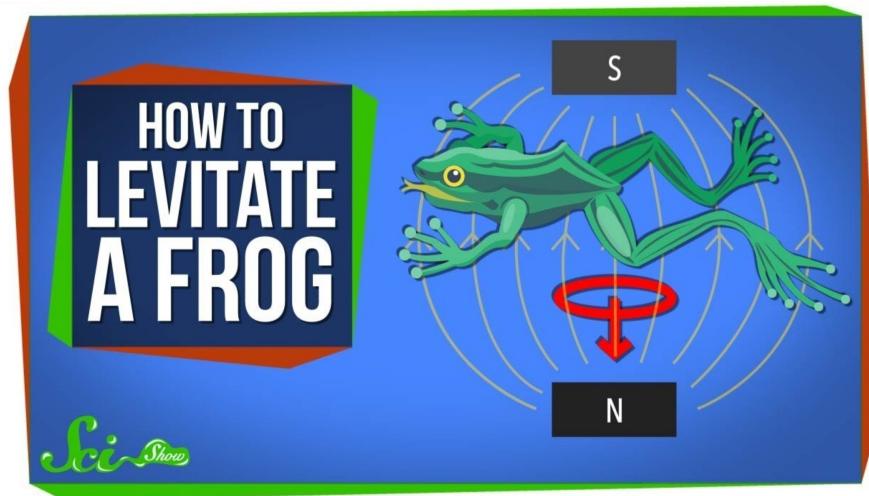


# Magnetism and Superconductivity

Lecture notes of Giorgio Calandra from *Prof. Ghiringhelli's course*



# Contents

<b>1</b>	<b>Introduction</b>	<b>4</b>
1.1	Magnetic Momenta and Larmor Precession . . . . .	4
1.1.1	Quantum Magnetic Momenta . . . . .	5
1.2	LS Coupling and Hund's Rules . . . . .	6
1.2.1	JJ Coupling . . . . .	6
1.2.2	Atomic Magnetic Momentum and Landé Factor . . . . .	6
1.3	Nuclear Magnetism . . . . .	7
1.4	Rare Earth magnetism . . . . .	7
1.5	3d Transition Metals . . . . .	8
1.6	Diamagnetism and Paramagnetism . . . . .	9
1.6.1	Semiclassical Description . . . . .	10
1.6.2	Diamagnetism . . . . .	10
1.6.3	Paramagnetism . . . . .	11
1.6.4	Qunatum Paramagnetism . . . . .	13
1.6.5	Magnetization, Energy, Specific Heat and Entropy . . . . .	13
<b>2</b>	<b>Magnetic Techniques</b>	<b>15</b>
2.1	Nuclear Magnetic Resonance . . . . .	15
2.1.1	Working Principle . . . . .	15
2.1.2	NMR Intensities and Timescales . . . . .	16
2.2	Mössbauer Spectroscopy . . . . .	18
2.3	Muon Spin Rotation . . . . .	19
<b>3</b>	<b>Magnetic Interaction</b>	<b>21</b>
3.1	Exchange Interaction . . . . .	21
3.2	Heisenberg Model . . . . .	23
3.3	Weiss Model for Ferromagnetism . . . . .	23
3.3.1	Exchange Interaction and Molecular Field . . . . .	26
3.4	Weiss model for Antiferromangetism . . . . .	26
3.4.1	AFM, Field Direction and Spin Flop . . . . .	28
3.4.2	Helical Order . . . . .	30
<b>4</b>	<b>Order and Broken Symmetry</b>	<b>32</b>
4.1	Broken Symmetry . . . . .	32
4.2	Landau Model of Phase Transitions . . . . .	34
4.3	Heisenberg and Ising Models . . . . .	35
4.3.1	The 1D Ising Model . . . . .	36
4.4	Phase Transition and Nematic Order . . . . .	37
4.5	Nematic Order . . . . .	37
<b>5</b>	<b>Real Materials</b>	<b>39</b>
5.1	Crystalline Structure . . . . .	39
5.2	Examples . . . . .	40
5.2.1	Perovskite . . . . .	40
5.2.2	Spinel Structure . . . . .	42
5.2.3	Cuprates . . . . .	42
5.2.4	Kagome lattice . . . . .	43

<b>6 Potential and Crystal Field</b>	<b>44</b>
6.1 Centrifugal Potential . . . . .	44
6.2 Electronic Configuration and Orbitals . . . . .	45
6.3 Orbitals and Crystal Field . . . . .	46
6.3.1 Sugano-Tanabe Diagrams . . . . .	48
6.4 Jahn-Teller Effect . . . . .	49
<b>7 Interactions</b>	<b>52</b>
7.1 Direct Exchange . . . . .	52
7.2 Indirect Exchange . . . . .	53
7.2.1 Super-Exchange . . . . .	53
7.2.2 Double Exchange . . . . .	54
7.2.3 Anisotropic Exchange . . . . .	55
7.3 Magnons . . . . .	55
7.3.1 FM Spin Waves . . . . .	55
7.3.2 AFM Spin Waves . . . . .	58
<b>8 Neutron Scattering Techniques</b>	<b>60</b>
8.1 AFM lattices . . . . .	60
8.2 Bragg Diffraction and Elastic Scattering . . . . .	61
8.2.1 Scattering Form Factor . . . . .	62
8.3 Neutron Scattering . . . . .	63
8.3.1 Neutron Sources . . . . .	64
8.4 INS and RIXS . . . . .	65
8.4.1 Inelastic Neutron Scattering . . . . .	65
8.4.2 Resonant Inelastic X-ray Scattering . . . . .	66
<b>9 Magnetism in Metals</b>	<b>68</b>
9.1 Quasi-Free Electron Model . . . . .	68
9.2 Pauli Paramagnetism . . . . .	70
9.3 Stoner Ferromagnetism . . . . .	72
9.3.1 Stoner Enhancement . . . . .	73
9.3.2 Stoner Excitations . . . . .	74
9.4 Landau Diamagnetism . . . . .	75
<b>10 Bose-Einstein Condensates</b>	<b>77</b>
10.1 Historical Introduction . . . . .	77
10.2 Bose-Einstein Statistic . . . . .	77
10.3 Bose-Einstein Condensation . . . . .	79
10.3.1 Below $T_C$ . . . . .	81
10.3.2 Specific Heat . . . . .	82
10.4 Experimental Realization of a BEC . . . . .	82
10.4.1 Rubidium in BEC . . . . .	83
10.4.2 Laser Cooling . . . . .	85
<b>11 Superfluidity</b>	<b>86</b>
11.1 Quantum Fluids . . . . .	86
11.2 The wave function . . . . .	88
11.3 Superflow . . . . .	89
11.3.1 The experiment . . . . .	89
11.3.2 Fountain Effect . . . . .	90
11.4 Flow Quantization and Vortices . . . . .	92
<b>12 Superconductivity</b>	<b>95</b>
12.1 History and Phenomenology . . . . .	95
12.1.1 The first observations . . . . .	96
12.1.2 The Meissner-Ochsenfeld Effect . . . . .	97
12.1.3 Type I and Type II Superconductors . . . . .	99
12.1.4 Isotope Effect and Superconducting Gap . . . . .	100

12.1.5	Superconductivity: a Nobel Story . . . . .	102
12.2	London Equations . . . . .	102
12.2.1	Stationary Case - DC Current . . . . .	102
12.2.2	Time Dependent Case - AC Current . . . . .	105
12.3	The Ginzburg-Landau Model . . . . .	106
12.3.1	The Condensation Energy . . . . .	106
12.3.2	The Landau Model of Phase Transitions . . . . .	108
12.4	The BCS Theory . . . . .	111
12.4.1	Two Particles Wave Function . . . . .	113
12.5	Superconductor Junctions . . . . .	115
12.5.1	Josephson Effect . . . . .	117

# Chapter 1

## Introduction

In this chapter, we are going to introduce a series of introductory concepts to study the behaviour of a magnet, represented by its magnetic dipole  $\mu$ , in a magnetic field  $B$ . A lot of concepts will be given for grants although, sometimes, they are not that clear even to me.

### 1.1 Magnetic Momenta and Larmor Precession

Let's consider a magnetic dipole (*e.g.* an atom) surrounded by a magnetic field. A torque  $G$  will act on the dipole trying to align it to the magnetic field. The torque is given by:

$$G = \mu \times B \quad (1.1)$$

while the potential energy of such a system is given by:

$$U_m = -\mu \cdot B \quad (1.2)$$

Since the total angular momentum of the system must be conserved the dipole starts precessing. If we consider an atom with one electron rotating around the nucleus with a radius  $r$  as in Fig.1.1 we have that the magnetic dipole is given by:

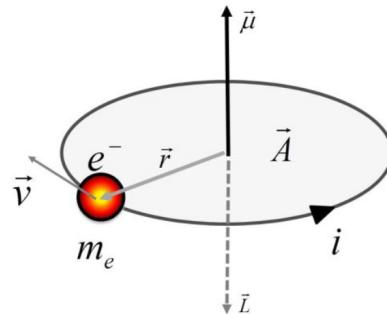
$$\mu = iA = i(\pi r^2) = -\frac{ev}{2\pi r}\pi r^2 = -\frac{em_e vr}{2m_e} = -\frac{eL}{2m_e} = \gamma L \quad (1.3)$$

where  $L$  is the (modulus) of the angular momentum of the electron and  $\gamma$  is the gyroscopic ratio. Since  $L$  must be quantised, by substituting  $L = n\hbar$  we get:

$$\mu = -n\frac{e\hbar}{2m_e} = -n\mu_B \quad (1.4)$$

where  $\mu_B = \frac{e\hbar}{2m_e}$  is the Bohr magneton that allows to express the gyroscopic ratio as  $\gamma = \frac{\mu_B}{\hbar}$ . In general, the cardinal equation of dynamic stays:

$$\frac{dL}{dt} = G = \Omega \times L \quad (1.5)$$



**Figure 1.1:** Atomic precession scheme.

in which  $\Omega$  is the angular velocity around the precession axis. In the case of a magnetic dipole Eq.(1.5) becomes:

$$\frac{d\mu}{dt} = \gamma \frac{dL}{dt} = \gamma G = \gamma \Omega \times L \quad (1.6)$$

This equation can be written in cartesian coordinates to get all the components separately. By considering a magnetic field  $B$  direct along the z-axis we get:

$$\begin{cases} \frac{d\mu_x}{dt} = \gamma B \mu \sin \theta \cos(\Omega t) \\ \frac{d\mu_y}{dt} = -\gamma B \mu \sin \theta \sin(\Omega t) \\ \frac{d\mu_z}{dt} = 0 \end{cases} \quad (1.7)$$

where  $\Omega = -\gamma B$  is the so-called Larmor frequency. Such a system can be solved and gives:

$$\begin{cases} \mu_x(t) = \mu \sin \theta \sin(\Omega t) \\ \mu_y(t) = \mu \sin \theta \cos(\Omega t) \\ \mu_z(t) = \mu \cos \theta \end{cases} \quad (1.8)$$

which is the set of equations which describes the precession motion of a dipole in a magnetic field. Moreover, the Larmor frequency can also be used to compute the magnetic field knowing that:

$$\Omega(B) = -\frac{\mu_B}{\hbar} B = \frac{e}{2m_e} B = 87.9 \times 10^9 B \text{ rad s}^{-1} \quad (1.9)$$

### 1.1.1 Quantum Magnetic Momenta

Since magnetic momentums are quantized, we need to keep in mind the discrete values of the orbital and spin momentums:

- Orbital momentum:

$$\begin{aligned} \hat{L}^2 &= l(l+1)\hbar^2 \\ L_z &= m_l \hbar \\ \mu_L^2 &= g_L^2 l(l+1) \mu_B^2 \quad g_L = 1 \\ \mu_z &= -m_l \mu_B \end{aligned} \quad (1.10)$$

- Spin momentum:

$$\begin{aligned} \hat{S}^2 &= s(s+1)\hbar^2 = \frac{3}{4}\hbar^2 \\ S_z &= m_s \hbar = \pm \frac{1}{2}\hbar \\ \mu_S^2 &= g_S^2 s(s+1) \mu_B^2 = \frac{3}{4}g_S^2 \mu_B^2 \quad g_S = 1 \\ \mu_{S_z} &= -m_s \mu_B = \mp \mu_B \end{aligned} \quad (1.11)$$

where  $g_L$  and  $g_S$  are the gyromagnetic ratios for the orbital and spin momenta respectively. Moreover, is worth remembering that the presence of a magnetic field, not only starts a precession motion of the atom around the axis of the field itself, but splits the energy levels of the system of a quantity

$$\Delta E = g_S \mu_B B \quad (1.12)$$

which is the well-known Zeeman splitting effect. In other words, two discrete levels arise with energy:

$$E = \pm \frac{1}{2} g_S \mu_B B$$

## 1.2 LS Coupling and Hund's Rules

It should be well known that the interaction between the angular momentum  $\mathbf{L}$  and the spin one  $\mathbf{S}$  is responsible for the so-called Spin-Orbit (S-O) interaction. In a very qualitative way, this can be explained by saying that an electron in an atom, in its reference frame, sees the nucleus rotate around it and therefore feels the presence of the magnetic field generated by the nucleus. The interaction between the magnetic field of the nucleus and the spin of the electron causes a variation in the energy. This energy splitting, also called spin-orbit splitting, is given by:

$$\Delta E_{SO} = \frac{Ze^2}{8\pi\varepsilon_0 m^2 c^2 r^3} \mathbf{S} \cdot \mathbf{L} \quad (1.13)$$

For our purposes is sufficient to remember that  $\Delta E_{SO} \propto \frac{Z}{r^3} \mathbf{S} \cdot \mathbf{L}$ . This expression makes evident that for atoms with medium-low values of the atomic number  $Z$  the S-O interaction is quite weak and the (residual) Coulombian interaction is dominant. Therefore, for small  $Z$  atoms, in first approximation the S-O interaction is neglected and the spin and orbital momentums of each electron in the atom are considered decoupled. In this case, the Coulombian interaction is responsible for the coupling between each spin and orbital momenta separately to give the total spin angular momentum  $\mathbf{S}$  and the total orbital angular momentum  $\mathbf{L}$ . Only after that, the S-O interaction can be considered by summing these two quantities to obtain the total angular momentum of the atom  $\mathbf{J} = \mathbf{L} + \mathbf{S}$  which is a conserved quantity. If the S-O interaction is small it can be considered a perturbation, then  $L$  and  $S$  are good quantum numbers and we can find the ground state using Hund's Rules:

- The minimum energy state is the one where the total spin is maximum (that is compatible with Pauli's exclusion principle)
- For a given spin quantum number, the minimum energy state is the one where the total angular momentum is maximum.
- For fixed spin and orbital momentum, the minimum energy state is either the one with  $J = |L - S|$  ( $J$  minimum) for less than half-filled shell or the one with  $J = |L + S|$  ( $J$  maximum) for more than half filled shell.

The ground state electronic configuration coming from Hund's rule or any other excited state configuration can be written with the Russell-Saunders notation:

$$^{2S+1}L_J$$

In general, since we are interested in the valance electrons for which the expectation value of the radius  $r$  is big, we will often neglect the S-O interaction ( $\Delta E_{SO} \propto \frac{1}{r^3}$ ).

### 1.2.1 JJ Coupling

If the S-O coupling is strong with respect to the Coulombian interaction, we have to use a different criterion to describe the angular momentum called JJ Coupling. What matters now is mostly  $J$ , while  $L$  and  $S$  are not good quantum numbers anymore. In this case, one has to calculate the total angular momentum for each electron ( $J_i = L_i + S_i$ ) and only then sum all of them to obtain the total angular momentum ( $J = \sum_{i=1}^N J_i$ ).

### 1.2.2 Atomic Magnetic Momentum and Landé Factor

An electron in an atom has a total magnetic momentum given by the contribution of the orbital momentum  $\mathbf{L}$  and the spin one  $\mathbf{S}$ , which gives:

$$\mu_{tot} = \mu_S + \mu_L = -\frac{\mu_B}{\hbar} (g_L \mathbf{L} + g_S \mathbf{S}) \quad (1.14)$$

Remembering that the eigenvalues for the two momenta separately are:

$$\begin{aligned} \mu_S &= g_S \mu_B \sqrt{S(S+1)} \\ \mu_L &= g_L \mu_B \sqrt{L(L+1)} \end{aligned} \quad (1.15)$$

one can find that the eigenvalue for the total magnetic momentum is:

$$\mu_{tot} = \mu_B g_J \sqrt{J(J+1)} \quad (1.16)$$

in which  $g_J$  is the so-called Landé factor defined by:

$$g_J = 1 + \frac{J(J+1) - L(L+1) + S(S+1)}{2J(J+1)} \quad (1.17)$$

Moreover, for an experimental comparison, is used to define the number of effective Bohr magnetons  $p_J$  as:

$$p_J = g_J \sqrt{J(J+1)} \quad (1.18)$$

So that we can write the total magnetic momentum as  $\mu_{tot} = \mu_B p_J$ .

### 1.3 Nuclear Magnetism

Nuclei, as well as electrons, have an angular momentum associated with both protons and neutrons (since both of them are fermions) and therefore they possess magnetic properties. Also for nucleons an analogous of the Bohr magneton can be defined and it's the nuclear magneton  $\mu_N$ :

$$\mu_N = \frac{e\hbar}{2m_p} \approx \frac{e\hbar}{2m_n} = 5.0508 \times 10^{-27} \text{ Am}^2 \quad (1.19)$$

where  $m_p$  is the mass of the proton that is almost equal to the neutron mass  $m_n$ . Since this expression differs from the one of the Bohr's magneton only for the considered mass, that for Bohr is the one of the electron, in general, the nuclear magneton is almost 2000 times smaller than the Bohr one. The nuclear magnetic momentum is:

$$\mu = g_I \mu_N I \quad (1.20)$$

Here  $I$  is the nuclear spin quantum number that can take values multiple of  $1/2$  and  $g_I$  is the gyromagnetic ratio that depends on the considered nucleon. Moreover, nuclei are influenced by magnetic fields and, as well as in the electron case, their energy levels are subject to Zeeman splitting:

$$\Delta E = g_I \mu_N B \quad (1.21)$$

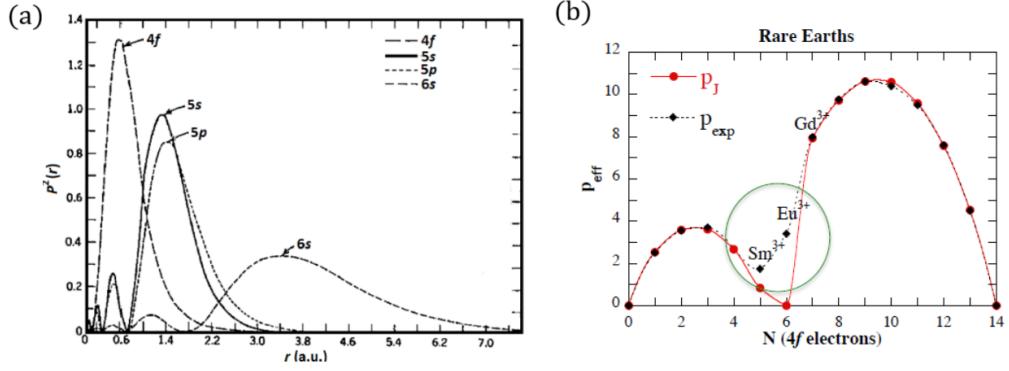
Once again, because of the higher mass of nucleons, the splitting for a given field is much lower than the one for electrons. Tab.1.1 reports some interesting values for a neutron, a proton, and other noticeable atoms. The last column indicates the frequency associated with the Zeeman splitting for a magnetic field of  $1T$ .

**Table 1.1:** Nuclear Magnetic Properties

Nucleus	$Z$	$N$	$I$	$\frac{\mu}{\mu_N}$	$g_I$	$\nu$ (MHz, for $B = 1$ T)
$n$	0	1	$1/2$	-1.913	-3.826	29.17
$p$	1	0	$1/2$	2.793	5.586	42.58
${}^2\text{H}$ (D)	1	1	1	0.857	0.8789	6.53
${}^{12}\text{C}$	6	6	0	0	0	0
${}^{13}\text{C}$	6	7	$1/2$	0.702	1.404	10.71

### 1.4 Rare Earth magnetism

The Rare Earth (or Lanthanides) are made by the elements of the periodic table that go from lanthanum ( $La$ ,  $Z = 57$ ) to lutetium ( $Lu$ ,  $Z = 71$ ) that correspond to the progressive filling of the  $4f$  subshell. Their atomic configuration in the fundamental state is of the type  $[Xe]4f^n5d^{0-1}6s^2$ . Fig.1.2(a) depicts the behaviour of the atomic radial wavefunctions related to the external orbitals of an element of the rare earth. For the ions of these kinds of elements, the presence of completely filled  $5s$  and  $5p$  orbitals creates an electrostatic shield (screening effect) that makes the



**Figure 1.2:** (a) Rare Earth radial wavefunctions. (b) Rare Earth experimental and theoretical behaviour.

outer electrons unsensible to the **crystalline field**. The crystalline field, or crystal field, is the electrostatic field generated by the positive ions in a lattice which influences the behaviour of the electrons in a neighbouring atom. For this reason, the ions will act as if they were isolated and they behave as expected by the theory as it is possible to see from Tab.1.2 in which the theoretical and experimental results are compared. As we can see from Fig.1.2(b) the theory for these elements is coherent with the experimental results in all cases except for the ions  $Sm^{3+}$  and  $Eu^{3+}$ . This is because these two elements behave a bit differently since, even at room temperature, thanks to thermal excitations their first excited levels are occupied. This makes the paramagnetic theory developed so far incorrect since it has been developed by considering all the atoms in their ground states.

**Table 1.2:** Electronic configurations and spectroscopic terms of rare earth ions.

Ion	Shell	$S$	$L$	$J$	Term	$p_J$	$p_{exp}$
$La^{3+}$	$4f^0$	0	0	0	$^1S_0$	0	0
$Ce^{3+}$	$4f^1$	$1/2$	3	$5/2$	$^2F_{5/2}$	2.54	2.51
$Pr^{3+}$	$4f^2$	1	5	4	$^3H_4$	3.58	3.56
$Nd^{3+}$	$4f^3$	$3/2$	6	$9/2$	$^4I_{9/2}$	3.62	3.30-3.70
$Pm^{3+}$	$4f^4$	2	6	4	$^5I_4$	2.68	-
$Sm^{3+}$	$4f^5$	$5/2$	5	$5/2$	$^6H_{5/2}$	0.85	1.74
$Eu^{3+}$	$4f^6$	3	3	0	$^7F_0$	0	3.4
$Gd^{3+}$	$4f^7$	$7/2$	0	$7/2$	$^8S_{7/2}$	7.94	7.98
$Tb^{3+}$	$4f^8$	3	3	6	$^7F_6$	9.72	9.77
$Dy^{3+}$	$4f^9$	$5/2$	5	$15/2$	$^6H_{15/2}$	10.63	10.63
$Ho^{3+}$	$4f^{10}$	2	6	8	$^5I_8$	10.6	10.40
$Er^{3+}$	$4f^{11}$	$3/2$	6	$15/2$	$^4I_{15/2}$	9.59	9.50
$Tm^{3+}$	$4f^{12}$	1	5	6	$^3H_6$	7.57	7.61
$Yb^{3+}$	$4f^{13}$	$1/2$	3	$7/2$	$^2F_{7/2}$	4.53	4.50
$Lu^{3+}$	$4f^{14}$	0	0	0	$^1S_0$	0	0

## 1.5 3d Transition Metals

The transition metals of the 3d group (3dTMs) are the elements of the periodic table that go from Scandium ( $Sc, Z = 21$ ) to the Zinc ( $Zn, Z = 30$ ) that correspond to the progressive filling of the 3d subshell. Their ground state configuration is  $[Ar]3d^n4s^{1-2}$ . For these elements, the distribution of the radial wavefunctions is different from the previous case and is such that the 3d subshell is the most external one. Therefore, differently from before, the screening of the other orbitals is not sufficient to prevent the effects of the crystal field and it becomes prevalent with respect to the S-O coupling. We can expect, then, that the behaviour of the paramagnetic ions in the crystal differs from the one of the free ions, predicted by the theory. In Tab.1.3 we can see the comparison between theoretical and experimental results which, as predicted, differ. In particular, we can

**Table 1.3:** Electronic Configurations and Spectroscopic Terms of  $3d$  Ions

Ion	Shell	$S$	$L$	$J$	Term	$p_J$	$p_{exp}$	$p_{L=0}$
$Ti^{3+}, V^{4+}$	$3d^1$	$1/2$	$2$	$3/2$	$^2D_{3/2}$	1.55	1.70	1.73
$V^{3+}$	$3d^2$	$1$	$3$	$2$	$^3F_2$	1.63	2.61	2.83
$Cr^{3+}, V^{2+}$	$3d^3$	$3/2$	$3$	$3/2$	$^4F_{3/2}$	0.77	3.85	3.81
$Mn^{3+}, Cr^{2+}$	$3d^4$	$2$	$2$	$0$	$^5D_0$	0	4.82	4.90
$Fe^{3+}, Mn^{2+}$	$3d^5$	$5/2$	$0$	$5/2$	$^6S_{5/2}$	5.92	5.82	5.92
$Fe^{2+}$	$3d^6$	$2$	$2$	$4$	$^5D_4$	6.7	5.36	4.90
$Co^{2+}$	$3d^7$	$3/2$	$3$	$9/2$	$^4F_{9/2}$	6.63	4.90	3.87
$Ni^{2+}$	$3d^8$	$1$	$3$	$4$	$^3F_4$	5.59	3.12	2.83
$Cu^{2+}$	$3d^9$	$1/2$	$2$	$5/2$	$^2D_{5/2}$	3.55	1.83	1.73
$Zn^{2+}$	$3d^{10}$	$0$	$0$	$0$	$^1S_0$	0	0	0

notice that the only elements for which the theoretical prediction fits the experimental result are the ones with a zero orbital angular momentum, *i.e.*  $L = 0$ . Everything goes as if, when the ion is in the crystalline field, there is no orbital angular momentum anymore. This phenomenon is called **quenching** of the orbital angular momentum. We convince ourselves about the quenching of the angular momentum by looking at the last column of Tab.1.3. It shows the theoretical results obtained by putting  $L = 0$  which are in perfect agreement, in the limit of the experimental error, with the experimental results.

## 1.6 Diamagnetism and Paramagnetism

Magnetic materials can be classified based on their behaviour when exposed to a magnetic field  $\mathbf{B}_0$ . The vector  $\mathbf{H}$  defined as:

$$\mathbf{B}_0 = \mu_0 \mathbf{H} \quad (1.22)$$

is used to describe the magnetic field inside a given material. In fact, being  $\mathbf{B}$  the total magnetic field inside a given material (considered homogeneous) we have that:

$$\mathbf{B} = \mathbf{B}_0 + \mu_0 \mathbf{M} = \mu_0 (\mathbf{H} + \mathbf{M}) \quad (1.23)$$

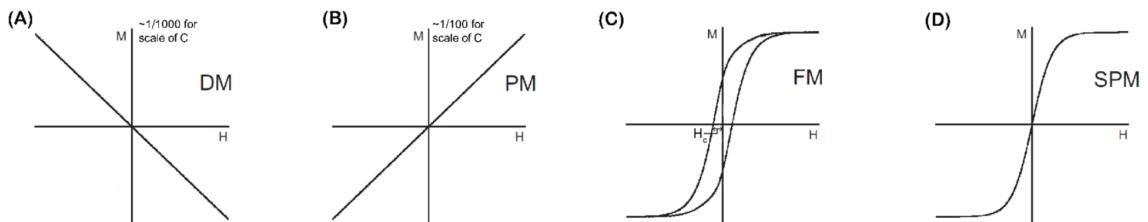
where  $\mathbf{M}$  is the magnetization of the material. It is useful to define the magnetic susceptibility  $\chi$ :

$$\mathbf{M} = \chi \mathbf{H} \quad (1.24)$$

Therefore we have:

$$\mathbf{B} = \mu_0 (1 + \chi) \mathbf{H} \quad (1.25)$$

This relation can also be written as  $\mathbf{B} = \mu \mathbf{H}$  where, since in principle  $\mathbf{B}$  and  $\mathbf{H}$  are not parallel in any material,  $\mu$  is a tensor. Materials can be defined depending on the sign of  $\chi$ : if  $\chi < 0$  the material is diamagnetic while if  $\chi > 0$  the material is paramagnetic. In particular, a material which has  $\chi = -1$ , *i.e.*  $\mathbf{M} = -\mathbf{H}$ , is said to be a perfect diamagnet. The A and B graphs in Fig.1.3 show the relations between  $\mathbf{M}$  and  $\mathbf{H}$  in the case of a diamagnetic (DM) and a paramagnetic (PM) material. From the graphs, it is possible to understand that both in the DM and in the PM cases, since the relation is linear, the susceptibility is constant. If the relation between  $\mathbf{H}$  and  $\mathbf{M}$  is not linear we are in the presence of a ferromagnetic (FM) material (graph C in Fig.1.3). In this case,



**Figure 1.3:** Magnetization in function of the magnetic field for different types of materials

for each value of  $\mathbf{H}$  two possible values of  $\mathbf{M}$  exist which depend on the previous history of the magnetization of the material. Furthermore, for FM materials the scale of the magnetization  $\mathbf{M}$  is much larger than the one of DM and PM materials. Some other classes of material exist such as the Super Paramagnetic (SPM) materials, defined by the relation in graph D of Fig.1.3 and the Anti-Ferromagnetic (AFM) materials which are macroscopically "dead" on a magnetic point of view, *i.e.* they have a zero magnetization, but they are microscopically made by strong magnets, *i.e.* the atoms which compose the AFM materials posses a strong dipole moment.

### 1.6.1 Semiclassical Description

To try to explain diamagnetism and paramagnetism a fully quantum description is necessary. Nevertheless, as usual in physics, let's start with a semiclassical description. Let's consider an atom described by an orbital angular momentum  $\mathbf{L}$  and a spin angular momentum  $\mathbf{S}$ . The interaction of the atom with a magnetic field  $\mathbf{B}$  can be written in a compact Hamiltonian by exploiting the vector potential  $\mathbf{A}$  of the magnetic field and the canonical momentum  $\mathbf{p} = m\mathbf{v} + q\mathbf{A}$ :

$$\mathcal{H} = \mathcal{H}_0 + \mu_B(\mathbf{L} + g_S\mathbf{S}) \cdot \mathbf{B} + \frac{e^2}{8m_e} \sum_{i=1}^Z (\mathbf{B} \times \mathbf{r}_i)^2 \quad (1.26)$$

where:

- The first term  $\mathcal{H}_0$  includes the kinetic energy and the Coulomb potential of the system.
- The second term is a paramagnetic term due to the intrinsic momentum of an atom. The momentum gets aligned to the field and it always results in an attraction force given by  $\mathbf{F} = -\nabla U_m$  with  $U_m = -\mu \cdot \mathbf{B}$ .
- The last term is a diamagnetic term that is always present (even when both  $\mathbf{L}$  and  $\mathbf{S}$  are zero). It is related to the Larmor precession which induces a change of  $\mathbf{L}$  and a decrease of  $\mu \propto -r^2 B$ . Since the energy in this case is proportional to  $r^2 B^2$  it is always positive and therefore the force is always negative: this term appertains to the system a repulsion term.

The magnetic susceptibility is a dimensionless quantity but often the molar susceptibility  $\chi_m$  or the mass susceptibility  $\chi_g$  are used because they are easier to determine with macroscopic measurements. Tab.1.4 reports the value of the susceptibility for some interesting molecules. As

**Table 1.4:** Magnetic susceptibility for Various Substances

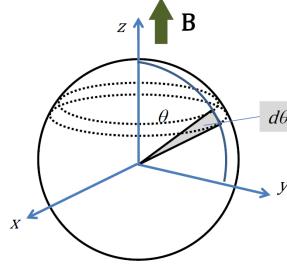
	$\chi(10^{-6})$	$\chi_m(10^{-10} m^3 mol^{-1})$
Water	-90	-16
Benzene	-7.2	-6.4
NaCl	-13.9	-3.75
Graphite $\perp$	-260	-31
Graphite $\parallel$	-3.8	-4.6
Al	+22	+2.2
Na	+7.3	+1.7

we can see, a lot of them tend to be diamagnetic since during the spontaneous formation of the molecule the atoms bond in such a way that they can complete the shell and therefore they have a zero spin angular momentum. This gave the molecule a diamagnetic behaviour.

### 1.6.2 Diamagnetism

Let's study the case of a diamagnetic material for which, since  $\chi < 0$ , the induced magnetization is opposite to the external field. The energy term associated with diamagnetism is the one in the Hamiltonian of Eq(1.26). The energy shift due to the diamagnetic term is proportional to  $\sum_{i=1}^Z (\mathbf{B} \times \mathbf{r}_i)^2$  and it is possible to prove that it is equal to:

$$\Delta E_0 = \frac{e^2 B^2}{12m_e} \sum_{i=1}^Z \langle GS | r_i^2 | GS \rangle \quad (1.27)$$



**Figure 1.4:** Magnetic dipoles distribution.

where GS stands for the ground state. We can use the thermodynamic potential of the Helmholtz free energy  $F$  to compute the magnetization  $M$  and then the susceptibility for  $N$  ions in a volume  $V$ :

$$\begin{aligned} F &= E - TS \\ M &= -\left(\frac{\partial F}{\partial B}\right)_{T=0,V} = -\frac{N}{V} \frac{\partial}{\partial B} \Delta E_0 \\ \chi &= \frac{M}{H} \simeq \frac{\mu_0 M}{B} = -\frac{N}{V} \frac{\mu_0 e^2}{6m_e} \sum_{i=1}^Z \langle GS | r_i^2 | GS \rangle \end{aligned} \quad (1.28)$$

We can now explicitly see that, as told before, the susceptibility does not depend on the field  $B$  and is therefore constant. In this relation, the most important dependence is given by the quantity  $r_i$ , *i.e.* the distance of each electron from its nucleus. More precisely, the expectation value  $\langle GS | r_i^2 | GS \rangle = \langle r_i^2 \rangle$  is the relevant quantity. The bigger the distance, the stronger the effect on the susceptibility and therefore we can neglect the influence of the core electrons on the susceptibility because for these ones the expectation value of  $r_i$  is very small. We can thus say that diamagnetism is mainly due to the electron in the outer shell and then we have:

$$\sum_{i=1}^Z \langle r_i^2 \rangle \simeq \sum_{i=1}^{Z_{eff}} \langle r_i^2 \rangle \simeq Z_{eff} r^2 \quad (1.29)$$

since for the outer shell the average distance can be considered constant. The average distance of the electrons is not important only for atoms. For example, in the Naphthalene molecule the outer electrons are shared between more carbon atoms and thus their motion is described by a larger "super-orbit" which gives the molecule a strong diamagnetic behaviour. In general, DM materials are the best ones for magnetic levitation because the induced magnetization  $M$  is always antiparallel to the external field resulting in a force that is always direct upward, even if the material rotates in the field<sup>1</sup>.

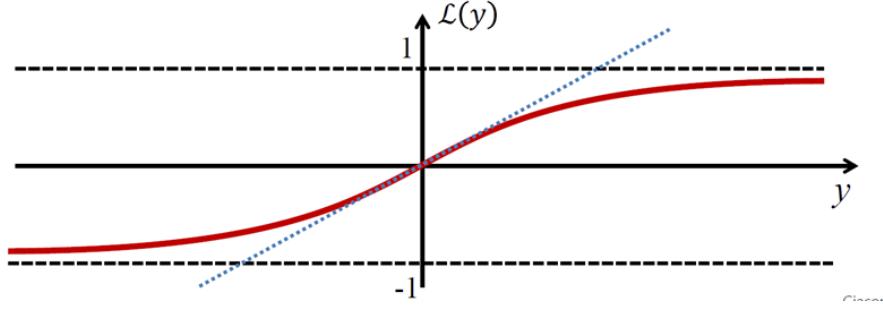
### 1.6.3 Paramagnetism

Let's now deal with paramagnetic materials which are the ones with a positive  $\chi$  and which are always attracted by a permanent magnet, *i.e.* are characterised by the presence of an attractive force. As told, this phenomenon is due to the partial alignment of atomic magnetic momenta to the external magnetic field because of the torque  $\mathbf{G}$  which acts on them. To avoid problems with the quantization of the angular and magnetic momenta and their projections in a given direction, we will start by assuming that these projections can take any possible values (no quantization) that in other words means that we will consider  $J = \infty$ . Let's consider a magnetic field  $\mathbf{B}$  direct along the  $z$ -axis and let's calculate the total magnetization along  $x$  due to  $n = \frac{N}{V}$  magnetic momenta, all identical in modulus  $\mu$  but differing in orientation. By referring to Fig.1.4, let's consider the circular ribbon on the sphere and thus the momenta oriented between  $\theta$  and  $\theta + d\theta$  with respect to  $z$ . The energy and the projection of the momentum along  $z$  will be:

$$\begin{aligned} U_i &= -\mu_i B \cos \theta \\ \mu_{z,i} &= \mu_i B \cos \theta \end{aligned}$$

---

<sup>1</sup>On YouTube you can find the video "How to levitate a Frog" in which a frog is made levitating by exploiting the huge amount of diamagnetic water molecule that every living being contains.



**Figure 1.5:** Langevin function.

If the momenta are randomly oriented in space, the statistical fraction of them forming an angle  $\theta$  with the z-axis is:

$$\frac{2\pi \sin \theta d\theta}{4\pi} = \frac{1}{2} \sin \theta d\theta$$

that is the normalized area of the ribbon. If we want to consider the case in which the temperature  $T$  is different from zero<sup>2</sup>, we need to use the Boltzmann statistic. This gives us:

$$\langle \mu_z \rangle = \frac{\int_0^\pi \mu \cos \theta e^{\frac{\mu B \cos \theta}{k_B T}} \frac{1}{2} \sin \theta d\theta}{\int_0^\pi e^{\frac{\mu B \cos \theta}{k_B T}} \frac{1}{2} \sin \theta d\theta} \quad (1.30)$$

By defining the quantities:

$$y = \frac{\mu B}{k_B T}, \quad x = \cos \theta \quad (1.31)$$

Eq.(1.30) becomes:

$$\langle \mu_z \rangle = \mu \frac{\int_{-1}^1 x e^{yx} dx}{\int_{-1}^1 e^{yx} dx} \quad (1.32)$$

This leads to the definition of the **Langevin function**  $\mathcal{L}(y)$ :

$$\mathcal{L}(y) = \frac{\langle \mu_z \rangle}{\mu} = \coth y - \frac{1}{y} \quad (1.33)$$

This function's behaviour is depicted in Fig.1.5. For small  $y$ , *i.e.* the magnetic energy is much smaller than the thermal one, the Langevin function can be expanded in series and simplified:

$$\mathcal{L}(y) \simeq \frac{y}{3} + \mathcal{O}(y^3) \quad (1.34)$$

We also note that  $M = n \langle \mu_z \rangle$  and the saturation magnetization, *i.e.* the magnetization when all the dipoles are aligned, is  $M_s = n\mu$ . This implies:

$$\frac{M}{M_s} = \frac{\langle \mu_z \rangle}{\mu} \simeq \frac{y}{3} = \frac{\mu B}{3k_B T} \quad (1.35)$$

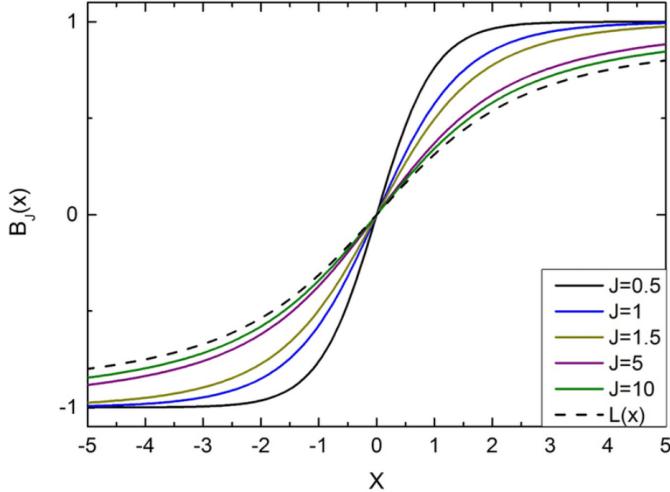
We can compute the susceptibility  $\chi$  for a paramagnet in the case of small  $y$ . This results in a proportionality relation which is known as **Curie's Law**:

$$\chi = \frac{M}{H} \simeq \frac{\mu_0 M}{B} = \frac{\mu_0}{B} \frac{\mu B}{3k_B T} M_s = \frac{n\mu_0 \mu^2}{3k_B T} \propto \frac{1}{T} \quad (1.36)$$

We note, then, that the susceptibility for a PM is constant with respect to the magnetic field  $B$  only in the case of small  $y$ , differently from the DM case.

---

<sup>2</sup>If  $T = 0$  there is no thermal excitation and even an infinitesimally small field  $B$  can align all the dipoles.



**Figure 1.6:** Brillouin function.

#### 1.6.4 Qunatum Paramagnetism

As told, we have deliberately ignored the quantization of the magnetic dipoles. Let's start considering quantization in the simple, but relevant, case in which  $J = \frac{1}{2}$ . In this case, we deal with only two possible projections of the moment along z:  $m_j = \pm \frac{1}{2} \Rightarrow \mu_{z,i} = \mp \mu_B$ . We can calculate the magnetization by using the statistical probability. At a finite temperature  $T$  this is given by the partition function:

$$\langle \mu_z \rangle = \langle g\mu_B m_J \rangle = \frac{\mu_B e^{\frac{\mu_B B}{k_B T}} + \mu_B e^{-\frac{\mu_B B}{k_B T}}}{e^{\frac{\mu_B B}{k_B T}} + e^{-\frac{\mu_B B}{k_B T}}} = \mu_B \tanh \frac{\mu_B B}{k_B T} \quad (1.37)$$

Similarly to what was done before we can define  $y = \frac{g\mu_B JB}{k_B T} = \frac{\mu_B B}{k_B T}$  for  $J = \frac{1}{2}$  and  $g = 2$  and we obtain:

$$\frac{M}{M_s} = \frac{\langle \mu_z \rangle}{\mu_B} = \tanh y \quad (1.38)$$

Such a function is similar to the Langevin function but for small  $y$  it can be expanded as  $\tanh y \approx y$  which means that the derivative at the origin is 1 instead of 1/3. By doing as before we find Curie's Law once more:

$$\chi = \frac{n\mu_0\mu_B^2}{k_B T} \propto \frac{1}{T} \quad (1.39)$$

In a similar way for any possible values of  $J$  we can find that the ratio between the two magnetizations is equal to the **Brillouin function**  $\mathcal{B}_J(y)$ :

$$\frac{M}{M_s} = \mathcal{B}_J(y) \quad (1.40)$$

with:

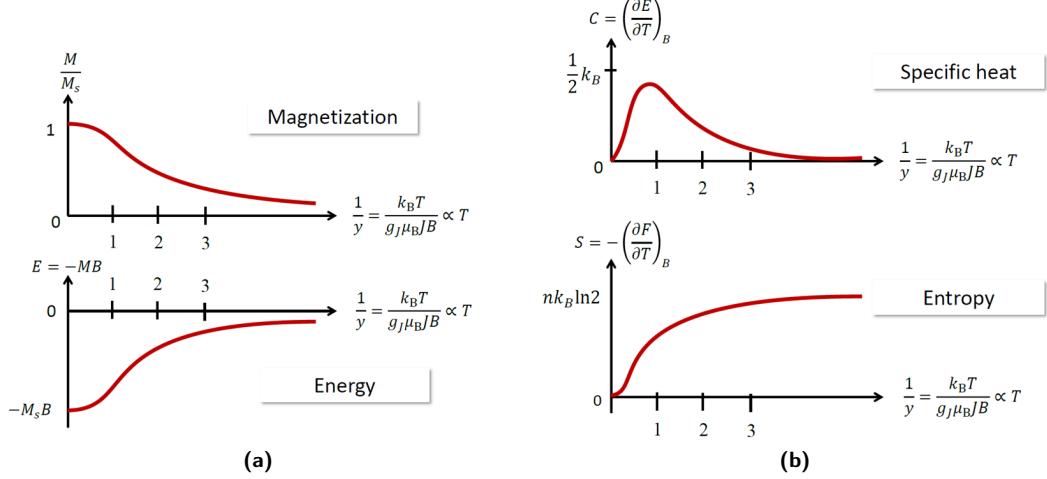
$$\begin{aligned} y &= \frac{g_J \mu_B J B}{k_B T} \\ \chi &= \frac{n\mu_0\mu_{eff}^2}{3k_B T} \\ \mu_{eff} &= g_J \mu_B \sqrt{J(J+1)} \end{aligned} \quad (1.41)$$

The trend of the Brillouin function for different values of  $J$  is depicted in Fig.1.6.

#### 1.6.5 Magnetization, Energy, Specific Heat and Entropy

Let's briefly discuss the behaviour of four different physical quantities in the case of a PM material depicted in Fig.1.7. We can notice that, since the magnetic energy associated with a material is given by  $E = -MB$  the energy trend is equal to the magnetization one but overturned with respect

to the horizontal axis. Moreover, since the specific heat of a material is given by  $C = \left(\frac{\partial E}{\partial T}\right)_B$  we can observe that the specific heat has a maximum for  $\frac{1}{y} \approx 1$  and both for high temperature and low temperature the specific heat tends to zero. The graph also makes evident that for very low and high temperatures the system is either really ordered or really disordered, respectively. This means that a small change in the temperature does not change too much the energy while in the middle region, a small variation of  $T$  is associated with a huge variation of the energy. This behaviour is reflected in the trend of the entropy depicted in Fig.1.7 as well.



**Figure 1.7:** Magnetization, Energy, Specific Heat and Entropy in function of  $\frac{1}{T}$

# Chapter 2

# Magnetic Techniques

## 2.1 Nuclear Magnetic Resonance

Nuclear Magnetic Resonance (NMR) exploits the non-zero nuclear spin of some elements (*e.g.*  $^1H$ ,  $^{13}C$ ,  $^{19}F$ ) by putting them in a very strong magnetic field to split their ground state into different energy levels separated by a tiny energy thanks to Zeeman splitting. Then, a suitable electromagnetic radiation, usually in the Radio Frequency (RF) spectrum, is used to induce transitions among the split levels. Since the resonance on the stimulated transitions is extremely sharp, by measuring the resonance frequency we can measure very accurately the effective magnetic field seen by the nucleus, which depends on the chemical environment where the nucleus is placed. This allows us to study the chemical composition of different materials by recognising the proper frequency. Moreover, we can also measure the volume-density distribution of the nuclei in the material through radiation absorption to make images in the so-called Magnetic Resonance Imaging (MRI).

We know that the nuclear magnetic moment and the Zeeman splitting are:

$$\mu = g_I \mu_N I$$
$$\Delta E = g_I \mu_N B$$

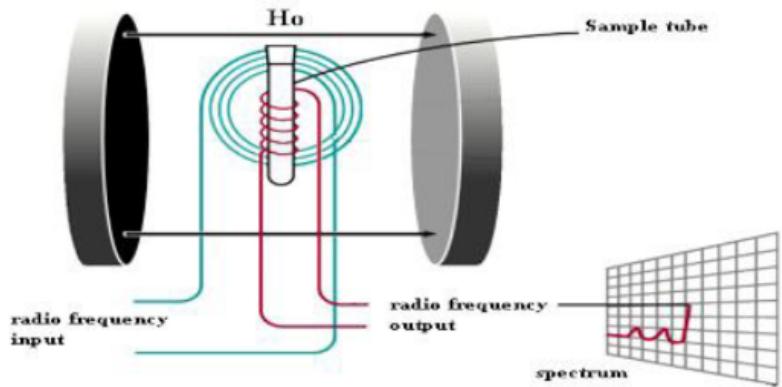
In Tab.2.1 some relevant quantities for different atoms are listed where  $\omega_L = -\gamma B$  is the Larmor precession frequency and  $\gamma$  is the gyromagnetic ratio. These atoms have very small magnetic momenta and, when placed in a strong magnetic field, they behave as paramagnetic atoms. Nevertheless, whatever the field, in general, the ratio  $\frac{\mu_B}{k_B T}$  stays always close to one since it's difficult and energetically expensive to cool down the system enough to make the ratio much smaller than one. By looking at the frequency column we can see that, as told before, almost all the frequencies are located in the RF spectrum ( $10^6 - 10^8 Hz$ ).

**Table 2.1:** Nuclear Magnetic Properties.

Mass	Element	Moment ( $\mu_N$ )	$I$	$g$	$\omega_L$ (MHz/T)	$\gamma$	Rel. NMR Sensitivity
1	H	2.79	1/2	5.58	42.6	267	1
2	D	0.85	1	0.85	6.5	41	
7	Li	3.25	3/2	2.17	16.5	103	
13	C	0.70	1/2	1.40	10.7	67	0.016
14	N	0.40	1	0.40	3.1	19	0.001
19	Na	2.62	1/2	5.25	40.4	25	0.83

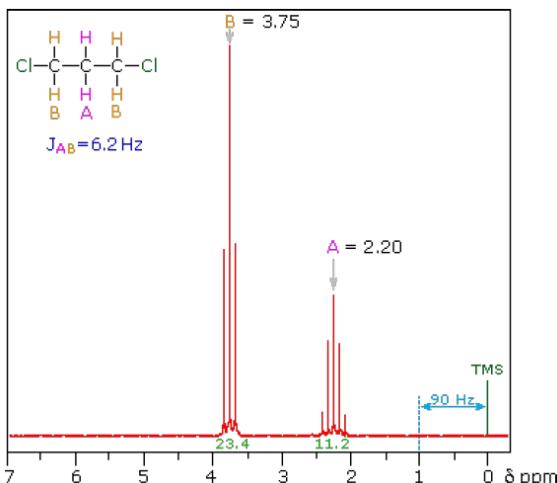
### 2.1.1 Working Principle

As depicted in Fig.2.1 the material to be examined is placed in a magnetic field ( $\sim 1 - 10T$ ) and two coils, which work as antennas, are connected and wrapped around the sample. The first antenna gives as an input a variable RF signal that is used to induce transitions among the nucleus energy levels. When the frequency is such that it stimulates a transition, an absorption event occurs



**Figure 2.1:** NMR Schematic Design.

and the following emission is detected by the other output antenna. Since, as seen before, the nuclear Zeeman splitting depends on the magnetic field and, in particular,  $\Delta E \propto B$ , the magnetic field must be very homogeneous to not include a position dependence on the measurements. This technique works particularly well and with very high accuracy due to the very small intrinsic linewidth ( $\sim 1\text{Hz}$ ) compared to the energy of the emitted signal ( $\sim 10^6 - 10^7\text{Hz}$ ). Because of this high accuracy, NMR is used to detect the chemical shifts of atoms in different compounds. In fact, when an atom bonds with others in the formation of a molecule, its energy levels are shifted depending on the atoms to which it is bonded which is the so-called chemical environment. NMR is often used in chemistry to determine, in organic compounds, where hydrogen and carbon bond to each other. In Fig.2.2 is possible to see the different peaks associated with different hydrogen atoms on a particular molecule.



**Figure 2.2:** Chemical Shift for  $C_3H_6Cl_2$  molecule.

### 2.1.2 NMR Intensities and Timescales

We can write the rate equations for our two-level system ( $I = \frac{1}{2}$ ) by considering that the transition probability is proportional to the transverse magnetic field. Moreover, the transition probability from one level to the other does not depend on the "direction" of the transition, *i.e.* it is the same for absorption and emission. Being  $W$  the RF power emitted by the coil we have:

$$\begin{cases} \frac{dN_+(t)}{dt} = WN_-(t) - WN_+(t) \\ \frac{dN_-(t)}{dt} = WN_+(t) - WN_-(t) \end{cases} \quad (2.1)$$

Then, the population imbalance  $n(t) = N_+(t) - N_-(t)$  is:

$$\frac{dn(t)}{dt} = -2Wn(t) \Rightarrow n(t) = n(0)e^{-2Wt} \quad (2.2)$$

Starting from an unbalanced population (because of the thermal energy) the RF phonons get absorbed until the two populations  $N_+$  and  $N_-$  become identical. The energy of the system at any time  $t$  is:

$$\begin{aligned} E(t) &= N_-(t)E_- + N_+(t)E_+ = N_-(t)E_- + N_+(t)(E_- + \hbar\omega) = \\ &= E_0(N_- + N_+) + \frac{1}{2}\hbar\omega n(t) \end{aligned} \quad (2.3)$$

where  $\hbar\omega$  is the energy splitting between the ground state and the first excited level and  $E_0$  is the mean value of energy. Thus, the derivative of the energy is:

$$\frac{dE(t)}{dt} = \frac{1}{2}\hbar\omega \frac{dn(t)}{dt} = -W\hbar\omega n(t) \quad (2.4)$$

From this equation we can clearly see that the energy varies until the population imbalance is different from zero: when the two populations become identical the energy varies no more. When the RF is on after a number of time constants  $\tau = \frac{1}{2W}$  the two populations get balanced and the energy is not absorbed anymore. However, in the absence of RF, thermal equilibrium implies:

$$\left(\frac{N_+}{N_-}\right)_0 = e^{-\frac{\hbar\omega}{k_B T}} \quad (2.5)$$

This means that the system, starting from  $n = 0$  will go back to the thermal equilibrium with a certain time constant  $T_1$  called **spin-lattice relaxation time**. This phenomenon can be explained qualitatively by imaging that all the nuclei when exposed to the magnetic field and excited by the RF signal start precessing (Larmor precession), and that when the RF is turned off they will slowly stop their motion. Mathematically, this means that the population imbalance becomes:

$$n(t) = n_0 \left(1 - e^{-\frac{t}{T_1}}\right) \quad (2.6)$$

The time scale for the spin-lattice relaxation time  $T_1$  is in the order of  $0.1 - 100s$  which is a huge time compared, for example, to the lifetime of an electron in a core level which is in the order of  $10^{-15}s$ . The spin-lattice relaxation phenomenon is spontaneous and always present so, for a better description of the absorption rate, we have to include the thermal effect also when the RF is on at its steady state. By combining the two exponentials:

$$\frac{dn(t)}{dt} = -2Wn(t) + \frac{n_0 - n(t)}{T_1} \quad (2.7)$$

that, for the steady state  $\frac{dn(t)}{dt} = 0$  becomes:

$$n(t) = \frac{n_0}{1 + 2WT_1} \Rightarrow \frac{dE(t)}{dt} = -W\hbar\omega n(t) = -n_0\hbar\omega \frac{W}{1 + 2WT_1} \quad (2.8)$$

For a large RF power  $W$ , the derivative of the absorbed energy goes asymptotically to:

$$\frac{dE(t)}{dt} \underset{W \rightarrow \infty}{\sim} -\frac{n_0\hbar\omega}{2T_1}$$

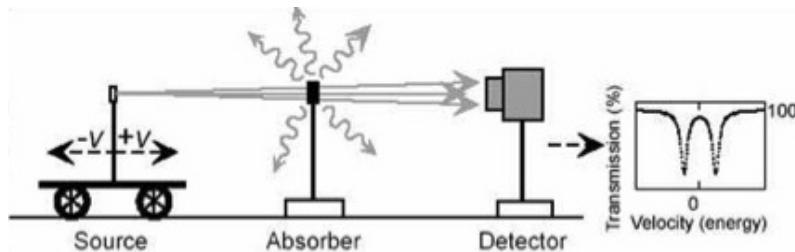
that means that we have a saturation in the energy when the power goes to infinity. There is also  $T_2$ , spin-spin relaxation time, which is the time for the collective magnetic system to lose coherence in the precession. We can imagine that all the nuclei precess together in phase when the RF is on the right frequency and when it is turned off this collective behaviour vanishes with time. Usually,  $T_2 < T_1$ , it qualitatively means that the nuclei lose coherence before stopping their precession.

One of the most famous and important applications of the NMR is the Magnetic Resonance Imaging (MRI) which is used in the medical field. The signal coming from the excitation of a  $^1H$  atom as a function of its position is collected and it is used to reconstruct the distribution of this element in the human body. This particular atom is used since it is part of the water molecule, present in

the whole body with different distributions. Finally, there is another technique that exploits the same principle of NMR but uses the Zeeman splitting of electron spin states which is the Electron Spin Resonance (ESR). In this case, the energy splitting is 1000 times larger (microwave radiation instead of RF) and the intrinsic linewidth is much broader since the electron's spin is more strongly coupled to the rest of the world than the nuclear one. ESR is used to study hyperfine interaction and chemical environment.

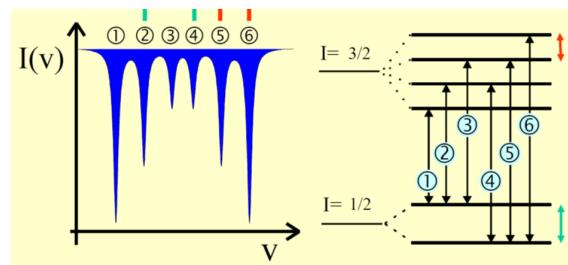
## 2.2 Mössbauer Spectroscopy

It's a spectroscopy technique that exploits another nuclear phenomenon: the  $\gamma$ -emission. When a  $\gamma$ -emission occurs inside a crystal, *i.e.* a photon and other particles are emitted by a nucleus in an excited state, the kickback to which the decaying nucleus is usually subject is amortized by the whole crystal resulting, then, in no loss of energy in the emitted photon ( $\gamma$ -ray). This phenomenon is the so-called Mössbauer effect. Because of this, the emitted photon has the same energy as the energy splitting of the nucleus energy levels and therefore, the photon can be used to stimulate a transition inside another identical nucleus. Since the energy levels of a nucleus depend on the magnetic environment in which it is placed, this technique can be also used to study magnetism. Differently from the NMR, in this case, the electromagnetic radiation used is in the x-rays/ $\gamma$ -rays spectrum and, as well as the NMR, the emitted radiation has very sharp peaks because of the sharpness of the nuclear resonance. The idea is that using a radioactive source of the same isotope to be studied, one can measure the absorption of the photons by the sample and measure tiny shifts of the resonance that are due to the magnetic and electronic environment of the absorbing nuclei. The experimental setup is depicted in Fig.2.3. The target to be studied



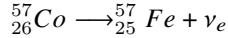
**Figure 2.3:** Mössbauer Spectroscopy experimental setup.

(absorber) is located between the detector and the emitting source that is placed onto a moving kart. The source and the studied material are made of the same material, nevertheless, the nuclei energy levels of the absorber are slightly different from the one of the source. This is because of, for example, a magnetic field, to which only the absorber is subjected. To "tune" the frequency of the emitted phonon, the source is made moving back and forth with a certain velocity  $v$ . Thanks to the Doppler effect, the frequency of the radiation impinging on the absorber will be slightly different depending on the velocity of the source. In this way, it is possible to vary continuously the frequency of the radiation and when the frequency is such to be equal to one of the nuclei energy splitting, an absorption event occurs and the radiation measured by the detector drops rapidly. By plotting and measuring the absorption intensity as a function of the velocity, it is possible to retrieve the energy splitting in the atomic nuclei and, for example, if present, the magnetic field seen by the nuclei. In general, the nucleus of an atom has several energy levels as depicted in



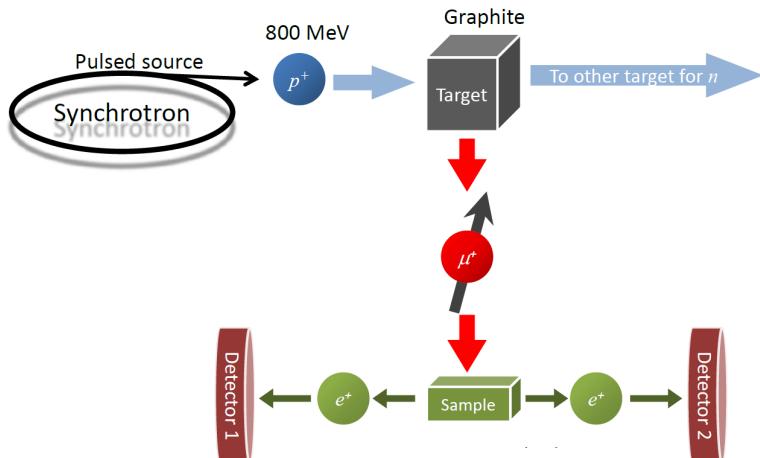
**Figure 2.4:** Mössbauer Spectroscopy energy splitting example.

Fig.2.4 which depend on the spin  $I$  of the nucleus. When doing a Mössbauer Spectroscopy these splittings result in several drops in the absorption that, thanks to the extremely sharp resonance of the nucleus, can be resolved easily. An example is the study of the iron isotope  $^{57}Fe$  which is present in nature with only the 2% of abundance. To produce and study it, the  $^{57}Co$  is used because it decays, through electronic capture, into  $^{57}Fe$ :



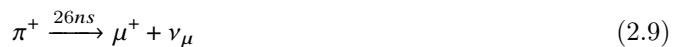
## 2.3 Muon Spin Rotation

Muon Spin Rotation ( $\mu$ SR) is another technique which exploits the spin magnetic moment of an elementary particle, the muon, and its precession. In this case, the precession frequency is measured in the time domain by detecting the emission direction of the decay product (positron  $e^+$ ) of the precessing muon. The frequency is proportional to the effective magnetic field in the position of the muon in the material and therefore measuring the frequency is a way of measuring the local magnetic field inside a material and thus the magnetic properties of the material. The muon  $\mu^-$  is an "exotic" particle with a very short lifetime ( $\sim \mu s$ ) but it is easy to produce. Muons are leptons, their mass is approximately  $105 MeV/c^2$  which is between the mass of the proton and the electron. They are negatively charged ( $q_\mu = -e$ ) and they have spin 1/2. It is this last property that allows them to precess when exposed to a magnetic field. More specifically, the antiparticle of the muon, the antimuon  $\mu^+$ , is used because it has a positive charge  $e$  and it is thus repelled by the nuclei of the material to be studied. The instability of muons is used to study materials since, as we will see, they decay into positrons and they emit them mostly in the direction of their magnetic dipole. The setup for  $\mu$ SR is depicted in Fig.2.5. Muons are produced thanks to a synchrotron. Initially,



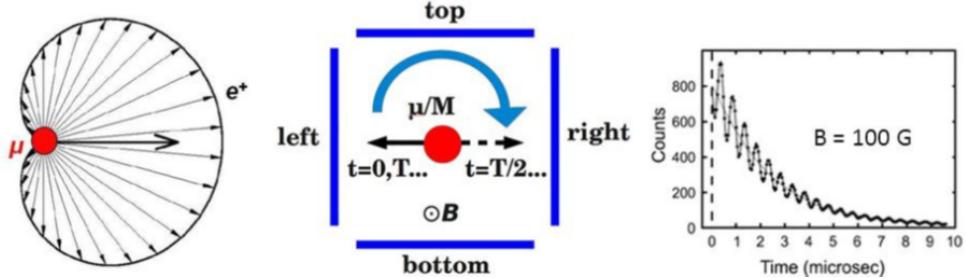
**Figure 2.5:**  $\mu$ SR experimental setup.

the synchrotron produces protons that are shot against a graphite target. The interaction between the graphite atoms and the protons produces pions  $\pi^+$  that quickly decay into antimuons  $\mu^+$ :



Because of the type of decay of the pions, the antimuons produced are 100% spin-polarized, *i.e.* they all have the same spin direction. In fact, pions have no spin and the neutrino has spin antiparallel to its momentum: for the momentum conservation law to be valid the emitted antimuon has to have spin antiparallel as well. This behaviour is a consequence of the parity law and of the neutrino characteristics. Thanks to an electric field the muons are then collected and directed towards the sample to study. Inside the material, because of inelastic scattering, the antimuons slow down and they stop into the lattice in a "convenient" site. In fact, being the antimuons positively charged particles they are repelled by the nuclei of the sample and they will stop their motion in an "empty" site of the lattice, as far as possible from the nuclei<sup>1</sup>. After a certain time, the antimuon will decay

<sup>1</sup>What really happens is that an antimuon loses energy by forming, with an electron of the material, the so-called Muonium, a hydrogen-like atom made by an antimuon in the centre and one electron which orbits around.



**Figure 2.6:**  $\mu$ SR probability distribution of emission, scheme and plot.

into a positron:

$$\mu^+ \xrightarrow{2.2\mu s} e^+ + \nu_e + \bar{\nu}_\mu$$

If there is a magnetic field or the target is spontaneously magnetized, the muon spins start to precess synchronously. Using two (or more) symmetric detectors we can monitor the precession, as an oscillating intensity on top of the exponential decrease of intensity due to the decay. Fig.2.6 shows, on the graph on the right, this oscillating behaviour on top of the exponential one. Luckily, the period of the oscillations of  $\mu^+$  is much shorter than the decay time resulting, then, in several oscillations along the exponential behaviours. Moreover, the figure shows on the left the probability distribution of the emission direction during the antimuon decay. In the picture, the thicker arrow represents the direction of the antimuon spin. It is possible to see that the probability is maximum in the very same direction of the spin and therefore, statistically, the majority of emitted positrons will be in the direction of the spin and, since the antimuon (and its spin) is precessing, the spin direction will change in time as shown in the middle picture in Fig.2.6. By using more than one detector the frequency  $\omega$  of the oscillations can be computed and then, since:

$$\omega = \frac{ge}{2m_\mu} B \quad (2.10)$$

we can find the field inside the material. This technique is often used to study the magnetic properties of very thin films since the implantation of the antimuon in the material does not damage the sample since it occurs well inside the material ( $> 1\mu m$ ).

# Chapter 3

# Magnetic Interaction

From the very beginning of the studies about magnetism on a microscopic scale, scientists agreed that the magnetic properties of matter can not come only from the contribution of the atomic currents but there must be something else, related to the atom itself. Today we know that this something is the spin and it contributes to the magnetic properties of material as we have seen so far. Nevertheless, some particular behaviours such as ferromagnetism, required an additional ingredient. It is well known, indeed, that the dipolar interaction between two atomic magnetic momenta can not explain the perfect dipole alignment that is seen in FM materials. A dipole would generate a field that would tend to align another close dipole in an antiparallel way resulting, then, in a zero magnetization. It is then clear that is necessary to consider a quantum description of the phenomenon that leads us to the exchange interaction. This interaction is the key of magnetism and it can explain the huge energies coming from the alignment of the dipoles in a ferromagnet. The exchange interaction is physically due to the electrostatic interaction between electrons of different atoms in a crystal but, from a mathematical and theoretical point of view, it is convenient for us to consider that this interaction is due to the interaction between spins.

## 3.1 Exchange Interaction

In general, every time two particles with spin 1/2 interact one with the other through a Hamiltonian of amplitude  $A$  of the type:

$$\hat{\mathcal{H}} = A \hat{\mathbf{S}}_a \cdot \hat{\mathbf{S}}_b \quad (3.1)$$

the energy of the system depends on the relative orientation of the two spins, *i.e.* on the total spin. Then we can consider:

$$\begin{aligned} \hat{\mathbf{S}}_{tot} &= \hat{\mathbf{S}}_a + \hat{\mathbf{S}}_b \\ \hat{\mathbf{S}}_{tot}^2 &= \hat{\mathbf{S}}_a^2 + \hat{\mathbf{S}}_b^2 + 2\hat{\mathbf{S}}_a \cdot \hat{\mathbf{S}}_b \end{aligned} \quad (3.2)$$

The total spin of each electron is given by:

$$\begin{aligned} \hat{\mathbf{S}}_a^2 &= \hbar^2 s(s+1) = \hbar^2 \frac{1}{2} \left( \frac{1}{2} + 1 \right) = \frac{3}{4} \hbar^2 \\ \hat{\mathbf{S}}_b^2 &= \hbar^2 s(s+1) = \hbar^2 \frac{1}{2} \left( \frac{1}{2} + 1 \right) = \frac{3}{4} \hbar^2 \end{aligned}$$

and the overall spin of the system is:

$$\hat{\mathbf{S}}_{tot}^2 = \hbar^2 s_{tot}(s_{tot}+1) = \begin{cases} 2\hbar^2 & s_{tot} = 1 \Rightarrow |\uparrow\uparrow\rangle \\ 0 & s_{tot} = 0 \Rightarrow |\uparrow\downarrow\rangle \end{cases}$$

Therefore we have (let's neglect the constant  $\hbar$ ):

$$\hat{\mathbf{S}}_a \cdot \hat{\mathbf{S}}_b = \frac{\hat{\mathbf{S}}_{tot}^2 + \hat{\mathbf{S}}_a^2 + \hat{\mathbf{S}}_b^2}{2} = \frac{1}{2} \left( \hat{\mathbf{S}}_{tot}^2 - \frac{3}{4} - \frac{3}{4} \right) = \begin{cases} +\frac{1}{4} & s_{tot} = 1 \Rightarrow |\uparrow\uparrow\rangle \\ -\frac{3}{4} & s_{tot} = 0 \Rightarrow |\uparrow\downarrow\rangle \end{cases} \quad (3.3)$$

This means that the system has two energy levels: one singlet ground level  $E_S = -\frac{3}{4}A$  in which the two spins are opposite, and one triplet excited level  $E_T = \frac{1}{4}A$  which corresponds to the case of equal spins. Now, what about the wave function of this system? Let us consider two electrons belonging to the same atom (*e.g.* the two external electrons of  $He$ ) and let's try to write their joint wave function as the product of the individual wave functions:

$$\Psi = \psi_a(\mathbf{r}_1)\psi_b(\mathbf{r}_2)S_aS_b$$

This wave function is not good for our case because it is not invariant to the exchange operator  $\hat{P}$ . In fact for fermions, the exchange operation must be antisymmetric, which means:

$$\hat{P}|\mathbf{x}_1, \mathbf{x}_2\rangle = -|\mathbf{x}_2, \mathbf{x}_1\rangle$$

It is easy to verify that this does not happen with such a wave function. Therefore, we need to better define the wave function in order to let it be antisymmetric. We should already know that this results in the definition of two wave functions:

$$\begin{aligned}\Psi_S &= \frac{1}{\sqrt{2}} [\psi_a(\mathbf{r}_1)\psi_b(\mathbf{r}_2) + \psi_a(\mathbf{r}_2)\psi_b(\mathbf{r}_1)] \chi_S = \tilde{\Psi}_S \chi_S \\ \Psi_T &= \frac{1}{\sqrt{2}} [\psi_a(\mathbf{r}_1)\psi_b(\mathbf{r}_2) - \psi_a(\mathbf{r}_2)\psi_b(\mathbf{r}_1)] \chi_T = \tilde{\Psi}_T \chi_T \\ \chi_S &= \frac{|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle}{2} \\ \chi_T &= |\uparrow\uparrow\rangle; \quad \frac{|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle}{2}; \quad |\downarrow\downarrow\rangle\end{aligned}$$

These two wave functions are overall odd to exchange, but one is symmetric in the spatial part and antisymmetric in the spin ( $\Psi_S$ , singlet) and the other one is antisymmetric in the spatial part and symmetric in the spin one ( $\Psi_T$ , triplet). We want now verify if the two states corresponding to the two different wave functions are different in energy. We see from their definition that the spin part of the joint wave functions,  $\chi_S$  and  $\chi_T$ , are normalized and thus, the difference in energy comes from the spatial part only and the effective Hamiltonian interaction  $\hat{\mathcal{H}}$ :

$$\begin{aligned}E_S &= \int \tilde{\Psi}_S^* \hat{\mathcal{H}} \tilde{\Psi}_S d\mathbf{r}_1 d\mathbf{r}_2 \\ E_T &= \int \tilde{\Psi}_T^* \hat{\mathcal{H}} \tilde{\Psi}_T d\mathbf{r}_1 d\mathbf{r}_2\end{aligned}$$

It is possible to demonstrate that in the difference between these two energies, only the mixed terms survive:

$$E_S - E_T = 2 \left\{ \int \psi_a^*(\mathbf{r}_1)\psi_b^*(\mathbf{r}_2)\hat{\mathcal{H}}\psi_a(\mathbf{r}_2)\psi_b(\mathbf{r}_1) + \int \psi_a^*(\mathbf{r}_2)\psi_b^*(\mathbf{r}_1)\hat{\mathcal{H}}\psi_a(\mathbf{r}_1)\psi_b(\mathbf{r}_2) \right\} \equiv 2J \quad (3.4)$$

where we defined the value of the integral as  $J$  which is the so-called **exchange integral** or exchange constant, which represents the exchange interaction.  $J$  is an integral over the spatial part of the wave functions only. Now, by substituting the result obtained in Eq.(3.3) in the Hamiltonian of Eq.(3.1) we can link the effective Hamiltonian to the total spin operator, assigning the energy difference to the spin part. It is possible to demonstrate that the following Hamiltonian respects the definitions of the wave function seen before and gives rise to the two energy  $E_S$  and  $E_T$  as in Eq.(3.4):

$$\hat{\mathcal{H}} = A\hat{\mathbf{S}}_a \cdot \hat{\mathbf{S}}_b = \frac{1}{4}(E_S + 3E_T) - (E_S - E_T)\hat{\mathbf{S}}_a \cdot \hat{\mathbf{S}}_b = -(E_S - E_T)\hat{\mathbf{S}}_a \cdot \hat{\mathbf{S}}_b \quad (3.5)$$

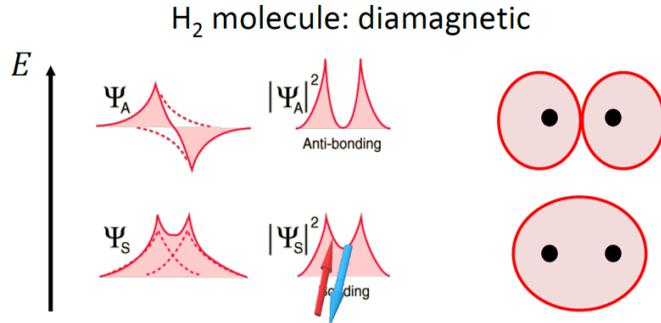
because the first term is zero if we substitute  $E_S = -3A/4$  and  $E_T = A/4$ . By using the definition of  $J$ :

$$\hat{\mathcal{H}} = -2J\hat{\mathbf{S}}_a \cdot \hat{\mathbf{S}}_b \quad (3.6)$$

Furthermore, from Eq.(3.4) we have:

$$J = \frac{E_S - E_T}{2} \quad (3.7)$$

In general, since  $E_S$  and  $E_T$  can be one higher or lower than the other,  $J$  can be either positive or negative. If  $J < 0$  then  $E_S < E_T$  and the ground state is a singlet state (antiparallel spins)



**Figure 3.1:** Molecular Hydrogen wave functions distribution.

while, if  $J > 0$  the ground state is a triplet state, the spins are parallel and the material can be ferromagnetic. Fig.3.1 shows the case of an  $H_2$  molecule. We can see that the bonding state, associated with a higher electron density between the nuclei, is the one with a symmetric spatial part of the wave functions, *i.e.* it is a singlet state. Therefore,  $J < 0$  and the molecule compound is diamagnetic. As we can see from the picture, the anti-bonding state is related to a lower electron density between the nuclei because of the strong Coulombian repulsion.

## 3.2 Heisenberg Model

The formalism for two generic electrons used above can be generalized to the atomic spin momenta interaction (and even to the atomic total momenta). For a lattice of magnetic atoms, one can consider the effect of the exchange interactions of each atomic spin with all the other atomic spin in the lattice. This is the idea behind the Heisenberg model that considers the interaction between atoms by writing a Hamiltonian:

$$\hat{\mathcal{H}} = - \sum_{i,j \neq i} J_{ij} \hat{\mathbf{S}}_i \cdot \hat{\mathbf{S}}_j = -2 \sum_{i,j > i} J_{ij} \hat{\mathbf{S}}_i \cdot \hat{\mathbf{S}}_j \quad (3.8)$$

where  $J_{ij}$  is the exchange interaction for each atom in the lattice. Although the Heisenberg model provides the ingredient for spontaneous magnetic order, it is very difficult to solve, even under simplified conditions. A different approach is required to solve the problem. We can start simplifying the model by considering only the exchange interaction for the nearest neighbours but once again, this is not enough. We need to apply a mean-field approach which consists of replacing the one-to-one interaction between neighbours with an effective magnetic field  $\mathbf{B}_{mf}$  called molecular or mean field. This field acts equally on all the magnetic momenta and it comes from the presence of all the other momenta present in the lattice, considered to be already ordered.

## 3.3 Weiss Model for Ferromagnetism

The use of a molecular field was introduced by Pierre-Ernest Weiss in 1907 as a mathematical expedient to explain the origin of ferromagnetism. In a ferromagnet, which has a spontaneous alignment of all the spins in the same direction, the complete Hamiltonian is:

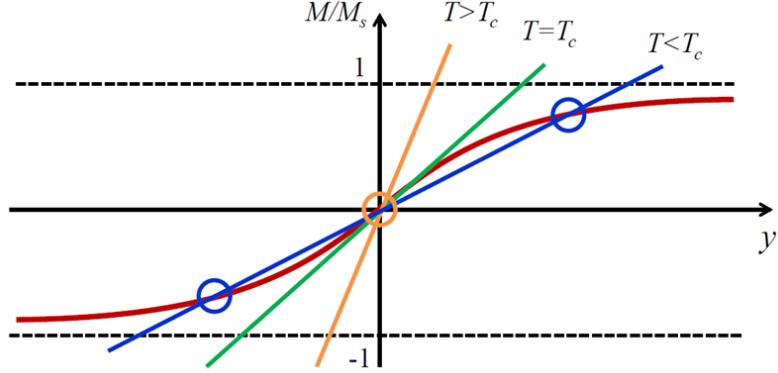
$$\hat{\mathcal{H}} = -2 \sum_{i,j > i} J_{ij} \mathbf{S}_i \cdot \mathbf{S}_j + g\mu_B \sum_i \mathbf{S}_i \cdot \mathbf{B} \quad (3.9)$$

where we considered  $\mathbf{J} = \mathbf{S} \Rightarrow \mathbf{L} = 0$  for simplicity. Then, we can define the molecular field felt by each site<sup>1</sup>, and in particular by the site  $i$ , as:

$$\mathbf{B}_{mf} = -\frac{2}{g\mu_B} \sum_{j \neq i} J_{ij} \mathbf{S}_j \quad (3.10)$$

---

<sup>1</sup>Here we are focusing on the site  $i$  so we have to add to the molecular field a factor of 2 because we are not counting the specular exchange from  $j$  to  $i$ . By doing this we can keep the summation over all  $j \neq i$ .



**Figure 3.2:** Graphical solution of the Weiss model  $\mathbf{B} = 0$ .

so that we can write the Hamiltonian as:

$$\hat{\mathcal{H}} = g\mu_B \sum_i \mathbf{S}_i \cdot (\mathbf{B} + \mathbf{B}_{mf}) \quad (3.11)$$

To get a solution for such a model we need to add two ingredients. Firstly, we will consider that the link between the molecular field  $\mathbf{B}_{mf}$  and the magnetization  $\mathbf{M}$  is as simple as possible, *i.e.* of direct proportionality<sup>2</sup>:  $\mathbf{B}_{mf} = \lambda \mathbf{M}$ . Secondly, we will assume we can apply again the statistical model used for the paramagnetism. It can be proven that, for a generic moment  $J$ , the relative magnetization  $M/M_s$  for a paramagnetic system is described by the Brillouin function  $B_J(y)$  that is:

$$\frac{M}{M_s} = B_J(y) = \frac{2J+1}{2J} \coth \frac{2J+1}{2J}y - \frac{1}{2J} \coth \frac{1}{2J}y \quad (3.12)$$

with  $y = \frac{g_J \mu_B J B}{k_B T}$ . In the expression of  $y$  the magnetic field is considered to be the total one that is:  $\mathbf{B}_{tot} = \mathbf{B} + \mathbf{B}_{mf} = \mathbf{B} + \lambda \mathbf{M}$ . We can obtain two equations for  $\mathbf{M}$ , the first one is the one related to the Brillouin function while the second comes from the expression of  $y$ :

$$\begin{cases} M = M_s B_J(y) \\ M = \frac{k_B T y}{g_J \mu_B J \lambda} - \frac{B}{\lambda} \end{cases} \quad (3.13)$$

This means that the solution is not just given by the Brillouin function once  $B$  and  $T$  are known, but it is the intersection of the Brillouin function with a straight line passing through the origin for  $B = 0$ . Although it is impossible to solve this system analytically, the solution can be found graphically. In particular, let's at first study the case of zero field ( $\mathbf{B} = 0$ ). The graphical solution in this case is shown in Fig.3.2. We notice that there exists a temperature threshold  $T_C$  below which there is only a solution for  $\mathbf{M} = 0$  and above which a non-trivial solution exists. Since we are studying the case of  $\mathbf{B} = 0$ , having a non-zero solution for the magnetization at  $T < T_C$  means that this magnetization is created spontaneously by the material. In other words, this means that the dipoles inside the material organize themselves in an ordered way spontaneously. This is one of the most peculiar characteristics of ferromagnet that is, at least qualitatively, well explained by the Weiss model. Moreover, the model allows us to find that particular threshold value of the temperature, the well-known **Curie temperature**. To find it, we notice that the slope of the line at  $T = T_C$  is equal to the slope of the Brillouin function for small values of  $y$ . Since for small values of  $y$  the Brillouin function can be developed as:

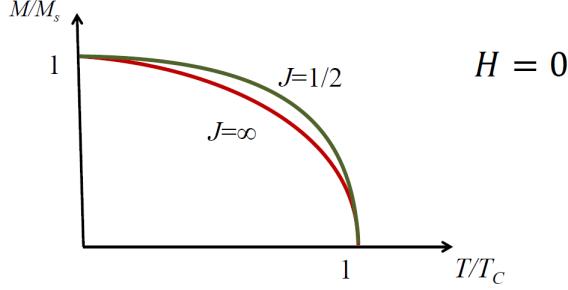
$$\frac{M}{M_s} = B_J(y) \approx \frac{1}{3} \frac{J+1}{J} y$$

Also, at  $B = 0$  we have:

$$\frac{M}{M_s} = \frac{k_B T}{M_s g_J \mu_B J \lambda} y$$

---

<sup>2</sup>Actually, this was the starting point of Weiss when he formulated his model. The previous part, related to the Hamiltonian and the exchange interaction, arises from considering the Heisenberg model too.



**Figure 3.3:** Magnetization in function of the temperature for the zero-field case.

Thus at  $T = T_C$  we get:

$$\frac{1}{3} \frac{J+1}{J} y = \frac{k_B T_C}{M_s g_J \mu_B J \lambda} y$$

from which we can finally obtain the expression of the **Curie temperature**:

$$T_C = \frac{g_J \mu_B \lambda (J+1)}{3k_B} M_s \quad (3.14)$$

Starting from the expression of the saturation magnetization  $M_s = n g_J \mu_B J$  the expression of  $T_C$  can be simplified by defining an effective magnetic momentum  $\mu_{eff}$ :

$$\mu_{eff} = g_J \mu_B \sqrt{J(J+1)} \quad (3.15)$$

Thus we obtain:

$$T_C = \frac{g \mu_B \lambda (J+1)}{3k_B} M_s = \frac{n \lambda \mu_{eff}^2}{3k_B} \quad (3.16)$$

In the Weiss model the link between the molecular field and the magnetization of the material is represented by the constant  $\lambda$ . This constant contains the "quantum information" related to the exchange interaction acting between dipoles. In fact, as we already told, ferromagnetism can not be explained in terms of dipolar interaction. Also, Weiss tried to estimate the magnitude of this interaction by estimating the molecular field. To do that we can approximate  $M$  with  $M_s$  ( $M \approx M_s$ ) and for  $J = \frac{1}{2}$  and  $T_C = 1000K$  we get a molecular field of about  $1500T$ . Such an impossible value means that the phenomenon can not be explained with classical arguments and requires a quantum description. It is worth analysing quickly the behaviour of the magnetization as a function of temperature. Fig.3.3 shows this very behaviour. We can notice that for  $T = T_C$  the curve has a discontinuous derivative meaning that the transition from  $T < T_C$  to  $T > T_C$  (ferro-para transition) is a phase transition of the second order. We will come back on this matter in the following.

Let's now study the case in which the field is different from zero but small. If  $\mathbf{B}$  is small, at  $T > T_C$  we can still use the approximation of the Brillouin function seen before. In this case, we can also find the expression of the susceptibility  $\chi$ . In fact:

$$\begin{aligned} \frac{M}{M_s} = B_J(y) &\approx \frac{g_J \mu_B (J+1)(B + \lambda M)}{3k_B T} = \frac{T_C}{\lambda M_s} \left( \frac{B + \lambda M}{T} \right) \Rightarrow \\ &\Rightarrow \frac{M}{M_s} \left( 1 - \frac{T_C}{T} \right) = \frac{T_C B}{T \lambda M_s} \Rightarrow M = \frac{T_C}{\lambda(T - T_C)} B \end{aligned}$$

Consequently, we can obtain the expression for the susceptibility:

$$\chi = \lim_{B \rightarrow 0} \frac{\mu_0 M}{B} = \frac{\mu_0 T_C}{\lambda(T - T_C)} = \frac{n \mu_0 \mu_{eff}^2}{k_B(T - T_C)} \quad (3.17)$$

This is a very important result known as **Curie-Weiss Law** which can be expressed as:

$$\chi \propto \frac{1}{T - T_C} \quad (3.18)$$

valid only for  $T > T_C$  when the material is not ferromagnetic anymore but paramagnetic. In general, also for  $\mathbf{B} \neq 0$  the solution of the system in Eq.(3.13) can be only found graphically. The graph is really similar to the previous one but, in this case, the set of lines is shifted in the positive direction of the horizontal axis. In this case, a non-zero magnetization is present now at any temperature  $T$ . Clearly, the magnetization in this case is not spontaneous anymore because of the presence of the external field.

### 3.3.1 Exchange Interaction and Molecular Field

The link between the exchange interaction  $J_{ij}$  and the molecular field  $B_{mf}$  can be found by using the Heisenberg model. Neglecting the sign and using scalars, and assuming that we are dealing with identical spin momenta  $S_j$  all oriented in the same direction (saturation), we can rewrite:

$$B_{mf} = \lambda M = \frac{2}{g\mu_B} \sum_{j \neq i} J_{ij} S_j = 2 \frac{z J_{ij} S_j}{g\mu_B}$$

$$M = M_s = n g \mu_B S_j$$

where the last equivalence of the first equation comes from Eq.(3.10). Here,  $z$  is the number of nearest neighbours with identical exchange  $J_{ij}$  and  $n$  is the number of dipoles per unit volume. By substituting the expression of  $M$  in the first equation we get:

$$\lambda = \frac{2z}{ng^2\mu_B^2} J_{ij} \quad (3.19)$$

As expected,  $\lambda$  is strongly related to the exchange interaction (direct proportionality). We have assumed so far that  $L = 0$  and  $J = S$ , so that the magnetization is provided purely by spin and that the exchange interaction is acting on the full magnetic moment  $J$ . This is a valid approximation for 3dTMs, but not for lanthanides ( $4f$  elements), where  $L$  can be as large as  $S$ . In order to unify things we can try to write the exchange interaction as a function of  $J$  instead of  $S$ . In such a case, we have to use a reduced value of  $J$ , corresponding to the actual contribution coming from  $S$ . This can be done considering that  $J$  is always a good quantum number for any atoms while  $S$  is not. Therefore, only the component of  $S$  parallel to  $J$  is conserved while the perpendicular one averages zero. Thus we get:

$$\hat{\mu} = g_J \mu_B \hat{J} = \mu_B \left( g_L \hat{L} + g_S \hat{S} \right) \quad (3.20)$$

It can be proofed that the component of  $S$  parallel to  $J$  is  $(g_J - 1)J$  then in the exchange interaction we can replace  $S_i \cdot S_j$  with  $(g_J - 1)^2 J_i \cdot J_j$ . By repeating all the calculations with these corrections we get:

$$\lambda = \frac{2z(g_J - 1)^2 J_{ij}}{ng_J^2\mu_B^2} \quad (3.21)$$

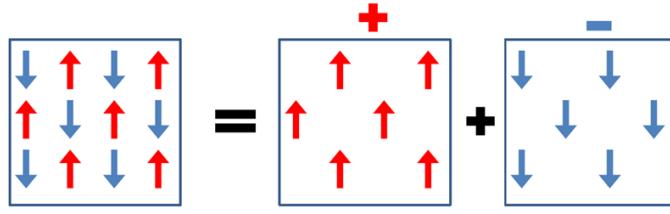
$$T_C = \frac{2z(g_J - 1)^2 J_{ij}}{3k_B} J(J+1) \quad (3.22)$$

## 3.4 Weiss model for Antiferromagnetism

AFM materials are the macroscopic demonstration of what happens when we have a negative exchange interaction. In an AFM material, each atomic dipole is opposite to the ones of the nearest neighbours resulting then in a null magnetization. This phenomenon was understood and introduced for the first time by the French physician Louis Eugène Félix Néel. As depicted in Fig.3.4 an AFM lattice can be seen as the superposition of two ferromagnetic sublattices, each with opposite dipole orientations. There will be a molecular field associated with each sublattice:

$$B_{mf}^+ = -|\lambda| M^- \quad B_{mf}^- = -|\lambda| M^+ \quad (3.23)$$

Here we notice that the molecular field related to the "positive" sublattice is associated with the magnetization of the "negative" one and vice versa. This is due to the fact that the mean-field felt by each sublattice is due to the dipoles of the other sublattice since the dipoles are displaced in an alternated way. Moreover, the minus sign derives from the fact that in both cases the molecular



**Figure 3.4:** Schematic representation of an AFM material.

magnetic field is opposite to the magnetization of the other sublattice. The mathematical solution is equal to the one of the FM material, but one for each sublattice:

$$\frac{M^\pm}{M_s} = B_J(y) \quad (3.24)$$

where in this case  $y = \left(-\frac{gJ\mu_B J}{k_B T} |\lambda| M^\mp\right)$ . Since the two sublattices must be identical, apart from the direction of the moments, we have that  $M^+ = M^- = M$  and then the solution remains almost the same ( $\frac{M}{M_s} = B_J(y)$ ) with  $y$  defined as:

$$y = \left(-\frac{gJ\mu_B J}{k_B T} |\lambda| M\right) \quad (3.25)$$

Also in this kind of material a critical temperature exists and it is called **Néel temperature**  $T_N$  in full analogy with the Curie temperature. It is the temperature below which the material is AFM and above which the material is PM. The expression is identical to the Curie temperature one:

$$T_N = \frac{n|\lambda|\mu_{eff}^2}{3k_B} = \frac{g\mu_B|\lambda|(J+1)}{3k_B} M_s \quad (3.26)$$

The two expressions in Eq.(3.24) and Eq.(3.26) are formally identical to the expressions of the ferromagnetism case. This is because the microscopic behaviour of AFM materials is similar to the one of ferromagnets. Despite the formal equivalence, the main difference is that the macroscopic magnetization here is equal to zero:

$$M = M^+ + M^- = 0 \quad (3.27)$$

However, we can take into account the order present in these materials by using a special order parameter called **staggered magnetization**  $M_{st}$  defined as:

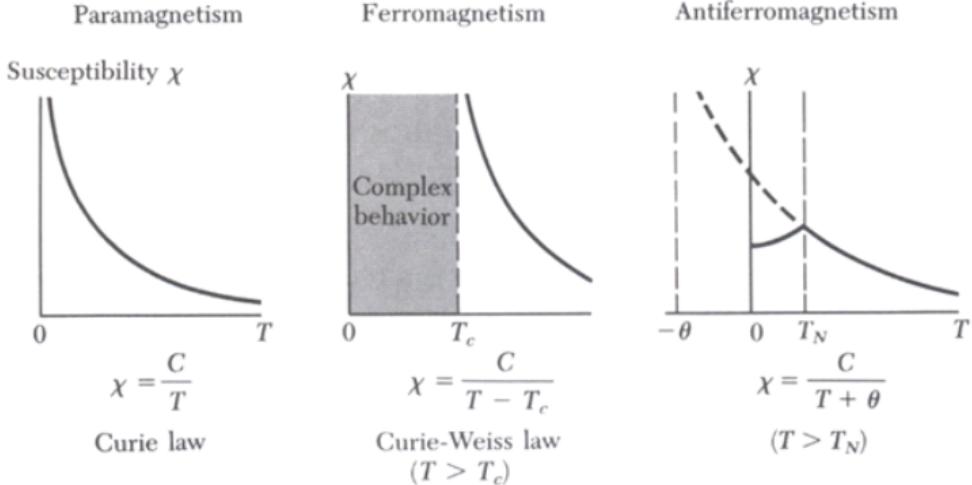
$$M_{st} = M^+ - M^- \quad (3.28)$$

In analogy with the FM case, we can study the behaviour of the magnetization in the case of a small field  $\mathbf{B}$ . For the negative sublattice the magnetization  $M^-$ , opposite to the magnetic field, is:

$$\begin{aligned} \frac{M^-}{M_s} &= B_J(y) \simeq \frac{T_N}{|\lambda|M_s} \left( \frac{B - |\lambda|M^+}{T} \right) \Rightarrow \\ &\Rightarrow \frac{M^-}{M_s} \left( 1 + \frac{T_N}{T} \right) = \frac{T_N B}{T |\lambda| M_s} \Rightarrow M = \frac{T_N}{|\lambda|(T + T_N)} B \end{aligned}$$

What happens is that, if there is a magnetic field, both sublattices want to align with the external field and, at the same time, want to stay antiparallel one to other. Globally, the result is a compromise between these two orientation trends which brings the system to have a small, but non-zero, magnetization. Therefore, even AFM materials when exposed to an external field exhibit weak magnetic properties. For this reason, we can compute the magnetic susceptibility for an antiferromagnet when  $T > T_N$  that is:

$$\chi = \lim_{B \rightarrow 0} \frac{\mu_0 M}{B} = \frac{\mu_0 T}{|\lambda|(T + T_N)} \quad (3.29)$$



**Figure 3.5:** Susceptibility as a function of the temperature for PM, FM and AFM.

The susceptibility respects the Curie-Weiss law also in this case but with a special critical temperature  $\theta = -T_N$ :

$$\chi \propto \frac{1}{T - \theta} = \frac{1}{T + T_N}$$

We can summarize the results obtained so far by looking at the graphs in Fig.3.5 in which the susceptibility behaviour as a function of the temperature is shown for paramagnetism, ferromagnetism and antiferromagnetism respectively. In particular, if we write the Curie-Weiss law using the variable  $\theta$ :

$$\chi \propto \frac{1}{T - \theta}$$

we can generalize this law to all the cases by defining:

- Paramagnetism:  $\theta = 0$
- Ferromagnetism:  $\theta = T_C > 0$
- Antiferromagnetism:  $\theta = -T_N < 0$

Moreover, it is worth noting that, in the AFM case the divergence behaviour for  $\chi$  happens for  $T = -T_N$  that has no physical meaning (negative temperature). Around  $T = T_N$  nothing special happens to the susceptibility that can be considered almost the same in the neighbourhood of  $T_N$  that is  $T_N \pm dT$ . It is also interesting to look at the trends of the inverse of  $\chi$  as a function of the temperature:

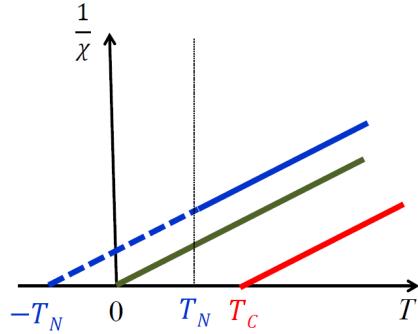
$$\frac{1}{\chi} \propto (T - \theta)$$

This is shown in Fig.3.6. We can generalize the result obtained for the different materials by writing the susceptibility as:

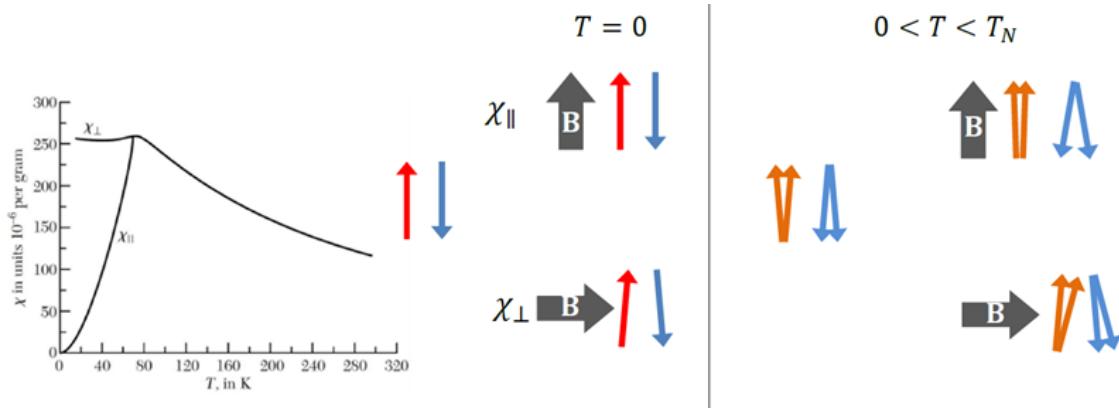
$$\chi = \frac{n\mu_0\mu_{eff}^2}{k_B(T - \theta)} \quad (3.30)$$

### 3.4.1 AFM, Field Direction and Spin Flop

Let's now study what happens microscopically to the dipoles of an AFM material when it is exposed to a magnetic field starting from the case of a weak field  $\mathbf{B}$ . Let's do this by referring to Fig.3.7. If  $T = 0$  and the field is parallel to the direction of the dipoles no torque acts on the dipoles and there is no reaction of the system. The magnetization is zero. If  $T = 0$  but the field is perpendicular to the dipoles, a torque will act on them trying to align them to the field and they will oppose resistance since they want to stay antiparallel. The balance between these two effects results in a small magnetization of the system. When  $0 < T < T_N$  the dipoles start oscillating



**Figure 3.6:** Inverse of the susceptibility as a function of the temperature for PM, FM and AFM.

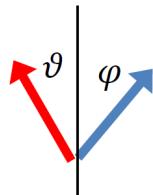


**Figure 3.7:** AFM dipoles alignment in weak field.

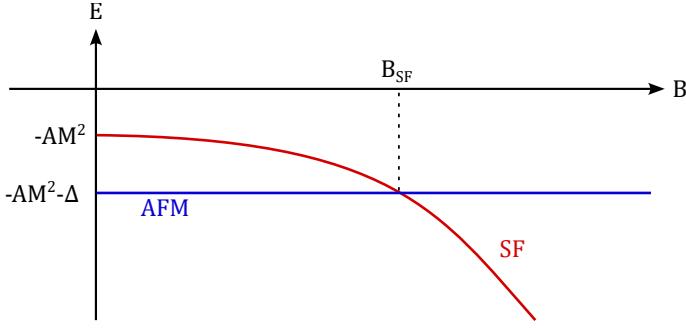
(the oscillation is represented by the two arrows in the figure). Now, when the field is parallel to the dipoles, since they oscillate there will be a small component of the magnetic momentum perpendicular to the field resulting, then, in a small torque different from zero. In particular, the antiparallel dipoles will increase the amplitude of their oscillations since the torque tends to rotate them while the parallel dipoles will decrease the oscillation amplitude since they already are in the "right" direction. Therefore, the system, even with a parallel field, has a small magnetization. For a perpendicular field, the situation is close to the one at  $T = 0$  but unlike before the dipoles oscillate. Nevertheless, the result is the same: a small magnetization of the system. The overall result is shown on the left side of Fig.3.7. In conclusion, we can say that for  $0 < T < T_N$ , regardless of the direction of the field, we obtain a small magnetization of the system.

What happens in the case of a strong magnetic field is definitely more interesting. At  $T \neq 0$ , the field manages to rotate significantly the dipoles bringing them into a configuration which results in an angle  $\theta$  and an angle  $\varphi$  of the two dipoles with respect to the direction of the field as shown in Fig.3.8. In particular, if the two angles are equal  $\theta = \varphi$  the configuration is called **Spin Flop**. In general, we can calculate the magnetic energy of the system that is:

$$E = -MB \cos \theta - MB \cos \varphi + AM^2 \cos(\theta + \varphi) \quad (3.31)$$



**Figure 3.8:** Dipoles orientation angles with respect to a vertical magnetic field.



**Figure 3.9:** Spin Flop and AFM energies as functions of the magnetic field.

where  $A$  is the exchange parameter. The last term is the one that minimizes the energy in the AFM standard configuration in which  $\theta = 0, \pi$  and  $\varphi = \pi, 0$ . This form of energy is not complete because it doesn't take into account the fact that, in real materials, there is a preferential direction for the magnetization. This happens because real materials are not completely isotropic and therefore there is always a little anisotropy that makes one direction preferred among all the possible ones. We need to include in the energy an anisotropy term that is:

$$-\Delta (\cos^2 \theta + \cos^2 \varphi)$$

In the AFM standard configuration, the energy is minimized, its value is  $E_{AFM} = -AM^2 - 2\Delta$  and it does not depend on the magnetic field  $B$ . In the Spin Flop configuration, the energy is:

$$E = -2MB \cos \theta + AM^2 \cos 2\theta - 2\Delta \cos^2 \theta$$

We can minimize the energy with respect to  $\theta$  and find the so-called Spin Flop angle:

$$\theta_{SF} = \theta_{min} = \arccos \frac{MB}{2(AM^2 - \Delta)} \quad (3.32)$$

Then we can easily find the minimum value of the energy for the Spin Flop configuration and compare it with the AFM configuration one:

$$E_{SF} = -AM^2 - \frac{M^2}{2(AM^2 - \Delta)} B^2 \quad (3.33)$$

$$E_{AFM} = -AM^2 - 2\Delta \quad (3.34)$$

In particular, we can see that there exists a particular value of the field  $B_{SF}$  for which the Spin Flop configuration becomes energetically favourable with respect to the AFM one. This threshold value of the field can be computed by imposing  $E_{SF} = E_{AFM}$  and it is:

$$B_{SF} = 2 \frac{\Delta}{M} \sqrt{\left( \frac{AM^2}{\Delta} - 1 \right)} \quad (3.35)$$

The situation is depicted in Fig.3.9. For  $B > B_{SF}$  the dipoles will arrange to get into the Spin Flop configuration since it is energetically favourable. Experimentally, this configuration change can be seen by studying the magnetization of the material that spikes when the configuration passes from the AFM to the Spin Flop one.

### 3.4.2 Helical Order

The Helical Order is another dipole configuration that allows to obtain an AFM order. In particular, if the considered material has a layered structure, the planes can have an FM internal order, but interact more weakly with the other planes. Then the 3D structure might not be FM but AFM. Also in this case is appropriate to write the expression of the energy for such a structure. Using  $J_1$  and  $J_2$  for the exchange interaction between a given plane and its first and second neighbouring planes, and being  $\theta$  the angle between the magnetizations of successive planes, we get:

$$E = -2NS^2(J_1 \cos \theta + J_2 \cos 2\theta)$$

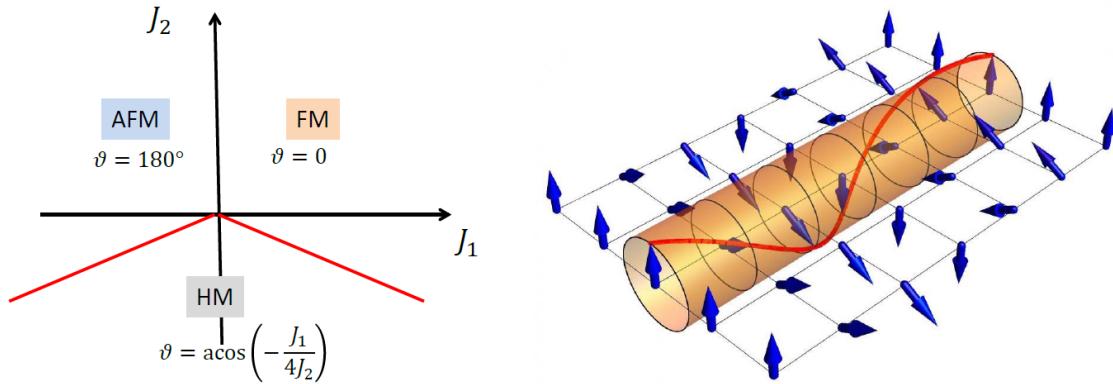
where  $N$  is the number of atoms in each plane and  $2\theta$  is the angle between the magnetization of the given plane and the one of the second neighbouring plane. By minimizing the energy we can find three different solutions:

$$\theta = 0; \quad \theta = \pi; \quad \theta = \arccos\left(-\frac{J_1}{4J_2}\right);$$

The particular configuration in which  $\theta = \arccos\left(-\frac{J_1}{4J_2}\right)$  is called Helical Magnetic order (HM). It can be obtained only if  $J_1 < 4|J_2|$  and  $J_2 < 0$ . In the HM order, the dipoles are not parallel and they displace in a helicoidal way as shown in Fig.3.10. In general, we have:

- FM:  $\theta = 0 \implies E_{FM} = -2NS^2(J_1 + J_2)$
- AFM:  $\theta = \pi \implies E_{AFM} = -2NS^2(-J_1 + J_2)$
- HM:  $\theta = \arccos\left(-\frac{J_1}{4J_2}\right) \implies E_{HM} = +2NS^2 \frac{\frac{J_1^2}{8} + J_2^2}{J_2}$

The first two solutions are obtained if  $J_2 > 0$  and if  $J_1 > 0$  then we have the FM order while if  $J_1 < 0$  we have the AFM one. These results are summarized on the left side of Fig.3.10.



**Figure 3.10:** Magnetic behaviour depending on  $J_1$  and  $J_2$  and Helical Magnetic order.

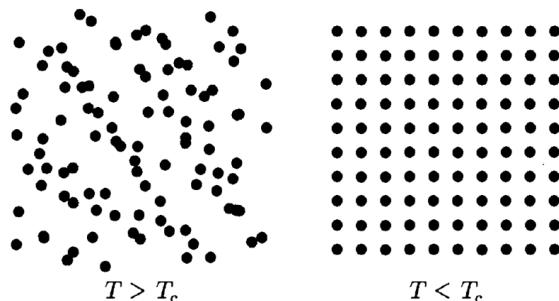
## Chapter 4

# Order and Broken Symmetry

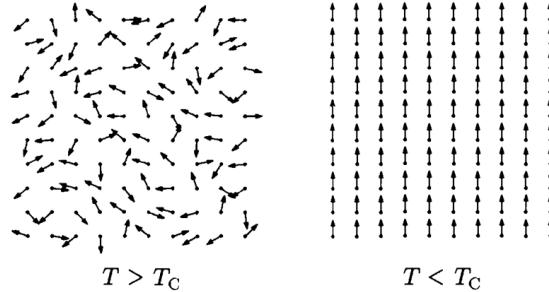
In this chapter, we will try to introduce some concepts regarding order and symmetry in condensed matter. Being these concepts deep and not easy to understand we will treat them in a non-formal way, trying to explain the physical picture related to them. A more complete and clear description of these phenomena can be found in "*Magnetism in Condensed Matter*" by Stephen Blundell.

### 4.1 Broken Symmetry

The first question that we are going to answer is: "What is order with respect to symmetry?". We do have an intuitive answer to this question. We can think that the higher the order of a system, the higher its symmetry but, unfortunately, this is not the right answer. Let's consider, for instance, the two phases of a substance: solid and liquid (*e.g.* a block of ice and a glass of water). As the liquid cools down below its critical temperature  $T_C$  the order in which the atoms are displaced increases and, at the same time, its symmetry decreases. This may seem surprising because the picture of the solid "looks" more symmetrical than the one of the liquid but, once again, the intuitive answer is not the right one. The atoms in the solid are all symmetrically lined up while in the liquid they are randomly displaced, as we can see in Fig.4.1. This is the key observation. The fact that the atoms in a liquid are randomly displaced makes any point of the liquid exactly the same as any other. If you average the system over time, each position is visited by atoms as often as any other. In short, the liquid system possesses complete translational and rotational symmetry. This does not happen for the solid. In the solid case, in fact, some residual symmetry is still present (discrete translational symmetry by the lattice vector  $\mathbf{R}$  and discrete rotational symmetry) but the high symmetry of the liquid state has been broken. In particular, during the transition from the liquid phase to the solid one, which happens at  $T = T_C$ , the symmetry of the system is broken. Many similar features are observed in condensed matter systems. As in this case, the parameter which drives the symmetry-breaking transition is often the temperature. This is the case, for example, also for the transition between a ferromagnetic and a paramagnetic state which happens at the Curie temperature  $T_C$ . For ferromagnet at  $T > T_C$ , all the directions are equivalent and each magnetic moment can point in any direction. This means that the system is characterized by a low order but a high symmetry. On the other end, at  $T < T_C$  the system *chooses* a unique direction for all the spins. The high symmetry state is broken while the order is

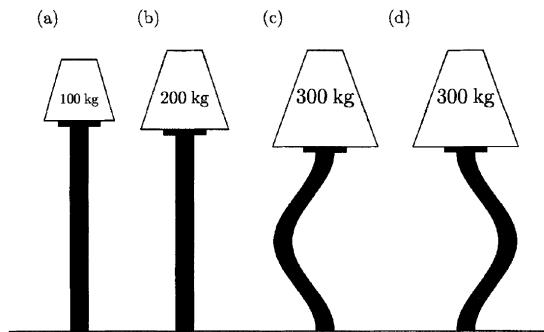


**Figure 4.1:** Schematic representation of solid and liquid phases.

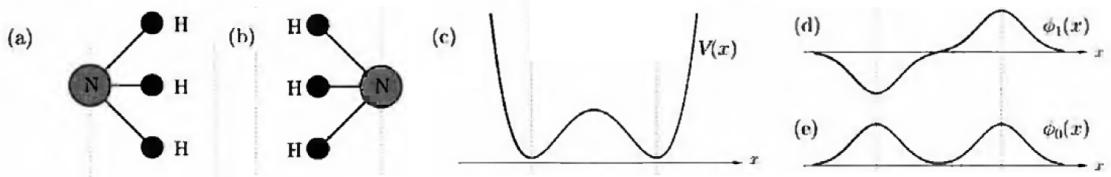


**Figure 4.2:** Schematic representation of FM and PM phases.

increased. We said that the system *chooses* a unique direction for all the spins. This is actually a rather surprising effect since the energy of the "up" configuration is completely identical to the "down" one. The theory behind this phenomenon can not explain why the system chooses either the "up" configuration or the "down" one. Indeed, equations-wise these two choices are completely the same. The key element to consider in this case is the anisotropy of the material that will make one configuration favourable with respect to the other or, for example, a random thermal fluctuation that will favour one of the two configurations. In general, the solutions of a symmetric problem do not have the same symmetry as the very problem they solve. This phenomenon is called **spontaneous symmetry breaking**. Another example of this is represented by the so-called Euler strut represented in Fig.4.3. It consists of a pillar on the top of which a weight is placed. While the weight increases the pillar will tend to bend in one direction or the other. For the problem itself, if the weight is perfectly balanced, the pillar is completely isotropic and well placed, any possible direction for the bending is absolutely the same. Once again, as in the magnetization case, there is no favourable direction. Nevertheless, the pillar will choose one direction among all the possible ones and the symmetry of the problem will be broken. As often happens in physics, the reality is more complicated than equations and the bending direction will not be randomly chosen but it will depend on some random thermal fluctuations or on a defect on the pillar itself. But then, how can we understand which configuration will be taken by the system among all the possible ones? To try to have an idea of this, let's consider another example: the ammonia molecule  $NH_3$ . This tetrahedral molecule exists in two different configurations depending on the plane in which the nitrogen atom is placed as shown in Fig.4.4(a) and (b). The energy of these two configurations is the same and, being  $x$  the position of the nitrogen atom, we can imagine that the potential energy curve  $V(x)$  has two minima associated with the two stable configurations. The two lowest energy eigenstates associated with this potential  $V(x)$  are shown in Fig.4.4(d) and (e) for the first excited state  $\phi_1(x)$  and the ground state  $\phi_0(x)$ . Therefore, since we have two possible eigenstates, the configurations represented in Fig.4.4(a) and (b) are not stable but metastable states. In particular, they can be described as a superposition of the two eigenstates



**Figure 4.3:** The Euler strut with different loads.



**Figure 4.4:** (a) (b) The two possible configurations for the ammonia molecule. (c) The potential energy which describes the ammonia. (d) (e) The wave functions of the first excited state  $\phi_1(x)$  and of the ground state  $\phi_0(x)$ .

of the potential  $V(x)$ . The configuration (a) and (b) are given by:

$$\begin{aligned}\phi_a(x) &= \frac{1}{\sqrt{2}} (\phi_0(x) - \phi_1(x)) \\ \phi_b(x) &= \frac{1}{\sqrt{2}} (\phi_0(x) + \phi_1(x))\end{aligned}$$

This means that the ammonia molecule, for example, prepared in the metastable state (a), will start to oscillate, *i.e.* tunnelling, between the state (a) and the (b) one, with a frequency that is given by the energy difference between  $\phi_1(x)$  and  $\phi_0(x)$ . This frequency for the ammonia molecule is  $2.4\text{GHz}$ . We have just seen that, in this case, the system does not choose any of the two possible configurations but oscillates between them. Theoretically, this tunnelling between configurations can also happen for a ferromagnet state, in which the direction of the spin would oscillate between "up" and "down". This phenomenon can happen for a very small amount of atoms ( $\simeq 10$ ) but it is extremely unlikely in reality. In real materials, in fact, we have a huge amount of atoms ( $\simeq 10^{23}$ ) and this tunnelling effect does not realize. It is actually true, also for the ferromagnetic case, that the ground state of the system with all the spins pointing in a single direction is not a true stationary state but a metastable one. Nevertheless, the lifetime of this metastable state is greater than the age of the universe and the state can be considered a full-fledged stable state. In fact, to access the other state, the one with all the spins reversed, one would have to flip every spin in the system simultaneously, which is extremely unlikely. To further and better discuss the idea of order and symmetry, let's now discuss and introduce some models related to symmetry and phase transitions.

## 4.2 Landau Model of Phase Transitions

This model was invented by the Russian physicist Lev Landau. By using some very general considerations in the scheme of Landau's model, it is possible, for example, to show how the transition from a ferromagnetic to a paramagnetic state can be seen as a phase transition. Landau used a thermodynamic approach, choosing as an order parameter the magnetization  $M$  of a ferromagnet. In particular, he decided to write the Helmholtz free energy  $F$  as a power series of  $M$ . The series must contain only even powers of  $M$  since, as told, there is no preferential direction in the magnetization, *i.e.* the energy for  $\pm M$  has to be the same. Thus, by considering a temperature dependence on the coefficients, we have:

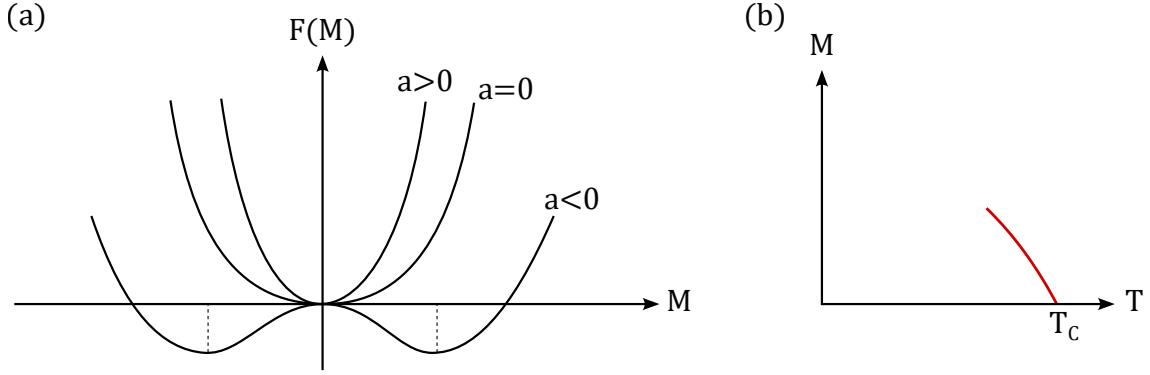
$$F(M, T) = F_0 + a(T)M^2 + b(T)M^4$$

where we decided to stop at the fourth power of  $M$  and  $F_0$  is a constant. Since we want the energy  $F$  to have two stable configurations at  $T < T_C$  and we want the magnetization to drastically decrease when  $T > T_C$  we need to have:

$$a : \begin{cases} > 0 & \text{if } T > T_C \\ < 0 & \text{if } T < T_C \end{cases}$$

Moreover, since we want  $\lim_{M \rightarrow \infty} F = +\infty$  we need the  $b$  term to be constant and positive:  $b(T) = b > 0$ . The behaviour based on the sign of  $a$  can be understood by looking at the shape of the quadratic function  $F = F(M)$  represented in Fig.4.5(a). A suitable form for  $a$  is then:

$$a(T) = a_0(T - T_C)$$



**Figure 4.5:** (a) Free energy as a function of the magnetization and (b) magnetization as a function of the temperature for  $T \approx T_C$ .

where  $a_0$  is a constant term. Then the energy becomes:

$$F(M, T) = F_0 + a_0(T - T_C)M^2 + bM^4$$

To find the minimum solution we need to derive with respect to  $M$  and solve the equation:

$$\frac{\partial F}{\partial M} = 2a_0(T - T_C)M + 4bM^3 = 0 \quad (4.1)$$

This equation has two solutions. The first trivial solution is  $M = 0$  and it only applies when  $T > T_C$ . The second more interesting solution, only valid for  $T < T_C$ , is:

$$M = \pm \sqrt{\frac{a_0}{2b}(T_C - T)} \quad (4.2)$$

Clearly, we need  $T < T_C$  for the square root to be positive. Thus the magnetization follows the curve in Fig.4.5(b). In particular, one can notice that the magnetization has a discontinuous derivative at  $T = T_C$ . Based on Landau's Theory of Phase Transition, this corresponds to a **second-order phase transition**. Landau's model can be used to study not only the magnetization of a ferromagnet but a lot of phenomena like, for example, superconductivity as we will see in Sec.12.3.

### 4.3 Heisenberg and Ising Models

We have already met the Heisenberg model which is a microscopic approach to try to explain the magnetic behaviour of solids. This model uses a Hamiltonian which considers the interactions among all the spins in a given solid. If we only consider the interactions between the nearest neighbour atoms this Hamiltonian is of the form:

$$H = - \sum_{i \neq j} JS_i \cdot S_j$$

where  $J$  is the constant exchange integral. The number of nearest neighbours (NN) that we need to take into account depends on the dimensionality of the lattice that we are considering. For example, in the case of a 1D lattice, we will have NN=2 and for a 2D square lattice NN=4. In the more realistic case of a 3D lattice, NN depends a lot on the type of lattice that we are considering (*e.g.* cubic, hexagonal, etc.). Although in our reality only 3 spatial dimensions exist, the Heisenberg model can be generalized to a  $n$  dimensional case by distinguishing between the dimensionality  $d$  of the lattice and the dimensionality  $D$  of the spins<sup>1</sup>. In the Heisenberg model for ferromagnetism,  $D = 3$  because the spin is a three-dimensional vector, *i.e.* it can take all the possible values in the space, while, as told,  $d = 1, 2, 3$  depending on the lattice. The Ising problem is a subset of the Heisenberg problem in which the spins are only allowed to point up or down (only the component along the z-axis is considered) and therefore  $D = 1$ . The Ising problem can be treated for different dimensionality  $d$  of the space.

<sup>1</sup>In general  $D$  is known as the dimensionality of the order parameter

### 4.3.1 The 1D Ising Model

Let's consider a one-dimensional chain of  $N$  sites with a positive exchange interaction  $J > 0$  with periodic boundary conditions, *i.e.* the chain is bent on itself to create a circle. The energy of such a system is:

$$\hat{\mathcal{H}} = -2J \sum_{i=1}^N S_i S_{i+1} \quad (4.3)$$

where 2 is the number of the nearest neighbours and  $J > 0$  so that the ground state (GS) corresponds to all the spins aligned upwards like in a ferromagnet as depicted in the upper part of Fig.4.6. Since the spins can take only two values  $\pm 1/2$ , the energy of the ground state is:

$$E_{GS} = -2N \frac{1}{2} \frac{1}{2} J = -\frac{NJ}{2}$$

Now, consider adding one "mistake" along the chain, a single defect as depicted in the lower part of Fig.4.6. The energy cost of this mistake is:

$$\Delta E = E_{GS} - E_{1def} = J \quad (4.4)$$

since we have to turn one favourable interaction (energy saving  $J/2$ ) into an unfavourable one (energy cost  $J/2$ ) resulting in an overall cost of  $J$ . However, there is an entropy gain equal to:

$$\Delta S = k_B \ln N$$

because we can put the defect in any one of the  $N$  sites along the chain, *i.e.* the possible microstate is equal to the number of places in which we can place the defect that is  $N$ . So the overall variation in the Helmholtz free energy is:

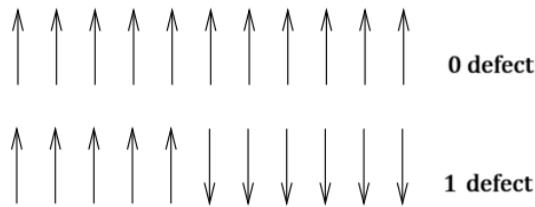
$$\Delta F = \Delta E - T\Delta S = J - k_B T \ln N \xrightarrow[N \rightarrow \infty]{} -\infty$$

This means that the defect can spontaneously form and, as a result, it is impossible to have an ordered chain. In other words, this also means that the critical temperature for a 1D chain of spins is zero. This tells us that the Ising model does not allow the creation of a spontaneous order and, for example, it can not explain the ferromagnetic spontaneous magnetization. The Ising model can be formalised in any dimension  $d$  but so far only the problem with a dimensionality smaller or equal to two has been solved by mathematicians. In particular, the solution of the 2D Ising model was one of the outstanding achievements of twentieth-century statistical physics and was solved by Lars Onsager in 1944.

Before proceeding, let's summarize the different dimensionality and the sensibility to this parameter of the Weiss, Ising and Heisenberg models. By considering the critical exponents defined by the relations:

$$\begin{aligned} M &\propto (T_C - T)^\beta \\ \chi &\propto (T - T_C)^\gamma \\ M &\propto H^{\frac{1}{\delta}} \end{aligned}$$

We can summarize the results obtained so far in Tab.4.1.



**Figure 4.6:** 1D chain of spins aligned ferromagnetically with zero and one defect.

**Table 4.1:** Critical exponents for various models.

Model	Weiss	Ising 2D	Ising 3D	Heisenberg 3D
D	any	1	1	3
d	any	2	3	3
$\beta$	1/2	1/8	0.326	0.367
$\gamma$	1	7/4	1.24	1.39
$\delta$	3	15	4.78	4.78

## 4.4 Phase Transition and Nematic Order

Phase transitions are strongly related to the order of a system as we have previously seen. From a graphical point of view phase transitions can be described thanks to phase diagrams. Fig.4.7, for instance, depicts the phase diagram for water. In general, not all phase transitions involve a change of symmetry. If, for example, the boundary line between the liquid and gas regions is terminated by a critical point, it is possible to avoid the sharp phase transition which leads to a change of symmetry. The sharp phase transition is the one which crosses the boundary line between two phases (path B in Fig.4.7). It can be avoided by taking the path through the phase diagram which turns around the critical point (path C in Fig.4.7). When this happens there is no change in symmetry. For water, for example, the transition between gas and liquid involves no change of symmetry because of this "alternative" path that the transition can follow. In contrast, the solid-liquid transition involves a change of symmetry and, consequently, there is no critical point for the melting curve. Traditionally, the way of labelling phase transition was theorized by Ehrenfest and is related to thermodynamics quantities. In particular, we can distinguish two orders of phase transition based on the derivative of the Helmholtz free energy with respect to temperature. As depicted in Fig.4.8 we have:

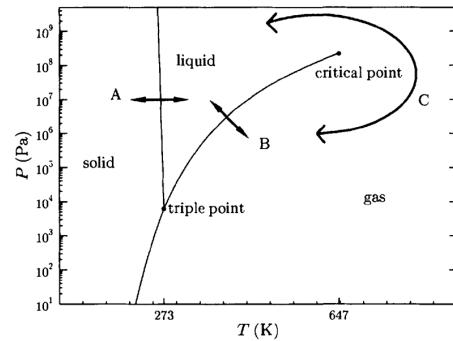
- I Order:  $\frac{\partial F}{\partial T}$  is discontinuous.
- II Order:  $\frac{\partial^2 F}{\partial T^2}$  is discontinuous.

where  $\frac{\partial F}{\partial T}$  is the specific heat of a material. Phase transitions can also be classified based on the latent heat  $Q_L$  and on the susceptibility  $\chi$ . In particular, we have:

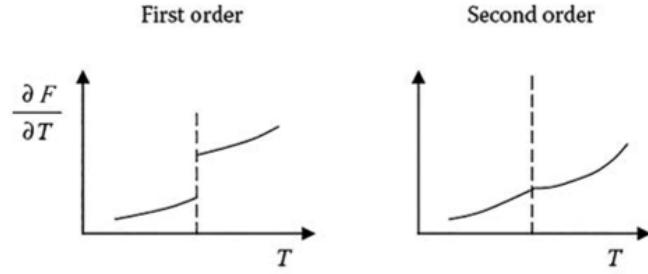
- I Order:  $Q_L \neq 0$
- II Order:  $Q_L = 0$  and  $\chi$  diverges at  $T = T_C$

## 4.5 Nematic Order

A nematic order is a type of order which does not depend on the orientation but only on the direction, *i.e.* the orientation is described only by a line, not by an arrow. It is often used to describe some phases of liquid crystal compounds. Let's consider, for example, a measurement



**Figure 4.7:** The phase diagram of water.



**Figure 4.8:** Ehrenfest criterion for phase transitions.

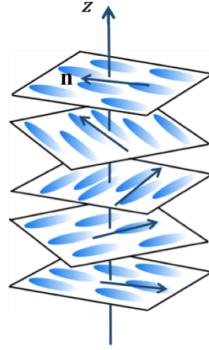
axis  $z$  and angle  $\theta$  between the  $z$ -axis and the orientation of the nematic system. The order parameter for such a system can not depend on, for instance,  $\cos \theta$  since otherwise it would change sign for the opposite value of  $\theta$  and it would depend on the direction. Hence,  $\cos^2 \theta$  looks already a better order parameter. Nevertheless, if we consider a group of molecules all displaced on a plane perpendicular to the measurement axis, as depicted in Fig.4.9, we would have  $\theta = 90^\circ$  and thus a null order parameter despite the high order of the system. Therefore we need to define a new order parameter  $S$  which has to be a function of  $\cos^2 \theta$  and be equal to 1 for a perfectly ordered system and 0 in a total disordered case. Since:

$$\langle \cos^2 \theta \rangle = \frac{2\pi \int_0^\pi \cos^2 \theta \sin \theta d\theta}{4\pi} = \frac{1}{6} [-\cos^2 \theta]_0^\pi = \frac{1}{3}$$

we can define our order parameter as:

$$S = \left\langle \frac{3 \cos^2 \theta - 1}{2} \right\rangle \quad (4.5)$$

which is equal to 1 for  $\theta = 0, \pi$ , equal to  $-1/2$  if  $\theta = \frac{\pi}{2}$  and null for any random value of  $\theta$ .



**Figure 4.9:** Nematic order example.

# Chapter 5

## Real Materials

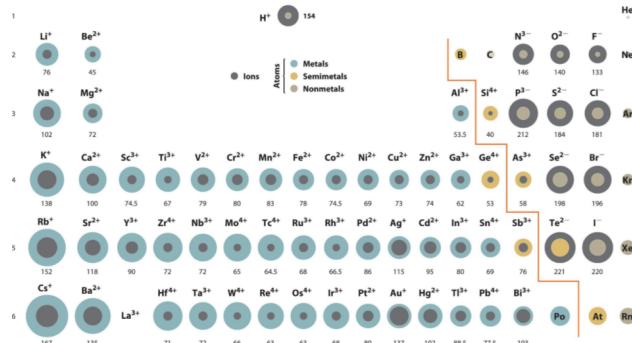
Magnetic materials are composed of a subset of the elements of the periodic table but they are mostly based on the 3d Transition Metals and Lanthanides. In general, the overall behaviour of a solid made of ions is different from the behaviour of the isolated ion. This means that single particle models (based on Bloch theorem and on quasi-free electron gas and bands) are usually not applicable. An atomic model is often better suited in which the electron-electron interactions on-site are accounted for by the Hartree-Fock-Slater atomic model, and the inter-site interactions are added perturbatively (*e.g.* Coulomb repulsion, exchange and hopping integrals). Magnetism then results from the Hamiltonian of the ions in a given lattice, with a given magnetic moment, interacting in a more or less complex way. Therefore, to study the magnetic behaviour of a real material we need to know the orbital and spin momenta of ions, the crystalline structure, the interactions present in the material and how these ingredients influence each other. In this chapter, we start with some common examples of structures of magnetic materials that will often be oxides of 3dTMs.

### 5.1 Crystalline Structure

In general, the size of the atoms which constitute a particular solid influences a lot the crystalline structure of the solid itself. We have three different ways to define the size of an atom based on its radius:

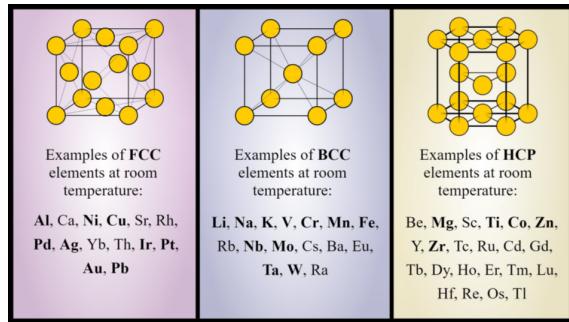
- Covalent radius: is half the distance of the covalent bond between two identical atoms' nuclei.
- Atomic radius: is the distance from a neutral atom's nucleus to its outermost valence shell.
- Ionic radius: is the distance between the nuclei of two identical atoms in an ionic solid, *i.e.* bonded by an ionic bond.

The last one is computed by considering ions as rigid spheres and is the most interesting one for our purposes since it is less subjected to variations along the periodic table with respect to the other two radii. The values of the ionic radius for all the elements of the periodic table are shown in Fig.5.1. As told, the atoms' size is one of the parameters which determines the crystalline structure



**Figure 5.1:** Ionic radius values.

of a solid. The structure, then, is really important for the general properties of the material like the magnetic behaviour, the specific heat and the conductivity. The most common structures, depicted in Fig.5.2, are three: the Faced Center Cubic (FCC), the Body Center Cubic (BCC) and the Hexagonal Closed Packed (HCP). The figure also reports the elements which can give rise to the

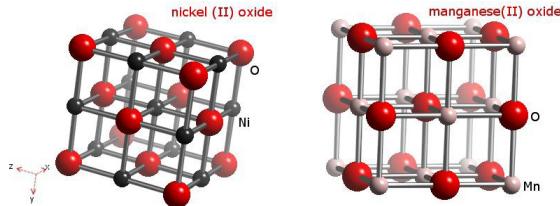


**Figure 5.2:** Common crystalline structures.

different structures. It is worth noting that despite *Fe*, *Co* and *Ni* are three consecutive elements in the periodic table they arrange themselves in three different crystalline structures: BCC, HCP and FCC respectively.

## 5.2 Examples

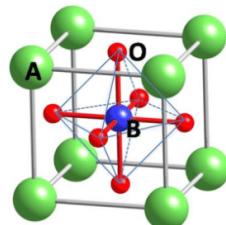
Let's start considering some relevant examples of real magnetic materials starting from simple oxides. The first one is the nickel oxide  $NiO$  made of  $Ni^{2+}$  ions ( $3d^8$ ) and  $O^{2-}$  ions ( $2p^6$ ). They are arranged in a FCC structure and exhibit an AFM behaviour. The manganese oxide  $MnO$  has the same structure and the same AFM behaviour and both the oxides are reported in Fig.5.3.



**Figure 5.3:** Simple oxides  $NiO$  and  $MnO$ .

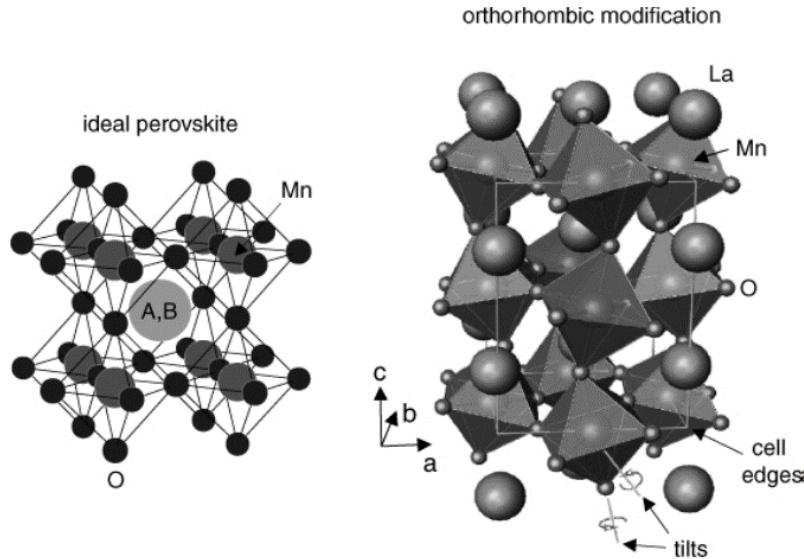
### 5.2.1 Perovskite

We saw in the previous case of the simple oxides that a 3dTM ion can be considered "caged" by six Oxygen atoms displaced in an octahedron structure which contains the metallic ion. Usually, a third atom can stabilize the octahedron and give rise to the so-called perovskite structure. This is the case, for example, of  $SrTiO_3$  which is made by  $Sr^{2+}$  (A site),  $Ti^{4+}$  (B site) and three  $O^{2-}$ . In general, the ideal perovskite structure is in the form reported in Fig.5.4 and can be indicated with the formula  $ABO_3$ . All the ions in the perovskite have an empty or completely filled external shell:



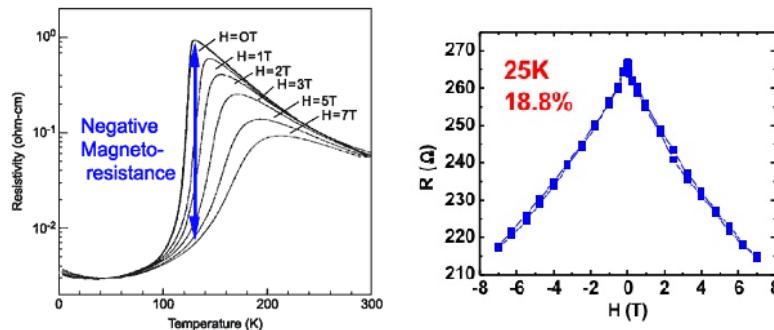
**Figure 5.4:** Pervoskite structure  $ABO_3$ .

the  $Sr^{2+}$  has a  $4d^0 5s^0$  configuration, the  $Ti^{4+}$  a  $3d^0$  and the  $O^{2-}$  is the completely filled one with a  $2p^6$  configuration. All electrons of the external shells of Strontium and Titanium are “stolen” by the Oxygen to complete the  $2p$  shell. Since all atoms have full shells we have an insulator material. Another example of perovskite structure is given by the  $LaAlO_3$ . So far we have seen the so-called ideal perovskite structure but, as we should have understood, reality is far more complicated. In real materials, the octahedral can be slightly rotated on the O-A axis or deformed as we will see in the next chapters. Depending on the size size of the metallic ion (site A) this rotation can change in magnitude. Fig.5.5 shows this modification (orthorhombic modification) in the case of  $LaMnO_3$ . Differently from before, in this case, the  $Mn^{3+}$  has a partially filled shell  $3d^4$ . Despite the open shell, the material is an insulator because the  $3d$  electrons have a localized nature. One of the most



**Figure 5.5:** Real perovskite deformation for  $LaMnO_3$ .

studied materials of the last years is the  $LaMnO_3$  which has a perovskite structure. In its general form, this oxide has molecular formula  $La_{1-x}Sr_xMnO_3$  where  $x \in [0; 1]$  and it is called LSMO. This material has the particular property of the *magneto-resistance*. The magneto-resistance effect is the influence of the resistivity of a certain material on the magnetic field which acts on it. Fig.5.6 shows the behaviour of the resistivity as a function of the temperature and the behaviour of the resistance as a function of the internal field of the material. This effect can be exploited to make

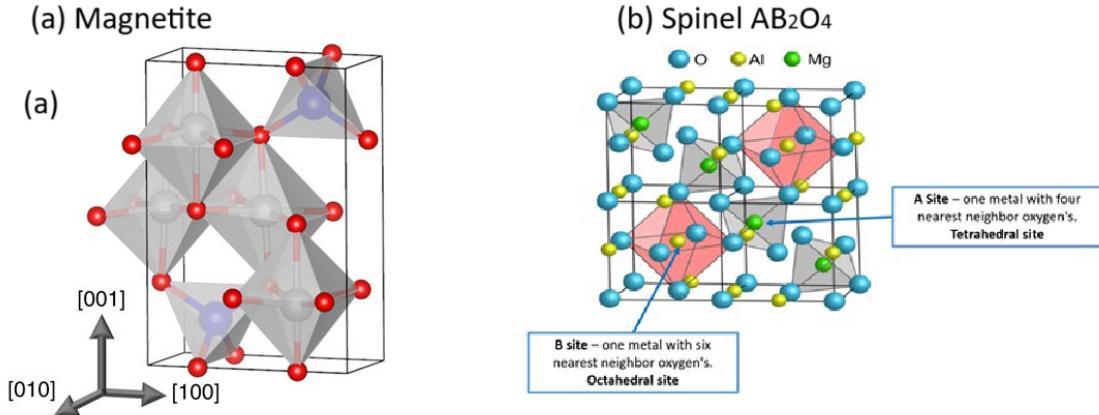


**Figure 5.6:** Magneto-resistance effect.

magnetic sensors which, by measuring the resistivity of a certain material, can compute the value of the field inside the material. LSMO is usually an insulator but by substituting some  $La^{3+}$  with  $Sr^{2+}$  the material can become a metal. In fact, not all oxides are insulators, although most of them are.

### 5.2.2 Spinel Structure

For the series "Nature is complicated" let's have a look at the magnetite structure. This compound with molecular formula  $Fe_3O_4$  is one of the most common magnetic materials in nature<sup>1</sup>. It is made by a mixture of octahedron and tetrahedron structures in which the iron ion is surrounded by oxygen atoms as we can see from Fig.5.7(a). Magnetite is an oxide with an FM behaviour even

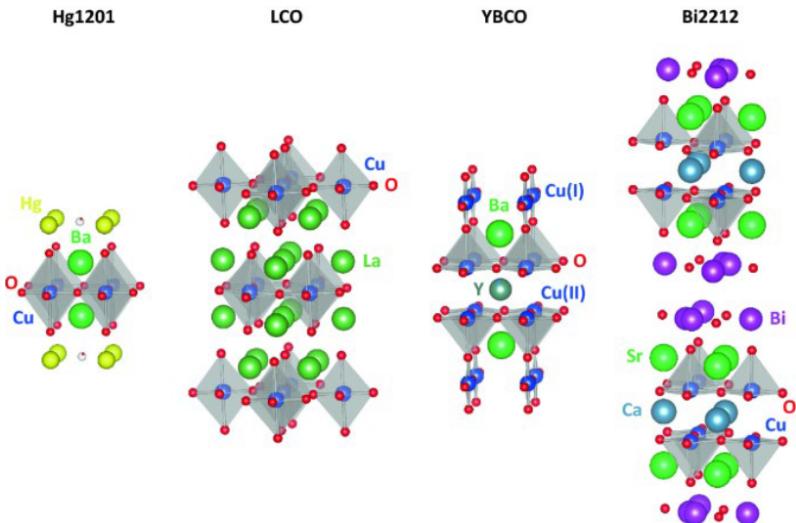


**Figure 5.7:** Magnetite structure.

if, microscopically its spins are not actually aligned. This is because this material has often the coexistence of two ions of the iron atom, the  $Fe^{2+}$  and the  $Fe^{3+}$ . In particular, we can generalise the expression of the molecular formula for magnetite by writing it as  $FeFe_2O_4$ . With this way of writing the formula, we can recognize the single ion  $Fe^{2+}$  (B site), two ions  $Fe^{3+}$  (A and B sites) and four  $O^{2-}$ . The general structure  $AB_2O_4$  is called spinel structure and it is represented in Fig.5.7(b).

### 5.2.3 Cuprates

Cuprates are the most interesting oxides. They are made of layers of copper oxides ( $CuO_2$ ) alternating with layers of other metal oxides. They can either have an octahedral structure or a pyramidal one. Fig.5.8 reports some examples of cuprates. Cuprates belong to the family of the



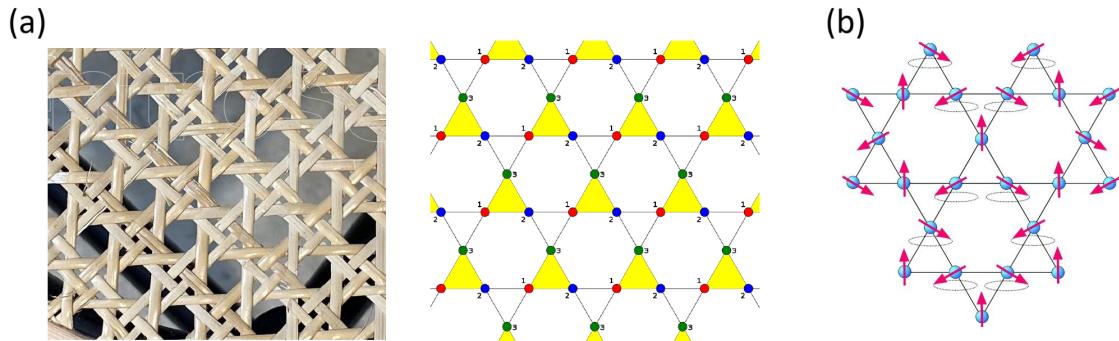
**Figure 5.8:** Cuprates examples.

so-called high-temperature superconductors and, in particular, they are the highest-temperature superconductors.

<sup>1</sup>Actually, once upon a time, the first phenomenological observations about magnetism were made by studying how magnetite interacts with other metals.

### 5.2.4 Kagome lattice

Some other interesting materials exist like, for example, the one which lattice is the so-called Kagome lattice. Kagome is a traditional Japanese woven bamboo pattern; its name is composed of the word *kago*, which means "basket" and *me*, meaning "eye(s)", referring to the pattern of holes in this particular woven basket. Fig.5.9(a) reports the structure of the lattice and the typical pattern of these baskets. An example of a compound which arranges itself in a kagome lattice is the *CoSn*. Apart from their unique structure, these materials are interesting since they represent a frustrated AFM material in which, because of the structure of the lattice, the spins cannot arrange themselves in a well-defined way (see Fig.5.9(b)).



**Figure 5.9:** Kagome lattice and AFM frustrated behaviour .

# Chapter 6

# Potential and Crystal Field

## 6.1 Centrifugal Potential

To better understand the effect of the crystalline field on a given atom is appropriate to remember some concepts related to the effective potential felt by an electron in an atom. In particular, since the effect of the crystalline field is strongly related to the radial part of the wave functions it is useful to remember the concept of the effective potential in the case of a central field, where the angular momentum is conserved. Even in classical mechanics, we have that the total energy is given by the sum of a potential energy and a kinetic one:

$$E_{tot} = U(r) + E_K = -\frac{D}{r} + \frac{1}{2}mv_\theta^2 + \frac{1}{2}mv_r^2$$

where  $U(r)$  is the central part of the potential energy which is attractive due to the Coulombian interaction between the nucleus and the electron, and the kinetic energy has been divided into two components: one tangential and one radial. Now, since the total angular momentum is  $L = mr\nu_\theta$  we can write the previous expression as:

$$E_{tot} = U(r) + E_K = -\frac{D}{r} + \frac{L^2}{2mr^2} + \frac{1}{2}mv_r^2$$

We can thus separate the two contributions to the potential energy defining the effective potential  $U_{eff}$  as:

$$U_{eff}(r) = -\frac{D}{r} + \frac{L^2}{2mr^2}$$

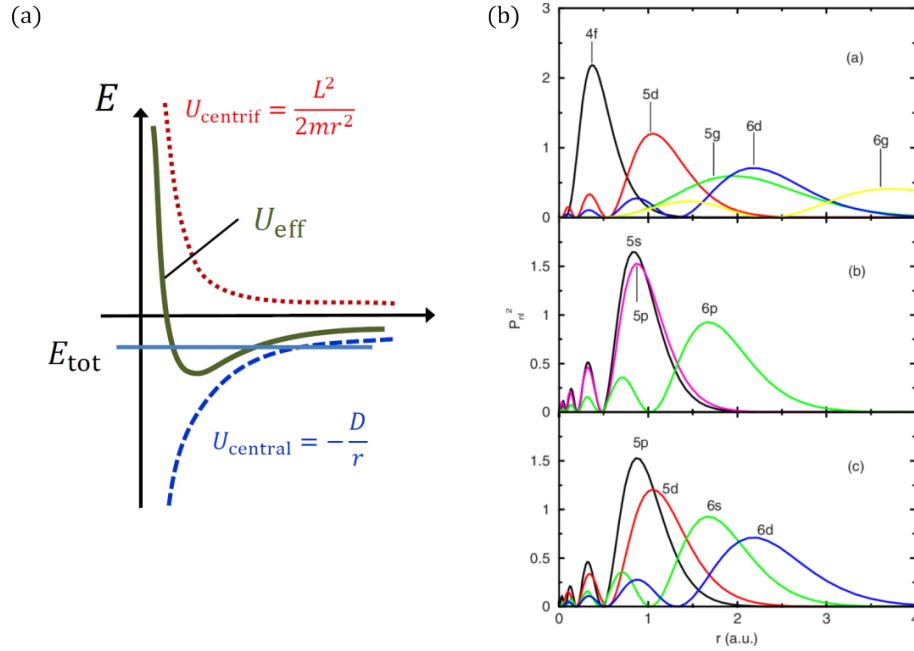
which is composed of a central attractive part and a centrifugal repulsive one as depicted in Fig.6.1(a). The total energy is then  $E_{tot} = U_{eff}(r) + \frac{1}{2}mv_r^2$  and determines the possible value of  $r$  that can be taken by the considered electron. The quantum description of such potential comes directly by substituting the continuous value of  $L^2$  with its quantum analogues  $L^2 \rightarrow l(l+1)\hbar^2$ . With this substitution, the effective potential becomes:

$$U_{eff}(r) = -\frac{D}{r} + \frac{l(l+1)\hbar^2}{2mr^2} \quad (6.1)$$

This means that for  $L = 0$ , which is the case of the  $s$  states, there is no centrifugal potential while for larger values of  $L$ , like in the  $d$  or  $f$  states, the centrifugal potential pushes the radial wave function away from the nucleus. The behaviours of the probability distributions as a function of the distance from the nucleus are depicted in Fig.6.1(b) for different orbitals. Such probability is proportional to the square modulus of the radial part of the wave function  $R_{nl}$ . More specifically:

$$P(r, r + dr) = 4\pi r^2 |R_{nl}|^2$$

Looking at the shape of this probability is another way of looking at the shape of the radial part of the orbitals.



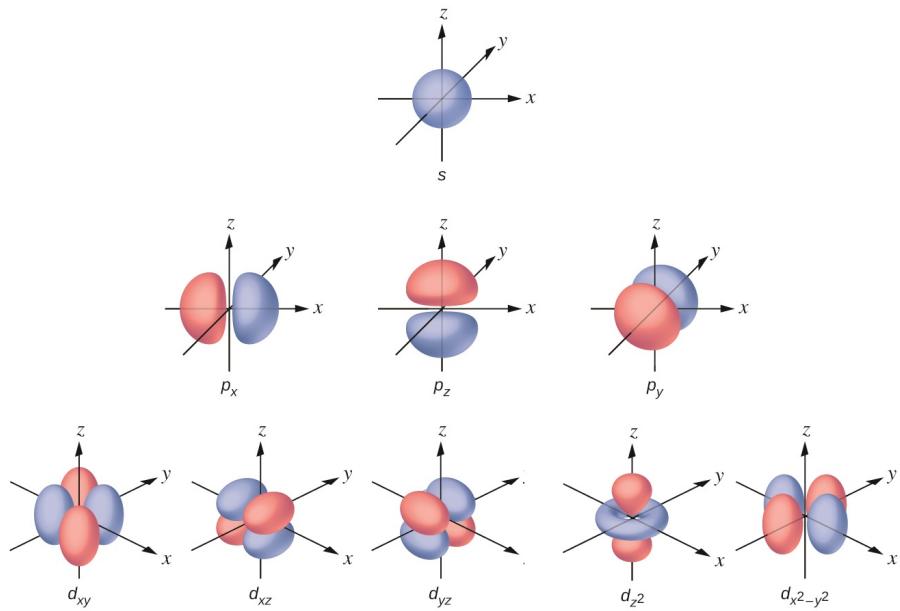
**Figure 6.1:** (a) Effective potential as a function of the distance. (b) Probability distribution as a function of the distance from the nucleus.

## 6.2 Electronic Configuration and Orbitals

We know that the electronic configuration of a given atom in "standard conditions" can be determined by using Hund's rules. Nevertheless, there are other possible configurations, higher in energy with respect to the GS one, in which the atom can be. We can use a table which has on its row all the possible values of the total spin  $M_s$  and on the columns all the possible values of  $M_l = -l, \dots, l$ . For example, let's consider the case of a  $2p^2$  subshell and let's build the table. In this case for each electron  $m_l = -1, 0, 1$  since  $l = 1$  and the total value, instead, can be  $M_l = -2, -1, 0, 1, 2$ . The result is reported in Tab.6.1. In the table, the parenthesis contains the value of  $m_l$  and the + or - indicates the value of  $m_s = \pm \frac{1}{2}$  for each electron. In the first line, for example, if  $M_l = 2$  we have both electrons must be in the subshell with  $m_l = 1$  and for Pauli's exclusion principle, since they are in the same shell, they must have opposite spin and therefore  $M_s = 0$  is the only possible value. Differently, for  $M_l = 1$  the electrons can be placed in different subshells and therefore they can also have the same spin which implies  $M_s = \pm 1$ . Moreover, always for  $M_l = 0$ , we can notice that three equivalent configurations correspond to  $M_s = 0$ . In general, then, there are many different configurations, not only the ground state one which is the one that we found by applying Hund's rules.

**Table 6.1:** Table for electronic configurations of  $2p^2$  subshell.

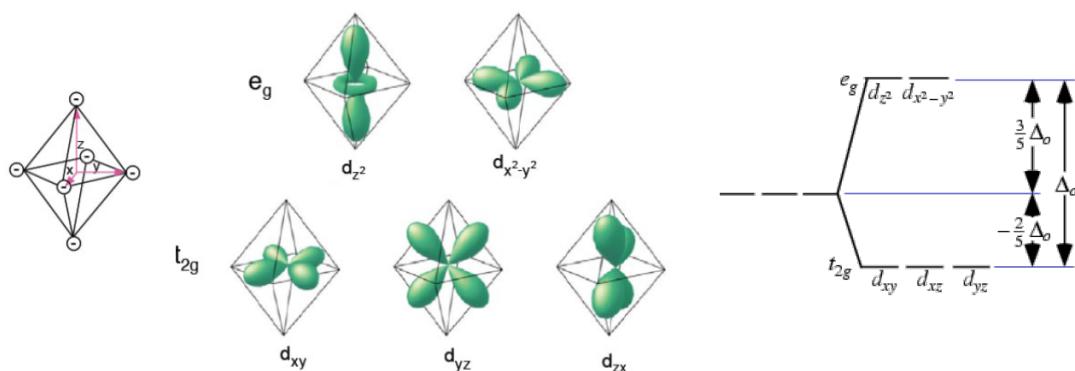
$M_l/M_s$	-1	0	1
2		(1,+); (1,-)	
1	(1,-)(0,-)	(1,+); (0,-); (1,-); (0,+)	(1,+),(0,+)
0	(1,-); (-1,-)	(1,+); (-1,-); (0,+); (0,-); (1,+); (-1,+) (1,-); (-1,+)	
-1	(0,-); (-1,-)	(0,+); (-1,-); (0,-); (-1,+)	(0,+); (-1,+)
-2		(-1,+); (-1,-)	



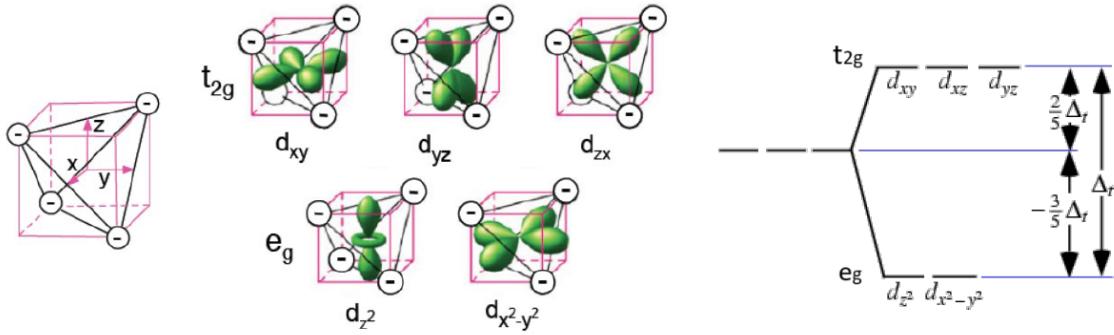
**Figure 6.2:** The angular distribution of the  $s$ ,  $p$  and  $d$  orbitals.

### 6.3 Orbitals and Crystal Field

Atomic orbitals are made by linear combination of harmonic functions and they are a really powerful tool to describe how atoms interact in solids. We know that based on the type of orbitals (*e.g.*  $s$ ,  $p_x$ ,  $p_y$ ,  $p_z$ , etc.) we will have different geometries in the space. Fig.6.2 list some of them. In particular, the  $d_{z^2}$  and  $d_{x^2-y^2}$  orbitals are grouped together and called  $e_g$  levels while the  $d_{xy}$ ,  $d_{xz}$  and  $d_{yz}$  levels are grouped together and called  $t_{2g}$  levels. These orbitals are the bottom ones in Fig.6.2. When an atom is placed into a solid and interacts with other atoms the orbitals determine how the energy levels are placed one with respect to the other. Let's, for example, consider an octahedral environment as depicted in Fig.6.3 in which a positive ion is caged by negative ions (*e.g.*  $O^{2-}$  ions). In such a situation the  $t_{2g}$  orbitals of the central positive ion better fit the distribution of the negative charges placed at the vertices of the octahedron. There is, indeed, a smaller overlap between the  $t_{2g}$  orbitals and the negative charges resulting in a smaller energy of the configuration. In the right part of Fig.6.3 the splitting of energy between the  $e_g$  and the  $t_{2g}$  levels is depicted. As told, because of the smaller overlap, the ground state level is  $t_{2g}$  while  $e_g$  is higher in energy by a quantity  $\Delta_0$ . The opposite happens, for example, in a tetrahedral configuration, depicted in Fig.6.4. In this case, the  $e_g$  orbitals have a smaller overlap with the negative charge distribution and therefore they constitute the ground state of the system while the  $t_{2g}$  constitute the first excited state. This is where the crystal field comes from: it arises mainly from electrostatic repulsion from the negatively charged electrons in the neighbouring orbitals. The crystal field also influences



**Figure 6.3:** Octahedral environment and orbitals displacement.



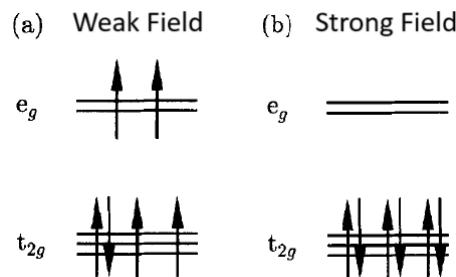
**Figure 6.4:** Tetrahedral environment and orbitals displacement.

how orbitals are filled with electrons. This is particularly evident for metal ions in which not all the  $3d$  electrons are present. In an octahedral configuration, in which the  $t_{2g}$  level is the lower in energy, the electrons present in the  $3d$  orbital will fill the lowest levels before filling the  $e_g$  levels. However, the precise order in which the orbitals fill depends on the competition between the crystal field energy  $\Delta_{CF}$  and the Coulomb energy cost of putting two electrons in the same orbital, which is known as the pairing energy  $J_{Hund}$  and it is usually positive. If  $\Delta_{CF} < J_{Hund}$  (weak-field case), then as the electrons are added to the system they will first singly occupy each orbital before any orbital becomes doubly occupied. On the other hand, if  $\Delta_{CF} > J_{Hund}$  (strong-field case), electrons will doubly occupy the lower energy orbitals before they consider scaling the dizzy energetic heights of the higher energy orbitals. These cases are illustrated in Fig.6.5. These two configurations are also called high-spin (weak-field) and low-spin (strong-field). In fact, in the weak-field case, we can see from Fig.6.5(a) that electrons are displaced to get a total spin  $S = 2$  (4 unpaired electrons) while in the strong-field case, all the electrons are paired and the spin  $S = 0$ .

In Section 1.5 we saw that the  $3d$  transition metal behave as if their total orbital angular momentum  $L$  is equal to zero. This phenomenon is called orbital quenching and it is due to the crystal field. In particular, for the ions of these elements, the crystal field is much stronger than the spin-orbit interaction which brings to a "re-organization" of the hierarchy of the interactions. Hence, Hund's third rule, which is based upon the fact that the spin-orbit interaction is the next most significant energy term after the Coulombian effect, is valid no more. The system, as shown by the experimental data, chooses the configuration with  $L = 0$ , *i.e.* with a quenched orbital momentum. From a more formal point of view this can be seen by remembering that:

$$\hat{\mathbf{L}} = -i\hbar\mathbf{r} \times \nabla$$

For such a system ( $3d$ TMs ions) can be proven that both the eigenvalues and the eigenfunctions are real which means that the ground state  $|GS\rangle$  is real. Therefore, since the expectation value  $\langle GS|\hat{\mathbf{L}}|GS\rangle$  must be real and the  $\hat{\mathbf{L}}$  operator is imaginary, the only possible result for the expectation value is zero. The orbital angular momentum is quenched.



**Figure 6.5:** Electronic configurations for the weak-field (a) and the strong-field (b) cases.

### 6.3.1 Sugano-Tanabe Diagrams

To keep into account the variation in the energy of the ground state with respect to the variation of the crystal field energy the graphical representation of the Sugano-Tanabe diagrams can be used. In particular, Tanabe-Sugano diagrams are used to predict absorptions in the ultraviolet (UV), visible and infrared (IR) electromagnetic spectrum of coordination compounds<sup>1</sup>. Some of these diagrams for 3dTM ions are shown in Fig.6.10 at the end of this chapter. We can clearly see the increase in the complexity of the diagrams with the increase in the number of electrons in the represented shell ( $d$  shell in this case). In general, similar electronic configurations have similar Sugano-Tanabe diagrams. For example, the diagram for the  $d^2$  (e.g.  $V^{3+}$  ion) is really similar to the  $d^8$  one (e.g.  $Ni^{2+}$  ion) while the  $d^4$  is similar to the  $d^6$ . Moreover, by looking at Fig.6.10 we can notice that both the y and the x-axis are scaled by a quantity  $B$ . This is one of the Racah parameters defined by the relation:

$$\begin{bmatrix} A \\ B \\ C \end{bmatrix} = \begin{bmatrix} 1 & 0 & -49 \\ 0 & 1 & -5 \\ 0 & 0 & 35 \end{bmatrix} \begin{bmatrix} F_0 \\ F_2 \\ F_4 \end{bmatrix}$$

where  $F_k$  are the Slater integrals defined as:

$$\begin{bmatrix} F_0 \\ F_2 \\ F_4 \end{bmatrix} = \begin{bmatrix} F^0 \\ \frac{1}{49}F^2 \\ \frac{1}{441}F^4 \end{bmatrix}$$

and lastly,  $F^k$  are the Slater-Condon parameters:

$$F^k = \int_0^\infty r_1^2 dr_1 \int_0^\infty r_2^2 dr_2 R^2(r_1) R^2(r_2) \frac{r_<^k}{r_>^{k+1}}$$

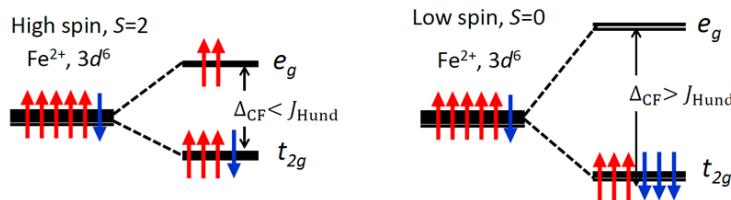
where this last integral expresses the superposition of the wave functions of two electrons on two given orbitals (can be the same orbital). Moreover,  $B$  is expressed in  $cm^{-1}$  which can be converted into an energy unit by using the relation for the wavenumber  $\tilde{\nu}$ :

$$\tilde{\nu} = \frac{1}{\lambda} = \frac{\omega}{2\pi c} = \frac{E}{2\pi\hbar c}$$

In the end, one obtains the conversion:

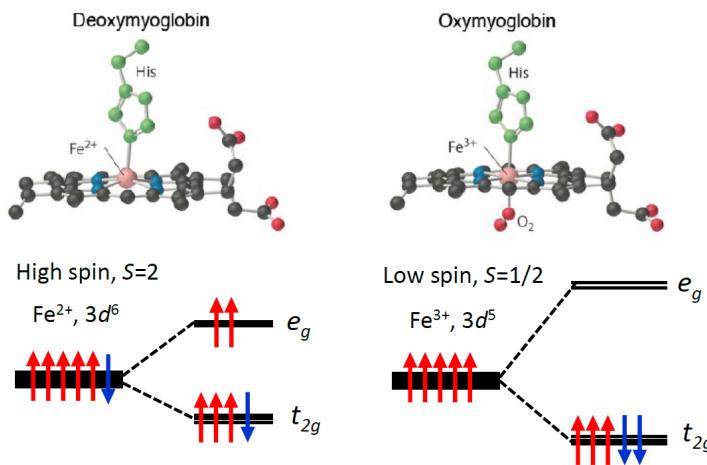
$$8066 cm^{-1} = 1 eV$$

These diagrams allow us to distinguish between the high-spin configuration and the low-spin one. Let's consider, for example, the case of the  $Fe^{2+}$  ion, which has 6 electrons in the  $3d$  shell ( $3d^6$ ). Its Sugano-Tanabe diagram is the one in Fig.6.10(d) in which a vertical line at  $\frac{\Delta_{CF}}{B} = 2$  separates the diagram into two parts. The left part corresponds to a ground state with a quintet configuration in which  $S = 2$  and therefore the spin degeneracy is  $2S + 1 = 5$ . The Russel-Sunders notation is  $^5T_2$  and this is a high-spin (HS) state. The right part, instead, corresponds to a singlet state in which  $S = 0$  and therefore this is a low-spin (LS) state. The energy levels are depicted in Fig.6.6. For the threshold value of the crystalline field  $\frac{\Delta_{CF}}{B} = 2$  we have a drastic change in the ground state configuration and therefore in the magnetic properties of the material. This drastic change is



**Figure 6.6:** HS and LS states for  $Fe^{2+}$ .

<sup>1</sup>A coordination complex is a chemical compound consisting of a central atom or ion, which is usually metallic, and a surrounding array of bound molecules or ions



**Figure 6.7:** Deoxymyoglobin and Oxymyoglobin molecules and their GS electronic configurations.

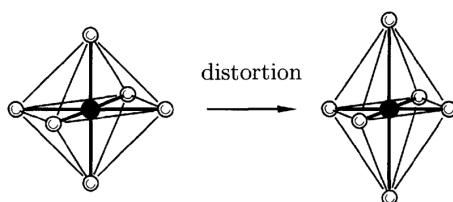
called the **HS to LS transition**. For example, this is the transition which happens when a bond with an oxygen molecule is added to the deoxymyoglobin molecule which becomes oxymyoglobin. In these two molecules, in fact, the oxygen bonds to an iron atom in the central position which passes from a 2+ oxidation state to a 3+ one. This changes the electronic configuration from being a high-spin  $S = 2$  to a low-spin one  $S = \frac{1}{2}$ . This change is due to the different displacement of electrons in the two cases because of the different energy splitting among the  $t_{2g}$  and the  $e_g$  levels which are influenced by the oxygen bond.

## 6.4 Jahn-Teller Effect

The Jahn-Teller effect is a spontaneous distortion of an orbital in a molecular system which decreases the energy of the system. It describes the geometrical distortion of molecules and ions that results from certain electron configurations. It is a consequence of the homonym theorem which states:

**Theorem 6.4.1 (Jahn-Teller Thorem)** *It is impossible for any quantum system to have an orbital degeneracy in the ground state*

From this theorem, we can directly understand that any molecule or complex ion in an electronically degenerate state will be unstable relative to a configuration of lower symmetry in which the degeneracy is absent. It means that it can sometimes be energetically favourable for, say, an octahedron to spontaneously distort as shown in Fig.6.8. From an energy levels point of view, this implies that energy levels are split and lowered in energy. The distortion raises the energy of certain orbitals while lowering the energy of others. In particular, in the cases of partially filled orbitals this effect can be highly significant while for completely full or completely empty orbitals the energy remains the same. For example, in the octahedral complex made of a  $Mn^{3+}$  with a ground state  $t_{2g}$  and a first excited state  $e_g$  (see Fig.6.9), these two levels that are 3 times and 2 times degenerate respectively, are split. As a result, the distortion lowers the energy because the singly occupied  $e_g$  level is lowered in energy. This effect can be phenomenologically understood by assuming that the distortion of the system can be quantified by a parameter  $Q$  which denotes the distance of distortion along a particular axis. Thus, the energy of the orbital can be written



**Figure 6.8:** Jahn-Teller effect for an octahedral complex.

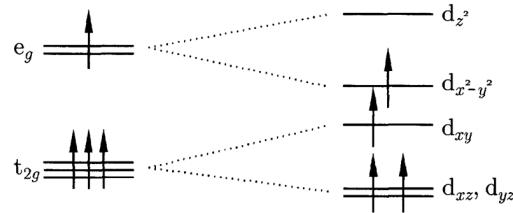
as the sum of a quadratic elastic term and a term which quantifies the lowering or the raising of the electronic energy:

$$E(Q) = \pm A Q + \frac{1}{2} M \omega^2 Q^2 \quad (6.2)$$

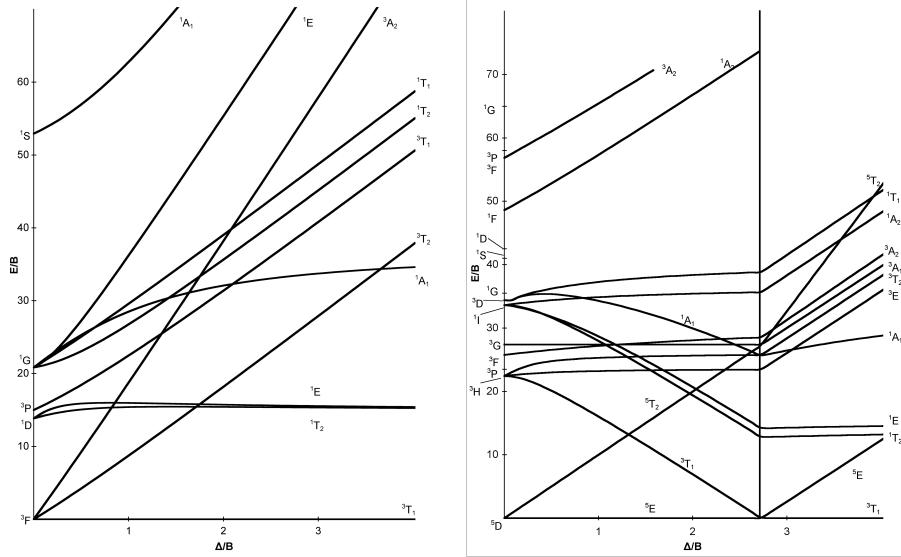
where the + or - sign apply to one or the other  $e_g$  orbitals which have been splitted. If both the split orbitals are occupied the overall effect is zero while if only one is partially occupied the system goes to one of the two configurations which has its own  $Q_{min}$ . For example, in the - case, the minimum of the energy is:

$$\frac{\partial E}{\partial Q} = 0 \Rightarrow Q_{min} = \frac{A}{M \omega^2}$$

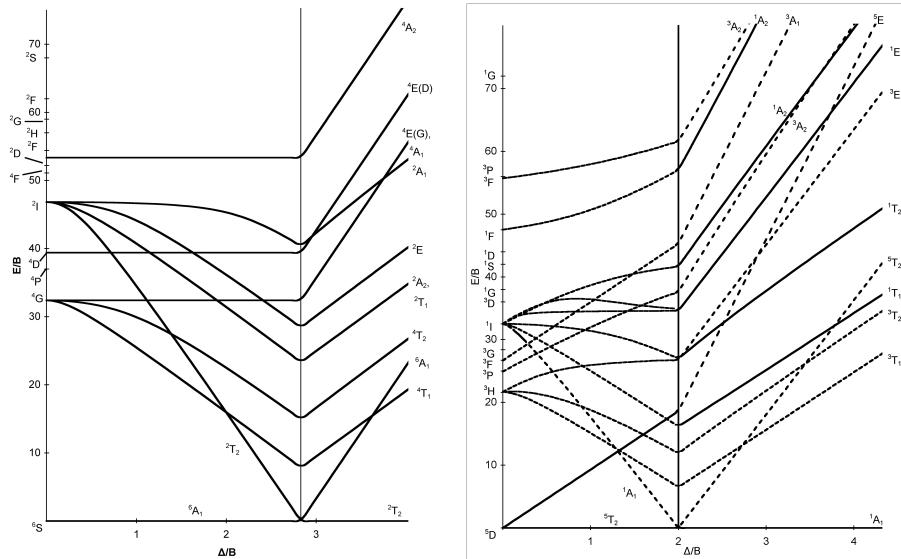
Which gives minimum energy of  $E_{min} = -\frac{A^2}{2M\omega^2}$  which is less than zero. This means that if only that orbital is occupied then the system can make a net energy saving by spontaneously distorting by  $Q_{min}$ .



**Figure 6.9:** Levels splitting because of Jahn-Teller effect for an octahedral complex of  $Mn^{3+}$ .



**(a)** Sugano-Tanabe diagram for  $d^2$  subshell. **(b)** Sugano-Tanabe diagram for  $d^4$  subshell.



**(c)** Sugano-Tanabe diagram for  $d^5$  subshell. **(d)** Sugano-Tanabe diagram for  $d^6$  subshell.

**Figure 6.10:** Sugano-Tanabe Diagrams.

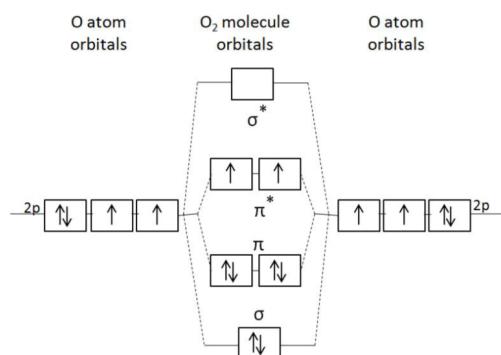
# Chapter 7

## Interactions

In this chapter, some of the different types of magnetic interaction will be discussed. These interactions, such as the exchange interactions, are the one which allows dipoles to communicate with each other and thus allow a material to get a particular magnetic order. We have already seen how the exchange interaction allows this kind of order and where it comes from and we will try to go further, analysing the other possible interactions.

### 7.1 Direct Exchange

The direct exchange is a group of interactions that do not have the need for an intermediary but proceed directly. An example of direct exchange is the exchange interaction. In this case, in fact, electrons on neighbouring magnetic atoms interact with each other with nothing more than a Coulombian repulsion. The exchange interaction can be either positive or negative. We have seen that for electrons belonging to the same atom  $J > 0$  because it is obvious that the electrons lower the total energy by occupying different orbitals, *i.e.* by increasing their average distance (first Hund's rule). Clearly, it is not always like this. In Section 3.1 we discussed the case of the hydrogen molecule in which the exchange interaction between electrons of two different hydrogen atoms brings to the creation of a bonding orbital and an antibonding one (Fig.3.1) since it lowers the energy of the system  $J < 0$ . In general, in molecules,  $J$  is often negative because of the overall gain in occupying the bonding molecular orbital with both electrons, which must then be in a singlet state with antiparallel spins like in the  $H_2$  case. Obviously, there are exceptions like the  $O_2$  molecule which has a triplet ground state ( $S = 1$ ) because there are 4 electrons per oxygen atom and only 12 state available, *i.e.* 8 electrons in total to place in 5 molecular orbitals as depicted in Fig.7.1 (two electrons remain unpaired). In general, direct exchange is often quite weak, in particular for lanthanides, where the atomic wave functions overlap very little. These materials are metal which means that the role of the conduction electrons should not be neglected, and a correct description needs to take account of both the delocalized and band character of electrons. Thus in many magnetic materials, it is necessary to consider some kind of indirect exchange interaction.



**Figure 7.1:** Molecular orbitals for oxygen molecule.

## 7.2 Indirect Exchange

Let's consider an ionic solid, like a metal. In metals, the exchange interaction between magnetic ions can be mediated by the conduction electrons. In fact, in metals made of lanthanides or in 3dTMs the  $3d$  and  $4f$  electrons are the ones responsible for generating the magnetic dipole of the atoms while the outer electrons ( $4s$ ,  $6s$  and  $5d$ ), which are the conduction electrons, provide the exchange thanks to the non-zero overlap of their wave functions with the ionic state one. The exchange interaction is thus indirect because it does not involve direct coupling between magnetic moments. This interaction is known as RKKY interaction in honour of their discoverer Ruderman, Kittel, Kasuya and Yosifa. They found that the effective exchange depends on the inter-ion distance with an oscillatory behaviour:

$$J_{RKKY} \propto \frac{\cos(2k_F r)}{r^3} \quad (7.1)$$

where  $r$  is the inter-ion distance and  $k_F$  is the Fermi momentum (assuming a spherical Fermi surface).

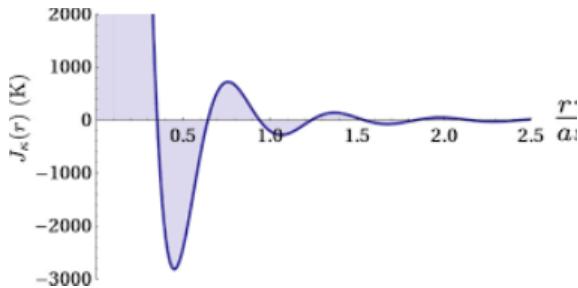


Figure 7.2:  $J_{RKKY}$  as a function of the inter-ion distance in lattice parameter units.

### 7.2.1 Super-Exchange

However, this is not always what happens when we deal with a metallic ion. In oxides, in fact, the situation is different. The exchange mechanism which is operative here is known as super-exchange. It can be defined as an indirect exchange interaction between non-neighbouring magnetic ions which is mediated by a non-magnetic ion which is placed between the magnetic ones. For example, let's consider the AFM manganese-oxide  $MnO$  depicted in Fig.7.3. For simplicity, we can assume that the magnetic moment of the manganese ion is due to a single unpaired electron and we want to consider the case in which  $J < 0$ , *i.e.* the unpaired electrons on two different manganese ions are opposite in spin. Let's suppose that the magnetic ion  $Mn^{2+}$  one on the left has one electron with spin up in the outer shell while the ion on the right has one electron with spin down (see the lower level on the left side of Fig.7.3). The oxygen, instead, has two electrons with opposite spin. We can think that either one electron on the oxygen ion  $O^{2-}$  is transferred to one of the two  $Mn^{2+}$  or that one electron from the  $Mn^{2+}$  goes to the  $O^{2-}$ . These two cases are represented on the second and third energy levels of Fig.7.3, respectively. In particular, we can notice that when the electron goes from the  $O^{2-}$  to the  $Mn^{2+}$  (in the figure the electron is transferred to the ion on the left) the energy of the system is smaller than the opposite case. This is because having electrons with opposite spins increases the energy of the system due to the Coulombian repulsion. On the other

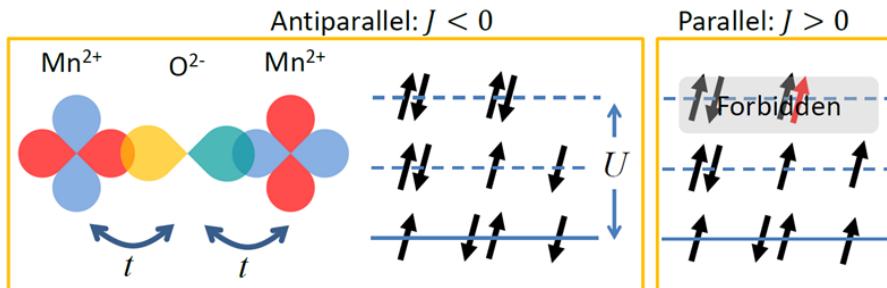
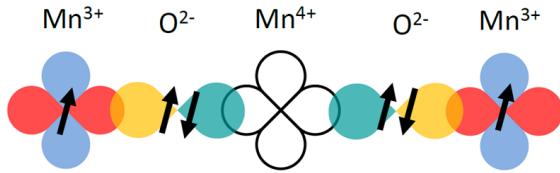


Figure 7.3: Super-exchange in  $MnO$ .



**Figure 7.4:** Double exchange in compounds with  $Mn$ .

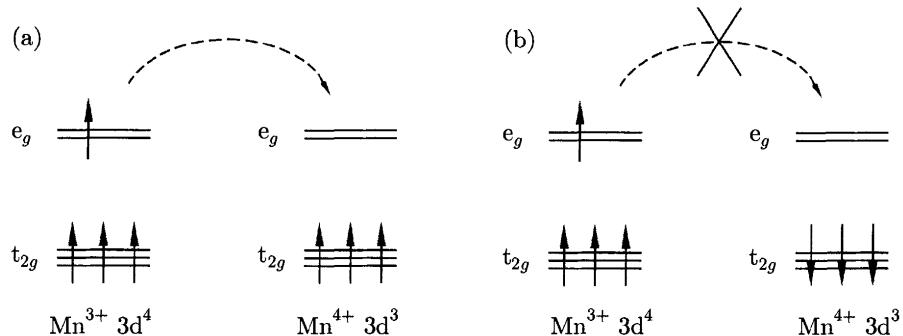
hand, if  $J > 0$  the electrons on the manganese ions would have the same spin and the electrons could only be transferred from the oxygen to the manganese and not vice versa since it would end up in a configuration with two electrons with the same spin in the same shell (see Fig.7.3). In general, the superexchange  $J$  is often negative and thus of antiparallel coupling, because it favours the virtual hopping of electrons from cation to cation via the oxygen, lowering the kinetic energy term of the total energy. Being  $t$  the hopping integral between the central non-magnetic ion ( $O^{2-}$ ) and the magnetic ion ( $Mn^{2+}$ ), it can be proven that the super-exchange is:

$$J \propto -\frac{t^2}{U} \quad (7.2)$$

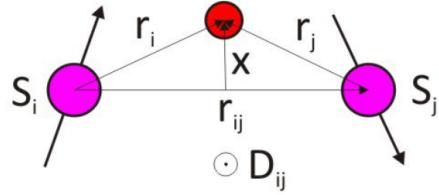
where  $U$  is the Coulombian repulsion energy.

### 7.2.2 Double Exchange

Oxides are not always antiferromagnetic. In some of them, in fact, it is possible to have a ferromagnetic exchange interaction ( $J > 0$ ) which occurs when the magnetic ions present in the oxides have a mixed valency, *i.e.* we have alternated sites with different valence. The alternated valence allows a full fluctuation of the charge from site to site and thus the parallel coupling is lowered in energy. An example of this phenomenon is given by compounds which contain  $Mn$  which can exist in oxidation state 3 or 4 as depicted in Fig.7.4. The empty sites of  $Mn^{4+}$  allow electrons to travel easily along the molecule only if the unpaired electrons on each  $Mn^{3+}$  have parallel spins. With antiparallel alignment, in fact, the on-site exchange energy would be paid for the charge fluctuation. Let's look at Fig.7.5 which represent the  $e_g$  and  $t_{2g}$  levels in  $Mn^{3+}$  and  $Mn^{4+}$  in the two cases of parallel and antiparallel spins. The  $e_g$  electron on a  $Mn^{3+}$  ion can hop to a neighbouring site only if there is a vacancy there of the same spin but since the neighbour is a  $Mn^{4+}$  which has no electron in its  $e_g$  shell, this does not represent a problem. However, there is a strong interaction between the  $e_g$  electron and the 3 electrons in the  $t_{2g}$  levels which wants to keep them all aligned as a consequence of the first Hund's rule. Thus is not energetically favourable for an  $e_g$  electron to hop to a neighbouring ion in which the  $t_{2g}$  spins are antiparallel to the  $e_g$  electron. Therefore, the ferromagnetic alignment is required to maintain the high-spin arrangement on both the donating and receiving ions. The parallel arrangement is favoured because it lowers the kinetic energy term of the Hamiltonian of the system. In conclusion, we can say that the double exchange interaction is responsible for the ferromagnetic behaviour of some oxides such as  $La_{a-x}Sr_xMnO_3$  ( $0 \leq x \leq 1$ ) or the more common magnetite ( $Fe_3O_4$ ).



**Figure 7.5:** Double exchange mechanism depending on the spin alignment.



**Figure 7.6:** Spin and  $\mathbf{D}$  vector orientation.

We have just seen that there exist two types of indirect exchange interactions which are responsible for two opposite magnetic behaviours and both interactions are present in oxides. The super-exchange makes an oxide antiferromagnetic while the double exchange makes another oxide ferromagnetic. Thus a question easily arises: how can we predict the magnetic behaviour of a given oxide? The Goodenough-Kanamori rule is an empirical rule which allows one to predict if an oxide is either ferromagnetic or antiferromagnetic. The Goodenough-Kanamori rule states that indirect exchange interactions are antiferromagnetic where the virtual electron transfer is between overlapping orbitals that are each half-filled, but they are ferromagnetic where the virtual electron transfer is from a half-filled to an empty orbital or from a filled to a half-filled orbital.

### 7.2.3 Anisotropic Exchange

There exists another kind of indirect exchange in which the spin-orbit interaction plays the same role as the oxygen atom in the super-exchange. Here the spin-orbit interaction creates an excited state in one of the magnetic ions and this excited state interacts with the ground state of the other neighbouring ion. This is known as anisotropic exchange interaction or also Dzyaloshinsky-Moryia (DM) interaction. The energy term due to this interaction is:

$$\hat{\mathcal{H}}_{DM} = \mathbf{D} \cdot \mathbf{S}_1 \times \mathbf{S}_2 \quad (7.3)$$

where the  $\mathbf{D}$  vector can take different orientations but it vanishes when the crystal field, *i.e.* the lattice, has an inversion symmetry with respect to the centre between the two magnetic ions. In AFM materials this interaction brings to a little FM component.

## 7.3 Magnons

We have seen so far that spins in magnetic ions can interact one with another through many different mechanisms. Since these spins are somehow "connected" by these interactions when the spins are perturbed they exhibit a collective behaviour which can give rise to the so-called spin waves. For a solid, at  $T = 0$  we know that atoms are perfectly ordered and (almost) no fluctuations perturb the system. At non-zero temperatures, the order is disturbed by thermally excited lattice vibrations which are quantized as phonons. For these quasiparticles, a dispersion relation, *i.e.* a relation between their frequency  $\omega$  (or their energy  $\hbar\omega$ ) and their wave vector  $k$ , can be derived. The same happens for a ferromagnet which at non-zero temperature can not be considered as perfectly ordered. At  $T \neq 0$  the order is disturbed by spin waves, quantised as magnons. Thus magnons play the same role in ferromagnets as phonons do in solids. In the following, we will try to derive the dispersion relation for magnons with a semiclassical approach.

### 7.3.1 FM Spin Waves

Let's start considering the ferromagnetic case of a one-dimensional ( $d = 1$ ) chain of  $N$  spins with dimensionality 3 ( $D = 3$ ), *i.e.* the spins can take any orientation in space. Since we are dealing with a ferromagnet the exchange interaction  $J$  between spins is positive. The general form of the energy for this system is:

$$U = -2J \sum_{p=1}^{N-1} \mathbf{S}_p \cdot \mathbf{S}_{p+1}$$

where here we used  $p$  to denote the p-th spin to not create confusion with the imaginary unit  $i$  that we will use in the following. The ground state energy for such a system is the energy for all the spins oriented in the same way and it is  $U_0 = -2J(N-1)S^2$ . When the temperature is different from zero all the spins are slightly tilted from their equilibrium position and the energy of a spin in the p-th site is given by:

$$U_p = -2JS_p \cdot (S_{p-1} + S_{p+1}) \quad (7.4)$$

We want to rewrite this energy term in a classical way and thus by considering the energy of a dipole  $\mu_p$  interacting with a local magnetic field  $\mathbf{B}_p$ :

$$U_p = -\mu_p \cdot$$

We can use the well-known relation for the dipole magnetic moment to extract the value of  $S_p$ :

$$\mu_p = -g\mu_B S_p \Rightarrow S_p = -\frac{\mu_p}{g\mu_B}$$

By replacing this result in Eq.(7.4) and rearranging the terms we get:

$$U_p = -\mu_i \cdot \underbrace{\left[ -\frac{2J}{g\mu_B} (S_{p-1} + S_{p+1}) \right]}_{\mathbf{B}_p} \quad (7.5)$$

This way of writing the equation gives us the possibility to understand the physical meaning of the local field  $\mathbf{B}_p$ . It is a fictitious magnetic field representing the effect on the p-th spin by the two nearest neighbours' spins. Classically, we can consider that the effect of the local field on the spin is a torque  $\tau = \mu_p \times \mathbf{B}_p$  which tries to align the dipole to the field. We can now write the second cardinal equation of dynamics by considering that the total angular momentum for each spin is  $\mathbf{M} = \hbar \mathbf{S}_p$ :

$$\begin{aligned} \tau &= \frac{d\mathbf{M}}{dt} = \hbar \frac{d\mathbf{S}_p}{dt} = \mu_p \times \mathbf{B}_p = -g\mu_B S_p \times \left[ -\frac{2J}{g\mu_B} (S_{p-1} + S_{p+1}) \right] = \\ &= 2J \begin{vmatrix} u_x & u_y & u_z \\ S_p^x & S_p^y & S_p^z \\ (S_{p-1}^x + S_{p+1}^x) & (S_{p-1}^y + S_{p+1}^y) & (S_{p-1}^z + S_{p+1}^z) \end{vmatrix} \end{aligned}$$

This vectorial equation gives rise to 3 scalar equations:

$$\begin{cases} \frac{dS_p^x}{dt} = \frac{2J}{\hbar} [S_p^y(S_{p-1}^z + S_{p+1}^z) - S_p^z(S_{p-1}^y + S_{p+1}^y)] \\ \frac{dS_p^y}{dt} = \frac{2J}{\hbar} [-S_p^x(S_{p-1}^z + S_{p+1}^z) + S_p^z(S_{p-1}^x + S_{p+1}^x)] \\ \frac{dS_p^z}{dt} = \frac{2J}{\hbar} [S_p^x(S_{p-1}^y + S_{p+1}^y) - S_p^y(S_{p-1}^x + S_{p+1}^x)] \end{cases} \quad (7.6)$$

If we now consider that  $z$  is the vertical direction and that the tilt angle  $\theta$  of each spin with respect to the z-axis is small we can approximate:

$$S_p^z \approx S = \text{const}; \quad S_p^x, S_p^y \ll S$$

Thus we get:

$$\begin{cases} \frac{dS_p^x}{dt} = \frac{2JS}{\hbar} [2S_p^y - S_{p-1}^y - S_{p+1}^y] \\ \frac{dS_p^y}{dt} = -\frac{2JS}{\hbar} [2S_p^x - S_{p-1}^x - S_{p+1}^x] \\ \frac{dS_p^z}{dt} = 0 \end{cases}$$

To solve this set of differential equations we need to provide a particular solution. Since we are looking for a wave which propagates along the chain we can look for a solution of the type:

$$\begin{aligned} S_p^x &= ue^{i(pka - \omega t)} \\ S_p^y &= ve^{i(pka - \omega t)} \end{aligned}$$

where  $u, v \in \mathbb{C}$ . By using these particular solutions the equation for the x-term becomes:

$$-i\omega ue^{i(pka - \omega t)} = \frac{2JS}{\hbar}(2 - e^{-ika} - e^{ika})ve^{i(pka - \omega t)}$$

So by simplifying the exponential and by using Euler's formula, we obtain both for  $x$  and  $y$ :

$$\begin{cases} -i\omega u = \frac{2JS}{\hbar}(2 - 2 \cos ka)v \\ -i\omega v = -\frac{2JS}{\hbar}(2 - 2 \cos ka)u \end{cases}$$

This system can be written in matrix form:

$$\begin{bmatrix} -i\omega & \frac{4JS}{\hbar}(1 - \cos ka) \\ -\frac{4JS}{\hbar}(1 - \cos ka) & i\omega \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

The system has non-trivial solutions only if the determinant of the coefficient matrix is equal to zero. Thus we have:

$$-\omega^2 + \left[ \frac{4JS}{\hbar}(1 - \cos ka) \right]^2$$

This gives us the dispersion relations for magnons in a 1D ferromagnetic chain:

$$\hbar\omega = 4JS(1 - \cos ka) \quad (7.7)$$

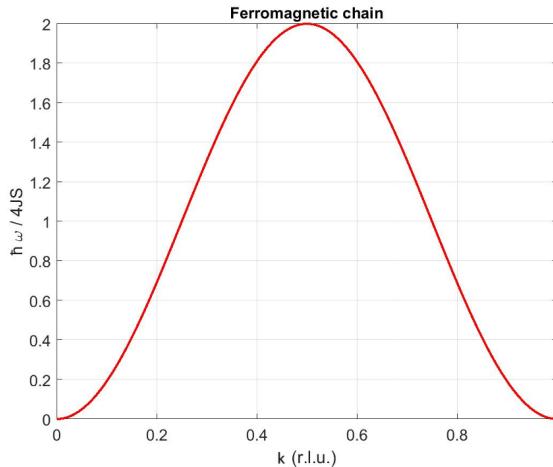
The dispersion relation is shown in Fig.7.7. We can notice that the maximum energy is  $8JS$  and it is proportional to the exchange interaction  $J$ . Moreover, the dispersion relation is quadratic for small values of  $k$ , *i.e.*  $\hbar\omega \sim_{k \rightarrow 0} k^2$ . The solution can now be found and it can be demonstrated that it is:

$$u = iv \quad (7.8)$$

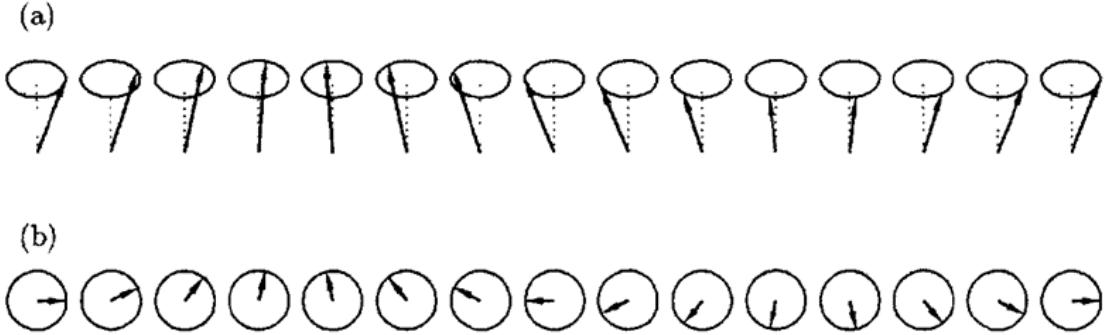
This means that the two phase terms associated with the  $x$  and  $y$  components have a phase shift of  $\frac{\pi}{2}$ . So we get:

$$\begin{aligned} \Re(S_p^x) &= u \cos(pka - \omega t) \\ \Re(S_p^y) &= u \sin(pka - \omega t) \end{aligned}$$

where here  $u \in \mathbb{R}$ . These two equations describe a circular motion of the spin in the  $xy$  plane. If we consider more than one spin, *i.e.* we change the p-coordinate of the site in the equation, we get the motion described in Fig.7.8(b). The equations describe a progressive wave made by the spin which moves rightward along the chain. This is a **spin wave**. Moreover, since the spin component along  $z$  has been considered constant this means that the overall motion in the space is a precession



**Figure 7.7:** Magnons dispersion relation for a 1D FM chain.



**Figure 7.8:** Spin Wave representation.

along the z-axis as depicted in Fig.7.8(a). The approximation we made about the smallness of the tilt angle  $\theta$  can be confirmed by proofing that:

$$\theta \propto \frac{1}{\sqrt{NS}}$$

where  $N$  is a very large number since it is the number of spins along the chain and  $S$  in the spin value that, for instance, is equal to  $\frac{1}{2}$  for an electron.

### 7.3.2 AFM Spin Waves

We can try now to do the same for a 1D chain of spins arranged in an antiferromagnetic way. Therefore, in this case, each spin is opposite to its two nearest neighbours since  $J < 0$ . We need to proceed as before but we need to change something in the dynamic equations in Eq.(7.6). In this case, we need to approximate:

$$S_p^z \simeq S; \quad S_{p+1}^z \simeq -S; \quad S_{p-1}^z \simeq -S;$$

Since the spins are oriented in opposite directions. What we obtain is:

$$\begin{cases} \frac{dS_p^x}{dt} = -\frac{2JS}{\hbar} [2S_p^y - S_{p-1}^y - S_{p+1}^y] \\ \frac{dS_p^y}{dt} = \frac{2JS}{\hbar} [2S_p^x - S_{p-1}^x - S_{p+1}^x] \\ \frac{dS_p^z}{dt} = 0 \end{cases}$$

In the AFM, as told, each site is different from its nearest ones. Since, for example, if the p-site is positive the p+1 and p-1 sites are negative we need to write another set of equations for the p+1 site. This results in:

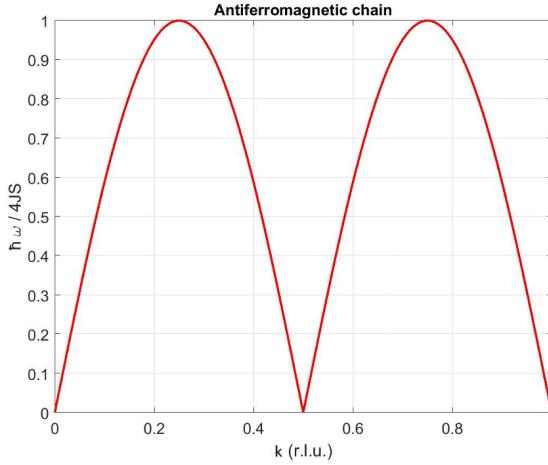
$$\begin{cases} \frac{dS_{p+1}^x}{dt} = \frac{2JS}{\hbar} [2S_{p+1}^y + S_p^y + S_{p+1}^y] \\ \frac{dS_{p+1}^y}{dt} = -\frac{2JS}{\hbar} [2S_{p+1}^x + S_p^x + S_{p+2}^x] \\ \frac{dS_{p+1}^z}{dt} = 0 \end{cases}$$

In order to get a propagating wave as a solution we don't need both full sets of equations but we can consider only a linear combination of them, such as:

$$\begin{cases} S_p^+ = S_p^x + iS_p^y \\ S_{p+1}^+ = S_{p+1}^x + iS_{p+1}^y \end{cases} \quad (7.9)$$

Now, in order to select the right particular solution we need to keep into account the fact that in the AFM the periodicity of the chain, i.e. the distance between equally oriented spins, is doubled with respect to the previous case. Thus we can choose as particular solutions:

$$\begin{aligned} S_p^+ &= ue^{i(2pka-\omega t)} \\ S_p^+ &= ue^{i[2(p+1)ka-\omega t]} \end{aligned}$$



**Figure 7.9:** Magnons dispersion relation for a 1D AFM chain.

By replacing in the linear combination what we have obtained so far we get:

$$\begin{cases} \frac{dS_p^+}{dt} = \frac{2JS}{\hbar}i[2S_p^+ + S_{p-1}^+ + S_{p+1}^+] \\ \frac{dS_{p+1}^+}{dt} = -\frac{2JS}{\hbar}i[2S_{p+1}^+ + S_p^+ + S_{p+2}^+] \\ \Rightarrow \begin{cases} -i\omega u = \frac{4JS}{\hbar}i(u + v \cos ka) \\ -i\omega v = \frac{4JS}{\hbar}i(v + u \cos ka) \end{cases} \end{cases}$$

For simplicity we can define the exchange frequency  $\omega_{ex} = \frac{4JS}{\hbar}$  and then rewrite the system in a matrix form:

$$\begin{bmatrix} \omega_{ex} - \omega & \omega_{ex} \cos ka \\ \omega_{ex} \cos ka & \omega_{ex} + \omega \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

By imposing the determinant to be equal to zero we get the dispersion relation for magnons in a 1D antiferromagnetic chain, depicted in Fig.7.9.

$$\hbar\omega = 4|J|S|\sin ka| \quad (7.10)$$

It is worth noting that when moving from the FM to the AFM case, both the period and the maximum energy of the magnons get halved. Moreover, the behaviour of the dispersion relation for small values of  $k$  is different. In the FM case, it was quadratic, now it is linear  $\hbar\omega \sim_{k \rightarrow 0} k$ .

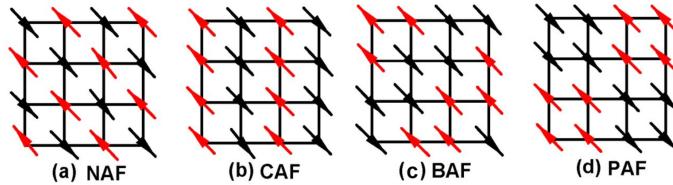
# Chapter 8

## Neutron Scattering Techniques

For such a variety of magnetic phenomena, the experimental techniques to investigate them become fundamentals. When we speak about experiments in magnetic materials the crucial point is to link macroscopic and measurable quantities (*e.g.* magnetization, resistivity, refractive index, etc.) to microscopic quantities (*e.g.* spin arrangements, lattice parameter, etc.) which gives us information about the nature of the studied material. In AFM materials this becomes quite complicated since one of the most important magnetic macroscopic quantities, the magnetization, is null for these materials. Here, one of the most effective techniques is neutron scattering which exploits the periodicity of the lattice to retrieve information about the magnetic order of the material.

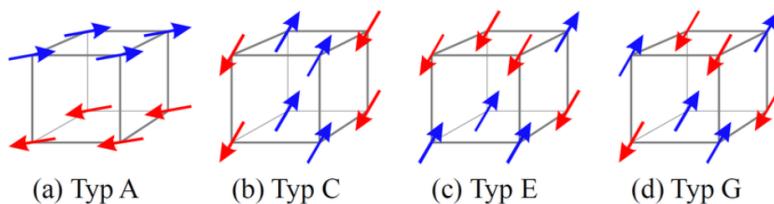
### 8.1 AFM lattices

Before dealing with the technique in itself let's have a look at the most relevant antiferromagnetic lattices. We can start by looking at the 2D most relevant AFM lattices which are represented in Fig.8.1. The most "natural" choice of AFM lattices is the Néel Antiferromagnetic (NAF) lattice



**Figure 8.1:** Most relevant 2D AFM lattices.

in Fig.8.1(a). Here the diagonal planes are FM and the overall order is AFM. The NAF lattice is the most natural one since it respects the symmetry between neighbouring sites and for this reason, the other three choices, the CAF, the BAF and the PAF lattices will never be realised in a perfectly squared geometry. Nevertheless, small anisotropy in the system can lead to the formation of such geometries. Clearly, other possible non-squared lattices exist, such as the hexagonal one, the frustrated triangular lattice or the already-discussed kagome lattice. The most relevant 3D geometries, instead, are reported in Fig.8.2. Also in this case we have the most symmetric lattice, which is the Type G one (Fig.8.2(c)), and other possible less symmetric choices. In all cases, a single atomic FM plane can be individuated. In the Type G lattice, for instance, the FM plane is the (1,1,1).



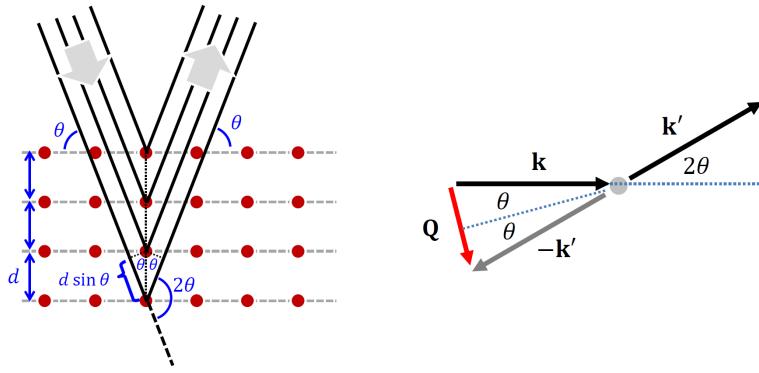
**Figure 8.2:** Most relevant 3D AFM lattices.

## 8.2 Bragg Diffraction and Elastic Scattering

As we said, to study the magnetic properties of AFM materials we will use the neutron scattering technique. To better understand this technique and the concepts at its very basis, it is worthwhile to review the Bragg diffraction and the concept of scattering vector. We know that when a plane wave is made impinging onto a crystal, the different crystalline planes will diffract the wave in such a way that, if a certain condition regarding the incidence angle is respected, the diffracted waves will have a constructive interference among each other. This "certain condition" for the angle to respect is called Bragg's condition and it is nothing more than the interference condition: the difference in the optical paths of the different diffracted waves must be an integer multiple of the wavelength. This results in the mathematical expression of Bragg's condition which is:

$$N\lambda = 2d \sin \theta \quad (8.1)$$

where  $d$  is the distance between two adjacent crystalline planes and  $\theta$  the incidence angle as depicted in Fig.8.3. The most important quantity here is the angle  $\theta$ . It is worth noting that



**Figure 8.3:** Bragg diffraction scheme.

with respect to the initial position the beam is deviated by an angle  $2\theta$  which means that, in a real experiment, to detect the diffracted waves we need to place our detector at an angle  $2\theta$  with respect to the sample. Moreover, we need to keep into account that the number of planes involved in the diffraction process is a huge number and therefore the resulting interference peaks are really narrow and intense. This can be understood by remembering that for a diffraction grating of  $N$  coherent sources equally spaced by  $d$  along a line, with equal intensity  $I_0$ , the intensity of the wave arising from the interference among all the  $N$  sources is:

$$I(\theta) = \frac{I_0}{r^2} \left( \frac{\sin \frac{N\pi d \sin \theta}{\lambda}}{\sin \frac{\pi d \sin \theta}{\lambda}} \right)^2$$

where  $r$  is the distance at which the intensity is measured and  $\theta$  is the incident angle. Hence, the intensity grows with  $N$  and that's why in our case, with a big number of crystalline planes involved, we have narrow and intense peaks. Going back to our analysis, we can relate the wave vector of the incident waves with the wave vector of the scattered waves. This is graphically shown on the right side of Fig.8.3 and it results in:

$$\mathbf{Q} = \mathbf{k} - \mathbf{k}' \quad (8.2)$$

the  $\mathbf{Q}$  vector is known as scattering vector. In general, in the case of punctiform scattering centres, it is possible to compute the elastic scattering rate. For a flux  $\Phi$  of incident particles (*e.g.* photons, neutron, electrons, etc.) which are *elastically* scattered by  $N$  scattering centers locate at position  $\mathbf{R}_j$ , the scattering rate  $R(2\theta, \phi)$  is proportional to:

$$R(2\theta, \phi) \propto \Phi \left| \sum_{j=1}^N f_j(\lambda, \theta) e^{-i\mathbf{Q} \cdot \mathbf{R}_j} \right|^2 \quad (8.3)$$

where here in the scattering rate the dependence on the total deviation angle  $2\theta$  is made explicit and the solid angle  $\phi$  is reported for completeness<sup>1</sup>. What matters here is the so-called form factor

<sup>1</sup>Usually, scattering is an isotropic phenomenon and there is no dependence on the solid angle.

$f_j(\lambda, \theta)$  which depends on the type of scattered particle and defines the relevant interaction for each scattering centre. Another important quantity in scattering phenomena is the cross-section  $\sigma(\lambda)$ , measured in Barn ( $1\text{barn} = 10^{-28}\text{m}^2$ ), which is the effective area centred on the scattering centre, *i.e.* if the scattering particle passes within the cross-section the scattering phenomenon occurs, otherwise it does not. It is defined as:

$$\sigma(\lambda) = 2\pi \int_{\theta=0}^{2\pi} |f(\lambda, \theta)|^2 \sin \theta d\theta \quad (8.4)$$

Moreover, the term  $\left| \sum_{j=1}^N f_j(\lambda, \theta) e^{-i\mathbf{Q} \cdot \mathbf{R}_j} \right|^2$  in the scattering rate can be simplified into  $\left| \sum_{j=1}^N e^{-i\mathbf{Q} \cdot \mathbf{R}_j} \right|^2$  if one assumes that the form factor  $f_j(\lambda, \theta)$  is a constant term (which is true in the neutron scattering case). We can now notice that this is the definition of the reciprocal lattice: the vectors  $\mathbf{Q}$  of the reciprocal lattice are those vectors which satisfy the relation  $\left| \sum_{j=1}^N e^{-i\mathbf{Q} \cdot \mathbf{R}_j} \right|^2$ . This can be also expressed by stating that the reciprocal lattice is the Fourier transform of the direct lattice: the x-ray diffraction provides the reciprocal lattice of the crystal under analysis.

### 8.2.1 Scattering Form Factor

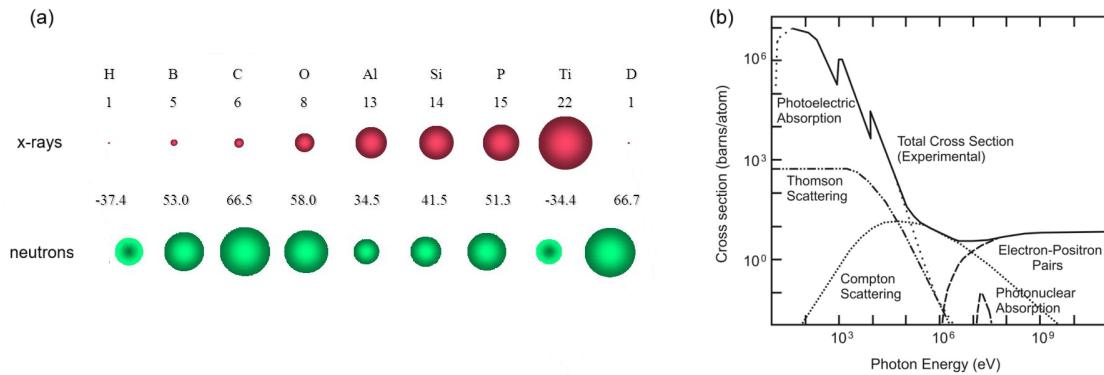
Let's start by considering the problem of studying the structure of an AFM lattice. In the lattice all atoms are identical and they all scatter in the same way. If the scattered particles are photons (x-ray photons in particular) there is no possibility to distinguish one atom from the other and, for instance, there is no chance to distinguish a spin-up atom from a spin-down one<sup>2</sup>. If we remember, in fact, all the light-matter interaction phenomena are controlled by a matrix element which, in the photon case, has no dependence on the spin of the atom. The solution, then, stays in the use of neutrons which have a magnetic dipole moment and can interact with the spin of the atom. Moreover, in the elastic scattering case, the form factor for neutrons is constant and equal to:

$$f(\lambda, \theta) = -b \quad (8.5)$$

where  $b$  is called scattering length, it is in the order of  $10^{-14}\text{m}$  and depends on the particular nucleus with which the neutron is interacting. For x-rays, instead:

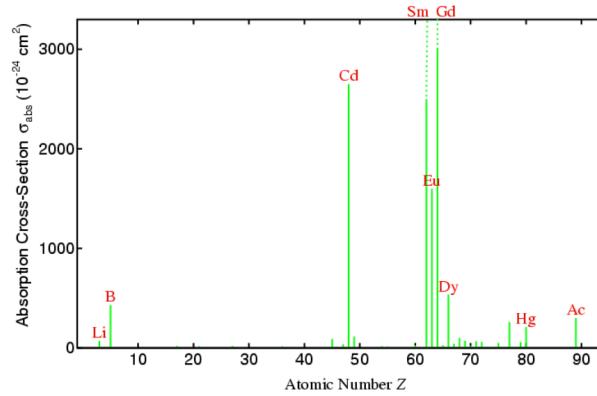
$$f(\lambda, \theta) = Zg(Q)r_e \quad (8.6)$$

where  $Z$  is the atomic number of the scattering centre,  $g(Q)$  is a particular function which is worth 1 at  $Q = 0$  and goes to 0 for  $Q \rightarrow \infty$  and  $r_e = 2.818 \cdot 10^{-15}\text{m}$  is the Thomson scattering length. These two expressions of the form factor strongly influence the behaviour of the cross-section which is almost constant for elastically scattered neutrons and strongly variable and dependent on the atomic number for photons as we can see from Fig.8.4.



**Figure 8.4:** (a) Cross-sections for some elements in the x-ray (red) and the neutron (green) case. (b) Cross-section as a function of the energy for photons.

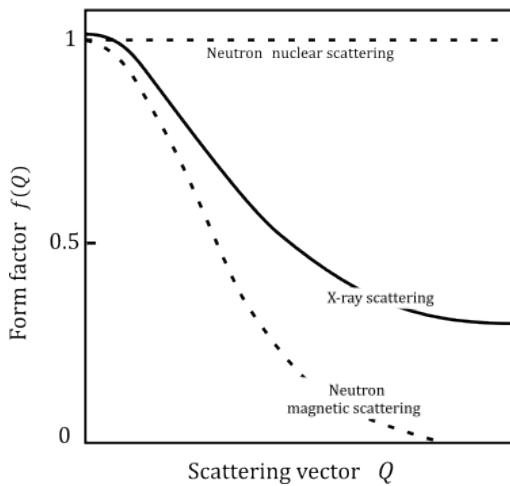
<sup>2</sup>To be fair x-rays can be used to study spin interaction in the so-called resonant diffraction which exploits the spin-orbit interaction but it is far more complicated than to use neutrons.



**Figure 8.5:** Absorption cross-section for elastic neutron scattering as a function of the atomic number.

For example, because of the strong dependence of the photon cross-section on  $Z$ , it is impossible to observe hydrogen atoms with x-ray scattering. Clearly, neither elastic neutron scattering is an infallible technique. The "enemy" of neutrons is the absorption. Some nuclei, in fact, can absorb neutrons with a certain energy and become radioactive in the so-called neutron capture. This phenomenon is strongly dependent on the type of nuclei and we can see from Fig.8.5 that is quite relevant for some lanthanoids such as *Sm*, *Eu* and *Gd*. For neutrons, there are two possible ways of scattering:

- Via atomic force interaction. This is a short-range interaction which acts between nucleons (protons and neutrons) and that is responsible for "keeping together" these particles in the nucleus. This interaction can be used to determine the size of nuclei but it does not depend on the spin and is not interesting for the study of magnetic properties. Actually, as can we see from Fig.8.6 this is the interaction which leads to the constant form factor seen above.
- Via spin interaction. This interaction comes from the fact that both neutrons and the atoms in the magnetic material have spin. For neutrons, it is due to the dipolar magnetic interaction of the particle with the atomic magnetic dipoles. The form factor for this type of scattering is far more similar to the one of x-rays (see Fig.8.6)



**Figure 8.6:** Neutron form factors as a function of the scattering vector compared with the x-rays form factor.

### 8.3 Neutron Scattering

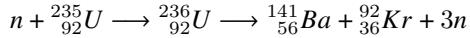
Now that we know that neutrons are the right particles to observe magnetic phenomena we need to understand how to produce neutrons in reality.

### 8.3.1 Neutron Sources

There are two main methods to produce neutrons:

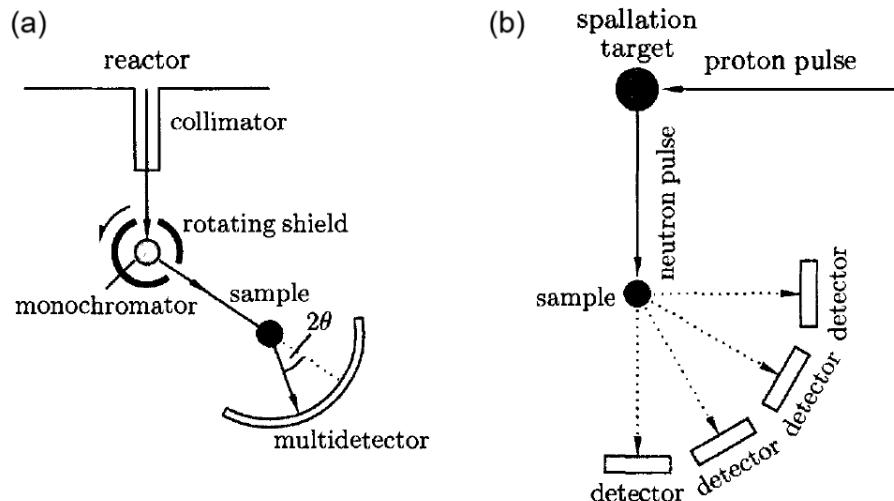
- Fission of  $^{235}U$
- Spallation

The first procedure consists of the fission reaction of the fissile isotope of Uranium which is only 0.7% of the Uranium present in nature. The nuclear reaction involved is:



To be performed such a procedure requires the presence of a nuclear reactor in which the fission reactor occurs. A neutron, coming from previous reactions is used to feed the reactor while the other remaining neutron, whose quantities depend on a probability low but it is often equal to two, are available to be used for neutron scattering. However, before being ready to be used the neutrons must be slowed down. Water is used to do that and the neutrons' kinetic energy at the end of the slowing process is  $\sim 25\text{meV}$  which is the energy corresponding to a room temperature condition. Eventually, neutrons are ready to be focused on the sample. The other method for neutron production is spallation. This method consists of bombarding with a high-speed proton a suitable atom (*e.g.* Tungsten  $W$ ) which absorbs the proton and emits a neutron<sup>3</sup>. This method requires the presence of a particle accelerator to accelerate enough the proton.

Regardless of the method used, once produced the neutrons have different velocities, *i.e.* different wavelengths, and the appropriate one has to be selected. To do that we use Bragg condition on a well-known crystal. In fact, by looking at Eq.(8.1) it is clear that, by knowing the distance  $d$  between the adjacent planes of the crystal and changing appropriately the angle  $\theta$  it is possible to "select" the wavelength of the outgoing wave and thus the velocity of the neutrons. This instrument is called single-crystal diffractometer (the "monochromator" in Fig.8.7). Fig.8.7

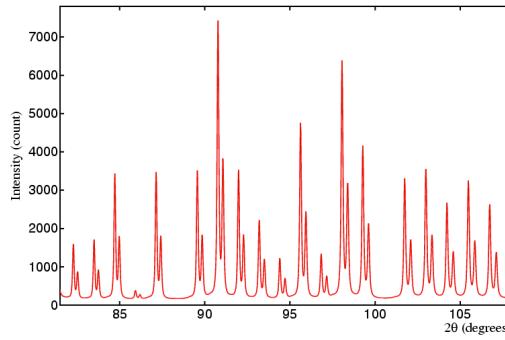


**Figure 8.7:** Neutron elastic scattering experimental setup in the two production cases.

shows the schematic experimental apparatus for neutron elastic scattering. After the single-crystal diffractometer, the neutrons are directed towards the sample to study where the elastic scattering occurs. Since scattering is an isotropic phenomenon the scattered waves go in all directions and the detectors used to detect these waves are usually huge in dimension, so huge that they usually cover all the surfaces of the room in which the sample is placed. The sample has to be oriented in such a way that the Bragg condition is respected in order to have diffraction. Sometimes a sample in powder form is used. This technique, called powder diffraction, exploits the fact that in the powder all the small fragments of the sample are oriented in a random way with respect to the incident beam and therefore some of them will have the right orientation to cause diffraction.

---

<sup>3</sup>The mechanism is much more complex and deeper than this but for our purposes, it is enough.



**Figure 8.8:** Scattering intensity as a function of the  $\theta$  angle for powder scattering.

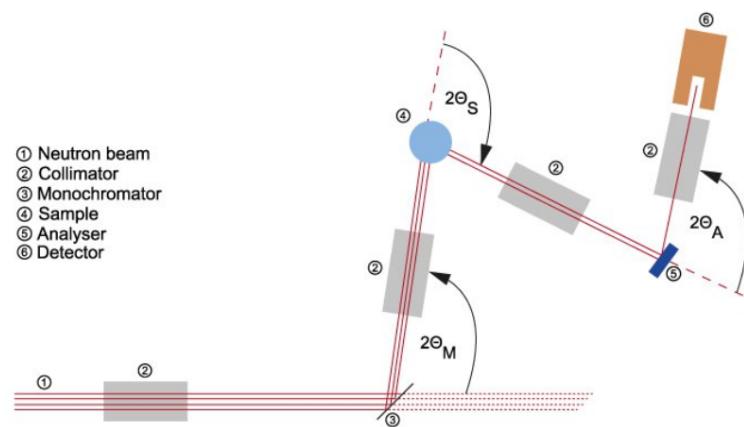
Moreover, this technique allows us to find all the possible scattering directions for the studied sample and the intensity of the scattered wave can be plotted as a function of the  $\theta$  angle (see Fig.8.8). The elastic neutron scattering technique was used for the first time to resolve the AFM magnetic structure of the already cited manganese oxide  $MnO$ . The invention of the technique and the study of  $MnO$  was worth a Nobel in 1994 to Clifford Shull and Bertram Brockhouse. To study  $MnO$  they knew from the susceptibility value that it was an antiferromagnet and they also knew the Néel temperature ( $T_N = 116K$ ). They cooled down the system thanks to liquid nitrogen and then studied it with neutron diffraction.

## 8.4 INS and RIXS

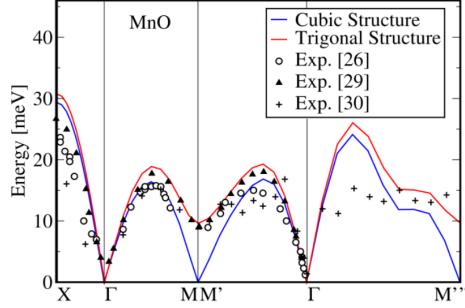
We know that in magnetic materials an excitation of the system can lead to a perturbation of the spins and therefore to a spin wave quantized as a magnon. Hence, to measure magnon dispersion relation we need somehow to perturb our system and to do so we use Inelastic Neutron Scattering (INS) or Resonant Inelastic X-ray Scattering (RIXS). In the INS, differently from the elastic case, the interaction between the spin of the neutron and the one of the atom perturb the system and creates a magnon. The RIXS, instead, is a quite new technique which exploits photons to create spin excitations by exploiting the S-O interaction.

### 8.4.1 Inelastic Neutron Scattering

INS is similar to neutron elastic diffraction, but here there is a net transfer of energy from the scattering particle to the sample. The intensity is non-zero far from Bragg peaks (the ones corresponding to  $2\theta$ ) and the neutron angular momentum can be transferred to create a magnon. Once neutrons are produced and selected into a monochromatic beam, they are sent towards the sample and after the inelastic scattering, the beam is not monochromatic anymore and to compute the energy loss corresponding to the different wavelengths an appropriate analyser is used. The



**Figure 8.9:** Three-axis spectrometer for INS.

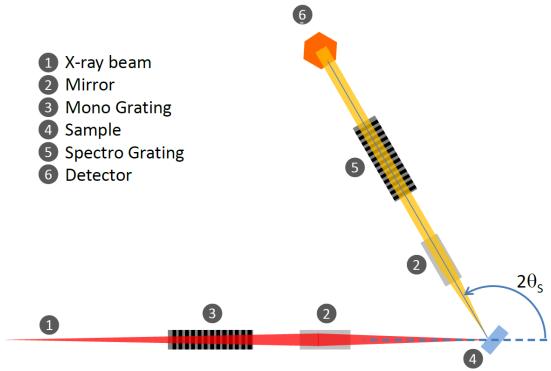


**Figure 8.10:** Energy as a function of the  $k$  vector for  $MnO$ .

experimental apparatus used for the INS is the so-called three-axis spectrometer shown in Fig.8.9. To move the axis of the spectrometer and therefore change the incident angle between the beam and the sample, an air-compressed system is used. The spectrometer is placed onto a marble floor and a flow of compressed air lifts the axis of the spectrometer a few micrometers from the ground to make it able to move without any friction. Fig.8.10 shows the result for an INS experiment on a  $MnO$  sample. We can recognize the typical sinusoidal behaviour of the magnon dispersion relation for an AFM material. Moreover, at the  $\Gamma$  point of the reciprocal lattice, for instance, we can see that the energy is zero which means that there is no energy loss in the scattered neutrons. This means that at these zero-energy points, also called Bragg points, the neutron is scattered elastically, *i.e.* it is scattered as in the Bragg diffraction. Furthermore, knowing the energy at the peaks, for example, between the  $\Gamma$  and the  $MM'$  one, makes it possible to retrieve a lot of information about the exchange integral  $J$  as we can deduce by looking at Eq.(7.10).

#### 8.4.2 Resonant Inelastic X-ray Scattering

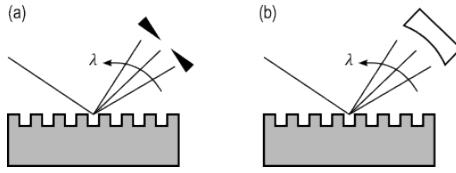
Despite being a very efficient technique, INS in some cases has some important issues. In fact, for some particular materials, such as the  $YBa_2Cu_3O_{6.15}$ , the energy of the peaks resulting from INS is really low and the experimental error associated with them is high, resulting in a non-accurate result. In these cases, RIXS is the best option to investigate the magnetic excitation in a material. The experimental setup for this technique is quite similar to the INS one since for RIXS a kind of three-axis spectrometer is used as well. It is worth noting that in this setup both a Mono Grating



**Figure 8.11:** Three-axis spectrometer for RIXS.

and a Spectro Grating are used. These two instruments exploit the same principle to do two opposite things. They both use a diffraction grating to spatially separate the incident light on the many different wavelengths. The first, the Mono Grating, selects only one of these wavelengths and makes the beam monochromatic and the other one, the Spectro Grating, is used to separate the beam coming out from the sample into all its spectrum to make it analyzable. The two working principles are shown in Fig.8.12. Let's now try to understand what happens in the RIXS for the case of copper  $Cu$  which has an electronic configuration  $[Ar]3d^{10}4s^1$ . Here the energy of the incident x-rays is appropriately chosen to excite the  $2p_{\frac{3}{2}}$  core level<sup>4</sup> (930eV x-rays) of the atom to

<sup>4</sup>In spectroscopy terms the  $2p_{3/2}$  is also called  $L_3$  level, notation that is often used for Auger spectroscopy.

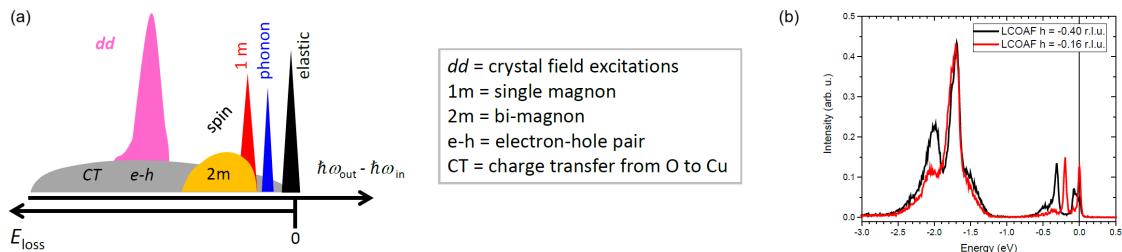


**Figure 8.12:** (a) Mono Gratin working principle and (b) Spectro Gratin working principle.

bring the electron on the  $3d$  level. The general picture of the physics phenomenon can be seen as if the incident photons transfer to the atom one unit of angular momentum ( $\Delta S = 1$ ). Consequently, the atom will pass from a magnetic state to a non-magnetic state for a small amount of time, *i.e.* the lifetime of that particular excited state. It is like if a temporary non-magnetic impurity is placed into the lattice which can cause a perturbation and therefore creates a magnon. The actual phenomenon which occurs is far more complicated and can be seen both from the energetic point of view and from the spin point of view.

- From the energetic point of view, we have that, when the atom is excited, *i.e.* the transition  $2p_{3/2} \rightarrow 4s^1$  occurs, the atom completes the shell and fastly decays through Auger decay into the  $3d$  level. The emitted photons have an energy equal to the energy difference between the initial and the final state and by detecting those photons it is possible to retrieve this energy. The point is that this energy gives a lot of information about the structure of the sample. In fact, the energy depends on the crystal field because the orbitals of the  $Cu$  atoms in the crystal are deformed by the Jahn-Teller effect and thus the energy levels are shifted depending on the crystal field. Knowing the energy gives us the possibility to learn a lot about the structure of the atoms.
- From the spin point of view, instead, a lot depends on the spin-orbit interaction. In the initial state  $2p_{3/2}$ , in fact, the S-O coupling is really strong and the spin quantum number is not a good quantum number. Therefore, the spin state is not well defined, *i.e.* it is a superposition of spin-up and spin-down states. On the contrary, the final state  $4s$  has a low S-O coupling and it is a well-defined spin state. However, we need to keep into account that the system will relax through Auger emission which means that the electron on the  $4s$  will relax in another surface level which in this case is the  $3d$ . For the spin conservation law, the spin must be conserved during the transition, both on the excitation and on the relaxation of the system. Therefore, depending on the spin "mixing" on the core level and on the well-defined spins of the surface levels, it may happen that during the relaxation  $4s \rightarrow 3d$  the spin one electron can flip. If this happens we have a magnetic excitation and therefore a magnon.

The spectrum of the emitted radiation can be detected and analysed quite easily. However, we need to keep into account that a huge variety of phenomena occur during the light-matter interaction. The spin-flip indeed can either happen or not depending on the mixing of states and on the energy of the photons. Moreover, lattice excitations, *i.e.* phonons, can be "switched on", elastic scattering can occur, two magnons can be excited, and if the sample is not completely pure but is oxidised it can happen that some charge flows from  $Cu$  atoms to oxygen ones and so on. This variety of phenomena is detected on the spectrum and we can see them in the schematic representation of Fig.8.13(a) and on the real spectrum of Fig.8.13(b).



**Figure 8.13:** Schematic representation of a RIXS experiment spectrum.

# Chapter 9

## Magnetism in Metals

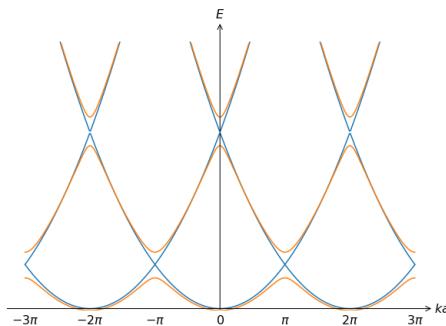
So far we mostly concentrate on the study of the magnetic phenomena due to atoms and their interactions (exchange interaction, super-exchange, etc.) actually due to the interactions between electrons belonging to the atoms themselves. This description works particularly well for those atoms in which the electrons in the open shells have a localized nature (*e.g.* oxides) such as the electrons in the  $d$  or the  $f$  shells. In this chapter we are going to study the magnetic phenomena due to those atoms in which the electrons are not tightly bonded but, instead, are described by a "quasi-free" particle model like, for example, metals. When these kinds of atoms are arranged in a crystalline solid the electrons are delocalised and can travel freely in the whole volume of the material and the lattice, where the atoms displaced in a periodic way, represent only a perturbation for the free electrons. Let's start our discussion by reviewing the free (or quasi-free) electron model.

### 9.1 Quasi-Free Electron Model

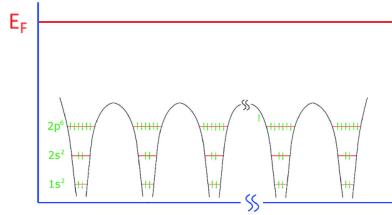
The quasi-free electron model describes those electrons which are loosely bonded to their original atoms and therefore are free to travel in the lattice. This work particularly well for material like metals for instance. Let's start considering the case in which electrons are completely free, *i.e.* there is no perturbative potential due to the lattice. In this case, the confinement of the electrons in the lattice can be represented by a "big potential box" which represents the whole lattice. The solution of the Schrodinger equation for these free electrons is represented by plane waves whose kinetic energy is given by:

$$E_K = \frac{\hbar^2 k^2}{2m} \quad (9.1)$$

The energy levels of the electrons in the box are quantized and, if we consider that the lattice is made by  $N$  atoms and that each atom contributes with only one electron, we will have that the number of occupied levels is  $\frac{N}{2}$ , *i.e.* two electrons per level. Since  $N$  is usually a huge number ( $\sim 10^{23}$ ) we can consider that these discrete levels give rise to a continuum of energy levels which is the band structure of a material. The band is indeed filled until the  $\frac{N^{th}}{2}$  level which corresponds to the Fermi level which has an energy  $E_F$  (Fermi energy). The dispersion relation  $E = E(k)$  of these free electrons is the one given in Eq.(9.1) and it is represented by a parabola. If we now introduce



**Figure 9.1:** Dispersion relation for the quasi-free electron model.



**Figure 9.2:** Periodic potential due to the lattice and Fermi energy of the electrons.

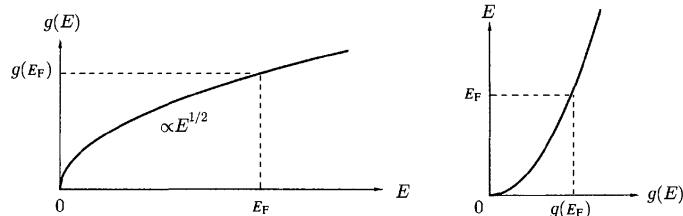
the presence of a periodic potential due to the presence of the atoms displaced in the lattice in a periodic way we will have a perturbation in the energy of the electrons. Electrons are indeed influenced by the potential energy of the atoms and their dispersion relation changes becoming the one in Fig.9.1 which consists in a parabola repeated in the  $k$ -space with the periodicity of the reciprocal lattice. The electrons are not completely free anymore but, since the periodic potential of the atoms is far smaller than the Fermi energy (see Fig.9.2), we can consider them as quasi-free electrons and add the potential as a perturbative term. For a parabolic dispersion relation, it is quite easy to compute the Density of States (DOS)  $g(E)$  for a 3D system which is:

$$g(E) = \frac{dn}{dE} = \left( \frac{2m_e}{\hbar^2} \right)^{\frac{3}{2}} \frac{1}{2\pi^2} \sqrt{E} \quad (9.2)$$

As usual,  $g(E)$  is the number of accessible states for a given energy. The DOS is often used to compute the number of states at the Fermi energy  $E_F$  that, as we will see, is the important parameter for the magnetisation of metal materials because it influences the distribution of electrons in the bands and therefore the overall spin distribution. At  $E = E_F$  the DOS is:

$$g(E_F) = \frac{3}{2} \frac{n}{E_F} \quad (9.3)$$

with  $n$  total number of electrons. Fig.9.3 shows two possible ways of representing the function  $g = g(E)$ . The "tilted" way, in which the energy  $E$  is on the vertical axis<sup>1</sup> is a quite useful



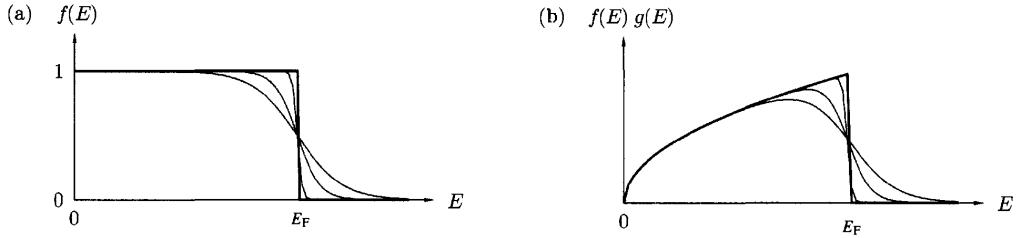
**Figure 9.3:** DOS representations.

representation which is often used in solid-state physics. What has been told and seen so far is valid at zero temperature where the electrons fill all the levels up to the Fermi energy, *i.e.* the electrons distribution is a step function which is worth 1 for  $E < E_F$  and zero otherwise. When  $T \neq 0$  the shape of the bands itself does not change but what changes is the way in which electrons are distributed along the bands. When  $T \neq 0$  we need to consider the appropriate distribution which is the Fermi-Dirac distribution  $f(E)$  which expresses the probability of a given energy level to be occupied and it is:

$$f(E) = \frac{1}{e^{\frac{E-\mu}{k_B T}} + 1} \quad (9.4)$$

where  $\mu$  is the chemical potential which we will consider equal to the Fermi energy ( $\mu \simeq E_F$ ) which is, in general, a quite good approximation. In general, the number of electrons  $dn$  in a small interval of energy  $dE$  can be computed as the product of the number of available states at that energy ( $g(E)$ ) times the probability of that state to be occupied ( $f(E)$ ). Therefore, the total

<sup>1</sup>The shape of  $g(E)$  in this tilted representation is obviously a parabola as well as the dispersion relation  $E(k)$ . Be careful to not confuse these two graphs: they are both parabolas but with completely different meanings.



**Figure 9.4:**  $f(E)$  and  $n(E) = f(E)g(E)$  functions.

number of electrons up to at a given energy  $n(E)$  will be:

$$n(E) = \int_0^E g(E)f(E)dE \quad (9.5)$$

The shape of  $f(E)$  and the one of  $n(E)$  are represented in Fig.9.4. This is it for what concerns the review of the quasi-free electron model and we can now deal with the first magnetic phenomenon due to electrons in metals which is the Pauli Paramagnetism

## 9.2 Pauli Paramagnetism

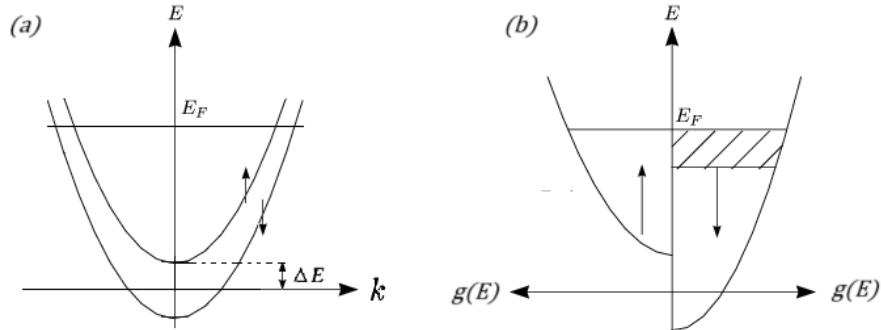
Let's start by applying a magnetic field to our quasi-free electrons. Since they are quasi-free we will assume that no exchange interactions occur between them. Let's consider an external magnetic field  $B_{ext}$  direct along  $z$ . The direction of the field makes natural the choice of the quantization axis for the system which will be the  $z$  axis as well. We will consider "spin-up" those electrons with spin parallel to the field and "spin-down" the ones with spin antiparallel to the field. Since the energy exchange between the spin and the field is:

$$E = g_S \mu_B m_S B_{ext} = \pm \mu_B B_{ext}$$

since  $m_S = \pm \frac{1}{2}$  and  $g_S = 2$ . Therefore, the electron will have a gain or a loss of energy of:

$$\Delta E = \mu_B B_{ext} \quad (9.6)$$

depending on their spin. The spin-up electrons will increase their energy of  $\Delta E$  while the spin-down ones will decrease their energy of the same amount. The energy as a function of the  $k$ -vector for both spin up and spin down electrons is depicted in Fig.9.5(a) while Fig.9.5(b) shows the behaviour of the DOS for the spin-up electrons on the right and the spin-down electrons on the left. In general,  $\Delta E \ll E_F$  and we can thus consider that the Fermi energy is not affected by this distortion of the bands and that it is the same for both types of electrons. Then, the electrons on the lowered parabola have a bigger interval of  $k$ -values available, *i.e.* have more states available and thus more electrons. Energetically, the lowering in energy of the spin-down electrons left some available and empty states below the Fermi level and, on the contrary, the same amount of state above the Fermi level will be occupied by the spin-up electrons because of their increase in energy. It is



**Figure 9.5:** Dispersion relations and DOS for the two electrons' populations.

energetically favourable for these latter electrons above the Fermi level to fill the empty states in the spin-down band by changing their spin. The spin-down electrons' population increases. There will be a **majority** state (the  $\downarrow$  parabola) and **minority** state (the  $\uparrow$  parabola). The unbalance in the electrons' population with opposite spin will cause a net magnetization of the system which gives rise to the so-called Pauli Paramagnetism. We can consider that the number of electrons which vary their energy are the ones close to the Fermi energy in an energy interval  $\Delta E$ . Hence, by considering that half of the electrons are spin up and the other half are spin down, we can write the variation in the electrons' population with opposite spin as<sup>2</sup>:

$$\begin{aligned}\Delta n^\downarrow &= \frac{1}{2}g(E_F)\Delta E \\ \Delta n^\uparrow &= -\frac{1}{2}g(E_F)\Delta E\end{aligned}\tag{9.7}$$

where  $g(E_F)$  is the DOS at the Fermi energy. We can therefore write the population of electrons with spin up and down as:

$$n^{\downarrow,\uparrow} = n + \Delta n^{\downarrow,\uparrow}$$

where  $n$  is the total number of electrons with no field. Then the net magnetization of the system is:

$$\begin{aligned}M &= \mu_B(n^\downarrow - n^\uparrow) = \mu_B(\Delta n^\downarrow - \Delta n^\uparrow) = \mu_B g(E_F) \mu_B \Delta E = \\ &= \mu_B g(E_F) \mu_B B_{ext} = \mu_B^2 g(E_F) B_{ext}\end{aligned}$$

We have obtained the contribution of the Pauli paramagnetism to the magnetization of a PM metal which in conclusion is:

$$M = \mu_B^2 g(E_F) B_{ext}\tag{9.8}$$

We notice that, as in the PM case, the magnetization is linear with the magnetic field but, unlike before, the magnetization seems to have no dependence on the temperature. In fact, we should remember that in PM materials the magnetization does not depend on the temperature only for a small value of the external field (see Fig.3.2) and otherwise it depends on the temperature through the Brillouin function. Moreover, the Pauli susceptibility  $\chi_P$  is:

$$\chi_P = \frac{M}{H} = \mu_0 \frac{M}{B_{ext}} = \mu_0 \mu_B^2 g(E_F)\tag{9.9}$$

Also here we notice the no dependence on the field and on the temperature which results in the impossibility of writing an analogous of the Curie law for the Pauli paramagnetism. Actually, it is really strange that the susceptibility does not depend on the temperature. On one hand, we used the approximation  $\Delta E \ll E_F$  which implies that we are working on the "linear" zone of the magnetization. In reality, this approximation is really appropriate since having  $\Delta E \approx E_F$  is impossible because usually in metals  $E_F \sim eV$  and to obtain such values for  $\Delta E$  we would need a non-realistic magnetic field. This can not be the problem. On the other hand, to properly consider the temperature dependence we need to include the Fermi distribution in the computation of the populations  $n^\downarrow$  and  $n^\uparrow$ . Let's do this. By using the Fermi distribution we get:

$$\begin{aligned}n^\downarrow &= \frac{1}{2} \int_0^\infty g(E + \Delta E) f(E) dE \\ n^\uparrow &= \frac{1}{2} \int_0^\infty g(E - \Delta E) f(E) dE\end{aligned}\tag{9.10}$$

Thus the magnetization becomes:

$$\begin{aligned}M &= \mu_B(n^\downarrow - n^\uparrow) = \frac{\mu_B}{2} \int_0^\infty [g(E + \Delta E) - g(E - \Delta E)] f(E) dE = \\ &= \frac{\mu_B}{2} \int_0^\infty \left[ \frac{g(E + \Delta E) - g(E - \Delta E)}{2\Delta E} \right] 2\Delta E f(E) dE = \\ &= \mu_B \Delta E \int_0^\infty \frac{dg}{dE} f(E) dE = \mu_B^2 B \left\{ [gf]_0^\infty - \int_0^\infty \frac{df(E)}{dE} g(E) dE \right\}\end{aligned}$$

---

<sup>2</sup>Since  $\Delta E \ll E_F$  we can approximate  $g(E^\downarrow) \approx g(E^\uparrow) \approx g(E_F)$

In which the last equivalence is obtained by integrating by parts. The first term of the integration vanishes because:

$$[gf]_0^\infty = g(0)f(0) - g(\infty)f(\infty) = 0$$

since  $g(0) = 0$  and  $f(\infty) = 0$ . Thus the magnetization takes the form:

$$M = -\mu_B^2 B \int_0^\infty \frac{df(E)}{dE} g(E) dE \quad (9.11)$$

Now we can distinguish two cases:

- If  $T = 0$  the Fermi distribution is a step function and its derivative is the Dirac Delta  $\delta(E - E_F)$  then the magnetization becomes:

$$M = \mu_B^2 B g(E_F) \quad (9.12)$$

which is the result obtained previously in Eq.(9.8).

- If  $T \neq 0$  the Fermi distribution has the form of Eq.(9.4) but, if  $k_B T \ll E$  ( $T$  small), it can be approximated with the Maxwell-Boltzmann distribution:

$$f(E) \simeq e^{-\frac{E-E_F}{k_B T}} \quad (9.13)$$

whose derivative is  $\frac{df(E)}{dE} = -\frac{1}{k_B T} f(E)$  and then we obtain a magnetization which is:

$$M \simeq \frac{\mu_B^2 B}{k_B T} \int_0^\infty f(E) g(E) dE = \frac{\mu_B^2 B}{k_B T} n \quad (9.14)$$

We managed to find a temperature dependence on the magnetization but, unfortunately, this result is not appropriate since we approximate the Fermi distribution with the Boltzmann one. This approximation, which is also called non-degenerate limit, is valid only in the cases in which the Fermi energy is really close to the valance band which is not the case for materials like metals. Therefore, we need to take this result with caution because in real material the proper value of the magnetization depends a lot on the shape of the band, which varies case by case. In conclusion, we can say that for Pauli paramagnetism for a metal, the magnetization is weakly dependent on the temperature.

### 9.3 Stoner Ferromagnetism

At the beginning of the previous section, we specified that we wouldn't have considered any kind of interaction between electrons because of their delocalized nature. We want now to consider a possible interaction among electrons by introducing a molecular field  $B_{mf}$  (or mean-field) similar to the one used in the Weiss model for ferromagnetism. In this case, we will not consider any external field applied to the system. The origin of this field is not discussed in the following because it is out of the scope of this course. As done during the discussion of the Weiss model we can choose a linear form for the molecular field with respect to the magnetization  $M$  of the system, which means:

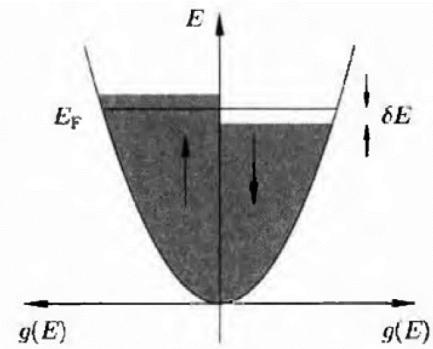
$$B_{mf} = \lambda M \quad (9.15)$$

The molecular field acts on the bands in the same way the external field did before. In this case, since the distortion of the bands is much smaller than before, we can compute the energy gain for the system by considering that a small number of electrons from the spin-up band are moved to the spin-down one as depicted in Fig.9.6. In particular, we will consider that a small quantity of electrons  $\delta n$  is "taken" from the spin-up band right below the Fermi energy and it is placed right above  $E_F$  in the spin-down band. We call  $\delta E$  the difference in energy from the energy of the electrons to the Fermi level. Since the electrons are considered to be free, the only energy cost is a kinetic term  $\Delta E_K$  which is:

$$\Delta E_K = \delta n \delta E = \frac{1}{2} g(E_F) \delta E \delta E \quad (9.16)$$

where  $\delta n$  is computed as in the Pauli paramagnetism case. This moving leads to a magnetization  $M$  which is:

$$M = \mu_B (n^\downarrow - n^\uparrow) = \mu_B g(E_F) \delta E \quad (9.17)$$



**Figure 9.6:** DOS spontaneous splitting.

because the two populations are:

$$n^\uparrow = n - \frac{1}{2}g(E_F)\delta E$$

$$n^\downarrow = n + \frac{1}{2}g(E_F)\delta E$$

We can now compute the gain in terms of potential energy due to the magnetization that is:

$$\Delta E_P = - \int_0^M \mu_0 B_{mf} dM' = - \int_0^M \mu_0 \lambda M' dM' = -\frac{1}{2} \mu_0 \lambda M^2 = -\frac{1}{2} \mu_0 \lambda [\mu_B g(E_F) \delta E]^2 \quad (9.18)$$

Finally, the energy balance is given by the sum of the kinetic cost and the potential gain:

$$\Delta E = \Delta E_K + \Delta E_P = \frac{1}{2}g(E_F) [1 - \mu_0 \lambda \mu_B^2 g(E_F)] (\delta E)^2 \quad (9.19)$$

We can have a spontaneous magnetization of the system only if the overall energy variation is negative. Since  $g(E_F)$  and  $(\delta E)^2$  are positive, the only way in which  $\Delta E < 0$  is:

$$\Delta E < 0 \Rightarrow \mu_0 \lambda \mu_B^2 g(E_F) > 1$$

We can define  $U = \mu_0 \mu_B^2 \lambda$  which is called Coulombain energy and the condition becomes:

$$U g(E_F) > 1 \quad (9.20)$$

which is known as the **Stoner criterion** for ferromagnetism. If  $U g(E_F) > 1$  the material can be a ferromagnet and has a spontaneous magnetization.

### 9.3.1 Stoner Enhancement

Let's look at Tab.9.1 which reports the value of the molar susceptibility for some relevant metals. First of all, we notice that noble metals (*Cu*, *Ag* and *Au*) have all weak diamagnetic behaviour.

**Table 9.1:** Molar Magnetic Susceptibility ( $\chi$ ) of Elements

Element	$\chi_m$ ( $10^{-10} m^3 mol^{-1}$ )
Li	1.78
Na	1.7
K	2.62
Cu	-0.68
Ag	-2.45
Au	-3.51
Bi	-36

This is because they all have a complete  $d$  shell (which as we have seen is the "magnetic shell", *i.e.* the one responsible for the magnetic properties) and thus they do not possess a magnetic moment. These metals are particularly interesting for various reasons. For example, while  $Cu$  is not magnetic its ion  $Cu^+$  has a strong AFM behaviour and it is fundamental in the already-cited cuprates. Gold, instead, is very famous for its metallic properties and for being a totally inert element. In fact, it does not react with almost anything (turpentine is the only acid that can attack gold), it does not oxidize and for this reason, it has been one of the most valuable metals since ancient times<sup>3</sup>. Another interesting metal is Bismuth ( $Bi$ ) which is one of the heaviest non-radioactive elements and exhibits a strong DM behaviour. We want now to try to find out why some simple metals behave like paramagnets while others like diamagnets. Is there a way in which we can predict the magnetic properties of a given metal?

Let's start considering the Stoner criterion of Eq.(9.20). We have that a metal is not ferromagnetic if:

$$Ug(E_F) < 1$$

In particular, we don't want the term  $Ug(E_F)$  to be zero so we will consider  $0 < Ug(E_F) < 1$ . This means that we will consider the case in which the interaction  $U$  is not big enough to give the metal a ferromagnetic character and, at the same time, being  $U$  different from zero, it can bring some magnetic behaviour. Moreover, we will consider also an external magnetic field  $B_{ext}$  which introduces in the energy form of Eq(9.19) a new energy term which is:

$$\Delta E = \frac{1}{2}g(E_F)[1 - Ug(E_F)](\delta E)^2 - MB_{ext} \quad (9.21)$$

From Eq(9.17) we can express the energy as a function of the magnetization and substitute it into Eq.(9.21). This gives rise to:

$$\delta E = \frac{M}{\mu_B g(E_F)} \Rightarrow \Delta E = \frac{M^2}{2\mu_B^2 g(E_F)} [1 - Ug(E_F)] - MB_{ext}$$

We can minimize this energy and we find:

$$\frac{\partial \Delta E}{\partial M} = 0 \Rightarrow M = \frac{\mu_B^2 B_{ext} g(E_F)}{1 - Ug(E_F)} \quad (9.22)$$

The resulting susceptibility  $\chi$  is:

$$\chi = \mu_0 \frac{M}{B_{ext}} = \frac{\mu_0 \mu_B^2 B_{ext} g(E_F)}{1 - Ug(E_F)} = \frac{\chi_P}{1 - Ug(E_F)} \quad (9.23)$$

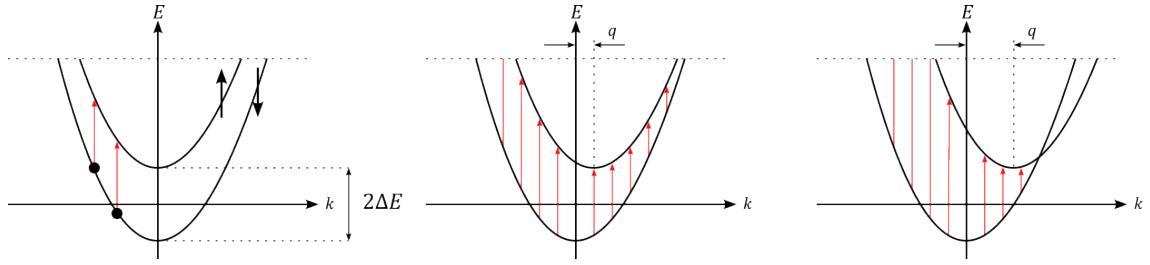
We can see from this equation that the term  $Ug(E_F)$  acts for the susceptibility as a "booster" in the sense that the bigger the term, the bigger the susceptibility. This enhancement for the susceptibility is the so-called **Stoner enhancement**. Also, depending on the case the value of  $Ug(E_F)$  can be smaller or bigger but in some cases, it can reach values close to one and then the value of the susceptibility of the system is strongly increased. Although this susceptibility expression can explain the strong PM behaviour of some metals it does not explain the DM behaviour. To explain this phenomenon we need to analyze the so-called Landau Diamagnetism. Before doing that, let's finish the study of the Stoner ferromagnetism looking at what happens when a magnetic system of this type is perturbed by a spin wave.

### 9.3.2 Stoner Excitations

Let's now consider the ferromagnetic case in which the Stoner criterion is fulfilled which means  $Ug(E_F) > 1$ . We want to try to describe how spin waves propagate in such a metal material. We know that spin waves are caused by a perturbation, *i.e.* an excitation, in the magnetization of the system. In metals, an excitation can be seen as an electron which moves from a low energy level to a higher one creating, as a consequence, a hole in the starting level. In a "normal" metal in which the bands are not split during this excitation the spin doesn't flip because of the conservation of

---

<sup>3</sup>For precision's sake it must be specified that recently scientists have found that nano-clusters of gold atoms can exhibit a small FM behaviour due to some surface effect which occurs when the clusters is placed at the interface with vacuum or other materials.

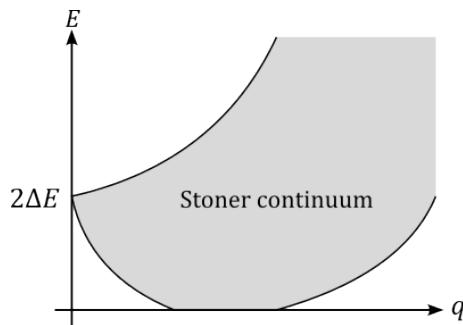


**Figure 9.7:** Energy dispersion relations representation of the spin-up and spin-down bands for different values of  $q$ .

the spin quantum number but, in a metal ferromagnet in which the bands are spin-split by the molecular field, it can happen that an electron pass from the spin-up band to the spin-down one or vice versa. In such a case, the magnetization of the system varies because it varies the number of particles in the sub-bands and this implies that we will have a magnetic excitation. These kinds of excitations are called **Stoner excitations**. We can try to qualitatively understand the dispersion relation of this phenomenon by considering how the energy required for the intra-band transition varies depending on the transferred momentum  $\mathbf{q} = \mathbf{k}' - \mathbf{k}$  where  $\mathbf{k}$  is the momentum in the initial band and  $\mathbf{k}'$  the one in the final band. Let's refer to Fig.9.7. If  $q = 0$  (Fig.9.7(a)) the spin-up and down bands are perfectly aligned and the electron transition is vertical. The energy required to go from the spin-down to the spin-up band in this case is well known since it is two times the energy splitting between the bands  $\Delta E$  and it is constant for the whole extension of the band. If  $q > 0$  we can imagine that parabolas representing the dispersion relations are translated one with respect to the other of a quantity  $q$  since  $k = q - k'$  (Fig.9.7(b)). In this case, there will be a continuum of possible required energies for the electron since it can make the transition to the upper band from any point of the lower band. If the value of  $q$  continues to increase there will be a point at which the two parabolas overlap (Fig.9.7(c)), *i.e.* the energy cost to make the transition is zero. This zero cost for the transition will last as long as the two parabolas overlap one with the other at any point. For very large values of  $q$  the parabolas will split apart once again and the energy cost to make the transition will start to increase again. What we have told so far can make us understand the dispersion relation for Stoner excitation represented in Fig.9.8. As reported in the figure, this continuum of energy values is called Stoner continuum.

## 9.4 Landau Diamagnetism

It is now the moment to study where the diamagnetic behaviour of some metals comes from. We will deal with this topic in a very qualitative way and we will not demonstrate the result since a full discussion of the topic will require a lot of time that, unfortunately, we do not have. As we said, in a metal free electrons are described by a Hamiltonian which contains only a kinetic term and their wave functions are simple plane waves. Nevertheless, if we apply an external magnetic field  $B_{ext}$  to the free electrons they will be influenced by it and in the Hamiltonian the canonical momentum operator  $(-i\hbar\nabla)$  will be replaced with the momentum of charge particle into a magnetic field  $(-i\hbar\nabla + e\mathbf{A})$ . This substitution gives rise to a quadratic Hamiltonian which has quantized energy



**Figure 9.8:** Dispersion relation for Stoner excitations.

eigenvalues, *i.e.* discrete energy levels. Therefore, the application of a magnetic field causes the electron to form discrete energy levels which are the so-called **Landau levels**. The idea is that as a magnetic field is applied, and the electron distribution breaks up into a series of Landau levels, the total energy of the system may change. This change in energy with the field is equivalent to a magnetization of the system. This phenomenon is known as **Landau diamagnetism**. Although the mathematical discussion of the phenomenon being rather complicated, the result is surprisingly simple and it is possible to demonstrate that the Landau diamagnetic susceptibility  $\chi_L$  is:

$$\chi_L = -\frac{1}{3}\chi_P = -\frac{1}{3}\mu_0\mu_B^2g(E_F) \quad (9.24)$$

Although successful this result does not explain the diamagnetic behaviour of a metal. In fact, we need to consider both contributions to the magnetic properties, the Pauli paramagnetism and the Landau diamagnetism, resulting in a total susceptibility of:

$$\chi = \chi_L + \chi_P = \frac{2}{3}\chi_P > 0 \quad (9.25)$$

This apparent problem is quickly resolved by considering that an electron inside a material, even if a metal, is not a free electron. Within a material, in fact, the electrons interact with a lot of particles and quasi-particles (*e.g.* phonons) and this interaction can be taken into account in our model by substituting the mass of the free electron  $m_e$  with its effective mass  $m^*$  which can be greater or smaller than  $m_e$ . If we include this effective mass in the model the form of the Landau diamagnetic susceptibility slightly changes:

$$\chi_L = -\frac{1}{3} \left( \frac{m_e}{m^*} \right)^2 \chi_P \quad (9.26)$$

Hence, if  $m^* < m_e$  by a certain amount, we can imagine that the diamagnetic contribution will become dominant on the paramagnetic one and therefore we will have an overall diamagnetic behaviour. At the beginning of our discussion, we said that Bismuth has a strong diamagnetic behaviour. For this metal, in fact, the effective mass is  $m^* = 0.01m_e$  and therefore the diamagnetic susceptibility is:

$$\chi_L = -\frac{1}{3} \left( \frac{m_e}{0.01m_e} \right)^2 \chi_P = -\frac{1}{3} 10^4 \chi_P$$

This is a clear example of a metal in which the diamagnetic contribution is much stronger than the paramagnetic one. In conclusion, one might wonder why we don't need to include the contribution of the effective mass on the Pauli paramagnetism. This is actually already done considering that the effective mass enters both in the Pauli and Landau terms through the density of state at the Fermi level  $g(E_F)$ . Despite that, in the Landau term, there is an additive contribution of the effective mass that can be understood by imaging that also the classical "circular orbits" of electrons placed into a magnetic field are affected by the effective mass.

# Chapter 10

## Bose-Einstein Condensates

In this and the following chapters we are going to introduce and discuss three different phenomena which are:

- Bose-Einstein condensates (BEC).
- Superfluidity of Helium.
- Superconductivity.

In summary, a Bose-Einstein condensate is a condensation of a system of non-interacting bosons in their lowest energy state. Superfluidity is the absence of any viscosity in the motion of a liquid while superconductivity is the absence of any resistance in the motion of a current in a solid. Despite being three different phenomena, they have some interesting characteristics in common. All of them are a particular "state" of matter which appears only below a certain critical temperature, they are related to a certain "order" and can therefore be seen as a phase transition and they have no classical explanation but are linked to the very quantum essence of nature. Although usually quantum mechanics manifest itself on the microscopic scale, Bose-Einstein condensation, superfluidity and superconductivity are macroscopic manifestations of quantum mechanics. The history of these three phenomena is tightly linked to low-temperature physics and to the development of low-temperature technologies.

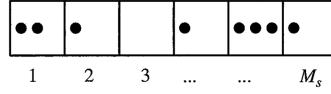
### 10.1 Historical Introduction

During the 20<sup>th</sup> century the huge progress in the achievement of low-temperature experimental conditions made possible the study of these phenomena. To have an idea of the importance of these topics, it is sufficient to look at Tab.10.1 which reports some relevant events in the BEC, superfluidity and superconductivity history. It is interesting to see that BEC, differently from the other two phenomena, were predicted theoretically in 1925 but we need to wait 70 years, until 1995, to see their experimental realization. On the contrary, superconductivity was observed for the first time in 1911 and it was a really puzzling observation but it took more than 40 years to have a quite efficient theory of superconductivity (BCS theory). Superfluidity had some issues during its development as well. There was a hint of this strange behaviour for  ${}^4He$  in 1927 but the effective observations happened only 11 years later, in 1938. Nowadays, BEC and superconductivity are two of the most attractive fields of scientific research. BEC research aims to manipulate this condensate of cold atoms and simulate phenomena with them that wouldn't be realizable otherwise (*e.g.* it is possible to create artificial lattices with BEC and manipulate them). Superconductivity represents an open problem both from an experimental and a theoretical point of view. This field became richer and richer also after the discoveries of the so-called high-temperature superconductors which are not explained by the already existing theories of superconductivity.

### 10.2 Bose-Einstein Statistic

Let's begin our discussion with BEC which is a quantum phenomenon in which a large number of bosons simultaneously occupy the ground state of a system. Let's consider a system of  $N$  identical

non-interacting bosons and  $M$  identical possible quantum states in which place our particles. We want to find the number of possible ways of displacing the  $N$  bosons in the  $M$  states. We can imagine that the situation is represented by Fig.10.1. We thus have  $N$  particles and  $M - 1$  separation walls



**Figure 10.1:**  $N$  bosons in  $M$  quantum states.

and, since particles and walls are identical, the number of ways in which the particles can be distributed  $W_s$  is:

$$W_s = \frac{[N + (M - 1)]!}{N!(M - 1)!} \quad (10.1)$$

which is the expression of the permutation of  $N + (M - 1)$  elements with repetitions of  $N$  and  $(M - 1)$  of them. It is possible to apply this combinatory rule to the thermodynamics of an ideal gas of  $N$  bosons occupying a volume  $V$ . This procedure, whose passages are not reported here<sup>1</sup>, can be used to compute the average occupation of a given state, *i.e.* the average number of particles per state, and the result is the well-known Bose-Einstein distribution  $f_{BE}(\varepsilon)$ :

$$\frac{N}{M} = f_{BE}(\varepsilon) = \frac{1}{e^{\beta(\varepsilon - \mu)} - 1} \quad (10.2)$$

where  $\beta = \frac{1}{k_B T}$  and  $\varepsilon$  is the energy of the state. Let's now try to see what happens in the case of a two-state system  $M = 2$  in which we have a ground state  $\varepsilon_0 = 0$  and an excited state  $\varepsilon_1 = \varepsilon$ . We want to compute the average occupation number of the ground state  $\langle N_{GS} \rangle$  both for a classic Maxwell-Boltzmann (MB) distribution and the quantum Bose-Einstein (BE) one. It is possible to demonstrate<sup>2</sup> that:

$$\begin{aligned} \langle N_{GS}^{(MB)} \rangle &= \frac{N}{1 + e^{-\beta\varepsilon}} \\ \langle N_{GS}^{(BE)} \rangle &= \frac{1}{e^{-\beta\varepsilon} - 1} + \frac{N + 1}{1 - e^{-(N+1)\beta\varepsilon}} \end{aligned}$$

where  $N$  is the total number of particles. The behaviour of these two functions for  $N = 100$  is represented in Fig.10.2(a). From the figure, we can see that, for any temperature, in the BE statistic we have a higher probability of the ground state being occupied with respect to the MB statistic. With MB statistics, the temperature has to be very low to get the ground state to have an appreciable filling fraction. This means that for a system of bosons, especially at low temperature, *i.e.*  $k_B T \ll \varepsilon$ , the ground state will be on average more populated. For example, at  $k_B T = \frac{N}{10}\varepsilon \gg \varepsilon$

<sup>1</sup>See James F. Annet - *Superconductivity, Superfluids and Condensates* for the detailed passages.

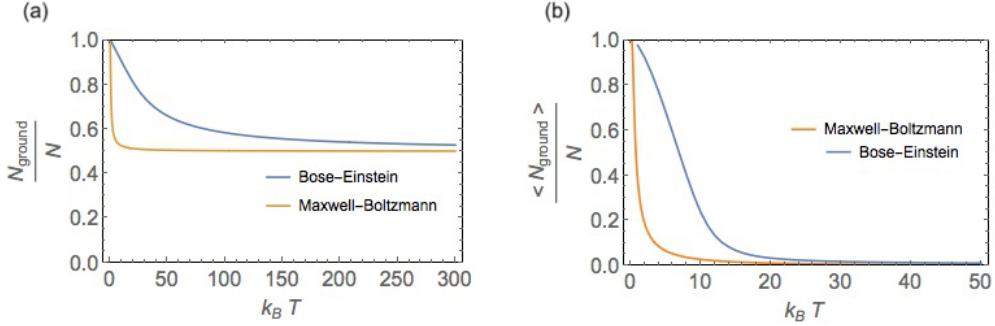
<sup>2</sup>See Matthew Schwartz - *Lecture 12: Bose-Einstein Condensation* for the detailed passages.

**Table 10.1:** Some of the key events in the BEC, superfluidity and superconductivity history.

---

1908	Liquefaction of ${}^4He$ at 4.2 K
1911	Superconductivity discovered in $Hg$ at 4.1 K
1925	Bose-Einstein condensation predicted
1927	Discontinuity of the specific heat found in ${}^4He$ at 2.2 K
1933	Meissner-Ochsenfeld effect observed
1938	Demonstration of superfluidity in ${}^4He$
1950	Ginzburg-Landau theory of superconductivity
1957	Bardeen-Cooper-Schrieffer (BCS) theory
1957	Abrikosov flux lattice
1962	Josephson effect
1963-4	Anderson-Higgs mechanism
1971	Superfluidity found in ${}^3He$ at 2.8 mK
1986	High-temperature superconductivity discovered (30-165 K)
1995	BEC achieved in atomic gases at $0.5\mu K$

---



**Figure 10.2:** (a) Ground state occupation for a two-state system ( $M = 2$ ) and (b) ground state occupation for a general  $M$  states system with  $M$  large.

it is possible to numerically find that 90% of the atoms are in the ground state for BE statistics, but only 52% for MB statistics. In a certain qualitative way, this demonstrates the possibility of Bose-Einstein condensation. Bose-Einstein condensation is a phase transition whereby the ground state becomes highly occupied.

Furthermore, we can see that both statistics reach a limit value for  $\frac{N_{GS}}{N} = \frac{1}{2}$  which means that half the particles of the system will occupy the ground state. When the number of available state  $M$  gets larger this limit value changes. In fact, it can be seen that the limit value is  $\frac{1}{M}$  which is equal to  $\frac{1}{2}$  in the previous case with  $M = 2$  and can be almost zero when  $M$  gets large, *i.e.* no particle at high temperature occupy the ground state of the system (see Fig.10.2(b)).

### 10.3 Bose-Einstein Condensation

The term "condensation" in BEC is used by analogy with the normal liquid-gas phase transition, in which liquid drops condense out of the gas to form a saturated vapour. In the same way, below the critical temperature  $T_C$  in BEC "normal gas" particles coexist in equilibrium with "condensed" particles. But, unlike a liquid droplet in a gas, here the "condensed" particles are not separated in space from the normal particles. Instead, they are separated in the momentum space. The condensed particles all occupy a single quantum state of zero momentum, while the normal particles all have finite momentum. We can now begin our study by system of  $N$  bosons in a volume  $V$  and try to derive the critical temperature  $T_C$  at which the condensate starts to form. By using the BE statistic of Eq.(10.2) we can compute the total number of particle  $N$  as:

$$N = \sum_{\mathbf{k}} \frac{1}{e^{\beta(\varepsilon_{\mathbf{k}} - \mu)} - 1} \quad (10.3)$$

In the thermodynamic limit,  $V \rightarrow \infty$ , the possible  $\mathbf{k}$  values become a continuum and therefore:

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

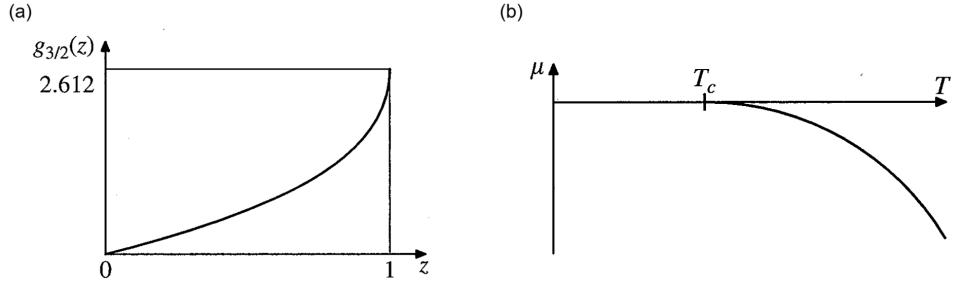
We can thus change the summation with an integration over all the  $\mathbf{k}$ -space and compute:

$$n = \frac{N}{V} = \frac{1}{(2\pi)^3} \int \frac{1}{e^{\beta(\varepsilon_{\mathbf{k}} - \mu)} - 1} d^3k \quad (10.4)$$

We can also change the integration variable and integrate not over the momentum but over the energies by performing the substitution  $(2\pi)^3 g(\varepsilon) d\varepsilon = d^3k$ . This gives us:

$$n = \int_0^{\infty} \frac{1}{e^{\beta(\varepsilon - \mu)} - 1} g(\varepsilon) d\varepsilon \quad (10.5)$$

This equation describes the particle density  $n(\mu, T)$  as a function of the temperature and chemical potential. Actually, in a real case scenario, we know  $n$  and we want to determine  $\mu$  as a function of  $n$  and  $T$ . Therefore, Eq.(10.5) should be seen as an implicit definition of the chemical potential.



**Figure 10.3:**  $g_{3/2}(z)$  function.

To solve this equation we define the fugacity  $z = e^{\beta\mu}$  and the variable  $x = \beta\varepsilon$  and we substitute the 3D density of state  $g(\varepsilon)$  which is:

$$g(\varepsilon) = \left(\frac{2m}{\hbar^2}\right)^{\frac{3}{2}} \frac{V}{2\pi^2} \sqrt{\varepsilon}$$

Thus we get:

$$n = \frac{(mk_B T)^{\frac{3}{2}}}{\sqrt{2\pi^2 \hbar^3}} \int_0^\infty \frac{ze^{-x}}{1 - ze^{-x}} \sqrt{x} dx \quad (10.6)$$

To calculate this integral we need to expand the fraction inside the integral as Taylor expansion<sup>3</sup>:

$$\frac{ze^{-x}}{1 - ze^{-x}} = ze^{-x} (1 + ze^{-x} + z^2 e^{-2x} + \dots) = \sum_{p=1}^{\infty} z^p e^{-px} \quad (10.7)$$

This series converges for  $z < 1$ . If we use such a series and we solve the integral using Gamma functions<sup>4</sup>, by rearranging the terms we obtain the final result for the particle density  $n$  as a function of the fugacity  $z$  which is:

$$n(z) = \left(\frac{mk_B T}{2\pi\hbar^2}\right)^{\frac{3}{2}} g_{3/2}(z) \quad (10.8)$$

where the function  $g_{3/2}(z)$  is defined by the series:

$$g_{3/2}(z) = \sum_{p=1}^{\infty} \frac{z^p}{p^{3/2}} \quad (10.9)$$

The behaviour of such a function is represented in Fig.10.3(a) in which we notice that  $g_{3/2}(z = 1) = 2.612$ . In  $z = 1$  it can be proven that the series converges. We notice that the convergence criterion  $z \leq 1$  implies that  $e^{\beta\mu} \leq 1 \Rightarrow \mu \leq 0$  which means that the chemical potential must be negative. This can also be seen by writing  $n(z)$  the other way around, i.e. writing  $g_{3/2}(z)$  as a function of  $n$ :

$$g_{3/2}(z) = \left(\frac{2\pi\hbar^2}{mk_B T}\right)^{\frac{3}{2}} n \quad (10.10)$$

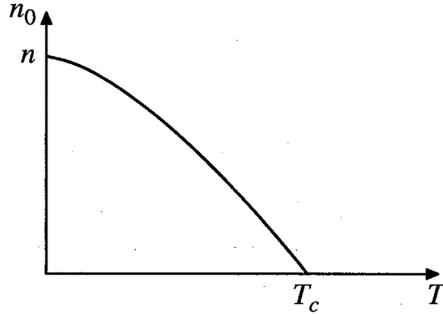
If the temperature is high than  $z$  will be small since  $z = e^{\beta\mu}$  and we can use a short expansion of the function  $g_{3/2}(z) \approx z + \dots$ . Thus we get:

$$g_{3/2}(z) \approx z = e^{\beta\mu} \Rightarrow \mu \approx -\frac{3}{2} k_B T \ln\left(\frac{mk_B T}{2\pi\hbar^2 n^{\frac{2}{3}}}\right) \quad (10.11)$$

We can notice that the chemical potential is negative and that it depends on the temperature. The function  $\mu = \mu(T)$  is sketched in Fig.10.3(b). However, on cooling the boson system to lower temperatures the value of  $z$  gradually increases until it eventually equals one. At this point, the

<sup>3</sup>The Taylor expansion for  $(1+x)^\alpha$  that with  $\alpha = -1$  becomes  $(1+x)^{-1} = 1 - x + x^2 - \dots$

<sup>4</sup>The Gamma functions are defined as  $\Gamma(t) = \int_0^\infty y^{t-1} e^{-y} dy$  and in particular  $\Gamma(\frac{3}{2}) = \frac{\sqrt{\pi}}{2}$ .



**Figure 10.4:** Particle density in the BEC as a function of the temperature.

chemical potential  $\mu$  becomes zero. The temperature where this happens, for a fixed density  $n$ , defines the critical temperature  $T_C$ :

$$T_C = \frac{2\pi\hbar^2}{k_B m} \left( \frac{n}{2.612} \right)^{\frac{2}{3}} \quad (10.12)$$

where we used  $g_{3/2}(z=1) = 2.612$ . Here we notice that:

- $T_C$  only depends on the number of particles per unit volume  $n$ , so it is a fixed number for a given system of bosons.
- $\mu = 0$  physically means that there is no energy cost in adding a boson to our system.

This  $T_C$  is the **BEC temperature**.

### 10.3.1 Below $T_C$

It is worth asking what happens when the temperature decreases below  $T_C$ . As soon as the chemical potential becomes zero, the number of particles in the lowest energy state becomes infinite. This can be seen by looking at the Bose-Einstein distribution for the  $\varepsilon_k = 0$  state which gives us the number of particles  $N_0$  in the lowest state. From Eq.(10.3) we get:

$$N_0 = N(\varepsilon_k = 0) = \frac{1}{e^{-\beta\mu} - 1} \xrightarrow{\mu=0} \infty \quad (10.13)$$

More precisely we can say that out of a total of  $N$  particles in the gas, below  $T_C$ , a macroscopic number  $N_0$  occupies the one quantum state with  $\varepsilon_k = 0$ . Below  $T_C$  we must take into account the  $\mathbf{k} = 0$  point separately, and so we must replace Eq.(10.3) by:

$$N = N_0 + \sum_{\mathbf{k} \neq 0} \frac{1}{e^{\beta\varepsilon_k} - 1} \quad (10.14)$$

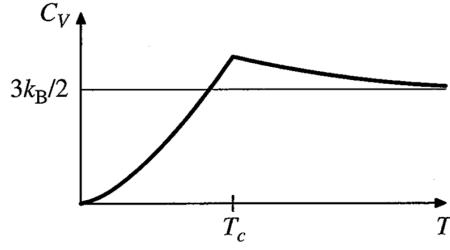
By proceeding as before, substituting the summation with an integral and solving it we can obtain the particle density for  $T < T_C$  which is:

$$n = n_0 + 2.612 \left( \frac{mk_B T}{2\pi\hbar^2} \right)^{\frac{3}{2}} \quad (10.15)$$

The particle density  $n$  can therefore be divided into a condensate density  $n_0$  and a normal one. We will have some particles which respect the normal Bose-Einstein distribution and some others that will form the BEC. The fraction of particles which will form the condensate at  $T < T_C$  can be written very compactly as:

$$\frac{n_0}{n} = 1 - \left( \frac{T}{T_C} \right)^{\frac{3}{2}} \quad (10.16)$$

We can recognize from Fig.10.4, which shows  $n_0$  as a function of the temperature, the classical behaviour of a phase transition. Once again, the formation of a BEC can be seen as a phase transition.



**Figure 10.5:** Specific heat for a system of bosons as a function of the temperature.

### 10.3.2 Specific Heat

Since we have a phase transition we want to know its order. We can do that by computing the specific heat for our system of bosons and looking for its discontinuity. We can compute the energy density  $u$  by multiplying the particle density of Eq.(10.5) for the energy  $\varepsilon$  of each particle:

$$u = \frac{U}{V} = \int_0^\infty \frac{\varepsilon}{e^{\beta(\varepsilon-\mu)} - 1} g(\varepsilon) d\varepsilon \quad (10.17)$$

The result of this integral can be found with a procedure similar to the one used before and we get if  $T > T_C$ :

$$u = \frac{3}{2} k_B T \frac{g_{5/2}(z)}{g_{3/2}(z)} \quad \text{for } T > T_C \quad (10.18)$$

while, if  $T < T_C$ :

$$u = \frac{3}{2} k_B T \frac{T^{5/2}}{T_C^{3/2}} \frac{g_{5/2}(1)}{g_{3/2}(1)} \quad \text{for } T < T_C \quad (10.19)$$

where the function  $g_{5/2}(z)$  is defined by:

$$g_{5/2}(z) = \sum_{p=1}^{\infty} \frac{z^p}{p^{5/2}} \quad (10.20)$$

We can notice that when  $T \gg T_C$  we get the usual classical energy of a three-dimensional gas  $u \approx \frac{3}{2} k_B T$ . Now that we know the energy density, we can compute the specific heat which is:

$$C_V = \frac{\partial u}{\partial T}$$

The result is shown in Fig.10.5. The specific heat as a cusp at  $T = T_C$ , i.e. the second order derivative of the energy is discontinuous. The phase transition is therefore a second-order phase transition.

## 10.4 Experimental Realization of a BEC

As we have seen in the historical introduction, low-temperature physics is tightly bonded to the experimental technique. When in 1995 BEC was finally achieved experimentally, this was done by using an extremely diluted gas of alkaline atoms, in particular, Rubidium. The problem in the realization of a BEC, in fact, does not stay only in the low temperature but also in the non-interacting condition. In fact, to obtain the BEC we studied, we need a non-interacting gas of bosons below  $T_C$ . Moreover, we need the wave functions of the bosons to overlap and at the same time, we don't want the bosons to be too close to each other. This means, in physical terms, that the De Broglie wavelength  $\lambda_{dB}$  associated with each boson must be comparable with the average distance  $d$  between bosons:

$$\lambda_{dB} = \frac{\hbar}{p} = \left( \frac{2\pi\hbar^2}{mk_B T} \right)^{\frac{1}{2}} \simeq d$$

From Tab.10.1 we can notice that in 1927 a discontinuity in the specific heat of Helium-4 was observed at  $T = 2.2K$ . At first, scientists thought that this was related to the formation of a BCE condensate in Helium. This is actually not true. For liquid Helium ( $LHe$ ) in fact, we have a density

of  $\rho = 145 \text{ kg/m}^3$ , a particle density of  $n = 10^{22} \text{ cm}^{-3}$  and a mass  $m \simeq 4m_p$  and the prediction of the critical temperature for the formation of a BEC using Eq.(10.12) gives as a result:

$$T_C = \frac{2\pi\hbar^2}{k_B m} \left( \frac{n}{2.612} \right)^{\frac{2}{3}} \simeq 3.1K$$

which differs from the experimental observation of  $T = 2.2K$ . Not only the critical temperature is different, but also the properties of  $LHe$  below  $2.2K$  are different from the one expected from a BEC. Today we know that discontinuity in the specific heat is associated with the appearance of a new phase, the superfluid one. The problem in using  $LHe$  as a candidate for a BCE stays in the density. In general, for BCE we need a very small density in order to avoid any possible interaction between particles. It has been observed that to create a BCE the required particle density is in the order of  $10^{11} - 10^{15} \text{ cm}^{-3}$ . If Helium is not a good candidate for a BEC, Rubidium is. Let's see why.

#### 10.4.1 Rubidium in BEC

First of all, we need to keep in mind that any good candidate for a BEC must be a neutral boson. This is because we need the particles to be non-interacting and having charged particles wouldn't satisfy that condition for sure. Neutral atoms are therefore our best choice. Rubidium  $Rb$  is a heavy alkali atom ( $m \simeq 87m_p$ ) with an electronic configuration  $[Kr]5s^1$ . This means that for  $Rb$  we have  $S = \frac{1}{2}$  and  $L = 0$ . Apparently, this doesn't look like a boson at all. The apparent problem is easily solved by considering the nuclear spin too. For Rubidium, in fact, we have  $I = \frac{3}{2}$  which gives an integer total atomic spin  $S_a = 1, 2$ . Then, not only  $Rb$  is a boson but it also has an electronic spin which means that it is responsive to a magnetic field, *i.e.* it has a magnetic dipole moment different from zero. The responsivity to a magnetic field is really important in the experimental realization of a BEC since magnetic fields are used in the confinement and cool-down process of the atom. Let's look at the spin configuration more in detail by referring to Tab.10.2. In the table

**Table 10.2:** Spin configurations for  $Rb$ .

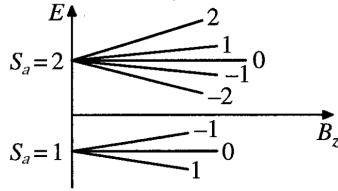
$S_a$	$M_{S_a}$	$m_n$	$m_e$	State $ m_n, m_e\rangle$
2	2	$\frac{3}{2}$	$\frac{1}{2}$	$ \frac{3}{2}; \frac{1}{2}\rangle$
2	1	$(\frac{3}{2}; \frac{1}{2})$	$(-\frac{1}{2}; \frac{1}{2})$	$\frac{1}{2} (\sqrt{3}  \frac{1}{2}; \frac{1}{2}\rangle +  \frac{3}{2}; +\frac{1}{2}\rangle)$
2	0	$(\frac{1}{2}; -\frac{1}{2})$	$(-\frac{1}{2}; \frac{1}{2})$	$\frac{1}{\sqrt{2}} ( \frac{1}{2}; -\frac{1}{2}\rangle +  -\frac{1}{2}; \frac{1}{2}\rangle)$
2	-1	$(-\frac{3}{2}; -\frac{1}{2})$	$(\frac{1}{2}; -\frac{1}{2})$	$\frac{1}{2} (\sqrt{3}  -\frac{1}{2}; -\frac{1}{2}\rangle +  -\frac{3}{2}; \frac{1}{2}\rangle)$
2	-2	$-\frac{3}{2}$	$-\frac{1}{2}$	$ -\frac{3}{2}; -\frac{1}{2}\rangle$
1	1	$(\frac{1}{2}; \frac{3}{2})$	$(\frac{1}{2}; -\frac{1}{2})$	$\frac{1}{2} ( \frac{1}{2}; \frac{1}{2}\rangle - \sqrt{3}  \frac{3}{2}; -\frac{1}{2}\rangle)$
1	0	$(\frac{1}{2}; -\frac{1}{2})$	$(-\frac{1}{2}; \frac{1}{2})$	$\frac{1}{\sqrt{2}} ( \frac{1}{2}; -\frac{1}{2}\rangle -  -\frac{1}{2}; \frac{1}{2}\rangle)$
1	-1	$(-\frac{1}{2}; -\frac{3}{2})$	$(-\frac{1}{2}; \frac{1}{2})$	$\frac{1}{2} ( -\frac{1}{2}; -\frac{1}{2}\rangle - \sqrt{3}  -\frac{3}{2}; \frac{1}{2}\rangle)$

$m_n = -I, \dots, 0, \dots, I$  and  $m_e = -S, \dots, 0, \dots, S$  where  $I$  and  $S$  are the nuclear and the electronic spin quantum numbers, respectively. We already know that in the presence of an external magnetic field, the levels are shifted in energy because of Zeeman splitting. However, even in the absence of an external field, the levels are slightly shifted because of the interaction of the nucleus with the external electrons. This is the so-called **hyperfine structure** of the atom. By keeping into account both the contributions from the hyperfine structure and the Zeeman shift, the Hamiltonian of the system will be:

$$\hat{\mathcal{H}} = A \hat{\mathbf{I}} \cdot \hat{\mathbf{S}} + \mu_B \hat{S}_z B_z \quad (10.21)$$

where  $A$  is the hyperfine interaction between the nucleus and the valance electrons. In general, the product of two any angular momenta can be computed as:

$$\begin{aligned} \hat{\mathbf{S}}_{tot} &= \hat{\mathbf{S}}_1 + \hat{\mathbf{S}}_2 \\ \hat{\mathbf{S}}_1 \cdot \hat{\mathbf{S}}_2 &= \frac{\hat{S}_{tot}^2 + \hat{S}_1^2 + \hat{S}_2^2}{2} \end{aligned}$$



**Figure 10.6:** Rubidium hyperfine and Zeeman splittings.

where, in general,  $S^2 = s(s + 1)$ . In our case, we have to compute  $\hat{\mathbf{I}} \cdot \hat{\mathbf{S}}$  where the total spin is the atomic spin  $S_a = 1, 2$ . Thus, with  $I = 3/2$  and  $S = 1/2$  we get:

$$I^2 = \frac{3}{2} \left( \frac{3}{2} + 1 \right) = \frac{15}{4}$$

$$S^2 = \frac{1}{2} \left( \frac{1}{2} + 1 \right) = \frac{3}{4}$$

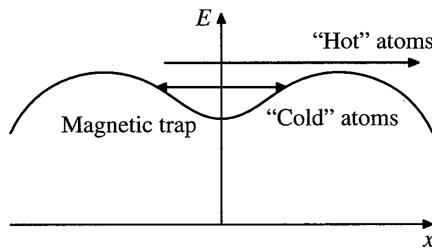
From which we can obtain:

$$\hat{\mathbf{I}} \cdot \hat{\mathbf{S}} = \frac{S_a^2 + I^2 + S^2}{2} = \begin{cases} \frac{1}{2} \left[ 6 - \frac{15}{4} - \frac{3}{4} \right] = \frac{3}{4} & S_a = 2 \\ \frac{1}{2} \left[ 2 - \frac{15}{4} - \frac{3}{4} \right] = -\frac{5}{4} & S_a = 1 \end{cases} \quad (10.22)$$

So, because of the hyperfine structure, we have two energy levels:

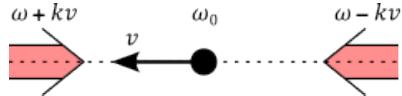
$$E = \begin{cases} E_2 = \frac{3}{4}A & S_a = 2 \\ E_1 = -\frac{5}{4}A & S_a = 1 \end{cases} \quad (10.23)$$

The weak magnetic field Zeeman further splits these  $S_a = 2$  and  $S_a = 1$  hyperfine levels into their different  $M_{S_a}$  states as shown in Fig.10.6. The magnetic field dependence of the quantum state energies is exploited in the magnetic atom trap. The trap is constructed by producing a region of space in which the magnetic field has a local minimum<sup>5</sup> as shown in Fig.10.7. Since different spin configurations correspond to different energies, it is possible to "select" atoms with the same spin by considering that only some of them will have the right amount of energy to be trapped in the magnetic field. Moreover, atoms must be produced at first, before sending them into the trap. Usually, this is done by heating a block of solid Rubidium inside an oven, to let it evaporate. Because of this production method, the atoms exiting the oven will have different kinetic energies according to a certain distribution, *i.e.* Maxwell-Boltzmann distribution. Only the atoms with the right kinetic energy will remain trapped. Therefore, it is fundamental to be able to slow down atoms to collect in the magnetic trap as many atoms as we can. To reduce atoms' kinetic energy the laser cooling technique is used.



**Figure 10.7:** Potential magnetic energy associated with a magnetic atom trap.

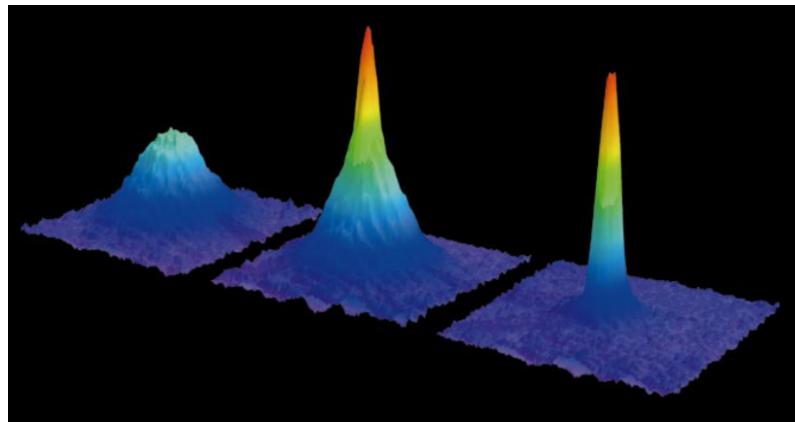
<sup>5</sup>The existence of a local minimum is surprisingly consistent with Maxwell's equations for a region of free space:  $\nabla \cdot \mathbf{B} = 0$  and  $\nabla \times \mathbf{B} = 0$ .



**Figure 10.8:** Laser Cooling scheme: two counterpropagating pulses on an atom.

### 10.4.2 Laser Cooling

The idea of laser cooling is, in principle, quite simple. Two counterpropagating laser pulses are made impinging on the atoms from both directions as represented in Fig.10.8. If the laser is resonant with the atomic transition, an absorption event occurs and the conservation of the total momentum implies that the atom will reduce its momentum. The initial momentum is given by the atomic momentum minus the one of the photon since the two momenta are opposite in direction (counterpropagating pulse). After the absorption, the photon disappears and as a result, the atomic momentum decreases. To the absorption event, a spontaneous emission event follows. However, spontaneous emission is isotropic and therefore it implies no change in the momentum of the atom, *i.e.* there is an equal probability of emission in any direction and the average momentum change is zero. One may think: "If we have two counterpropagating lasers impinging on the atoms, one laser will reduce the velocity but the other will increase it, "pushing" the atom from the back". Doppler effect prevents this, granting that the atom is slowed down only, and not accelerated. In fact, because of the Doppler effect, the frequency of the photons seen by the atom is equal to the resting frequency of the photon  $\omega$  plus (or minus) a term due to the relative motion between the photon and the atom. As a result, for example in the case in which an atom and photon move one towards the other the frequency of the photon seen by the atom will be  $\omega + kv$  where  $k$  is the wave vector of the photon and  $v$  the velocity of the atom. For the photon that is "chasing" the atom, instead, the frequency will be  $\omega - kv$ . We can thus tune the laser in such a way that only the photon facing the atom is resonant with the atom itself and can therefore cause an absorption. In conclusion, the atoms can be slowed down and trapped by the magnetic field. Once the atoms have been trapped and the system has been cooled down, we need to find a way to recognize the formation of a BEC. In fact, from an experimental point of view, we need to know if the BEC is formed or if we need to decrease the temperature more. To detect the formation of the BEC its properties are used. We know that the atoms that will form the BEC will all have a zero-momentum  $k$  because of the definition of BEC itself. If we turn off the lasers and the magnetic confinement, the atoms will fall under the effect of gravity. If the BEC is formed the atoms will have  $k = 0$  and during the falling, they will not expand and will remain in a cylindrical shape. Otherwise, a normal group of atoms while falling would spread because of the internal forces acting between them which provide a non-zero momentum. This can be easily seen in Fig.10.9<sup>6</sup>. The perfectly cylindrical shape is ideal since there always will be a bit of spreading in the cloud of atoms because of the uncertainty principle  $\Delta x \Delta p$ .



**Figure 10.9:** Observation of BEC by absorption imaging.

---

<sup>6</sup>This picture comes from the Nobel lecture: *when atoms behave as waves: Bose-Einstein condensation and the atom laser* from Wolfgang Ketterle who won the Nobel prize in 2001 for the realization of the Bose-Einstein condensate in 1995.

# Chapter 11

## Superfluidity

There exist only two superfluids that can be studied in the laboratory and they are two isotopes of Helium:  $^3He$  and  $^4He$ . Superfluidity was firmly established in  $^4He$  in 1938 by Kapitsa who won the Nobel prize for his discovery 40 years later. Despite being isotopes of the same element, superfluidity of  $^4He$  and  $^3He$  is due to two different phenomena and has different properties:

$$\begin{array}{lll} {}^4He : & S_a = 0 & T_{SF} = 2.17K \\ {}^3He : & S_a = 1/2 & T_{SF} = 2mK \end{array}$$

In the following, we will limit ourselves to the study of superfluidity in  ${}^4He$ . We will use the short  $LHe$  to indicate liquid Helium and we will always refer to the  $A = 4$  isotope. The short SF will be used for the words superfluidity and superfluid indistinctly.

### 11.1 Quantum Fluids

With the term quantum fluid, we indicate a substance which remains fluid, *i.e.* gas or liquid, at such low temperature that the effects of quantum mechanics play a dominant role. This fluidity of the substance at very low temperatures is the combination of a light nuclear mass and the relatively weak Van der Waals interaction between particles.  $LHe$  is a quantum fluid. As well as BCE, also in quantum fluids, not all the particles become quantum fluid when the temperatures decrease enough but only a fraction of them. The fraction of particles which become a quantum fluid (or a superfluid) is zero above a certain critical temperature  $T_C$  and increases for  $T < T_C$ . In this fraction of particles which becomes SF, a zero viscosity  $\eta = 0$  is observed and these particles' phase coexist with the normal phase of the other fraction of particles which is not SF. The first question we can ask ourselves is: why Helium? Why is Helium the only noble gas which becomes superfluid?? The answer to this question lies, as for BEC, in the ratio between the average distance between atoms and their associated de Broglie wavelength. Let's compare, for example,  $He$  and  $Ne$ , two noble gases. The distance between atoms is determined by the attraction and repulsive potentials. Both for Helium and Neon, the overall potential energy can be expressed as a function of the interatomic distance  $r$ :

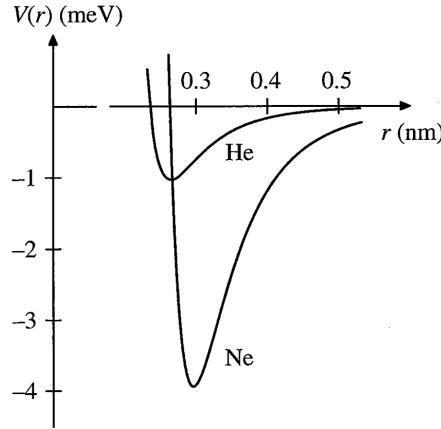
$$V(r) = ae^{-br} - \frac{C}{r^6} \quad (11.1)$$

where  $a$ ,  $b$  and  $C$  are constant which can be calculated or determined experimentally<sup>1</sup>. Fig.11.1 shows this potential for  $He$  and  $Ne$ . We can see from the graph that the energy corresponding to the minimum of the potential is far higher in absolute value for  $Ne$  than for  $He$  while the equilibrium position is approximately the same. More specifically, being  $d$  the equilibrium position and  $\epsilon_0$  the corresponding energy, we have:

$$\begin{array}{llll} He : & d = 0.265nm & \epsilon_0 = 1.03meV & m = 4m_p \\ Ne : & d = 0.296nm & \epsilon_0 = 3.94meV & m = 20m_p \end{array}$$

---

<sup>1</sup>Close to the equilibrium point this potential can be also approximated with a Lennard-Jones type potential  $V(r) = \epsilon \left( \frac{d^{12}}{r^{12}} - \frac{d^6}{r^6} \right)$



**Figure 11.1:** Potential energy as a function of the interatomic distance  $r$  for  $He$  and  $Ne$ .

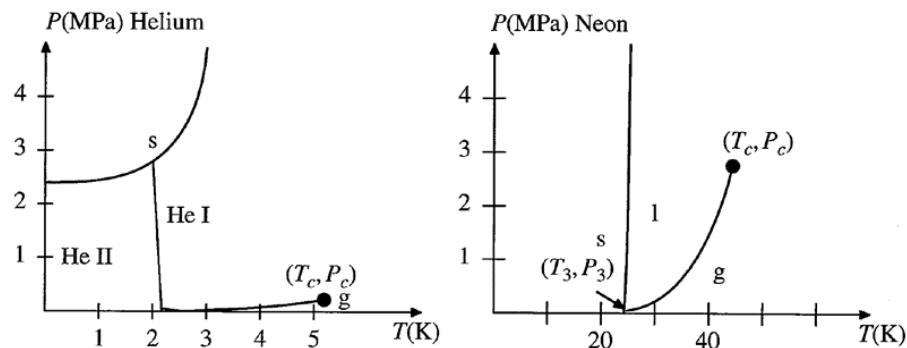
The weaker attraction for Helium, combined with the lighter atomic mass makes Helium a quantum liquid, while Neon is practically purely classic. This can be better seen by looking at the de Broglie wavelength and comparing it with the equilibrium position  $d$  which essentially represents the size of the atom. We have that:

$$\lambda_{dB} = \left( \frac{2\pi\hbar^2}{mk_B T} \right)^{\frac{1}{2}} = \begin{cases} \lambda_{dB}^{(He)} \simeq 0.4 \text{ nm} & \text{at } T = 4K \\ \lambda_{dB}^{(Ne)} \simeq 0.07 \text{ nm} & \text{at } T = 24K \end{cases}$$

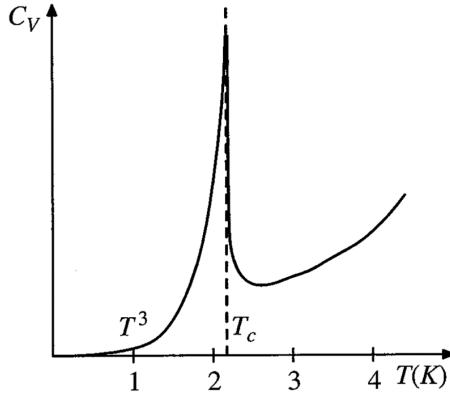
where the two chosen temperatures are the ones at which the two elements liquify. We see that for  $He$  the de Broglie wavelength is comparable with the average distance between atoms  $d$  and this allows the wave functions of different atoms to overlap one with the other. This does not happen for Neon in which  $\lambda_{dB}$  is much smaller than  $d$ : Neon becomes a solid before having the possibility to become SF. This can also be seen in terms of phase diagrams as shown in Fig.11.2. In the case of Neon, we have a "classical" phase diagram in which a critical and a triple point are present. In particular, all the phases are well distinguished one from the other and at  $T = 0$  only the solid phase exists. The first difference we notice in the phase diagram of Helium is this one: at  $T = 0$  the material can be liquid at relatively high pressures. Moreover, for Helium two liquid phases exist and there is no triple point, *i.e.* no point in which solid, liquid and gas coexist. The transition between the liquid phases I and II is sharp and happens at a very well-defined temperature, *i.e.* the critical temperature  $T_c$ . We can attribute the fact that at  $T = 0$  Helium is not solid to zero-point oscillations. Let's suppose to have a solid. The atoms in the solid vibrate with an energy which can be roughly estimated as:

$$E_0 = \frac{3}{2}\hbar\omega_0 \quad (11.2)$$

We can estimate the frequency for an FCC lattice as  $\omega = \sqrt{\frac{4k}{m}}$  where  $k$  is computed by using a Lennard-Jones potential approximation  $k = \frac{1}{2} \frac{d^2 V(r)}{dr^2} = \frac{36\epsilon_0}{d^2}$ . This gives us the zero-point oscillation



**Figure 11.2:** Phase diagrams  $P$ - $T$  for  $He$  and  $Ne$ .



**Figure 11.3:** Specific heat of  $He$ .

energy that in the two cases of Helium is  $E_0^{(He)} \simeq 7\text{meV}$  and for Neon  $E_0^{(Ne)} \simeq 4\text{meV}$ . In the case of  $He$ , we had that the bottom of the potential well was at about  $1\text{meV}$  which is much lower than the zero-point oscillation energy: the oscillations, even at  $T = 0$ , destroy the solid order and make Helium a liquid. This does not happen for  $Ne$  which has a bottom of the potential well comparable with the zero-point oscillation energy.

## 11.2 The wave function

The existence of two liquid phases for  $LHe$  was first observed as a discontinuity of the specific heat (see Fig.11.3). It is evident from the graph the presence of a phase transition at a well-defined critical temperature. It is worth noting that the behaviour of  $C_V$ , in this case, differs from the one of the BEC: there we had a finite but different derivative at  $T = T_C$ , here the derivative of  $C_V$  at  $T = T_C$  is infinite. Moreover, the shape of  $C_V(T)$  resembles the Greek letter  $\lambda$  and it is therefore called lambda-shape. The critical temperature is also called lambda temperature. The shape of  $C_V$  can be approximated with a power series<sup>2</sup>:

$$C_V = \begin{cases} C(T) + A_+ |T - T_C|^{-\alpha} & \text{for } T > T_C \\ C(T) + A_- |T - T_C|^{-\alpha} & \text{for } T < T_C \end{cases} \quad (11.3)$$

where  $A_+$ ,  $A_-$  and  $\alpha$  are constants and  $C(T)$  is a smooth function of  $T$  near  $T_C$ . It is possible to determine experimentally that  $\alpha = -0.009$ . We have already faced one of these types of exponents in Sec.4.3 which are the so-called critical exponents. These exponents can tell us something about the phase transition that they are describing and in this particular case, the critical exponent  $\alpha$  is the same as another physical model called the *3D XY model*. The 3D XY model describes a sort of magnetic order in which:

- The spin dimensionality is 2, differently from the Heisenberg model in which we had  $D = 3$ . The dimensionality of the spin-lattice is  $d = 3$ .
- The density of the spin counts and it is a parameter of the model.
- Spins are oriented with respect to a particular horizontal direction with an angle  $\theta$  which is the order parameter of the model.

In this model, the wave function which describes the system is associated with the order parameter  $\theta$  as a function of the position  $\mathbf{r}$ :

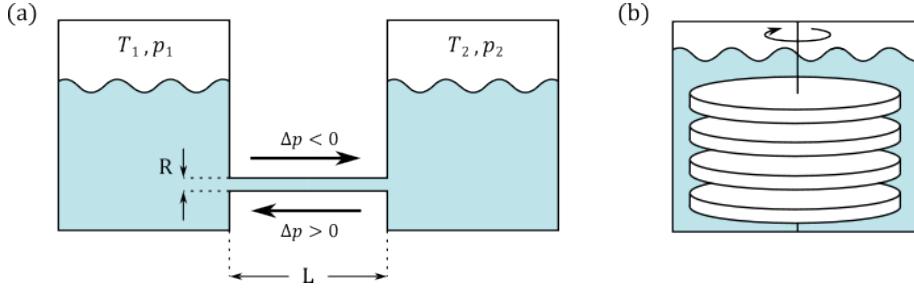
$$\psi_0(\mathbf{r}) = |\psi_0(\mathbf{r})| e^{i\theta(\mathbf{r})}$$

In the model,  $|\psi_0(\mathbf{r})|^2$  is the particle density  $n(\mathbf{r})$ . We can choose by analogy this wave function to describe the Helium atoms in the SF and therefore we can impose:

$$n(\mathbf{r}) = |\psi_0(\mathbf{r})|^2 \quad (11.4)$$

---

<sup>2</sup>Actually, at  $T \rightarrow 0$ :  $C_V \propto T^3$  while at  $T \approx T_C$  the specific heat is very close to being a logarithmic function  $C_V \sim \ln |T - T_C|$ .



**Figure 11.4:** (a) Two tanks filled with *LHe* connected by a small pipe and (b) disks dip into a *LHe* tank.

and the complex macroscopic wave function which describes the state of the Helium atoms as a function of their position is:

$$\psi_0(\mathbf{r}) = \sqrt{n(\mathbf{r})} e^{i\theta(\mathbf{r})} \quad (11.5)$$

We can see that with this definition if  $|\psi_0(\mathbf{r})| = 0$  it is not possible to define the phase  $\theta$  while if  $|\psi_0(\mathbf{r})| \neq 0$  the phase  $\theta$  of the wave function can be seen as the natural order parameter of the system. In fact,  $|\psi_0(\mathbf{r})| = 0$  means no density of particles in the SF state, *i.e.*  $n(\mathbf{r}) = 0$ , which corresponds to an undefined order parameter for the SF state.

### 11.3 Superflow

Superflow arises whenever the macroscopic wave function phase  $\theta(\mathbf{r})$  varies in space. Let's consider the usual formula for the current density for particle flow:

$$\mathbf{j}_0 = \frac{\hbar}{2mi} [\psi_0^*(\mathbf{r}) \nabla \psi_0(\mathbf{r}) - \psi_0(\mathbf{r}) \nabla \psi_0^*(\mathbf{r})] \quad (11.6)$$

This is the number of particles flowing per unit area per second in the SF phase<sup>3</sup>. Using  $\psi_0(\mathbf{r}) = \sqrt{n_0} e^{i\theta}$  we have:

$$\begin{aligned} \nabla \psi(\mathbf{r}) &= e^{i\theta} \nabla (\sqrt{n_0}) + i\sqrt{n_0} e^{i\theta} \nabla \theta \\ \nabla \psi^*(\mathbf{r}) &= e^{-i\theta} \nabla (\sqrt{n_0}) - i\sqrt{n_0} e^{-i\theta} \nabla \theta \end{aligned}$$

which gives a current density:

$$\mathbf{j}_0 = \frac{\hbar}{m} n_0 \nabla \theta \quad (11.7)$$

Since, in general,  $\mathbf{j}_0 = n_0 \mathbf{v}_s$  we get that the velocity of the particles in the SF is:

$$\mathbf{v}_s = \frac{\hbar}{m} \nabla \theta \quad (11.8)$$

this is the **superfluid velocity**. Unlike fluid flow in normal liquids, superflow is a movement of particles without dissipation, *i.e.* with zero viscosity.

#### 11.3.1 The experiment

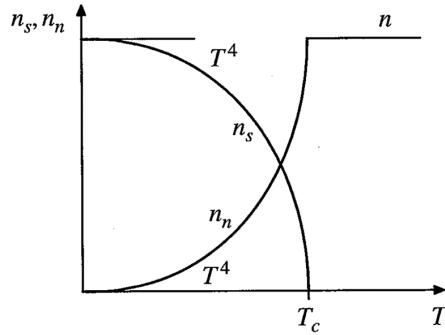
Kapitsa showed that *LHe II* could flow through narrow capillaries without any apparent resistance. The experimental setup, whose scheme is shown in Fig.11.4(a), consists of two equal tanks filled with *LHe* and connected by a small pipe. In normal conditions, at  $T > T_C$ , the flow of liquid from one tank to the other is only driven by a pressure difference. If  $\Delta p = p_2 - p_1 = 0$  no movement occurs in the fluid. If, instead,  $\Delta p < 0 \Rightarrow p_1 > p_2$  a motion occurs from the left to the right. It is quite easy to demonstrate that:

$$\frac{\Delta p}{L} = \eta \frac{v}{R^2} \quad (11.9)$$

where  $L$  and  $R$  are the length and the radius of the pipe, respectively. Kapitsa found that at  $T < T_C$  there was a velocity  $v$  different from zero even when  $\Delta p = 0$ . The only explanation to

---

<sup>3</sup>Here the subscript "0" stands for SF phase.



**Figure 11.5:** Temperature dependence of the superfluid and normal components of liquid *He* II.

obtain a non-null velocity with a zero difference in pressure is to have zero viscosity  $\eta = 0$ . At  $T < T_C$  we are entered in the SF phase. Kapitsa made another observation. He took a series of disks packed one on top to the other with a little space in between. He placed the disks into a tank filled with *LHe*, attaching them to a wire and letting them free to rotate as in Fig.11.4(b). A first measurement was taken with the empty tank to retrieve the elastic constant of the wire and the reference oscillation frequency of the disks  $\omega_0$ . This frequency depends on the mass and on the inertia momentum of the block of disks. Then he put *LHe* I, *i.e.* at  $T > T_C$ , and he computed  $\omega_I$  which is different from  $\omega_0$ . In fact, when *LHe* is placed into the tank, the fluid will fill the empty spaces between the disks and because of the fluid viscosity, the disk will drag with it a thin layer of fluid along its motion, effectively increasing the inertia of the disk and thus decreasing the frequency. When the temperature is cooled down below  $T_C$  and the *LHe* enters its II liquid phase, Kapitsa observed that the frequency  $\omega_{II}$  changed. Now, only a fraction of the liquid has a non-null viscosity and can be dragged by the disks, the other part, the SF part, does not influence the motion, and the frequency, at all. In conclusion, we can say that:

$$\omega_0 > \omega_{II} > \omega_I$$

This experiment confirms the two-fluid model of liquid Helium in which, as we said at the beginning, the SF phase (I) and the normal liquid phase (II) coexist together at  $T < T_C$ . In fact, if all the fluid will become superfluid, we would have  $\omega_{II} \simeq \omega_0$ . As we did for BEC, we can write the total particle density as the sum of a SF particle density and a normal one:

$$n = n_s + n_n \quad (11.10)$$

The SF component  $n_s$  flows with zero viscosity. Both these two densities depend on the temperature with the trend shown in Fig.11.5. At very low temperatures, near  $T = 0$ , it is found empirically that almost all the fluid is in the II liquid phase, so  $n_s \simeq n$  and  $n_n \simeq 0$ . In this temperature region, experiments show that  $n_s(T) \sim T^4$ . Moreover, for  $T \simeq T_C$ ,  $n_s \sim (T_C - T)^{0.67}$  which is in perfect agreement with the theoretical predictions based on the 3D XY model. Since we have two particle densities, to describe the *LHe* hydrodynamics we need two current densities. Overall:

$$\mathbf{j} = \mathbf{j}_s + \mathbf{j}_n \quad (11.11)$$

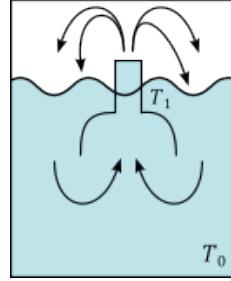
where clearly  $\mathbf{j}_s = n_s \mathbf{v}_s$  and  $\mathbf{j}_n = n_n \mathbf{v}_n$ .

### 11.3.2 Fountain Effect

In the previous section, the main observation of the experiment in Fig.11.4(a) was that in a SF we can have a motion of fluid even with a zero difference in the pressure of the two tanks. The origin of the motion can be attributed to the difference in temperature  $\Delta T = T_2 - T_1$ . In particular, we observe a motion from the coldest tank to the hottest one, *i.e.* a motion from the left to the right if  $\Delta T > 0$ . So, to summarize we have:

$$\text{Normal fluid : } \Delta p < 0 \Rightarrow v_n > 0 \quad (11.12)$$

$$\text{Superfluid : } \Delta T > 0 \Rightarrow v_s > 0 \quad (11.13)$$



**Figure 11.6:** Fountain effect for a baker of liquid  $He$  II.

where by convention we defined  $v > 0$  when the motion occurs from the left to the right. The fact that a superflow is due to a temperature difference gives rise to the spectacular **fountain effect**. Fig.11.6 shows a schematic representation of this peculiar effect of SF. Let's consider a baker filled with  $LHe$  II, in which a tube filled with a dense material (*e.g.* sand) lets pass only the fluid in the SF state. Usually, a small wire is wrapped around the tube and a current is made to flow in the wire to increase the temperature at the top of the tube. The temperature difference between the baker and the top of the tube will cause a motion of the SF causing a sort of spontaneous fountain. The fraction of the liquid in the normal state will remain still. Sometimes this kind of motion is mistaken for the non-existent spontaneous motion. Clearly, this is not the case since energy is provided to the system to maintain the temperature difference. However, in the motion of the SF there is no heat exchange and therefore no entropy variation and the motion can go ahead forever if  $\Delta T > 0$ . For such a system we can write the expression of the Gibbs free energy  $G$  which is:

$$G = U + pV - ST = H - ST \quad (11.14)$$

where  $H$  is the enthalpy. We want to use the Gibbs free energy since it is linked to the chemical potential through the relation:

$$G = N_m\mu \quad (11.15)$$

with  $N_m$  number of moles. We can write the differential expression for the energy:

$$dG = -SdT + Vdp + \mu dn \quad (11.16)$$

since for our system, we consider a constant number of particles we have  $\mu dn = 0$ . By dividing everything by the number of moles we get:

$$\frac{1}{N_m}dG = d\mu = -\frac{V}{N_m}\frac{S}{V}dT + \frac{V}{N_m}dp = -V_m s dT + V_m dp \quad (11.17)$$

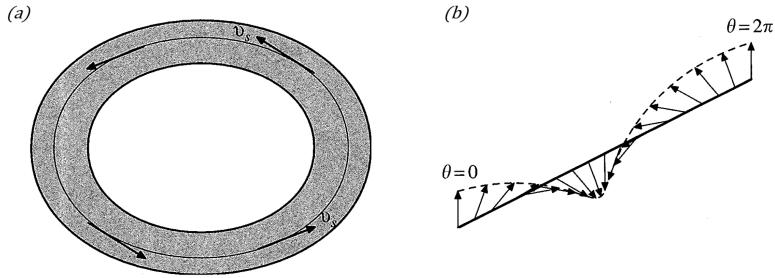
where we define the molar volume  $V_m = V/N_m$  and the entropy per unit volume  $s = S/V$ . To compute the expression of  $\mu(p, T)$  we need to integrate along a path in the  $p - T$  plane. By choosing a path made of a vertical stretch at  $T = 0$  from 0 to  $p$  and a horizontal one at fixed  $p$  from 0 to  $T$ , we get:

$$\mu(p, T) = \mu(0, 0) + \int_0^p V_m(p', 0)dp' - \int_0^T V_m s(p, T')dT' = \mu(0, 0) + V_m p - V_m p_f \quad (11.18)$$

where we defined the fictitious pressure  $p_f = \int_0^T s(p, T')dT'$ , known as fountain pressure, and we considered  $V_m$  constant since we are dealing with a liquid. The integral which we summarized in the fountain pressure  $p_f$  can be numerically computed by knowing the value of the specific heat  $C_V$  and, in principle, its value is known for any  $p$  and  $T$ . For example, we have:

$$\begin{aligned} p_f &= 0.692 \text{ bar} && \text{at } T_C \\ p_f &= 0 && \text{at } T = 0 \end{aligned}$$

The value of the fountain pressure is very high at  $T_C$ . With such pressure, we could obtain a fountain jet 50m high. We can write Newton's law for our system by considering that the overall



**Figure 11.7:** (a) Motion of a superfluid in a close tube and (b) winding number of the superfluid phase  $\theta$ .

potential energy of the fluid in the fountain is given by the gravitational potential energy and the Gibbs free energy which is equal to the chemical potential. We get:

$$\begin{aligned} M_m \frac{d\mathbf{v}}{dt} &= \mathbf{F} = -\nabla U = -\nabla(Mgz + \mu) \\ \Rightarrow \rho \frac{d\mathbf{v}}{dt} &= -\nabla(p - p_f + \rho gz) \end{aligned} \quad (11.19)$$

where  $M_m$  is the molar mass and we divided the first equation by the molar volume and we neglect the constant  $\mu(0,0)$  since it vanishes in the gradient.  $z$  is the height of the fluid in the fountain. If we consider the movement on a horizontal plane we can set  $\Delta z = 0$ . Now, we can analyse the two cases  $\Delta T = 0$  and  $\Delta T \neq 0$ :

- If  $\Delta T = 0 \Rightarrow \Delta p_f = 0$  and therefore:

$$\rho \int \frac{d\mathbf{v}}{dt} = - \int \nabla p \simeq -\Delta p$$

which is the classical fluid motion driven by pressure difference.

- If  $\Delta T = T_2 - T_1 > 0 \Rightarrow \Delta p_f > 0$  and if we consider the relevant case in which  $\Delta p = 0$  we get:

$$\rho \int \frac{d\mathbf{v}}{dt} = - \int \nabla(-p_f) \simeq \Delta p_f$$

This means that we get a fluid motion, driven by the temperature, from the left to the right.

Superfluidity is responsible for another spectacular effect. When SF Helium is placed inside a baker, this will tend to exit the baker itself. This is because of a very thin layer of fluid which deposits on the surfaces of the baker and that works like a "ladder" for the superfluid which can exit the baker. The deposition of a microscopic wetting layer of fluid on the surfaces of a container occurs for many fluids and it is typically a few micrometers thick. For example, looking at wine in a glass, one can see a thin fluid layer coating the sides of the glass. In the case of a superfluid, however, there is no viscosity to prevent it from flowing up through the wetting layer and out the rim of the beaker<sup>4</sup>.

## 11.4 Flow Quantization and Vortices

Superfluidity properties are not over. The existence of the macroscopic wave function in  $He II$ , in fact, leads to a quantization of the superflow. To see how, let's consider the definition of the superfluid velocity  $\mathbf{v}_s$ :

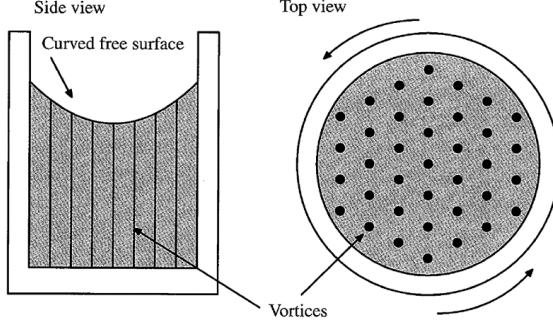
$$\mathbf{v}_s = \frac{\hbar}{m} \nabla \theta \quad (11.20)$$

In general, any field that can be expressed as a function of a gradient (e.g. electric field) can be proven to be irrotational<sup>5</sup>, which means:

$$\nabla \times \mathbf{v}_s = 0 \quad (11.21)$$

<sup>4</sup>A video of this phenomenon can be found at the link: <https://www.youtube.com/watch?v=2Z6UJbwxBZI>

<sup>5</sup>To prove this is sufficient to remember the geometric relation for which the curl of any gradient is zero.



**Figure 11.8:** Superfluid in a rotating cylindrical tank.

If we then consider the motion of a SF along any close tube, as the one in Fig.11.7(a), we have that the circulation  $\kappa$  of  $\mathbf{v}_s$  along any close line is zero:

$$\kappa = \oint \mathbf{v}_s \cdot d\mathbf{r} = 0 \quad (11.22)$$

because the field  $\mathbf{v}_s$  is irrotational. Nevertheless, we want the angle  $\theta$  to vary in order to have a motion of the fluid. We can respect both conditions. Let us rewrite the integral as:

$$\kappa = \oint \mathbf{v}_s \cdot d\mathbf{r} = \frac{\hbar}{m} \oint \nabla\theta \cdot d\mathbf{r} = \frac{\hbar}{m} \Delta\theta \quad (11.23)$$

where  $\Delta\theta$  is the change in the phase angle  $\theta$  after going around the tube. From Fig.11.7(b) we can see that the  $\theta$  angle remains the same after a complete rotation along the tube, in the non-trivial case in which  $\theta$  is not constant if the angle variation  $\Delta\theta$  is an integer multiple of  $2\pi$ . We can therefore impose:

$$\Delta\theta = 2\pi n \quad (11.24)$$

where  $n$  is an integer. The circulation of the flow is then quantized as:

$$\kappa = \frac{\hbar}{m} n \quad (11.25)$$

the integer  $n$  is the number of times that  $\theta$  winds through  $2\pi$  around the closed loop. For this reason,  $n$  is called **winding number**. In Fig.11.7(b), for instance,  $n = 1$ . The presence of a quantized circulation has been confirmed by several experiments. This quantization, for example, is visible on a macroscopic scale in the form of vortices which appear in the superfluids. Let's consider a cylindrical container which is spinning around its central axis (see Fig.11.8). The container is filled with  $LHe$ . If the motion occurs at  $T > T_C$  nothing particularly interesting happens but when we cool down the system at  $T < T_C$  the fraction of  $LHe$  in the superfluid state must respect the quantization. On a side point of view in which the vortices create a sort of triangular lattice. A vortex is a circulating flow which can satisfy the condition  $\nabla \times \mathbf{v}_s = 0$ . By rewriting this condition in circular coordinates one can find that this corresponds to:

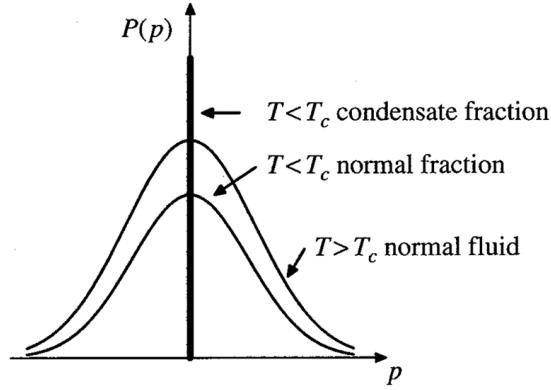
$$\nabla \times \mathbf{v}_s = 0 \Rightarrow \frac{1}{r} \frac{\partial}{\partial r} (r v_\phi) = 0$$

where  $v_\phi$  is the part of the velocity tangent to the vortex. From this expression, one can find that the velocity around the vortex must be of the form:

$$\mathbf{v}_s = \mathbf{v}_\phi = \frac{\kappa}{2\pi r} \mathbf{e}_\phi \quad (11.26)$$

with  $\mathbf{e}_\phi$  tangent versor. This equation is valid at any  $\mathbf{r} \neq 0$ . For  $\mathbf{r} \approx 0$ , in the so-called vortex core, the singularity on the velocity is solved by imposing  $\mathbf{v}_s = 0$ .

Many other observations can be made about superfluidity but we will limit ourselves to the last



**Figure 11.9:** Momentum density  $P(p)$ .

one<sup>6</sup>. We want to answer the question: "Is there a BEC in a superfluid?". Well, the answer is yes but the reason is complicated and stays in the distribution probability of the momentum of a particle in a superfluid which is reported in Fig.11.9. This distribution is made by two contributions. First a contribution from the condensate, which is a Dirac delta function at zero momentum corresponding to  $N_0$  particles in the BEC. Second, there is a contribution from the remaining  $N - N_0$  particles which is a smooth function of the momentum  $\mathbf{p}$ . This picture has been confirmed experimentally since the momentum distribution  $P(\mathbf{p})$  can be measured by neutron scattering. For liquid Helium at very low temperatures, both experiment and theory agree that, even at  $T = 0$ , the BEC part is only a 10% of the total fluid.

---

<sup>6</sup>For a deeper and better analysis of what is told so far I suggest the reading of James F. Annet - *Superconductivity, Superfluids and Condensates*. In particular, in the following, the quasi-particle excitation will not be discussed at all.

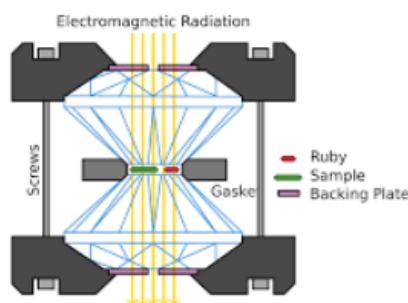
# Chapter 12

## Superconductivity

Superconductivity is probably one of the most interesting and fascinating subjects of low-temperature physics and, for sure, it is the most spoken topic of physics by non-specialised people. Superconductivity arouses interest because of its potential applications which go from quantum computers to transport and communication.

### 12.1 History and Phenomenology

Let's begin our story from the end. It is not unusual to see superconductivity news on media and newspapers, especially for what concerns quantum computing. The very first realization of a working quantum computer, in fact, was made with a superconducting circuit which exploited the Josephson effect, an effect typical of superconductors. Nowadays, the efficient and usable quantum computers are still made with superconductors. However, superconductivity is a phenomenon which occurs at low temperatures: how low the temperature has to be, makes the difference between feasible and not-feasible applications. A magnetic levitation train, for example, wouldn't be that useful if to make it work we must cool it down to  $2K$ . This is why researchers in the last two decades looked for superconductors with higher and higher critical temperatures. One of the latest trustable results (2014), for example, showed a critical temperature of  $203K$  under the application of a huge pressure of about  $200GPa$ . Clearly, despite being a very interesting scientific result, this is not very useful on the application side. One thing is to realize such pressure in a small volume of material in laboratory conditions, and quite another is to realize this in a macroscopic material, like a wire or a tube for instance. The method used in labs to realize this high-pressure condition is quite interesting. Two diamonds with a flat tip are pushed one against the other with an electromechanical system and a Rubi chip is used to measure the pressure (see Fig.12.1). Rubi's optical properties, in fact, change with the applied pressure and by impinging with appropriate electromagnetic radiation, it is possible to measure the frequency of the light diffracted by the Rubi and retrieve the pressure. The sample is placed between the two diamond's tips and a conducting ring is used to inject a current in the sample to study. However, we need to be very careful when we read about superconductivity and, in particular, about superconductivity at high temperatures.



**Figure 12.1:** High-pressure realization system.

In 2020 a group of scientists from the USA claimed to have obtained superconductivity at room temperature in a carbonaceous sulfur hydride<sup>1</sup>. Apparently, the results of the experiment were really good but nobody in the scientific community was able to reproduce these results. Some time later, the results were demonstrated to be fake and the article was retracted in September 2022. The scientist who found out the incorrectness of the data was Jorge Eduardo Hirsch who is also the inventor of the H-index, a numerical value to evaluate the quality of a researcher's work. The same group of scientists published another article in March 2023 claiming not only to have reached superconductivity at room temperature but also at ambient pressure. Unsurprising, the article was retracted a few weeks ago<sup>2</sup>, in August 2023. All these examples demonstrate how superconductivity is still alive and arouses a high scientific interest which, however, needs some precautions.

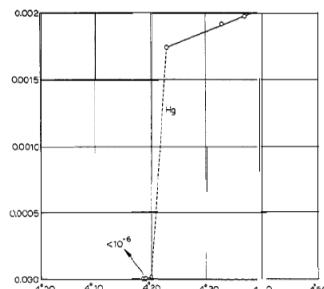
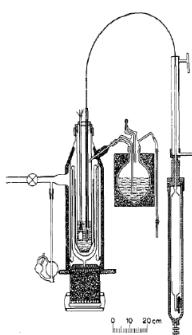
### 12.1.1 The first observations

The story of superconductivity is tightly bonded to the development of low-temperature physics, as we told in the introduction of Chapter 10. Physics of low temperature began in 1823 when Micheal Faraday managed to liquidy Chlorine *Cl*. Despite not being very skilled in mathematics, Faraday was a really good expert in laboratory techniques and by 1845 he had liquefied all known gases except six, named "permanent gases" (*O*, *H*, *N*, *CO*, *CH<sub>4</sub>* and *NO*). Another good contribution to low-temperature physics comes from James Dewar, a Scottish scientistic who worked at the University of Edinburgh. He is the inventor of the dewar, a double-walled flask of metal or silvered glass with a vacuum between the walls. This flask, thanks to the insulation property of the vacuum, kept the liquids cold for longer periods of time and helped revolutionize low-temperature research. By 1884 Dewar managed to "liquefy and solidify air". Through distillation processes Dewar was able to separate the elements which compose the air in the atmosphere and "play" with them, liquifying and solidifying them. Unfortunately, he wasn't able to liquefy Helium since, because of *He* light mass, it is not present in the atmosphere<sup>3</sup>. Helium liquefaction was achieved by Heike Kamerlingh Onnes on the 10<sup>th</sup> of July 1908. To achieve such a result, Onnes did not work alone. He understood that he would have needed a lot of *He*, which is not easy to find, a big laboratory and a lot of good technicians. He managed to obtain substantial state funding which he used to found a school for lab technicians and to build the lab itself. Onnes, among a lot of topics, was particularly interested in the study of electrical resistivity and in particular how it behaves with the decrease of the temperature. It was known at the time that resistivity depends on temperature. Onnes found out, by studying different metals, that the resistivity decreases with temperature until a plateau, at  $T \approx 0$ , which is due to the scattering of the electrons with the impurities of the lattice. Among all metals, he decided to study Mercury *Hg* and something completely unexpected

HEIKE KAMERLINGH ONNES

Investigations into the properties of substances at low temperatures, which have led, amongst other things, to the preparation of liquid helium

*Nobel Lecture, December 11, 1913*



**Figure 12.2:** Resistivity and *LHe* preparation scheme from Onnes Nobel lecture in December 1913.

<sup>1</sup> *Nature* - Vol 586, pages 373-377(2020).

<sup>2</sup> Few weeks ago with respect to the moment in which I am writing this lecture notes, which is October 2023.

<sup>3</sup> More on the life of James Dewar can be found at <https://www.aps.org/publications/apsnews/201201/physicshistory.cfm>

happened. The **Age of Superconductivity** was begun<sup>4</sup>. Thanks to the availability of liquid Helium, in 1911 Onnes and Holst observed in Mercury an abrupt decrease of resistivity, practically zero on the limit of the experimental precision allowed at the time. Onnes won the Nobel Prize only two years after for his amazing discovery. This was the first of a long series of Nobel prizes related to superconductivity. Fig.12.2 shows a part of Onnes Nobel lecture in which is possible to see the experimental behaviour of the resistivity as a function of the temperature for Mercury. We can see a clear and net transition from a discrete value to a null one at about  $4.2K$ . The data set for this experiment is made by only 7 experimental points. Apparently, they were very well chosen. In a more accurate measurement, it is possible to see that the transition is a bit more rounded for a real material. This was not visible in Onnes's measurement because of the lack of experimental points due to the technology limitations of the time. The reason for the more rounded transition stays in the fact that real materials are non-homogeneous and thus not all the regions of the materials become superconductive at the same time. In the following years, scientists started to investigate other materials to find the critical temperature at which the resistivity becomes zero. Tab.12.1 shows some of these metals with the corresponding critical temperature and year of discovery.

**Table 12.1:** Critical temperature of some superconducting elements

Element	Critical Temperature (K)	Year of Discovery
Mercury	4.15	1911
Tin	3.69	1913
Lead	7.26	1913
Tantalum	4.38	1928
Niobium	9.2	1930
Aluminium	1.19	1933
Vanadium	4.3	1934

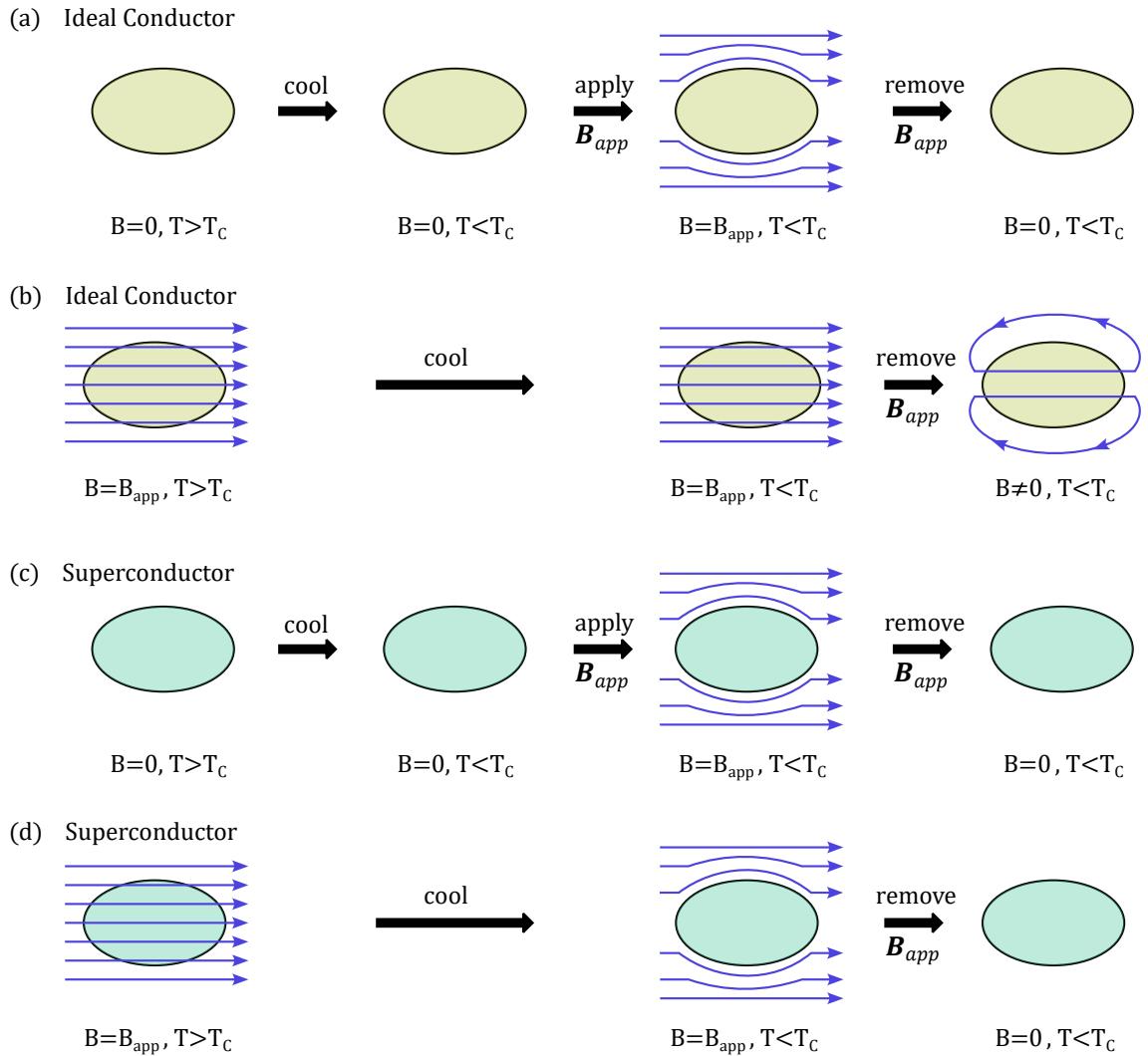
### 12.1.2 The Meissner-Ochsenfeld Effect

It is important to underline the fact that, below  $T_C$ , the resistivity does not limit itself to diminish, even if by a lot, but cancels out completely, as revealed by experiments carried out on rings made of superconducting material. If we induce currents in these rings by varying the magnetic flux inside them, for example by removing a magnet inserted in the central hole of the ring thus giving rise to an electromotive force, these currents never decay. Within the experimental errors, it can be said that the time constant of their possible decay cannot be less than 100 000 years. Nevertheless, the null resistivity is only half of the story for superconductivity. A material is a superconductor if it has zero resistivity *and* it is a **perfect diamagnet**, *i.e.* the magnetization  $\mathbf{M}$  is equal in modulus and opposite in direction to the internal field  $\mathbf{H}$ :

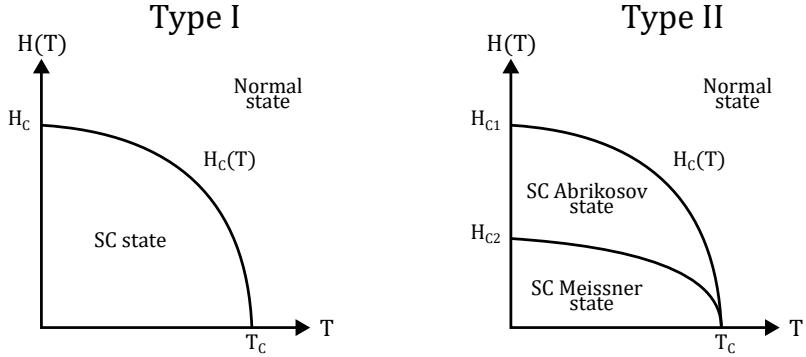
$$\mathbf{M} = -\mathbf{H} \Rightarrow \chi_m = -1$$

This leads to a complete annulment of the internal field inside the superconductor. Experimentally, we see that when a superconductor is cooled down in the presence of an external magnetic field, when the temperature goes below  $T_C$  suddenly the magnetic field outside the superconductor increases, which means that the magnetic field is expelled by the material. This is called **Meissner-Ochsenfeld effect**, it was discovered in 1933 by the two scientists who named the effect. To understand the importance of this effect and how this characterizes in a unique way superconductor materials let's look at Fig.12.3. The figure shows the difference in the magnetic behaviour for an ideal conductor, *i.e.* a conductor which has  $\rho = 0$ , and a superconductor, *i.e.* a conductor which has  $\rho = 0$  and  $\chi_m = -1$ . We can see in Fig.12.3(a) and (c) that an ideal conductor and a superconductor behave the same when they are first cooled down and then subjected to an external magnetic field is applied to them. When the external field is turned on, *i.e.* there is a temporal variation in the field, a current is induced into the material but, since the resistivity is zero, the value of the current is exactly the one necessary to shield the external field. As a result, the field is expelled by the material. When the field is turned off, *i.e.* the field varies in the opposite way, the previous

<sup>4</sup>Not only Onnes gave birth to the Age of Superconductivity, but also to the so-called "Big Science" which is the science made by large groups of researchers.



**Figure 12.3:** Meissner-Ochsenfeld effect: (a) An ideal conductor is first cooled down and then subjected to an external field and the opposite situation (b) in which the ideal conductor is subjected to the field and then cooled down. (c) Now, a superconductor is first cooled down and then subjected to an external field and (d) vice versa.



**Figure 12.4:** Field as a function of the temperature for Type I and Type II superconductors.

induced current is destroyed by another opposite induced current. The situation changes when we first apply the magnetic field and then cool down the system. For an ideal conductor (Fig.12.3(b)), since there is no variation in the external field when the system is cooled down, we have no current induction and therefore the field is not expelled by the material. However, when we turn the field off, the field varies and induces a current which will generate a field inside the material opposite in direction with respect to the external one<sup>5</sup>. These currents would then continue to flow indefinitely ( $\rho = 0$ ) and the material would become a permanent magnet. For a superconductor, instead, the situation is different. We can see in Fig.12.3(d) that when the superconductor is cooled down in the presence of an external magnetic field, as soon as the temperature decreases below  $T_C$  the magnetic field is expelled by the material. When the field is removed by the material nothing special happens. In conclusion, we can affirm that the magnetic behaviour of superconductors is a further aspect of their very peculiar nature that cannot be traced back simply to the cancellation of resistivity. A superconductor is both a perfect conductor and a perfect diamagnet.

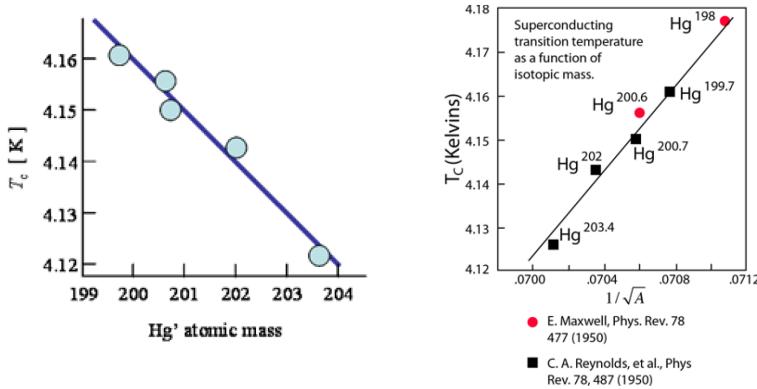
### 12.1.3 Type I and Type II Superconductors

Some years after the observation of the Meissner effect another important property of superconductors was observed. Lev Shubnikov found out that not all superconductors behave the same and for some types of superconductors an external magnetic field, above certain values, can penetrate inside the material. Moreover, it has been observed that a field strong enough destroys the superconducting state. This value of the magnetic field is called **critical field** and this phenomenon represents a serious problem for applications, especially for the manufacture of superconducting magnets. Even a current greater than a certain value, called **critical current**, destroys the superconducting state, and this too is a major technological problem. The behaviour with respect to the applied magnetic fields leads to a distinction between two types of superconductors. In the normal superconductors called **Type I**, we observe the existence of a unique critical field  $H_C$ , whose value depends on T, as shown in Fig.12.4, below which the material remains superconducting and there is no flux penetration. For this reason, this regime is called the Meissner state. For higher values of the field, the material returns to the normal state and the applied field penetrates freely. There are, however, materials in which the magnetic behaviour is more complicated. In these cases, there exists a lower critical field,  $H_{C1}$ , below which the magnetic flux does not penetrate. Above this, however, there is no transition to the normal state, to get to that we must reach a value  $H_{C2}$  said higher critical field. For applied fields between these two values, there is a partial penetration of the magnetic flux into the material. These were the superconductors observed by Shubnikov in 1936. These materials are called **Type II** superconductors. In between the two critical fields, the magnetic flux lines penetrate some points of the material that, locally, is not in the superconducting state. Around these flow tubes, there are areas of the material that are in the superconducting state and in which supercurrents circulate which prevent the field from penetrating the superconducting area. These supercurrent vortices are called **Abrikosov vortices** and for this reason, this regime is called the Abrikosov state. Alexei Abricokov was a Russian scientist who in 1957 formulated the theory of Type II superconductors explaining the origin of these vortices which were first observed by Shubnikov. Most superconductors are of Type II. Superconductors of the

<sup>5</sup>Just to be sure that everybody is on the same wavelength, the induced current comes from an electromotive force which is given by the Faraday-Neumann-Lenz law:  $f_{em} = -\frac{\partial}{\partial t}\Phi_{\Sigma}(\mathbf{B})$

**Table 12.2:** Upper critical field  $H_{C2}$  of superconducting compounds

Substance	Critical Temperature (K)	Upper Critical Field (Oe)
$Nb_3Ti$	8-10	90 000-130 000
$V_3Ga$	14.5	210 000-230 000
$Nb_3Sn$	17-18	220 000-250 000
$V_3Si$	17	230 000
$Nb_3Ga$	20	340 000
$Nb_3Ge$	21-24	370 000-400 000



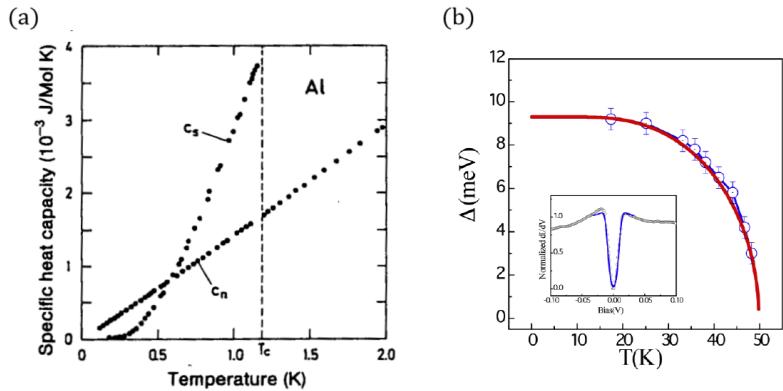
**Figure 12.5:** Isotope effect.

second type are more interesting for applications since the value of  $H_{C2}$  is much higher than the critical field in Type I superconductors and also much higher than  $H_{C1}$ . Tab.12.2 shows the value of the upper critical field  $H_{C2}$  for some relevant compounds. The field in the table is measured in Oersted (Oe)<sup>6</sup>

#### 12.1.4 Isotope Effect and Superconducting Gap

In 1950 another effect characteristic of superconductors was discovered and it is the so-called isotope effect. It has been observed that the critical temperature of superconductors depends on the atomic mass of the element. The relation between  $T_C$  and  $A$  is shown in Fig.12.5. It is worth noting that in a solid there is a particular phenomenon which is related to the atomic mass of the atoms forming the solid: the lattice vibrations. This is a clear hint for the fact that phonons could play a relevant role in the superconductivity phenomenon since their energy is related to the atomic mass and therefore they could affect the critical temperature. This observation will be of crucial importance for the following development of the superconductivity theory which will end with the BCS Theory. This theory states that phonons are the "trigger" particles to switch on superconductivity. Another relevant characteristic of superconductors is the existence of a superconducting gap. In a superconducting material below  $T_C$ , we observe the appearance of an energy gap which separates the superconducting electrons from the conduction band. The presence of a gap was hypothesized because of the discontinuity of the specific heat of superconductors. In superconductors, in fact, the specific heat  $C_s$  shows an abrupt increase at  $T_C$  and then it drops exponentially as visible in Fig.12.6(a). If the material is released from the superconducting state, for example by applying a fairly intense magnetic field, the trend of the normal behaviour of the specific heat is recovered. The presence of a discontinuity in the specific heat suggests that the transition to the superconducting state can be considered as a phase transition while the exponential trend suggests that there is a gap of empty states that separates the fundamental state of the superconducting electrons from the first excited state. The electrons in the fundamental state, in order to reach the first excited state, must overcome this step of energy and it is therefore understandable that this process of thermal activation, regulated by the Boltzmann factor, has an exponential character. Fig.12.6(b) shows the dependence of the energy gap on the temperature.

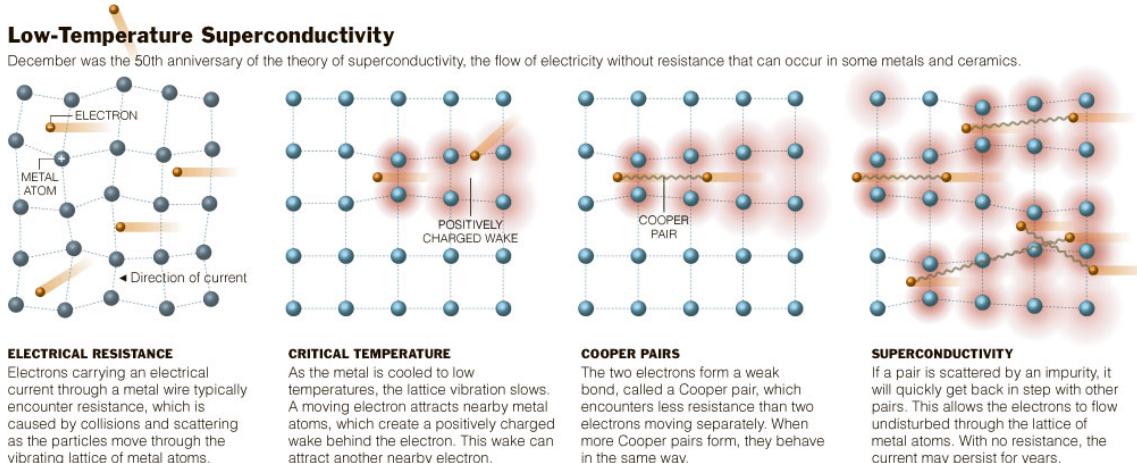
<sup>6</sup>1 Oe =  $1000/(4\pi) A/m \approx 79.577 A/m$



**Figure 12.6:** (a) Specific heat as a function of the temperature for a normal conductor and a superconductor.  
(b) Energy gap as a function of the temperature of a superconductor.

We can see that for  $T > T_C$  the gap is zero, *i.e.* the material is a normal metal, while when  $T$  decreases below  $T_C$  the gap starts to increase until reaching its maximum value at  $T = 0$ .

As well as the isotope effect, the superconducting gap is perfectly explained by the BCS Theory. All the phenomenology of the so-called conventional superconductor is beautifully summarized in the BCS Theory. This theory takes its name from the initials of physicists John Bardeen, Leon Cooper and Robert Schrieffer, who proposed it in 1957 and for which they received the Nobel Prize for Physics in 1972. It is one of the most brilliant theories of solid-state physics and its beauty and elegance still amuse us nowadays. The theory explain superconductivity thanks to the formation of the so-called **Cooper pairs** which are pairs of two electrons, bounded by an attractive interaction. Classically, this can be understood by imagining that, as it passes through the lattice of a solid, an electron, on account of its negative charge, leaves behind a deformation trail affecting the positions of the ion cores. This trail is associated with an increased density of positive charge due to the ion cores and thus has an attractive effect on a second electron. The lattice deformation therefore causes a weak attraction between the pairs of electrons. Unfortunately, this classical vision, despite being helpful in having a physical picture of the phenomenon, is not correct. A more detailed description of the phenomenon shows that the two electrons in the pair have opposite momentum  $\mathbf{k}$ , *i.e.* it is like they are moving in opposite directions. Fig.12.7 summarizes what is told so far. Quantum mechanically speaking, the BCS Theory attributes the cause of superconductivity to pairs of electrons formed by the interaction between two electrons via an exchange of a phonon. This is the attractive interaction cited before which led to the formation of Cooper pairs. After the formulation of the BCS theory, nothing particularly interesting happened in the field of superconductivity and the scientific interest on the subject dropped drastically until



Sources: Oak Ridge National Laboratory; Philip W. Phillips

JONATHAN CORUM/THE NEW YORK TIMES

**Figure 12.7:** The physical picture of BCS Theory by steps.

1986. In April 1986 a supernova lights up on the night of superconductivity: high-temperature superconductivity was observed in a cuprate (*LaBaCuO*) compound by Georg Bednorz and Karl Müller. This raised the interest of research and started the quest for higher critical temperatures. More and more scientists were interested in studying exotic superconductive compounds to find higher and higher critical temperatures. BCS Theory reliability has been tested by the discovery of the high- $T_C$  superconductors. The theory, in fact, does not allow the existence of a superconductive state above  $30 - 40K$  and nowadays theoretical physicists are still working on the formulation of a unique theory for superconductivity.

### 12.1.5 Superconductivity: a Nobel Story

We went through the whole story of superconductivity and we arrived at the present day, the very point from where we started a few pages ago. Superconductivity is a marvellous story not only physics-wise. Some of the greatest characters of superconductivity had the most strange and troubled life. However, superconductivity is also a story of Nobel prizes. Since everything began in 1911 a lot of Nobel prizes were assigned in the field of superconductivity. All of them are listed in Tab.12.3. In the following sections, we will describe the theories of superconductivity, starting from the phenomenological description by means of London's equations, moving to the thermodynamic Ginzburg-Landau model and ending with the most complete approach of the BCS Theory.

**Table 12.3:** Key discoveries and Nobel prizes in superconductivity.

Year of Discovery	Year of Nobel Prize	Discoverers	Reason for the Nobel
1911	1913	Onnes	Discovery of superconductivity in $Hg$
1938	1978	Kapitsa	Discovery of superfluidity in $^4He$
1947	1962	Landau	Theory of Superfluidity in $^4He$
1950	2003	Ginzburg	Ginzburg-Landau Theory of Superconductivity
1957	2003	Abrikosov Bardeen,	Type II Superconductor Theory
1957	1972	Cooper and Schrieffer	BCS Theory
1962	1973	Josephson	Josephson effect and supercurrents properties
1972	1996	Lee, Osheroff and Richardson	Discovery of superfluidity in $^3He$
1972	2003	Legget	Theory of Superfluidity in $^3He$
1986	1987	Bednorz and Müller	Discovery of High- $T_C$ superconductors

## 12.2 London Equations

The London equations are a set of two equations formulated by the brothers Fritz and Heinz London in 1935. These two equations are obtained by playing with the Maxwell equations applied to a medium with no resistivity and by introducing, without being able to justify it, an ad hoc hypothesis to explain the Meissner-Ochsenfeld effect. The equations can be derived both in the stationary and in the time-variable case. Let's start with the stationary case of a DC current.

### 12.2.1 Stationary Case - DC Current

The first equation we want to write is the Newton equation applied to the case of a constant electric field  $\mathbf{E}$  acting on a superconducting material, *i.e.* with no resistivity forces<sup>7</sup>:

$$\mathbf{F} = -e\mathbf{E} = m \frac{d\mathbf{v}_s}{dt} \quad (12.1)$$

<sup>7</sup>The equation in the general case of  $\rho \neq 0$  should include a viscous friction term proportional to the drift velocity:  $\frac{m}{\tau} \mathbf{v}_d$

where  $\mathbf{v}_s$  is the velocity of the supercarriers in the superconductor. Moreover, we know that in general, we can write the current density as:

$$\mathbf{j}_s = -en_s \mathbf{v}_s \quad (12.2)$$

where  $n_s$  is the supercarrier density. By substituting this result in Eq.(12.1) we get:

$$\frac{d\mathbf{j}_s}{dt} = \frac{e^2 n_s}{m} \mathbf{E} \quad (12.3)$$

This means that, because of the constant acceleration acting of the supercarriers, the current will increase as long as the field acts on the system which leads to a divergent behaviour. Nevertheless, Eq.(12.3) also tells us that, once the supercurrent is started, it will flow forever since we obtain a constant derivative of  $\mathbf{j}_s$  when  $\mathbf{E} = 0$ . Now, the third Maxwell equation, *i.e.* the Faraday's law, is:

$$\nabla \times \mathbf{E} = -\frac{d\mathbf{B}}{dt} \quad (12.4)$$

where  $\mathbf{B}$  is an external magnetic field. If we apply the curl to both sides of Eq.(12.3) and we use Faraday's law we get:

$$\begin{aligned} \nabla \times \frac{d\mathbf{j}_s}{dt} &= \frac{e^2 n_s}{m} \nabla \times \mathbf{E} = -\frac{e^2 n_s}{m} \frac{d\mathbf{B}}{dt} \\ \Rightarrow \frac{d}{dt}(\nabla \times \mathbf{j}_s) &= -\frac{e^2 n_s}{m} \frac{d\mathbf{B}}{dt} \end{aligned} \quad (12.5)$$

By rearranging the term we obtain the **First London Equation**:

$$\frac{d}{dt} \left( \nabla \times \mathbf{j}_s + \frac{e^2 n_s}{m} \mathbf{B} \right) = 0 \quad (12.6)$$

The term inside parentheses is the expression of the total field inside the material which is the sum of the field induced by the supercurrents and the external field. This equation is derived from Maxwell's equations by only imposing that the resistivity is null  $\rho = 0$  in Newton's law. Despite the fact that the equation predicts that the external field will be balanced by the induced currents, it does not include the Meissner effect. The effect, as told, is included with an ad hoc hypothesis. Since we know that the internal field inside a superconductor is null, we can impose:

$$\nabla \times \mathbf{j}_s + \frac{e^2 n_s}{m} \mathbf{B} = 0 \quad (12.7)$$

Which is the **Second London Equation**. The equation means that if there is a magnetic field in the superconductor there are also currents flowing along closed lines wound around the field lines. It should be pointed out that there are no arguments based on classical electromagnetism capable of justifying this ansatz whose legitimacy occurs *a posteriori* because, by means of it, it is possible to describe correctly the Meissner effect. To try to describe this effect let's define a constant:

$$\Lambda_L = \frac{m}{e^2 n_s} \quad (12.8)$$

and let's rewrite Eq.(12.3) and Eq.(12.7) as:

$$\mathbf{E} = \Lambda_L \frac{d\mathbf{j}_s}{dt} \quad (12.9)$$

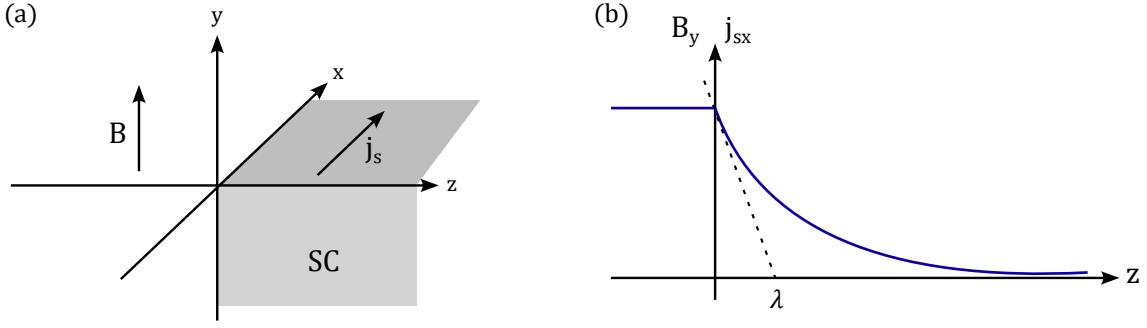
$$\mathbf{B} = -\Lambda_L \nabla \times \mathbf{j}_s \quad (12.10)$$

We can now use the fourth Maxwell equation, *i.e.* the Ampere's law:

$$\nabla \times \mathbf{B} = \mu_0 \mathbf{j}_s \quad (12.11)$$

By applying the curl to both sides of this equation and using Eq.(12.10) we get:

$$\nabla \times (\nabla \times \mathbf{B}) = \mu_0 \nabla \times \mathbf{j}_s = -\frac{\mu_0}{\Lambda_L} \mathbf{B} \quad (12.12)$$



**Figure 12.8:** Interface between the vacuum and a superconductor with an external field along the y-axis.

Or by taking the curl of Eq.(12.10) and using the Ampere's law:

$$\begin{aligned}\nabla \times \mathbf{B} &= -\Lambda_L \nabla \times (\nabla \times \mathbf{j}_s) \\ \Rightarrow \nabla \times (\nabla \times \mathbf{j}_s) &= -\frac{1}{\Lambda_L} \nabla \times \mathbf{B} = -\frac{\mu_0}{\Lambda_L} \mathbf{j}_s\end{aligned}\quad (12.13)$$

By remembering the vectorial relation for the rotor of a generic vector  $\nabla \times \nabla \times \mathbf{V} = \nabla(\nabla \cdot \mathbf{V}) - \nabla^2 \mathbf{V}$  and by also noticing that both  $\mathbf{B}$  and  $\mathbf{j}_s$  have a null divergence, *i.e.*  $\nabla \cdot \mathbf{B} = 0$  and  $\nabla \cdot \mathbf{j}_s = 0$ , we can rewrite the two equations obtained so far as:

$$\begin{cases} \nabla^2 \mathbf{B} - \frac{\mu_0}{\Lambda_L} \mathbf{B} = 0 \\ \nabla^2 \mathbf{j}_s - \frac{\mu_0}{\Lambda_L} \mathbf{j}_s = 0 \end{cases} \quad (12.14)$$

These two equations have the same form of the Poisson equation  $\nabla^2 V + \frac{\rho_c}{\epsilon_0} = 0$  and can be solved numerically. We are now interested in considering a particular and simplified case in which the equations have an analytical solution. We want to consider the case represented in Fig.12.8(a) in which the space is divided into two parts: for  $z > 0$  we have a superconductor while for  $z < 0$  we have the vacuum. An external magnetic field is applied along the  $y$  direction. Because of the symmetry of the problem, the solution must be translationally invariant along  $x$  and  $y$  and therefore we can simplify our problem by considering:

$$\begin{aligned}\mathbf{B} &= (0, B_y^0, 0) && \text{for } z < 0 \\ \mathbf{j}_s &= (j_{sx}^0, 0, 0) && \text{for } z > 0\end{aligned}$$

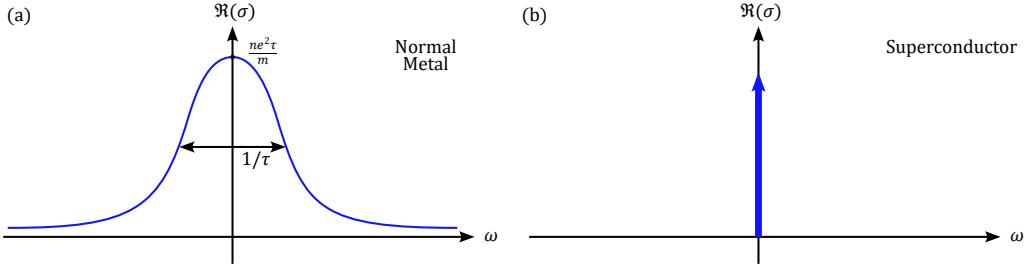
while the direction of  $\mathbf{j}_s$  along  $x$  has been chosen by considering that it has to compensate for the external field along  $y$ . We can write the system of Eq.(12.14) for this particular case:

$$\begin{cases} \frac{\partial^2}{\partial z^2} B_y(z) - \frac{\mu_0}{\Lambda_L} B_y(z) = 0 \\ \frac{\partial^2}{\partial z^2} j_{sx}(z) - \frac{\mu_0}{\Lambda_L} j_{sx}(z) = 0 \end{cases} \Rightarrow \begin{cases} B_y(z) = B_y^0 e^{-\frac{z}{\lambda}} \\ j_{sx}(z) = j_{sx}^0 e^{-\frac{z}{\lambda}} \end{cases} \quad (12.15)$$

where  $\lambda$  is the **London penetration length** which is defined as:

$$\lambda = \sqrt{\frac{\Lambda_L}{\mu_0}} \quad (12.16)$$

$\lambda$  is a property of the superconductor material which depends only on the density of supercarriers  $n_s$  through  $\Lambda_L$ . In particular,  $\lambda$  decreases when the density of the supercarriers increases. The smaller the  $\lambda$  value, the quicker an external field is "killed" in the superconductor. This can be seen also referring to Fig.12.8(b) which shows the behaviour of the field and the current density as expressed by Eq.(12.15). This solution does not tell us that the field is zero inside the material but that it drops exponentially inside it. Moreover, for the same reason, we see that the supercurrent does not flow in the whole material but only at its surface. We can now ask ourselves what happens if the current and the field are not constant but vary with time.



**Figure 12.9:** (a) Conducibility as a function of the frequency for a normal metal and (b) for a superconductor.

### 12.2.2 Time Dependent Case - AC Current

In the case of a variable electric field for a generic metal, we have:

$$\mathbf{j}(\omega) = \sigma(\omega)\mathbf{E}(\omega) \quad (12.17)$$

where both  $\mathbf{j}$ ,  $\mathbf{E}$  and the conducibility  $\sigma$  are complex number whose depend on the frequency  $\omega$ . This is the frequency which describe the change of  $\mathbf{j}$  and  $\mathbf{E}$  in time:

$$\begin{aligned} \mathbf{j} &= j_0 e^{-i\omega t} \\ \mathbf{E} &= E_0 e^{-i\omega t} \end{aligned} \quad (12.18)$$

From the Drude model it is possible to derive the dependence of the conducibility on the frequency which results in:

$$\sigma(\omega) = \frac{ne^2\tau}{m} \frac{1}{1 - i\omega\tau} \quad (12.19)$$

where the first part  $\sigma(\omega = 0) = \frac{ne^2\tau}{m}$  is the conducibility in the stationary case. The real part of this function is:

$$\Re(\sigma(\omega)) = \frac{ne^2}{m} \frac{\tau}{1 + \omega^2\tau^2}$$

This is a Lorentzian function with a FWHM of  $1/\tau$  and with a maximum at  $\omega = 0$  which corresponds to  $\frac{ne^2\tau}{m}$ . The real part of the conducibility as a function of the frequency is shown in Fig.12.9(a). In general, the real part of the susceptibility is related to a resistor component in the circuit while the imaginary part to an inductor or a capacitor component. For a superconductor, we have that  $\rho = 0$  which implies  $\Re(\sigma) \rightarrow \infty$ . We will have then that the FWMH of the Lorentzian becomes infinitesimally small and the maximum goes to infinity since  $\tau \rightarrow \infty$ . The real part of the conducibility becomes a Dirac Delta function for a superconductor (see Fig.12.9(b)). Therefore, for  $\omega \neq 0$  the real part of the conducibility is zero, *i.e.* there is no resistive component, and the conducibility is purely imaginary. We can write a general expression for a superconductor:

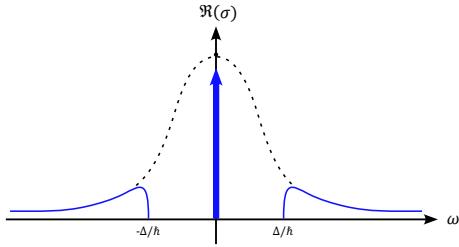
$$\sigma(\omega) = \frac{\pi e^2 n_s}{m} \delta(\omega) - \frac{e^2 n_s}{i\omega m} \quad (12.20)$$

The experiments show that  $\Re(\sigma)$  for a superconductor is not zero for any  $\omega \neq 0$ . Below and above a certain value of the frequency  $\Delta/\hbar$  we observe that  $\Re(\sigma)$  goes back to its "normal" behaviour. This is another evidence of the presence of an energy gap above which the material goes back to its normal behaviour. This is shown in Fig.12.10. If we now consider the expression  $\mathbf{j} = \sigma\mathbf{E}$  and we apply the curl to both sides of the equation and then we substitute what we have seen so far for the complex conducibility of a superconductor, we get:

$$\begin{aligned} \nabla \times (\mathbf{j}_s e^{-i\omega t}) &= \sigma \nabla \times (\mathbf{E}_0 e^{-i\omega t}) \\ &= -\sigma \frac{d}{dt} (\mathbf{B}_0 e^{-i\omega t}) \\ &= i\omega \sigma \mathbf{B}_0 e^{-i\omega t} \end{aligned} \quad (12.21)$$

For  $\omega \neq 0$  we have  $\sigma(\omega) = \Im(\sigma(\omega)) = -\frac{e^2 n_s}{i\omega m}$ . Thus we get:

$$\begin{aligned} (\nabla \times \mathbf{j}_s) e^{-i\omega t} &= -\frac{e^2 n_s}{m} \mathbf{B}_0 e^{-i\omega t} \\ \Rightarrow \nabla \times \mathbf{j}_s &= -\frac{e^2 n_s}{m} \mathbf{B}_0 \end{aligned} \quad (12.22)$$



**Figure 12.10:** Real behaviour of the conducibility as a function of the frequency for a superconductor.

which is the second London equation. Moreover, it is possible, by expressing the magnetic field as a function of the vector potential  $\mathbf{A}$ , to link  $\mathbf{j}_s$  to  $\mathbf{A}$  but this is not particularly interesting for our purposes<sup>8</sup>.

## 12.3 The Ginzburg-Landau Model

We can now move to a second theoretical approach to superconductivity which is related to thermodynamics aspects and which considers superconductivity as a phase transition, as suggested by the discontinuity of the specific heat we observed in Fig.12.6(a). The Ginzburg-Landau Model of superconductivity is based on the Landau model for second-order phase transitions and therefore requires spotting a suitable order parameter for the system. On the trail of what we have done for superfluidity, we choose a complex order parameter in the form of a wave function  $\Psi(\mathbf{r}, t)$ :

$$\Psi(\mathbf{r}, T) = |\Psi(\mathbf{r}, T)|e^{i\theta(\mathbf{r}, T)} \quad (12.23)$$

The choice of this order parameter is attributed to Landau and it was a real stroke of genius. Today, we are able to justify this ad hoc hypothesis since we know that  $|\Psi(\mathbf{r}, t)|^2 = n_s$ .

### 12.3.1 The Condensation Energy

The other choice we have to make regards the thermodynamic potential we want to use to describe the phase transition. Since we want to include in the model the magnetic effect related to superconductivity, we need an energy that depends on the magnetic energy of the system. The easiest choice would be to use the magnetic energy density  $u_m$ :

$$du_m = \mu_0(\mathbf{H} \cdot d\mathbf{M} + \mathbf{H} \cdot d\mathbf{H})$$

Unfortunately, the use of the magnetic energy of the system turns out to not be a good choice. The best solution is, once again, the use of the Gibbs free energy  $G$ :

$$G(T, \mathbf{H}) = U - TS - \mu_0 V \mathbf{H} \cdot \mathbf{M} \Rightarrow dG = -SdT - \mu_0 V M \cdot d\mathbf{H} \quad (12.24)$$

We want now to estimate the **condensation energy** of the system which is a measure of the gain in free energy per unit volume in the superconducting state compared with the normal state at the same temperature. We are therefore dealing with a transition from a normal state to a superconducting state. The quantity we want to compute is the energy difference between the superconductive state and the normal one at  $H = 0$ , i.e.  $G_s(T, 0) - G_n(T, 0)$ , because this will allow us to determine if the superconductive state is the most stable one. To do that, we can imagine going through the transition with the path indicated in Fig.12.11. We start from  $H = 0$  at a fixed temperature  $T < T_C$  and we increase the value of  $H$  until the intersection with the line which distinguishes the normal phase from the superconducting one. We have that:

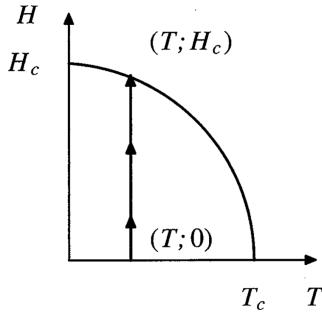
$$\Delta G = G_s(T, H_C(T)) - G_s(T, 0) = \int dG = \int_0^{H_C(T)} -\mu_0 V M dH \quad (12.25)$$

where we considered that we are going through a transition at a constant temperature, i.e.  $dT = 0$ . Now, if we consider that for a superconductor  $M = -H$  we get<sup>9</sup>:

$$\Delta G = - \int_0^{H_C(T)} \mu_0 V M dH = \int_0^{H_C(T)} \mu_0 V H dH = \frac{\mu_0 H_C^2 V}{2} \quad (12.26)$$

<sup>8</sup>More about this can be found on James F. Annet - *Superconductivity, Superfluids and Condensates*.

<sup>9</sup>This is an ad hoc hypothesis introduced in the model.



**Figure 12.11:** Path for the estimation of the condensation energy in a Type I superconductor.

We can now do the same for two generic points in the normal state and we find:

$$G_n(T, H_C(T)) - G_n(T, 0) = - \int_0^{H_C(T)} \mu_0 V M dH \approx 0 \quad (12.27)$$

The value of this integral is zero because we can consider that the magnetization for a generic superconductor in the normal state is practically zero if we neglect the small diamagnetic or paramagnetic behaviour of the metal. Thus we get:

$$G_n(T, H_C(T)) = G_n(T, 0) \quad (12.28)$$

It is now important to notice that the Gibbs free energy has a well-defined and unique value for any point  $(T; H)$  of the phase diagram. Therefore, along the line  $H_C(T)$  which separates the normal and the superconductive state, the value of the energy is the same because the field  $H$  and temperature  $T$  are the same. This means:

$$G_n(T, H_C(T)) = G_s(T, H_C(T)) \quad (12.29)$$

By putting everything together we find:

$$\begin{aligned} \Delta G_c &= G_s(T, 0) - G_n(T, 0) = G_s(T, H_C(T)) - \frac{\mu_0 H_C^2 V}{2} - G_n(T, H_C(T)) \\ &\Rightarrow \Delta G_c = -\frac{\mu_0 H_C^2 V}{2} < 0 \end{aligned}$$

This energy difference is smaller than zero which means that the superconductive state at  $T < T_C$  has a smaller energy than the normal one, *i.e.* is a favourable state for the system. By dividing by the volume, we get an estimation of the condensation energy  $\epsilon$ :

$$\epsilon = \frac{\Delta G_c}{V} = -\frac{\mu_0 H_C^2}{2} \quad (12.30)$$

The value of this energy can be estimated. For example, for Niobium:

$$\begin{aligned} \text{Nb : } \quad T_c &= 9 \text{ K}, \quad H_C = 160 \text{ kAm}^{-1} \Rightarrow B = \mu_0 H_C \approx 0.2 \text{ T} \\ \epsilon &= -\frac{\mu_0 H_C^2}{2} \approx 16.5 \text{ kJm}^{-3} \end{aligned}$$

This condensation energy corresponds to an energy of  $2 \mu\text{eV}$  per atom. This is the situation for Type I superconductors. In Type II superconductors everything is more complicated since we need to consider both the higher and the lower critical field and the magnetization will not always be equal to  $-H$  ( $M \neq -H$  in the Abrikosov regime). However, we can assign to the integral of the free energy a fictitious thermodynamic critical field  $\int M dH = \frac{1}{2} \tilde{H}_C^2$  and use it to compute the other thermodynamic quantities. For example, one can find that the variation in entropy between the two states is:

$$S_s(T, H_C(T)) - S_n(T, H_C(T)) = -\mu_0 H_C \frac{dH_C(T)}{dT} = \begin{cases} \neq 0 & \text{if } H_C(T) \neq 0 \\ = 0 & \text{if } H_C(T) = 0 \end{cases}$$

So we have that if  $H_C \neq 0 \Rightarrow \Delta S \neq 0$  during the transition there is an exchange of heat. We have a latent heat and therefore a first-order phase transition. If, instead,  $H_C = 0 \Rightarrow \Delta S = 0$  there is no latent heat and the transition is a second-order one. This can be physically understood by thinking that the latent heat during the transition comes from the dissipation of the current when the resistivity passes from  $\rho = 0$  to  $\rho \neq 0$  and vice versa. If the critical field is zero there is no current and therefore no dissipation.

### 12.3.2 The Landau Model of Phase Transitions

We now have all the ingredients to use the Landau model of phase transitions. First of all, we define our complex order parameter as:

$$\Psi = \begin{cases} 0 & \text{if } T > T_C \\ \Psi(T) & \text{if } T < T_C \end{cases} \quad (12.31)$$

We can consider the Helmozt free energy  $F$ :

$$F = G - \mu_0 V \mathbf{H} \cdot \mathbf{M} \quad (12.32)$$

In particular, we will consider the case in which  $H = 0$ , which means that  $F = G$ , and the case of an infinitely extended superconductor, *i.e.* we are describing the bulk properties of the superconductor. Moreover, is better to take the free energy density  $f = F/V$ . If we now expand  $f$  as a series of powers of the order parameter as done for superfluidity, we get:

$$f_s(T) = f_n(T) + a(T)|\Psi|^2 + \frac{1}{2}b(T)|\Psi|^4 + \dots \quad (12.33)$$

where  $f_s(T)$  and  $f_n(T)$  are the superconducting and the normal state free energy densities, respectively. The parameters  $a(T)$  and  $b(T)$  are assumed to be smooth functions of the temperature. We must also assume that  $b(T)$  is always positive, since otherwise the free energy density would have no minimum, which would be unphysical. Plotting  $f_s - f_n$  as a function of  $\Psi$ , is easy to see that there are two possible curves, depending on the sign of the parameter  $a(T)$  (see Sec.4.2). Since we want our energy to have symmetrical minima for  $T < T_C$  we can define:

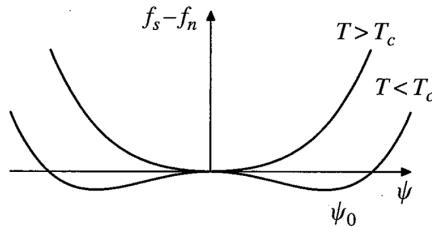
$$\begin{aligned} a(T) &= \bar{a}(T - T_C) & \bar{a} \in \mathbb{R}, \bar{a} > 0 \\ b(T) &= \bar{b} & \bar{b} \in \mathbb{R}, \bar{b} > 0 \end{aligned} \quad (12.34)$$

The plot of the free energy difference as a function of the order parameter is shown in Fig.12.12. By doing as in Sec.4.2, *i.e.* minimizing the expression of the energy, it is possible to find the order parameter corresponding to the minimum energy configuration, which is:

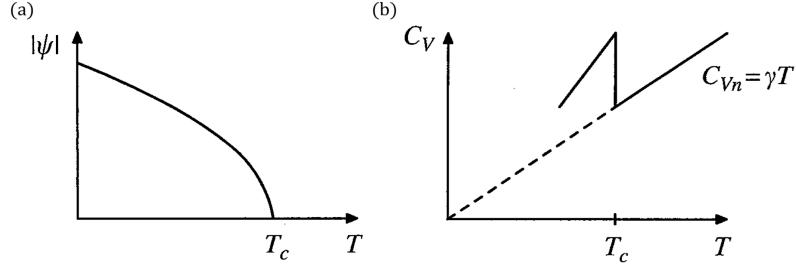
$$|\Psi| = \sqrt{\frac{\bar{a}}{\bar{b}}(T_C - T)} \quad (12.35)$$

for  $T < T_C$ . The behaviour of  $\Psi$  with temperature is shown in Fig.12.13(a) from which we recognize the classical trend of a phase transition. The corresponding value of the energy density at the minima is:

$$f_s(T) - f_n(T) = -\frac{\bar{a}^2(T - T_C)^2}{2\bar{b}} \quad (12.36)$$



**Figure 12.12:** Free energy difference between the normal and the superconducting state as a function of the order parameter.



**Figure 12.13:** (a) Order parameter and (b) specific heat as functions of the temperature.

But this is nothing more than the condensation energy we computed in the previous section:

$$\epsilon = \frac{\Delta G}{V} = f_s(T) - f_n(T) = -\frac{\bar{a}^2(T - T_C)^2}{2b} = -\frac{\mu_0 H_C^2}{2} \quad (12.37)$$

This expression can be used to express the value of the critical field as a function of the two parameters of the Landau model:

$$H_C(T) = \frac{\bar{a}}{\sqrt{\mu_0 b}}(T_C - T) \quad (12.38)$$

which is valid for  $T \simeq T_C$ . This means that the critical field close to  $T_C$  has a **linear trend** with temperature as shown in Fig.12.14(a). Moreover, from this free energy, we can also obtain other relevant physical quantities such as entropy or heat capacity. For example, by differentiating  $f$  with respect to  $T$  gives the entropy per unit volume  $s = S/V$ :

$$s_s(T) - s_n(T) = -\frac{\bar{a}^2}{b}(T_C - T) \quad (12.39)$$

We can see that there is no discontinuity in the entropy at  $T = T_C$ , *i.e.* no latent heat, which confirms that for  $H = 0$  superconductivity corresponds to a second-order thermodynamic phase transition. Differentiating the entropy we find the specific heat:

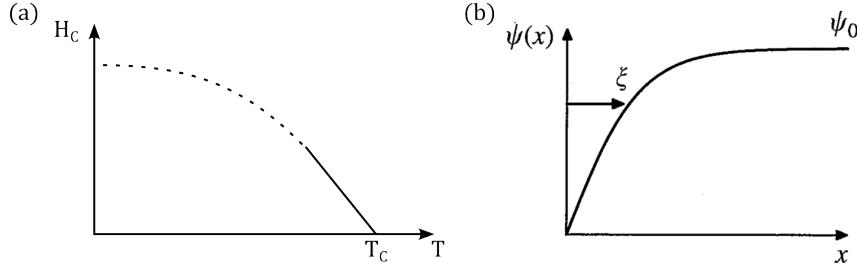
$$T \frac{\partial s}{\partial T} = C_V \Rightarrow C_{V_s} - C_{V_n} = \begin{cases} T \frac{\bar{a}^2}{b} & T < T_c \\ 0 & T > T_c \end{cases} \quad (12.40)$$

which means that at  $T = T_C$  the specific heat has a discontinuity  $\Delta C_V = T_C \frac{\bar{a}^2}{b}$ . This is shown in Fig.12.13(b). So far we described the case of  $H = 0$  in a bulk superconductor but what happens if we set a boundary for the extension of the superconductor and/or we have an external field? Clearly, the situation gets more and more complicated but essentially what one has to do is to include a position dependence on the order parameter  $\Psi(T) \rightarrow \Psi(\mathbf{r}, T)$  as in the general case of Eq.(12.23). Once this is done we can rewrite the expression of the free energy by considering two additive terms:

$$f_s(T) = f_n(T) + a(T)|\Psi|^2 + \frac{1}{2}b(T)|\Psi|^4 + \frac{1}{2m^*}|(-i\hbar\nabla - 2e\mathbf{A})\Psi|^2 + \frac{|\nabla \times \mathbf{A}|^2}{2\mu_0} \quad (12.41)$$

where:

- The fourth term is the quantum expression of the kinetic energy of the electric supercurrent that is flowing. To understand where this term comes from, let us remember that the kinetic energy is equal to  $p^2/2m$  and that, for an electron whose charge is  $-e$  in the presence of a magnetic field whose vector potential is  $\mathbf{A}$ , instead of the normal momentum we must consider the canonical momentum  $(\mathbf{p} - e\mathbf{A})$ . The quantum operator of the momentum is obtained by substituting to  $\mathbf{p}$  the expression  $-i\hbar\nabla$ . This term refers to the spatial fluctuations of the order parameter and assumes its minimum value when this is constant. Fluctuations therefore have an energy cost that must be taken into account. The term  $-e$  represents the charge of the electron, multiplied by two for taking into account the fact that the carriers are pairs of electrons, while  $m^*$  is the effective mass, in turn, multiplied by two.



**Figure 12.14:** (a) The critical field is linear with  $T$  for  $T \simeq T_C$ . (b) Order parameter of a superconductor near a surface.

- If a magnetic field is present close to the surface of the superconductor (not inside due to the Meissner-Ochsenfeld effect), then the fifth term must also be added, remembering that  $\mathbf{B}$  is the curl of  $\mathbf{A}$ .

We are trying to understand how the order parameter  $\Psi$  changes while passing from a zero value outside the superconductor, for example in the vacuum, to its value in the bulk, let's call it  $\Psi_0$ , which is the one we computed previously. We can consider a simple model for the interface between a normal metal and a superconductor. Suppose the interface lies in the  $yz$  plane separating the normal metal in the  $x < 0$  region from the superconductor in the  $x > 0$  region. We have in this case  $\Psi(\mathbf{r}) = \Psi(x)$ . Through a series of calculations which involve a procedure similar to the one done previously and the use of Schrodinger's equation, it is possible to demonstrate that  $\Psi$  evolves as:

$$\Psi(x) = \Psi_0 \tanh\left(\frac{x}{\sqrt{2}\xi(T)}\right) \quad (12.42)$$

where  $\xi(T)$  is called **Ginzburg-Landau coherence length** and it is defined as:

$$\xi(T) = \sqrt{\frac{\hbar^2}{2m^*|a(T)|}} \quad (12.43)$$

As we can see from Fig.12.14(b), this quantity shows how fast we pass from a normal state outside to a fully superconductive state in the bulk, *i.e.* how fast we go from  $\Psi = 0$  to  $\Psi = \Psi_0$ . We can also understand from this behaviour that the superconductive properties are not the same in the whole material but, for example, are different at the interface. The Ginzburg-Landau coherence length makes it possible to distinguish Type I and Type II superconductors. In fact, substituting the expression for  $a(T)$  and rearranging the terms, we can write:

$$\xi = \xi_0 \left( \frac{T_C}{T_C - T} \right)^{1/2} \xrightarrow[T \rightarrow T_C]{} \infty \quad (12.44)$$

The divergence behaviour of  $\xi$  at  $T = T_C$  implies that no coherence, *i.e.* no superconductive state, is possible at  $T = T_C$ . Furthermore, in the context of the Ginzburg-Landau model, it is possible to derive the London penetration length  $\lambda$  which one finds to be:

$$\lambda \propto \left( \frac{1}{T_C - T} \right)^{1/2} \quad (12.45)$$

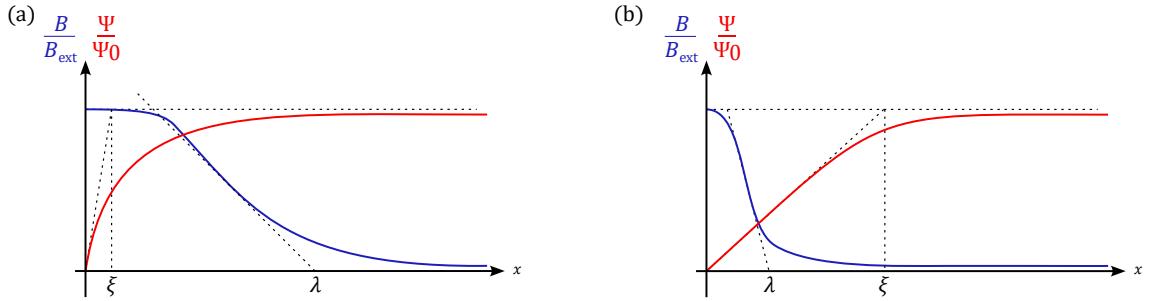
which has the same temperature dependence of  $\xi$ . This means that the ratio between these two quantities is constant:

$$\kappa = \frac{\lambda(T)}{\xi(T)} = \text{const} \quad (12.46)$$

This number  $\kappa$  only depends on the property of the superconductor material and its value can tell us something about the penetration of an external magnetic field into the superconductor, *i.e.* the type of the superconductor. We have that:

- $\kappa \ll \frac{1}{\sqrt{2}}$  for a Type I superconductor.

- $\kappa \gg \frac{1}{\sqrt{2}}$  for a Type II superconductor.



**Figure 12.15:** Magnetic field and order parameter behaviours as functions of the distance from the interface for (a) a Type II and (b) a Type I superconductor.

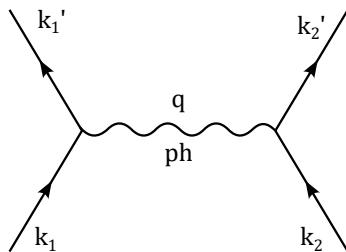
Let's refer to Fig.12.15 which depicts the behaviour of the magnetic field into a superconductor as found using London's equations and the order parameter behaviour found with the Ginzburg-Landau model. We have that if the magnetic field rapidly decays while entering the material (Fig.12.15(b)) and, at the same time, the order parameter, *i.e.* the superconductive state, is not strong at the surface, there is no penetration of the magnetic flux and the superconductive state is not destroyed by the field. This is the case of a Type I superconductor and, as we can see from the figure, in this case,  $\lambda \ll \xi$ . The other case, instead, is the one in which the magnetic field highly penetrates the material (Fig.12.15(a)). For these materials, which are Type II superconductors, it is convenient to create magnetic vortices which minimise the energy and, at the same time, do not destroy the superconductive state. In this case,  $\lambda \gg \xi$  and therefore  $\kappa$  is large.

## 12.4 The BCS Theory

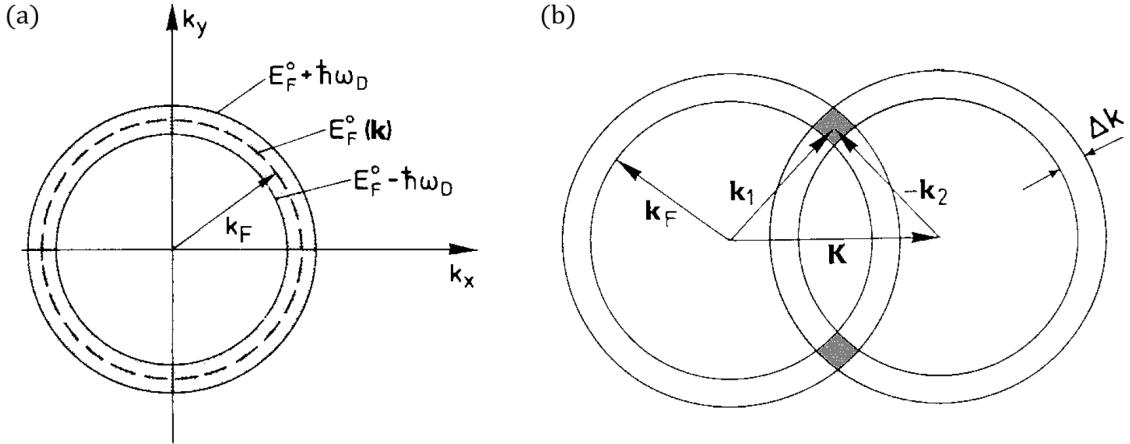
It is finally time to tackle the BCS Theory. We have already seen a qualitative and classic explanation of the theory and the end of Sec.12.1.4 and it is now the moment to deal with a more quantitative description. Despite being quantitative the dissertation we are going to make is not complete at all and for a more detailed description, it is suggested the reading of H. Ibach, H. Lüth - *Solid State Physics*. We will consider the case of a simple metal with a spherical Fermi surface. It is possible to say that the theory is summarized by the use of three ingredients:

- An electron-electron interaction: we already know that the interaction we are looking for is related to a phonon exchange.
- The solution to the "Cooper problem": consists of demonstrating that however weak the interaction between the two electrons is, they can form a stable pair.
- Many bodies problem: once a way to create pairs has been found, the theory assumes that all and only the electrons close to the Fermi surface can form pairs.

We know that lattice deformation can be seen as the superposition of phonons which the electron, due to its interaction with the lattice, continuously emits and absorbs. The formation of a pair is then due to this continuous exchange of a photon between the two electrons forming the pair. If we have two electrons with wave vector  $\mathbf{k}_1$  and  $\mathbf{k}_2$ , when they scatter with a phonon  $\mathbf{q}$  they will end in a new state with wave vectors  $\mathbf{k}'_1$  and  $\mathbf{k}'_2$ . This exchange is represented in Fig.12.16 in the form of a Feynman diagram. It is worth noting that in the picture no directionality is involved.



**Figure 12.16:** Feynman diagram of the phonon exchange between two electrons forming a Cooper pair.



**Figure 12.17:** (a) Fermi sphere in two dimensions for one electron with energy  $E$  in a interval of  $2\hbar\omega_D$ . (b) Electron pair collisions for which  $\mathbf{k}_1 + \mathbf{k}_2 = \mathbf{k}'_1 + \mathbf{k}'_2 = \mathbf{K}$ .

The arrows do not represent the direction of the wave vectors and the phonon in the middle has no direction since it can be transferred from one electron to the other with the same probability. In the following, we will ignore the spin of the two electrons although the spin plays a relevant role and it is possible to demonstrate that the two electrons in the pair need to have opposite spin which leads to a total spin of the pair equal to zero<sup>10</sup>. With non-easy calculation, it is possible to derive the effective interaction  $V_{eff}$  between the Cooper electrons mediated by the phonon, which is:

$$V_{eff}(\mathbf{q}, \omega) = |g_{\mathbf{q},\lambda}|^2 \frac{1}{\omega^2 - \omega_{\mathbf{q},\lambda}^2} \quad (12.47)$$

where  $\omega$  and  $\omega_{\mathbf{q},\lambda}$  are the frequencies of the electrons and the phonon respectively,  $g_{\mathbf{q},\lambda}$  is a matrix element and  $\lambda$  is a label to denote the mode of the phonon. We are interested in an attractive potential which corresponds to the case in which  $\omega < \omega_{\mathbf{q},\lambda}$ . This means that the phonons which play a relevant role are the high-energy ones because otherwise, we would get a repulsive interaction. The matrix element  $g_{\mathbf{q},\lambda}$  contains inside the information related to the electron-phonon exchange interaction. It is possible to see:

$$g_{\mathbf{q},\lambda} \propto \sqrt{\frac{m}{M}} \quad (12.48)$$

where  $m$  is the mass of the electron and  $M$  is the mass of the atoms. Although dependent on the mass of the ions, this matrix has an order of magnitude of  $\sim 10^{-3}$  which is a small number and can not be responsible for the isotope effect<sup>11</sup>. We want now to consider only the phonons with maximum energy which correspond to the Debye frequency  $\omega_D$ . This implies that we will consider:

$$V_{eff} \propto |g_{eff}|^2 \frac{1}{\omega^2 - \omega_D^2} \quad (12.49)$$

Moreover, if we consider only the electrons close to the Fermi surface, it is always true that  $\omega \ll \omega_D$ . Therefore, let's consider two electrons close to the Fermi energy. We want to write the two particles wave function for them but first, we need to define their wave vector values,  $\mathbf{k}_1$  and  $\mathbf{k}_2$ . To understand which are the optimal values of  $\mathbf{k}$  to form a Cooper pair, let's consider our Fermi sphere which is completely filled up to the  $k_F$  value. Let's consider two electrons in a small energy interval  $2\hbar\omega_D$  centred at  $\hbar k_F$  as in Fig.12.17(a). We will assume that all the electrons with  $E < E_F$  can not participate in the formation of the pair and that the two electrons that are forming the pair do not interact in any way with the other electrons below  $E_F$ . To summarise, we will consider only two electrons which will form a couple and ignore all the others. For the conservation of the total momentum, we have that before and after the phonon exchange, the overall momentum must be the same:

$$\mathbf{k}_1 + \mathbf{k}_2 = \mathbf{k}'_1 + \mathbf{k}'_2 = \mathbf{K} \quad (12.50)$$

<sup>10</sup>There also exist cases in which the spin of the two electrons is parallel leading to a total spin of 1. These cases are really rare and they will be ignored.

<sup>11</sup>We saw that the isotope effect depends somehow on the mass of the atoms. However, the dependence which led to this effect is contained inside the phonon energy  $\omega_{\mathbf{q},\lambda}$  and not in the matrix element.

where  $\mathbf{K}$  is a constant value. Since we are close to the Fermi surface we can assume  $|\mathbf{k}_1| \approx |\mathbf{k}_2| \approx k_F$ , therefore we can represent the situation in the reciprocal space with two circles of almost the same radius as in Fig.12.17(b). In the figure, we see that the two vectors  $\mathbf{k}_1$  and  $\mathbf{k}_2$  are in between a small circular area of thickness  $\Delta k$ . The figure shows that the only possible configurations in which the sum of  $\mathbf{k}_1$  and  $\mathbf{k}_2$  gives as a result a constant vector  $\mathbf{K}$  is when the two vectors are in the shaded area, which is the intersection between the two circles. This area can be seen as a measure of the number of phonon exchange processes between the two electrons, *i.e.* the strength of the attractive interaction. Since these processes reduce the total energy of the system, we want this area to be maximised. The area is maximised when the two circles perfectly overlap one with the other which corresponds to the case in which  $\mathbf{K} = 0$ . Therefore we find that the two wave vectors must be equal and opposite one another, so that:

$$\mathbf{k}_2 = -\mathbf{k}_1 = \mathbf{k} \Rightarrow \mathbf{K} = 0 \quad (12.51)$$

We can also estimate the distance between the two electrons forming the pair. We can suppose that a phonon during the exchange has a lifetime comparable with the period of its oscillations which is  $T = \frac{2\pi}{\omega_D} \approx 10^{-13} \text{ s}$ . The electrons move with a velocity which is the Fermi velocity  $v_F \approx 10^8 \text{ cm/s}$ . The distance between the two electrons will then be the space covered during one phonon oscillation which is  $d = T v_F \approx 10^{-5} \text{ cm} = 100 \text{ nm}$ . On an atomic scale, this is a huge distance and we can therefore understand why the Coulomb repulsion between the two electrons does not affect the stability of the pair.

#### 12.4.1 Two Particles Wave Function

We can now write the two-particle wave function  $\Psi(\mathbf{r}_1, \mathbf{r}_2)$ . First of all, we know that  $\Psi(\mathbf{r}_1, \mathbf{r}_2)$  must obey the Schrödinger equation:

$$-\frac{\hbar^2}{2m}(\nabla_1^2 + \nabla_2^2)\Psi(\mathbf{r}_1, \mathbf{r}_2) + V(\mathbf{r}_1, \mathbf{r}_2)\Psi(\mathbf{r}_1, \mathbf{r}_2) = E\Psi(\mathbf{r}_1, \mathbf{r}_2) \quad (12.52)$$

In this equation,  $E$  is the energy of the two electrons and it can be written as the sum of the two Fermi energies plus a term  $\varepsilon$  which represents the energy gain or loss in the formation of the pair:

$$E = 2E_F^0 + \varepsilon \quad (12.53)$$

where  $E_F^0$  is the Fermi energy at zero temperature. In this case, the two-particle wave function consists of two plane waves:

$$\Psi \propto \left( e^{i\mathbf{k}_1 \cdot \mathbf{r}_1} \right) \left( e^{i\mathbf{k}_2 \cdot \mathbf{r}_2} \right) = e^{i\mathbf{k} \cdot (\mathbf{r}_1 - \mathbf{r}_2)} \quad (12.54)$$

since the two electrons in the pair have opposite wave vectors. This equation implies that the two electrons must have opposite spin. In fact, this two-particle wave function is symmetric in spatial coordinates  $(\mathbf{r}_1, \mathbf{r}_2)$  under the exchange of electrons 1 and 2, but the whole wave function including spins must be antisymmetric (Pauli's exclusion principle general formulation). Therefore, an antisymmetric spin part is required, *i.e.* antiparallel spins, in order to preserve the correct symmetry. By doing a lot of calculations, it is possible to demonstrate<sup>12</sup>:

$$\varepsilon = \frac{2\hbar\omega_D}{1 - \exp\left\{-\frac{2}{V_0 Z(E_F^0)}\right\}} \quad (12.55)$$

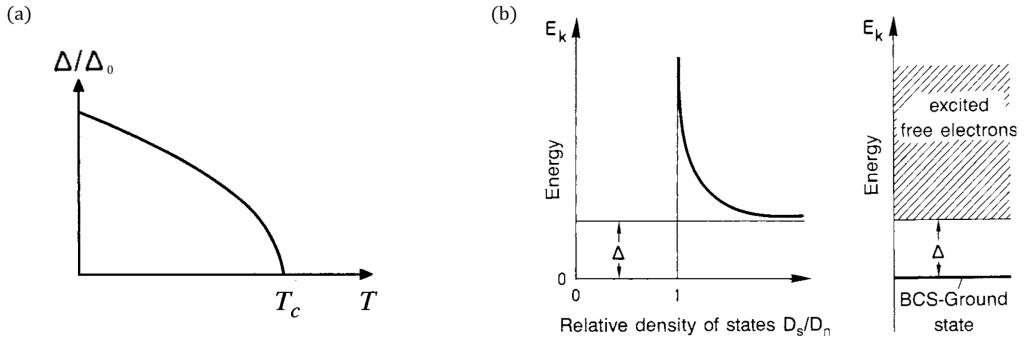
where  $V_0$  is a positive constant energy term and  $Z(E_F) = \frac{1}{2}g(E_F^0)$  is half of the density of states at the Fermi energy. In the case of a weak interaction,  $V_0 Z(E_F^0) \ll 1$ , it follows that:

$$\varepsilon \approx -2\hbar\omega_D e^{-\frac{2}{V_0 Z(E_F^0)}} \quad (12.56)$$

In which appears clearly that  $\varepsilon < 0$  and therefore there exists a two-electrons bound state whose energy is lower than of the fully occupied Fermi sea by an amount of  $\varepsilon = E - 2E_F^0 < 0$ . This is the solution of the Cooper problem which leads to the formation of the Cooper pairs. The last problem we need to solve is the so-called "many bodies problem" which consists of explaining the

---

<sup>12</sup>See H. Ibach, H. Lüth - *Solid-State Physics* for the detailed passages.



**Figure 12.18:** (a) Energy gap as a function of the temperature. (b) Density of state and energy gap.

energy variation of the system after the formation of, not only a single pair but an arbitrarily large number of pairs. We can imagine, in fact, that when the pair is formed our system is deprived of two fermions while a boson (the pair) appears. This causes a change in the dispersion relation of the electron of the material. Since bosons, differently from fermions, can stay in the same quantum state, *i.e.* in the same energy level, the new pairs, as soon as they form, will "fall" into a lower energy state, equal for all the couples. This leads to the formation of an energy gap for the system since the levels close to the Fermi level are depleted and a lower energy level is filled. This picture makes us understand a more subtle aspect of superconductivity. Cooper pairs live only inside the material and each time we try to excite a pair above the energy gap the pair gets destroyed. The only way to observe a pair is to make it tunneling from one superconductor to another and this is the so-called Josephson effect. To compute the condensation energy of the system, *i.e.* the energy difference per unit volume between the normal and the superconductive state, one can exploit the use of Pauli'matrices to build "creation" and "annihilation" operators for the Cooper pair. Then, the wave function is written as a superposition of the two states  $\mathbf{k}$  and  $-\mathbf{k}$  of the two electrons, and it is used in the Hamiltonian written with the "creation" and "annihilation" operators. A very long series of computations leads to the result:

$$\frac{W_{BCS} - W_n}{V} = -\frac{1}{2} Z(E_F^0) \Delta^2 \quad (12.57)$$

where  $\Delta$  is the energy gap of a superconductor which depends on temperature as shown in Fig.12.18(a). Once again, we find the behaviour of a phase transition. The energy gap at zero temperature is given by:

$$\Delta_0 = \Delta(T=0) \simeq 2\hbar\omega_D e^{-\frac{1}{V_0 Z(E_F^0)}} \quad (12.58)$$

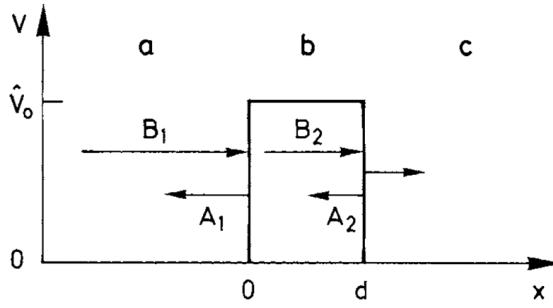
Fig.12.18(b) also shows the energy dependence as a function of the density of states. In this vision, the order parameter for the superconductor could also be the very energy gap  $\Delta$ . While we go through the critical temperature the superconductive state becomes more fragile since it gets easier and easier to break the pairs and excite one electron. This is because the gap becomes smaller with the increase of the temperature and therefore it costs less energy to break a couple and excite one electron. The energy which breaks the couples is the thermal energy. It is possible to demonstrate that the thermal energy at  $T = T_C$  is:

$$k_B T_C = 1.14 \hbar \omega_D e^{-\frac{1}{V_0 Z(E_F^0)}} \quad (12.59)$$

By using Eq.(12.58) and dividing it by Eq.(12.59) one gets:

$$\frac{\Delta_0}{k_B T_C} = \frac{2}{1.14} \simeq 1.764 \quad (12.60)$$

This equation makes it possible to relate the energy gap at  $T = 0$  to the critical temperature and this is one of the most astonishing results of the BCS Theory. Knowing the critical temperature of a material we can predict the width of the energy gap. Moreover, the measurements of  $\Delta_0$  or  $T_C$  it is also used to predict the value of the so-called coupling constant  $V_0 Z(E_F^0)$  which appears in almost all the equations. According to Eq.(12.59) and Eq.(12.58) the energy gap  $\Delta_0$  and the critical temperature  $T_C$  are proportional to the phonon cut-off frequency  $\omega_D$ . It can be seen that



**Figure 12.19:** Wave-mechanical description of tunnelling through a potential barrier of height  $V_0$  and width  $d$ .

this frequency is inversely proportional to the square root of the atomic mass of the atoms  $M$ . This implies:

$$\Delta_0, T_C \propto \omega_D \propto \frac{1}{\sqrt{M}} \quad (12.61)$$

This is the origin of the isotope effect. The agreement between the expected  $M^{-1/2}$  dependence and experiments is very good (see Fig.12.5).

## 12.5 Superconductor Junctions

Josephson effect is a tunnelling effect which occurs when two superconductors are placed at a small distance one from the other, separated by a vacuum or by a thin insulating layer. Before dealing with it let's remember the general tunneling effect. Let's consider the case of an atomic particle (*e.g.* an electron) of mass  $m$  which penetrates a potential barrier of of height  $V_0$  and width  $d$  as in Fig.12.19. Classically, if the energy  $E$  of the particle is smaller than the height  $V_0$  of the barrier, there is no chance for the particle to surmount the barrier. However, a quantum mechanical description of the phenomenon shows a finite probability of finding the electron in a region behind the barrier. One has to solve the Schrodinger equation in the  $a$ ,  $b$  and  $c$  regions:

$$\frac{d^2\psi}{dx^2} + \frac{2m}{\hbar^2}(E - V(x))\psi = 0 \quad (12.62)$$

where:

$$V(x) = \begin{cases} 0 & a \text{ region} \\ V_0 & b \text{ region} \\ 0 & c \text{ region} \end{cases} \quad (12.63)$$

By solving the equation in the different regions one obtains the following solutions:

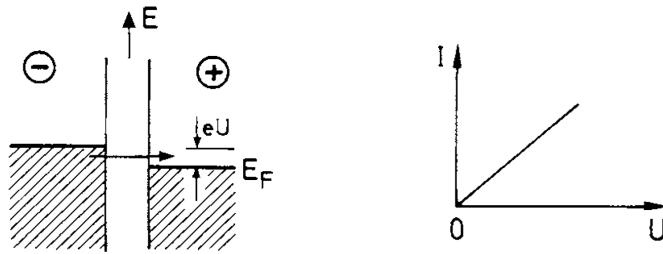
$$\begin{aligned} \psi_a &= A_1 e^{ikx} + B_1 e^{-ikx} \\ \psi_b &= A_2 e^{ik'x} + B_2 e^{-ik'x} \\ \psi_c &= e^{-ikx} \end{aligned} \quad (12.64)$$

with:

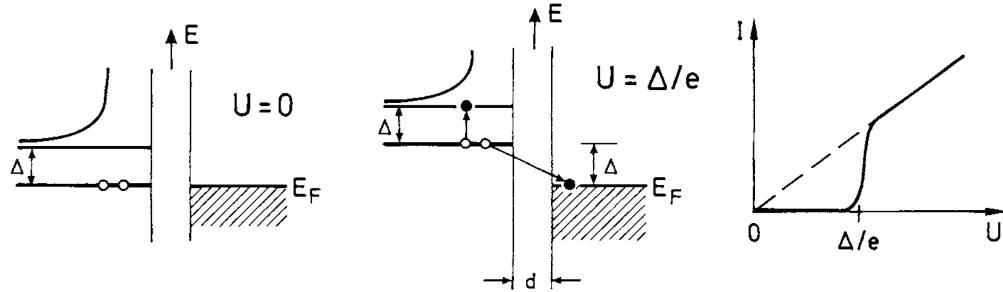
$$\begin{aligned} k &= \sqrt{\frac{2mE}{\hbar^2}} \\ k' &= \sqrt{\frac{2m(E - V_0)}{\hbar^2}} \end{aligned}$$

Clearly, if  $E < V_0$ ,  $k'$  is a complex number and gives a decay exponential behaviour for the probability of finding the electron in the  $c$  region behind the barrier:

$$\psi_c \psi_c^* \sim \exp\left\{-\frac{2}{\hbar} d \sqrt{2m(V_0 - E)}\right\} \quad (12.65)$$

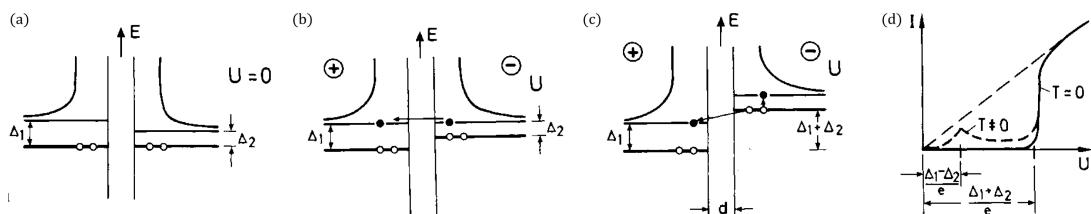


**Figure 12.20:** Schematic representation of the tunnelling of a single electron through an insulating layer in the metal-metal case.

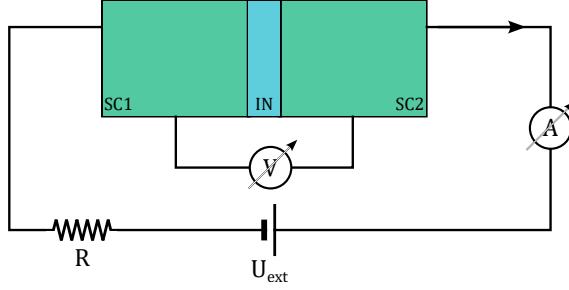


**Figure 12.21:** Schematic representation of the tunnelling of a single electron through an insulating layer in the superconductor-metal case.

So far, nothing new. The electrons are able to tunnel through an insulating layer between two conductors provided the layer has a sufficiently small thickness  $d$ . This is actually what happens in Scanning Tunnel Microscopy (STM) experiments in which the role of the insulating layer is played by a vacuum layer, and the roles of the metals are played by the STM tip and the sample. Let's start by considering the case of two metals separated by a thin insulating layer (see Fig.12.20) between which is applied a voltage  $U$ . The Fermi levels of the two metals are shifted by  $eU$  relative to one another. The occupied electronic states in the negatively biased metal perfectly match the empty states at the same energy in the other metal, and a current can flow through the barrier. For a small  $U$  the I-U characteristic is almost linear. If the negatively biased metal is in a superconducting state (see Fig.12.21) the situation changes. In a superconductor, in fact, we have an energy gap  $\Delta$  between the BCS many-body ground state, *i.e.* the state occupied by Cooper pairs, and the continuum of one-particle states. To make a current flow we need to break the Cooper pairs: one of the two electrons will tunnel towards the empty states of the normal metal and the other will go into the continuum of one-electron states in the superconductor. The couple is split apart. At  $U = 0$  the pair breaking is not possible since there is not enough energy for the process. For any small voltages  $U < \Delta/e$ , Cooper pairs can not break and current flow is thus impossible. However, when the bias reaches the value  $\Delta/e$  the Fermi level of the normal metal lies below the BCS ground state of a quantity  $\Delta$  and the breaking up of a Cooper pair and the current flow becomes possible. At this threshold voltage, the current jumps and then increases almost linearly with  $U$  as we can see on the right side of Fig.12.21. The tunnelling between two superconductors separated by an insulator can be treated in a similar way. Let's consider two



**Figure 12.22:** Schematic representation of the tunnelling of a single electron through an insulating layer in the superconductor-superconductor case.



**Figure 12.23:** Schematic representation of a superconductor junction circuit for Josephson effect.

superconductors with different energy gaps  $\Delta_1 > \Delta_2$ . For small voltages, for example,  $U = 0$ , either the empty states or the full states match in energy at  $T = 0$  and no tunnel current can flow (Fig.12.22(a)). When a voltage  $U = (\Delta_1 - \Delta_2)/e$  is applied at a finite temperature  $0 < T < T_C$ , the Cooper pairs can not break up yet but the non-bounded electrons, *i.e.* single electrons, can tunnel (Fig.12.22(b)). From Fig.12.22(d) we can see that a current suddenly sets in as soon as  $U > 0$  and reaches a maximum for  $U = (\Delta_1 - \Delta_2)/e$ . If we increase the voltage  $U$  up to  $U = (\Delta_1 + \Delta_2)/e$  the Cooper pairs in the right-hand superconductor can break up; one electron tunnels to the left and the other is simultaneously excited above the gap  $\Delta_2$  (Fig.12.22(c)). By measuring the I-U characteristics of the tunnelling between two superconductors at  $T \neq 0$ , one can determine the gap energies  $\Delta_1$  and  $\Delta_2$  as in Fig.12.22(d).

### 12.5.1 Josephson Effect

The Josephson effect is an effect which occurs in a superconductor junction and it is related to the phase difference of the order parameters  $\Psi$  which describe the two superconductors. Let's consider the circuit in Fig.12.23 in which two superconductors are connected through an insulating layer. The circuit is then closed and a voltage is applied to it while measuring both the current flowing through the junction and the voltage drop among it. From what we saw in the previous section, we know that if a bias  $U_{eff} \neq 0$  big enough is applied to the system, electrons should be able to tunnel from one superconducting piece to the other because of the breaking of the Cooper pairs. The wave functions associated with these superconductors have, in general, different phases. Each superconductor will be described by its order parameter with the associated phase term. Within each superconductor, the phase remains constant but a phase jump occurs during the transition from one superconductor to the other because of the presence of the insulating layer. Brian Josephson was a PhD student when he found out that a difference in the phases could give rise to a spontaneous superconducting current that could flow in both directions, even in the absence of an externally applied voltage difference. This phenomenon is called **Josephson effect**. Josephson demonstrated that the direction of the current could be determined by the difference between the two phases and in particular:

$$I = I_C \sin(\theta_2 - \theta_1) \quad (12.66)$$

where  $\theta_1$  and  $\theta_2$  are the phases of the superconductors SC1 and SC2 in Fig.12.23 and  $I_C$  is the critical current of the junction<sup>13</sup>. The current  $I$  is due to the tunnelling of the Cooper pairs and not of the single electrons as seen before. If, instead, the current  $I > I_C$  because, for example, a voltage  $U_{ext}$  is applied to the system, we start to see a drop of potential among the two superconductors due to the tunnelling of single electrons. Josephson found out that for  $I > I_C$  the voltage drop  $U$  across the superconductors is given by:

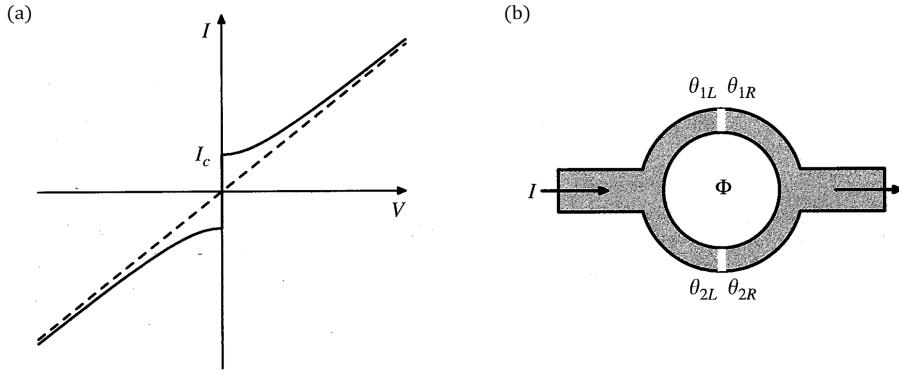
$$-qU = \hbar \left( \frac{d\theta_2}{dt} - \frac{d\theta_1}{dt} \right) \quad (12.67)$$

which means that if  $U \neq 0$  then the phase difference will vary with time. It is possible to demonstrate that the phase difference  $\Delta\theta = \theta_2 - \theta_1$  depends on time as:

$$\Delta\theta(t) = \Delta\theta_0 + \frac{2eU}{\hbar}t \quad (12.68)$$

---

<sup>13</sup>This critical current has nothing to do with the critical current which destroys the superconductive state in a general superconductor.  $I_C$  is a critical current for the junction.



**Figure 12.24:** (a) I-V characteristic for a Josephson junction. (b) Schematic representation of a SQUID device

which means that the current will vary with time as well:

$$I(t) = I_C \sin \Delta\theta = I_C \sin \left( \Delta\theta_0 + \frac{2eU}{\hbar} t \right) \quad (12.69)$$

The current oscillates in time with a frequency  $\omega$  which is:

$$\omega = \frac{2eU}{\hbar} \quad (12.70)$$

The I-V characteristic for a Josephson junction is shown in Fig.12.24(a). The Josephson effect is also at the heart of many different applications of superconductivity such as the most famous superconductive qubits. One of the simplest devices to make is a Superconducting Quantum Interference Device (SQUID) which is a simple superconducting ring in which there are two weak links and a magnetic flux  $\Phi$  concatenated to the ring (see Fig.12.24(b)). With the term "weak link" one can mean either a tunnel barrier (*e.g.* an insulator) or a thin normal metallic spacer. The current through the junctions depends on the phase difference across them and so the overall current flowing through the SQUID is given by the sum of the currents at each Josephson junction:

$$I = I_{C1} \sin \Delta\theta_1 + I_{C2} \sin \Delta\theta_2 \quad (12.71)$$

It is possible to demonstrate that the concatenated flux  $\Phi$  depends on the phase difference as well:

$$\Phi = \frac{\hbar}{2e} (\Delta\theta_1 - \Delta\theta_2) \quad (12.72)$$

We can rewrite this expression using the definition of the flux quantum of the Abrikosov vortices  $\Phi_0 = h/(2e)$ :

$$\Phi = \frac{\Phi_0}{2\pi} (\Delta\theta_1 - \Delta\theta_2) \quad (12.73)$$

This equation implies that because of the magnetic flux, the two phase differences are not equal one to the other. We have from Eq.(12.73):

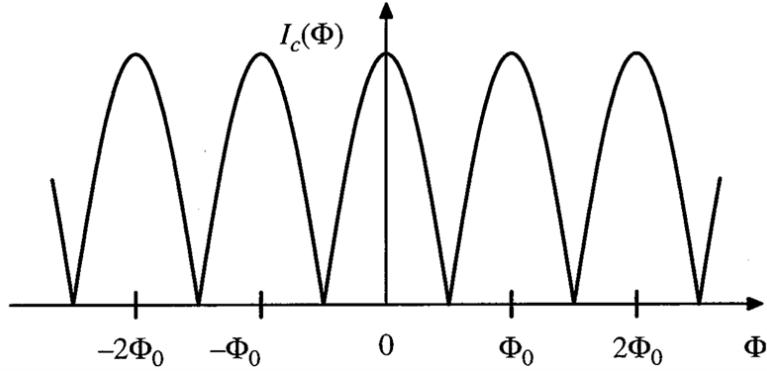
$$\Delta\theta_1 - \Delta\theta_2 = \frac{2\pi\Phi}{\Phi_0} \quad (12.74)$$

This means that we can write the two phase differences as:

$$\begin{aligned} \Delta\theta_1 &= \Delta\theta + \frac{\pi\Phi}{\Phi_0} \\ \Delta\theta_2 &= \Delta\theta - \frac{\pi\Phi}{\Phi_0} \end{aligned} \quad (12.75)$$

because with this definition the relation of Eq.(12.74) is satisfied without any loss of generality<sup>14</sup>. With this definition, in a perfectly balanced junction in which  $I_{C1} = I_{C2}$ , we can rewrite the current

<sup>14</sup>This part differs from what reported in James F. Annet - *Superconductivity, Superfluids and Condensates*. I think that in the version of the book that I have there is an error in the derivation of Eq.(12.72).



**Figure 12.25:** Modulation of the critical current in a SQUID ring.

of Eq.(12.71) as:

$$\begin{aligned}
 I &= I_{C1} \sin \Delta\theta_1 + I_{C2} \sin \Delta\theta_2 = \\
 &= I_C \sin \left( \Delta\theta + \frac{\pi\Phi}{\Phi_0} \right) + I_C \sin \left( \Delta\theta - \frac{\pi\Phi}{\Phi_0} \right) \\
 &= 2I_C \sin \Delta\theta \cos \left( \frac{\pi\Phi}{\Phi_0} \right) = I_C(\Phi) \sin \Delta\theta
 \end{aligned} \tag{12.76}$$

where  $I_C(\Phi)$  is the critical current which is modulated by a factor which depends on the net flux through the ring. Since the SQUID current will always have the same sign as the driving current we can write:

$$I_C(\Phi) = I_0 \left| \cos \left( \frac{\pi\Phi}{\Phi_0} \right) \right| \tag{12.77}$$

The behaviour of the critical current as a function of the magnetic flux is depicted in Fig.12.25. This is the pattern of an ideal Fraunhofer interference pattern, exactly analogous to the interference pattern one observes in optics with Young's two-slits experiment. Here the two Josephson junctions are playing the role of the two slits, and the interference is between the supercurrents passing through the two halves of the ring. The supercurrents shift in phase because of the presence of the magnetic field. The SQUID device provides a simple but highly accurate way of measuring the magnetic flux.