathbf{k}, \sigma} c_{\mathbf{k}', \sigma'}^\dagger \rangle \\ &= \langle N \rangle^2 - \sum_{\mathbf{k}, \sigma} \langle c_{\mathbf{k}, \sigma}^\dagger c_{-\mathbf{k}, -\sigma}^\dagger \rangle \langle c_{\mathbf{k}, \sigma} c_{-\mathbf{k}, -\sigma} \rangle + \sum_{\mathbf{k}, \sigma} \langle c_{\mathbf{k}, \sigma}^\dagger c_{\mathbf{k}, \sigma} \rangle \langle c_{\mathbf{k}, \sigma} c_{\mathbf{k}, \sigma}^\dagger \rangle \\ &= \langle N \rangle^2 + 4 \sum_{\mathbf{k}} u_{\mathbf{k}}^2 v_{\mathbf{k}}^2 \end{aligned}$$

So that we find

$$\delta N^2 = 4 \sum_{\mathbf{k}} u_{\mathbf{k}}^2 v_{\mathbf{k}}^2 \quad (6.12)$$

Alternative explicit derivation starting from

$$\langle N^2 \rangle = \langle \phi_0 | \sum_{\mathbf{k}, \mathbf{k}', \sigma, \sigma'} c_{\mathbf{k}, \sigma}^\dagger c_{\mathbf{k}, \sigma} c_{\mathbf{k}', \sigma'}^\dagger c_{\mathbf{k}', \sigma'} | \phi_0 \rangle$$

If $(\mathbf{k}', \sigma') \neq (\mathbf{k}, \sigma)$ and $\neq (-\mathbf{k}, -\sigma)$, the terms commute $\rightarrow v_{\mathbf{k}}^2 v_{\mathbf{k}'}^2$

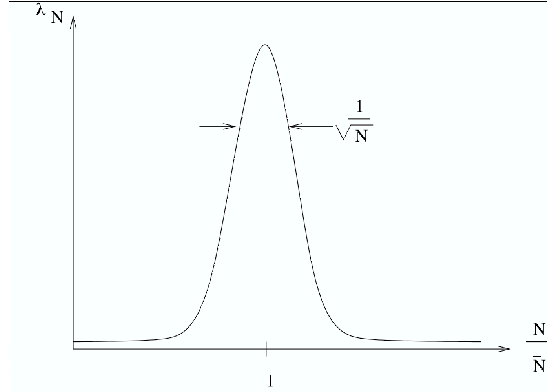
If $(\mathbf{k}', \sigma') = (\mathbf{k}, \sigma) \Rightarrow c_{\mathbf{k}, \sigma}^\dagger \underbrace{c_{\mathbf{k}, \sigma} c_{\mathbf{k}, \sigma}^\dagger}_{1 - c_{\mathbf{k}, \sigma}^\dagger c_{\mathbf{k}, \sigma}} c_{\mathbf{k}, \sigma} \rightarrow v_{\mathbf{k}}^2$

If $(\mathbf{k}', \sigma') = (-\mathbf{k}, -\sigma) \Rightarrow c_{\mathbf{k}, \uparrow}^\dagger c_{\mathbf{k}, \uparrow} c_{-\mathbf{k}, \downarrow}^\dagger c_{-\mathbf{k}, \downarrow} (u_{\mathbf{k}} + v_{\mathbf{k}} c_{\mathbf{k}, \uparrow}^\dagger c_{-\mathbf{k}, \downarrow}^\dagger) \rightarrow v_{\mathbf{k}}^2$

$$\begin{aligned} \langle N^2 \rangle &= \left(\sum_{\mathbf{k}, \sigma} v_{\mathbf{k}}^2 \right)^2 - \underbrace{\sum_{\mathbf{k}, \sigma} (v_{\mathbf{k}}^4 - v_{\mathbf{k}}^2)}_{(\mathbf{k}', \sigma') = (\mathbf{k}, \sigma)} - \underbrace{\sum_{\mathbf{k}, \sigma} (v_{\mathbf{k}}^4 - v_{\mathbf{k}}^2)}_{(\mathbf{k}', \sigma') = (-\mathbf{k}, -\sigma)} \\ &= \left(\sum_{\mathbf{k}} 2v_{\mathbf{k}}^2 \right)^2 - 2 \sum_{\mathbf{k}} 2v_{\mathbf{k}}^4 + 2 \sum_{\mathbf{k}} 2v_{\mathbf{k}}^2 \\ &= \bar{N} + 4 \sum_{\mathbf{k}} v_{\mathbf{k}}^2 (1 - v_{\mathbf{k}}^2) \end{aligned}$$

$$\delta N^2 = 4 \sum_{\mathbf{k}} u_{\mathbf{k}}^2 v_{\mathbf{k}}^2$$

Conclusion \bar{N} grows as N , but δN only grows as \sqrt{N} . So, if one writes $|\phi_0\rangle = \sum_N \lambda_N |\phi_N\rangle$, where $|\phi_N\rangle$ is an eigenstate of \hat{N} with eigenvalue N , the λ_N are distributed according to



So the expectation value of any observable calculated in $|\phi_0\rangle$ is the same as that calculated in $|\phi_{\bar{N}}\rangle$ in the limit $N \rightarrow +\infty$. 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 $\epsilon_{\mathbf{k}}$ lies between $\epsilon_F - \omega_D$ and ϵ_F). There are thus $2N$ values of \mathbf{k} such that $\epsilon_F - \omega_D < \epsilon_{\mathbf{k}} < \epsilon_F + \omega_D$, and the Fermi sea consists in filling up the N values such that $\epsilon_F - \omega_D < \epsilon_{\mathbf{k}} < \epsilon_F$ with an \uparrow spin or a \downarrow spin.

If we construct a wave function only built out of pairs to gain potential energy, with a well-defined number of particles, a natural choice would be

$$|\phi_N\rangle = \frac{1}{N!} \left(\sum_{\mathbf{k}} g_{\mathbf{k}} c_{\mathbf{k},\uparrow}^\dagger c_{-\mathbf{k},\downarrow}^\dagger \right)^N |0_c\rangle$$

If we choose

$$g_{\mathbf{k}} = \begin{cases} 1 & \text{when } \epsilon_{\mathbf{k}} < \epsilon_F \\ 0 & \text{otherwise} \end{cases}$$

we get back the Fermi sea. Starting from this wave function, one could look for the $g_{\mathbf{k}}$ that minimize the energy. But this wave function is *very* complicated to evaluate. It contains of the order of $(2N)^N$ terms!

Now, if we add various $|\phi_N\rangle$ to build

$$\begin{aligned} |\phi\rangle &= \sum_N |\phi_N\rangle \\ &= \sum_N \frac{1}{N!} \left(\sum_{\mathbf{k}} g_{\mathbf{k}} c_{\mathbf{k},\uparrow}^\dagger c_{-\mathbf{k},\downarrow}^\dagger \right)^N |0\rangle \end{aligned}$$

this function can be rewritten as a coherent state

$$|\phi\rangle = e^{\sum_{\mathbf{k}} g_{\mathbf{k}} c_{\mathbf{k},\uparrow}^{\dagger} c_{-\mathbf{k},\downarrow}^{\dagger}} |0\rangle$$

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 because of the Pauli principle

$$\Rightarrow |\phi\rangle = \prod_{\mathbf{k}} (1 + g_{\mathbf{k}} c_{\mathbf{k},\uparrow}^{\dagger} c_{-\mathbf{k},\downarrow}^{\dagger}) |0\rangle$$

If one writes $g_{\mathbf{k}} = \frac{v_{\mathbf{k}}}{u_{\mathbf{k}}}$, one gets back the BCS wave function. This wave function is much easier to manipulate. This is the trick that got Bardeen, Cooper and Schrieffer the Nobel prize!

6.3.4 Elementary excitations

It is clear from the form of H_{MF} that an elementary excitation consists in adding one particle $\alpha_{\mathbf{k},\sigma}^{\dagger}$. This costs an energy

$$E_{\mathbf{k}} = \sqrt{\epsilon_{\mathbf{k}}^2 + \Delta^2}$$

This energy is greater than or equal to Δ . Applying the operator on $|\phi_0\rangle$, we can see that it breaks the pair, replacing the superposition of a pair and two-holes by a superposition of an electron and a hole.

There is thus an *energy gap* Δ 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: a superconductor admits currents while having a gap. One needs to investigate the properties of the electron gas in this state.

6.4 Finite temperatures

At finite temperature, one can still use mean-field theory, but now the gap involves the thermal mean value of $\langle c_{\mathbf{k},\uparrow}^{\dagger} c_{-\mathbf{k},\downarrow}^{\dagger} \rangle$.

$$\begin{aligned} c_{\mathbf{k},\uparrow}^{\dagger} c_{-\mathbf{k},\downarrow}^{\dagger} &= (u_{\mathbf{k}} \alpha_{\mathbf{k},\uparrow}^{\dagger} + v_{\mathbf{k}} \alpha_{-\mathbf{k},\downarrow}) (u_{\mathbf{k}} \alpha_{-\mathbf{k},\downarrow}^{\dagger} - v_{\mathbf{k}} \alpha_{\mathbf{k},\uparrow}) \\ &= -u_{\mathbf{k}} v_{\mathbf{k}} \alpha_{\mathbf{k},\uparrow}^{\dagger} \alpha_{\mathbf{k},\uparrow} + u_{\mathbf{k}} v_{\mathbf{k}} \alpha_{-\mathbf{k},\downarrow} \alpha_{-\mathbf{k},\downarrow}^{\dagger} + u_{\mathbf{k}}^2 \alpha_{\mathbf{k},\uparrow}^{\dagger} \alpha_{-\mathbf{k},\downarrow}^{\dagger} - v_{\mathbf{k}}^2 \alpha_{-\mathbf{k},\downarrow} \alpha_{\mathbf{k},\uparrow} \\ &= u_{\mathbf{k}} v_{\mathbf{k}} (1 - \alpha_{\mathbf{k},\uparrow}^{\dagger} \alpha_{\mathbf{k},\uparrow} - \alpha_{-\mathbf{k},\downarrow}^{\dagger} \alpha_{-\mathbf{k},\downarrow}) + u_{\mathbf{k}}^2 \alpha_{\mathbf{k},\uparrow}^{\dagger} \alpha_{-\mathbf{k},\downarrow}^{\dagger} - v_{\mathbf{k}}^2 \alpha_{-\mathbf{k},\downarrow} \alpha_{\mathbf{k},\uparrow} \\ \langle c_{\mathbf{k},\uparrow}^{\dagger} c_{-\mathbf{k},\downarrow}^{\dagger} \rangle &= u_{\mathbf{k}} v_{\mathbf{k}} (1 - 2f(E_{\mathbf{k}})) \end{aligned}$$

since $\langle \alpha_{\mathbf{k},\uparrow}^\dagger \alpha_{\mathbf{k},\uparrow} \rangle = \langle \alpha_{-\mathbf{k},\downarrow}^\dagger \alpha_{-\mathbf{k},\downarrow} \rangle = \frac{1}{e^{\beta E_{\mathbf{k}}+1}}$ et $\langle \alpha_{\mathbf{k},\uparrow}^\dagger \alpha_{-\mathbf{k},\downarrow}^\dagger \rangle = \langle \alpha_{-\mathbf{k},\downarrow} \alpha_{\mathbf{k},\uparrow} \rangle = 0$

So the gap equation is given by :

$$\Delta = \frac{V}{N} \sum_{\mathbf{k}} \frac{\Delta}{2E_{\mathbf{k}}} (1 - 2f(E_{\mathbf{k}}))$$

$$\Delta = \frac{V}{N} \sum_{\mathbf{k}} \frac{\Delta}{2E_{\mathbf{k}}} \tanh \frac{\beta E_{\mathbf{k}}}{2} \quad (6.13)$$

The non-vanishing solution thus satisfies:

$$1 = \frac{V}{N} \sum_{\mathbf{k}} \frac{1}{2E_{\mathbf{k}}} \tanh \frac{\beta E_{\mathbf{k}}}{2}$$

When T increases, $\beta \rightarrow 0$ and this equation is satisfied for smaller and smaller values of Δ . In particular, for $T = +\infty$, there is no non-vanishing solutions. In fac, there is a temperature T_c where $\Delta(T \geq T_c) = 0$ defined by

$$1 = \frac{V}{N} \sum_{\mathbf{k}} \frac{1}{2|\epsilon_{\mathbf{k}}|} \tanh \frac{\beta_c \epsilon_{\mathbf{k}}}{2} \quad (\beta_c = \frac{1}{T_c})$$

$$1 = Vg(\epsilon_F) \int_0^{\omega_D} \tanh \frac{\beta_c \epsilon}{2} \frac{d\epsilon}{\epsilon}$$

$$\frac{1}{Vg(\epsilon_F)} = \int_0^{\frac{\omega_D}{2T_c}} \frac{\tanh u}{u} du \quad (u = \frac{\beta_c \epsilon}{2})$$

Let us suppose that $T_c \ll \omega_D$. It is a reasonable approximation as we expect that $T_c \simeq \Delta(0) \ll \omega_D$ because thermal fluctuations destroy the gap if $T \gg \Delta(0)$.

$$\begin{aligned} \int_0^{\frac{\omega_D}{2T_c}} \frac{\tanh u}{u} du &= \int_0^\eta \frac{\tanh u}{u} du + \int_\eta^{\frac{\omega_D}{2T_c}} \frac{\tanh u}{u} du \\ &\simeq \int_0^\eta \frac{\tanh u}{u} du + \int_\eta^{\frac{\omega_D}{2T_c}} \frac{du}{u} \quad (\eta \gg 1 \Rightarrow \tanh u \simeq 1) \\ &= \ln \frac{\omega_D}{T_c} + \int_0^\eta \frac{\tanh u}{u} du - \ln \eta - \ln 2 \end{aligned}$$

The second term depends very little on $\eta \gg 1$

$$\lim_{\eta \rightarrow +\infty} \underbrace{\int_0^\eta \frac{\tanh u}{u} du - \ln \eta - \ln 2}_{\gamma} = \ln \frac{2\gamma}{\pi} \approx \ln 1.13 \quad (6.14)$$

where γ is the Euler constant.

$$T_c \simeq 1.13 \omega_D e^{\frac{-1}{Vg(\epsilon_F)}} \quad (6.15)$$

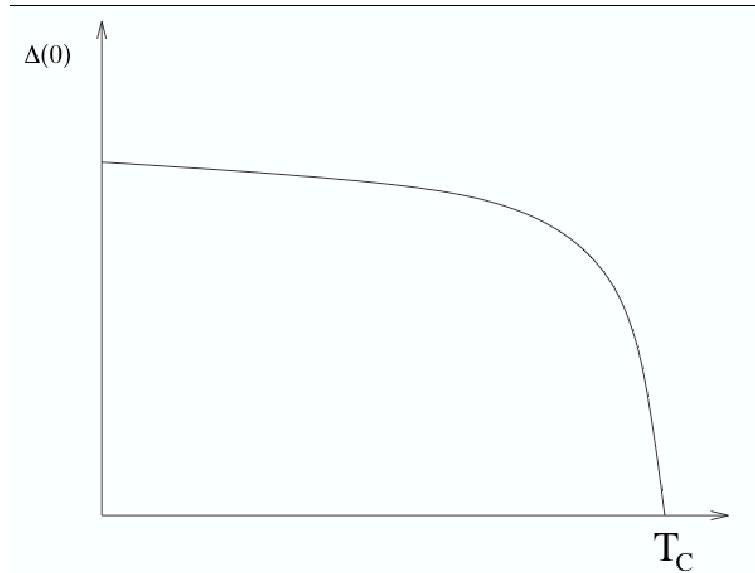
This formula cannot be used directly to predict T_c since ω_D , V and $g(\epsilon_F)$ are adjustable parameters. However, if we compare it to the formula for Δ , we get

$$\frac{\Delta}{k_B T_c} = \frac{2}{1.13} = 1.76$$

Δ 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:

$$\Delta(T) \simeq 3.2 T_c \sqrt{1 - \frac{T}{T_c}} \quad (6.16)$$

The exponent $\frac{1}{2}$ is typical of mean field theory.



6.5 Electrodynamics of superconductors

6.5.1 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 λ called the "penetration length".

Let us look in more detail at the consequences of these properties on the constitutive equations:

Vanishing resistivity: In a metal, one generally assumes that the motion of an electron is governed by the equation:

$$m \frac{d\mathbf{v}}{dt} + m \frac{\mathbf{v}}{\tau} = -e\mathbf{E} \quad (6.17)$$

where τ 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 $\mathbf{v} = -\frac{e\tau}{m}\mathbf{E}$, and of the general solution of the equation without the right hand side:

$$\begin{aligned} \mathbf{v} &= \mathbf{v}_0 e^{-\frac{t}{\tau}} \\ \Rightarrow \mathbf{v} &= \mathbf{v}_0 e^{-\frac{t}{\tau}} - \frac{e\tau}{m}\mathbf{E} \end{aligned}$$

After a time $t \gg \tau$, one thus gets:

$$\mathbf{v} = \frac{e\tau}{m}\mathbf{E}.$$

Finally, using $\mathbf{j} = -nev$

$$\mathbf{j} = \frac{ne^2\tau}{m}\mathbf{E} \quad (6.18)$$

or $\mathbf{j} = \sigma\mathbf{E}$, with $\sigma = \frac{ne^2\tau}{m}$. This is the constitutive relation of *metals*.

In superconductors, $\sigma = \infty$ ($R = 0$) $\Rightarrow \tau = +\infty$

$$m \frac{d\mathbf{v}}{dt} = -e\mathbf{E}$$

But $\mathbf{j} = -n_s e\mathbf{v}$, where n_s is the superfluid density, i.e. the density of electrons that take part into superconductivity, so that

$$\begin{aligned} m \left(\frac{1}{n_s e} \right) \frac{d\mathbf{j}}{dt} &= e\mathbf{E} \\ \mathbf{E} &= \frac{m}{n_s e^2} \frac{d\mathbf{j}}{dt} \end{aligned}$$

This relation only fixes the current up to a constant. It is thus not sufficient to describe the response of superconductors.

Meissner effect: We can combine this equation with the following Maxwell equation

$$\mathbf{E} = -\frac{\partial \mathbf{A}}{\partial t}. \quad (6.19)$$

Note that this equation relates the electric field (a physical quantity) to the vector potential, in a fashion which is clearly not gauge invariant. It is therefore only valid in a specific gauge that we will discuss later. We find

$$\begin{aligned} \frac{m}{n_s e^2} \frac{d}{dt} \mathbf{j} &= -\frac{\partial \mathbf{A}}{\partial t} \\ \frac{m}{n_s e^2} \mathbf{j} + \mathbf{A} &= \mathbf{c} \end{aligned}$$

where \mathbf{c} is a constant of integration independent of the time. The Meissner effect is a consequence of this equation if $\nabla \times \mathbf{c} = \mathbf{0}$. Indeed, Maxwell's equations imply

$$\begin{aligned} \nabla \times \mathbf{B} &= \mu_0 \mathbf{j} \\ \nabla \times \nabla \times \mathbf{B} &= \mu_0 \nabla \times \mathbf{j} \end{aligned}$$

But $\nabla \times \nabla \times \mathbf{B} = -\Delta \mathbf{B} + \underbrace{\nabla(\nabla \cdot \mathbf{B})}_{=0}$ and $\nabla \times \mathbf{j} = -\frac{n_s e^2}{m} \underbrace{\nabla \times \mathbf{A}}_{\mathbf{B}}$. So the equation relating \mathbf{j} to \mathbf{A} implies

$$\Delta \mathbf{B} = \frac{\mu_0 n_s e^2}{m} \mathbf{B} + \nabla \times \mathbf{c}$$

If $\nabla \times \mathbf{c} = \mathbf{0}$, we get

$$\Rightarrow \Delta \mathbf{B} = \frac{n_s e^2 \mu_0}{m} \mathbf{B}$$

Let us define

$$\frac{1}{\lambda^2} = \frac{\mu_0 n_s e^2}{m}$$

We thus have:

$$\Delta \mathbf{B} = \frac{1}{\lambda^2} \mathbf{B} \quad (6.20)$$

Let us adopt a geometry where $\mathbf{B} = \mathbf{B}(z)$. Integrating this equation leads to:

$$\mathbf{B} = \mathbf{B}(0) e^{-\frac{z}{\lambda}}$$

The magnetic field \mathbf{B} decreases exponentially over a characteristic length given by λ . Since $\rho = 0$ in a metal, $\nabla \cdot \mathbf{j} = 0$. Thus, $\frac{1}{c} \nabla \cdot \mathbf{A} = \nabla \cdot \mathbf{c}$. If one chooses a gauge in which $\nabla \cdot \mathbf{A} = 0$ and $\mathbf{A} \rightarrow \mathbf{0}$ in the bulk, one can choose $\mathbf{c} = \mathbf{0}$. Under these conditions, \mathbf{j} and \mathbf{A} are related by:

$$\mathbf{j} = -\frac{n_s e^2}{m} \mathbf{A} = -\frac{1}{\lambda^2 \mu_0} \mathbf{A} \quad (6.21)$$

This equation is known as London equation (F. et H. London, 1935). This equation is valid in the London gauge (in the sense of choosing correctly our gauge for \mathbf{A}):

- $\nabla \cdot \mathbf{A} = 0$.
- $\mathbf{A} \rightarrow \mathbf{0}$ 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 (as $\mathbf{F} = e\mathbf{v} \times \mathbf{B}$ is orthogonal to \mathbf{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.

6.5.2 Response to a static magnetic field:

Let us consider a static field in the gauge $\nabla \cdot \mathbf{A} = 0$.

$$\mathbf{p} = m\mathbf{v} + e\mathbf{A}$$

$$\begin{aligned} \Rightarrow H_{cin} &= \frac{1}{2}m \sum_i v_i^2 = \frac{1}{2m} \sum_i (\mathbf{p}_i - e\mathbf{A}_i)^2 \\ &= \frac{1}{2m} \sum_i (-i\hbar\nabla_i - e\mathbf{A}(\mathbf{r}_i))^2 \\ &= \underbrace{\frac{H_{cin}^0}{\frac{1}{2m} \sum_i \mathbf{p}_i^2}} + \underbrace{\frac{ie\hbar}{2m} \sum_i (\nabla_i \cdot \mathbf{A}(\mathbf{r}_i) + \mathbf{A}(\mathbf{r}_i) \cdot \nabla_i)}_{H_1} + O(A^2) \end{aligned}$$

The term $O(A^2)$ can be ignored for the calculation of \mathbf{j} to first order.

Let us introduce the Fourier transform of the vector potential $\mathbf{A}(\mathbf{r}) = \sum_{\mathbf{q}} \mathbf{a}(\mathbf{q}) e^{i\mathbf{q}\cdot\mathbf{r}}$, and let us determine the form of H_1 in second quantization. For the term $\nabla_i \cdot \mathbf{A}(\mathbf{r}_i)$, we compute:

$$\begin{aligned} \langle k' \nabla \cdot \mathbf{A}(\mathbf{r}) | k \rangle &= \sum_{\mathbf{q}} \int \frac{e^{-i\mathbf{k}'\cdot\mathbf{r}}}{\sqrt{V}} \nabla \cdot (\mathbf{a}(\mathbf{q}) e^{i\mathbf{q}\cdot\mathbf{r}} \frac{e^{i\mathbf{k}\cdot\mathbf{r}}}{\sqrt{V}}) d\mathbf{r} \\ &= \sum_{\mathbf{q}} \frac{1}{V} \int e^{-i\mathbf{k}'\cdot\mathbf{r}} i(\mathbf{k} + \mathbf{q}) \cdot \mathbf{a}(\mathbf{q}) e^{i(\mathbf{k}+\mathbf{q})\cdot\mathbf{r}} d\mathbf{r} \end{aligned}$$

Since the field is assumed to be transverse ($\nabla \cdot \mathbf{A} = 0$), we have $\mathbf{q} \cdot \mathbf{a}(\mathbf{q}) = 0$, and we get :

$$\sum_{\mathbf{q}} i\mathbf{k} \cdot \mathbf{a}(\mathbf{q}) \delta_{\mathbf{k}', \mathbf{k}+\mathbf{q}}$$

The contribution to H_1 is thus given by :

$$-\frac{e\hbar}{2m} \sum_{\mathbf{k}, \mathbf{q}, \sigma} \mathbf{k} \cdot \mathbf{a}(\mathbf{q}) c_{\mathbf{k}+\mathbf{q}, \sigma}^\dagger c_{\mathbf{k}, \sigma}$$

For the term $\mathbf{A}(\mathbf{r}_i) \cdot \nabla_i$, we obtain the same result such that

$$H_1 = \frac{-e\hbar}{m} \sum_{\mathbf{k}, \mathbf{q}, \sigma} \mathbf{k} \cdot \mathbf{a}(\mathbf{q}) c_{\mathbf{k}+\mathbf{q}, \sigma}^\dagger c_{\mathbf{k}, \sigma}$$

The current operator can be written :

$$\mathbf{j}(\mathbf{r}) = \frac{e}{2m} \sum_{\sigma} \left\{ \Psi^\dagger(\mathbf{r}, \sigma) (-i\hbar\nabla - e\mathbf{A}) \Psi(\mathbf{r}, \sigma) + [(i\hbar\nabla - e\mathbf{A}) \Psi^\dagger(\mathbf{r}, \sigma)] \Psi(\mathbf{r}, \sigma) \right\}$$

where $\Psi(\mathbf{r}, \sigma)$, $\Psi^\dagger(\mathbf{r}, \sigma)$ are field operators (see Appendix on second quantization and exercises). They can be expanded in Fourier series:

$$\Psi^\dagger(\mathbf{r}, \sigma) = \sum_{\mathbf{k}_1} e^{-i\mathbf{k}_1 \cdot \mathbf{r}} c_{\mathbf{k}_1, \sigma}^\dagger; \quad \Psi(\mathbf{r}, \sigma) = \sum_{\mathbf{k}_2} e^{i\mathbf{k}_2 \cdot \mathbf{r}} c_{\mathbf{k}_2, \sigma}$$

as well as their gradient:

$$\nabla \Psi^\dagger(\mathbf{r}, \sigma) = \sum_{\mathbf{k}_1} (-i\mathbf{k}_1) e^{-i\mathbf{k}_1 \cdot \mathbf{r}} c_{\mathbf{k}_1, \sigma}^\dagger; \quad \nabla \Psi(\mathbf{r}, \sigma) = \sum_{\mathbf{k}_2} i\mathbf{k}_2 e^{i\mathbf{k}_2 \cdot \mathbf{r}} c_{\mathbf{k}_2, \sigma}$$

It is convenient to decompose the current into two contributions:

$$\mathbf{j}(\mathbf{r}) = \mathbf{J}_1(\mathbf{r}) + \mathbf{J}_2(\mathbf{r})$$

with

$$\mathbf{J}_1(\mathbf{r}) = \frac{-ie\hbar}{2m} \sum_{\sigma} \left\{ \Psi^\dagger(\mathbf{r}, \sigma) \nabla \Psi(\mathbf{r}, \sigma) - (\nabla \Psi^\dagger(\mathbf{r}, \sigma)) \Psi(\mathbf{r}, \sigma) \right\}$$

and

$$\mathbf{J}_2(\mathbf{r}) = \frac{-e^2}{m} \mathbf{A} \sum_{\sigma} \Psi^\dagger(\mathbf{r}, \sigma) \Psi(\mathbf{r}, \sigma)$$

The Fourier transform of $\mathbf{J}_1(\mathbf{r})$ is given by:

$$\begin{aligned} \mathbf{J}_1(\mathbf{q}) &= \int d\mathbf{r} e^{-i\mathbf{q} \cdot \mathbf{r}} \mathbf{J}_1(\mathbf{r}) \\ &= \frac{-ie\hbar}{2m} \sum_{\sigma} \left(\int d\mathbf{r} e^{-i\mathbf{q} \cdot \mathbf{r}} \sum_{\mathbf{k}_1, \mathbf{k}_2} \left[e^{-i\mathbf{k}_1 \cdot \mathbf{r} + i\mathbf{k}_2 \cdot \mathbf{r}} (i\mathbf{k}_2) c_{\mathbf{k}_1, \sigma}^\dagger c_{\mathbf{k}_2, \sigma} - e^{-i\mathbf{k}_1 \cdot \mathbf{r} + i\mathbf{k}_2 \cdot \mathbf{r}} (-i\mathbf{k}_1) c_{\mathbf{k}_1, \sigma}^\dagger c_{\mathbf{k}_2, \sigma} \right] \right) \\ &= \frac{-ie\hbar}{2m} \sum_{\mathbf{k}, \sigma} \left[i(\mathbf{k} + \mathbf{q}) c_{\mathbf{k}, \sigma}^\dagger c_{\mathbf{k}+\mathbf{q}, \sigma} + i\mathbf{k} c_{\mathbf{k}, \sigma}^\dagger c_{\mathbf{k}+\mathbf{q}, \sigma} \right] \\ &= \frac{e\hbar}{m} \sum_{\mathbf{k}, \sigma} \left(\mathbf{k} + \frac{1}{2} \mathbf{q} \right) c_{\mathbf{k}, \sigma}^\dagger c_{\mathbf{k}+\mathbf{q}, \sigma} \end{aligned}$$

Besides,

$$\begin{aligned}\mathbf{J}_2(\mathbf{r}) &= \frac{-e^2}{m} \mathbf{A} \underbrace{\sum_{\sigma} \Psi^{\dagger}(\mathbf{r}, \sigma) \Psi(\mathbf{r}, \sigma)}_{\text{density}} \\ &= -\frac{e^2 n(\mathbf{r})}{m} \mathbf{A}\end{aligned}$$

\mathbf{J}_2 is proportional to \mathbf{A} , and this relation is similar to the London equation, but it involves the *total* density and not the superconducting density. Therefore, the term \mathbf{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 \mathbf{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 $\mathbf{J}_1(\mathbf{q})$ in the limit $\mathbf{q} \rightarrow 0$ (large wavelength). In this case, a simpler calculation is possible.

Let us start by writing the operators H_1 and $\mathbf{J}_1(\mathbf{q})$ in terms of the Bogoliubov operators:

$$\begin{aligned}H_1 &= \frac{-e\hbar}{m} \sum_{\mathbf{k}, \mathbf{q}} \mathbf{k} \cdot \mathbf{a}(\mathbf{q}) \left[(u_{\mathbf{k}+\mathbf{q}} \alpha_{\mathbf{k}+\mathbf{q}, \uparrow}^{\dagger} + v_{\mathbf{k}+\mathbf{q}} \alpha_{-\mathbf{k}-\mathbf{q}, \downarrow}) \right. \\ &\quad \times (u_{\mathbf{k}} \alpha_{\mathbf{k}, \uparrow} + v_{\mathbf{k}} \alpha_{-\mathbf{k}, \downarrow}^{\dagger}) \underbrace{-}_{\mathbf{k} \rightarrow -\mathbf{k}-\mathbf{q}} (u_{\mathbf{k}} \alpha_{-\mathbf{k}, \downarrow}^{\dagger} - v_{\mathbf{k}} \alpha_{\mathbf{k}, \uparrow}) \\ &\quad \left. \times (u_{\mathbf{k}+\mathbf{q}} \alpha_{-\mathbf{k}-\mathbf{q}, \downarrow} + v_{\mathbf{k}+\mathbf{q}} \alpha_{\mathbf{k}+\mathbf{q}, \uparrow}^{\dagger}) \right] \\ &= \frac{-e\hbar}{m} \sum_{\mathbf{k}, \mathbf{q}} \mathbf{k} \cdot \mathbf{a}(\mathbf{q}) \left[(u_{\mathbf{k}} u_{\mathbf{k}+\mathbf{q}} + v_{\mathbf{k}} v_{\mathbf{k}+\mathbf{q}}) \right. \\ &\quad \times (\alpha_{\mathbf{k}+\mathbf{q}, \uparrow}^{\dagger} \alpha_{\mathbf{k}, \uparrow} - \alpha_{-\mathbf{k}, \downarrow}^{\dagger} \alpha_{-\mathbf{k}-\mathbf{q}, \downarrow}) \\ &\quad + (v_{\mathbf{k}} u_{\mathbf{k}+\mathbf{q}} - u_{\mathbf{k}} v_{\mathbf{k}+\mathbf{q}}) \\ &\quad \left. \times (\alpha_{\mathbf{k}+\mathbf{q}, \uparrow}^{\dagger} \alpha_{-\mathbf{k}, \downarrow}^{\dagger} - \alpha_{-\mathbf{k}-\mathbf{q}, \downarrow} \alpha_{\mathbf{k}, \uparrow}) \right]\end{aligned}$$

In the limit $\mathbf{q} \rightarrow \mathbf{0}$, the second term disappears, and the first term gives:

$$\begin{aligned}H_1 &= \frac{-e\hbar}{m} \sum_{\mathbf{k}} \mathbf{k} \cdot \mathbf{a}(\mathbf{0}) (\alpha_{\mathbf{k}, \uparrow}^{\dagger} \alpha_{\mathbf{k}, \uparrow} - \alpha_{-\mathbf{k}, \downarrow}^{\dagger} \alpha_{-\mathbf{k}, \downarrow}) \\ &\Rightarrow \begin{cases} E_{\mathbf{k}, \uparrow} = E_{\mathbf{k}} - \frac{e\hbar}{m} \mathbf{k} \cdot \mathbf{a}(\mathbf{0}) \\ E_{\mathbf{k}, \downarrow} = E_{\mathbf{k}} + \frac{e\hbar}{m} \mathbf{k} \cdot \mathbf{a}(\mathbf{0}) \end{cases}\end{aligned}$$

Similarly, in the limit $\mathbf{q} \rightarrow \mathbf{0}$, we have:

$$\mathbf{J}_1(\mathbf{q} \rightarrow \mathbf{0}) = \frac{e\hbar}{m} \sum_{\mathbf{k}} \mathbf{k} (\alpha_{\mathbf{k}, \uparrow}^{\dagger} \alpha_{\mathbf{k}, \uparrow} - \alpha_{-\mathbf{k}, \downarrow}^{\dagger} \alpha_{-\mathbf{k}, \downarrow})$$

But $\langle \alpha_{\mathbf{k},\sigma}^\dagger \alpha_{\mathbf{k},\sigma} \rangle = f(E_{\mathbf{k},\sigma})$

$$\Rightarrow \langle \mathbf{J}_1 \rangle = \frac{e\hbar}{m} \sum_{\mathbf{k}} \mathbf{k} [f(E_{\mathbf{k},\uparrow}) - f(E_{\mathbf{k},\downarrow})]$$

Besides,

$$\begin{aligned} f(E_{\mathbf{k},\uparrow}) - f(E_{\mathbf{k},\downarrow}) &\simeq \frac{2e\hbar}{m} \mathbf{k} \cdot \mathbf{a}(\mathbf{0}) \left(-\frac{\partial f}{\partial E_{\mathbf{k}}} \right) \\ \langle \mathbf{J}_1(\mathbf{q} \rightarrow \mathbf{0}) \rangle &= \frac{2e^2\hbar^2}{m^2} \sum_{\mathbf{k}} (\mathbf{k} \cdot \mathbf{a}(\mathbf{0})) \mathbf{k} \left(-\frac{\partial f}{\partial E_{\mathbf{k}}} \right) \end{aligned}$$

This means that $\langle \mathbf{J}_1(\mathbf{q} \rightarrow \mathbf{0}) \rangle$ is parallel to $\mathbf{a}(\mathbf{0})$. Indeed, for a decomposition $\mathbf{k} = \alpha \mathbf{a}(\mathbf{0}) + \beta \mathbf{q}$, with $\mathbf{q} \cdot \mathbf{a}(\mathbf{0}) = 0$, the terms in the sum is odd in β . Focusing on $k \approx k_F$, we finally obtain

$$\langle \mathbf{J}_1(\mathbf{q} \rightarrow \mathbf{0}) \rangle \simeq \frac{2e^2\hbar^2 k_F^2}{m^2} \frac{1}{3} \sum_{\mathbf{k}} \left(-\frac{\partial f}{\partial E_{\mathbf{k}}} \right) \mathbf{a}(\mathbf{0})$$

Proof: let us assume that $\mathbf{N} = \sum_{\mathbf{k}} (\mathbf{k} \cdot \mathbf{a}(\mathbf{0})) \mathbf{k} h(k)$, where h is an arbitrary function:

$$\begin{aligned} \mathbf{N} \cdot \mathbf{a}(\mathbf{0}) &= \sum_{\mathbf{k}} (\mathbf{a}(\mathbf{0}) \cdot \mathbf{k})^2 h(k) \\ &= \int dk \int 2\pi \sin \theta (\mathbf{a}(\mathbf{0}) k \cos \theta)^2 d\theta k^2 h(k) \\ &= \int 2\pi \sin \theta \cos^2 \theta d\theta \int dk a^2(0) k^2 k^2 h(k) \\ &= 2\pi \frac{2}{3} \int dk a^2(0) k^4 h(k) \\ &= \frac{1}{3} 2\pi \int_0^\pi \sin \theta d\theta \int dk a^2(0) k^4 h(k) \\ &= \frac{1}{3} \int d\mathbf{k} a^2(0) k^2 h(k) \end{aligned}$$

We want to make the electronic density appear. We now that $g(\varepsilon) = A\sqrt{\varepsilon}$ in three dimensions. Therefore, taking into account the spin degree of freedom,

$$n = 2 \times \int_0^{\varepsilon_F} g(\varepsilon) d\varepsilon = 2A \int_0^{\varepsilon_F} \sqrt{\varepsilon} d\varepsilon = 3A\varepsilon_F^{3/2} = 3N\varepsilon_F g(\varepsilon_F)$$

. Inserting $\frac{\hbar^2 k_F^2}{2m} = \varepsilon_F$ and $\sum_{\mathbf{k}} = Ng(\varepsilon_F) \int d\varepsilon$, we get:

$$\begin{aligned} \langle \mathbf{J}_1(\mathbf{q} \rightarrow \mathbf{0}) \rangle &\simeq \frac{2e^2}{3m} \varepsilon_F g(\varepsilon_F) N \int_{-\infty}^{+\infty} d\varepsilon \left(-\frac{\partial f}{\partial E_{\mathbf{k}}} \right) \mathbf{a}(\mathbf{0}) \\ &\simeq \frac{e^2}{m} n N \int_{-\infty}^{+\infty} d\varepsilon \left(-\frac{\partial f}{\partial E_{\mathbf{k}}} \right) \mathbf{a}(\mathbf{0}) \end{aligned}$$

The total current can thus be written:

$$\begin{aligned} \langle \mathbf{j} \rangle &= \langle \mathbf{J}_1 + \mathbf{J}_2 \rangle \\ &= -\frac{e^2 n}{m} N \mathbf{a}(\mathbf{0}) \left[1 - \int_{-\infty}^{+\infty} d\epsilon \left(-\frac{\partial f}{\partial E_{\mathbf{k}}} \right) \right] \end{aligned}$$

with $E = \sqrt{\epsilon^2 + \Delta^2}$.

If $\Delta = 0$

$$\int_{-\infty}^{+\infty} d\epsilon \left(\frac{\partial f}{\partial E_{\mathbf{k}}} \right) = \int_{-\infty}^{+\infty} dE \left(\frac{\partial f}{\partial E} \right) = f(+\infty) - f(-\infty) = -1$$

We recover that there is no macroscopic current in a normal metal: $\langle J \rangle = \mathbf{0}$

If $\Delta \neq 0$

$$-\int_{-\infty}^{+\infty} d\epsilon \left(\frac{\partial f}{\partial E_{\mathbf{k}}} \right) = 2 \int_{\Delta}^{+\infty} dE \left(-\frac{\partial f}{\partial E} \frac{E}{\sqrt{E^2 - \Delta^2}} \right) < 1$$

If one defines the density of normal fluid by

$$n_n = n \int_{-\infty}^{+\infty} d\epsilon \left(\frac{\partial f}{\partial E_{\mathbf{k}}} \right)$$

and the superfluid density by

$$n_s = n - n_n$$

one recovers London equation

$$\mathbf{j} = -\frac{e^2 n_s}{m} \mathbf{a}(\mathbf{0}) \quad (6.22)$$

As $f(E) = \frac{1}{e^{\beta E} + 1}$, in the limit $\beta \rightarrow +\infty$, $\partial_E f = \delta(E)$. Therefore, at $T = 0$, with $\Delta \neq 0$, 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 \mathbf{J}_1 cancels anyway the contribution of \mathbf{J}_2 coming from the electrons that do not participate to superconductivity. n is thus the density of electrons which are at ω_D of ϵ_F .

6.5.3 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 ξ_0 that is *not* related to the penetration depth λ . 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

$$\epsilon_F - \Delta_{\text{SC}} < \frac{p^2}{2m} < \epsilon_F + \Delta_{\text{SC}} \quad \text{with} \quad \epsilon_F = \frac{p_F^2}{2m}$$

Let us define

$$\frac{p_2^2}{2m} = \epsilon_F + \Delta_{\text{SC}}$$

and

$$\frac{p_1^2}{2m} = \epsilon_F - \Delta_{\text{SC}}$$

$$\frac{1}{2m}(p_2^2 - p_1^2) = \frac{1}{2m} \underbrace{(p_2 - p_1)}_{\Delta p} \underbrace{(p_2 + p_1)}_{2p_F} = 2\Delta_{\text{SC}}$$

$$\Delta p = 2 \frac{\Delta_{\text{SC}} m}{p_F} = \frac{2\Delta_{\text{SC}}}{v_F}$$

$$\boxed{\Delta x \geq \frac{\hbar v_F}{2\Delta_{\text{SC}}} \simeq \xi_0} \quad (6.23)$$

It is thus impossible to localize a Cooper pair over a distance smaller than ξ_0 . Perturbations of the superconducting state must therefore act over lengths of order ξ_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, as shown in 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 \mathbf{H} .

If we cool 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_{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 $\int \mathbf{j} \cdot \mathbf{A}_{\text{ext}} d\tau$ whose density is $\frac{H^2}{8\pi}$. So the energy per volume unit in the superconducting state is

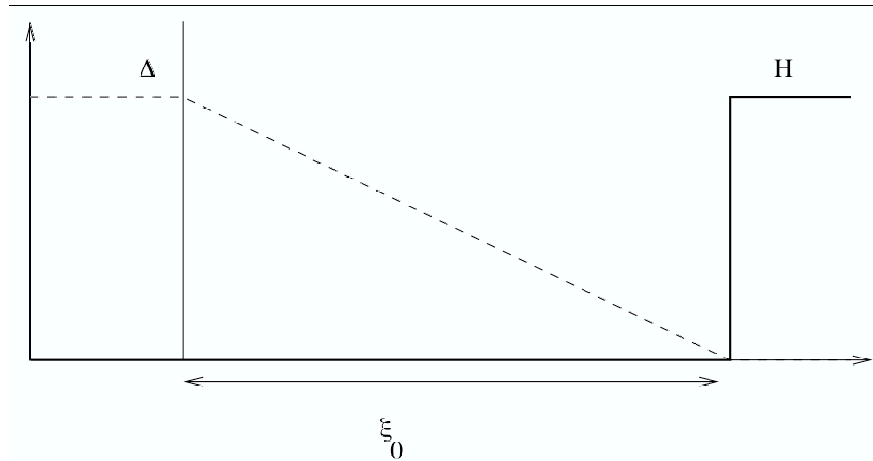
$$\frac{H^2}{8\pi} - |E_{\text{cond}}|$$

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 ξ_0 and λ_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-Ginzburg theory, which we introduce in the next section. Nonetheless, the basic idea can be explained in simple terms. Creating an interface between a normal region and a superconducting reduces the magnetic energy at a cost in superconducting energy. Let us consider the two limiting scenarii.

$\xi_0 \gg \lambda$ In this regime, we see a progressive disappearance of the superconducting density. On the other hand, we expect the magnetic field to be directly immediately suppressed as soon as the superconducting density is non-zero. You can find a schematic below.

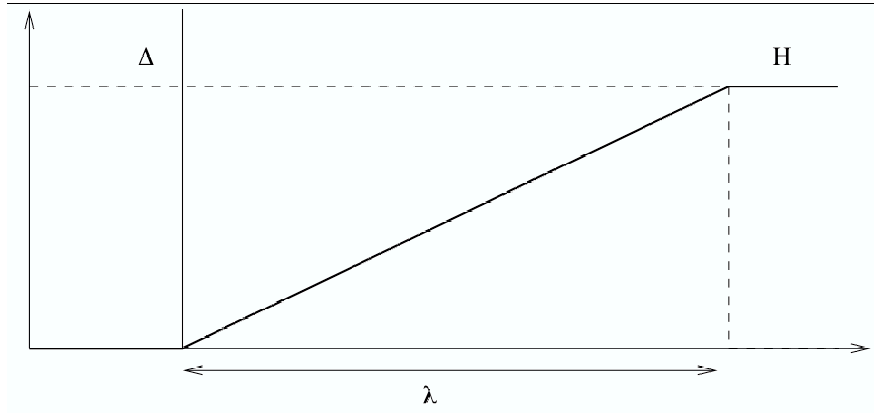


Over the transition region, the energy cost per surface area is

$$\frac{E}{S} = \frac{H^2}{8\pi}\xi_0 - |E_{\text{cond}}|\frac{\xi_0}{2}$$

Close to $H = H_c$, it is positive, and therefore unfavorable, compared to having directly a transition at the boundary of the material. The system will therefore avoid creating such normal to superconducting interfaces.

$\xi_0 \ll \lambda$ In that situation, the scenario is opposite: the transition between superconducting region and metallic region is abrupt, with a large penetration of the external magnetic field.



In that case, the energy cost of the interface is

$$\frac{E}{S} = \frac{H^2 \lambda}{8\pi} - |E_{\text{cond}}| \lambda$$

This is negative at $H = H_c$. The superconductor will therefore prefer to have a large number of these interfaces, and the magnetic field will penetrate the superconductor for values of the field much smaller than H_c , through the presence of normal regions. To increase the surface of interfaces while keeping their volume as small as possible, the magnetic field penetrates as small cylinders known as *vortices*.

6.6 The Landau-Ginzburg theory

The most efficient description of the macroscopic properties of superconductors has been developed before the BCS theory. It is due to Landau and Ginzburg who postulated in 1951 the form of the free energy as a function of a *complex* order parameter that describes:

1. The superconducting density $|\Psi|^2$;
2. The coupling to the electromagnetic field through the change of its phase under a gauge transformation.

$$\mathbf{A}' = \mathbf{A} + \nabla\chi \Rightarrow \Psi' = \Psi e^{ie^* \frac{\chi}{\hbar}}$$

The free-energy density can thus be written :

$$f = f_n + \alpha |\Psi|^2 + \frac{\beta}{2} |\Psi|^4 + \frac{1}{2m} |(-i\hbar\nabla - e^* \mathbf{A}) \Psi|^2 + \frac{B^2}{8\pi}$$

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. e^* is the effective charge of the quasiparticle.

While this functional was introduced by Landau and Ginzburg, Gorkov showed in 1959, using Green functions, 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 Ψ with our superconducting parameter Δ naturally leads to this form. Indeed, $\Delta = \frac{V}{N} \sum_{\mathbf{k}} \langle c_{\mathbf{k},\uparrow}^\dagger c_{-\mathbf{k},\downarrow}^\dagger \rangle$ 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 chose a real solution. However, when there is a magnetic field, the phase of the wave function is related to the choice of gauge. Under a gauge transformation $\mathbf{A}' = \mathbf{A} + \nabla\chi(\mathbf{r})$, the wavefunction of a single-particle of charge e transforms as $\Phi(\mathbf{r}) = e^{\frac{ie}{\hbar}\chi(\mathbf{r})}\Psi(\mathbf{r})$.

Proof: the Schrödinger equation of a free particle in a field is given by:

$$H\Psi = \frac{1}{2m}(\mathbf{p} - e^*\mathbf{A})^2\Psi = E\Psi$$

Let us perform the gauge transformation $\mathbf{A}' = \mathbf{A} + \nabla\chi(\mathbf{r})$. The Schrödinger equation

$$\frac{1}{2m}(\mathbf{p} - e^*\mathbf{A}')^2\Phi = E\Phi$$

is satisfied by $\Phi(\mathbf{r}) = e^{\frac{ie^*}{\hbar}\chi(\mathbf{r})}\Psi(\mathbf{r})$, as

$$\begin{aligned} (-i\hbar\nabla - e^*\mathbf{A}')\Phi(\mathbf{r}) &= (-i\hbar\nabla - e^*\mathbf{A}')e^{\frac{ie^*}{\hbar}\chi(\mathbf{r})}\Psi(\mathbf{r}) \\ &= -e^*\mathbf{A}'\Phi(\mathbf{r}) - i\hbar\nabla(e^{\frac{ie^*}{\hbar}\chi(\mathbf{r})}\Psi(\mathbf{r})) \\ &= -e^*\mathbf{A}'\Phi(\mathbf{r}) - (i\hbar\frac{ie^*}{\hbar}\nabla\chi(\mathbf{r}))\Phi(\mathbf{r}) + e^{\frac{ie^*}{\hbar}\chi(\mathbf{r})}(-i\hbar\nabla)\Psi(\mathbf{r}) \\ &= -e^*\mathbf{A}\Phi(\mathbf{r}) + e^{\frac{ie^*}{\hbar}\chi(\mathbf{r})}(-i\hbar\nabla)\Psi(\mathbf{r}) \\ &= e^{\frac{ie^*}{\hbar}\chi(\mathbf{r})}(\mathbf{p} - e^*\mathbf{A})\Psi(\mathbf{r}) \end{aligned}$$

To get the Hamiltonian, one must again apply $(-i\hbar\nabla - e^*\mathbf{A}')$, performing the scalar product. Let us suppose that we have a vector, and not a scalar:

$$(-i\hbar\nabla - e^*\mathbf{A}')e^{\frac{ie^*}{\hbar}\chi(\mathbf{r})}\mathbf{F}(\mathbf{r}) = -e^*\mathbf{A}'e^{\frac{ie^*}{\hbar}\chi(\mathbf{r})}\mathbf{F}(\mathbf{r}) - i\hbar\nabla(e^{\frac{ie^*}{\hbar}\chi(\mathbf{r})}\mathbf{F}(\mathbf{r}))$$

But $\nabla \cdot (p\mathbf{A}) = p\nabla \cdot \mathbf{A} + \mathbf{A} \cdot \nabla p$

$$\begin{aligned} \nabla \left(e^{\frac{ie^*}{\hbar}\chi(\mathbf{r})} \mathbf{F}(\mathbf{r}) \right) &= e^{\frac{ie^*}{\hbar}\chi(\mathbf{r})} \nabla \cdot \mathbf{F}(\mathbf{r}) + \mathbf{F}(\mathbf{r}) \cdot \nabla e^{\frac{ie^*}{\hbar}\chi(\mathbf{r})} \\ &= e^{\frac{ie^*}{\hbar}\chi(\mathbf{r})} \nabla \cdot \mathbf{F}(\mathbf{r}) + e^{\frac{ie^*}{\hbar}\chi(\mathbf{r})} \frac{ie^*}{\hbar} \nabla \chi(\mathbf{r}) \cdot \mathbf{F}(\mathbf{r}) \end{aligned}$$

$$(-i\hbar\nabla - e^*\mathbf{A}') e^{\frac{ie^*}{\hbar}\chi(\mathbf{r})} \mathbf{F}(\mathbf{r}) = e^{\frac{ie^*}{\hbar}\chi(\mathbf{r})} (\mathbf{p} - e^*\mathbf{A}) \cdot \mathbf{F}(\mathbf{r})$$

$$H'\Phi = e^{\frac{ie^*}{\hbar}\chi(\mathbf{r})} H\Psi = e^{\frac{ie^*}{\hbar}\chi(\mathbf{r})} E\Psi = E\Phi$$

This property implies that the kinetic energy

$$\int \Psi^*(\mathbf{r}) \frac{1}{2m} (\mathbf{p} - e\mathbf{A})^2 \Psi(\mathbf{r}) d\mathbf{r}$$

is invariant under a gauge transformation. Besides, since $\mathbf{p} - e\mathbf{A}$ is hermitian, one can also write

$$\begin{aligned} E_{cin} &= \frac{1}{2m} \int \Psi^*(\mathbf{r}) (\mathbf{p} - e^*\mathbf{A})^2 \Psi(\mathbf{r}) d\mathbf{r} \\ &= \frac{1}{2m} \langle \Psi | (\mathbf{p} - e^*\mathbf{A})^2 | \Psi \rangle \\ &= \frac{1}{2m} \langle \Psi | (\mathbf{p} - e^*\mathbf{A})^\dagger (\mathbf{p} - e^*\mathbf{A}) | \Psi \rangle \\ &= \frac{1}{2m} \| (\mathbf{p} - e^*\mathbf{A}) | \Psi \rangle \|^2 \\ &= \frac{1}{2m} \int \left| \left(\mathbf{p} - \frac{e^*\mathbf{A}}{c} \right) \Psi(\mathbf{r}) \right|^2 d\mathbf{r} \end{aligned}$$

Similarly, one can show that, for the gap Δ , the form of the energy that is invariant under a gauge transformation is given in term of the gradient by :

$$\frac{1}{2m} |(-i\hbar\nabla + 2e\mathbf{A})\Delta|^2$$

Neglecting the normal fraction, the Landau-Ginzburg free energy is thus a functional of the fields $\mathbf{A}(\mathbf{r})$ and $\Psi(\mathbf{r})$ given by

$$F[\mathbf{A}, \Psi] = \int d\mathbf{r} \left(\alpha |\Psi|^2 + \frac{\beta}{2} |\Psi|^4 + \frac{1}{2m} |(-i\hbar\nabla + 2e\mathbf{A})\Psi|^2 + \frac{B^2}{8\pi} \right) \quad (6.24)$$

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 satisfy, 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 $\delta\mathbf{A}$ of the vector potential :

$$\begin{aligned}
F[\mathbf{A} + \delta\mathbf{A}] &= \int \frac{(\nabla \times (\mathbf{A} + \delta\mathbf{A}))^2}{8\pi} d\mathbf{r} + \int d\mathbf{r} \left(\alpha |\Psi|^2 + \frac{\beta}{2} |\Psi|^4 \right) \\
&+ \int \frac{d\mathbf{r}}{2m} (-i\hbar \nabla \Psi + 2e(\mathbf{A} + \delta\mathbf{A})\Psi) (i\hbar \nabla \Psi^* + 2e(\mathbf{A} + \delta\mathbf{A})\Psi^*) \\
&\simeq F[\mathbf{A}] + \int \frac{\delta\mathbf{A}}{2m} \left[2e\Psi (i\hbar \nabla \Psi^* + 2e\mathbf{A}\Psi^*) \right. \\
&+ \left. 2e\Psi^* (-i\hbar \nabla \Psi + 2e\mathbf{A}\Psi) \right] d\mathbf{r} \\
&+ \int \frac{(\nabla \times \mathbf{A}) \cdot (\nabla \times \delta\mathbf{A})}{4\pi}
\end{aligned}$$

But

$$\begin{aligned}
\nabla \cdot (\mathbf{X} \times \mathbf{Y}) &= (\nabla \times \mathbf{X}) \cdot \mathbf{Y} - \mathbf{X} \cdot (\nabla \times \mathbf{Y}) \\
\mathbf{B} \cdot \nabla \times \delta\mathbf{A} &= (\nabla \times \mathbf{B}) \cdot \delta\mathbf{A} + \nabla \cdot (\delta\mathbf{A} \times \mathbf{B})
\end{aligned}$$

and

$$\int \nabla \cdot (\delta\mathbf{A} \times \mathbf{B}) = \int \int_S \mathbf{n} \cdot (\delta\mathbf{A} \times \mathbf{B}) d\sigma$$

We limit ourselves to $\delta\mathbf{A} = \mathbf{0}$ at the surface of the superconductor. Then,

$$\int \frac{\mathbf{B} \cdot \nabla \times \delta\mathbf{A}}{4\pi} = \int \frac{\nabla \times \mathbf{B} \cdot \delta\mathbf{A}}{4\pi}$$

At lowest order in $\delta\mathbf{A}$, using $\nabla \times \mathbf{B} = \mu_0 \mathbf{j}$, we obtain

$$F[\mathbf{A} + \delta\mathbf{A}] - F[\mathbf{A}] = \int \delta\mathbf{A} \cdot \left(\frac{\mu_0}{4\pi} \mathbf{j} + \frac{ie\hbar}{m} (\Psi \nabla \Psi^* - \Psi^* \nabla \Psi) + \frac{4e^2}{m} \Psi^* \Psi \right)$$

In order for $F[\mathbf{A} + \delta\mathbf{A}] - F[\mathbf{A}]$ to vanish for all $\delta\mathbf{A}$, we must have

$$\frac{\mu_0}{4\pi} \mathbf{j} = \frac{ie\hbar}{m} (\Psi^* \nabla \Psi - \Psi \nabla \Psi^*) - \frac{4e^2}{m} \mathbf{A} \Psi^* \Psi$$

Let us now make a variation around Ψ . For a function of a complex variable $f(z, \bar{z})$, the statement that the variations vanish is equivalent to saying that if $z = re^{i\theta}$ ($r \neq 0$),

$$\left\{ \begin{array}{l} \frac{\partial f}{\partial r} = 0 \\ \frac{\partial f}{\partial \theta} = 0 \end{array} \right\} \Rightarrow \left\{ \begin{array}{l} \frac{\partial f}{\partial z} e^{i\theta} + \frac{\partial f}{\partial \bar{z}} e^{-i\theta} = 0 \\ \frac{\partial f}{\partial z} ire^{i\theta} + \frac{\partial f}{\partial \bar{z}} (-ir)e^{-i\theta} = 0 \end{array} \right\} \Rightarrow \left\{ \begin{array}{l} \frac{\partial f}{\partial z} = 0 \\ \frac{\partial f}{\partial \bar{z}} = 0 \end{array} \right.$$

Let us thus express F as a function of (Ψ, Ψ^*) , and let us write $\frac{\partial F}{\partial \Psi^*} = 0$

$$F = \int d\mathbf{r} \left(\alpha \Psi \Psi^* + \frac{\beta}{2} \Psi^2 \Psi^{*2} + \frac{1}{2m} (-i\hbar \nabla + 2e\mathbf{A}) \Psi (i\hbar \nabla + 2e\mathbf{A}) \Psi^* + \frac{B^2}{8\pi} \right)$$

$$\begin{aligned} F(\Psi, \Psi^* + \delta\Psi^*) - F(\Psi, \Psi^*) &= \int d\mathbf{r} \left[\alpha \Psi + \beta |\Psi|^2 \Psi + \frac{1}{2m} (-i\hbar \nabla \Psi + 2e\mathbf{A} \Psi) \right. \\ &\quad \left. \times (i\hbar \nabla + 2e\mathbf{A}) \right] \delta\Psi^* = 0 \end{aligned}$$

An integration by part of the last term leads to

$$\nabla \Psi \cdot \nabla \delta\Psi^* = \nabla \cdot (\delta\Psi^* \nabla \Psi) - \delta\Psi^* \underbrace{\nabla \cdot \nabla \Psi}_{\Delta \Psi}$$

since $\nabla(p\mathbf{A}) = p\nabla \cdot \mathbf{A} + \mathbf{A} \cdot \nabla p$. So,

$$\begin{aligned} \int \nabla \Psi \cdot \nabla \delta\Psi^* &= \int \nabla \cdot (\delta\Psi^* \nabla \Psi) - \int \delta\Psi^* \Delta \Psi \\ &= \int_S \delta\Psi^* \mathbf{n} \cdot \nabla \Psi - \int \delta\Psi^* \Delta \Psi \end{aligned}$$

But $\int_S \delta\Psi^* \mathbf{n} \cdot \nabla \Psi = 0$ if one limits oneself to $\delta\Psi^*$ that vanish at the surface. So,

$$\int -i\hbar \nabla \Psi \cdot i\hbar \nabla \delta\Psi^* = \int [(-i\hbar)^2 \nabla^2 \Psi] \delta\Psi^*$$

Besides,

$$\begin{aligned} \int \Psi \nabla \delta\Psi^* &= - \int \delta\Psi^* \nabla \Psi + \int \nabla (\Psi \delta\Psi^*) \\ &= - \int \delta\Psi^* \nabla \Psi + \int_S \mathbf{n} (\Psi \delta\Psi^*) d\sigma \end{aligned}$$

But $\int_S \mathbf{n} (\Psi \delta\Psi^*) d\sigma = 0$ if $\delta\Psi^* = 0$ at the surface. Thus,

$$\begin{aligned} &\int d\mathbf{r} (-i\hbar \nabla \Psi + 2e\mathbf{A} \Psi) (i\hbar \nabla + 2e\mathbf{A}) \delta\Psi^* \\ &= \int d\mathbf{r} [(-i\hbar \nabla)^2 \Psi + (2e\mathbf{A})^2 \Psi - i\hbar (2e\mathbf{A}) \nabla \Psi + 2e\mathbf{A} (-i\hbar) \nabla \Psi] \delta\Psi^* \\ &= \int d\mathbf{r} \delta\Psi^* (-i\hbar \nabla + 2e\mathbf{A})^2 \Psi \end{aligned}$$

In order to ensure that $F(\Psi, \Psi^* + \delta\Psi^*) - F(\Psi, \Psi^*) = 0$ at first order in $\delta\Psi^*$, one must have

$$\alpha \Psi + \beta |\Psi|^2 \Psi + \frac{1}{2m} (-i\hbar \nabla + 2e\mathbf{A})^2 \Psi = 0$$

Finally, the minimization of the functional with respect to the order parameter Ψ and to the vector potential \mathbf{A} leads to the following pair of equations, known as the Landau-Ginzburg equations:

$$\frac{\mu_0}{4\pi} \mathbf{j} = \frac{e\hbar}{m} (\Psi^* \nabla \Psi - \Psi \nabla \Psi^*) - \frac{4e^2 \mathbf{A}}{m} \Psi^* \Psi \quad (6.25)$$

$$\alpha \Psi + \beta |\Psi|^2 \Psi + \frac{1}{2m} (-i\hbar \nabla + 2e\mathbf{A})^2 \Psi = 0 \quad (6.26)$$

These equations constitute the basis of the description of the macroscopic properties of superconductors. They contain two characteristic lengths that depend on α and β . 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, Ψ satisfies the equation

$$\alpha \Psi + \beta |\Psi|^2 \Psi - \frac{\hbar^2}{2m} \Delta \Psi = 0$$

According to Landau's theory of phase transitions, $\alpha = \alpha_0 \times (T - T_c)$, 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 Ψ 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 Ψ 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 Ψ only depends on z . We are looking for a non-constant solution in the $z > 0$ half-space. We can therefore multiply by Ψ' and integrate to obtain

$$\begin{aligned} -\frac{\hbar^2}{2m} \Psi'' \Psi' + \alpha \Psi \Psi' + \beta \Psi^3 \Psi' &= 0 \\ -\frac{\hbar^2}{4m} \Psi'^2 + \frac{\alpha}{2} \Psi^2 + \frac{\beta}{4} \Psi^4 &= cst \end{aligned}$$

For this equation to have a solution $\Psi^2 = \frac{-\alpha}{\beta}$ for $z \rightarrow +\infty$, one must have :

$$\frac{-\alpha^2}{2\beta} + \frac{\beta \alpha^2}{4 \beta^2} = cst \Leftrightarrow cst = -\frac{\alpha^2}{4\beta}.$$

We can rewrite the equation as

$$\begin{aligned}\frac{\hbar^2}{4m}\Psi'^2 &= \frac{\beta}{4}\Psi^4 + \frac{\alpha}{2}\Psi^2 + \frac{\alpha^2}{4\beta} \\ &= \frac{\beta}{4}\left(\Psi^4 + \frac{2\alpha}{\beta}\Psi^2 + \frac{\alpha^2}{\beta^2}\right) \\ &= \frac{\beta}{4}\left(\Psi^2 + \frac{\alpha}{\beta}\right)^2\end{aligned}$$

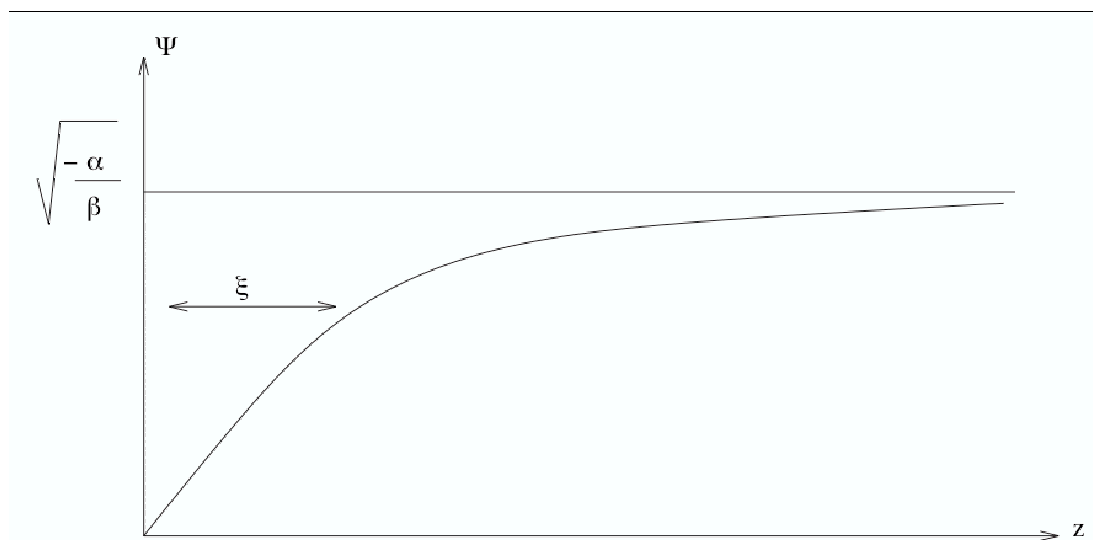
$$\Psi' = -\frac{\sqrt{\beta m}}{\hbar}\left(\Psi^2 + \frac{\alpha}{\beta}\right)$$

$$\Psi = \sqrt{\frac{-\alpha}{\beta}}\tanh\frac{z}{\sqrt{2}\xi}$$

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-Ginzburg 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. Ψ is nearly constant, and \mathbf{j} is approximately given by

$$\frac{\mu_0}{4\pi} \mathbf{j} = \frac{4e^2}{m} |\Psi_0|^2 \mathbf{A}$$

To see this, one writes locally

$$\mathbf{A}(\mathbf{r}) = \underbrace{\mathbf{A}_0}_{\text{uniform}} + \delta \mathbf{A}(\mathbf{r})$$

and one expands in $\delta \mathbf{A}(\mathbf{r})$. To zeroth order, Ψ is uniform according to the second Landau-Ginzburg equation, and the first equation then leads to the result.

Now this equation is nothing but London equation, with

$$\frac{1}{\lambda^2} = \frac{16\pi^2 e^2 - \alpha}{m \beta}$$

Like ξ , $\lambda \sim \frac{1}{\sqrt{T_c - T}}$ when $T \rightarrow T_c$.

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:

$$\begin{aligned} \frac{H_c^2}{8\pi} &= - \left(\alpha |\Psi|^2 + \frac{\beta}{2} |\Psi|^4 \right) \\ &= - \left(\alpha \frac{-\alpha}{\beta} + \frac{\beta \alpha^2}{2 \beta^2} \right) = \frac{\alpha^2}{\beta} \\ H_c^2 &= \frac{4\pi \alpha^2}{\beta} \end{aligned}$$

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-Ginzburg 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 Ψ very small.

To first order in Ψ , the equation reads

$$\frac{1}{2m} (-i\hbar \nabla + 2e\mathbf{A})^2 \Psi = -\alpha \Psi$$

This equation is similar to the Schrödinger equation of a particle of charge $-2e$ in a magnetic field H described by \mathbf{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 ω_c is the cyclotron frequency

$$\omega_c = \frac{2eH}{m}$$

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

$$\begin{aligned} \frac{1}{2} \hbar \frac{2eH_{c2}}{m} &= -\alpha \\ H_{c2} &= -\frac{\alpha m}{\hbar e} \end{aligned}$$

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

$$-\frac{\alpha m}{\hbar e} > \sqrt{\frac{4\pi}{\beta}}(-\alpha)$$

or

$$\sqrt{\beta} > \frac{\sqrt{4\pi}\hbar e}{m}$$

But

$$\begin{aligned} \kappa = \frac{\lambda}{\xi} &= \frac{\sqrt{m}\sqrt{\beta}}{4\sqrt{\pi}e\sqrt{-\alpha}} \frac{\sqrt{2m(-\alpha)}}{\hbar} \\ &= \frac{m\sqrt{\beta}}{\hbar e 2\sqrt{2}\sqrt{\pi}} = \sqrt{\beta} \frac{m}{\sqrt{4\pi}\hbar e} \frac{1}{\sqrt{2}} \end{aligned}$$

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.

6.7 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:

$$\nabla\Psi = (\nabla|\Psi|)e^{i\varphi} + |\Psi|i(\nabla\varphi)e^{i\varphi} = (\nabla|\Psi|)e^{i\varphi} + \Psi i\nabla\varphi$$

so that

$$\begin{aligned} \Psi^*\nabla\Psi - \Psi\nabla\Psi^* &= \Psi^*(\nabla|\Psi|)e^{i\varphi} + |\Psi|^2i\nabla\varphi - [\Psi(\nabla|\Psi|)e^{-i\varphi} + \Psi|\Psi|(-i)(\nabla\varphi)e^{-i\varphi}] \\ &= \nabla|\Psi|(\Psi^*e^{i\varphi} - \Psi e^{-i\varphi}) + |\Psi|^22i\nabla\varphi \\ &= |\Psi|^22i\nabla\varphi \end{aligned}$$

$$\frac{\mu_0}{4\pi}\mathbf{j} = -\frac{2e\hbar}{m}|\Psi|^2\left(\nabla\varphi + \frac{2e}{\hbar}\mathbf{A}\right)$$

But deep inside a superconductor, $\mathbf{j} = \mathbf{0}$. Indeed, $\mathbf{B} = \mathbf{0} \Rightarrow \nabla \times \mathbf{B} = \mathbf{0} \Rightarrow \mathbf{j} = \mathbf{0}$. So

$$\oint \mathbf{j} \cdot d\mathbf{l} = 0 \Rightarrow \underbrace{\oint \nabla\varphi \cdot d\mathbf{l}}_{2n\pi} = \frac{2e}{\hbar c} \underbrace{\oint \mathbf{A} \cdot d\mathbf{l}}_{\int \nabla \times \mathbf{A} \cdot d\mathbf{S}}$$

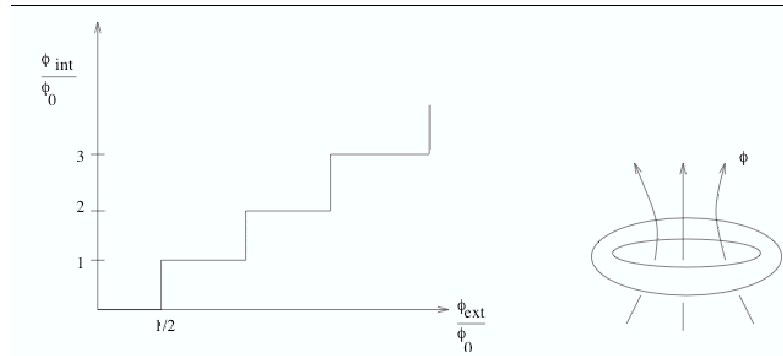
The first term must be equal to $2\pi n$ to ensure that the order parameter is well defined. Besides, $\int \nabla \times \mathbf{A} \cdot d\mathbf{S} = \Phi$, where Φ is the flux of the magnetic field through the contour since $\nabla \times \mathbf{A} = \mathbf{B}$. This leads to :

$$\Phi = n\frac{\pi\hbar}{e} = n\phi_0$$

with $\phi_0 = \frac{\pi\hbar}{e} = \frac{h}{2e}$.

Consequences:

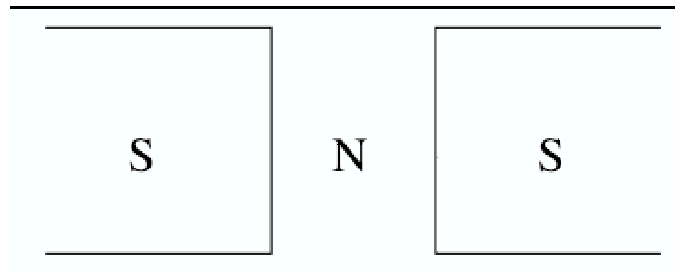
- Penetration of the field in type II superconductors : below H_{c2} , the sample is pierced by tubes of normal state. According to this argument, the flux of the magnetic field through such a tube must be a multiple of ϕ_0 . Now, we know that it is advantageous for the system to maximize its ratio surface/normal volume. So to get a flux $2\phi_0$ through the sample, it is better to create two vortices of flux ϕ_0 than one of flux $2\phi_0$. So the magnetic flux penetrates as vortices of flux $\phi_0 = \frac{h}{2e}$.
- If the charge of the effective charge carriers was e , as first assumed by Landau and Ginzburg, and not $2e$, the flux would be quantized in units of $\frac{h}{e}$ and not $\frac{h}{2e}$. 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 $\frac{h}{2e}$. 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 ϕ_0 .



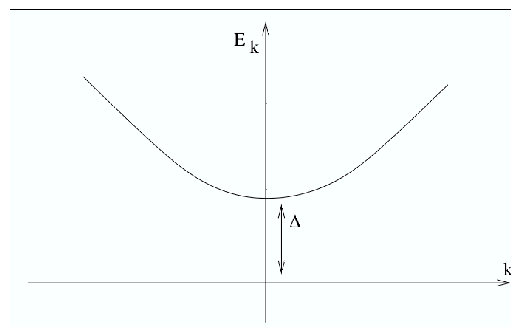
- Persistent currents : if one induces a current in a superconducting ring, this current must create a magnetic field whose flux is of the form $n\phi_0$. To decay, this current must overcome a potential barrier to reach the states $(n - 1)\phi_0$, $(n - 2)\phi_0$, ... Now, this barrier is very high 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.

6.8 Josephson effect:

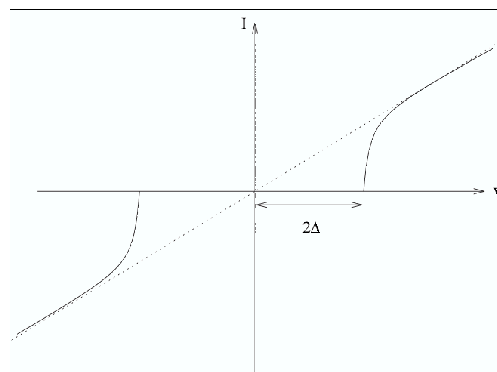
Let us consider two superconductors connected by a normal region



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 2Δ in total (Δ 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 2Δ .

In 1962, Josephson 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-Ginzburg 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 \mathbf{A} is equal to $\mathbf{0}$, and without any voltage, a current can flow if the phase of the order parameter is not uniform:

$$\frac{\mu_0}{4\pi} \mathbf{j} = -\frac{2e\hbar}{m} |\Psi|^2 \nabla \phi$$

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 Ψ . The simplest calculation consists in going back to the Landau-Ginzburg equation for Ψ :

$$-\frac{\hbar^2}{2m} \Delta \Psi + \alpha \Psi + \beta |\Psi|^2 \Psi = 0$$

Inside the junction, Ψ changes very fast, and the term $\Delta \Psi$ dominates

$$\Rightarrow \Delta \Psi = 0$$

Let us suppose that the phase is equal to ϕ_1 in the left superconductor, and to ϕ_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 $\Psi(\mathbf{r}) = e^{i\phi_1} |\Psi_0|$ far on the left and $\Psi(\mathbf{r}) = e^{i\phi_2} |\Psi_0|$ far on the right.

Let us look for Ψ as

$$\Psi(\mathbf{r}) = |\Psi_0| [f(\mathbf{r}) e^{i\phi_1} + (1 - f(\mathbf{r})) e^{i\phi_2}]$$

where $f(\mathbf{r})$ 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

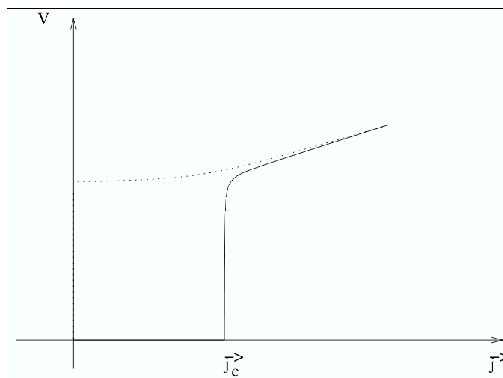
$$\begin{aligned} \frac{\mu_0}{4\pi} \mathbf{j} &= \frac{ei\hbar}{m} (\Psi^* \nabla \Psi - \Psi \nabla \Psi^*) \\ \Psi^* &= |\Psi_0| (f(\mathbf{r})e^{-i\phi_1} + (1-f(\mathbf{r}))e^{-i\phi_2}) \\ \nabla \Psi &= |\Psi_0| \nabla f(\mathbf{r})(e^{i\phi_1} - e^{i\phi_2}) \\ \Psi^* \nabla \Psi &= |\Psi_0|^2 \nabla f(\mathbf{r}) [f(\mathbf{r})e^{-i\phi_1} + (1-f(\mathbf{r}))e^{-i\phi_2}] (e^{i\phi_1} - e^{i\phi_2}) \\ &= |\Psi_0|^2 \nabla f(\mathbf{r}) [f - fe^{-i\phi_1+i\phi_2} - (1-f) + (1-f)e^{-i\phi_2+i\phi_1}] \\ \Psi \nabla \Psi^* &= (\Psi^* \nabla \Psi)^* \\ \Psi^* \nabla \Psi - \Psi \nabla \Psi^* &= |\Psi_0|^2 \nabla f(\mathbf{r}) [-f(e^{-i\phi_1+i\phi_2} - e^{i\phi_1-i\phi_2}) \\ &\quad + (1-f)(e^{-i\phi_2+i\phi_1} - e^{i\phi_2-i\phi_1})] \\ &= |\Psi_0|^2 \nabla f(\mathbf{r}) (e^{-i\phi_2+i\phi_1} - e^{i\phi_2-i\phi_1}) \\ &= 2i \sin(\phi_1 - \phi_2) |\Psi_0|^2 \nabla f(\mathbf{r}) \\ \frac{\mu_0}{4\pi} \mathbf{j} &= -\frac{2e\hbar}{m} |\Psi_0|^2 \nabla f \sin(\phi_1 - \phi_2) \end{aligned}$$

The maximum current is reached when $\phi_1 - \phi_2 = \frac{\pi}{2}$, and it is of the order of

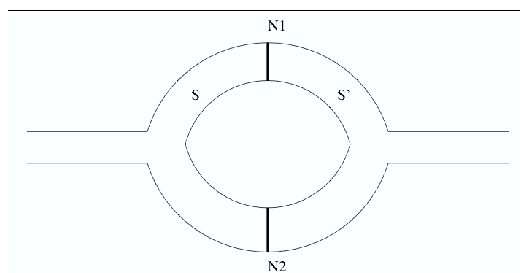
$$\frac{\mu_0}{4\pi} j_c = -\frac{2e\hbar}{m} |\Psi_0|^2 \frac{1}{L}$$

where L is the width of the junction (∇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 Josephson junctions according to the following geometry,



one can show that the maximal current is a periodic function of the flux ϕ through the ring, of periodicity ϕ_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 7

Appendix - Second quantization

7.1 Introduction

When dealing with systems of interacting particles, one must work with many-particle wave functions. We assume that particles of the same type (e.g., electrons) are indistinguishable. Consequently, the state of a system of identical particles must be invariant under the exchange of the coordinates of any two particles. To illustrate, consider a system of two point-like particles. Since they are indistinguishable, exchanging their coordinates cannot alter any physical properties. The two wavefunctions can thus differ at most by a phase factor:

$$\psi(x_1, x_2) = e^{i\alpha}\psi(x_2, x_1).$$

What happens if we exchange their position twice?

To answer this question precisely, one must carefully define the exchange process. We can adopt the reference frame of one particle. From this perspective, a double exchange is equivalent to the other particle tracing a full circle around the origin. In three dimensions, this path can be continuously deformed to a point by moving it out of the original plane. A double exchange must therefore recover the original wavefunction, implying:

$$e^{2i\alpha} = 1 \Rightarrow e^{i\alpha} = \pm 1$$

We thus recover the requirement that wavefunctions must be either fully symmetric (bosons) or fully antisymmetric (fermions) under particle permutation. Therefore, if one denotes by $\{\psi_i(x)\}$ a basis of one-particle wavefunctions, one cannot simply use the set of functions

$$\Psi(x_1, \dots, x_N) = \psi_{i_1}(x_1) \dots \psi_{i_N}(x_N)$$

as a basis of the N -particle Hilbert space. One must instead build N particle wave functions that are totally symmetric or totally antisymmetric when exchanging the coordinates of particles.

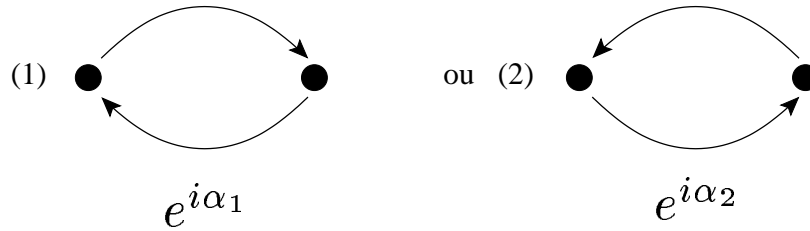


Figure 7.1: The two possible ways to exchange particles in two-dimensions. In three dimensions (or more), these paths can both be deformed to a point, and they become equivalent.

Remark: This argument relies on several implicit assumptions. First, it presupposes that no subtleties arise from changing reference frames; this assumption requires refinement for charged particles in a magnetic field. Second, the particles are treated as point-like; internal structure would necessitate defining how it transforms under permutation. We will largely overlook these complications in this course.

Nonetheless, statistics beyond fermions and bosons are possible. For instance, in two dimensions, two distinct topologically inequivalent paths exist for exchanging two particles (see Fig. 7.1). From the perspective of one particle, the exchange can correspond to the other particle circling clockwise or counterclockwise. Since the particles are confined to the plane, these two paths cannot be continuously deformed into one another without crossing the reference particle. They are topologically distinct. The system returns to its initial state only after applying both exchanges successively, implying:

$$e^{i(\alpha_1+\alpha_2)} = 1 \Leftrightarrow \alpha_2 = -\alpha_1$$

However, α_1 need not be 0 or π , allowing for phase factors other than ± 1 . Particles exhibiting such exotic statistics are called *anyons*. Anyonic statistics describe excitations in a two-dimensional electron gas under a strong magnetic field, such as in the fractional quantum Hall effect. Even more complex statistics are possible, including *non-Abelian anyons*, where the exchange operation is represented by a matrix acting on a degenerate space of multiple particle types.

7.2 Bosons

7.2.1 Fock space

We consider an orthonormal basis of the one-particle Hilbert space $\{\psi_i(x)\}$. We want to construct bosonic 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

$$\Psi_{ii}(x_1, x_2) = \psi_i(x_1)\psi_i(x_2)$$

is symmetric and is thus acceptable. For two different states i and j , one must build the wave function

$$\Psi_{ij}(x_1, x_2) = \frac{1}{\sqrt{2}} [\psi_i(x_1)\psi_j(x_2) + \psi_j(x_1)\psi_i(x_2)]$$

to get a symmetric wave function. Note that if we have M possible orbitals, we have only $M + \frac{M(M-1)}{2} = \frac{M(M+1)}{2}$ possible two-body states, instead of M^2 for distinguishable particles.

More generally, a N -particle wave function is uniquely defined by the number of times N_j each one-particle wave function j appears. To build a fully symmetric, we can write

$$|N_1, N_2, \dots\rangle = \Psi(x_1, \dots, x_N) = \frac{1}{\sqrt{\mathcal{N}}} \sum_{P \in \mathcal{S}_N} \psi_{i_{P(1)}}(x_1) \dots \psi_{i_{P(N)}}(x_N) \quad (7.1)$$

where \mathcal{S}_N is the group of permutations of 1 to N and \mathcal{N} is a normalization factor.

Proof that Eq. (7.1) is invariant under permutation.

Let σ a permutation of \mathcal{S}_N . The action of σ on Ψ is

$$\sigma\Psi(x_1, \dots, x_N) = \Psi(x_{\sigma(1)}, \dots, x_{\sigma(N)}),$$

such that

$$\begin{aligned} \sigma\Psi(x_1, \dots, x_N) &= \frac{1}{\sqrt{\mathcal{N}}} \sum_{P \in \mathcal{S}_N} \psi_{i_{P(1)}}(x_{\sigma(1)}) \dots \psi_{i_{P(N)}}(x_{\sigma(N)}) \\ &= \frac{1}{\sqrt{\mathcal{N}}} \sum_{P \in \mathcal{S}_N} \psi_{i_{P(\sigma^{-1}(1))}}(x_1) \dots \psi_{i_{P(\sigma^{-1}(N))}}(x_N). \end{aligned}$$

To proceed, we need to remember that $\sigma\mathcal{S}_N = \mathcal{S}_N$, i.e. that σ simply redistributes the permutations, such that

$$\sigma\Psi(x_1, \dots, x_N) = \frac{1}{\sqrt{\mathcal{N}}} \sum_{\tilde{P} \in \mathcal{S}_N} \psi_{i_{\tilde{P}(1)}}(x_1) \dots \psi_{i_{\tilde{P}(N)}}(x_N) = \Psi(x_1, \dots, x_N).$$

The set of all $|N_1, N_2, \dots\rangle$ is the *Fock space*. It can also be convenient to write Eq. (7.1) as

$$\Psi(x_1, \dots, x_N) = \frac{1}{\sqrt{\mathcal{N}'}} \sum_{P \in \mathcal{S}'_N} \psi_{i_{P(1)}}(x_1) \dots \psi_{i_{P(N)}}(x_N) \quad (7.2)$$

where \mathcal{S}'_N is the set of permutations that does not let the sequence i_1, \dots, i_N invariant, without double-counting (i.e., if several permutations lead to the same sequence, we only counts them once). Mathematically, \mathcal{S}'_N is the coset $\mathcal{S}_N / (S_{N_1} \times \dots \times S_{N_M})$, where N_j is the number of times the orbital j is occupied. \mathcal{N}' is then simply the number of elements of \mathcal{S}'_N and is equal to :

$$\boxed{\mathcal{N}' = \frac{N!}{N_1! N_2! \dots}} \quad (7.3)$$

with the standard convention $N_i! = 1$ if $N_i = 0$, and of course $\sum_i N_i = N$.

Proof of normalization: As the basis is orthonormal,

$$\int dx \psi_i(x)^* \psi_j(x) = \delta_{i,j}.$$

We can write

$$\langle N_1, \dots | N_1, \dots \rangle = \frac{1}{\mathcal{N}'} \sum_{P, Q \in \mathcal{S}'_N} \langle \psi_{P(1)} | \psi_{Q(1)} \rangle \dots \langle \psi_{P(N)} | \psi_{Q(N)} \rangle.$$

As there is no repeat, $\langle \psi_{P(1)} | \psi_{Q(1)} \rangle \dots \langle \psi_{P(N)} | \psi_{Q(N)} \rangle \neq 0$ if and only if $P = Q$, and therefore

$$\langle N_1, \dots | N_1, \dots \rangle = \frac{1}{\mathcal{N}'} \sum_{P \in \mathcal{S}'_N} 1 = \frac{|\mathcal{S}'_N|}{\mathcal{N}'}$$

Proof of equivalence: To prove the equivalence of the definitions, it is enough to prove that Eq. (7.2) is invariant under all permutations. For any permutation σ in \mathcal{S}_N , it is also quite straightforward that $\sigma \mathcal{S}'_N \approx \mathcal{S}'_N$. From there, one can use the same proof as for the previous expression.

Example : 3 particles using ψ_1 twice and ψ_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 : $\frac{3!}{2!1!} = 3$

7.2.2 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.

$$H = \frac{p^2}{2m} + \frac{1}{2}m\omega^2x^2 \quad [x, p] = i\hbar$$

1. Change of variables :
$$\begin{cases} X = \sqrt{\frac{m\omega}{\hbar}}x \\ P = \frac{1}{\sqrt{m\hbar\omega}}p \end{cases}$$

$$\Rightarrow H = \frac{\hbar\omega}{2}(X^2 + P^2) \quad [X, P] = i$$

2.

$$\begin{cases} a = \frac{1}{\sqrt{2}}(X + iP) \\ a^\dagger = \frac{1}{\sqrt{2}}(X - iP) \end{cases} \Rightarrow \begin{cases} X = \frac{1}{\sqrt{2}}(a^\dagger + a) \\ P = \frac{i}{\sqrt{2}}(a^\dagger - a) \end{cases}$$

3.

$$[a, a^\dagger] = 1$$

$$a^\dagger a = \frac{1}{2}(X^2 + P^2 - 1) \Rightarrow H = \left(a^\dagger a + \frac{1}{2}\right) \hbar\omega$$

4. Let us define $N = a^\dagger a$, 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 ν is an eigenvalue of N , then $\nu \geq 0$.
Indeed, $\langle \varphi | a^\dagger a | \varphi \rangle = \|a|\varphi\rangle\|^2 \geq 0$
- 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,

$$\begin{aligned} a^\dagger a a |\varphi_\nu\rangle &= (a a^\dagger - 1) a |\varphi_\nu\rangle = a \underbrace{a^\dagger a |\varphi_\nu\rangle}_{\nu |\varphi_\nu\rangle} - a |\varphi_\nu\rangle \\ &= (\nu - 1) a |\varphi_\nu\rangle \end{aligned}$$

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.

- $a^\dagger|\varphi_\nu\rangle$ is never equal to zero.
Indeed, $\|a^\dagger|\varphi_\nu\rangle\|^2 = \langle\varphi_\nu|aa^\dagger|\varphi_\nu\rangle = (\nu+1)\langle\varphi_\nu|\varphi_\nu\rangle$.
Now $\nu \geq 0 \Rightarrow \nu+1 > 0$.
- $a^\dagger|\varphi_\nu\rangle$ is an eigenvector of N with eigenvalue $\nu+1$.
Indeed, $a^\dagger aa^\dagger|\varphi_\nu\rangle = a^\dagger(a^\dagger a + 1)|\varphi_\nu\rangle = (\nu+1)a^\dagger|\varphi_\nu\rangle$

Finally, ν can take all positive integer values, including 0.

- 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 $|n\rangle$.
- Let $|n\rangle$ and $|n+1\rangle$ be two normalized states. One knows from the above that $a^\dagger|n\rangle \propto |n+1\rangle$. To find the proportionality coefficient, let us write :

$$a^\dagger|n\rangle = c_{n+1}|n+1\rangle$$

Let us take the norm of each member :

$$\begin{aligned} |c_{n+1}|^2 \underbrace{\langle n+1|n+1\rangle}_1 &= \langle n|aa^\dagger|n\rangle = \langle n|1+a^\dagger a|n\rangle = n+1 \\ \Rightarrow c_{n+1} &= \sqrt{n+1} \end{aligned}$$

Similarly, writing $a|n\rangle = c_{n-1}|n-1\rangle$, one finds :

$$|c_{n-1}|^2 = \langle n|a^\dagger a|n\rangle = n \quad \Rightarrow \quad c_{n-1} = \sqrt{n}$$

Thus,

$$\begin{cases} a^\dagger|n\rangle &= \sqrt{n+1}|n+1\rangle \\ a|n\rangle &= \sqrt{n}|n-1\rangle \end{cases}$$

7.2.3 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^\dagger defined by:

$$a_i|N_1 \dots N_i \dots\rangle = \begin{cases} \sqrt{N_i}|N_1 \dots N_i - 1 \dots\rangle & \text{if } N_i \geq 1 \\ 0 & \text{otherwise} \end{cases} \quad (7.4)$$

$$a_i^\dagger|N_1 \dots N_i \dots\rangle = \sqrt{N_i + 1}|N_1 \dots N_i + 1 \dots\rangle \quad (7.5)$$

Let us investigate the properties of these operators:

$$\begin{aligned} a_i^\dagger a_i |N_1 \dots N_i \dots\rangle &= \sqrt{N_i} a_i^\dagger |N_1 \dots N_i - 1 \dots\rangle \\ &= N_i |N_1 \dots N_i \dots\rangle \\ \Rightarrow a_i^\dagger a_i &= N_i \text{ diagonal} \end{aligned}$$

$$\begin{aligned} a_i a_i^\dagger |N_1 \dots N_i \dots\rangle &= \sqrt{N_i + 1} a_i |N_1 \dots N_i + 1 \dots\rangle \\ &= (N_i + 1) |N_1 \dots N_i \dots\rangle \\ \Rightarrow a_i a_i^\dagger &= N_i + 1 \end{aligned}$$

This implies that :

$$\begin{aligned} a_i a_i^\dagger - a_i^\dagger a_i &= 1 \\ \left. \begin{aligned} a_i a_j &= a_j a_i \\ a_i^\dagger a_j &= a_j a_i^\dagger \\ a_i^\dagger a_j^\dagger &= a_j^\dagger a_i^\dagger \end{aligned} \right\} \text{ if } i \neq j \end{aligned}$$

Or, to summarize

$$\boxed{[a_i, a_j^\dagger] = \delta_{ij}} \quad (7.6)$$

$$\boxed{[a_i, a_j] = [a_i^\dagger, a_j^\dagger] = 0} \quad (7.7)$$

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 . For example, $\hat{f}_a^{(1)}$ could be a potential $V(x_a)$ or a differential operator ∂_{x_a} . We define the matrix elements of the operator $\hat{F}^{(1)}$ 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:

$$\begin{cases} \langle N_1 \dots N_i \dots | \hat{F}^{(1)} | N_1 \dots N_i \dots \rangle &= \sum_i f_{ii}^{(1)} N_i \\ \langle \dots N_i \dots N_k - 1 \dots | \hat{F}^{(1)} | \dots N_i - 1 \dots N_k \dots \rangle &= f_{ik}^{(1)} \sqrt{N_i N_k} \end{cases}$$

Proof

$$\langle N_1 N_2 \dots | \hat{F}^{(1)} | N_1 N_2 \dots \rangle = \sum_i f_{ii}^{(1)} N_i$$

$|N_1 N_2 \dots\rangle$ 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 is the index of the function ϕ_i of the variable x_a . Since the variables x_a enter all the functions ϕ_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 side 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 :

$$\begin{aligned} |\dots N_i - 1 \dots N_k \dots\rangle &\rightarrow \sqrt{\frac{N_k! (N_i - 1)!}{N!}} \\ \langle \dots N_i \dots N_k - 1 \dots | &\rightarrow \sqrt{\frac{N_i! (N_k - 1)!}{N!}} \end{aligned}$$

So the coefficient of $f_{ik}^{(1)}$ is given by :

$$\underbrace{\frac{N!}{\dots N_k! (N_i - 1)! \dots}}_{\text{number of terms in the ket}} \underbrace{N_k}_{\text{number of operators in } \hat{F}^{(1)} \text{ that give } f_{ik}^{(1)}} \underbrace{\sqrt{\frac{\dots N_k! (N_i - 1)! \dots}{N!}} \sqrt{\frac{\dots (N_k - 1)! N_i! \dots}{N!}}}_{\text{normalization coefficients}}$$

which finally leads to $\sqrt{N_i N_k}$.

The operator $\hat{F}^{(1)}$ can thus be written :

$$\boxed{\hat{F}^{(1)} = \sum_{ik} f_{ik}^{(1)} a_i^\dagger a_k} \quad (7.8)$$

Proof

- Diagonal elements

$$\langle N_1 N_2 \dots | a_i^\dagger a_k | N_1 N_2 \dots \rangle = 0 \quad \text{if } i \neq k$$

$$\begin{aligned} \Rightarrow \langle N_1 N_2 \dots | \hat{F}^{(1)} | N_1 N_2 \dots \rangle &= \sum_i f_{ii}^{(1)} \langle N_1 \dots | a_i^\dagger a_i | N_1 \dots \rangle \\ &= \sum_i f_{ii}^{(1)} N_i \quad \text{Q.E.D.} \end{aligned}$$

- Off-diagonal elements

$$\begin{aligned} \langle \dots N_i N_k - 1 \dots | a_j^\dagger a_l | \dots N_i - 1 N_k \dots \rangle &= 0 \quad \text{except if } l = k \text{ and } i = j \\ \langle \dots N_i N_k - 1 \dots | a_i^\dagger a_k | \dots N_i - 1 N_k \dots \rangle &= \sqrt{N_i N_k} \quad \text{Q.E.D.} \end{aligned}$$

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 :

$$\hat{F}^{(2)} = \frac{1}{2} \sum_{iklm} (\hat{f}^{(2)})_{lm}^{ik} a_i^\dagger a_k^\dagger a_l a_m \quad (7.9)$$

$$\text{with } (\hat{f}^{(2)})_{lm}^{ik} = \int \int \psi_i^*(x_1) \psi_k^*(x_2) \hat{f}^{(2)} \psi_l(x_2) \psi_m(x_1) dx_1 dx_2$$

The proof is left as an exercise, and follows straightforwardly from the one-body example. Finally, the Fock space itself can be described simply in terms of creation and annihilation operators. Indeed, using repeatedly the identity :

$$|N_1 \dots N_i + 1 \dots \rangle = \frac{a_i^\dagger}{\sqrt{N_i + 1}} |N_1 \dots N_i \rangle$$

one can write :

$$|N_1 \dots N_i \dots \rangle = \prod_i \frac{(a_i^\dagger)^{N_i}}{\sqrt{N_i!}} |0 \rangle \quad (7.10)$$

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.

7.3 Fermions

In the case of fermions, the wave function must be fully antisymmetric. Such a wave function can only be built out of different one-particle wavefunctions as Slater determinants:

$$\Psi(x_1, \dots, x_N) = \frac{1}{\sqrt{N!}} \begin{vmatrix} \psi_{p_1}(x_1) & \psi_{p_1}(x_2) & \dots & \psi_{p_1}(x_N) \\ \psi_{p_2}(x_1) & \psi_{p_2}(x_2) & \dots & \psi_{p_2}(x_N) \\ \vdots & \vdots & & \vdots \\ \psi_{p_N}(x_1) & \psi_{p_N}(x_2) & \dots & \psi_{p_N}(x_N) \end{vmatrix} \quad (7.11)$$

Due to the Pauli principle, in Fock space, each state is defined by an occupation number equal to 0 or 1:

$$|N_1 \dots N_i \dots\rangle$$

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 wavefunctions once for all and keep this ordering. The determinant can then be calculated according to the chosen convention. In this section, we will take the convention

$$p_1 < p_2 < \dots < p_N.$$

We now turn to the one-body operators, and define the matrix elements as for the bosons by

$$f_{ij}^{(1)} = \int \psi_i^*(x) \hat{f}^{(1)} \psi_j(x) dx.$$

Let us consider the example of an operator $F^{(1)}$ whose only non-zero matrix element is $f_{i_0 j_0}^{(1)} = 1$, with $i_0 < j_0$. Under the action of $F^{(1)}$, the orbital j_0 , if present, is replaced by the orbital i_0 , if originally absent, in Eq. (7.11).

$$F^{(1)} |N_1 \dots N_{i_0} = 0 \dots N_{j_0} = 1 \dots\rangle = \frac{1}{\sqrt{N!}} \begin{vmatrix} \psi_{p_1}(x_1) & \psi_{p_1}(x_2) & \dots & \psi_{p_1}(x_N) \\ \vdots & \vdots & & \vdots \\ \psi_{i_0}(x_1) & \psi_{i_0}(x_2) & \dots & \psi_{i_0}(x_N) \\ \vdots & \vdots & & \vdots \\ \psi_{p_N}(x_1) & \psi_{p_N}(x_2) & \dots & \psi_{p_N}(x_N) \end{vmatrix}$$

This state is proportional to $|N_1 \dots N_{i_0} = 1 \dots N_{j_0} = 0 \dots\rangle$, but the ordering does not necessarily satisfy our chosen convention. In fact,

$$\langle \dots N_{i_0} = 1 \dots N_{j_0} = 0 \dots | F^{(1)} | \dots N_{i_0} = 0 \dots N_{j_0} = 1 \dots \rangle = f_{i_0 j_0}^{(1)} (-1)^{\sum_{n=i_0+1}^{j_0-1} N_n}.$$

The phase factor is called a *string*. Consequently, to represent this operator with creation and annihilation operators, by analogy to bosons, they must be defined by:

$$c_i^\dagger |N_1, \dots, N_i, \dots\rangle = (-1)^{\sum_{n=1}^{i-1} N_n} (1 - N_i) |N_1, \dots, N_i + 1, \dots\rangle \quad (7.12)$$

$$c_i |N_1, \dots, N_i, \dots\rangle = (-1)^{\sum_{n=1}^{i-1} N_n} N_i |N_1, \dots, N_i - 1, \dots\rangle \quad (7.13)$$

This leads to the following anticommutation relations (Clifford algebra):

$$\{c_i^\dagger, c_k\} \equiv c_i^\dagger c_k + c_k c_i^\dagger = \delta_{ik} \quad (7.14)$$

$$\{c_i^\dagger, c_k^\dagger\} = \{c_i, c_k\} = 0 \quad (7.15)$$

Proof of the commutation relations.

$$i = k$$

$$\begin{aligned} c_i^\dagger c_i |0_i\rangle &= 0 \\ c_i c_i^\dagger |0_i\rangle &= c_i (-1)^{\sum(1, i-1)} |1_i\rangle = |0_i\rangle \\ c_i^\dagger c_i |1_i\rangle &= c_i^\dagger (-1)^{\sum(1, i-1)} |0_i\rangle = |1_i\rangle \\ c_i c_i^\dagger |1_i\rangle &= 0 \\ &\Rightarrow c_i^\dagger c_i + c_i c_i^\dagger = 1 \end{aligned}$$

$$i < k$$

$$\begin{aligned} c_i^\dagger c_k^\dagger |N_1 \dots\rangle &= (-1)^{N_1 + \dots + N_{k-1}} c_i^\dagger (1 - N_k) |N_1 \dots N_k + 1 \dots\rangle \\ &= (-1)^{N_1 + \dots + N_{k-1}} (-1)^{N_1 + \dots + N_{i-1}} (1 - N_i) (1 - N_k) |\dots N_i + 1 \dots N_k + 1 \dots\rangle \\ c_k^\dagger c_i^\dagger |N_1 \dots\rangle &= (-1)^{N_1 + \dots + N_{i-1}} c_k^\dagger (1 - N_i) |N_1 \dots N_i + 1 \dots N_k\rangle \\ &= (-1)^{N_1 + \dots + N_{i-1}} (-1)^{N_1 + \dots + N_{i+1} + N_{i+1} + \dots + N_{k-1}} \\ &\quad (1 - N_i) (1 - N_k) |\dots N_i + 1 \dots N_k + 1 \dots\rangle \\ &\Rightarrow c_i^\dagger c_k^\dagger = -c_k^\dagger c_i^\dagger \Rightarrow c_i^\dagger c_k^\dagger + c_k^\dagger c_i^\dagger = 0 \\ &\quad \text{or } \{c_i^\dagger, c_k^\dagger\} = 0 \end{aligned}$$

Similarly for $\{c_i, c_k^\dagger\}$ or $\{c_i, c_k\}$

Some remarks on anticommutation relations :

- The order of the orbitals is completely arbitrary. It is possible to reorder our convention, changing the definition of the string. While mathematically straightforward, it is a highly non-local transformation. The string is the core difference between fermions and hard-core bosons/spin operators.
- $c_i^\dagger c_i^\dagger = c_i c_i = 0$. This is Pauli's principle: two fermions cannot occupy the same quantum state.

- Besides, one indeed has:

$$\begin{aligned} c_i^\dagger c_k |\dots N_i = 0 \dots N_k = 1 \dots\rangle &= (-1)^{N_1 + \dots + N_{i-1} + N_{i+1} + \dots + N_{k-1}} c_i^\dagger |\dots N_i = 0 \dots N_k = 0 \dots\rangle \\ &= (-1)^{2 \times (N_1 + \dots + N_{i-1}) + N_{i+1} + \dots + N_{k-1}} |\dots N_i = 1 \dots N_k = 0 \dots\rangle \\ &= (-1)^{N_{i+1} + \dots + N_{k-1}} |\dots N_i = 1 \dots N_k = 0 \dots\rangle \end{aligned}$$

- Finally, $N = \sum_i c_i^\dagger c_i$ satisfies :

$$N |N_1 \dots\rangle = \sum_i N_i |N_1 \dots\rangle \Rightarrow N = \text{number of particles}$$

With these definitions, the expressions that we have derived for bosons still apply:

$$\hat{F}^{(1)} = \sum_{ik} f_{ik}^{(1)} c_i^\dagger c_k \quad (7.16)$$

and

$$\hat{F}^{(2)} = \frac{1}{2} \sum_{iklm} (\hat{f}^{(2)})_{lm}^{ik} c_i^\dagger c_k^\dagger c_l c_m \quad (7.17)$$

$$\text{with } (\hat{f}^{(2)})_{lm}^{ik} = \int \int \psi_i^*(x_1) \psi_k^*(x_2) \hat{f}^{(2)} \psi_l(x_2) \psi_m(x_1) dx_1 dx_2$$

For the two-body matrix elements, **the order of the indices now matters** since the operators anticommute.

7.4 Field operator

It is very useful to use linear combinations of creation and annihilation operators called *field operators* and defined by :

$$\begin{aligned} \Psi(x) &= \sum_i \varphi_i(x) a_i \\ \Psi^\dagger(x) &= \sum_i \varphi_i^*(x) a_i^\dagger \end{aligned} \quad (7.18)$$

Since the operator a_i^\dagger creates one particle in the state of wave function $\varphi_i(x)$, the operator $\Psi^\dagger(x_0)$ 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.

$$\langle x | \Psi^\dagger(x_0) | 0 \rangle = \delta(x - x_0) \quad (7.19)$$

Proof

$$\begin{aligned} \langle x | a_i^\dagger | 0 \rangle &= \varphi_i(x) \\ \Rightarrow \langle x | \Psi^\dagger(x_0) | 0 \rangle &= \sum_i \varphi_i^*(x_0) \langle x | a_i^\dagger | 0 \rangle \\ \sum_i \varphi_i^*(x_0) \varphi_i(x) &= \sum_i \varphi_i^*(x_0) \varphi_i(x) \\ &= \delta(x - x_0) \end{aligned}$$

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 :

$$\hat{F}^{(1)} = \int \Psi^\dagger(x) \hat{f}^{(1)} \Psi(x) dx \quad (7.20)$$

$$\hat{F}^{(2)} = \frac{1}{2} \int \int \Psi^\dagger(x) \Psi^\dagger(x') \hat{f}^{(2)} \Psi(x') \Psi(x) dx dx' \quad (7.21)$$

where $\hat{f}^{(1)}$ and $\hat{f}^{(2)}$ act on the functions of x and x' .

Indeed,

$$\hat{F}^{(1)} = \sum_{ij} a_i^\dagger a_j \int \varphi_i^*(x) \hat{f}^{(1)} \varphi_j(x) dx$$

or

$$\hat{F}^{(1)} = \sum_{ij} f_{ij}^{(1)} a_i^\dagger a_j$$

The same applies to $\hat{F}^{(2)}$.

The commutation relations of the field operators can be deduced from those of the operators a_i^\dagger and a_i . They are of course different for fermions and bosons :

- Bosons :

$$\boxed{[\Psi(x), \Psi^\dagger(x')] = \delta(x - x')} \quad (7.22)$$

$$[\Psi(x), \Psi(x')] = [\Psi^\dagger(x), \Psi^\dagger(x')] = 0 \quad (7.23)$$

Proof:

$$\begin{aligned} [\Psi(x), \Psi^\dagger(x')] &= \sum_{ij} \varphi_i(x) \varphi_j^*(x') \underbrace{[a_i, a_j^\dagger]}_{\delta_{ij}} \\ &= \sum_i \varphi_i(x) \varphi_i^*(x') \\ &= \delta(x - x') \end{aligned}$$

- Fermions :

$$\boxed{\{\Psi(x), \Psi^\dagger(x')\} = \delta(x - x')} \quad (7.24)$$

$$\{\Psi(x), \Psi(x')\} = \{\Psi^\dagger(x), \Psi^\dagger(x')\} = 0 \quad (7.25)$$

7.5 Application: interacting electrons

Let us consider the problem of an electron gas in a potential $U(\mathbf{r})$ interacting through the Coulomb interaction $V(\mathbf{r}_1, \mathbf{r}_2) = \frac{e^2}{|\mathbf{r}_1 - \mathbf{r}_2|}$. We will consider that the electrons live in a box of volume Ω . Let us try to determine the form of the Hamiltonian in second quantization.

7.5.1 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 σ 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 :

$$\varphi_{\mathbf{k}}(\mathbf{r}) = \frac{1}{\sqrt{\Omega}} e^{i\mathbf{k} \cdot \mathbf{r}}.$$

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 $|\downarrow\rangle$

$$\begin{aligned} S^z|\uparrow\rangle &= \frac{\hbar}{2}|\uparrow\rangle & S^{\dagger}|\uparrow\rangle &= 0 & S^{-}|\uparrow\rangle &= |\downarrow\rangle \\ S^z|\downarrow\rangle &= -\frac{\hbar}{2}|\downarrow\rangle & S^{\dagger}|\downarrow\rangle &= |\uparrow\rangle & S^{-}|\downarrow\rangle &= 0 \end{aligned}$$

$$S^x = \frac{S^{\dagger} + S^{-}}{2} \quad S^y = \frac{S^{\dagger} - S^{-}}{2i}$$

Description with vectors and matrices

$$\begin{aligned} |\uparrow\rangle &\rightarrow \begin{pmatrix} 1 \\ 0 \end{pmatrix} & |\downarrow\rangle &\rightarrow \begin{pmatrix} 0 \\ 1 \end{pmatrix} \\ S^z &= \frac{\hbar}{2} \underbrace{\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}}_{\sigma_z} & S^x &= \frac{\hbar}{2} \underbrace{\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}}_{\sigma_x} & S^y &= \frac{\hbar}{2} \underbrace{\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}}_{\sigma_y} \end{aligned}$$

$\sigma_x, \sigma_y, \sigma_z$: 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 σ can take two values, say $\frac{1}{2}$ et $-\frac{1}{2}$. One can then define two wave functions by :

$$\begin{cases} \eta_{+\frac{1}{2}}(\frac{1}{2}) = 1 & \eta_{+\frac{1}{2}}(-\frac{1}{2}) = 0 \\ \eta_{-\frac{1}{2}}(\frac{1}{2}) = 0 & \eta_{-\frac{1}{2}}(-\frac{1}{2}) = 1 \end{cases}$$

or, more compactly,

$$\eta_{\sigma_1}(\sigma) = \delta_{\sigma\sigma_1}$$

These functions build an orthonormal basis:

$$\sum_{\sigma} \eta_{\sigma_1}^*(\sigma) \eta_{\sigma_2}(\sigma) = \delta_{\sigma_1\sigma_2}$$

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 eigenfunctions are thus labeled by two indices:

$$\varphi_{\mathbf{k}\sigma}(\mathbf{r}, \sigma') = \frac{1}{\sqrt{\Omega}} e^{i\mathbf{k}\cdot\mathbf{r}} \eta_{\sigma}(\sigma')$$

The creation and annihilation operators are thus given by :

$$c_{\mathbf{k}\sigma}^\dagger \quad \text{and} \quad c_{\mathbf{k}\sigma}$$

Similarly, the field operators are defined by :

$$\begin{aligned} \Psi(\mathbf{r}, \sigma) &= \sum_{\mathbf{k}, \sigma'} \varphi_{\mathbf{k}\sigma'}(\mathbf{r}, \sigma) c_{\mathbf{k}\sigma'} \\ &= \sum_{\mathbf{k}, \sigma'} \frac{1}{\sqrt{\Omega}} e^{i\mathbf{k}\cdot\mathbf{r}} \delta_{\sigma\sigma'} c_{\mathbf{k}\sigma'} \\ &= \sum_{\mathbf{k}} \frac{1}{\sqrt{\Omega}} e^{i\mathbf{k}\cdot\mathbf{r}} c_{\mathbf{k}\sigma} \\ &= \sum_{\mathbf{k}} \varphi_{\mathbf{k}}(\mathbf{r}) c_{\mathbf{k}\sigma} \end{aligned}$$

and

$$\Psi^\dagger(\mathbf{r}, \sigma) = \sum_{\mathbf{k}} \varphi_{\mathbf{k}}^*(\mathbf{r}) c_{\mathbf{k}\sigma}^\dagger$$

7.5.2 Operators

There are two one-body operators :

Kinetic energy

$$\begin{aligned} H_{\text{kin}} &= \sum_i -\frac{\hbar^2 \nabla_i^2}{2m} \equiv \sum_i -\frac{\hbar^2 \Delta_i}{2m} \\ \Delta_i \varphi &= \frac{\partial^2 \varphi}{\partial x_i^2} + \frac{\partial^2 \varphi}{\partial y_i^2} + \frac{\partial^2 \varphi}{\partial z_i^2} \end{aligned}$$

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.

7.5.3 Hamiltonian in second quantization

$$H_{\text{kin}} = \sum_{\mathbf{k}_1 \mathbf{k}_2 \sigma_1 \sigma_2} c_{\mathbf{k}_1 \sigma_1}^\dagger c_{\mathbf{k}_2 \sigma_2} \int d\mathbf{r} \sum_{\sigma} \varphi_{\mathbf{k}_1 \sigma_1}^*(\mathbf{r}, \sigma) \left(-\frac{\hbar^2 \Delta}{2m} \right) \varphi_{\mathbf{k}_2 \sigma_2}(\mathbf{r}, \sigma)$$

$$\text{OR} \quad \begin{aligned} \varphi_{\mathbf{k}_1 \sigma_1}^*(\mathbf{r}, \sigma) &= \frac{1}{\sqrt{\Omega}} e^{-i\mathbf{k}_1 \cdot \mathbf{r}} \delta_{\sigma \sigma_1} \\ \varphi_{\mathbf{k}_2 \sigma_2}(\mathbf{r}, \sigma) &= \frac{1}{\sqrt{\Omega}} e^{i\mathbf{k}_2 \cdot \mathbf{r}} \delta_{\sigma \sigma_2} \end{aligned}$$

$$\begin{aligned} \Rightarrow H_{\text{kin}} &= \sum_{\mathbf{k}_1 \mathbf{k}_2 \sigma} c_{\mathbf{k}_1 \sigma}^\dagger c_{\mathbf{k}_2 \sigma} \int \frac{d\mathbf{r}}{\Omega} e^{-i\mathbf{k}_1 \cdot \mathbf{r}} \frac{\hbar^2 \mathbf{k}_2^2}{2m} e^{i\mathbf{k}_2 \cdot \mathbf{r}} \\ &= \sum_{\mathbf{k}_1 \mathbf{k}_2 \sigma} c_{\mathbf{k}_1 \sigma}^\dagger c_{\mathbf{k}_2 \sigma} \frac{\hbar^2 \mathbf{k}_2^2}{2m} \delta_{\mathbf{k}_1 \mathbf{k}_2} \end{aligned}$$

$$\boxed{H_{\text{kin}} = \sum_{\mathbf{k} \sigma} \frac{\hbar^2 \mathbf{k}^2}{2m} c_{\mathbf{k} \sigma}^\dagger c_{\mathbf{k} \sigma}}$$

$$\begin{aligned} H_{\text{pot}} &= \sum_{\mathbf{k}_1 \mathbf{k}_2 \sigma_1 \sigma_2} c_{\mathbf{k}_1 \sigma_1}^\dagger c_{\mathbf{k}_2 \sigma_2} \int \frac{d\mathbf{r}}{\Omega} \sum_{\sigma} e^{-i\mathbf{k}_1 \cdot \mathbf{r}} U(\mathbf{r}) e^{i\mathbf{k}_2 \cdot \mathbf{r}} \delta_{\sigma \sigma_1} \delta_{\sigma \sigma_2} \\ &= \sum_{\mathbf{k}_1 \mathbf{k}_2 \sigma} c_{\mathbf{k}_1 \sigma}^\dagger c_{\mathbf{k}_2 \sigma} \int \frac{d\mathbf{r}}{\Omega} e^{-i\mathbf{k}_1 \cdot \mathbf{r}} U(\mathbf{r}) e^{i\mathbf{k}_2 \cdot \mathbf{r}} \end{aligned}$$

Let us define the Fourier transform of the potential by :

$$\tilde{U}(\mathbf{k}) = \int \frac{d\mathbf{r}}{\Omega} e^{-i\mathbf{k} \cdot \mathbf{r}} U(\mathbf{r})$$

This leads to :

$$H_{\text{pot}} = \sum_{\mathbf{k}_1 \mathbf{k}_2 \sigma} c_{\mathbf{k}_1 \sigma}^\dagger c_{\mathbf{k}_2 \sigma} \tilde{U}(\mathbf{k}_1 - \mathbf{k}_2)$$

or :

$$\boxed{H_{\text{pot}} = \sum_{\mathbf{k} \mathbf{q} \sigma} \tilde{U}(\mathbf{q}) c_{\mathbf{k} + \mathbf{q} \sigma}^\dagger c_{\mathbf{k} \sigma}}$$

The operator :

$$\rho_{-\mathbf{q}} = \sum_{\mathbf{k} \sigma} c_{\mathbf{k} + \mathbf{q} \sigma}^\dagger c_{\mathbf{k} \sigma}$$

is the Fourier transform of the density operator. Indeed,

$$\begin{cases} \rho_{\mathbf{q}} &= \int d\mathbf{r} \rho(\mathbf{r}) e^{-i\mathbf{q} \cdot \mathbf{r}} \\ \rho(\mathbf{r}) &= \sum_{\sigma} \Psi^\dagger(\mathbf{r}, \sigma) \Psi(\mathbf{r}, \sigma) \end{cases}$$

$$\begin{aligned} \Rightarrow \rho_{\mathbf{q}} &= \int \frac{d\mathbf{r}}{\Omega} e^{-i\mathbf{q} \cdot \mathbf{r}} \sum_{\mathbf{k}_1 \mathbf{k}_2 \sigma} e^{-i\mathbf{k}_1 \cdot \mathbf{r}} e^{i\mathbf{k}_2 \cdot \mathbf{r}} c_{\mathbf{k}_1 \sigma}^\dagger c_{\mathbf{k}_2 \sigma} \\ &= \sum_{\mathbf{k} \sigma} c_{\mathbf{k} \sigma}^\dagger c_{\mathbf{k} + \mathbf{q} \sigma} \end{aligned}$$

One can thus write :

$$H_{\text{pot}} = \sum_{\mathbf{q}} \tilde{U}(\mathbf{q}) \rho_{-\mathbf{q}}$$

$$H_{\text{Coul}} = \frac{1}{2} \sum_{\substack{\mathbf{k}_1 \mathbf{k}_2 \mathbf{k}_3 \mathbf{k}_4 \\ \sigma_1 \sigma_2 \sigma_3 \sigma_4}} c_{\mathbf{k}_1 \sigma_1}^\dagger c_{\mathbf{k}_2 \sigma_2}^\dagger c_{\mathbf{k}_3 \sigma_3} c_{\mathbf{k}_4 \sigma_4} \\ \times \int d\mathbf{r} d\mathbf{r}' \sum_{\sigma \sigma'} \varphi_{\mathbf{k}_1 \sigma_1}^*(\mathbf{r}, \sigma) \varphi_{\mathbf{k}_2 \sigma_2}^*(\mathbf{r}', \sigma') V(\mathbf{r} - \mathbf{r}') \varphi_{\mathbf{k}_3 \sigma_3}(\mathbf{r}', \sigma') \varphi_{\mathbf{k}_4 \sigma_4}(\mathbf{r}, \sigma)$$

The integral is given by :

$$\int \frac{d\mathbf{r} d\mathbf{r}'}{\Omega^2} V(\mathbf{r} - \mathbf{r}') e^{-i\mathbf{k}_1 \cdot \mathbf{r}} e^{-i\mathbf{k}_2 \cdot \mathbf{r}'} e^{i\mathbf{k}_3 \cdot \mathbf{r}'} e^{i\mathbf{k}_4 \cdot \mathbf{r}} \underbrace{\sum_{\sigma \sigma'} \delta_{\sigma_1 \sigma} \delta_{\sigma_2 \sigma'} \delta_{\sigma_3 \sigma'} \delta_{\sigma_4 \sigma}}_{\delta_{\sigma_1 \sigma_4} \delta_{\sigma_2 \sigma_3}}$$

or, defining $\mathbf{R} = \mathbf{r} - \mathbf{r}'$:

$$\int \frac{d\mathbf{R} d\mathbf{r}'}{\Omega^2} V(\mathbf{R}) e^{-i\mathbf{k}_1 \cdot (\mathbf{R} + \mathbf{r}')} e^{-i\mathbf{k}_2 \cdot \mathbf{r}'} e^{i\mathbf{k}_3 \cdot \mathbf{r}'} e^{i\mathbf{k}_4 \cdot (\mathbf{R} + \mathbf{r}')}$$

$$= \underbrace{\int \frac{d\mathbf{R}}{\Omega} V(\mathbf{R}) e^{-i(\mathbf{k}_1 - \mathbf{k}_4) \cdot \mathbf{R}}}_{\tilde{V}(\mathbf{k}_1 - \mathbf{k}_4)} \underbrace{\int \frac{d\mathbf{r}'}{\Omega} e^{-i(\mathbf{k}_1 + \mathbf{k}_2 - \mathbf{k}_3 - \mathbf{k}_4) \cdot \mathbf{r}'}}_{\delta_{\mathbf{k}_1 + \mathbf{k}_2, \mathbf{k}_3 + \mathbf{k}_4}}$$

$$H_{\text{Coul}} = \frac{1}{2} \sum_{\mathbf{k} \mathbf{k}' \mathbf{q} \sigma \sigma'} \tilde{V}(\mathbf{q}) c_{\mathbf{k} + \mathbf{q} \sigma}^\dagger c_{\mathbf{k}' - \mathbf{q} \sigma'}^\dagger c_{\mathbf{k}' \sigma'} c_{\mathbf{k} \sigma}$$

With the help of the density operators, this can also be written :

$$H_{\text{Coul}} = \frac{1}{2} \sum_{\mathbf{q}} \tilde{V}(\mathbf{q}) (\rho_{\mathbf{q}} \rho_{-\mathbf{q}} - N)$$

Nota Bene : This expression is valid with the definition $\tilde{V}(\mathbf{q}) = \int \frac{d\mathbf{r}}{\Omega} e^{-i\mathbf{q} \cdot \mathbf{r}} V(\mathbf{r})$.

If one defines :

$$V(\mathbf{q}) = \int d\mathbf{r} e^{-i\mathbf{q} \cdot \mathbf{r}} V(\mathbf{r})$$

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 :

$$H_1 = \frac{1}{2\Omega} \sum_{\mathbf{k} \mathbf{k}' \mathbf{q} \sigma \sigma'} V(\mathbf{q}) c_{\mathbf{k} + \mathbf{q} \sigma}^\dagger c_{\mathbf{k}' - \mathbf{q} \sigma'}^\dagger c_{\mathbf{k}' \sigma'} c_{\mathbf{k} \sigma}$$

This the form that is generally used. The inverse Fourier transform is then given by :

$$V(\mathbf{r}) = \frac{1}{\Omega} \sum_{\mathbf{q}} V(\mathbf{q}) e^{i\mathbf{q} \cdot \mathbf{r}}$$

or, for Coulomb potential :

$$\frac{e^2}{r} = \frac{1}{\Omega} \sum_{\mathbf{q}} \frac{4\pi e^2}{\mathbf{q}^2} e^{i\mathbf{q} \cdot \mathbf{r}}.$$